



ÉCOLE POLYTECHNIQUE

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

OPTIMAL SOCIAL SECURITY WITH IMPERFECT TAGGING

Oliver DENK
Jean-Baptiste MICHAU

Cahier n° 2013-04

DEPARTEMENT D'ECONOMIE

Route de Saclay
91128 PALAISEAU CEDEX
(33) 1 69333033

<http://www.economie.polytechnique.edu/>
<mailto:chantal.poujouly@polytechnique.edu>

Optimal Social Security with Imperfect Tagging*

Oliver Denk
OECD

Jean-Baptiste Michau[†]
Ecole Polytechnique

February 2013

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 against disability. After deriving the first-order conditions to this problem, we calibrate the model to the U.S. economy and run a numerical simulation. The government should offer back-loaded incentives and make strategic use of the difference 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 can asymptotically be implemented by making 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 Joseph Altonji, Nick Barr, Tim Besley, Raj Chetty, David Cutler, Giulio Fella, Mikhail Golosov, Henrik Kleven, Guy Laroque, Emmanuel Saez, Jeremy Sandford, Monica Singhal and to seminar participants at Berkeley, CREST, Harvard and the London School of Economics for useful comments and suggestions. The findings, interpretations and conclusions expressed in the paper are those of the authors and do not necessarily represent the views of the OECD.

[†]Corresponding author: jean-baptiste.michau@polytechnique.edu.

1 Introduction

Two of the most pressing issues in public finance across industrialized nations are the rising costs of providing disability insurance and pensions 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 and Duggan, 2006; Li and Maestas, 2008). Furthermore, both the disability insurance and pension programs can be seen as complementary ways of providing insurance to workers against the risk that they lose their ability to earn income from labor. Indeed, in the U.S., these two programs are the 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 1889.¹

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 benefits² 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. Most European countries have even larger disability insurance programs (as a share of their GDP). Despite these gigantic numbers and the potentially large welfare implications of the disability risk (Chandra and Samwick, 2006), very little is known about the optimal design of insurance against this risk.

To approach this problem, we rely on a framework that incorporates two important dimensions. First, we have a dynamic setup in order to capture the fact that disability could potentially hit any worker at any age. Second, we assume that medical impairments can only be imperfectly observed by the government. This assumption is hard to dispute given that almost 70% of workers receiving disability insurance payments fall into the following three categories: mental disorders (33.4%), musculoskeletal system and connective tissue, e.g. back pain, (26.4%), and nervous system and sense organs (9.5%). Our contribution is therefore to characterize the optimal provision of insurance against the disability risk in a dynamic framework with imperfectly observable health. We however restrict ourselves to the risk of permanent disability. Our aim is to uncover the key

¹There is also, of course, a leisure component to the provision of pensions which will be captured by our theoretical framework.

²7 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.

qualitative and quantitative features that must be satisfied by an optimal policy.

Let us now describe our theoretical framework. At any age, a worker faces the risk of being hit by an irreversible disability shock. The government would like to provide insurance against disability while inducing the able to work up to a general retirement age. To enhance the provision of insurance, the government could rely on its imperfect information on health 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"³ as disabled and therefore eligible for this higher consumption level. Thus, the only information that the government has about the health of an individual is whether or not this individual has been awarded a tag.

Importantly, 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 government still wants to provide the able and tagged with incentives to work. Thus, the optimal allocation of resources is found by setting up a dynamic mechanism design problem where the able, whether tagged or not, are induced to work until some retirement age to be determined. Inducing the tagged to participate is costly, since they face a strong temptation to claim to be disabled. Hence, they should only be induced to work up to an early retirement age.⁴

The first-order conditions to the planner's problem provide a number of qualitative insights. First, as in some other dynamic contracting problems, the optimal allocation is characterized by back-loaded incentives. The reward for participation is higher for more experienced workers. Not only does it induce the old to participate, it also induces the young who are motivated by the prospect of high consumption once they will have accumulated a lot of work experience. The consumption of participating workers is therefore increasing with age and jumps upward when a tag is awarded.

A more fundamental insight is that, when an agent stopped working before being awarded the tag, the optimal policy exploits the difference between the age at which this agent stopped working, i.e. the age at which he claimed to be disabled, and the age at which he was awarded the tag. The idea is that someone 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 a long time. The former should therefore be rewarded with high consumption while the latter should be punished with low consumption. This key insight, which is specific to a dynamic framework with imperfect tagging, shows how the planner can use imperfect information on health in order to improve the trade-off

³The term was originally introduced by Akerlof (1978) who performed the first analysis of the optimal use of tagging in the design of welfare programs.

⁴As we argue in the text, such an early retirement age could be seen as 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.

between insurance and incentives to work.

To illustrate these features and to have a more quantitative sense of the main characteristics of the optimal policy, we calibrate the model to U.S. data and perform a numerical simulation. The first finding that we obtain is that the implementation of the optimal policy makes it desirable to decrease the strictness of the disability test so as to reduce the number of disabled individuals who are not awarded the tag. However, crucially, doing so is only desirable because the able and tagged are induced to work. When such incentives are not provided, it becomes desirable to increase the strictness of the disability test in order to reduce the number of able workers who are thrown out of the labor force by the award of a tag. We also show that it is important to implement an early retirement age for the tagged as inducing them to work until the general retirement age would be excessively costly and could result in a welfare loss compared to a situation where they do not work at all.

These numerical results are obtained assuming that the strictness of the disability test is chosen to minimize the total number of classification errors (while allowing for a preference between rejection and award errors). Although this is a natural and realistic benchmark, it is interesting to characterize 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 case, the first-best allocation of resources can asymptotically be implemented. The idea is to set a very high disability threshold after the retirement age, so that those who do not become tagged are almost surely able to work, and to punish individuals who claimed to be disabled in the past but fail to be awarded the tag. If the punishment is sufficiently severe, no able worker would ever claim to be disabled; and, if the disability threshold is sufficiently high, the punishment would almost never fall onto a truly disabled worker. 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, as we argue empirically, 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 of resources. We do not investigate how it could be implemented in a decentralized market economy where the government is constrained to use fiscal instruments instead of directly choosing individuals' consumption levels and labor supplies. Note that, while optimal allocations are typically unique, there usually exist multiple ways of implementing them. Thus, in general, results about optimal allocations are more robust than about their implementation.

Related Literature. Our paper builds on two strands of the literature. The first focuses on the optimal design of insurance against the occurrence of unobservable idiosyncratic

shocks to productivity. More specifically, we build on the seminal work of Diamond and Mirrlees (1978) which characterizes the optimal provision of social insurance against the risk of permanent disability when health is unobservable. As inducing the able to work is costly, they find that the general retirement age should be smaller than the first-best retirement age of the full-information economy. Thus, old-age pensions do not merely allow people to enjoy leisure at the end of their lives, they should also be designed to provide insurance against the risk of permanent disability.

Over the last decade, a renewed interest in the optimal provision of insurance against productivity shocks has led to the emergence of the New Dynamic Public Finance literature. Golosov, Kocherlakota and Tsyvinski (2003) showed that, within this family of problems, the optimal incentive-feasible allocation of consumption is always characterized by inverse Euler equations. However, it is not possible to obtain such a general characterization of the optimal allocation of time between work and leisure. Thus, simulations of the optimal allocation of resources have only been realized in special cases. Albanesi and Sleet (2006) solved the optimal insurance problem when productivity shocks are i.i.d.. More recently, Farhi and Werning (2012) have been able to solve the AR(1) case. It turns out that the case of permanent disability shocks *à la* Diamond and Mirrlees (1978) can also be fully solved. Indeed, Golosov and Tsyvinski (2006) were the first to perform a numerical simulation of the optimal insurance policy with permanent disability shocks. However, they only allowed for an intensive labor supply margin, which implies that able workers never wish to retire (even though their working hours decline as their productivity falls). By contrast, following Diamond and Mirrlees (1978), we only have an extensive labor supply margin, which eventually induces all agents to retire so as to enjoy leisure.⁵

The main contribution of Golosov and Tsyvinski (2006) is their demonstration that the optimal allocation of resources can be implemented in a decentralized economy by an asset test which specifies that an agent is eligible for disability benefits if and only if his assets are below a specific threshold. Their numerical simulation shows that implementing the asset test generates a consumption-equivalent welfare gain of 0.5% compared to the optimal policy with hidden savings.

Within this literature, Grochulski and Kocherlakota (2010) showed that, for a broad family of social insurance problems, the optimal allocation can be implemented by a linear tax on wealth together with a social security system which consists of history-dependent taxes or transfers after retirement. However, their paper focuses almost exclusively on the implementation within a decentralized economy of an optimal allocation that they do not fully characterize. Their approach to the optimal design of social security systems

⁵Note that Michau (2012) and Shourideh and Troshkin (2012) also perform numerical simulations of an optimal policy with a retirement margin. However, both papers focus on redistribution across ex-ante heterogeneous individuals and do not allow for permanent disability shocks.

should therefore be seen as complementary to ours.

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 imperfectly observe. Indeed, Diamond and Sheshinski (1995), Parsons (1996) and Salanie (2002) showed that, even if the government's information is very imperfect, welfare can be improved by enhancing the consumption of those who are tagged as disabled. Our work is particularly related to that of Parsons (1996) which stressed that the able who are mistakenly tagged as disabled should be incentivized to work. However, all these models are static and none of them performs a numerical simulation of the optimal policy. They therefore do not offer any quantitative evaluation of the welfare gains that could be generated by the imperfect observability of health.

There has recently been renewed interest in the use of tagging for the optimal design of insurance or redistribution policies. For example, it has been shown that significant welfare gains could be generated by making taxes dependent on age (Weinzierl, 2011), on gender (Alesina, Ichino and Karabarbounis, 2011) or even on height (Mankiw and Weinzierl, 2010). However, given the significant size of these welfare gains, it seems puzzling that such tags are not more widely used in practice. Interestingly, Weinzierl (2012) argues that concerns about horizontal equity, which follow from the equal sacrifice principle, imply that tags should only be used in practice if the information they provide is strongly correlated with hidden productivity. He concludes that disability tags are acceptable, while gender, race or height tags are not.

Our work combines these two strands of the literature on optimal social insurance by introducing imperfect tagging into the dynamic mechanism design approach of Diamond and Mirrlees (1978).

While there is a considerable literature on optimal unemployment insurance, relatively little is known about the optimal design of disability insurance. In addition to the work mentioned above, the literature nevertheless includes some important contributions on the topic. Benitez-Silva, Buchinsky and Rust (2006) relied 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 (2011) considered 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 (2012) propose a structural model of labor supply in a life-cycle setting where workers are subject to both disability and wage shocks. Relying on an empirical

estimation of their structural parameters, they argue that decreasing the strictness of the disability test would enhance social welfare. Thus, it is desirable to increase the number of disabled workers who are awarded the tag, despite a resulting worsening of incentives.

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 conditions that characterize the optimal allocation of resources. Then, in the following section, we calibrate the model. Section 4 is devoted to the numerical simulation of the optimal policy and to the evaluation 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 our theoretical framework. We first present the setup of our model, then give the planner's problem and, finally, describe the first-order conditions which characterize the optimal allocation of resources.

2.1 Setup

There is a mass 1 of agents who face a deterministic life span equal to H . Time is continuous, which is necessary to be able to rely on a first-order condition to determine the retirement age. Everyone derives instantaneous utility $u(c)$ from consuming c at a given point in time, where $u'(\cdot) > 0$, $u''(\cdot) < 0$, $\lim_{c \rightarrow 0^+} u(c) = -\infty$ and $\lim_{c \rightarrow 0^+} u'(c) = +\infty$.⁶ At a given age, an individual can either be able or disabled. Only the able can work. Their productivity evolves deterministically over the life-cycle and is equal to γ_t for a worker of age t . Productivity is a continuous function of age and is minimal at H . We will later assume, in the calibration section, that γ_t follows an inverted U-shape. To generate a retirement decision, we impose that labor supply is indivisible and that, at each instant, workers face a fixed utility cost $b > 0$ of working.⁷ Agents discount the future at rate ρ . Finally, the planner can transfer resources across time at an exogenous risk-free interest rate which, for simplicity, is assumed to be equal to ρ .

At the individual level, the only source of risk in this economy is the stochastic

⁶Finkelstein, Luttmer and Notowidigdo (2013) 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.

⁷It would be straightforward to add an intensive margin. Indeed, in a similar setup, Golosov and Tsyvinski (2006) only allowed for an intensive margin of labor supply. However, Liebman, Luttmer and Seif (2009) provided some empirical evidence that most of the labor supply response to changes in the level of Social Security benefits occurs at the extensive margin.

occurrence of disability shocks. Disability is an absorbing state which implies that once a worker has been hit by the shock, he never regains his ability to work. The occurrence of the shock over the life cycle is determined by a c.d.f. which is denoted by $F(t)$, where $t \in [0, H]$, while the corresponding p.d.f. is denoted by $f(t)$. Thus, at age t a fraction $F(t)$ of the population is disabled.

