Optimal Social Security
with Imperfect Tagging*

Oliver Denk
OECD
Oliver.DENK@oecd.org

Jean-Baptiste Michau
Ecole Polytechnique
Jean-Baptiste.Michau@polytechnique.edu

March 2012

Abstract
Workers are exposed to the risk of permanent disability. We rely on a dynamic mechanism design approach to determine how imperfect information on health should optimally be used to improve the trade-off between inducing the able to work and providing insurance to the truly disabled. After deriving the first-order conditions to this problem, we calibrate the model and run a numerical simulation. The government should offer back-loaded incentives and make a strategic use of the gap in timing between the age at which disability occurs and the age of eligibility to disability benefits. Also, the able who are (mistakenly) tagged as disabled should be encouraged to work until some early retirement age. This makes a decrease in the strictness of the disability test desirable which would reduce the number of disabled who are not awarded the tag and, hence, improve insurance. Finally, we show how the first-best allocation of resources could asymptotically be implemented by making a strategic use of the disability test.

Keywords: Disability insurance, Imperfect tagging, Optimal policy, Social Security

JEL Codes: E62, H21, H55, J26

*We are grateful to Nick Barr, Tim Besley, Maitreesh Ghatak, Henrik Kleven, Alan Manning, Christopher Pissarides, Jeremy Sandford, Monica Singhal and to seminar participants at the London School of Economics for useful comments and suggestions.
1 Introduction

Two of the most pressing issues in public finance across industrialized nations are the rising costs of providing disability insurance and pension to an aging population. These two problems are in fact closely related as disability insurance is often used as a stepping stone towards retirement (Autor Duggan 2006, Li Maestas 2008). Furthermore, both programs can be seen as complementary ways of providing insurance to workers against the risk that they permanently lose their ability to work before they wish to retire. Indeed, in the U.S., the disability insurance and pension programs are the two main pillars of the "Social Security" system which is meant to provide security against the risk of being unable to work.

The disability insurance program relies on imperfect information on health to provide a decent income to those who are likely to be truly disabled. However, it is clearly not possible to provide perfect insurance against the disability risk as some agents who are truly disabled fail to qualify. Thus, systematic eligibility to old-age pensions beyond a certain age is justified as another, complementary, way of providing insurance. Indeed, this is what motivated Bismarck to invent pension programs as early as 1889.

In 2007, the U.S. Social Security system provided income to almost 50 million individuals for a total cost of $585 billion (4.2% of GDP) of which 9 million received disability benefits\(^1\) for a total cost of $99 billion (0.7% of GDP) (SSA 2008). By contrast, in 2007, the total cost of unemployment insurance was only $32 billion, about a third of the size of the disability insurance program. Despite these gigantic numbers and the potentially large welfare implications of the permanent disability risk (Chandra Samwick 2006), very little is known about the optimal design of insurance against this risk in a dynamic context with imperfectly observable health. The aim of this paper is to characterize such an optimal policy, to provide the key intuitions and quantitative insights and, finally, to give an order of magnitude of the potential welfare gains to be expected.

Let us now describe our theoretical framework. The government could rely on its imperfect information on health to enhance the provision of insurance against the disability risk by giving higher consumption to those who seem to be unable to work. More precisely, those who seem to be in poor health are "tagged"\(^2\) as disabled and therefore eligible for this higher consumption level. However, tagging is imperfect and some classification errors are unavoidable. Hence, some workers who are able to work are awarded the tag, while others who are truly disabled are rejected. Recognizing this problem, the planner still wants to provide the able and tagged with incentives to work. Thus, the optimal allocation is found by setting up a dynamic mechanism design problem where the able,

\(^1\)7 million of those where disabled workers, which represents about 4.4% of the population between the ages of 25 and 64. The other beneficiaries are the spouses and children of disabled workers.

\(^2\)The term was originally introduced by Akerlof (1978).
whether tagged or not, are induced to work until some retirement age to be determined. Intuitively, providing incentives to the tagged is more costly, as their outside option is more attractive, and, hence, it is optimal to let them retire earlier than the untagged. Since the adjustment is done on both margins, retirement age and consumption levels, the able and tagged should also get higher pensions.

By this channel, the optimal policy provides some support for the implementation of a health-dependent retirement age. However, it should be emphasized that this retirement age depends on health as observed by the government but only applies to the able, who are, by definition, in good health.

The first-order conditions to this problem relate inverse marginal utilities across ages and across states, i.e. being tagged or untagged. Thus, the need to preserve incentives prevents the standard equalization of marginal utilities of consumption. The optimal allocation is characterized by back-loaded incentives. The consumption of the able is increasing with age and jumps when a tag is awarded. The optimal policy also makes use of the difference in timing between the award of the tag and the occurrence of disability. The idea is that someone untagged who claims to be disabled is likely to say the truth if he becomes tagged shortly after stopping to work but is probably lying if he remains untagged for long. The former should therefore be rewarded with high consumption while the latter should be punished. To illustrate these features and to have a more quantitative sense of the main characteristics of the optimal policy, we calibrate the model with U.S. data and perform a numerical simulation.

One of the important differences between the current U.S. situation and the optimal policy is that the able who are tagged are currently not induced to work whereas they should be until some early retirement age. When such incentives are provided, it becomes desirable to lower the strictness of the disability test in order to decrease the number of truly disabled who are denied the tag. Our estimation suggests that these changes would generate welfare gains of about 0.2% of consumption. We also show that it is important to implement the optimal health-dependent retirement age as inducing the able and tagged to work until the general retirement age would be excessively costly and could result in a welfare loss.

These numerical results are obtained assuming that the strictness of the disability test is chosen to minimize the total number of classification errors, but allowing for a preference between rejection and award errors. Although, this is a natural and realistic benchmark, we might be interested in the optimal policy when the strictness of the disability test at each age is directly under the control of the planner. We show that, in this setup, the first-best allocation of resources could asymptotically be implemented. The idea is to set a very high threshold after the retirement age, so that the untagged are almost surely able to work, and to severely punish those who claimed to be disabled in the past. While it
might not be realistic to believe that such an extreme policy is implementable in practice, this result nevertheless suggests that significant welfare gains can be obtained by setting the disability threshold strategically and, hence, by moving beyond the minimization of classification errors which characterizes the current U.S. policy.

It is important to emphasize that, in this paper, we exclusively focus on the determination of the optimal incentive-feasible allocation. We do not investigate how it could be implemented in a decentralized market economy where the government could only use fiscal instruments instead of choosing individuals’ consumption directly. Note that, while optimal allocations are typically unique, there usually exist multiple ways of implementing them. Thus, in general, results about allocations are more robust than about implementation.

Our paper builds on two strands of the literature. First the seminal work of Diamond and Mirrlees (1978) determines the optimal provision of social insurance against the risk of permanent disability with unobservable health. As inducing the able to work is costly, they find that the general retirement age should be smaller than the first-best retirement age. Their work therefore gives a justification for the provision of old-age pensions as an imperfect insurance against the risk of permanent disability.

Despite its generality, the Diamond Mirrlees (1978) model has never been used quantitatively to investigate optimal retirement policies. Thus, one contribution of this paper is to provide some quantitative results based on the Diamond Mirrlees approach to retirement. More generally, while, following Golosov, Kocherlakota and Tsyvinski (2003), important developments have been made on the optimal provision of social insurance against the stochastic evolution of workers’ skills, only few quantitative results have been obtained. An important exception is Golosov and Tsyvinski (2006) who found that, with unobservable health, the welfare gains generated by the possibility to tax savings amount to 0.5% of consumption. Although they focus on the risk of permanent disability, as we do, their model only allows for an intensive margin to labor supply and, hence, it cannot say anything about the optimal retirement age.

The second strand of the literature on which we build traces back to Akerlof (1978) who argued that, in the presence of asymmetric information, incentive compatibility constraints could be relaxed by relying on some publicly available information correlated with agents’ private information. This general principle naturally applies to disability insurance and retirement programs where health is the hidden information which the government can nevertheless imperfectly observe. Indeed, Diamond and Sheshinski (1995), Parsons (1996) and Salanie (2002) showed that welfare could be improved by giving more to those who seem to be disabled, even if the government’s information is very imperfect.

3Cremer, Lozachmeur and Pestieau (2004a) made a similar point in a simplified setup which allows for heterogeneous productivity among workers.
In particular, the work of Parsons (1996), which insists that the able who are tagged should be incentivized to work, is closely related to this paper. However, all these models are static and do not give any quantitative evaluation of the welfare gains to be expected from the imperfect observability of health.

Thus, our work combines these two approaches to optimal social insurance by introducing imperfect tagging into the dynamic mechanism design approach of Diamond and Mirrlees (1978).

While, following the seminal contribution of Shavell and Weiss (1979), there has been a considerable literature on optimal unemployment insurance, little is known about the optimal design of disability insurance. In addition to the work mentioned above, relevant contributions include Benitez-Silva, Buchinsky and Rust (2006) who rely on a careful empirical analysis of the tagging process to propose an optimal statistical screening rule which would result in fewer classification errors. Kleven and Kopczuk (2009) consider an environment where the government needs to impose some complexity into the system in order to obtain imperfect information on health. This has the adverse consequence of reducing take-up. They therefore characterize the optimal trade-off between complexity and take-up. Low and Pistaferri (2008) propose a structural model of labor supply in a life-cycle setting where workers are subject to both disability and productivity shocks. Relying on an empirical estimation of their structural parameters, they argue that increasing the strictness of the disability test would enhance welfare.

Our paper is also related to the work of Cremer, Lozachmeur and Pestieau (2004b, 2007). In their setup workers are heterogeneous in both productivity and health. Labor is supplied along the extensive margin which allows them to endogenize the retirement age. They argue that welfare can be improved by resorting to disability testing. However, to gain tractability, they considerably simplify the dynamic structure of the model by assuming that agents are \textit{ex-ante} heterogeneous in terms of their ability to work, instead of having the uncertainty about disability gradually unfolding over time. Also, they assume that disability testing is perfect and that the only reason why the government makes a limited use of audits is that they are costly. By contrast, we have no such costs in our model as these might be negligible compared to the welfare gains generated by imperfect tagging.

In a way, our contribution is to look at the welfare gains generated by the integration of disability insurance and pension programs, but restricting attention to permanent disabilities. Along similar lines, Stiglitz and Yunn (2005) argued that large benefits could be expected from the integration of unemployment insurance and pension programs. While we focus on imperfectly observable health, a number of papers have recently argued in favor of allowing policies to rely more extensively on observable characteristics correlated with hidden information such as productivity. For example, it has been shown that sig-
significant welfare gains could be generated by making taxes dependent on age (Weinzierl 2008), on gender (Alesina Ichino Karabarounis 2008) or even on height (Mankiw and Weinzierl 2007).

