

# Welfare programs and motivation bias of social workers\*

Pierre C. Boyer<sup>†</sup> and Jean-Marie Lozachmeur<sup>‡</sup>

January 30, 2013

## Preliminary and Incomplete

### Abstract

This paper studies optimal disability and welfare benefits with imperfect disability tagging. Labor supply is affected by the levels of both disability and welfare benefits. The tagging process is implemented by social workers that may have some altruism toward claimants. All the individuals that are not working are eligible for welfare benefits. We analyze the optimal structure of benefits and the implications of a reform aimed at raising the standard for being eligible for disability benefits.

*Keywords:* Welfare programs; Disability insurance; Social workers; Altruism.

*JEL classification:* H21; H51; H53.

---

\*We thank Vincenzo Galasso and Hans Peter Grüner for very helpful comments and discussions. Boyer thanks the DONDENA Center at Bocconi University for its hospitality. Boyer gratefully acknowledges the Collaborative Research Center 884 for financial support. The usual disclaimer applies.

<sup>†</sup>University of Mannheim, Department of Economics. E-mail: pierre.boyer@uni-mannheim.de

<sup>‡</sup>Toulouse School of Economics (CNRS-GREMAQ and IDEI). E-mail: jean-marie.lozachmeur@tse-fr.eu

# 1 Introduction

Disability insurance (DI) programs are under scrutiny due to their rapid increase in size (number of recipients) and their costs on the public finance. In the United States, since its creation in 1956, the DI roll has grown at a fast rate to become one of the largest social insurance programs (see Bound, 1989; Autor and Duggan, 2006). In December 2010, the US Social Security Administration reported 8.295 million of disabled workers enrolled in the program receiving 105.132 billion dollars of benefits.<sup>1</sup> Similar trends have been observed in Europe (see, e.g., OECD, 2003). Moreover, current pension reforms aimed at increasing the statutory retirement age may lead to an increase in the population of old workers that could become eligible for disability benefits.<sup>2</sup>

The problems associated with disability programs are linked to the difficulties to screen or tag correctly the claimants. The disadvantages of imperfect tagging are the perverse incentives to people to be identified as needy, i.e. to be tagged, the horizontal inequity of such a system with people of identical health status receiving different levels of transfers, and its costs of administration (see Akerlof, 1978). In addition, voters' support for disability programs would be easier to gain if it could be shown that the recipients were always needy. There is evidence that significant tagging errors occur in welfare systems. Type I errors (rejection errors) occur if an eligible individual applying for benefits is rejected instead of being tagged into the disability program. Type II errors (award errors) occur if a truly ineligible individual applying for benefit is accepted into the disability program. In two studies reviewed by Parsons (1991), reconsideration of initial determinations of DI allowance by either the same review panel one year later or a separate team of medical experts revealed substantial type I and type II errors. Parsons (1996) details that an internal study of the Social Security Administration found that approximately 20 percent of the initial allowances of DI were denied by a blind review whereas 20 percent of the initial denials were allowed.

In the case of disability programs, social workers, i.e. low-tier bureaucrats or caseworkers, are in charge of part of the screening of individuals into the program according to the standards given by the government. The success of a disability program depends on the social

---

<sup>1</sup>In the US, the Social Security Administration defines disability under Social Security based on the inability to work. It considers an individual disabled if (i) (s)he cannot do work that (s)he did before; (ii) (s)he decides that (s)he cannot adjust to other work because of her/his medical condition(s); and (iii) her/his disability has lasted or is expected to last for at least one year or to result in death (source: web page of the Social Security Administration accessed February 9, 2011).

<sup>2</sup>As documented by Berkel and Börsch-Supan (2004) for Germany, the majority of workers exiting the labor force before the official retirement age do it because of disability.

workers that are in charge of its implementation. The better the screening of claimants, the less likely perverse incentives on the labor supply and horizontal inequity will arise. From their interactions with the claimants, the social workers are better informed about the disability of the individuals than the head of the government, i.e. the designer of the programs. When the designer of the programs and social workers have conflicting interests, the asymmetric information generated within the hierarchy opens the possibility of distortions in the implementation of the programs. There is abundant evidence that substantial part of social workers working in street-level bureaucracies are concerned about the well-being of the claimants and provide services to clients with no concerns for cost control (see, e.g., Prendergast, 2007; Buurman and Dur, 2008). The social workers have some discretion in their activities. The level of effort they exert in order to screen an individual will affect the likelihood of this person being effectively tagged. In this paper, we assume that social workers have their own preferences with a bias in favor of the well-being of claimants. When they exert a costly effort, they can become advocate of the claimants by collecting more evidence such that claimants are more likely to be tagged into the disability program. We then study the implications of altruistic social workers on the implementation of welfare programs.

We consider the framework of Diamond and Sheshinski (1995) where there are disability and welfare benefits. Labor supply is endogenous and affected by the levels of both disability and welfare benefits. All individuals that are not working are eligible for welfare benefits. The designer of welfare programs has access to a tagging technology that screens the claimants into the different programs. We extend the paper by Diamond and Sheshinski (1995) as we consider that the tagging technology is affected by the costly effort made by social workers. We assume that the effort made by social workers is driven by their altruism toward claimants. An altruistic social worker derives an additional utility when the claimant gets a higher utility level. When the consumption of the claimant differ whether he supplies labor or is in a welfare program, the social worker tries to influence the tagging process in order to give the highest utility level to the claimant. Therefore, the tagging technology is not only linked to the level of disability of the claimant but is also dependant on the consumption levels in the different programs. In our model, we look at the role of social workers on the implementation of welfare programs. We do not deal with optimal contract to align the preferences of the social workers with the one of the principal as we consider fixed wage contract. Yet, the government can control the effort made by the social workers through the effective consumptions of the individuals in the different programs and the disability standard.

Our first result shows that the altruism of the social workers cannot lead to the imple-

mentation of first-best allocations. The intuition behind this result is the following: first-best allocations are such that all individuals get the same consumption level. A tagging technology implementing the first-best allocations would require to tag exclusively the individuals that pass a disability threshold defined by the government, i.e. tagging would occur with probability 1 for disability levels above the optimal threshold and zero otherwise. When the level of altruism is not extremely high, it is never the case that the tagging occur with probability 1 as the effort of social workers is too costly to achieve such level of tagging. When the altruism is so high that the social workers' disutility of effort does not matter anymore, the social workers would exert enough effort in order to tag perfectly all claimants. Nevertheless, the social workers do not respect a minimal threshold of disability before tagging the claimants. As the choice to supply labor or to claim disability benefits is endogenous, the only way to avoid that all individuals claim and get tagged into the program requires to leave a rent to workers. This implies that the government has to give different consumption levels to workers and non-workers. In this case, we obtain the allocations offered in a second-best world without tagging. A very high altruism toward claimants makes the tagging technology no longer welfare enhancing and results in the implementation of the second-best allocations without tagging technology. Therefore, the government will never be able to fully align the preferences of the social workers with his objective.

We then present the optimal second-best allocation of disability and welfare benefits conditional on the level of altruism of social workers. In particular, we show that the presence of altruism has two effects on social welfare:

- (i) it increases welfare as it reduces the error that an individual gets the welfare benefits instead of being tagged into the disability program (decreases error of type I);
- (ii) it decreases welfare as it increases the error that an individual that should work is tagged into the disability program and does not work (increases error of type II).