In order to have a pure social insurance problem, with *ex-ante* identical individuals, we shall assume that the planner attaches zero weight on those who became disabled before starting to work.⁸ 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 consumption, 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.

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 led 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 the 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 are 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.

However, the assumption of unobservable health is very strong. Indeed, in most countries, including the U.S., disability insurance programs do rely on imperfect information on the work ability of their applicants. They typically run a medical test and "tag" as disabled those whose health appears to be below some threshold. However, as the information is imperfect, some errors are unavoidable leading to the occurrence of gaps and leakages. Gaps occur when some disabled individuals remain untagged; while leakages occur when some able are tagged.⁹

⁸When determining the planner's optimal allocation, this is equivalent to imposing the normalization $F(0) = 0$.

⁹In 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

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, in the U.S., as in many other countries, the disability insurance program 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 binary tagging process.

More formally, let θ denote the outcome of the test for a given individual. Thus, θ could be thought of as his apparent health. Its c.d.f. over the population is $G_A(\theta)$ for an able and $G_D(\theta)$ for a disabled. The respective p.d.f.s are denoted by $g_A(\theta)$ and $g_D(\theta)$. An individual is tagged if his θ 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 $G_A(\cdot)$ first-order stochastically dominates $G_D(\cdot)$, i.e. $G_A(\theta) < G_D(\theta)$ for all θ , and that the two distributions satisfy the monotone likelihood ratio condition, i.e. $g_A(\theta)/g_D(\theta)$ is increasing in θ . Furthermore, we assume that, for a given individual, θ 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 of leakages. See Figure 1.

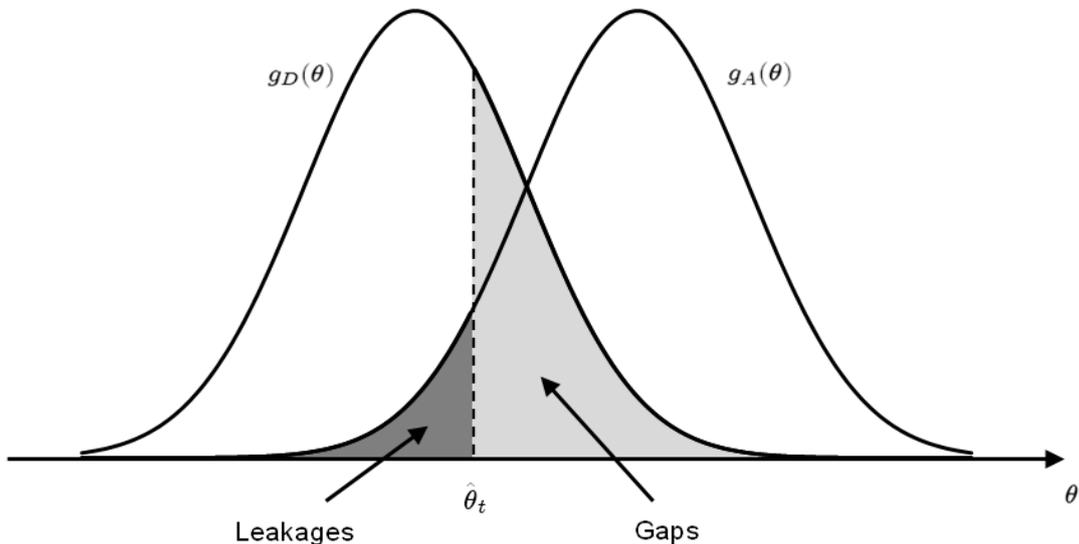


Figure 1: Trade-off between gaps and leakages

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 concern, while the opposite is true for gaps. We therefore assume an age-dependent disability standard $\hat{\theta}_t$. Since there is a mass 1 of individuals, gaps is equivalent to the probability of being disabled *and* untagged, and leakages to the probability of being able *and* tagged.

standard equal to $\hat{\theta}_t$ at age t , which is non-decreasing with age.¹⁰ 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 at which they become 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:

$$f(i, j) = f(j|i)f(i), \quad (1)$$

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:

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

A fraction $G_A(\hat{\theta}_0)$ of individuals obtain the tag at age 0. To understand the second and fourth cases, i.e. $0 < j < i$ and $i < j < H$, note that the only way by which an agent can become tagged if he does not simultaneously become disabled is that the disability standard $\hat{\theta}_j$ increases sufficiently so that his own constant θ falls below this 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 ε . The corresponding probability density is equal to $[G_A(\hat{\theta}_{j+\varepsilon}) - G_A(\hat{\theta}_j)] / \varepsilon$ with ε 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 having already been 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 seems 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

¹⁰Note that, for any given individual, θ 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. It follows that the disability standard $\hat{\theta}_t$ is only non-decreasing *relative* to $g_A(\theta)$ and $g_D(\theta)$.

able, $i = H$, (2) simplifies to:

$$f(j|i = H) = \begin{cases} G_A(\hat{\theta}_0) & \text{if } j = 0 \\ g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} & \text{if } 0 < j < H \\ 1 - G_A(\hat{\theta}_H) & \text{if } j = H \end{cases} . \quad (3)$$

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 θ . All they know is whether they are eligible for the tag or not. This assumption implies 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 θ , would imply that they can predict at age 0 when they would become eligible for the tag conditional on remaining able to work. 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. 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 health condition of an individual evolve over time.

Importantly, the structure of the model implies that increasing the number of medical tests imposed on a given individual does not elicit additional information. By contrast, Low and Pistaferri (2012) find it desirable to increase the frequency of reassessments (even though this only generates modest welfare gains). Crucially, they assume that the outcomes of the tests are independent from each other. By contrast, our model implies that once an individual has been awarded a tag, he remains eligible forever. Thus, the amount of information that the government can obtain on the health of its citizens is not fundamentally constrained by the cost of performing medical tests. For simplicity, we therefore assume throughout that these costs are negligible.

2.2 Planner's problem

The problem of the social planner is to maximize the expected lifetime utility of workers at age 0 subject to a resource constraint and to a set of incentive compatibility constraints which ensure that the able choose to work (up to some retirement age). For each individual, the planner observes his age and, if applicable, the age of award of the tag and the age at which the individual stopped working. However, the planner cannot directly observe whether or not an individual is able to work as this is private information. Hence, the age of occurrence of disability is only revealed by the incentive compatibility constraints (unless disability occurs after retirement). Before formally deriving the

planner's objective and constraints, we need to specify his control variables.

The first set of control variables includes all the labor supply decisions which, here, consist of the retirement ages of the able.^{11,12} The planner will make the retirement age of each worker conditional on all the information that he has about him. The relevant information set consists of whether the individual has been tagged and, if so, the age j of the award. We denote by RU the retirement age of the untagged and by $RT(j)$ the retirement age of able workers who got tagged at age j .¹³ Note that those whose apparent health is lower, i.e. who have a lower θ , will be tagged earlier. This implies that j is a sufficient statistic for the apparent health of the able and tagged and, hence, $RT(j)$ could be seen as a health-dependent retirement age.

Parsons (1996) was the first to emphasize 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. While, in our dynamic setup, it might not be desirable to induce the able and tagged to work until RU , the planner nevertheless has the flexibility to induce them to work until some earlier retirement ages given by $\{RT(j)\}_{j \in [0, RU]}$.¹⁴ Inducing the able and tagged to participate is not unrealistic. Indeed, in some 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 planner who might be tempted to remove the tag from those who reveal that they are able to work.

The remaining control variables are the consumption levels. Again, they should be made conditional upon all the information that the planner has about each individual. The consumption levels chosen by the planner are therefore a function of the age t of the agent and, if applicable, of the age j of award of the tag and of the age r of retirement. Note that an individual retires either when he becomes disabled or when he reaches the retirement age (which is either equal to RU or to $RT(j)$). The planner therefore needs to determine the consumption at age t of Working individuals who are Untagged, $\{c^{WU}(t)\}_{t \in [0, RU]}$, of Working individuals who are Tagged, $\{c^{WT}(t, j)\}_{j \in [0, RU], t \in [j, RT(j)]}$, of Non-working individuals who are Untagged, $\{c^{NU}(t, r)\}_{r \in [0, RU], t \in [r, H]}$, and, finally, of Non-working individuals who are Tagged, $\{c_N^{NT}(r, j)\}_{r \in [0, RU], j \in (r, H]}$ and

¹¹The disabled trivially retire when they lose their ability to work.

¹²Note that, even if workers have low productivity when young, we do not allow them to postpone entry into the labor force. One external justification for this is on-the-job learning, which makes early work at low productivity an investment into the future. 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. Thus, here, as in Diamond and Mirrlees (1978), the retirement age summarizes the labor supply decision of an individual.

¹³Clearly, those who only get tagged after RU retire at RU . We can therefore consider that $RT(j) = RU$ whenever $j \geq RU$.

¹⁴In theory, $RT(j)$ could be larger than RU . However, given that it is more costly to induce a worker to participate when he has been awarded a tag, we clearly expect $RT(j) < RU$. This is confirmed by our numerical simulation.

$\{c_T^{NT}(r, j)\}_{j \in [0, RU], r \in [j, RT(j)]}$. 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, $c_N^{NT}(r, j)$ and $c_T^{NT}(r, j)$, 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 these consumption levels over time. In other words, for individuals who are both retired and tagged, age does not provide any information on whether the agent was able to work when he retired. Thus, allowing consumption to depend on age would not help the social planner relax any incentive compatibility constraints. We can therefore safely omit the age t from these last two consumption functions.

Importantly, in this section, the disability standards $\{\hat{\theta}_t\}_{t \in [0, H]}$ are assumed to be exogenously determined. This assumption will be relaxed in the last section of the paper where they become additional control variables of the planner.

Now that we have defined the planner's control variables, we can derive his objective and constraints. *Ex post*, a given individual is characterized by the age i at which he became disabled and the age j at which he became tagged. The *ex-ante* probability density of becoming individual (i, j) is equal to $f(i, j)$, as defined by (1), (2) and (3).

Let us now derive the *ex-post* lifetime utility $v(i, j)$ of individual (i, j) . If an agent retires before becoming tagged, i.e. $\min\{i, RU\} < j$, his utility is:

$$\begin{aligned} v(i, j) &= \int_0^{\min\{i, RU\}} e^{-\rho t} [u(c^{WU}(t)) - b] dt \\ &\quad + \int_{\min\{i, RU\}}^j e^{-\rho t} u(c^{NU}(t, \min\{i, RU\})) dt \\ &\quad + \int_j^H e^{-\rho t} u(c_N^{NT}(\min\{i, RU\}, j)) dt. \end{aligned} \quad (4)$$

From age 0 to $\min\{i, RU\}$ the individual is working and untagged, he consumes $c^{WU}(t)$ at age t and gets disutility b from working. From age $\min\{i, RU\}$ to j , he is retired and untagged and gets the corresponding consumption level $c^{NU}(t, \min\{i, RU\})$ at age t . Finally, from age j to H , his consumption level is that of a retired and tagged who retired before becoming tagged. Now, if an agent becomes tagged before retirement or if he becomes tagged and retires simultaneously, i.e. $j \leq \min\{i, RU\}$, his utility is:

$$\begin{aligned} v(i, j) &= \int_0^j e^{-\rho t} [u(c^{WU}(t)) - b] dt \\ &\quad + \int_j^{\min\{i, RT(j)\}} e^{-\rho t} [u(c^{WT}(t, j)) - b] dt \\ &\quad + \int_{\min\{i, RT(j)\}}^H e^{-\rho t} u(c_T^{NT}(\min\{i, RT(j)\}, j)) dt. \end{aligned} \quad (5)$$

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

The objective of the social planner is to maximize the *ex-ante* expected lifetime utility of workers which is equal to:

$$\mathbb{E}[v(i, j)] = \int_0^H \int_0^H v(i, j) f(i, j) di dj, \quad (6)$$

since each individual faces a likelihood $f(i, j)$ of becoming individual (i, j) with lifetime utility $v(i, j)$.

To derive the resource constraint of the planner's problem, we need to know the lifetime budget deficit $z(i, j)$ generated by individual (i, j) . For $\min\{i, RU\} < j$, we have:

$$\begin{aligned} z(i, j) &= \int_0^{\min\{i, RU\}} e^{-\rho t} [c^{WU}(t) - \gamma_t] dt \\ &\quad + \int_{\min\{i, RU\}}^j e^{-\rho t} c^{NU}(t, \min\{i, RU\}) dt \\ &\quad + \int_j^H e^{-\rho t} c_N^{NT}(\min\{i, RU\}, j) dt, \end{aligned} \quad (7)$$

where we have used the fact that working agents, who get disutility b from work in (4), produce γ_t units of consumption goods at age t . Similarly, for $j \leq \min\{i, RU\}$, we have:

$$\begin{aligned} z(i, j) &= \int_0^j e^{-\rho t} [c^{WU}(t) - \gamma_t] dt \\ &\quad + \int_j^{\min\{i, RT(j)\}} e^{-\rho t} [c^{WT}(t, j) - \gamma_t] dt \\ &\quad + \int_{\min\{i, RT(j)\}}^H e^{-\rho t} c_T^{NT}(\min\{i, RT(j)\}, j) dt. \end{aligned} \quad (8)$$

Thus, the planner's resource constraint is:

$$\mathbb{E}[z(i, j)] \leq 0. \quad (9)$$

The resource constraint therefore imposes that the expected lifetime consumption of individuals does not exceed the amount that they are expected to produce.