In section 2, we present the theoretical model. We first describe the setup, then turn to the planner’s problem before giving the first-order condition characterizing the optimal policy. Then, in the following section, we calibrate the model. Section 4 is devoted to the numerical simulation and to the description of the corresponding welfare gains. Finally, in section 5, we describe how the first-best allocation of resources can be implemented if the government sets the strictness of the disability test strategically. The paper ends with a conclusion.

2 Model

This section describes the theoretical framework used to determine the optimal Social Security system with imperfectly observable health. We first present the setup, then give the planner’s problem and, finally, present the conditions which characterize the optimum.

2.1 Setup

All agents face a deterministic life span equal to $H$. Time is continuous, which is necessary to obtain a first-order condition for the retirement age. Resources could be safely transferred from one period to the next at an exogenous interest rate. For simplicity, we take this interest rate to be equal to the agents’ discount rate $\rho$. Everyone derives instantaneous utility $u(c)$ from consuming $c$, where $u' > 0$ and $u'' < 0$. At a given age, people are either able or disabled. Only the able can work. Their productivity evolves deterministically over time and is equal to $\gamma_t$ for a worker of age $t$. We will later assume, in the calibration section, that $\gamma_t$ follows an inverted U-shape. Labor supply is indivisible$^5$, which is a necessary assumption in a model of endogenous retirement. Working generates an instantaneous utility cost of $b$.

As the main justification for the provision of pension is to insure workers against the

\footnote{Finkelstein, Luttmer and Notowidigdo (2009) have recently provided some evidence that the marginal utility of consumption of the elderly declines as health deteriorates. However, we do not yet know whether the marginal utility of consumption differs between periods of employment and leisure. Thus, for simplicity, our specification assumes a constant marginal utility of consumption across all states.}

\footnote{It would be fairly straightforward to add an intensive margin. Indeed, in a similar setup, Golosov and Tsyvinski (2006) assume a continuous labor supply. However, in models that include both margins, such as Rogerson Wallenius (2008), Prescott Rogerson Wallenius (2009) or Michau (2009), most of the action occurs at the extensive margin. Furthermore, Liebman Luttmer and Seif (2009) provide some empirical evidence that most of the labor supply response to changes in the level of Social Security benefits occurs at the extensive margin.}
loss of their ability to work, we assume that disability hits people stochastically over time and that it is an absorbing state. The corresponding c.d.f. is denoted by $F(t)$ and the p.d.f. by $f(t)$, where $t \in [0, H]$. Thus, at age $t$ a fraction $F(t)$ of the population is disabled.

In order to have a well-defined social insurance problem, with ex-ante identical individuals, we shall assume that the planner attaches a zero weight on those who became disabled before starting to work.\textsuperscript{6} Such unfortunate individuals should certainly be taken care of, but outside the Social Security system which we investigate. This reflects the current U.S. situation where eligibility to Social Security requires some employment history.

With a diminishing marginal utility of income, the first-best allocation of resources is characterized by the provision of full insurance against the disability risk. Consumption should therefore be constant across all states and, hence, independent of whether an individual is able to work or not. Able workers should eventually retire to enjoy some leisure and typically choose to do so once their productivity becomes small.

If health is private information, this allocation of resources is not incentive compatible as able people have an incentive to masquerade as disabled in order to retire earlier and to save the disutility cost of working. This has lead Diamond and Mirrlees (1978) to characterize, within the above framework, the optimal provision of Social Security with unobservable health. They found that the consumption level of disabled should be sufficiently low to induce the able to work. Furthermore, incentives are back-loaded, i.e. the disabled should be provided with higher consumption if they stopped working at a more advanced age. But the most remarkable feature of the optimal policy is that it puts everyone into retirement before the first-best retirement age. The intuition for this result is that there is eventually so many disabled that it would be too costly, from a welfare perspective, to push their consumption level down in order to induce the able to work.

It could, however, be objected that the assumption of unobservable health is too extreme. More realistically, the government can obtain some imperfect information on the work ability of its citizens. It can run a medical test and "tag" as disabled those who fail the test. However, as the information is imperfect, some errors are made leading to the occurrence of gaps and leakages. Gaps occur when some disabled individuals are untagged; while leakages occur when some able are tagged.\textsuperscript{7}

\begin{footnotesize}
\textsuperscript{6}When determining the planner's optimal allocation, this is equivalent to imposing the normalization $F(0) = 0$.
\textsuperscript{7}In most of the existing literature on misclassifications in disability insurance programs (see, e.g., Benitez-Silva, Buchinsky and Rust, 2006), rejection (award) error is referred to as the probability of being disabled (able) conditional on being untagged (tagged), and type I (II) error as the probability of being untagged (tagged) conditional on being disabled (able). We, in contrast, define gaps as the number of individuals who are disabled and untagged, and leakages as the number of individuals who are able and tagged. Since there is a mass 1 of individuals, gaps is equivalent to the probability of being
\end{footnotesize}
Note that it would certainly be welfare enhancing for the government to use more detailed information on health. For instance, it could assign each disability applicant a probability of being truly unable to work. However, for reasons which are beyond the scope of our analysis, the US Social Security system, as in many countries, relies on a simple tagging process where individuals are either classified as disabled or not. We therefore follow most of the literature on the topic and constrain the government to rely on a simple tagging process.

More formally, let \( \theta \) denote the outcome of the test for a given individual. Thus, \( \theta \) could be thought of as his apparent health. Its c.d.f over the population is \( G_A(\theta) \) for the able and \( G_D(\theta) \) for the disabled. The respective p.d.f.s are denoted by \( g_A(\theta) \) and \( g_D(\theta) \). An individual is tagged as disabled if his \( \theta \) falls below a threshold \( \hat{\theta} \) which determines the disability standard. Thus, an able individual is tagged with probability \( G_A(\hat{\theta}) \) and a disabled with probability \( G_D(\hat{\theta}) \). Following Diamond and Sheshinski (1995), we assume that \( G_A \) first-order stochastically dominates \( G_D \), i.e. \( G_A(\theta) < G_D(\theta) \) for all \( \theta \) and that the two distributions satisfy the monotone likelihood ratio condition, i.e. \( g_A(\theta)/g_D(\theta) \) is increasing in \( \theta \). Furthermore, we assume that, for a given individual, \( \theta \) remains fixed throughout his life except for a drop when he becomes disabled. When determining the disability standard \( \hat{\theta} \), the government faces a trade-off between the number of gaps and leakages. See Figure 1.

![Figure 1: Trade-off between gaps and leakages](image)

Note that the share of disabled is very small among young individuals, but is much larger among senior people. Thus, as age increases, leakages become a smaller source of concern, while the opposite is true for gaps. We therefore assume an age-dependent disabled and untagged, and leakages to the probability of being able and tagged.
threshold, i.e. equal to $\hat{\theta}_t$ at age $t$, which is non-decreasing with age. Hence, at age $t$, the number of gaps is equal to $[1 - G_D(\hat{\theta}_t)]F(t)$ and that of leakages to $G_A(\hat{\theta}_t)[1 - F(t)]$.

We are now in a position to derive the joint p.d.f. of the ages at which people become disabled and tagged. Note that the structure of the problem implies that being tagged is an absorbing state. Let $i$ and $j$ stand for the ages at which an individual becomes disabled and tagged, respectively. We can consider that $i = H$ if someone dies while still able and $j = H$ if he dies untagged. Let $f(i, j)$ denote the joint p.d.f. of $(i, j)$. From Bayes’ law:

$$f(i, j) = f(j|i)f(i),$$

where $f(i)$ is the previously defined exogenous p.d.f. of ages at which people become disabled. With an obvious abuse of notation, the p.d.f. of getting tagged at age $j$ given that disability occurs at $i$, for $0 < i < H$, is given by:

$$f(j|i) = \begin{cases} 
G_A(\hat{\theta}_0) & \text{if } j = 0 \\
g_A(\hat{\theta}_j) \frac{d\theta_j}{d\theta} & \text{if } j < i \\
G_D(\hat{\theta}_i) - G_A(\hat{\theta}_i) & \text{if } j = i \\
g_D(\hat{\theta}_j) \frac{d\theta_j}{d\theta} & \text{if } i < j < H \\
1 - G_D(\hat{\theta}_H) & \text{if } j = H 
\end{cases} \quad (2)$$

A fraction $G_A(\hat{\theta}_0)$ of individuals obtain the tag at time 0. To understand the second and fourth cases, i.e. $j < i$ and $i < j < H$, note that the only way by which an agent could become tagged if he does not simultaneously become disabled is that the threshold $\hat{\theta}_j$ increases sufficiently so that his own constant $\theta$ falls below the threshold. For an able worker, this occurs with probability $G_A(\hat{\theta}_{j+\varepsilon}) - G_A(\hat{\theta}_j)$ over a time interval of length $\varepsilon$. The corresponding probability density is equal to $[G_A(\hat{\theta}_{j+\varepsilon}) - G_A(\hat{\theta}_j)]/\varepsilon$ with $\varepsilon$ tending to 0. The same argument applies for a disabled. The third case, $j = i$, gives the probability of becoming tagged when the disability occurs. This is equal to the probability of being tagged once disabled, $G_D(\hat{\theta}_i)$, minus the probability of being already tagged before becoming unable to work, $G_A(\hat{\theta}_i)$. Thus, the p.d.f. $f(i, j)$ is degenerate since a mass of agents become disabled and tagged simultaneously. In fact, this sounds sensible as the occurrence of disability should certainly lead to a deterioration of the apparent health observed by the government. Finally, the last case, $j = H$, corresponds to the probability of dying untagged. For completeness, note that for someone dying able, $i = H$, (2) simplifies to:

$$f(j|i = H) = \begin{cases} 
G_A(\hat{\theta}_0) & \text{if } j = 0 \\
g_A(\hat{\theta}_j) \frac{d\theta_j}{d\theta} & \text{if } j < H \\
1 - G_A(\hat{\theta}_H) & \text{if } j = H 
\end{cases} \quad (3)$$
Here, the last three cases of (2) boil down to a single one, i.e. $j = H$.