It is inevitable that some errors of classification are made, then the final effect of altruism on social welfare depends on the weight given by the government to the two types of errors.

Finally, we extend the model to endogenize the choice of the disability standard. A disability standard is a level of disability at which the claimant should be tagged into the disability program. In practice, the disability standard represents the stringency of the requirements to enter the disability program. A reform aimed at raising the disability standard can be driven by the objectives to increase the incentives to work, to reduce the number of recipients, and the costs on public finance. In an environment without altruism and given a consumption

schedule, an increase of the disability standard implies a decrease of the errors of type II and an increase of the errors of type I. This would result in more workers and more non-workers in the general welfare program, reducing the size of the disability program (see Diamond and Sheshinski, 1995). This result is significantly modified when social workers are altruist. The key element is the implication of an increase of the standard on the discretionary power of the social workers in the tagging process. We show that when an increase of the disability standard does not reduce the discretionary powers of the social workers, there exists a level of altruism for which the social workers undo the initial government policy with higher effort. When the increase in disability standards reduces the discretion of the social workers, raising disability standard can lead to a decrease in the accuracy of the tagging process as the social workers' incentives to provide an effort decrease.

The questions relative to potential reforms and the implementation of disability insurance programs have been extensively studied in the literature (see, for a review, Autor and Duggan, 2006). For instance, recent papers have studied empirically and theoretically the implications of increasing of denial rate (see Gruber and Kubik, 1997), the extend to which monitoring extra effort in screening of reintegration reports reduce the long-term sickness absenteeism (see, De Jong, Lindeboom and van der Klaauw, 2011), the importance of complexity in public programs (see Kleven and Kopczuk, 2011), or the trends in employment and earnings of allowed and rejected applications to DI (see von Wachter, Song and Manchester, 2011). We contribute to this literature by studying the implications of the motivation bias of the social workers on the design of disability insurance programs.

The remainder is organized as follows. Section 2 specifies the model. Section 3 presents the benchmark cases. Section 4 studies the implications of social workers' altruism on welfare. Section 5 extends the model to allow the choice of disability standards. Section 6 shows simulations of the model. The last section contains concluding remarks.

## 2 The model

**Individuals.** There is a continuum of individuals of mass 1. We take work to be a 0–1 variable and do not model varying hours or work intensity.

The utility function of an individual who decides to work is  $U_a = u(c_a) - \theta$ , where  $c_a$  is the consumption of a private good and  $\theta$  denotes the disutility of working.  $u(\cdot)$  satisfies

$u'(\cdot) > 0$ ,  $u''(\cdot) < 0$ ,  $\lim_{c \rightarrow 0} u'(c) = \infty$ , and  $\lim_{c \rightarrow \infty} u'(c) = 0$ . We assume that  $\theta \in \mathbb{R}_+$  is distributed according to a cumulative distribution  $F(\theta)$  and a density function  $f(\theta)$  continuous and positive for all values of  $\theta$ . The disutility of working  $\theta$  is private information for the individuals with respect to the social planner. The productivity of all workers is normalized to 1.

We consider two welfare programs when the individuals are not working: (i) either they are tagged into the disability program and consume  $c_d$  or (ii) they are not tagged as disable and are in a general welfare program where they consume  $c_b$ . We assume that, when there is a disability program, disability benefits are at least as large as welfare benefits such that  $c_d \geq c_b$ . The utility of an individual who does not work is then either  $U_d = u(c_d)$  or  $U_b = u(c_b)$ .

Given the utility functions and the consumption levels for workers and non-workers, we can analyze which individuals would prefer to work instead of receiving welfare benefits or disability benefits (if eligible). We define  $\theta_b$  and  $\theta_d$  the threshold values which equate utilities of working and non-working such that

$$\theta_b = \max\{0, u(c_a) - u(c_b)\} \text{ and } \theta_d = \max\{0, u(c_a) - u(c_d)\}.$$

All individuals with disutilities below the thresholds will decide to work. Since disability benefits are greater than welfare benefits, the threshold for the individuals who apply for disability benefits is lower than the threshold for those who apply for welfare benefits. Thus, we have  $\theta_b \geq \theta_d$ .

**Social workers.** A number of identical, risk-neutral social workers are employed by the government. Their sole responsibility is to contribute to the screening of the welfare applicants. The caseload of each social worker is drawn randomly from the pool of welfare applicants. All variables of relevance are defined on a per client basis and do not depend on the size of a social worker's caseload. Therefore, we need not specify either the size of the caseload of each social worker or the number of social workers (see Boadway, Marceau and Sato, 1999).

The disutility of working  $\theta$  is private information for the individuals with respect to the social planner whereas it is perfectly observable for the social worker. The social worker does not directly tag the claimant into the program but can collect evidence in favor of tagging an individual. As we describe below the social worker will be able to affect the tagging technology when he is making an effort.

Social workers have their own preferences and we assume that they care about the utility of the claimants. The utility of a social worker increases when the claimant gets higher utility

such that his utility per applicant is represented by

$$U_{SW} = \underline{w} + \beta [p(e, \theta)u(c_d) + (1 - p(e, \theta)) \max\{u(c_a) - \theta, u(c_b)\}] - \varphi(e), \quad (1)$$

where  $\underline{w}$  represents a fixed wage,  $\beta \geq 0$  represents the altruism (intrinsic motivation) of the social worker,  $e$  is the effort chosen by the social worker and  $\varphi(e)$  is the disutility of exerting an effort with  $\varphi(0) = 0$ ,  $\varphi'(e) > 0$  when  $e \neq 0$  and  $\varphi''(e) > 0$ . The probability that the claimant is tagged into the disability program is denoted by  $p(e, \theta)$ .

As soon as the consumptions of the individuals differ whether he supplies labor or is in a welfare program, the social worker will try to influence the probability of tagging in order to give the highest utility level to the claimant.

Notice that the level of effort exerted when the social worker is not altruist is zero. We normalize  $\underline{w} = 0$  such that the fixed cost of employing social workers is null.

In our model the social worker cannot be induced to exert an effort by making wages contingent on their information as would be done in the literature on incentives in bureaucratic organizations (e.g., Tirole, 1992; Laffont and Tirole, 1993; Dixit, 2002; Besley and Ghatak, 2005). In particular Boadway, Marceau and Sato (1999) study the case where the government has to provide incentive for the social workers to make more effort such that tagging errors are reduced. We take a different approach and look at the case where social workers have intrinsic motivation associated with altruism toward claimants. The bias due to altruism cannot be controlled by monetary transfers but by the levels of benefits and standards for entering welfare programs. Low-tier bureaucrats are usually working in a low-power incentives environment. Recent papers have argued that this incentive structure might be optimal with altruistic social workers as the use of external incentives reduces or eliminates the effects of intrinsic motivations.<sup>3</sup> Moreover, if the output of a social worker cannot be measured, a natural response will be to hire agents who are highly motivated to get benefits to clients (see Prendergast, 2008). Therefore, pay-for-performance can have important effects on the self-selection of agents into street-level bureaucracies. In particular, flat wages are predicted to be attractive mainly to highly altruistic, i.e. client oriented, agents whereas pay-for-performance would attract agents with neutral stance towards clients (see Buurman and Dur, 2008). Social welfare departments attempt to identify the extent to which social workers identify with their clients, and there is evidence that such altruism affects performance (see Prendergast 2007 for

---

<sup>3</sup>See, e.g., Bowles and Polanía-Reyes, 2012; Francois and Vlassopoulos, 2008; Frey and Jegen, 2001; Perry, Hondeghem and Wise, 2010. See, however, Naegelen and Mougeot (2011) who show that high-powered contract can be reintroduced even with self-motivated agents.

details).<sup>4</sup>

**The tagging technology.** The social workers have some discretion in their activities. When they provide an effort, they can become advocate of the claimants by collecting more evidence such that claimants are more likely to be tagged. Following this observation, we assume that the level of effort exerted in order to tag an individual into a particular program affects the likelihood of this person being effectively tagged. Therefore, the tagging technology depends on the effort of the social workers that is itself a function of the level of benefits.