Finally, the optimal allocation of resources must satisfy a set of incentive compatibility constraints which ensure that able individuals choose to work (provided that they have

not yet reached the retirement age). Able individuals can either be untagged or tagged. Hence, the set of incentive compatibility constraints could be divided into two subsets, one for the untagged and the other for the tagged. The first subset imposes that the untagged who are able and younger than RU choose to work. More formally, let $v_t(i, j)$ denote the *ex-post* lifetime utility of individual (i, j) from age t onwards.¹⁵ For each age $t \in [0, RU)$, we must impose the following incentive compatibility constraint:

$$\mathbb{E}[v_t(i, j)|i > t, j > t] \geq \mathbb{E}[v_t(t, j)|i > t, j > t]. \quad (10)$$

It requires that, at age t , the expected utility from working until the retirement age chosen by the planner or until disability occurs, $\mathbb{E}[v_t(i, j)|i > t, j > t]$, is never smaller than the expected utility from retiring at t , $\mathbb{E}[v_t(t, j)|i > t, j > t]$.

The second subset imposes that agents tagged at age j who are able and younger than $RT(j)$ choose to work. Thus, for each tag age $j \in [0, RU)$ and each age $t \in [j, RT(j))$, the following incentive compatibility constraint must hold:

$$\mathbb{E}[v_t(i, j)|i > t] \geq \mathbb{E}[v_t(t, j)|i > t]. \quad (11)$$

This requires that, at age t , the expected utility from working until $RT(j)$ or until disability occurs, $\mathbb{E}[v_t(i, j)|i > t]$, is never smaller than the expected utility from retiring at t , $\mathbb{E}[v_t(t, j)|i > t]$. Note that, in (11), j is not a random variable since its value is already known at age t (as $j \leq t$).¹⁶ Interestingly, this last subset of constraints is formally identical to the one imposed by Diamond and Mirrlees (1978). This is due to the fact that, once an agent is tagged, the planner cannot rely on any additional information about his health and therefore acts as if health was completely unobservable.

The difference between $c^{NT}(r, j)$ depending on whether the individual retires first, i.e. $c_N^{NT}(r, j)$ if $r < j$, or becomes tagged at retirement or before, i.e. $c_T^{NT}(r, j)$ if $j \leq r$, is explained by the fact that the latter consumption levels enter the incentive compatibility constraints of the tagged while the former do not.

The planner's problem is:

$$\max \mathbb{E}[v(i, j)] \quad (12)$$

subject to:

- Resource constraint:

$$\mathbb{E}[z(i, j)] \leq 0;$$

¹⁵Thus, $v_0(i, j)$ is equal to $v(i, j)$ as defined by (4) and (5).

¹⁶Also, note that $\mathbb{E}[v_t(t, j)|i > t] = v_t(t, j)$ since, once an agent is retired and tagged, the occurrence of disability can no longer affect his welfare.

- Incentive compatibility constraints for the untagged:

$$\forall t \in [0, RU), \quad \mathbb{E}[v_t(i, j)|i > t, j > t] \geq \mathbb{E}[v_t(t, j)|i > t, j > t];$$

- Incentive compatibility constraints for the tagged:

$$\forall j \in [0, RU), \forall t \in [j, RT(j)), \quad \mathbb{E}[v_t(i, j)|i > t] \geq \mathbb{E}[v_t(t, j)|i > t].$$

The control variables are $c^{WU}(\cdot)$, $c^{WT}(\cdot)$, $c^{NU}(\cdot)$, $c_N^{NT}(\cdot)$, $c_T^{NT}(\cdot)$, $RT(\cdot)$ and RU . The fully detailed planner's problem is given in Appendix A.

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 and Kopczuk (2011), 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, can be used as a screening device to reduce, or even eliminate, leakages. However, this possibility does not arise in our framework where agents can know the outcome of the test, through their private doctor for instance, before applying.

2.3 First-order conditions

The optimal allocation of resources is the solution to a constrained optimization problem. If we consider that the control variables are the utility levels of the agents, 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 necessary and sufficient.

The remaining control variables of the planner are the retirement ages. In that respect, a key feature of our model is that labor supply is indivisible which, at any given point in time, creates a non-convexity into workers' labor supply problem. As emphasized by Mulligan (2001) and Ljungqvist and Sargent (2006), this non-convexity is easily overcome in a life-cycle framework. Indeed, agents convexify their labor supply problem by working for a fraction of their lives and enjoying leisure for the remaining fraction. Hence, here, as in Diamond and Mirrlees (1978), we have a "time averaging" model of the labor supply.¹⁷ In that context, the planner's optimal retirement ages are characterized by first-order conditions.

¹⁷Ljungqvist and Sargent (2006) established some equivalence results between lotteries and time averaging models of indivisible labor.

Finally, we conjecture that all the incentive compatibility constraints are binding. If they were not, then welfare could be improved by lowering the consumption levels of the able while increasing those of the disabled.¹⁸ Our numerical simulation confirms that all the Lagrange multipliers are positive.

Importantly, the incentive compatibility constraints imply that, for agents who have not yet reached the planner's retirement ages, able individuals are all working and non-working individuals are all disabled.

We now review the first-order conditions that characterize the planner's optimal allocation of resources. The consumption levels at t of the able and disabled who became tagged at age j are related by:

$$\frac{d}{dt} \frac{1}{u'(c^{WT}(t, j))} = \left[\frac{1}{u'(c^{WT}(t, j))} - \frac{1}{u'(c_T^{NT}(t, j))} \right] \frac{f(t)}{1 - F(t)}. \quad (13)$$

This condition is identical to the original inverse Euler equation derived in Diamond and Mirrlees (1978). Indeed, once the tag has been awarded, the planner does not have any further information on the health of the people, which implies that we are back to the unobservable health benchmark.

What is the general intuition for these inverse Euler equations? Recall that the incentive compatibility constraints are linear in utilities. Hence, to preserve incentives to work, resources shifted to the next point in time must increase the utility in the good state, i.e. able and working, as much as in the bad state, i.e. disabled and non-working. 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, such transfers of utilities across time should be done at minimum budgetary cost. It follows that, at the optimum, the planner equates the marginal resource cost of providing utility at different points in time. Indeed, the inverse marginal utility of consumption is precisely the marginal resource cost of providing utility, i.e. $1/u'(c) = 1/(du/dc) = dc(u)/du$ where $c(\cdot) \equiv u^{-1}(\cdot)$.¹⁹

Finally, note that $f(t)/(1 - F(t))$ is the probability density that a tagged agent becomes disabled at t given that he was able up to then. The first-order condition (13) imposes that the lower is this probability of becoming disabled, the lower is the consump-

¹⁸Golosov and Tsyvinski (2004) provide a formal proof in a simpler context with unobservable health.

¹⁹Let $u^{WT}(t, j) \equiv u(c^{WT}(t, j))$, $u_T^{NT}(t, j) \equiv u(c_T^{NT}(t, j))$ and $c(\cdot) \equiv u^{-1}(\cdot)$, which implies that $c'(u^{WT}(t, j)) = 1/u'(c^{WT}(t, j))$ and $c'(u_T^{NT}(t, j)) = 1/u'(c_T^{NT}(t, j))$. Relying on these notations, the above interpretation is easier to follow if the first-order condition (13) is written as:

$$c'(u^{WT}(t - dt, j)) = \left(1 - \frac{f(t)dt}{1 - F(t)} \right) c'(u^{WT}(t, j)) + \frac{f(t)dt}{1 - F(t)} c'(u_T^{NT}(t, j)),$$

which shows that the marginal resource cost of providing utility today, at $t - dt$, must be equal to the expected marginal resource cost of providing utility tomorrow, at t .

tion level of non-working agents $c_T^{NT}(t, j)$ for a given path of $c^{WT}(t, j)$. This enhances incentives to work at little cost in terms of insurance.

The boundary condition associated with (13) is:

$$\lim_{t \rightarrow RT(j)-} c^{WT}(t, j) = c_T^{NT}(RT(j), j). \quad (14)$$

As workers approach the retirement age, incentives to participate only need to be provided for a short period of time. Hence, consumption can be almost perfectly smoothed over time.

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

$$\frac{b}{u'(c^{WT}(RT(j), j))} = \gamma_{RT(j)}. \quad (15)$$

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 working and non-working agents of age i , who are not tagged, are related by:

$$\begin{aligned} \frac{d}{di} \frac{1}{u'(c^{WU}(i))} &= \left[\frac{1}{u'(c^{WU}(i))} - \frac{1}{u'(c^{NU}(t, i))} \right] \\ &\times \frac{[1 - G_D(\hat{\theta}_t)] f(i)}{[1 - G_D(\hat{\theta}_t)] [F(t) - F(i)] + [1 - G_A(\hat{\theta}_t)] [1 - F(t)]}, \end{aligned} \quad (16)$$

for any $t \geq i$. The interpretation is similar to that of equation (13), 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 is still not tagged at age t . 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 $c^{NU}(t, i)$ be. Indeed, if the absence of a tag at age t reveals that an agent probably lied about becoming disabled at i , then this agent should be punished with a low consumption level $c^{NU}(t, i)$. This improves incentives at little cost in terms of insurance. This key insight shows how the imperfect tag can be used in a dynamic setup to extract information on the true health status of an individual. The boundary condition associated with (16) is:

$$\lim_{t \rightarrow RU-} c^{WU}(t) = c^{NU}(t, RU), \forall t \in [RU, H]. \quad (17)$$

Again, as workers approach the retirement age, their consumption converges to the level

that they will obtain when they retire.

Similarly, the consumption levels of the non-working and tagged, who stopped working before becoming tagged, and of the working and untagged are linked by:

$$\begin{aligned} \frac{d}{di} \frac{1}{u'(c^{WU}(i))} &= \left[\frac{1}{u'(c^{WU}(i))} - \frac{1}{u'(c_N^{NT}(i, j))} \right] \\ &\times \frac{g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i)}{g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + \left[G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j) + g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [F(j) - F(i)]}, \end{aligned} \quad (18)$$

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 became 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 got tagged at j , the lower should $c_N^{NT}(i, j)$ be. The corresponding boundary condition is:

$$\lim_{t \rightarrow RU-} c^{WU}(t) = c_N^{NT}(RU, j), \forall j \in (RU, H]. \quad (19)$$

Together with (17), this implies that being awarded the tag after retirement does not make any difference to those who worked until the maximum retirement age RU .

The consumption levels of the newly tagged, working and non-working, are related to that of the working and untagged by the following condition:

$$\begin{aligned} \frac{1}{u'(c^{WU}(j))} &= \frac{1}{u'(c^{WT}(j, j))} - \left[\frac{1}{u'(c^{WT}(j, j))} - \frac{1}{u'(c_T^{NT}(j, j))} \right] \\ &\times \frac{\left[G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j)}{g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + \left[G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j) \right] f(j)}, \end{aligned} \quad (20)$$

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 first-order condition imposes 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 static context, relates inverse marginal utilities.²⁰

All these first-order conditions relating inverse marginal utilities show that the planner wants to equalize the marginal resource cost of providing utility to the agents across time

²⁰In a discrete time version of our model, this first-order condition (20) becomes exactly identical to the one derived by Parsons (1996). Note that a discrete time model cannot be analytically tractable with an extensive labor supply margin.

and across states. This general principle nests the standard inverse Euler equation derived by Diamond and Mirrlees (1978), condition (13), the first-order condition of Parsons (1996), condition (20), as well as the two first-order conditions which are specific to this paper, (16) and (18).

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

$$\frac{b}{u'(c^{WU}(RU))} = \gamma_{RU}. \quad (21)$$

Again, as for condition (15), 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.

The Lagrange multiplier associated with the resource constraint is equal to $u'(c^{WU}(0))$. The Lagrange multipliers of the incentive compatibility constraints of the newly tagged, given by (11) with $t = j$, are:

$$u'(c^{WU}(0)) \left[\frac{1}{u'(c^{WT}(j, j))} - \frac{1}{u'(c^{WU}(j))} \right], \quad (22)$$

with $j \in [0, RU)$, and those of the previously tagged, given by (11) with $t > j$, are:

$$u'(c^{WU}(0)) \frac{d}{dt} \frac{1}{u'(c^{WT}(t, j))}, \quad (23)$$

with $j \in [0, RU)$ and $t \in (j, RT(j))$. Binding constraints imply that these multipliers are positive and, hence, that the consumption of tagged workers should initially be higher than that of the untagged and it should then be increasing over the life-cycle (until age $RT(j)$). 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 multipliers of the incentive compatibility constraints of the untagged, given by (10), are:

$$u'(c^{WU}(0)) \frac{d}{dt} \frac{1}{u'(c^{WU}(t))}, \quad (24)$$

with $t \in [0, RU)$, which implies that the consumption of the working and untagged should also be increasing over the life-cycle (until age RU).