Importantly, we will assume throughout this paper that able individuals do not know the value of their fixed $\theta$. All they know is whether they are eligible for the tag or not. While this assumption is somewhat restrictive, it is reasonable that, conditional on remaining able to work, agents cannot predict when they will become eligible for the tag. Note that the alternative benchmark, where people would know their $\theta$, would imply that they could predict at age 25 when they would become eligible for the tag conditional on remaining able. One way to think about our assumption is that people get a private medical check-up every year and that their doctor advises them to apply for the tag once they become eligible for it. In other words, the fixed $\theta$ is just an irrelevant modeling device and all that matters at age $t$ is the threshold $\hat{\theta}_t$ together with the probability of being awarded the tag, which is equal to $G_A(\hat{\theta}_t)$ for an able person and to $G_D(\hat{\theta}_t)$ for a disabled. It is important to emphasize that this approach provides a reduced form that captures the dynamic trade-off between gaps and leakages; it certainly does not pretend to give a realistic representation of the very complicated process by which the true and apparent physical condition of an individual evolve over time.

The problem of the social planner is to maximize the expected utility of workers at time 0 subject to the resource constraint and to the incentive compatibility constraints which ensure that the able choose to work. The imperfect information on health should make it possible to increase welfare by relaxing the incentive compatibility constraints. Note that, as the planner attaches a zero weight on individuals who became disabled before time 0, all agents could be considered to be initially able to work. Ex-post, a given individual is characterized by when he became disabled, i.e. age $i$, and when he became tagged, i.e. age $j$, where the ex-ante probability density of being individual $(i, j)$ is given by $f(i, j)$, as defined in (1).

One of the key control variables of interest in this paper is the retirement age of the able. In order to exploit the imperfect information on health, the planner will make this retirement age conditional on when someone got tagged, i.e. conditional on $j$. We denote by $RT(j)$ the retirement age of an able worker who got tagged at age $j$. Those whose apparent health is lower, i.e. lower $\gamma$, will be tagged earlier. This implies that $j$ is a sufficient statistic for the apparent health of the able and tagged and, hence, $RT(j)$ is a health-dependent retirement age. Even the untagged will eventually retire when their productivity becomes low as they want to enjoy some leisure. We denote by $RU$ the

\footnote{Also, note that, strictly speaking, $\theta$ is only fixed with respect to the distributions $g_A(\theta)$ and $g_D(\theta)$, but that these distributions could well shift over time. In particular, we might expect the apparent health of both the able and disabled to deteriorate as people get older, which is not a problem provided that both distributions shift by the same amount. Similarly, the threshold $\hat{\theta}_t$ is only increasing relative to $g_A(\theta)$ and $g_D(\theta)$.}

\footnote{The disabled trivially retire when they lose their ability to work.}
retirement age of the untagged.\textsuperscript{10}

It is important to emphasize that the tagged who are able to work do not retire immediately. Instead, those tagged at age $j$ are induced to work until $RT(j)$, provided that they remain able to work until that age. Parsons (1996) insisted that with only imperfect information on health, and therefore the possibility of leakages, there is no reason to force all the tagged to become inactive. This is not unrealistic and, indeed, in many countries, those officially registered as disabled are offered incentives to work, which could be seen as an illustration of this. Obviously, this requires commitment from the government which might be tempted to untag those who reveal that they are able to work.

The remaining control variables are the consumption levels corresponding to the different histories. Where appropriate, these should be allowed to depend on the age at which disability occurred or at which the tag was awarded. However, the planner only observes $j$ while $i$ is revealed by the incentive compatible policy. Hence, after an able worker retires, the planner cannot know whether he remains able to work or not. Therefore, when determining the optimal consumption levels, all retired agents could be considered disabled. Thus, an agent qualifies for the consumption of a disabled from age $r = \min\{i, RT(j)\}$ if tagged at $j < RU$ or from age $r = \min\{i, RU\}$ otherwise, where $r$ stands for his effective retirement age. The planner needs to determine the consumption at age $t$ of the able who are untagged, $\{c^{AU}(t)\}_{t \in [0, RU]}$, of the able who became tagged at $j$, $\{c^{AT}(t, j)\}_{j \in [0, RU], t \in [j, RT(j)]}$, of the untagged who retired at $r$, $\{c^{DU}(t, r)\}_{r \in [0, RU], t \in [r, H]}$, and of the retired at $r$ and tagged at $j$, $\{c^{DT}(t, r, j)\}_{j \in [0, RU], r \in [j, RT(j)]]}$. In this last case, for reasons that will subsequently become clear, we distinguish whether the individual retired first, $r < j$, or was tagged either first or when retiring, $j \leq r$. Note that these last two consumption functions, $c^{DT}(r, j)$ and $c^{DT}(r, j)$, should also depend on age, $t$. However, as the discount rate is equal to the interest rate, there is nothing to be gained from distorting their consumption levels over time. In other words, in this case, age does not provide any information on whether the agent is able to work or not and allowing consumption to depend on age would not help the social planner to relax any incentive compatibility constraint.

Let $v(i, j)$ stand for the \emph{ex-post} lifetime utility of an $(i, j)$ individual who became disabled at $i$ and tagged at $j$. If an agent retires before becoming tagged, i.e. $\min\{i, RU\} <$

\textsuperscript{10}Clearly, those who only get tagged after $RU$ retire at $RU$. 10
$j$, his utility is:

$$v(i, j) = \int_0^{\min\{i, RU\}} e^{-\rho t} \left[ u(c^{AU}(t)) - b \right] dt$$

$$+ \int_{\min\{i, RU\}}^j e^{-\rho t} u(c^{DU}(t, \min\{i, RU\})) dt$$

$$+ \int_j^H e^{-\rho t} u(c^{DT}(\min\{i, RU\} ; j)) dt.$$  

(4)

From age 0 to $\min\{i, RU\}$ the worker is able and untagged, he consumes $c^{AU}(t)$ at age $t$ and gets disutility $b$ from working. From age $\min\{i, RU\}$ to $j$, he is disabled and untagged and gets the corresponding consumption level where, again, from the perspective of the planner the agent became disabled at $\min\{i, RU\}$. Finally, from age $j$ to $H$, his consumption level is that of a disabled and tagged who became disabled first. Now, if an agent becomes tagged before retirement or if he becomes disabled and tagged simultaneously, i.e. $j \leq \min\{i, RU\}$, his utility is:

$$v(i, j) = \int_0^j e^{-\rho t} \left[ u(c^{AU}(t)) - b \right] dt$$

$$+ \int_0^{\min\{i, RT(j)\}} e^{-\rho t} \left[ u(c^{AT}(t, j)) - b \right] dt$$

$$+ \int_{\min\{i, RT(j)\}}^H e^{-\rho t} u(c^{DT}(\min\{i, RT(j)\} ; j)) dt.$$  

(5)

From age 0 to $j$, the worker is able and untagged; from $j$ to $\min\{i, RT(j)\}$, he is able and tagged; and from $\min\{i, RT(j)\}$ to $H$, he is disabled and tagged. Note that the able and tagged are induced to work until age $\min\{i, RT(j)\}$ and, hence, get disutility $b$ from work.

2.2 Planner’s problem

The planner solves the following problem:

$$\max E[v(i, j)] = \int_0^H \int_0^H v(i, j) f(i, j) d\mu dj$$

(6)

subject to:

- Resource constraint,
- Incentive compatibility constraint at age $t$ for the untagged, $\forall t \in [0, RU)$,
- Incentive compatibility constraint at age $t$ for those tagged at $j$, 

11
The control variables are \( c^{AU}(\cdot), c^{AT}(\cdot), c^{DU}(\cdot), c^{DT}(\cdot), RT(\cdot), \) and \( RU. \) The full planner’s problem is given in the appendix. The objective of the planner is to maximize the \textit{ex-ante} expected lifetime utility, where each individual has probability \( f(i, j) \) of becoming individual \((i, j)\) with lifetime utility \( v(i, j)\). The resource constraint imposes that the expected lifetime consumption of individuals does not exceed the amount that they are expected to produce, where the working agents, those getting disutility \( b \) from work in (4) and (5), produce \( \gamma_t \) units of consumption goods at age \( t \) (see equations (A2) and (A3) of the appendix). The first set of incentive compatibility constraints imposes that the untagged who are able choose to work until \( RU \). Similarly, the second set of incentive compatibility constraints ensures that, even when tagged at \( j \), the able still choose to work until \( RT(j) \). Note that this last set of constraints is formally identical to the one imposed by Diamond and Mirrlees (1978), as, once an agent is tagged, the government cannot rely on any additional information about his health and therefore acts as if health was completely unobservable.

The difference between \( c^{DT}(r, j) \) depending on whether the individual retires first, i.e. \( c^{DT}_D(r, j) \) for \( r < j \), or becomes tagged at retirement or before, i.e. \( c^{DT}_T(r, j) \) if \( j \leq r \), is explained by the fact that the latter consumption level enters the incentive compatibility constraint of the tagged while the former does not.

It should be emphasized that the generality of the planner’s problem implies that, once the optimal allocation has been derived, there is no additional screening mechanism which could further improve welfare. In Parsons (1996) and Kleven Kopczuk (2009), individuals applying for the tag cannot know in advance whether they are going to be successful or not. However, the disabled have a higher probability of being awarded the tag than the able. Thus, a high cost of applying for disability benefits, through fees or complexity, could be used as a screening device to reduce, or even eliminate, leakages. However, this possibility does not arise in our framework where agents know the outcome of the test, thanks to their private doctor for instance, before applying.

\[ \forall j \in [0, RU), \forall t \in [j, RT(j)). \]

2.3 First-order conditions

The planner’s problem can be solved using Lagrange multipliers. If we take the utility levels of the various agents to be the control variables, rather than their consumption levels, then the objective and the incentive compatibility constraints are linear while the resource constraint is convex. Hence, the corresponding first-order conditions are both

\[ \text{Importantly, the planner does not control the disability standard } \{ \theta_t \}_{t \in [0, H]} \text{ which is exogenously determined. This assumption will be relaxed towards the end of the paper.} \]
necessary and sufficient. Since the planner is trying to provide social insurance optimally against a certain type of stochastic evolution of workers' skills, it should not be surprising that most of these conditions take the form of inverse Euler equations. Indeed, Golosov, Kocherlakota and Tsyvinski (2003) showed that inverse Euler equations characterize the optimum in a wide class of social insurance problems.

The remaining first-order conditions correspond to the optimal retirement age. In that respect, the key feature of our model is that labor supply is indivisible. As we know since Hansen (1985) and Rogerson (1988), this could lead agents to determine their labor supply from lotteries in order to convexify their production possibility set. However, this possibility does not apply to our framework which, following Diamond and Mirrlees (1978) or Mulligan (2001), could be seen as a "time averaging" model where agents can convexify their labor supply problem by alternating spells of work and leisure. More specifically, an agent will supply labor when his productivity is high and enjoy leisure, during retirement, once his productivity has deteriorated.