When the social worker is not exerting effort, i.e.  $e = 0$ , the claimant is tagged into the disability with probability  $p(0, \theta) = \underline{p}(\theta)$ . We assume that  $\underline{p}'(\theta) > 0$ ,  $\underline{p}(0) = 0$ , and  $\underline{p}''(\theta) \leq 0$ . The probability of being tag as disable when no effort is made  $\underline{p}(\theta)$  is similar to the tagging technology presented in Diamond and Sheshinski (1995).

When the social worker exerts a positive effort the probability of being tagged becomes  $p(e, \theta) > p(0, \theta) = \underline{p}(\theta)$ . We assume that  $\frac{\partial p(\cdot)}{\partial e} > 0$ ,  $\frac{\partial^2 p(\cdot)}{\partial e^2} \leq 0$ ,  $\frac{\partial p(\cdot)}{\partial \theta} \geq 0$ ,  $\frac{\partial^2 p(\cdot)}{\partial \theta^2} < 0$  and  $\frac{\partial^2 p(\cdot)}{\partial e \partial \theta} < 0$ . This last assumption specifies that the disability level and the effort are substitute in the tagging technology. When the claimant has a high disability type, the effort of the social worker will play a marginal role in the tagging of the individual.

For instance, the following tagging technology would satisfy our assumptions

$$p(e, \theta) = 1 - (1 - \underline{p}(\theta))(1 - e), \quad (2)$$

with  $\underline{p}(\theta) = 1 - \exp\{-\gamma\theta\}$  and  $\gamma > 0$ .

**Optimal level of effort.** The optimal level of effort chosen by a social worker  $e^* := e^*(\theta, c_a, c_d, c_b) = \operatorname{argmax}_e U_{SW}$  is given implicitly by the first order condition

$$\beta \frac{\partial p(\cdot)}{\partial e^*} [u(c_d) - \max\{u(c_a) - \theta, u(c_b)\}] - \varphi'(e^*) = 0. \quad (3)$$

Individuals with type between 0 and  $\theta_d$  do not apply to disability benefits and the social workers do not exert any effort for these types.

When  $\theta \in [\theta_d, \theta_b]$ ,  $\max\{u(c_a) - \theta, u(c_b)\} = u(c_a) - \theta$ . In this case, we define the optimal level of effort  $e_1 := e_1(c_d, c_a, \theta, \beta)$  by  $\beta \frac{\partial p(\cdot)}{\partial e_1} [u(c_d) - u(c_a) + \theta] - \varphi'(e_1) = 0$ .

---

<sup>4</sup>In the case of job placement, Behncke, Frölich and Lechner (2010) find that when the unemployed and caseworkers belong to the same social group, the chances of job placement increases by 3 percentage points.

For any  $\theta$  such that  $\theta > \theta_b$ ,  $\max\{u(c_a) - \theta, u(c_b)\} = u(c_b)$  and the optimal level of effort  $e_2 := e_2(c_d, c_b, \theta, \beta)$  is given implicitly by  $\beta \frac{\partial p(\cdot)}{\partial e_2} [u(c_d) - u(c_b)] - \varphi'(e_2) = 0$ .

The optimal effort function is then

$$e^*(c_a, c_b, c_d, \theta, \beta) = \begin{cases} 0 & \text{if } \theta < \theta_d \\ e_1(c_d, c_a, \theta, \beta) & \text{if } \theta \in [\theta_d, \theta_b] \\ e_2(c_d, c_b, \theta, \beta) & \text{if } \theta > \theta_b \end{cases}$$

The effort function is increasing in  $c_d$  whereas it is decreasing in  $c_a$  and  $c_b$ . By continuity we have

- (i)  $e_1(c_d, c_a, \theta_b, \beta) = e_2(c_d, c_b, \theta_b, \beta)$  and then  $p(e_1(c_d, c_a, \theta_b, \beta), \theta_b) = p(e_2(c_d, c_b, \theta_b, \beta), \theta_b)$ ;
- (ii)  $e_1(c_d, c_a, \theta_d, \beta) = 0$  and then  $p(e_1(c_d, c_a, \theta_d, \beta), \theta_d) = \underline{p}(\theta_d)$ .

The effort is non monotone in the type  $\theta$  of the individuals. The derivative of the effort function is strictly positive in  $\theta = \theta_d$  ( $\frac{\partial e_1}{\partial \theta} |_{\theta=\theta_d} > 0$  as  $\frac{\partial p(\cdot)}{\partial e} + \frac{\partial^2 p(\cdot)}{\partial e \partial \theta} [u(c_d) - u(c_a) + \theta] > 0$  for  $\theta = \theta_d$  since  $\frac{\partial p(\cdot)}{\partial e} > 0$ ) and can turn negative for other  $\theta \in ]\theta_d, \theta_b]$ .<sup>5</sup> The effort function is strictly decreasing in  $\theta$  when  $\theta > \theta_b$ .

The effort made by altruistic social workers has an effect on the two types of errors:

- (i) it reduces the number of non-workers that gets the welfare benefits instead of being tagged into the disability program (i.e. decreases error of type I);
- (ii) it increases the number of workers that shifts into the disability program (i.e. increases error of type II).

The following Proposition presents the relationship between the degree of altruism, the level of effort made by the social worker and the associated tagging technology.

**Proposition 1** *Given a consumption schedule  $c_a > c_d > c_b$ , the level of effort is nondecreasing in the altruism of the social worker and there exists a level of altruism  $\bar{\beta}$  such that*

- (i) *when  $\beta > \bar{\beta}$ , the minimal associated level of effort  $\bar{e}$  for all  $\theta$  implies that  $p(\bar{e}, \theta) = 1$ ;*
- (ii) *when  $0 < \beta \leq \bar{\beta}$ , we have  $p(e, \theta) < 1$  for all  $\theta$  and there exists  $\bar{\theta}$  such that for all  $\theta \geq \bar{\theta}$ ,  $e(\theta) = 0$  and then  $p(0, \theta) = \underline{p}(\theta)$ .*

---

<sup>5</sup>In the specification of equation (2) for  $\theta \in ]\theta_d, \theta_b]$ ,  $\frac{\partial e_1}{\partial \theta} > 0$  if and only if  $\frac{1}{\gamma} > u(c_d) - u(c_a) + \theta$ .

The altruism reduces the cost of exerting an effort, therefore the effort function is nondecreasing in the level of altruism.

Proposition 1 (i) states that when the social workers are very altruistic all the claimants are going to be tagged in the disability program for any level of disability  $\theta$ . In this case the cost of providing an effort for the social worker is so small that he is going to exert enough effort to tag all claimants in order to insure that they have access to the disability benefits.