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

Proposition 1 *An optimal Social Security system with imperfect tagging should implement the allocation of resources which is characterized by the first-order conditions (13), (14), (15), (16), (17), (18), (19), (20), (21) together with the resource constraint, (9), the incentive compatibility constraints for the untagged, (10), and for the tagged, (11).²¹*

²¹The full expressions for these three set of constraints to the planner's problem are given in Appendix

To gain additional insights about this optimal incentive-feasible allocation we need to perform a numerical simulation. But, before that, the model needs to be 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. As in Golosov and Tsyvinski (2006), productivity γ_t at each age t is determined by fitting a quadratic approximation to 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. $\gamma_{50} = 1.47$.

3.2 Preferences

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

$$u(c) = \frac{c^{1-\phi} - 1}{1-\phi}. \quad (25)$$

We set the coefficient of relative risk aversion ϕ equal to 2. The annual discount rate ρ , which also equals the annual interest rate, is set at 0.02. The fixed cost b of working is calibrated such that, in the unobservable health case, the able retire at age 65. This exercise yields $b = 1.092$.²²

3.3 Distribution of the disability age

To determine the likelihood of being disabled at age t , $F(t)$, we rely on 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.²³ We make use of the following question:

A by (A2), (A3) and (A4), respectively.

²²Note that choosing b such that, in the unobservable health benchmark, the retirement age is equal to 65 is equivalent to choosing b such that, in the planner's optimal allocation, the able and untagged retire at age 67.3 (cf. Table 2 below).

²³This is the same data source as used by Low and Pistaferri (2012). 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.

"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.

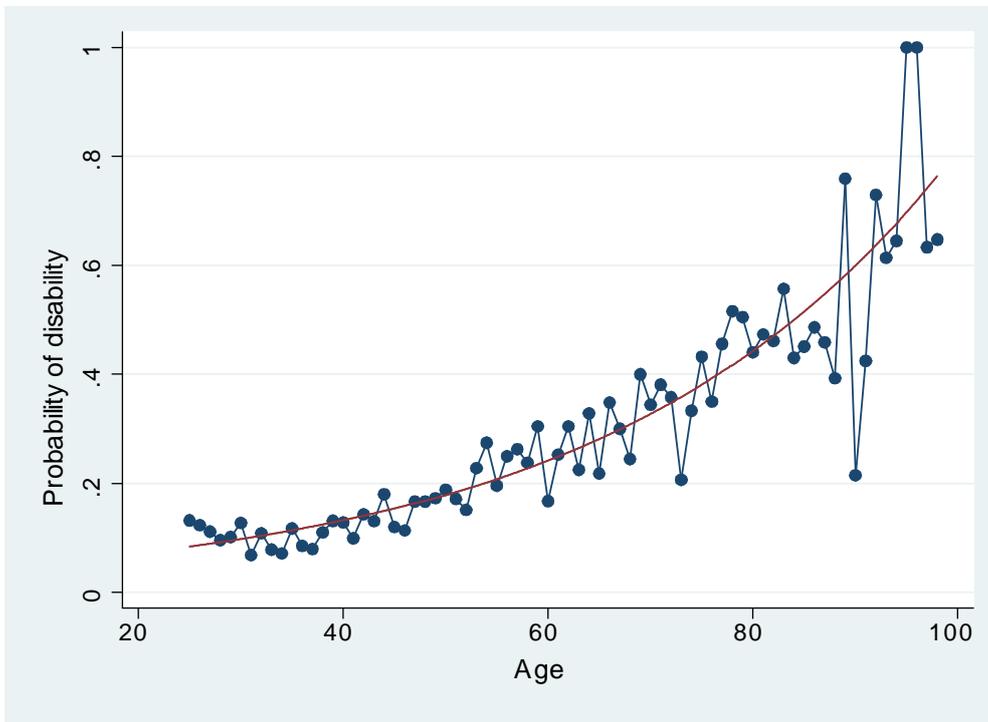


Figure 2: Distribution of disability age

At face value, our definition of disability may seem rather mild. However, in the context of our model, disability should not be interpreted too narrowly. Indeed, any individual whose productivity is virtually equal to zero can be considered to be 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).

An obvious concern with self-reported disability is that some workers, such as those applying for disability insurance, might exaggerate the severity of their health problems. However, Benitez-Silva, Buchinsky, Chan, Cheidvasser and Rust (2004) show that self-reported disability is in fact an unbiased predictor of the true disability status of individuals as measured by the U.S. Social Security Administration.

3.4 Trade-off between gaps and leakages

The outcome of the health test for both disabled, $g_D(\theta)$, and able, $g_A(\theta)$, individuals is assumed to be normally distributed. The two distributions are characterized by a difference in means equal to μ and a standard deviation of 1.²⁴ Although the actual means of the two distributions are inconsequential (cf. footnote 10), 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 μ , information is required on individuals' ability to work, i.e. able or disabled, as well as their official disability classification, 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 our sample, as these people have reached the full retirement age and are shifted to the retired worker portion of the U.S. Social Security system.^{25,26}

An individual's disability classification, i.e. tagged or untagged, is a random variable following a Bernoulli distribution, where the probability of being tagged depends on his age $t \in \{25, \dots, 65\}$ and on his ability to work. An agent of age t is awarded the tag when his test outcome falls below $\hat{\theta}_t$. The structure of our model therefore implies that:

$$\Pr(\text{Tagged}|\text{Age} = t, \text{Ability}) = \Phi \left(\sum_{s=25}^{65} \hat{\theta}_s \mathcal{I}(s = t) - \frac{\mu}{2} \mathcal{I}(\text{Able}) + \frac{\mu}{2} \mathcal{I}(\text{Disabled}) \right), \quad (26)$$

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.²⁷ Rearranging terms, a simple probit regression of disability classification on a set of age dummies and on self-reported ability status can be employed to obtain an estimate of

²⁴Alternatively, we could fix μ and calibrate the standard deviation. However, fixing the standard deviation is particularly suitable in our context as with $\mu = 0$ the problem collapses to the unobservable health case treated by Diamond and Mirrlees (1978).

²⁵Within the sample of people aged 25-65, a small proportion of individuals receives 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.

²⁶In our PSID data set and across all age groups (from age 25 to 65), out of 9 327 individuals, 84.2% report to be able and untagged, 9.7% to be disabled and untagged, 2.4% to be able and tagged and 3.6% to be disabled and tagged.

²⁷For instance, it follows from the normality of $g_A(\theta)$ that:

$$\Pr(\text{Tagged}|\text{Age} = 45, \text{Able}) = \Phi \left(\frac{\hat{\theta}_{45} - \mu/2}{1} \right).$$

The expression (26) is just a generalization that encompasses all possible ages and ability statuses.

μ and of $\{\hat{\theta}_t\}_{t \in [25,65]}$. Doing so, we obtain $\mu = 1.2329$. As shown in Figure 3, the estimated path of $\hat{\theta}_t$ is increasing with age. This is consistent with our assumption of a non-decreasing disability standard. The McFadden's pseudo R^2 for this regression is 19.9%.

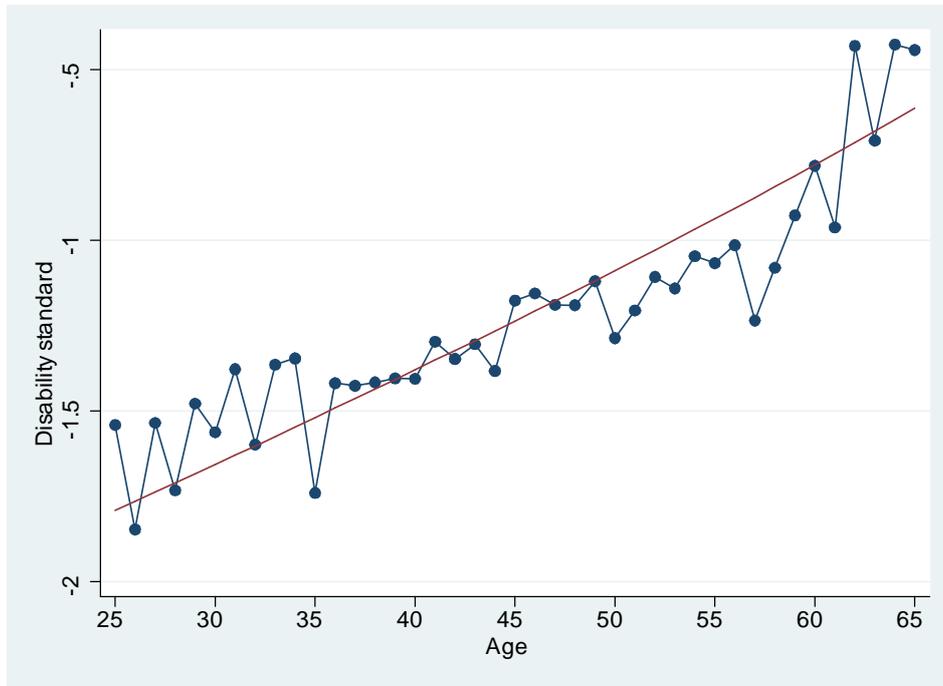


Figure 3: Disability standard

Note that both our theoretical model and empirical strategy rely on the assumption that the difference in means, μ , is the same at every age. To establish the validity of this claim, we run a probit regression where μ 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 μ is constant, we obtain a p-value of 0.813.

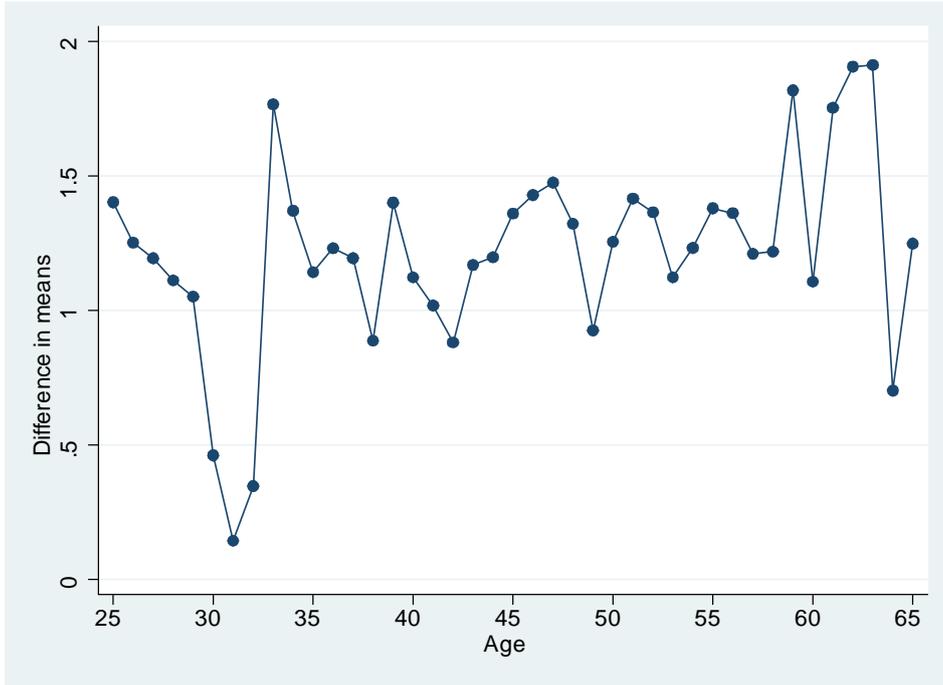


Figure 4: Difference in means

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 optimal allocation 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 , $\hat{\theta}_t$, is set.

4.1 Minimizing gaps and leakages

We consider the benchmark case where the path of the disability standard is set so 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 . 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:

$$\min_{\{\hat{\theta}_t\}_{t \in [0, H]}} \int_0^H \left\{ p_G F(t) \left[1 - G_D(\hat{\theta}_t) \right] + p_L [1 - F(t)] G_A(\hat{\theta}_t) \right\} dt, \quad (27)$$

where $F(t) \left[1 - G_D(\hat{\theta}_t) \right]$ and $[1 - F(t)] G_A(\hat{\theta}_t)$ 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:

$$p_G F(t) g_D(\hat{\theta}_t) = p_L [1 - F(t)] g_A(\hat{\theta}_t). \quad (28)$$

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

$$\hat{\theta}_t = \frac{1}{\mu} \ln \left[\frac{F(t)}{1 - F(t)} \right] + \frac{1}{\mu} \ln \left[\frac{p_G}{p_L} \right]. \quad (29)$$

Recall that the probit regression (26) from the last section yields the estimates for the disability standards displayed in Figure 3. To see whether they are consistent with the minimization of gaps and leakages, we add an age-specific error term to (29) and run an OLS regression of $\hat{\theta}_t$ on a constant and on the fitted values²⁸ 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 standard should increase slightly more rapidly with age than it currently does. However, if we run a constrained regression, which imposes that the slope coefficient must 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 policy of the U.S. Social Security system.