Finally, we conjecture that all the incentive compatibility constraints are binding. If they were not, then welfare could be improved by lowering the consumption level of the able. Our numerical implementation confirms that all the Lagrange multipliers are positive.

The consumption levels at $t$ of the able and disabled who became tagged at age $j$ are related by:

$$\frac{d}{dt} \frac{1}{u'(c^{AT}(t, j))} = \left[ \frac{1}{u'(c^{AT}(t, j))} - \frac{1}{u'(c^{DT}(t, j))} \right] \frac{f(t)}{1 - F(t)}. \quad (7)$$

As, once the tag has been awarded, the government does not have any further information on the health of the people, this condition corresponds to the original inverse Euler equation derived in Diamond and Mirrlees (1978). The general intuition for these inverse Euler equations is that, to preserve incentives to work, resources shifted to the next period must increase the utility in the good state, i.e. Able, as much as in the bad state, i.e. Disabled. Note that, as a result, more resources need to be allocated to the good state, where marginal utility is low, than to the bad state, where it is high. However, this transfer of utilities across time should be done at minimum cost to the

---

12 See Ljungqvist and Sargent (2006, 2008 and 2009) for detailed comparisons, and some equivalence results, between lotteries and time averaging models of indivisible labor.

13 Note that, even if workers have low productivity when young, we do not allow them to postpone entry into the labor market. One external justification for this is human capital accumulation, which makes early work at low productivity an investment into the future thanks to on-the-job learning effects. Hence, postponing entry does not increase the starting productivity of a worker and age 0 could be seen as a normalization of the age at which work begins.

14 See the appendix of Golosov and Tsyvinski (2004) for a formal proof in a simpler context à la Diamond-Mirrlees with unobservable health.

15 Remember that the incentive compatibility constraints are linear in utilities.
government. The expected resource cost of a marginal increase in utility at a given
time should therefore be equal to the expected resource cost of a marginal decrease in
utility at another time. But note that the inverse marginal utility of consumption is
precisely the increase in consumption associated to a given marginal increase in utility,
i.e. \(1/u'(c) = 1/(du/dc) = dc(u)/du\) where \(c(\cdot) \equiv u^{-1}(\cdot)\). This explains why
consumption should optimally follow an inverse Euler equation. Finally, the last difficulty
is that condition (7) is written in continuous time which implies that the terms are grouped in
a specific way. It says that the increase in the resource cost of a marginal postponement of
utility conditional on remaining in the good state, i.e. the left hand side, should exactly
compensate the expected drop in the marginal resource cost of utility associated with
a change of status from able to disabled, where the coefficient on the right gives the
probability density with which a tagged agent becomes disabled given that he was able
up to then.\(^{16}\) Note that the lower is this probability, i.e. the more unlikely it is that
an agent who worked until \(t\) truly becomes disabled at \(t\), the lower should
\(c_D^{DT}(t, j)\) be for a given path of \(c^{AT}(t, j)\). This improves incentives to work at little
cost in terms of insurance.

The boundary condition associated with (7) is:

\[
c^{AT}(RT(j), j) = c^{DT}_T(RT(j), j).
\]  

(8)

At age \(RT(j)\) the agent retires and, hence, consumption could be smoothed without
adverse incentive effects on labor supply.

The optimal retirement age \(RT(j)\) of an able worker who became tagged at age \(j\) solves:

\[
\frac{b}{u'(c^{AT}(RT(j), j))} = \gamma_{RT(j)}.
\]  

(9)

The agent keeps working until his marginal rate of substitution between leisure and
consumption equals his marginal product of labor. Indeed, the marginal utility cost of
working one more unit of time is \(b\) while the marginal product from doing so is \(\gamma_{RT(j)}\) at
age \(RT(j)\).

The consumption levels of the able and disabled of age \(i\), who are not tagged, are

\(^{16}\)All this might be simpler to see if condition (7) is written in terms of utilities:

\[
\frac{dc'(u^{AT}(t, j))}{dt} = \left[c'(u^{AT}(t, j)) - c'(u^{DT}_T(t, j))\right] \frac{f(t)}{1 - F(t)},
\]

where consumption levels are backed out from utilities using the function \(c(\cdot) \equiv u^{-1}(\cdot);\) for instance,
\(c(u^{AT}(t, j)) = u^{-1}(u^{AT}(t, j)) = c^{AT}(t, j)\) where \(u^{AT}(t, j)\) is the utility at age \(t\) of an able who became
tagged at \(j\).
related by:
\[
\frac{d}{dt} \frac{1}{u'(c^{AU}(i))} = \left[ \frac{1}{u'(c^{AU}(i))} - \frac{1}{u'(c^{DU}(t, i))} \right] \times \left[ 1 - G_D(\hat{\theta}_t) \right] f(i) \times \left[ 1 - G_D(\hat{\theta}_t) \right] [F(t) - F(i)] + \left[ 1 - G_A(\hat{\theta}_t) \right] [1 - F(t)],
\]
for any \( t \geq i \). The interpretation is similar to that of equation (7), except that the coefficient on the right stands for the probability density with which an agent became disabled at age \( i \) given that he was previously able and that he will only be tagged after \( t \). Note that the lower is this probability, i.e. the more unlikely it is that an agent truly became disabled at \( i \) given that he is still untagged at \( t \), the lower should \( c^{DU}(t, i) \) be. This new insight shows how the imperfect tag should be used in a dynamic setup to extract information on the true health status of individuals. The boundary condition associated with (10) is:
\[
c^{AU}(RU) = c^{DU}(t, RU), \forall t \in [RU, H].
\]
Again, there is nothing to be gained by distorting the consumption level of individuals after retirement.

Similarly, the consumption levels of the disabled and tagged who became disabled first and of the able and untagged are linked by:
\[
\frac{d}{dt} \frac{1}{u'(c^{AU}(i))} = \left[ \frac{1}{u'(c^{AU}(i))} - \frac{1}{u'(c^{DT}(i, j))} \right] \times g_D(\hat{\theta}_j) \frac{d\theta_j}{dj} f(i) \times g_A(\hat{\theta}_j) \frac{d\theta_j}{dj} [1 - F(j)] + \left[ G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j) + g_D(\hat{\theta}_j) \frac{d\theta_j}{dj} [F(j) - F(i)],
\]
where we must have \( j > i \). The coefficient on the right stands for the probability density with which an agent became disabled at \( i \) given that he was previously able and that he becomes tagged at \( j \). Again, the lower is this probability, i.e. the more unlikely it is that an agent truly became disabled at \( i \) given that he gets tagged at \( j \), the lower should \( c^{DT}(i, j) \) be. The corresponding boundary condition is:
\[
c^{AU}(RU) = c^{DT}(RU, j), \forall j \in [RU, H].
\]
Together with (11), this implies that being awarded the tag after retirement does not make any difference to those who worked until the maximum retirement age \( RU \).

Note that, for a given \( i \), the last two inverse Euler equations, (10) and (12), hold for
any \( t \geq i \) and any \( j > i \), respectively. This means that the expected drop in the marginal cost of providing utility induced by a change of status from able to disabled should be constant over time. Thus, the fact that the right hand side of (10) is equal for any \( t \geq i \) could be seen as another set of inverse Euler equations. Similarly for the right hand side of (12) which is independent of \( j \).

The consumption levels of the newly tagged, able and disabled, are related to that of the able and untagged by the following condition:

\[
\frac{1}{u'(c^{\text{AU}}(j))} = \frac{1}{u'(c^{\text{AT}}(j, j))} - \left[ \frac{1}{u'(c^{\text{AT}}(j, j))} - \frac{1}{u'(c^{\text{DT}}(j, j))} \right] \left[ G_D(\dot{\theta}_j) - G_A(\dot{\theta}_j) \right] f(j) \times \frac{g_A(\dot{\theta}_j) \frac{d\theta_j}{dj} [1 - F(j)] + \left[ G_D(\dot{\theta}_j) - G_A(\dot{\theta}_j) \right] f(j),
\]

where the coefficient on the right corresponds to the probability density with which an agent becomes disabled at age \( j \) given that he was previously able and that he becomes tagged at \( j \). This says that, at the optimum, the resource cost of a marginal increase in utility in the two states observed by the planner, i.e. tagged and untagged, should be equalized. Interestingly, although not dynamic, this condition, which was originally derived by Parsons (1996) in a simpler static context, relates inverse marginal utilities. This shows that the planner wants to equalize across time and states the marginal resource cost of providing utility to the agents. This general principle nests the standard inverse Euler equation derived by Diamond and Mirrlees (1978), condition (7), the first-order condition of Parsons (1996), condition (14), as well as the two first-order conditions which are specific to this paper, (10) and (12).

Finally, the first-order condition pinning down the optimal retirement age of the untagged is:

\[
b \frac{1}{u'(c^{\text{AU}}(RU))} = \gamma_{RU}.
\]

Again, as for condition (9), the interpretation is that, at the retirement age, the marginal rate of substitution between leisure and consumption should be equal to the marginal product of labor. Note that the formal derivation of (15) relies on the conjecture that \( \lim_{j \to RU} RT(j) = RU \) which is both intuitive and consistent with (15) together with (8), (9) and (14).

The Lagrange multiplier associated with the resource constraint is equal to \( u'(c^{\text{AU}}(RU)) \). The Lagrange multiplier of the incentive compatibility constraint of the newly tagged is

\[
u'(c^{\text{AU}}(RU)) \left[ \frac{1}{u'(c^{\text{AT}}(j, j))} - \frac{1}{u'(c^{\text{AU}}(j))} \right]
\]
and that of the previously tagged is\footnote{These multipliers are associated with constraint (A7) of the appendix for $s = j$ and $s > j$, respectively.}

\begin{equation}
    u'(c^{AU}(RU)) \frac{d}{dt} \frac{1}{u'(c^{AU}(t, j))}.
\end{equation}

Binding constraints imply that the multipliers are positive and, hence, that the consumption of the able who are tagged should initially be higher than that of the untagged and it should then be increasing over time. It is indeed common in dynamic contract theory that back-loaded incentives are optimal as they maintain incentives to work over time. Similarly, the Lagrange multiplier of the incentive compatibility constraint of the untagged is

\begin{equation}
    u'(c^{AU}(RU)) \frac{d}{dt} \frac{1}{u'(c^{AU}(t))},
\end{equation}

which implies that the consumption of the able and untagged should also be increasing over time.

We now have a full set of conditions determining the optimum allocation.

**Proposition 1** The optimal Social Security system with imperfect tagging is characterized by the first-order conditions (7), (8), (9), (10), (11), (12), (13), (14), (15) together with the resource constraint, (A4), the incentive compatibility constraints for the untagged, (A6), and for the tagged, (A7).