Proposition 1 (ii) comes from the substitutability between effort and the level of disability. For a sufficiently high disability level, the increase in the probability of being tagged through an effort does not compensate the cost of effort for social workers. Intuitively, the claimant disability level is high enough to ensure a high probability of the application to succeed without the help of the social worker. An example of the tagging technology in the case of Proposition 1 (ii) is depicted in Figure 1.

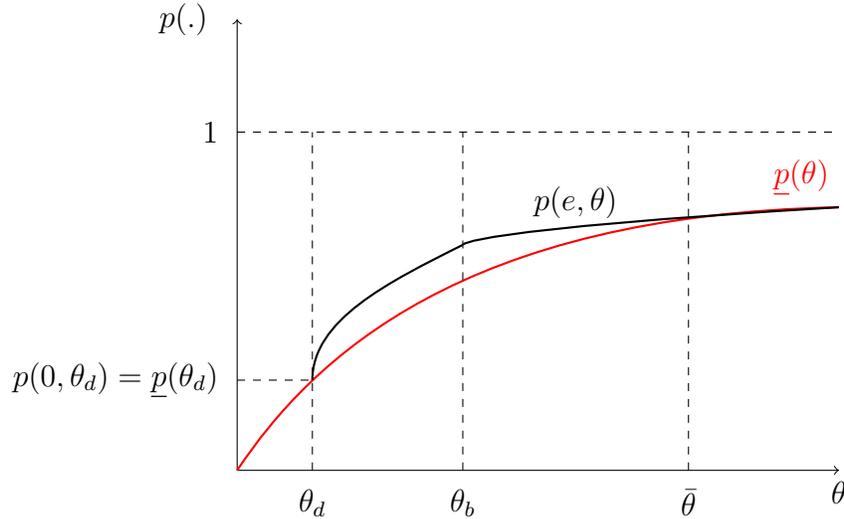


Figure 1: Example of the tagging technology with  $0 < \beta \leq \bar{\beta}$ .

**Social planner.** A benevolent social planner designs the welfare programs. The social planner has a utilitarian objective such that he maximizes

$$\begin{aligned} & \max_{c_a, c_d, c_b} \int_0^{\theta_d(c_d, c_a)} [u(c_a) - \theta] f(\theta) d\theta \\ & + \int_{\theta_d(c_d, c_a)}^{\theta_b(c_b, c_a)} [p(e_1(\theta, c_d, c_a), \theta) u(c_d) + (1 - p(e_1(\theta, c_d, c_a), \theta)) [u(c_a) - \theta]] f(\theta) d\theta \end{aligned}$$

$$+ \int_{\theta_b(c_b, c_a)}^{\infty} [p(e_2(\theta, c_d, c_b), \theta)u(c_d) + (1 - p(e_2(\theta, c_d, c_b), \theta))u(c_b)]f(\theta)d\theta \quad (4)$$

subject to the budget constraint

$$\begin{aligned} & [1 - c_a] \left[ \int_0^{\theta_d(c_d, c_a)} f(\theta)d\theta + \int_{\theta_d(c_d, c_a)}^{\theta_b(c_b, c_a)} [1 - p(e_1(\theta, c_d, c_a), \theta)]f(\theta)d\theta \right] \\ & - c_d \left[ \int_{\theta_d(c_d, c_a)}^{\theta_b(c_b, c_a)} p(e_1(\theta, c_d, c_a), \theta)f(\theta)d\theta + \int_{\theta_b(c_b, c_a)}^{\infty} p(e_2(\theta, c_d, c_b), \theta)f(\theta)d\theta \right] \\ & - c_b \left[ \int_{\theta_b(c_b, c_a)}^{\infty} (1 - p(e_2(\theta, c_d, c_b), \theta))f(\theta)d\theta \right] \geq 0 \quad (BC). \end{aligned}$$

In the following, we denote  $\lambda$  the lagrange multiplier associate to (BC). The maximization problem presented above also takes into account the implications of the choice of the consumption levels on labor supply decisions. In particular, as  $u(c_a) - \theta_d = u(c_d)$ , we have  $\theta_d = \theta_d(c_a, c_d)$  with  $\frac{\partial \theta_d}{\partial c_a} = u'(c_a)$  and  $\frac{\partial \theta_d}{\partial c_d} = -u'(c_d)$ . Similarly, we have  $u(c_a) - \theta_b = u(c_b)$  then  $\theta_b = \theta_b(c_b, c_a)$  with  $\frac{\partial \theta_b}{\partial c_a} = u'(c_a)$  and  $\frac{\partial \theta_b}{\partial c_b} = -u'(c_b)$ .

In the next sections, we determine how the social planner reacts to the tagging technology in term of the generosity of welfare programs and the implication of altruism on welfare. Finally, we look at the optimal choice of disability standard.

## 3 Benchmark cases

### 3.1 First-best environment

We consider the first-best environment in which the social planner knows the type of the individuals without the help of the social workers. In such environment  $\theta_b$  and  $\theta_d$  would then be control variables of the maximization problem. In a first-best world there are no reasons to have two welfare programs and the optimal allocations are such that the marginal utilities of all types are equalized:  $u'(c_a^*) = u'(c_d^*) = u(c_b^*)$  then  $c^* := c_a^* = c_d^* = c_b^*$ .

All individuals with disutility below the cutoff  $\theta^*$  should work. The cutoff is determined by comparing the utility gain from extra work,  $u(c_a^*) - \theta^* - u(c_d^*)$ , with the value of extra net consumption as a consequence of work, which is the sum of marginal product and the change in consumption which results from the change in status:  $u(c_a^*) - \theta^* - u(c_d^*) = -u'(c_a^*)(1 - c_a^* + c_d^*)$ .

The level of the tagging technology that would replicate the first-best allocation would be

such that

$$p^* = \begin{cases} 0 & \text{if } \theta \in [0, \theta^*] \\ 1 & \text{if } \theta > \theta^* \end{cases}$$

An example of this tagging technology is depicted in Figure 2.<sup>6</sup>

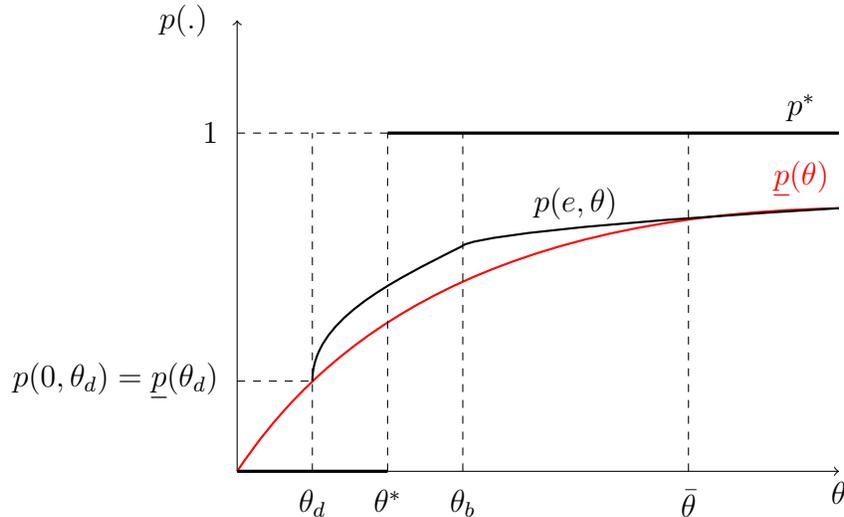


Figure 2: Example of second-best and first-best tagging technologies.