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 (29) with a relative price of gaps and leakages of 1.2.²⁹

The numerical simulation of the next subsection 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. Thus, if the optimal allocation is implemented, then it is desirable to decrease the strictness of the disability test.

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

²⁸We 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.

²⁹This measure of the strictness of the disability test is based on our PSID data set and, hence, implicitly, on the fraction of tagged individuals in the population. Since take-up is not systematic, this measure is not readily comparable to other estimates of the disability standard 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. Indeed, Weinzierl (2012) argues that most people are attached to the equal sacrifice principle which requires that tags should be as strongly correlated with the underlying income-earning ability as possible.

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 (29) with $p_G/p_L = 2.5$.

The consumption of the working and untagged, $c^{WU}(t)$, 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.

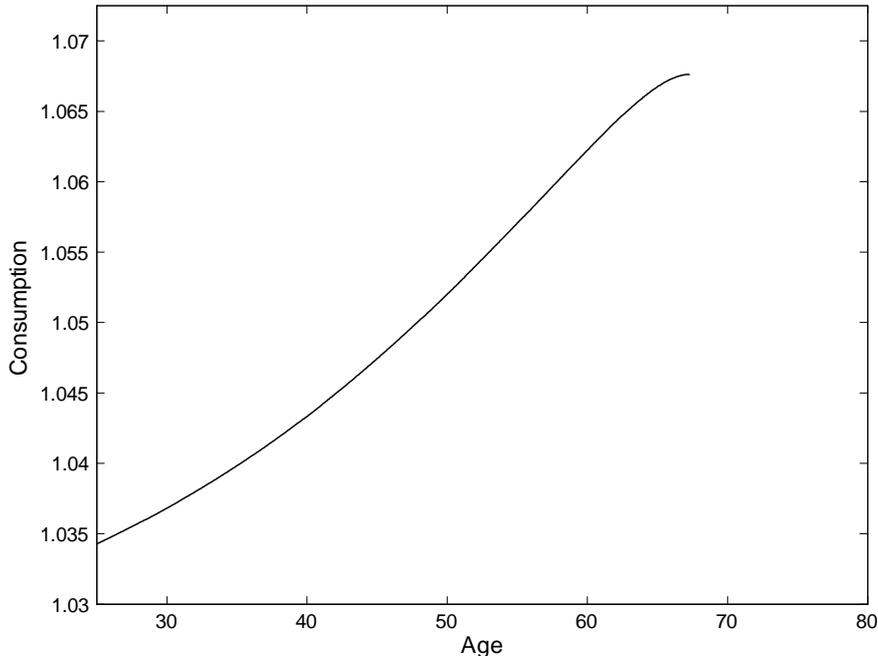


Figure 5: Consumption of the working and untagged

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, insurance against disability can, to a great extent, be provided by raising the consumption of the non-working tagged, which does not reduce incentives to work for the able and untagged. Hence, 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 $c^{NU}(t, r)$, the consumption of a non-working untagged individual as a function of his current age t and of the age r at which he ceased to work, with $t \geq r$. Recall that an untagged agent only retires before RU if he becomes disabled, i.e. $r = \min\{i, RU\}$. For a given retirement age r , the consumption of the agent is falling with age t and is minimal at 80 years old. To understand this pattern, which follows from the first-order condition (16), 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 level 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 at 80 years old, at the end of the life-cycle, serve as a threat and are therefore not welfare-reducing.

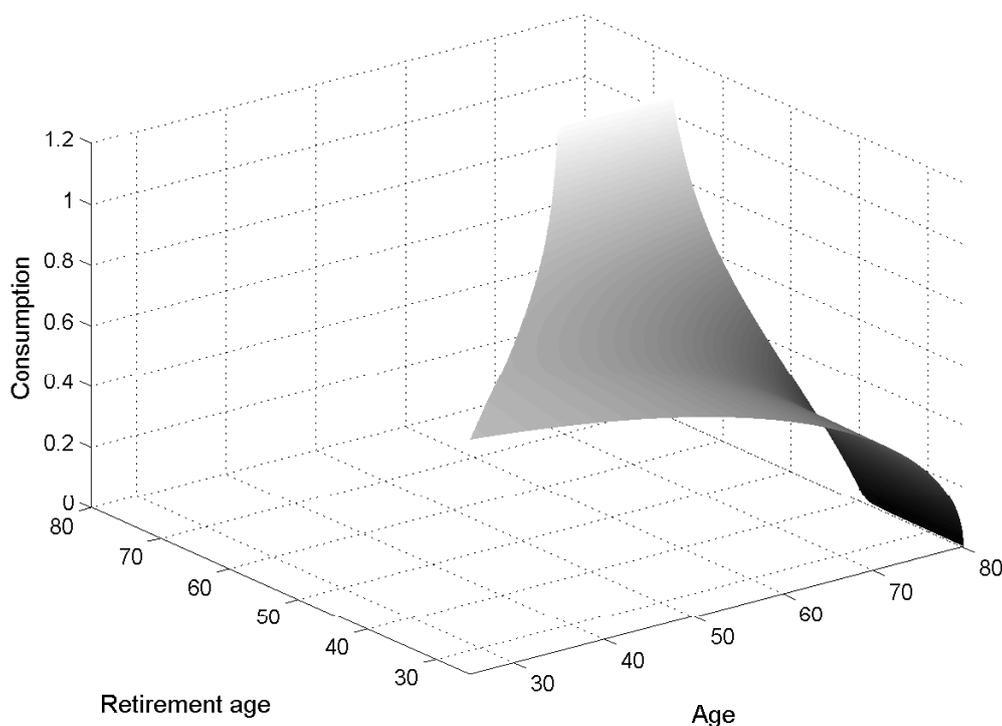


Figure 6: Consumption of the non-working and untagged

Figure 7 gives the consumption of a working and tagged individual, $c^{WT}(t, j)$, as a

function of his current age t and of the age j at which he became tagged, with $t \geq j$. For any given tag age j , consumption is increasing with age t . Again, the need to maintain incentives to work now and in the future makes back-loaded incentives particularly attractive.

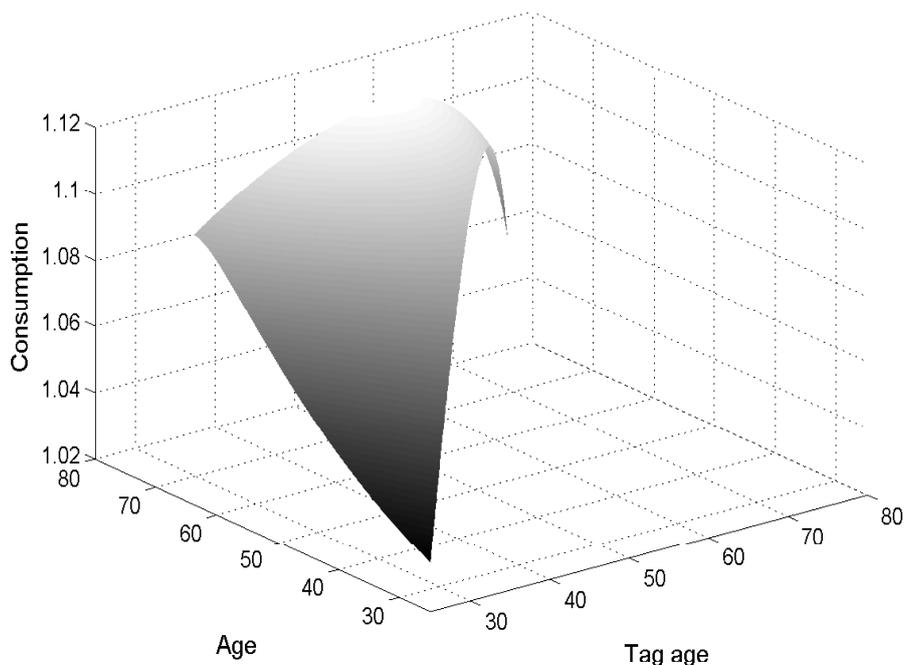


Figure 7: Consumption of the working and tagged

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 non-working and tagged which means that an even higher consumption level 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 of all the tagged is lower than that of the untagged.

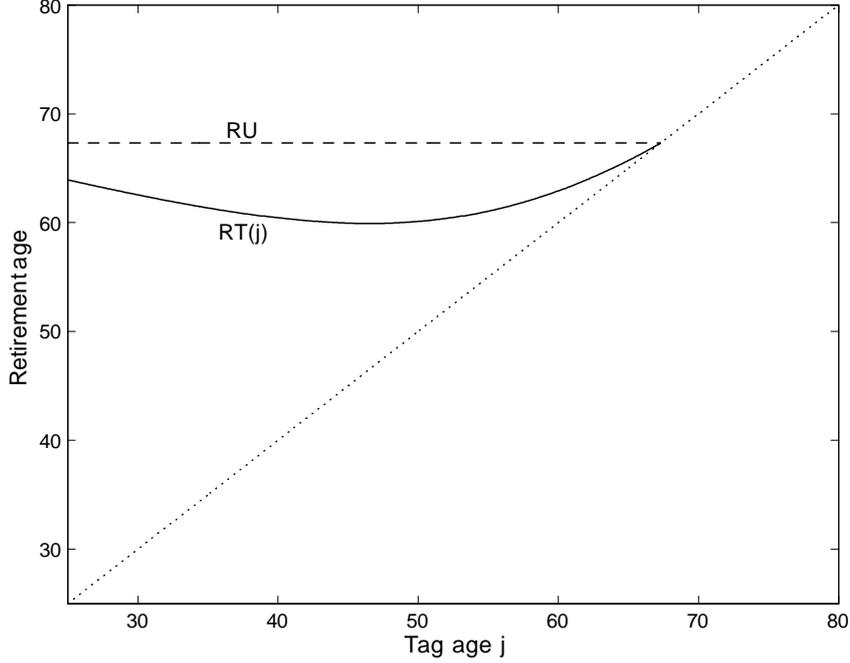


Figure 8: Retirement age

To understand why the retirement age is a U-shaped function of the tag age, recall from the first-order condition (20) 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, which implies that $c^{WU}(j)$ and $c^{WT}(j, j)$ are almost equal to each other. As j rises, $c^{WU}(j)$ increases which raises $c^{WT}(j, j)$. This makes back-loaded incentives so costly that it is optimal to reduce the retirement age. For higher tag ages, the tagged are more likely to be truly disabled and, hence, in equation (20), the increase in $c_T^{NT}(j, j)$ also contributes to match the increase in $c^{WU}(j)$. Thus, the increase in $c^{WT}(j, j)$ 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 $c^{WT}(j, j)$, 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 $c^{NT}(r, j)$, the consumption of the non-working and tagged who ceased to work at r and became tagged at j . Two sections are clearly distinguishable: $c_T^{NT}(r, j)$, $r \geq j$, on the left and $c_N^{NT}(r, j)$, $j > r$, 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 $c^{NT}(r, j)$ from $\max\{r, j\}$ until they die at age 80.

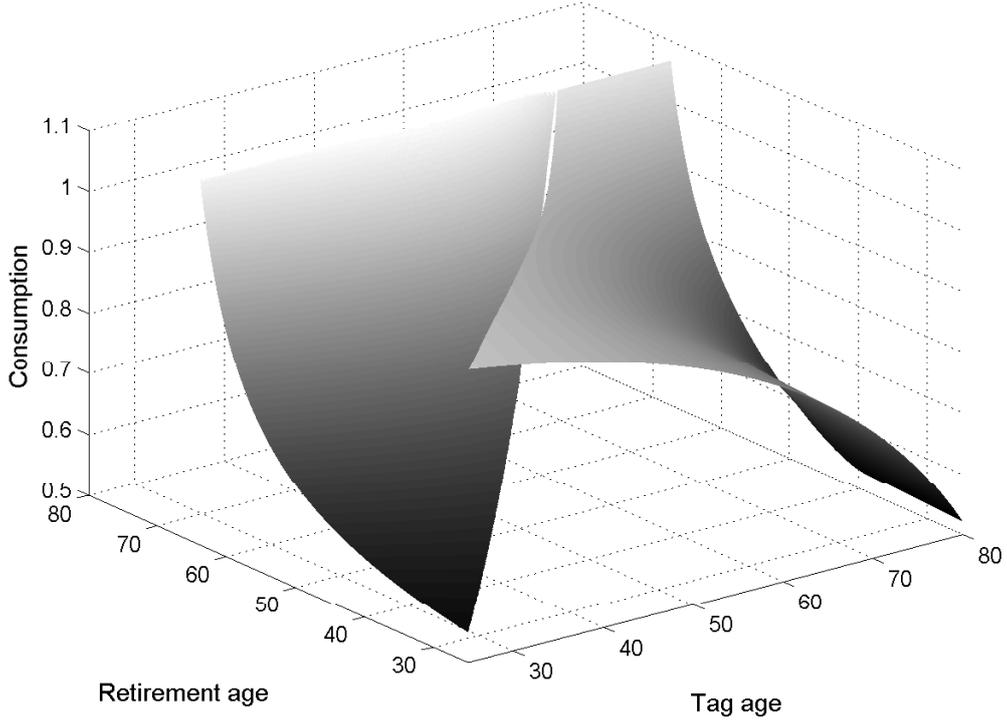


Figure 9: Consumption of the non-working and tagged

As argued above, it is desirable to provide back-loaded incentives to the working and tagged. But having an increasing consumption level for working individuals is not the only way to do so. In addition, the consumption of the non-working could be made higher, the later they ceased to work. This explains why, for a fixed tag age j , $c_T^{NT}(r, j)$ is increasing in r .