To gain additional insights about this Social Security system we need to perform a numerical simulation. But, before that, the model needs to be properly calibrated.

### 3 Calibration

This section describes the calibration of the distributions and parameters of the model. The discussion is divided into four parts: agents’ skill profile, their preferences, the distribution of the disability age and, finally, the trade-off between gaps and leakages.

#### 3.1 Skill profile

All individuals are assumed to enter the labor market at the age of 25 and die on their 80th birthday. Following Golosov and Tsyvinski (2006), productivity $\gamma_t$ at each age $t$ is determined by fitting a quadratic approximation through the data in Rios-Rull (1996). The resulting skill profile is characterized by a productivity of 1 at age 25 and 75, i.e. $\gamma_{25} = \gamma_{75} = 1$, and by a peak of 1.47 at age 50, i.e. $\gamma_{50} = 1.47$.\footnote{These multipliers are associated with constraint (A7) of the appendix for $s = j$ and $s > j$, respectively.}
3.2 Preferences

Agents are assumed to exhibit constant relative risk aversion so that:

$$u(c) = \frac{c^{1-\phi} - 1}{1 - \phi}. \quad (18)$$

We pick the coefficient of relative risk aversion $\phi = 2$. The annual discount rate $\rho$, which also equals the annual interest rate, is set at 0.02. The fixed cost of working $b$ is calibrated such that, in the unobservable health case, the able retire at age 65. This exercise yields $b = 1.092$.

3.3 Distribution of the disability age

To determine the likelihood of being disabled at age $t$, $F(t)$, we take cross-sectional data from the 2003 wave of the Panel Study of Income Dynamics (PSID) that surveys a representative sample of the U.S. population.\(^{18}\) We make use of the following question:\(^{19}\)

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

Specified answers are "yes" and "no"; accordingly we define any respondent who answers "yes" as disabled. At each age, the probability of being disabled is then set equal to the fraction of people answering "yes", using cross-sectional weights to correct for over- or under-representation of certain groups. The result is depicted in Figure 2. To obtain a smooth estimation of the disability distribution, we fit an exponential function through the resulting time series with the data points weighted by the number of observations for each age.

\(^{18}\)This is the same data source as used by Low and Pistaferri (2008). Other authors such as Benitez-Silva, Buchinsky and Rust (2006) chose to work with the Health and Retirement Study (HRS) instead. However, this is not an alternative for us as it only covers individuals over the age of 50.

\(^{19}\)As is the case with most other studies in the field, the underlying presumption here is that self-reported disability status is a valid measure of true disability status. This hypothesis finds support in Benitez-Silva, Buchinsky, Chan, Cheidvasser and Rust (2004).
At face value, our definition of disability may seem rather mild. However, it is to be stressed that, in our model, disability should not be interpreted too narrowly. Indeed, any individual whose productivity is virtually equal to zero should be considered as disabled. With, for instance, less than 40% of all 75-year-olds unable to work, it yields, if anything, numbers which are below what one might plausibly expect. Moreover, these figures are in line with those used in related papers (see, e.g., Golosov and Tsyvinski 2006).

### 3.4 Trade-off between gaps and leakages

The test outcome for both disabled, $g_D(\theta)$, and able, $g_A(\theta)$, individuals is assumed to be normally distributed with a difference in means equal to $\mu$ and a standard deviation of 1.\(^{20}\) Although the actual means of the two distributions are inconsequential (cf. footnote 8), for clarity, we adopt the normalization that they sum up to 0. Thus, the means of $g_A(\theta)$ and $g_D(\theta)$ are $\mu/2$ and $-\mu/2$, respectively.

To obtain an estimate of $\mu$, information is required on individuals’ ability to work, i.e. able or disabled, as well as their disability benefit status, tagged or untagged. For this, the disability data from above are combined with information on the sources of individuals’ revenue (which for 2003 are provided in the 2005 wave of the PSID). Everyone above the age of 65 is excluded from the sample, as these people have reached the full retirement

\(^{20}\)Alternatively, we could fix $\mu$ and calibrate the standard deviation. However, fixing the variance is particularly suitable in our context as with $\mu = 0$ the problem collapses to the unobservable health case treated by Diamond and Mirrlees (1978).
age and are shifted to the retired worker portion of the U.S. Social Security system.\footnote{Within the sample of people aged 25-65, a small proportion of individuals receive other types of Social Security benefits, such as retirement, survivor’s or dependent benefits. We exclude them on the grounds that the U.S. Social Security program may place disabled individuals with certain employment histories or family structures in a Social Security category other than disability benefits. Hence, we cannot know whether, absent these other benefits, they would get disability benefits.}

Disability benefit status, i.e. tagged or untagged, is a random variable following a Bernoulli distribution, where the probability of being tagged depends on an individual’s age $t \in \{25, ..., 65\}$ and on his ability to work. As an agent of age $t$ is awarded benefits when his test outcome is below $\hat{\theta}_t$, we have:

$$
\Pr(\text{Tagged}|\text{Age} = t, \text{Ability}) = \Phi \left( \sum_{s=25}^{65} \hat{\theta}_s \mathcal{I}(s = t) - \frac{\mu}{2} \mathcal{I}(\text{Able}) + \frac{\mu}{2} \mathcal{I}(\text{Disabled}) \right), \quad (19)
$$

where $\Phi(\cdot)$ is the c.d.f. of the standard normal distribution and $\mathcal{I}(\cdot)$ is the indicator function which is equal to 1 if the condition in brackets is satisfied and to 0 otherwise. Rearranging terms, a simple probit regression of disability benefit status on a set of age dummies and ability status can be employed to back out an estimate for $\mu$ and $\left\{\hat{\theta}_t\right\}_{t \in [25,65]}$. Doing so, we obtain $\mu = 1.2329$. As shown in Figure 3, the estimated path of the threshold, $\hat{\theta}_t$, is increasing with age. The McFadden’s pseudo $R^2$ for this regression is 19.9%.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{disability_standard.png}
\caption{Disability standard}
\end{figure}

Note that both our theoretical model and empirical strategy rely on the assumption that the difference in means, $\mu$, is the same at every age. To establish the validity of
this claim, we run a probit regression where \( \mu \) is allowed to be age-specific. It can be seen from Figure 4 that the resulting estimates do not exhibit any systematic pattern with respect to age. Indeed, when we test the hypothesis that \( \mu \) is constant, we obtain a p-value of 0.813.

![Figure 4: Difference in means](image)

4 Numerical results

In order to provide some quantitative insights about the optimal Social Security system with imperfect tagging, this section presents a numerical simulation of the model and an evaluation of the corresponding welfare gains. But, before turning to the results, we need to describe how the disability standard for each age \( t, \hat{t} \), is set.

4.1 Minimizing gaps and leakages

We consider the benchmark case where the path of the disability standard is set such as to minimize the total number of gaps and leakages, but allowing for a preference between the two. This preference is captured by defining a price of gaps, \( p_G \), and of leakages, \( p_L \), where, for instance, a higher price of gaps, i.e. \( p_G > p_L \), implies that gaps should be avoided more than leakages.

More formally, the disability standard is set by solving:

\[
\min_{\{\hat{t}_t\} \in [0,H]} \int_0^H \left\{ p_G F(t) \left[ 1 - G_D(\hat{t}_t) \right] + p_L \left[ 1 - F(t) \right] G_A(\hat{t}_t) \right\} dt, \tag{20}
\]
where \( F(t) \left[ 1 - G_D(\hat{\theta}_t) \right] \) and \( \left[ 1 - F(t) \right] G_A(\hat{\theta}_t) \) correspond to the total number of gaps and of leakages at age \( t \), respectively. In fact, this reduces to a static optimization problem for any given age, which yields the following first-order condition:

\[
p_G F(t) g_D(\hat{\theta}_t) = p_L \left[ 1 - F(t) \right] g_A(\hat{\theta}_t).
\]  

The marginal benefit from increasing \( \hat{\theta}_t \) is less gaps, the marginal cost more leakages. At the optimum, these, weighted by their respective prices, have to equate. Making use of the normality of the distribution of the test outcome, \( g_A(\theta) \) and \( g_D(\theta) \), we have:

\[
\hat{\theta}_t = \frac{1}{\mu} \ln \left( \frac{F(t)}{1 - F(t)} \right) + \frac{1}{\mu} \ln \left( \frac{p_G}{p_L} \right).
\]  

(22)

Recall that the probit regression (19) from the last section yields the age-specific estimates for the disability standards displayed in Figure 3. To see whether these are consistent with the minimization of gaps and leakages, we add an age-specific error term to (22) and run an OLS regression of \( \hat{\theta}_t \) on the fitted values\(^{22}\) of \( \ln \left( \frac{F(t)}{1 - F(t)} \right) \). We then test the hypothesis that the slope coefficient is equal to our previous estimate of \( 1/\mu = 1/1.2329 = 0.8111 \) and obtain a p-value of 0.028. In fact, the point estimate of the slope coefficient is 0.6907 which suggests that, to minimize the number of classification errors, the disability threshold should increase slightly more rapidly with age than it currently does. However, if we run a constrained regression, which imposes that the slope coefficient should be equal to \( 1/\mu = 1/1.2329 = 0.8111 \), we obtain the smooth line in Figure 3. As it provides a good fit to the empirically estimated \( \hat{\theta}_t \), we shall consider that the minimization of gaps and leakages is a good approximation to the current U.S. Social Security policy.

Finally, the constant coefficient of the constrained regression implies a relative price of gaps and leakages equal to 1.1998. Hence, in our subsequent evaluation of the welfare gains, we shall consider that the current disability standard in the U.S. is given by (22) with a relative price of gaps and leakages of 1.2.\(^{23}\)

The numerical simulation reported below assumes that the planner controls the relative price \( p_G/p_L \) and sets it to maximize welfare. A simple grid search reveals that the optimal price ratio is approximately equal to 2.5.

The minimization of gaps and leakages corresponds to a natural benchmark where the government makes a non-strategic use of its imperfect information on health. Furthermore, several arguments may be advanced in support of such a policy being constrained

---

\(^{22}\)We use the smoothed representation of \( F(t) \) as displayed in Figure 2 since the decision to award the tag should be based on the disability distribution prevailing in the entire population.

\(^{23}\)This measure gives an idea about the total number of tagged individuals in the population. Since take-up is not systematic, it is not readily comparable to other estimates found in the literature which are exclusively based on the applicants to disability insurance.
optimal. For one, the government might not be able to directly control doctors because their professional ethics may dictate them that they should make as few classification errors as possible. If so, the role of the government will be reduced to specifying the relative importance of gaps and leakages. Alternatively, one may think that the only tagging policy that is politically acceptable is one that minimizes gaps and leakages.