### 3.2 Second-best without tagging technology

The first-best allocations cannot be implemented in a second-best world as no individuals would work when  $c_a = c_d = c_b$ . Without a tagging technology, there cannot be two welfare programs as the planner does not have the possibility to differentiate among the individuals that are applying to the welfare benefits. This implies that  $c_d = c_b$  and then  $\theta_d = \theta_b$ . The decision to enter the welfare program is driven only by the incentives to work and the threshold determining the disability of the individuals that are not working is given by  $\theta_b = u(c_a) - u(c_b)$ . Therefore, at the second-best level of consumptions the planner has to leave a rent to the workers  $c_a^{SB} > c_b^{SB}$ . The maximization problem of the social planner is such that

$$\max_{c_a, c_b} \int_0^{\theta_b(c_b, c_a)} [u(c_a) - \theta] f(\theta) d\theta + \int_{\theta_b(c_b, c_a)}^{\infty} u(c_b) f(\theta) d\theta$$

subject to  $[1 - c_a]F(\theta_b(c_b, c_a)) - c_b[1 - F(\theta_b(c_b, c_a))] \geq 0$ .

<sup>6</sup>In Figure 2, the ranking of  $\theta^*$  with respect to  $\theta_d$  and  $\theta_b$  could differ depending on the parameters of the model.

The second-best values of  $c_a^{SB}$  and  $c_b^{SB}$  are defined, respectively, by the first order conditions

$$[u'(c_a) - \lambda]F(\theta_b) + \lambda(1 - c_a + c_b)f(\theta_b)u'(c_a) = 0,$$

$$[u'(c_b) - \lambda](1 - F(\theta_b)) - \lambda(1 - c_a + c_b)f(\theta_b)u'(c_b) = 0,$$

and the budget constraint:  $[1 - c_a]F(\theta_b) - c_b[1 - F(\theta_b)] = 0$ .

### 3.3 Non-altruistic social workers

We present the case where social workers are not altruist, i.e.  $\beta = 0$ . In this case the social workers are not exerting any effort and  $p(0, \theta) = \underline{p}(\theta)$ , for all  $\theta$ . This is the case studied in Diamond and Sheshinski (1995). We discuss the first order conditions under the assumptions that  $\theta_b > \theta_d > 0$  and  $c_d > c_b$ .

The solution of Problem 4 subject to the budget constraint (*BC*) is

$$[u'(c_a) - \lambda]F(\theta_d) + [u'(c_a) - \lambda] \int_{\theta_d}^{\theta_b} (1 - \underline{p}(\theta))f(\theta)d\theta$$

$$+ \lambda(1 - c_a + c_d)u'(c_a)\underline{p}(\theta_d)f(\theta_d) + \lambda(1 - c_a + c_b)u'(c_a)(1 - \underline{p}(\theta_b))f(\theta_b) = 0,$$

$$[u'(c_d) - \lambda] \int_{\theta_d}^{\infty} \underline{p}(\theta)f(\theta)d\theta - \lambda(1 - c_a + c_d)u'(c_d)\underline{p}(\theta_d)f(\theta_d) = 0,$$

$$[u'(c_b) - \lambda] \int_{\theta_b}^{\infty} (1 - \underline{p}(\theta))f(\theta)d\theta - \lambda(1 - c_a + c_b)u'(c_b)(1 - \underline{p}(\theta_b))f(\theta_b) = 0,$$

and the budget constraint.

At the optimum, we have that  $\theta_d > 0$ , then  $\theta_b > 0$  and  $c_d > c_b$ . Indeed, we have that  $\underline{p}(\theta)$  is increasing in  $\theta$  and the marginal utility of non-workers exceeds that of workers at the optimum without a disability program. Intuitively, when the tagging technology screens well the truly needy, the presence of a disability program increases the social welfare.

For the levels of consumptions  $c_a$ ,  $c_d$  and  $c_b$  defined above, an altruistic social worker will exert some effort as  $\frac{\partial p(0, \theta)}{\partial e} > 0$  for all  $\theta \geq \theta_d$ . This implies that, given  $\theta \geq \theta_d$ , the probability of being tagged  $p(\cdot)$  will increase such that (i) there will be more non-workers into the disability program, and (ii) there are less workers than what would be optimal. In this case the consumption levels defined above are no longer optimal.

## 4 Altruistic social workers

### 4.1 Altruism and first-best decentralization

We first look at the possibility to decentralize the first-best allocations by the use of the interaction between the altruism of social workers and the tagging technology.

**Proposition 2** *For any level of altruism  $\beta$ , the first-best allocations can never be implemented.*

A sketch of the proof of Proposition 2 suffices. The first-best allocations can only be supported by a perfect tagging technology  $p^*$  defined in Section 3.1. Therefore, from Proposition 1 (i), as soon as  $\beta \leq \bar{\beta}$ , we cannot have the first-best level of tagging as  $p(e, \theta) < 1$  for all  $\theta$ . Suppose now that  $\beta > \bar{\beta}$ . For this level of altruism the tagging technology is such that  $p(\bar{e}, \theta) = 1$ , for all  $\theta \geq \theta_d$ . Nevertheless the first-best allocations equalize the consumption levels of workers and non-workers. This would imply that  $\theta_d = 0$ , therefore no individual will work and they would all be tagged as disable.

This first result shows that the altruism of the social workers cannot be directly used to implement first-best allocations. This statement is true even when the altruism is so high that the social workers' disutility of effort does not matter anymore. In this case the only way to avoid the tagging of all individuals requires to leave a rent to workers such that  $c_a > c_d$ . This is exactly what we would have in a second-best world without tagging as presented in Section 3.2. A very high altruism toward claimants makes the tagging technology no longer welfare enhancing and results in the implementation of the second-best allocations without tagging technology.

### 4.2 Altruism and optimal second-best welfare programs

We now analyze the case where  $\beta \in ]0, \bar{\beta}]$ . The solution of Problem 4 subject to the budget constraint ( $BC$ ) is

$$\begin{aligned}
 & [u'(c_a) - \lambda]F(\theta_d) + [u'(c_a) - \lambda] \int_{\theta_d}^{\theta_b} (1 - p(e_1, \theta))f(\theta)d\theta \\
 & + \lambda \left[ (1 - c_a + c_b)u'(c_a)(1 - p(e_2, \theta_b))f(\theta_b) + (1 - c_a + c_d)u'(c_a) \underbrace{p(e_1, \theta_d)}_{p(\theta_d)} f(\theta_d) \right]
 \end{aligned}$$

$$+ \int_{\theta_d}^{\theta_b} \frac{\partial p}{\partial e} \frac{\partial e_1}{\partial c_a} [u(c_d) - u(c_a) + \theta - \lambda(1 - c_a + c_d)] f(\theta) d\theta = 0, \quad (5)$$

$$[u'(c_d) - \lambda] \left( \int_{\theta_d}^{\theta_b} p(e_1, \theta) f(\theta) d\theta + \int_{\theta_b}^{\infty} p(e_2, \theta) f(\theta) d\theta \right) - \lambda(1 - c_a + c_d) u'(c_d) \underbrace{p(e_1, \theta_d)}_{p(\theta_d)} f(\theta_d) \\ + \int_{\theta_d}^{\theta_b} \frac{\partial p}{\partial e} \frac{\partial e_1}{\partial c_d} [\theta - \theta_d - \lambda(1 - c_a + c_d)] f(\theta) d\theta + \int_{\theta_b}^{\infty} \frac{\partial p}{\partial e} \frac{\partial e_2}{\partial c_d} [u(c_d) - u(c_b) - \lambda(c_d - c_b)] f(\theta) d\theta = 0, \quad (6)$$