For an individual who stopped working before becoming tagged, his consumption once tagged is lower the later he became tagged, i.e. $c_N^{NT}(r, j)$ is decreasing in j . This follows from (18). The intuition for this is similar to that for $c^{NU}(t, r)$. 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 non-working who only get tagged much later serves as a threat to the able and untagged who might be tempted to claim to be disabled.

Turning to the diagonal of Figure 9, it is apparent that a higher consumption level is awarded if retirement occurs before the award of the tag. To understand this, note that a newly tagged worker who stops working gets consumption $c_T^{NT}(\cdot)$ immediately, while an untagged worker who stops working initially obtains $c^{NU}(\cdot)$ and is only likely to rapidly qualify for $c_N^{NT}(\cdot)$ if he is truly disabled. Thus, high values of $c_N^{NT}(\cdot)$ are not detrimental to the provision of incentives to work, while high values of $c_T^{NT}(\cdot)$ are.

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 stopped working. For this solution to work, the outcome θ of the test for a given individual should be exogenous to his action. A (computationally intensive) alternative would be to impose additional constraints to the planner's problem ensuring that individuals are always happy to be awarded the tag as soon as they are eligible. This would eliminate the discontinuity of $c^{NT}(\cdot)$. However, within our framework, imposing such extra constraints is not necessary and would come at the cost of reduced welfare.

The U.S. disability insurance program imposes a five-month waiting period out of the labor force before an individual can apply for the award of a tag. While we do not investigate the implementation of the optimal allocation in a decentralized economy, the above results show that the consumption levels of individuals who stop working should initially be as smooth as possible. This suggests that, in a decentralized economy, the waiting period can only be justified if agents have enough savings to sustain a decent consumption level before the award of disability benefits.

While it might not be justified to have a mandatory waiting period for all, the government should nevertheless make use of the information revealed by the existence of a waiting period for some workers. Indeed, Figure 9 shows that the provision of insurance should be more generous to those who stop working a few months before being awarded the tag.

4.3 Welfare gains

Our numerical simulations allow us to evaluate the welfare implications of the implementation of the optimal allocation. 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 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.³⁰ 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 $RU = RT(j) \equiv R, \forall j \in [0, RU)$. The optimal retirement age is then pinned down by the following condition,

$$b \left[\frac{1 - G_A(\hat{\theta}_R)}{u'(c^{WU}(R))} + \int_0^R \frac{g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj}}{u'(c^{WT}(R, j))} dj \right] = \gamma_R, \quad (30)$$

which replaces (15) and (21). A weighted average of the marginal rates of substitution

³⁰Again, 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.

between leisure and consumption should be equal to the marginal rate of transformation at retirement.

A policy yields consumption-equivalent 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 Table 1.

Table 1: Welfare gains compared to unobservable health

	Fixed retirement age	Health-dependent retirement age	First-best
$p_G/p_L = 2.5$	0.45%	0.64%	2.98%
$p_G/p_L = 1.2$	0.41%	0.56%	2.98%

In the first line the planner sets the optimal price of gaps and leakages.³¹ If, however, doctors are out of control and the government has to stick with the current disability standards, then the relevant results are those of the second line. The welfare gains generated by the imperfect information on health, equal to 0.64% when $p_G/p_L = 2.5$, are moderate but non-negligible. More than two thirds of these gains could be reaped with a fixed retirement age for all.

Clearly, from equation (29), as most people are able to work, the disability standards are 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 can come from two sources: improved insurance against the disability risk or improved incentives to work. The statistics on the average retirement age in Table 2, for the case $p_G/p_L = 2.5$, suggest that at least some of the gains come from better incentives to work.

Table 2: Retirement ages

	Unobservable health	Fixed retirement age	Health-dependent retirement age	First-best
Average retirement age	61.5	61.9	62.2	64.1
Maximum retirement age	65	65.4	67.3	68.4

³¹Note that the optimal relative price with a fixed retirement age is also approximately equal to 2.5.

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 at the onset of disability. 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 untagged 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 our optimal policy and that observed in the U.S. is that the able and tagged are currently not incentivized to work.³² 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 imposed.

Compared to the unobservable health case, the optimal policy under the constraint that all the tagged retire immediately, yields a welfare gain of 0.46% when $p_G/p_L = 1.2$.³³ It follows that switching from an optimal policy where the able and tagged do not work, i.e. $RT(j) = j$, to our optimal policy where they are incentivized to work up to a health-dependent retirement age generates welfare gains of only $0.56\% - 0.46\% = 0.10\%$. These welfare gains are even negative, and equal to $0.41\% - 0.46\% = -0.05\%$, with a common retirement age for all. 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; which somewhat qualifies the main message of Parsons (1996).

If the optimal relative price of gaps and leakages of 2.5 can be enforced, then our optimal policy generates a welfare gain of $0.64\% - 0.46\% = 0.18\%$ compared to the optimal policy under the constraints $RT(j) = j$ and $p_G/p_L = 1.2$. 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 for the sharp reduction in labor supply that this entails, the general retirement age RU needs to be pushed up to

³²The 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 and Knight, 2008). However, a similar policy in the U.S., Ticket to Work, failed to increase participation (Autor and Duggan, 2006, 2007).

³³The 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%.

72.1, which results in an average retirement age of only 61.0.

The continuous rise in the disability rolls over the past three decades has induced many policy makers to suggest a rise in the strictness of the disability standards. We find the opposite result that welfare would be increased by decreasing the strictness of the disability tests. However, crucially, this is only true provided that the able and tagged are induced to work up to some early retirement age.³⁴

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 our optimal policy consists in making strategic use of the difference between the age of occurrence of disability and the age of award of the tag. However, lacking a good benchmark representation of the current U.S. situation, it is not possible to isolate the corresponding welfare gains.

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 for gaps and leakages. Although this is a rather natural choice for the disability standard, we might be interested in determining the optimal allocation when the whole path of $\hat{\theta}_t$ is directly under the control of the planner. In fact, it turns out to be possible to asymptotically implement the first-best, perfect information, allocation by setting the disability standards $\{\hat{\theta}_t\}_{t \in [0, H]}$ 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 keep supplying labor until they reach the first-best retirement age, R^{FB} .

To prove that such an allocation can be asymptotically implemented, we propose a policy that does the job.³⁵ The planner should optimally award the tag as follows:

$$\hat{\theta}_t = \begin{cases} -\infty & \text{if } t \in [0, R^{FB}) \\ \hat{\theta} & \text{if } t = R^{FB} \\ +\infty & \text{if } t \in (R^{FB}, H] \end{cases}, \quad (31)$$

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

³⁴Low and Pistaferri (2012) have also reached the conclusion that the strictness of the disability standards should be decreased. In their structural model of the labor market and of the U.S. disability insurance program, they find that the welfare gains generated by a better provision of insurance against disability more than offset the corresponding reduced incentives to work.

³⁵Note that the precise characterization of such a policy, and in particular of the optimal path of $\hat{\theta}_t$, is not unique.

set:

$$c^{WU}(t) = c, \forall t \in [0, R^{FB}), \quad (32)$$

$$c^{NU}(t, r) = c, \forall r \in [0, R^{FB}), \forall t \in [r, R^{FB}] \quad (33)$$

$$c^{NT}(r, j) = \begin{cases} \delta & \text{if } r \in [0, R^{FB}) \text{ and } j > R^{FB} \\ c & \text{otherwise} \end{cases} \quad (34)$$

for some constants c and δ . The consumption of the working and tagged is irrelevant and does not need to be specified as people can only get tagged after retirement. Note that the consumption level δ only applies to those who retired before R^{FB} , who therefore claimed, rightly or wrongly, to be disabled, and who failed to be awarded the tag at R^{FB} . But, thanks to the monotone likelihood ratio property satisfied by $g_A(\theta)$ and $g_D(\theta)$, for a sufficiently high standard $\hat{\theta}$, 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 δ .

Proposition 2 *A policy characterized by (31), (32), (33) and (34) can be used to implement, asymptotically, the first-best allocation of resources. For that, choose δ , 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 $\hat{\theta} \rightarrow +\infty$, which implies $\delta \rightarrow 0$, $c \rightarrow c^{FB}$ and the ex-ante expected lifetime utility of workers converges to the first-best level.*

This proposition is formally proved in Appendix B. In a nutshell, the optimal policy is to shoot the liars. In particular, it should be emphasized that the low value of δ 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.

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 how high $\hat{\theta}$ is, the government can never be entirely sure that someone untagged at age R^{FB} 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 $\hat{\theta} = \bar{\theta}_D$, $\delta = 0$ 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 those who claimed to be disabled before R^{FB} and 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 the New Dynamic Public Finance literature, can 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.³⁶ 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 being able at age R^{FB} , is $G_A(\hat{\theta})$. In other words, they do not have any private information about whether they will be eligible at R^{FB} . While, as a first-order approximation to reality, this assumption is reasonable, a small departure from it could prevent the implementation of the 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} if the threshold $\hat{\theta}$ is very high.

While the first-best implementation result is primarily a theoretical result, it nevertheless suggests that the government can obtain substantial welfare gains by moving beyond the minimization of gaps and leakages. For instance, if, starting from a lower level, the disability standard was increasing even more rapidly with age than it currently does³⁷, then the tag would often be awarded later in life. This would be welfare-enhancing as the threat of not being tagged when old would deter the temptation to claim to be disabled when young. Moreover, few young and able workers would be tagged which would make it unnecessary to give them special rewards for participating in the labor market.

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 against disability, the planner offers back-loaded incentives and makes strategic use of the difference between the age of occurrence of disability and the age of 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 provide them with 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

³⁶See also Varian (1980).

³⁷Note that this implies raising the price of gaps relative to that of leakages as age increases.

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 against disability. Moreover, additional welfare gains could be obtained by moving beyond the minimization of classification errors and by setting the disability standards 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 determine how it could be implemented in a decentralized economy with private capital markets. 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) showed 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. Similarly, Diamond and Mirrlees (1995) allowed for hidden private savings within their original framework. However, to the extent that the main features of our optimal policy could generate significant welfare gains, the government should avoid imposing restrictions on policy instruments that would prevent the realization of these gains. This suggests that our main qualitative insights would remain relevant, even under reasonable additional constraints on policy instruments.

References

- [1] Adam, S., Bozio, A., Emmerson, C., Greenberg, D. and Knight, G. (2008), 'A cost-benefit analysis of Pathways to Work for new and repeat incapacity benefits claimants', Research Report No 498, Department for Work and Pensions.
- [2] Akerlof, G.A. (1978), 'The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs and Manpower Planning', *American Economic Review*, 68(1), 8-19.
- [3] Albanesi, S. and Sleet, C. (2006), 'Dynamic Optimal Taxation with Private Information', *Review of Economic Studies*, 73(1), 1-30.