4.2 Numerical simulation

All numerical simulations are achieved by solving a discretized version of the system of equations which characterizes the optimal allocation. The disability standard used for the reported simulation is determined from (22) with $\frac{p_G}{p_L} = 2.5$.

The consumption of the able and untagged, $c^{AU}(t)$, is plotted in Figure 5. Increasing consumption with age renders incentives back-loaded. This has the dual advantage of not only inducing the old and able to work, but also the young and able since by working they maintain the prospect of high consumption when old. As previously discussed, this consumption pattern is imposed by the incentive compatibility constraint for the untagged.

![Figure 5: Consumption of the able and untagged](image)

The maximum retirement age of the economy, that of the able and untagged, $RU$, is 67.3 years. This is relatively high compared to the corresponding age of 65 prevailing with unobservable health. In fact, with partially observable health the consumption level needed to induce the able and untagged to work is not so high. As a result, their marginal
rate of substitution between leisure and consumption is relatively low and it is optimal to let them retire rather late.

Figure 6 depicts $c^{DU}(t, r)$, the consumption of a disabled and untagged individual as a function of his current age $t$ and of the age $r$ at which he ceased to work\(^{24}\) (henceforth, "disability age"), with $t \geq r$. Once an untagged agent has become disabled, his consumption is falling with age and is minimal at $H$. To understand this pattern, which follows from condition (10), note that the planner wants to give high consumption to the truly disabled while deterring the able from claiming to be unable to work. To find the best compromise between these two goals, the planner exploits the fact that a truly disabled is unlikely to remain untagged for long. Thus, consumption is initially high to provide insurance. It then decreases over time as this lower consumption is unlikely to affect the truly disabled but would be likely to apply to an able person who claimed to be disabled. The very low consumption levels near $H$ serve as a threat and are therefore not welfare reducing.

Figure 6: Consumption of the disabled and untagged

Figure 7 gives the consumption of an able and tagged as a function of his current age $t$ and of the age $j$ at which he became tagged (henceforth, "tag age"), with $t \geq j$. For any given tag age, consumption is increasing over time. Again, the need to maintain incentives to work now and in the future makes back-loaded incentives particularly attractive.

\(^{24}\)Remember that individuals stop to work either when they become disabled or when they reach the retirement age. In this last case, from the perspective of the mechanism design problem, they can be considered as disabled from this retirement age onwards.
Figure 7: Consumption of the able and tagged

Figure 8 shows the retirement age of the able and tagged, $RT(j)$, as a function of the age $j$ at which the tag was awarded. The informative nature of the tag implies that the proportion of disabled will always be higher among the tagged than among the untagged. Higher consumption should therefore be provided to the disabled and tagged which means that even higher consumption is needed to induce the able and tagged to work. But this increases their marginal rate of substitution between leisure and consumption. It is therefore not surprising that the optimal retirement age for all tagged is lower than that of the untagged.
To understand why the retirement age is a U-shaped function of the tag age, recall from condition (14) that the expected marginal resource cost of providing utility should be the same whether the agent is newly tagged or untagged. But, initially, the tagged are very likely to be able to work and, hence, $c^{AT}(j, j)$ should follow the shape of $c^{AU}(j)$, i.e. they are both increasing in $j$. But this makes back-loaded incentives so costly that it is optimal to reduce the retirement age. Later, when the tagged are more likely to be truly disabled, with age rising, the increase in $c^{DT}(r, j)$ also contributes to match the increase in $c^{AU}(j)$. Hence, the increase in $c^{AT}(j, j)$ can be kept smaller, making back-loaded incentives cheaper and allowing the retirement age to be raised. This intuition concurs with the concave shape of $c^{AT}(j, j)$, which is apparent along the diagonal in Figure 7. Note that a reasonable approximation of the optimal policy might be to implement an early retirement age of 62 for all those who got tagged before 57.

Figure 9 shows $c^{DT}(r, j)$, the consumption of the disabled and tagged who ceased to work at $r$ and became tagged at $j$. Two sections are clearly distinguishable: $c^{DT}_T(r, j)$, $r \geq j$, on the left and $c^{DT}_D(r, j)$, $j > r$, on the right. This discontinuity is due to the incentive compatibility constraint for the tagged which only applies on the left. It should be emphasized that, while previous graphs were displaying instantaneous consumption levels, this one reports permanent consumption levels. Indeed, individuals consume $c^{DT}(r, j)$ from $\max\{r, j\}$ until they die at $H$. 
As argued above, it is desirable to provide back-loaded incentives to the able and tagged. But, having an increasing consumption level for the able is not the only way to do so. Alternatively, the consumption of the disabled could be made higher, the later they cease to work. This explains why, for a fixed tag age \(j\), \(c^D_T(r, j)\) is increasing in \(r\).

For an individual who is disabled and untagged, consumption after retirement will be lower the later he becomes tagged. This follows from (12). The intuition for this is similar to that for \(c^{DU}(t, r)\). If someone is truly disabled, he is likely to be awarded the tag shortly after stopping to work. In this case, the insurance motive commands a high consumption level. A low consumption level for the disabled who only get tagged much later serves as a threat to the able and untagged who might be tempted to deviate.

Turning to the diagonal of Figure 9, it is apparent that a higher consumption level is awarded if disability occurs before the award of the tag. To understand this, note that a newly tagged worker who deviates gets consumption \(c^D_T(\cdot)\) immediately, while an untagged worker who deviates initially obtains \(c^{DU}(\cdot)\) and is only likely to rapidly qualify for \(c^D_D(\cdot)\) if he is truly disabled. Thus, \(c^D_I(\cdot)\) could be made higher than \(c^D_T(\cdot)\) while still inducing the able to work.

It can be checked that the only situation where agents are not happy to be tagged as soon as they become eligible, is when disability and eligibility occur simultaneously. The solution to this problem is to impose a compulsory health check to individuals who have just become disabled. For this solution to work, the outcome \(\theta\) of the test for a given individual should be exogenous to his action. A (computationally feasible) alternative
would be to impose additional constraints to the planner’s problem ensuring that indi-
viduals are always happy to be awarded the tag as soon as they are eligible. This would 
eliminate the discontinuity of $c^{DT}(\cdot)$. However, imposing extra constraints does not seem 
essential and would come at the cost of reduced welfare.

### 4.3 Welfare gains

Our numerical simulations allow us to evaluate the welfare associated with the optimal 
policy. To get an idea about the gains generated by imperfect tagging, we take the 
unobservable health case, analyzed by Diamond and Mirrlees (1978) and Golosov and 
Tsyvinski (2006), as the reference. We also consider the first-best allocation which gives 
us an upper bound to the welfare gains that could be obtained.

A key characteristic of the Social Security system that we propose is that it implements 
a health-dependent retirement age.\(^{25}\) In order to assess the importance of this feature, we 
also compute the welfare obtained when the retirement age of the able has to be the same 
for all. More formally, the planner’s problem remains the same except that we impose 
$RU = RT(j) \equiv R, \forall j \in [0, H]$. The optimal retirement age is then pinned down by the 
following condition,

$$
\begin{align*}
\frac{b}{1 - G_A(\hat{\theta}_R)} + \int_0^R \frac{g_A(\hat{\theta}_j) \frac{d\theta_j}{dR}}{u'(c^{AU}(R))} dj & = \gamma_R, \\
\end{align*}
$$

which replaces (9) and (15). A weighted average of the marginal rates of substitution 
between leisure and consumption should be equal to the marginal rate of transformation.

A policy yields welfare gains of $x\%$ if its level of welfare can be matched in the 
unobservable health case by proportionally increasing consumption by $x\%$ in every state 
of the world. The results are reported in the following table.

<table>
<thead>
<tr>
<th></th>
<th>Fixed retirement age</th>
<th>Health-dependent retirement age</th>
<th>First-best</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_G/p_L = 2.5$</td>
<td>0.45%</td>
<td>0.64%</td>
<td>2.98%</td>
</tr>
<tr>
<td>$p_G/p_L = 1.2$</td>
<td>0.41%</td>
<td>0.56%</td>
<td>2.98%</td>
</tr>
</tbody>
</table>

In the first line the planner sets the optimal price of gaps and leakages.\(^{26}\) If, however, 
doctors are out of control and the government has to stick with the current disability

---

\(^{25}\)Again, it should be stressed that the retirement age is dependent on health as observed by the 
government but that it only applies to the able, who are, by definition, in good health.

\(^{26}\)Note that the optimal relative price with a fixed retirement age is also approximately equal to 2.5.
standards, then the relevant results are that of the second line. The welfare gains generated by the imperfect information on health are moderate but non-negligible. More than two thirds of these gains could be reaped with a fixed retirement age.

Clearly, from equation (22), as most people are able to work, the disability standard is quite low when almost equal weights are put on gaps and leakages, i.e. when $p_G/p_L = 1.2$. This implies that few people are tagged and, hence, only a limited use of the imperfect information on health could be made. This explains why the corresponding welfare gains are larger with $p_G/p_L = 2.5$.

The welfare improvements generated by the optimal policy could come from two sources: improved insurance against the disability risk or improved incentives to work. The following statistics on the average retirement age, for the case $p_G/p_L = 2.5$, suggest that at least some of the gains come from better incentives to work.

<table>
<thead>
<tr>
<th>Table 2: Retirement age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unobservable health</td>
</tr>
<tr>
<td>Average retirement age</td>
</tr>
<tr>
<td>Maximum retirement age</td>
</tr>
</tbody>
</table>

The average retirement age is the average age at which people cease to work, conditional on being able at 25. In all four scenarios, almost a quarter of the population retires as disability occurs. In the health-dependent retirement age case, about two thirds of the remaining three quarters of the population reach the maximum retirement age $RU$, which is smaller than the first-best retirement age as relatively high consumption is needed to induce the able and tagged to work until $RU$.

We have so far focused on the, rather theoretical, unobservable health benchmark. While the current U.S. Social Security system already uses imperfect information on health, one of the key differences between the planner’s policy and that observed in the U.S. is that the able and tagged are currently not incentivized to work.\textsuperscript{27} To evaluate the welfare gains generated by this feature of the optimal policy, we solved a modified planner’s problem where the constraint $RT(j) = j$, for all $j \in [0, RU)$, is added.

Compared to the unobservable health case, the optimal policy under the constraint that all tagged retire immediately, yields a welfare gain of 0.46% when $p_G/p_L = 1.2$.\textsuperscript{28}

\textsuperscript{27}The UK has recently experimented with a policy, Pathways to Work, encouraging employment among disability recipients. Preliminary evaluations suggest very high returns on investment both to the beneficiaries and to the taxpayer (Adam Bozio Emmerson Greenberg Knight 2008). However, a similar policy in the U.S., Ticket to Work, failed to increase participation (Autor Duggan 2006, 2007).