$$[u'(c_b) - \lambda] \int_{\theta_b}^{\infty} (1 - p(e_2, \theta)) f(\theta) d\theta - \lambda(1 - c_a + c_b) u'(c_b) (1 - p(e_2, \theta_b)) f(\theta_b) \\ + \int_{\theta_b}^{\infty} \frac{\partial p}{\partial e} \frac{\partial e_2}{\partial c_b} [u(c_d) - u(c_b) - \lambda(c_d - c_b)] f(\theta) d\theta = 0, \quad (7)$$

$$[1 - c_a] F(\theta_d) + \int_{\theta_d}^{\theta_b} [1 - p(e_1, \theta)] f(\theta) d\theta - c_d \left[ \int_{\theta_d}^{\theta_b} p(e_1, \theta) f(\theta) d\theta + \int_{\theta_b}^{\infty} p(e_2, \theta) f(\theta) d\theta \right] \\ - c_b \left[ \int_{\theta_b}^{\infty} (1 - p(e_2, \theta)) f(\theta) d\theta \right] = 0. \quad (8)$$

As we show in the Appendix, the introduction of altruism does not change the desirability of a disability program in addition to the welfare program. As in Diamond and Sheshinski (1995), the fact that  $p(\cdot)$  increases with  $\theta$  is sufficient to justify the existence of the disability program, i.e.  $c_d > c_b$  and  $\theta_d > 0$ .

It is interesting to compare the first order conditions when the social workers are altruist with the one without altruism derived in Diamond and Sheshinski (1995). The first point to notice is that the form of the first order conditions are very similar. The main difference are (i) the probabilities that individuals are tagged  $p(e_1, \theta)$  and  $p(e_2, \theta)$  and (ii) the introduction of new terms associated with the reaction of the social workers to a change in consumption in the different groups: these terms are the last lines of equations (5), (6) and (7). These terms capture the direct reaction of social workers to a change in the consumption profiles.

We have that  $\Delta_b := u(c_d) - u(c_b) - \lambda(c_d - c_b) > 0$  from the concavity of  $u(\cdot)$  and  $u'(c_d) \geq \lambda$  such that  $u(c_d) - u(c_b) > u'(c_d)(c_d - c_b) \geq \lambda(c_d - c_b)$ . Similarly, we have  $\Delta_a := u(c_d) - u(c_a) - \lambda(1 - c_a + c_d) < 0$  and  $u(c_d) - u(c_a) + \theta - \lambda(1 - c_a + c_d) < 0$  for at least some of the individuals with type  $\theta \in [\theta_d, \theta_b]$ .

From the social planner perspective, the introduction of altruism creates a tension between the redistributive objective and the budget constraint. This tension is clear in the equation (6)

defining  $c_d$ : on one side altruism helps the planner to screen the non-workers in the disability program and this reduces the number of deserving disabled receiving general welfare benefits; on the other side altruism increases the risk that workers enter in the disability program and this tightens the budget constraint.

The total effect of altruism on the social welfare is studied in the following section.

### 4.3 Comparative statics: altruism and social welfare

We denote by  $\mathcal{L}$  the value function at the optimal effort and consumptions levels  $c_a$ ,  $c_b$  and  $c_d$  defined by (5), (7) and (6). Using the envelope theorem, we have that

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \beta} &= \int_{\theta_d}^{\theta_b} \frac{\partial p}{\partial e_1} \frac{\partial e_1}{\partial \beta} [u(c_d) - u(c_a) + \theta - \lambda(1 - c_a + c_d)] f(\theta) d\theta \\ &+ \int_{\theta_b}^{\infty} \frac{\partial p}{\partial e_2} \frac{\partial e_2}{\partial \beta} [u(c_d) - u(c_b) - \lambda(c_d - c_b)] f(\theta) d\theta, \end{aligned} \quad (9)$$

or equivalently,

$$\frac{\partial \mathcal{L}}{\partial \beta} = \Delta_a \int_{\theta_d}^{\theta_b} \frac{\partial p}{\partial e_1} \frac{\partial e_1}{\partial \beta} dF(\theta) + \int_{\theta_d}^{\theta_b} \frac{\partial p}{\partial e_1} \frac{\partial e_1}{\partial \beta} \theta dF(\theta) + \Delta_b \int_{\theta_b}^{\infty} \frac{\partial p}{\partial e_2} \frac{\partial e_2}{\partial \beta} dF(\theta). \quad (10)$$

The effort made by the social workers implies that some people are moving out of the labor force to the disability program. This shift has a net social value of  $\Delta_a + \theta := u(c_d) - u(c_a) + \theta - \lambda(1 - c_a + c_d)$ . At the same time some people are moving from the welfare benefits to the disability program. This has a social value of  $\Delta_b := u(c_d) - u(c_b) - \lambda(c_d - c_b)$ .

Thus, at the optimum, with an increase of the altruism of the social workers, social welfare goes up from additional non-workers added to the disability rolls, i.e.  $\int_{\theta_b}^{\infty} \frac{\partial p}{\partial e} \frac{\partial e_2}{\partial \beta} [u(c_d) - u(c_b) - \lambda(c_d - c_b)] f(\theta) d\theta > 0$ , while it goes down from at least some of the workers switching to disability benefits, i.e. for some  $\theta$ ,  $\int_{\theta_d}^{\theta_b} \frac{\partial p}{\partial e} \frac{\partial e_1}{\partial \beta} [\theta - \theta_d - \lambda(1 - c_a + c_d)] f(\theta) d\theta < 0$ .

It is inevitable that some errors of classification are made, therefore the total effect is going to depend on the weight given by the social planner to the two types of errors.

In the following section, we enrich our model by endogenizing the choice of the disability standard  $\theta^{**}$  and we look at the effect of altruism on this choice.

## 5 Disability standard

We extend our model to include the possibility that the government is fixing the disability standard. A disability standard is a level of disability  $\theta^{**}$  at which the claimant should be tagged into the disability program. In practice, the disability standard represents the stringency of the requirements to enter into the disability program. Everything else equal, as the disability standard increases the number of eligible recipients and effective recipients will decrease. A reform aimed at raising the disability standard can be driven by the objectives to increase the incentives to work, to reduce the number of recipients, and the costs on public finance.

In order to introduce a disability standard, we follow Diamond and Sheshinski (1995) and consider the bivariate random vector  $(\tilde{\theta}, \theta)$  that has a joint distribution  $P$  which admits a joint p.d.f.  $g_{\tilde{\theta}\theta}$  with respect to the Lebesgue measure. We assume that  $g_{\tilde{\theta}\theta}$  is strictly positive on its support. We denote by  $g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta)$  the corresponding conditional p.d.f. of the observed disutility,  $\tilde{\theta}$ , conditional on the true disability being  $\theta$ . The conditional c.d.f.  $G(\theta^{**}; \theta)$  associated with  $P$  is defined as  $G(\theta^{**}; \theta) := P(\tilde{\theta} \leq \theta^{**} | \theta) = \int_0^{\theta^{**}} g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) d\tilde{\theta}$ .