- [4] Alesina, A., Ichino, A. and Karabarbounis, L. (2011), ‘Gender-Based Taxation and the Division of Family Chores’, *American Economic Journal: Economic Policy*, 3(2), 1-40.
- [5] Autor, D.H. and Duggan, M.G. (2006), ‘The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding’, *Journal of Economic Perspectives*, 20(3), 71-96.
- [6] Autor, D.H. and Duggan, M.G. (2007), ‘Distinguishing Income from Substitution Effects in Disability Insurance’, *American Economic Review Papers and Proceedings*, 97(2), 119-124.
- [7] Benitez-Silva, H., Buchinsky, M., Chan, H.M., Cheidvasser, S. and Rust, J. (2004), ‘How Large is the Bias in Self-Reported Disability?’ *Journal of Applied Econometrics*, 19(6), 649-70.
- [8] Benitez-Silva, H., Buchinsky, M. and Rust, J. (2006), ‘How Large are the Classification Errors in the Social Security Disability Award Process?’, Working Paper, SUNY-Stony Brook.
- [9] Chandra, A. and Samwick, A.A. (2006), ‘Disability Risk and the Value of Disability Insurance’, in *Health at Older Ages: The Causes and Consequences of Declining Disability Among the Elderly*, edited by D.M. Cutler and D.A. Wise, Chicago: Chicago University Press.
- [10] Diamond, P.A. and Mirrlees, J.A. (1978), ‘A Model of Social Insurance with Variable Retirement’, *Journal of Public Economics*, 10, 295-336.
- [11] Diamond, P.A. and Mirrlees, J.A. (1986), ‘Payroll-Tax Financed Social Insurance with Variable Retirement’, *Scandinavian Journal of Economics*, 88(1), 25-50.
- [12] Diamond, P.A. and Mirrlees, J.A. (1995), ‘Social Insurance with Variable Retirement and Private Saving’, Working Paper, MIT.
- [13] Diamond, P.A. and Sheshinski, E. (1995), ‘Economic Aspects of Optimal Disability Benefits’, *Journal of Public Economics*, 57, 1-23.
- [14] Farhi, E. and Werning, I. (2012), ‘Insurance and Taxation over the Life Cycle’, *Review of Economic Studies*, forthcoming.
- [15] Finkelstein, A., Luttmer, E.F.P. and Notowidigdo, M.J. (2013), ‘What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption’, *Journal of the European Economic Association*, 11(1), 221-258.

- [16] Golosov, M., Kocherlakota, N. and Tsyvinski, A. (2003), ‘Optimal Indirect and Capital Taxation’, *Review of Economic Studies*, 70(3), 569-587.
- [17] Golosov, M. and Tsyvinski, A. (2004), ‘Designing Optimal Disability Insurance: A Case for Asset Testing’, NBER Working Paper 10792.
- [18] Golosov, M. and Tsyvinski, A. (2006), ‘Designing Optimal Disability Insurance: A Case for Asset Testing’, *Journal of Political Economy*, 114(2), 257-279.
- [19] Grochulski, B. and Kocherlakota, N. (2010), ‘Nonseparable Preferences and Optimum Social Security Systems’, *Journal of Economic Theory*, 145, 2055-2077.
- [20] Kleven, H.J. and Kopczuk, W. (2011), ‘Transfer Program Complexity and the Take-Up of Social Benefits’, *American Economic Journal: Economic Policy*, 3(1), 54-90.
- [21] Li, X. and Maestas, N. (2008), ‘Does the Rise in the Full Retirement Age Encourage Disability Benefits Applications? Evidence from the Health and Retirement Study’, Working Paper, Michigan Retirement Research Center.
- [22] Liebman, J.B., Luttmer, E.F.P. and Seif, D.G. (2009), ‘Labor Supply Responses to Marginal Social Security Benefits: Evidence from Discontinuities’, *Journal of Public Economics*, 93, 1208-1223.
- [23] Ljungqvist, L. and Sargent, T. (2006), ‘Do Taxes Explain European Unemployment? Indivisible Labor, Human Capital, Lotteries, and Savings’, in *NBER Macroeconomics Annuals 2006*, edited by D. Acemoglu, K. Rogoff and M. Woodford, Cambridge, MA: MIT Press.
- [24] Low, H., and Pistaferri, L. (2012), ‘Disability Insurance and the Dynamics of the Incentive-Insurance Tradeoff’, Working Paper, University of Cambridge and Stanford.
- [25] Mankiw, N.G. and Weinzierl, M. (2010), ‘The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution’, *American Economic Journal: Economic Policy*, 2(1), 155-176.
- [26] Michau, J.B. (2012), ‘Optimal Redistribution with Intensive and Extensive Labor Supply Margins: A Life-Cycle Perspective’, Working Paper, Ecole Polytechnique.
- [27] Mirrlees, J.A. (1974), ‘Notes on Welfare Economics, Information and Uncertainty’, in *Essays in Equilibrium Behavior and Uncertainty*, edited by M. Balch, D. McFadden and S. Wu, Amsterdam: North Holland.

- [28] Mirrlees, J.A. (1999), ‘The Theory of Moral Hazard and Unobservable Behaviour: Part I’, *Review of Economic Studies*, 66(1), 3-21.
- [29] Mulligan, C. (2001), ‘Aggregate Implications of Indivisible Labor’, *Advances in Macroeconomics*, 1(1).
- [30] Parsons, D.O. (1996), ‘Imperfect ‘Tagging’ in Social Insurance Programs’, *Journal of Public Economics*, 62, 183-207.
- [31] Rios-Rull, J.V. (1996), ‘Life-Cycle Economies and Aggregate Fluctuations’, *Review of Economic Studies*, 63(3), 465-89.
- [32] Salanie, B. (2002), ‘Optimal Demogrants with Imperfect Tagging’, *Economic Letters*, 75, 319-324.
- [33] Shourideh, A. and Troshkin, M. (2012), ‘Providing Efficient Incentives to Work: Retirement Ages and the Pension System’, Working Paper, Wharton and Cornell.
- [34] SSA (U.S. Social Security Administration) (2008), *Social Security Bulletin: Annual Statistical Supplement*, Washington DC: Social Security Administration.
- [35] Varian, H.R. (1980), ‘Redistributive Taxation as Social Insurance’, *Journal of Public Economics*, 14, 49-68.
- [36] Weinzierl, M. (2011), ‘The Surprising Power of Age-Dependent Taxes’, *Review of Economic Studies*, 78(4), 1490-1518.
- [37] Weinzierl, M. (2012), ‘Why do we Redistribute so Much but Tag so Little? Normative Diversity, Equal Sacrifice and Optimal Taxation’, Working Paper, Harvard.

A The planner's problem

As explained in the text, cf. (12), the planner maximizes the *ex-ante* expected lifetime utility of agents subject to a resource constraint and to a set of incentive compatibility constraints which ensures that the able choose to work until they reach the relevant retirement age. In this appendix, we provide the detailed expressions of the planner's problem. His objective, which can be derived by substituting (1), (2), (3), (4) and (5) into (6), is to maximize:

$$\begin{aligned}
\mathbb{E} [v(i, j)] = & \tag{A1} \\
& \int_0^{RU} e^{-\rho t} [u(c^{WU}(t)) - b] [1 - G_A(\hat{\theta}_t)] [1 - F(t)] dt \\
& + \int_0^{RU} \int_j^{RT(j)} e^{-\rho t} [u(c^{WT}(t, j)) - b] g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(t)] dt dj \\
& + \int_0^{RT(0)} e^{-\rho t} [u(c^{WT}(t, 0)) - b] G_A(\hat{\theta}_0) [1 - F(t)] dt \\
& + \int_0^{RU} \int_i^H e^{-\rho t} u(c^{NU}(t, i)) [1 - G_D(\hat{\theta}_t)] f(i) dt di \\
& + \int_{RU}^H e^{-\rho t} u(c^{NU}(t, RU)) \left[[1 - G_A(\hat{\theta}_t)] [1 - F(t)] + [1 - G_D(\hat{\theta}_t)] [F(t) - F(RU)] \right] dt \\
& + \int_0^{RU} \int_i^H \left[\int_j^H e^{-\rho t} dt \right] u(c_N^{NT}(i, j)) g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) dj di \\
& + \int_{RU}^H \left[\int_j^H e^{-\rho t} dt \right] u(c_N^{NT}(RU, j)) \\
& \quad \times \left[g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) + g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [F(j) - F(RU)] \right] dj \\
& + \int_0^{RU} \int_j^{RT(j)} \left[\int_i^H e^{-\rho t} dt \right] u(c_T^{NT}(i, j)) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) di dj \\
& + \int_0^{RT(0)} \left[\int_i^H e^{-\rho t} dt \right] u(c_T^{NT}(i, 0)) G_A(\hat{\theta}_0) f(i) di \\
& + \int_0^{RU} \left[\int_{RT(j)}^H e^{-\rho t} dt \right] u(c_T^{NT}(RT(j), j)) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(RT(j))] dj \\
& + \left[\int_{RT(0)}^H e^{-\rho t} dt \right] u(c_T^{NT}(RT(0), 0)) G_A(\hat{\theta}_0) [1 - F(RT(0))] \\
& + \int_0^{RU} \left[\int_j^H e^{-\rho t} dt \right] u(c_T^{NT}(j, j)) [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) dj.
\end{aligned}$$

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.

Similarly, the full expression for the resource constraint is:

$$\begin{aligned}
& \int_0^{RU} e^{-\rho t} [c^{WU}(t) - \gamma_t] [1 - G_A(\hat{\theta}_t)] [1 - F(t)] dt \tag{A2} \\
& + \int_0^{RU} \int_j^{RT(j)} e^{-\rho t} [c^{WT}(t, j) - \gamma_t] g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(t)] dt dj \\
& + \int_0^{RT(0)} e^{-\rho t} [c^{WT}(t, 0) - \gamma_t] G_A(\hat{\theta}_0) [1 - F(t)] dt \\
& + \int_0^{RU} \int_i^H e^{-\rho t} c^{NU}(t, i) [1 - G_D(\hat{\theta}_t)] f(i) dt di \\
& + \int_{RU}^H e^{-\rho t} c^{NU}(t, RU) \left[[1 - G_A(\hat{\theta}_t)] [1 - F(t)] + [1 - G_D(\hat{\theta}_t)] [F(t) - F(RU)] \right] dt \\
& + \int_0^{RU} \int_i^H \left[\int_j^H e^{-\rho t} dt \right] c_N^{NT}(i, j) g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) dj di \\
& + \int_{RU}^H \left[\int_j^H e^{-\rho t} dt \right] c_N^{NT}(RU, j) \\
& \quad \times \left[g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) + g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [F(j) - F(RU)] \right] dj \\
& + \int_0^{RU} \int_j^{RT(j)} \left[\int_i^H e^{-\rho t} dt \right] c_T^{NT}(i, j) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) di dj \\
& + \int_0^{RT(0)} \left[\int_i^H e^{-\rho t} dt \right] c_T^{NT}(i, 0) G_A(\hat{\theta}_0) f(i) di \\
& + \int_0^{RU} \left[\int_{RT(j)}^H e^{-\rho t} dt \right] c_T^{NT}(RT(j), j) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(RT(j))] dj \\
& + \left[\int_{RT(0)}^H e^{-\rho t} dt \right] c_T^{NT}(RT(0), 0) G_A(\hat{\theta}_0) [1 - F(RT(0))] \\
& + \int_0^{RU} \left[\int_j^H e^{-\rho t} dt \right] c_T^{NT}(j, j) [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) dj \leq 0.
\end{aligned}$$

The incentive compatibility constraint inducing the able and untagged of age $s \in [0, RU)$ to work is:

$$\begin{aligned}
& \int_s^{RU} e^{-\rho t} [u(c^{WU}(t)) - b] [1 - G_A(\hat{\theta}_t)] [1 - F(t)] dt \\
& + \int_s^{RU} \int_j^{RT(j)} e^{-\rho t} [u(c^{WT}(t, j)) - b] g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(t)] dt dj \\
& + \int_s^{RU} \int_i^H e^{-\rho t} u(c^{NU}(t, i)) [1 - G_D(\hat{\theta}_t)] f(i) dt di \\
& + \int_s^{RU} e^{-\rho t} u(c^{NU}(t, RU)) \left[[1 - G_A(\hat{\theta}_t)] [1 - F(t)] + [1 - G_D(\hat{\theta}_t)] [F(t) - F(RU)] \right] dt \\
& + \int_s^{RU} \int_i^H \left[\int_j^H e^{-\rho t} dt \right] u(c_N^{NT}(i, j)) g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) dj di \\
& + \int_s^{RU} \left[\int_j^H e^{-\rho t} dt \right] u(c_N^{NT}(RU, j)) \\
& \quad \times \left[g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) + g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [F(j) - F(RU)] \right] dj \\
& + \int_s^{RU} \int_j^{RT(j)} \left[\int_i^H e^{-\rho t} dt \right] u(c_T^{NT}(i, j)) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} f(i) di dj \\
& + \int_s^{RU} \left[\int_{RT(j)}^H e^{-\rho t} dt \right] u(c_T^{NT}(RT(j), j)) g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(RT(j))] dj \\
& + \int_s^{RU} \left[\int_j^H e^{-\rho t} dt \right] u(c_T^{NT}(j, j)) [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) dj \\
& \geq \int_s^H e^{-\rho t} u(c^{NU}(t, s)) \left[[1 - G_A(\hat{\theta}_t)] [1 - F(t)] + [1 - G_D(\hat{\theta}_t)] [F(t) - F(s)] \right] dt \\
& + \int_s^H \left[\int_j^H e^{-\rho t} dt \right] u(c_N^{NT}(s, j)) \\
& \quad \times \left[g_A(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [1 - F(j)] + [G_D(\hat{\theta}_j) - G_A(\hat{\theta}_j)] f(j) + g_D(\hat{\theta}_j) \frac{d\hat{\theta}_j}{dj} [F(j) - F(s)] \right] dj.
\end{aligned} \tag{A3}$$

The left hand side is equal to $\int_s^H \int_s^H v_s(i, j) f(i, j) di dj = \mathbb{E}[v_s(i, j) | i > s, j > s]$ $\left(\int_s^H \int_s^H f(i, j) di dj \right)$, while the right hand side is equal to $\int_s^H \int_s^H v_s(s, j) f(i, j) di dj = \mathbb{E}[v_s(s, j) | i > s, j > s]$ $\left(\int_s^H \int_s^H f(i, j) di dj \right)$.³⁸ As could be seen from the right hand side of the constraint, agents who deviate and claim to be disabled get consumption $c^{NU}(t, s)$ until tagged at $j > s$ and $c_N^{NT}(s, j)$ thereafter.