\textsuperscript{28}The optimal relative price with immediate retirement of all tagged is $p_G/p_L = 0.9$. The corresponding welfare gain, compared to unobservable health, is 0.47%.
Using the numbers from Table 1, it follows that, with the current disability standard unchanged, the gains from inducing the able and tagged to work are small with a health-dependent retirement age, about 0.10%, and negative with common retirement age for all, about -0.05%. In this latter case, the costs of inducing work until the general retirement age are so large that they more than absorb all the benefits from encouraging work in the first place. This shows that inducing the able and tagged to work is only desirable up to a point, i.e. up to an early retirement age.

If the optimal relative price of gaps and leakages of 2.5 could be enforced, then the optimal policy is associated with a welfare gain of 0.18% compared to the immediate retirement of the tagged policy. It is therefore desirable to decrease the strictness of the disability test but, crucially, the able and tagged should be induced to work. Indeed, with $p_G/p_L = 2.5$, the policy of immediate retirement of the tagged generates a welfare loss of 0.45% compared to the unobservable health case. This illustrates the possibility that no information on health could be preferable to some badly used information. The problem with $p_G/p_L = 2.5$ when $RT(j) = j$ is that about 30% of the population retires when awarded the tag. To compensate the sharp reduction in labor supply that this entails, the general retirement age needs to be pushed up to 72.1, which results in an average retirement age of only 61.0.

In addition to the 0.18% that could be gained by inducing the able and tagged to work, another major welfare enhancing change recommended by the optimal policy consists in making a more strategical use of the gap in timing between the occurrence of disability and the award of the tag. However, lacking a good benchmark representation of the current U.S. situation, the corresponding welfare gains are harder to evaluate.

5 First-best implementation

We have so far considered the optimal Social Security system when the government chooses a path of $\hat{\theta}_t$ that minimizes the total number of classification errors but allowing for different prices of gaps and leakages. Although this is a rather natural choice for the disability threshold, we might be interested in determining the optimal allocation when $\hat{\theta}_t$ is under the control of the planner. In fact, it turns out to be possible to implement the first-best, perfect information, allocation asymptotically by setting the thresholds \$\hat{\theta}_t\$ \$t \in [0,H]\$ strategically. Remember that in a first-best allocation perfect insurance is provided and, hence, all agents enjoy a constant consumption stream, $c^{FB}$, while the able

\[29\] A number of other studies on the topic, such as Low and Pistaferri (2008), have reached the opposite conclusion that the strictness of the test should be increased. However, these only consider a \textit{ceteris paribus} change in the disability standard, while we simultaneously allow for other changes to the current U.S. policy such as increased incentives to work for the tagged. As implied by the previous footnote, without such changes the disability standard should indeed be decreased slightly.
To prove that such an allocation can be asymptotically implemented, we propose a policy that does the job.\footnote{The precise characterization of such a policy, and in particular of the optimal path of $\hat{\theta}_t$, is not unique. However, the underlying logic is always the same.} The planner should optimally award the tag as follows:

$$
\hat{\theta}_t = \begin{cases} 
-\infty & \text{if } t \in [0, R^{FB}) \\
\hat{\theta} & \text{if } t = R^{FB} \\
+\infty & \text{if } t \in (R^{FB}, H] 
\end{cases}, \tag{24}
$$

where $\hat{\theta}$ is a constant to be determined. Hence, the only uncertainty is whether people get tagged at the general retirement age, $R^{FB}$, or immediately after. Using this simple device, it is possible to deter deviations by setting consumption appropriately. In particular, we set:

$$
c^{AU}(t) = c, \forall t \in [0, R^{FB}), \tag{25}
c^{DU}(t, r) = c, \forall r \in [0, R^{FB}), \forall t \in [r, R^{FB}] \tag{26}
c^{DT}(r, j) = \begin{cases} 
\delta & \text{if } r \in [0, R^{FB}) \text{ and } j > R^{FB} \\
c & \text{otherwise} 
\end{cases}, \tag{27}
$$

for some constant $c$ and $\delta$. The consumption of the able and tagged is irrelevant and does not need to be specified as people can only get tagged after retirement. Note that the consumption level $\delta$ only applies to those who retired before $R^{FB}$, who therefore claimed, rightly or wrongly, to be disabled, and who failed to get tagged at $R^{FB}$. But, thanks to the monotone likelihood ratio property satisfied by $g_{A}(\theta)$ and $g_{D}(\theta)$, for a sufficiently high threshold $\hat{\theta}$, it is almost exclusively able people who fail to get tagged at $R^{FB}$. Thus, if they claimed to be disabled before $R^{FB}$, it is possible to punish a random subset of them by setting a sufficiently low value of $\delta$.

**Proposition 2** A policy characterized by (24), (25), (26) and (27) could be used to implement, asymptotically, the first-best allocation of resources. For that, choose $\delta$, as a function of $c$, to be the highest value such that all the incentive compatibility constraints of the untagged are satisfied. The consumption level $c$ should then be determined from the resource constraint. The first-best allocation obtains as $\hat{\theta} \to +\infty$, which implies $\delta \to 0$ and $c \to c^{FB}$.

In a nutshell, the optimal policy is to shoot the liars. In particular, it should be emphasized that the low value of $\delta$ is not welfare reducing as it is essentially off the equilibrium path. Note that every eligible person is trivially happy to be awarded the tag. Also, in this context, there is nothing to be gained from a health-dependent age of retirement.
The reason why the first-best allocation can only be implemented asymptotically is that $g_A(\theta)$ and $g_D(\theta)$ have the same support. Thus, no matter the severity of the test, the government can never be entirely sure that someone untagged is able to work. If, on the contrary, the upper limit of the support of $g_D(\theta)$, say $\bar{\theta}_D$, is lower than that of $g_A(\theta)$, then the first-best policy can be exactly implemented by setting $\bar{\theta} = \bar{\theta}_D$, $\delta = 0$ and $c = c^{FB}$. In other words, if there exists a disability test which only able people could fail, then the optimal policy is to shoot the previously allegedly disabled who fail the test at age $R^{FB}$.

An interesting feature is that the first-best allocation can always be asymptotically implemented, independently of the quality of the information on health. In terms of our previous calibration, where $g_A(\theta)$ and $g_D(\theta)$ are both assumed to be normal, all that is required is that the difference in means be strictly positive, i.e. $\mu > 0$. More generally, this shows that a small departure from the assumption of unobservable skills, which is pervasive in New Dynamic Public Finance, could have considerable consequences for the determination of the optimal policy.

Proposition 2 is reminiscent of a similar result derived by Mirrlees (1974, 1999) in the context of moral hazard. While the formal, mathematical, argument is very similar, it is interesting to note that this result is applicable to a hidden information framework in which the private information, on health, is partially observable by the government.

It should be emphasized that the first-best implementation heavily relies on the assumption that workers believe that their probability of being awarded the tag, conditional on remaining able up to age $R^{FB}$, is $G_A(\theta)$. In other words, they do not have any private information about when they might become eligible. While, as a first-order approximation to reality, this assumption is reasonable, a small departure from it could have important consequences when implementing the, extreme, first-best policy. Indeed, an able individual whose apparent health is already very bad at age 50 might be tempted to deviate being confident that he will get tagged at $R^{FB}$.

While it might not be reasonable to believe in the practical relevance of the first-best policy, the result nevertheless suggests that the government can obtain substantial welfare gains by moving beyond the minimization of gaps and leakages. For instance, if the disability threshold was increasing even more rapidly with age than it currently does, then the tag would often be awarded late in life. This would be welfare enhancing as the threat of not being tagged when old deters the temptation to claim to be disabled when young while few young and able workers would be tagged which makes it unnecessary to give them special rewards for participating to the labor market.

---

31 See also Varian (1980).

32 Note that this is equivalent to raising the price of gaps relative to that of leakages as age increases.
6 Conclusion

In this paper, we have characterized, within a general framework, the optimal Social Security system in a dynamic setting with imperfectly observable health. In order to induce the able to work, while providing insurance to the truly disabled, the planner offers back-loaded incentives and makes a strategic use of the difference in timing between the occurrence of disability and the award of the tag. The able who are tagged should be encouraged to work. But, as they are eligible for generous disability benefits, it is necessary to give them higher consumption and higher pensions than if they were untagged. It is therefore also desirable to let them retire earlier than others. Indeed, our simulation finds a general retirement age of 67.3 for the untagged and close to 62 for those tagged before age 57.

In many industrialized countries, both disability insurance and pension programs are subject to financial distress. It is commonly argued that the strictness of the disability test should be raised, to deal with the former problem, and that the statutory retirement age should be increased, to deal with the latter. A different solution emerges when the two problems are treated jointly rather than in isolation. To increase labor supply, the key is to offer the able and tagged proper incentives to work until some early retirement age. This would even make it desirable to decrease the strictness of the test which, by reducing the number of gaps, would improve the provision of insurance to the truly disabled. Moreover, additional welfare gains could be obtained by moving beyond the minimization of classification errors and by setting the disability standard and consumption levels strategically.

In this paper, we have derived the optimal incentive-feasible allocation by relying on the revelation principle. It would now be very interesting to know how it could be implemented in a decentralized economy with private capital markets. Golosov and Tsyvinski (2006) showed that asset-testing could be used to implement the optimal allocation with unobservable health. Things might not be as trivial with imperfect tagging. If the policy instruments needed for implementation turn out to be excessively complex, then implementation constraints might have to be added to the planner’s problem. Diamond and Mirrlees (1986) show a potentially useful direction by solving the same problem as in their previous paper but imposing that the consumption of the able should be constant over time, reflecting the impossibility of implementing age-dependent payroll taxes.