We keep the assumptions made previously about the tagging technology. Then, without effort  $e = 0$ , for the disability standard  $\theta^{**}$ , the probability that an individual with type  $\theta$  is tagged into the disability program is  $p(\theta^{**}, \theta) = 1 - G(\theta^{**}; \theta)$ . When the social worker exerts a positive effort the probability of being tagged becomes  $p(e, \theta^{**}, \theta)$ . We also assume the following monotone likelihood ratio property: for all  $\theta_1 < \theta_2$ ,  $\frac{\partial}{\partial \theta^{**}} \left[ \frac{g_{\tilde{\theta}|\theta_2}(\theta^{**}|\theta_2)}{g_{\tilde{\theta}|\theta_1}(\theta^{**}|\theta_1)} \right] \geq 0$ . This property ensures that  $p(e, \theta^{**}, \theta_1) < p(e, \theta^{**}, \theta_2)$  for all  $\theta_1 < \theta_2$ .

Without altruistic social workers and given a consumption schedule, an increase of the disability standard implies a decrease of the errors of type II and an increase of the errors of type I (see Diamond and Sheshinski, 1995).

The first order condition of Problem 4 with respect to  $\theta^{**} > 0$  is such that

$$\begin{aligned} & \int_{\theta_d}^{\theta_b} \left[ \frac{\partial p}{\partial e_1} \frac{\partial e_1}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} \right] [u(c_d) - u(c_a) + \theta - \lambda(1 - c_a + c_d)] f(\theta) d\theta \\ & + \int_{\theta_b}^{\infty} \left[ \frac{\partial p}{\partial e_2} \frac{\partial e_2}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} \right] [u(c_d) - u(c_b) - \lambda(c_d - c_b)] f(\theta) d\theta = 0, \end{aligned} \quad (11)$$

or equivalently,

$$\Delta_a \int_{\theta_d}^{\theta_b} \left[ \frac{\partial p}{\partial e_1} \frac{\partial e_1}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} \right] dF(\theta) + \int_{\theta_d}^{\theta_b} \left[ \frac{\partial p}{\partial e_1} \frac{\partial e_1}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} \right] \theta dF(\theta)$$

$$+\Delta_b \int_{\theta_b}^{\infty} \left[ \frac{\partial p}{\partial e_2} \frac{\partial e_2}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} \right] dF(\theta) = 0. \quad (12)$$

Without altruism, as in Diamond and Sheshinski (1995),  $\frac{\partial p}{\partial e} \frac{\partial e}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} = \frac{\partial p}{\partial \theta^{**}} = -g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta)$  and equation (12) would become

$$\Delta_a \int_{\theta_a}^{\theta_b} g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) dF(\theta) + \int_{\theta_a}^{\theta_b} \theta g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) dF(\theta) + \Delta_b \int_{\theta_b}^{\infty} g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) dF(\theta) = 0. \quad (13)$$

At the optimum, as in equation (9), an increase of the standard would balance the positive welfare effect of having more workers switching from the disability roll and the negative welfare effect from non-workers in the general welfare program instead of in the disability roll. The introduction of altruism may affect significantly this balance.

In order to characterize the implications of the altruism on the choice of the disability standard, we specify the possible reactions of the social workers to the disability standard and the implications for the tagging technology.

**Crowding out between effort and disability standard.** This case corresponds to the situation where the effort of the social worker and the disability standard are substitute in the tagging technology. There is a crowding out effect of disability standard on the incentives to exert effort. This case is described by  $\frac{\partial^2 p}{\partial e \partial \theta^{**}} < 0$ . In practice, when the disability standard increases and the discretionary power of the social worker becomes lower, then the incentives to exert a costly effort decreases. This implies that  $\frac{\partial e}{\partial \theta^{**}} < 0$  and  $\frac{\partial p(\cdot)}{\partial \theta^{**}} < 0$ , for all  $\theta$ .

An increase of the disability standard has a direct effect on the probability of a claimant being tagged but also an indirect effect: social workers see their discretionary power reduced by the standard. The effort they are providing is becoming less and less relevant in the tagging outcome when the disability standard is increased. Therefore, a reduction of the discretionary power of an altruistic social workers lead to a decrease in their costly effort.

At the optimum, an increase of the disability standard implies a decrease of the errors of type II and an increase of the errors of type I. These effects are similar to the case without altruism. Notice, however, that the magnitude of the effects varies due to the reaction of the social workers to the standard. From equation (11), the optimal standard does not need to increase as much as in a situation where the social workers are not altruist in order to achieve an equivalent reduction of the disability rolls, everything else being equal.

**Crowding in between effort and disability standard.** This case corresponds to the situation where the effort of the social worker and the disability standard are complement in the tagging technology. The increase of the disability standard reinforces the value of the effort for the social workers. Formally, this case is described by  $\frac{\partial^2 p}{\partial e \partial \theta^{**}} > 0$ . When the disability standard increases, the incentives to exert a costly effort increases. This implies that  $\frac{\partial e}{\partial \theta^{**}} > 0$ , for all  $\theta$ . The total effect on the probability of being tagged is going to depend on the level of effort of altruistic social workers.

In practice, the discretionary power of the social workers is not affected by the standard. Given a consumption schedule, the increase of the standard goes in the opposite direction of the preferred outcome for the social workers: it depresses the chance that a claimant gets into the disability program. Therefore, altruistic social workers try to undo the initial policy of the government. The final implications of an increase of the standard depend on the effective effort of the social workers. We illustrate different cases with an example.

*Example.* We assume that the tagging technology is represented by equation (2) and that the social workers' disutility of effort is quadratic such that  $\varphi(e) = \psi \frac{e^2}{2}$ , with  $\psi > 0$ . Then, the optimal effort is given by:

$$(i) \text{ for } \theta \in [\theta_d, \theta_b], e_1 = \frac{\beta}{\psi}(1 - \underline{p}(\theta^{**}, \theta))[u(c_d) - u(c_a) + \theta];$$

$$(ii) \text{ for } \theta > \theta_b, e_2 = \frac{\beta}{\psi}(1 - \underline{p}(\theta^{**}, \theta))[u(c_d) - u(c_b)].$$

To satisfy the assumption that the tagging technology  $p(e, \theta^{**}, \theta)$  is increasing in  $\theta$  for all  $\theta$ , we assume that the altruism and the marginal cost of effort are such that the optimal level of effort  $e < \frac{1}{2}$ .

We can now compute the effect of the standard on the effort of the social workers. The effort and the standard are complement in the tagging technology  $\frac{\partial^2 p}{\partial e \partial \theta^{**}} = -\frac{\partial \underline{p}(\theta^{**}, \theta)}{\partial \theta^{**}} > 0$ , for all  $\theta$ . This implies that the effort function is strictly increasing in  $\theta^{**}$ .