Finally, the able who obtained the tag at age j should be incentivized to work until

³⁸Recall that $v_s(i, j)$ denotes the *ex-post* lifetime utility of individual (i, j) from age s onwards.

age $RT(j)$. The corresponding constraint at age $s \geq j$ is:

$$\begin{aligned}
& \int_s^{RT(j)} e^{-\rho t} [u(c^{WT}(t, j)) - b] [1 - F(t)] dt \\
& + \int_s^{RT(j)} \left[\int_i^H e^{-\rho t} dt \right] u(c_T^{NT}(i, j)) f(i) di \\
& + \left[\int_{RT(j)}^H e^{-\rho t} dt \right] u(c_T^{NT}(RT(j), j)) [1 - F(RT(j))] \\
& \geq \left[\int_s^H e^{-\rho t} dt \right] u(c_T^{NT}(s, j)) [1 - F(s)].
\end{aligned} \tag{A4}$$

The left hand side is equal to $\int_s^H v_s(i, j) f(i) di = \mathbb{E}[v_s(i, j) | i > s] \left(\int_s^H f(i) di \right)$, while the right hand side is equal to $\int_s^H v_s(s, j) f(i) di = \mathbb{E}[v_s(s, j) | i > s] \left(\int_s^H f(i) di \right) = v_s(s, j) [1 - F(s)]$.³⁹ Since, once an agent is tagged, the planner 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 $c^{WU}(\cdot)$, $c^{WT}(\cdot)$, $c^{NU}(\cdot)$, $c_N^{NT}(\cdot)$, $c_T^{NT}(\cdot)$, $RT(\cdot)$ and RU subject to (A2), (A3) $\forall s \in [0, RU)$ and (A4) $\forall j \in [0, RU)$, $\forall s \in [j, RT(j))$.

B Proof of Proposition 2

The probability that an agent gets tagged at R^{FB} is given by:

$$f(j = R^{FB} | i) = \begin{cases} G_D(\hat{\theta}) & \text{if } i \leq R^{FB} \\ G_A(\hat{\theta}) & \text{if } i > R^{FB} \end{cases}.$$

It follows, by (31), that the probability that an agent gets tagged immediately after R^{FB} is equal to:

$$f(j > R^{FB} | i) = 1 - f(j = R^{FB} | i).$$

The *ex-post* lifetime utility of an (i, j) individual who becomes disabled before R^{FB} and who only gets tagged after R^{FB} , i.e. $i < R^{FB}$ and $j > R^{FB}$, is given by:

$$v(i, j) = \int_0^i e^{-\rho t} [u(c) - b] dt + \int_i^{R^{FB}} e^{-\rho t} u(c) dt + \int_{R^{FB}}^H e^{-\rho t} u(\delta) dt.$$

³⁹Once an agent is retired and tagged, the occurrence of disability can no longer affect his welfare. Hence, $\mathbb{E}[v_s(s, j) | i > s] = v_s(s, j)$.

The *ex-post* lifetime utility of all other individuals is equal to:

$$v(i, j) = \int_0^{\min\{i, R^{FB}\}} e^{-\rho t} [u(c) - b] dt + \int_{\min\{i, R^{FB}\}}^H e^{-\rho t} u(c) dt.$$

The *ex-ante* expected lifetime utility of workers is equal to:

$$\begin{aligned} \mathbb{E}[v(i, j)] &= \int_0^H [v(i, j = R^{FB})f(j = R^{FB}|i) + v(i, j > R^{FB})f(j > R^{FB}|i)] f(i) di \\ &= \int_0^{R^{FB}} [v(i, j = R^{FB})G_D(\hat{\theta}) + v(i, j > R^{FB}) [1 - G_D(\hat{\theta})]] f(i) di \\ &\quad + \int_{R^{FB}}^H [v(i, j = R^{FB})G_A(\hat{\theta}) + v(i, j > R^{FB}) [1 - G_A(\hat{\theta})]] f(i) di \\ &= \left(\int_0^H e^{-\rho t} dt \right) u(c) - \left(\int_{R^{FB}}^H e^{-\rho t} dt \right) [u(c) - u(\delta)] [1 - G_D(\hat{\theta})] F(R^{FB}) \\ &\quad - \left[\int_0^{R^{FB}} \left(\int_0^i e^{-\rho t} b dt \right) f(i) di + \left(\int_0^{R^{FB}} e^{-\rho t} b dt \right) [1 - F(R^{FB})] \right]. \end{aligned} \quad (\text{B1})$$

Similarly, the resource constraint, $\mathbb{E}[z(i, j)] \leq 0$, can be written as:

$$\begin{aligned} &\left(\int_0^H e^{-\rho t} dt \right) c - \left(\int_{R^{FB}}^H e^{-\rho t} dt \right) [c - \delta] [1 - G_D(\hat{\theta})] F(R^{FB}) \\ &\quad - \left[\int_0^{R^{FB}} \left(\int_0^i e^{-\rho t} \gamma_t dt \right) f(i) di + \left(\int_0^{R^{FB}} e^{-\rho t} \gamma_t dt \right) [1 - F(R^{FB})] \right] \leq 0. \end{aligned}$$

Note that the last term, in the square bracket, is the average lifetime production in the first-best allocation. It must therefore be equal to the average lifetime consumption in the first-best allocation. Hence, the resource constraint can be rewritten as:

$$\left(\int_0^H e^{-\rho t} dt \right) c - \left(\int_{R^{FB}}^H e^{-\rho t} dt \right) [c - \delta] [1 - G_D(\hat{\theta})] F(R^{FB}) \leq \left(\int_0^H e^{-\rho t} dt \right) c^{FB},$$

or, equivalently, as:

$$\left(\int_0^H e^{-\rho t} dt \right) [c - c^{FB}] \leq \left(\int_{R^{FB}}^H e^{-\rho t} dt \right) [c - \delta] [1 - G_D(\hat{\theta})] F(R^{FB}). \quad (\text{B2})$$

Finally, the incentive compatibility constraint of the untagged of age t , where $t \in [0, R^{FB})$, is:⁴⁰

$$\mathbb{E}[v_t(i, j)|i > t] \geq \mathbb{E}[v_t(t, j)|i > t],$$

⁴⁰ As agents cannot be tagged before they reach the retirement age R^{FB} , there is no incentive compatibility constraints for the tagged.

where $v_t(i, j)$ denotes the *ex-post* lifetime utility of individual (i, j) from age t onwards. Proceeding as in the derivation of the objective, it can easily be shown that:

$$\begin{aligned}\mathbb{E}[v_t(i, j)|i > t] &= \left(\int_t^H e^{-\rho s} ds \right) u(c) \\ &\quad - \left(\int_{R^{FB}}^H e^{-\rho s} ds \right) [u(c) - u(\delta)] [1 - G_D(\hat{\theta})] \frac{F(R^{FB}) - F(t)}{1 - F(t)} \\ &\quad - \left[\int_t^{R^{FB}} \left(\int_t^i e^{-\rho s} b ds \right) \frac{f(i)}{1 - F(t)} di + \left(\int_t^{R^{FB}} e^{-\rho s} b ds \right) \frac{1 - F(R^{FB})}{1 - F(t)} \right].\end{aligned}$$

We also have:

$$\begin{aligned}\mathbb{E}[v_t(t, j)|i > t] &= \int_t^{R^{FB}} \left[v_t(t, j = R^{FB}) G_D(\hat{\theta}) + v_t(t, j > R^{FB}) [1 - G_D(\hat{\theta})] \right] \frac{f(i)}{1 - F(t)} di \\ &\quad + \int_{R^{FB}}^H \left[v_t(t, j = R^{FB}) G_A(\hat{\theta}) + v_t(t, j > R^{FB}) [1 - G_A(\hat{\theta})] \right] \frac{f(i)}{1 - F(t)} di \\ &= \left(\int_t^H e^{-\rho s} ds \right) u(c) \\ &\quad - \left(\int_{R^{FB}}^H e^{-\rho s} ds \right) [u(c) - u(\delta)] [1 - G_D(\hat{\theta})] \frac{F(R^{FB}) - F(t)}{1 - F(t)} \\ &\quad - \left(\int_{R^{FB}}^H e^{-\rho s} ds \right) [u(c) - u(\delta)] [1 - G_A(\hat{\theta})] \frac{1 - F(R^{FB})}{1 - F(t)}.\end{aligned}$$

Hence, after straightforward simplifications, the incentive compatibility constraint for the untagedged of age t is:

$$\begin{aligned}\left(\int_{R^{FB}}^H e^{-\rho s} ds \right) [u(c) - u(\delta)] [1 - G_A(\hat{\theta})] [1 - F(R^{FB})] \\ \geq \left[\int_t^{R^{FB}} \left(\int_t^i e^{-\rho s} b ds \right) f(i) di + \left(\int_t^{R^{FB}} e^{-\rho s} b ds \right) [1 - F(R^{FB})] \right].\end{aligned}$$

The right hand side of the constraint is decreasing in t , while the left hand side is independent of t . Hence, if this constraint is satisfied for $t = 0$, then it is satisfied for all values of $t \in [0, R^{FB})$. We can therefore focus on this constraint at time 0, which can be written as:

$$[u(c) - u(\delta)] [1 - G_A(\hat{\theta})] \geq X, \quad (\text{B3})$$

where, by definition:

$$X = \frac{\left[\int_0^{R^{FB}} \left(\int_0^i e^{-\rho s} b ds \right) \frac{f(i)}{1 - F(R^{FB})} di + \left(\int_0^{R^{FB}} e^{-\rho s} b ds \right) \right]}{\int_{R^{FB}}^H e^{-\rho s} ds}.$$

Note that $X > 0$.

For any given value of $\hat{\theta}$, the consumption levels defined by (32), (33) and (34) characterize an incentive-feasible allocation of resources provided that c and δ satisfy the resource constraint (B2) and the incentive compatibility constraint (B3). Let us therefore consider that c and δ are jointly determined by these two constraints holding with equality. It immediately follows that, as $\hat{\theta}$ tends to $+\infty$, δ tends to 0 and c tends to c^{FB} .⁴¹

Finally, we need to show that welfare converges to its first-best level. Using the incentive compatibility constraint (B3), the second term of the objective (B1) can be written as:

$$\left(\int_{R^{FB}}^H e^{-\rho t} dt \right) [u(c) - u(\delta)] [1 - G_D(\hat{\theta})] F(R^{FB}) = \left(\int_{R^{FB}}^H e^{-\rho t} dt \right) X \frac{1 - G_D(\hat{\theta})}{1 - G_A(\hat{\theta})} F(R^{FB}).$$

The monotone likelihood ratio property implies that $[1 - G_D(\hat{\theta})] / [1 - G_A(\hat{\theta})]$ tends to 0 as $\hat{\theta}$ tends to $+\infty$. It immediately follows that the value of the objective (B1) converges to its first-best level as $\hat{\theta}$ tends to $+\infty$.

⁴¹The incentive constraint implies that $c > \delta$. Thus, by the resource constraint $c \geq c^{FB}$. For simplicity, we can write the resource constraint as $c - \delta = [c^{FB} - \delta] Y$ where $Y \in [0, 1)$. This expression can be rewritten as $c - \delta = [c^{FB} - \delta] / [1 - Y]$. We also know that $\delta \geq 0$, since consumption cannot be negative. Combining these last four expressions yields $0 \leq c - c^{FB} = [c - \delta] Y = [c^{FB} - \delta] Y / [1 - Y] \leq c^{FB} Y / [1 - Y]$. But, Y tends to 0 as $\hat{\theta}$ tends to infinity. Thus, the right hand side of the inequality tends to 0 and, hence, c tends to c^{FB} as $\hat{\theta}$ tends to infinity. If $G_A(\hat{\theta})$ tends to one and $u(c)$ tends to $u(c^{FB})$, then the incentive compatibility constraint implies $u(\delta)$ must tend to $-\infty$. Hence, δ tends to 0 as $\hat{\theta}$ tends to infinity.