References


A The planner’s problem

As explained in the text, cf. equation (6), the planner maximizes the \textit{ex-ante} expected lifetime utility of agents subject to a resource constraint and to a set of incentive compatibility constraints which insures that those who are able to work choose to do so until they reach the relevant retirement age. In this appendix, we give explicitly the equations of the planner’s problem. His objective, which could be derived using (1), (2), (3), (4), (5) and (6), is to maximize:

\[
E[v(i, j)] = \int_0^{RU} e^{-\rho t} \left[u(c_{AT}(t)) - b\right] \left[1 - G_A(\theta_i)\right] [1 - F(t)] dt \\
+ \int_0^{RT(0)} \int_j e^{-\rho t} \left[u(c_{AT}(t, j)) - b\right] g_A(\theta_j) \frac{\partial \theta_j}{\partial j} [1 - F(t)] dt dj \\
+ \int_0^{RT(0)} e^{-\rho t} \left[u(c_{AT}(t, 0)) - b\right] G_A(\theta_0) [1 - F(t)] dt \\
+ \int_0^{RT(0)} \int_i e^{-\rho t} u(c_{DU}(t, i)) \left[1 - G_D(\theta_i)\right] f(i) dt di \\
+ \int_0^{H} e^{-\rho t} u(c_{DU}(t, RU)) \left[1 - G_D(\theta_i)\right] [1 - F(t)] + \left[1 - G_D(\theta_i)\right] [F(t) - F(RU)] dt \\
+ \int_0^{H} \int_i \int_j e^{-\rho t} u(c_{DT}(i, j)) g_D(\theta_j) \frac{\partial \theta_j}{\partial j} f(i) dj di \\
+ \int_0^{H} \int_j e^{-\rho t} u(c_{DT}(RU, j)) \left[g_A(\theta_j) \frac{\partial \theta_j}{\partial j} [1 - F(j)] + \left[G_D(\theta_j) - G_A(\theta_j)\right] f(j) + g_D(\theta_j) \frac{\partial \theta_j}{\partial j} [F(j) - F(RU)]\right] dj \\
+ \int_0^{RU} \int_j e^{-\rho t} u(c_{DT}(i, j)) g_A(\theta_j) \frac{\partial \theta_j}{\partial j} f(i) dj \\
+ \int_0^{RU} \int_j e^{-\rho t} u(c_{DT}(i, 0)) G_A(\theta_0) f(i) di \\
+ \int_0^{RU} \int_j e^{-\rho t} u(c_{DT}(RT(j), j)) g_A(\theta_j) \frac{\partial \theta_j}{\partial j} [1 - F(RT(j))] dj \\
+ \int_0^{H} e^{-\rho t} u(c_{DT}(RT(0), 0)) G_A(\theta_0) [1 - F(RT(0))] \\
+ \int_0^{RU} \int_j e^{-\rho t} u(c_{DT}(j, j)) \left[G_D(\theta_j) - G_A(\theta_j)\right] f(j) dj. \tag{A1}
\]

Note that, when deriving this expression, care should be taken of the fact that \(f(i, j)\) is not a standard probability density function. In particular, a mass of agents become
disabled and tagged simultaneously. This justifies the existence of a specific term, i.e. the last term of (A1), corresponding to these people.

The resource constraint could be derived in the same way. Let $z(i, j)$ stand for the lifetime budget deficit generated by individual $(i, j)$. The counterpart to equation (4), for $\min \{i, RU\} < j$, is:

$$ z(i, j) = \int_0^{\min \{i, RU\}} e^{-pt} \left[ c^U(t) - \gamma_i \right] dt + \int_{\min \{i, RU\}}^j e^{-pt} c^U(t, \min \{i, RU\}) dt + \int_j^H e^{-pt} c^D_D(t, \min \{i, RU\}, j) dt. $$

(A2)

Similarly, the counterpart to (5), for $j \leq \min \{i, RU\}$, is:

$$ z(i, j) = \int_0^j e^{-pt} \left[ c^U(t) - \gamma_i \right] dt + \int_j^{\min \{i, RT(j)\}} e^{-pt} \left[ c^T(t, j) - \gamma_i \right] dt + \int_j^H e^{-pt} c^D_D(t, \min \{i, RT(j)\}, j) dt. $$

(A3)

Thus, the resource constraint is:

$$ E[z(i, j)] = \int_0^H \int_0^H z(i, j) f(i, j) didi dj \leq 0. $$

(A4)
The full expression is:

\begin{align}
&\int_0^R U e^{-pt} \left[ c_{AU}(t) - \gamma_I \right] \left[ 1 - G_A(\hat{\theta}_I) \right] [1 - F(t)] dt \\
&+ \int_0^R \int_j^{RT(j)} e^{-pt} \left[ c_{AT}(t, j) - \gamma_I \right] g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(t)] dt dj \\
&+ \int_0^{RT(0)} e^{-pt} \left[ c_{AT}(t, 0) - \gamma_I \right] G_A(\hat{\theta}_0) [1 - F(t)] dt \\
&+ \int_0^R \int_i^H e^{-pt} c_{DU}(t, i) \left[ 1 - G_D(\hat{\theta}_i) \right] f(i) dt di \\
&+ \int_j^H e^{-pt} c_{DU}(t, RU) \left[ \left[ 1 - G_A(\hat{\theta}_I) \right] [1 - F(t)] + \left[ 1 - G_D(\hat{\theta}_i) \right] [F(t) - F(RU)] \right] dt \\
&+ \int_0^R \int_i^H \left[ \int_j^H e^{-pt} dt \right] c_{DT}^{DU}(i, j) g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) dj di \\
&+ \int_0^H \left[ \int_j^H e^{-pt} dt \right] c_{DT}^{DU}(RU, j) \\
&\quad \times \left[ g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + \left[ G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j) + g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [F(j) - F(RU)] \right] dj \\
&+ \int_0^R \int_j^{RT(j)} \left[ \int_i^H e^{-pt} dt \right] c_{DT}^{DU}(i, j) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) dj di \\
&+ \int_0^{RT(0)} \left[ \int_i^H e^{-pt} dt \right] c_{DT}^{DU}(i, 0) G_A(\hat{\theta}_0) f(i) di \\
&+ \int_0^H \left[ \int_{RT(j)}^H e^{-pt} dt \right] c_{DT}^{DU}(RT(j), j) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(RT(j))] dj \\
&+ \int_0^{RT(0)} \left[ \int_j^H e^{-pt} dt \right] c_{DT}^{DU}(RT(0), 0) G_A(\hat{\theta}_0) [1 - F(RT(0))] \\
&+ \int_0^H \left[ \int_j^H e^{-pt} dt \right] c_{DT}^{DU}(j, j) \left[ G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j) dj \leq 0.
\end{align}
The incentive compatibility constraint inducing the able and untagged of age \( s \in [0, RU) \) to work is:

\[
\int_s^{RU} e^{-pt} \left[ u(c^{AU}(t)) - b \right] \left[ 1 - G_A(\hat{\theta}_t) \right] [1 - F(t)] \, dt \\
+ \int_s^{RU} \int_j^{RT(j)} e^{-pt} \left[ u(c^{AT}(t, j)) - b \right] g_A(\hat{\theta}_j) \frac{\partial \hat{\theta}_j}{\partial j} [1 - F(t)] \, dt \, dj \\
+ \int_s^{RU} \int_{j}^{H} e^{-pt} u(c^{DU}(t, i)) \left[ 1 - G_D(\hat{\theta}_i) \right] f(i) \, dt \, di \\
+ \int_H^{RU} e^{-pt} u(c^{DU}(t, RU)) \left[ \left[ 1 - G_A(\hat{\theta}_t) \right] [1 - F(t)] + \left[ 1 - G_D(\hat{\theta}_t) \right] [F(t) - F(RU)] \right] \, dt \\
+ \int_s^{RU} \int_{j}^{H} \left[ \int_{j}^{H} e^{-pt} dt \right] u(c^{DT}(i, j)) g_D(\hat{\theta}_j) \frac{\partial \hat{\theta}_j}{\partial j} f(i) \, dj \, di \\
+ \int_s^{RU} \int_{j}^{H} e^{-pt} \left[ \int_{j}^{H} \left[ \int_{j}^{H} e^{-pt} dt \right] u(c^{DT}(i, j)) g_A(\hat{\theta}_j) \frac{\partial \hat{\theta}_j}{\partial j} [1 - F(j)] \right] \, dj \\
+ \int_s^{RU} \int_{j}^{RT(j)} \left[ \int_{j}^{H} e^{-pt} dt \right] u(c^{DT}(RT(j), j)) g_A(\hat{\theta}_j) \frac{\partial \hat{\theta}_j}{\partial j} [1 - F(RT(j))] \, dj \\
+ \int_s^{RU} \int_{j}^{H} e^{-pt} \left[ \int_{j}^{H} e^{-pt} dt \right] u(c^{DT}(j, j)) \left[ G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j) \, dj \\
\geq \int_s^{H} e^{-pt} u(c^{DU}(t, s)) \left[ \left[ 1 - G_A(\hat{\theta}_t) \right] [1 - F(t)] + \left[ 1 - G_D(\hat{\theta}_t) \right] [F(t) - F(s)] \right] \, dt \\
+ \int_s^{H} \left[ \int_{j}^{H} e^{-pt} dt \right] u(c^{DT}(s, j)) \\
\times g_A(\hat{\theta}_j) \frac{\partial \hat{\theta}_j}{\partial j} [1 - F(j)] + \left[ G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j) + g_D(\hat{\theta}_j) \frac{\partial \hat{\theta}_j}{\partial j} [F(j) - F(s)] \, dj.
\]  

The left hand side just corresponds to the expected utility of an able and untagged of age \( s \). It could be written more compactly as \( E \left[ v(i, j) \right] i > s, j > s \) \( \equiv \int_s^{H} \int_s^{H} v(i, j) f(i, j) \, dj \). The right hand side is equal to the expected utility that these agents would obtain by deviating and claiming to be disabled. They would get consumption \( c^{DU}(t, s) \) until tagged at \( j > s \) and \( c^{DT}(s, j) \) thereafter.
Finally, the able who obtained the tag at age $j$ should be incentivized to work until age $RT(j)$. The corresponding constraint at age $s \geq j$ is:

$$
\int_s^{RT(j)} e^{-\rho t} \left[ u(c_{AT}(t, j)) - \bar{b} \right] [1 - F(t)] dt \\
+ \int_s^{RT(j)} \left[ \int_i^H e^{-\rho t} dt \right] u(c_{DT}(i, j)) f(i) di \\
+ \left[ \int_{RT(j)}^H e^{-\rho t} dt \right] u(c_{DT}(RT(j), j)) [1 - F(RT(j))] \\
\geq \left[ \int_s^H e^{-\rho t} dt \right] u(c_{DT}(s, j)) [1 - F(s)].
$$

(A7)

Since, once an agent is tagged, the government cannot rely on any additional information about his health, this constraint is formally identical to the incentive compatibility constraint imposed in Diamond and Mirrlees (1978).

The planner’s problem is to maximize (A1) with respect to $c_{AU}(\cdot), c_{AT}(\cdot), c_{DU}(\cdot), c_{DT}(\cdot), c_{DT}(\cdot), RT(\cdot)$ and $RU$ subject to (A5), (A6) $\forall s \in [0, RU)$ and (A7) $\forall j \in [0, RU), \forall s \in [j, RT(j)]$. 

42