Equation (11) becomes

$$\begin{aligned} \Delta_a \int_{\theta_d}^{\theta_b} [1 - 3e_1] g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) dF(\theta) + \int_{\theta_d}^{\theta_b} \theta [1 - 3e_1] g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) dF(\theta) \\ + \Delta_b \int_{\theta_b}^{\infty} [1 - 3e_2] g_{\tilde{\theta}|\theta}(\tilde{\theta}|\theta) dF(\theta) = 0, \end{aligned} \quad (14)$$

$$\text{as } \frac{\partial p}{\partial e} \frac{\partial e}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} = \frac{\partial p}{\partial \theta^{**}} [1 - 3e].$$

Therefore, two cases are possible:

- (i) if  $\frac{1}{3} > e$  for all  $\theta$ , then  $\frac{\partial p}{\partial e} \frac{\partial e}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} < 0$ . At the optimum, an increase of the disability standard implies a decrease of the errors of type II and an increase of the errors of type I. The effects are similar to the case without altruism. Notice, however, that the magnitude of the effects varies due to the reaction of the social workers to the standard. From equation (11), the optimal standard has to increase more than in a situation where the social workers are not altruist in order to achieve an equivalent reduction of the disability rolls, everything else being equal;
- (ii) if for some  $\theta$ ,  $\frac{1}{2} > e > \frac{1}{3}$  then  $\frac{\partial p}{\partial e} \frac{\partial e}{\partial \theta^{**}} + \frac{\partial p}{\partial \theta^{**}} > 0$ . As soon as the effort is sufficiently high, at the optimum, an increase of the disability standard imply an increase of the errors of type II and a decrease of the errors of type I. These effects are opposite to the case without altruism. The altruistic social workers react to the standard with an effort that is high enough to undo the policy of the planner. The final effect on the roll can be opposite of the initial objective of the planner.

## 6 Simulations

This section applies our theoretical model to a numerical example with a logarithmic utility function.

## 7 Concluding remarks

In this paper we study the implications of altruism on the design of welfare programs. The starting point of the analysis was the observation that social workers, responsible for the implementation of welfare programs, have a motivation bias in favor of the claimants they interact with.

In an environment with disability and general welfare programs, we find that the altruism of the social workers cannot be used to implement the first-best policy. We then characterize the second-best consumption schedule of the social planner given a level of altruism. We find that the total effect of altruism on social welfare is ambiguous as it reinforces the award errors (type II) whereas it reduces the rejection errors (type I). We illustrate this situation with numerical simulations.

Finally, we look at the effect of a reform aimed at raising disability standard when it is implemented by altruistic social workers. If the increases of disability standard is reducing

the discretionary power of the social workers, increasing disability standard is followed by a reduction of the effort made by social workers in their activity. With respect to the errors of type I and type II, the government has to be careful while calibrating the implications of the reform: the decrease of the effort accompanying the increase in the standard leads to an increase in type I and type II errors more than if the effort had remained constant. This implies that a reduction in the discretionary power of the social workers leads to a perverse effect in term of type I errors. On the contrary, when the social workers are sufficiently altruist and their discretionary power is not affected by the increases of the standard, social workers are going to exert more effort in order to tag applicants and to undo government initial policy. The final implications of an increase of the standard depend on the effective effort of the social workers.

## Appendix

**Value of the Lagrange multiplier when social workers are altruist.** If we sum the first order conditions, equations (5), (6), and (7) divided respectively by  $u'(c_a)$ ,  $u'(c_d)$ , and  $u'(c_b)$  we get

$$\begin{aligned} & \left[1 - \frac{\lambda}{u'(c_a)}\right]F(\theta_d) + \left[1 - \frac{\lambda}{u'(c_a)}\right] \int_{\theta_d}^{\theta_b} (1 - p(e_1, \theta))f(\theta)d\theta \\ & + \left[1 - \frac{\lambda}{u'(c_d)}\right] \left( \int_{\theta_d}^{\theta_b} p(e_1, \theta)f(\theta)d\theta + \int_{\theta_b}^{\infty} p(e_2, \theta)f(\theta)d\theta \right) \\ & + \left[1 - \frac{\lambda}{u'(c_b)}\right] \int_{\theta_b}^{\infty} (1 - p(e_2, \theta))f(\theta)d\theta = 0, \end{aligned}$$

then

$$\begin{aligned} & F(\theta_d)u'(c_a)^{-1} + \int_{\theta_d}^{\theta_b} \left( (1 - p(e_1, \theta))u'(c_a)^{-1} + p(e_1, \theta)u'(c_d)^{-1} \right) f(\theta)d\theta \\ & + \int_{\theta_b}^{\infty} \left( p(e_2, \theta)u'(c_d)^{-1} + (1 - p(e_2, \theta))u'(c_b)^{-1} \right) f(\theta)d\theta = \lambda^{-1}. \end{aligned}$$

This expression is such that the inverse of the Lagrangian multiplier equals the sum of the average of the inverses of the marginal utilities of consumption. The same result holds in Diamond and Sheshinski (1995) because  $\frac{1}{u'(c_a)} \frac{\partial e_1}{\partial c_a} = -\frac{1}{u'(c_d)} \frac{\partial e_1}{\partial c_d}$  and  $\frac{1}{u'(c_d)} \frac{\partial e_2}{\partial c_d} = -\frac{1}{u'(c_b)} \frac{\partial e_2}{\partial c_b}$ .

**Proof that  $c_d > c_b$  when social workers are altruist.** We sum of the first order conditions in  $c_d$  et  $c_b$  at the point  $c_d = c_b$ . This implies that  $\theta_d = \theta_b$  and  $p(e_1(\cdot)) = p(e_2(\cdot)) = \underline{p}(\theta)$ ,

therefore:

$$[u'(c_d) - \lambda](1 - F(\theta_d)) = (1 - c_a + c_d)f(\theta_d)u'(c_d).$$

This equation is exactly the one defining the second-best value of  $c_d$  without tagging technology.

If we plug this equation in equation (6), the first order condition with respect to  $c_d$ , at the point  $c_b = c_d$  we have that

$$[u'(c_d) - \lambda] \left( \int_{\theta_d}^{\theta_b} p(e_1, \theta) f(\theta) d\theta + \int_{\theta_b}^{\infty} p(e_2, \theta) f(\theta) d\theta \right) > \lambda(1 - c_a + c_d)u'(c_d) \underbrace{p(e_1, \theta_d)}_{\underline{p}(\theta_d)} f(\theta_d),$$

as  $p(\cdot)$  increases with  $\theta$ . This result is in line with the one of Diamond and Sheshinski (1995).

Then, we have  $c_d > c_b$  as  $u'(c_d) \geq \lambda$ , with equality when  $\theta_d = 0$ .

## References

- 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.
- Autor, D.H and M.G. Duggan, 2006. The Growth in the Social Security Disability Rolls: A Fiscal Crisis Unfolding, *Journal of Economic Perspectives*, 20(3), 71–96.
- Behncke, S., Frölich, M. and M. Lechner, 2010. A caseworker like me - Does the similarity between the unemployed and their caseworkers increase job placement?, *Economic Journal*, 120, 1430–1459.
- Besley, T. and M. Ghatak, 2005. Competition and Incentives with Motivated Agents, *American Economic Review*, 95(3), 616–636.
- Berkel, B. and A. Börsch-Supan, 2004. Pension Reform in Germany: The Impact on Retirement Decisions, *FinanzArchiv*, 60(3), 393–421.
- Boadway, R., N. Marceau and M. Sato, 1999. Agency and the design of welfare systems, *Journal of Public Economics*, 1–30.
- Bowles, S. and S. Polanía-Reyes, 2012. Economic Incentives and Social Preferences: Substitutes or Complements?, *Journal of Economic Literature*, 50, 368–425.

- Buurman, M and R. Dur, 2008. Incentives and the Sorting of Altruistic Agents into Street-Level Bureaucracies, *Tinbergen Institute Discussion Paper*.
- De Jong, P., M. Lindeboom and B. van der Klaauw, 2011. Screening disability insurance applications, *Journal of the European Economic Association*, 9(1), 106–129.
- Delfgaauw, J. and R. Dur, 2008. Incentives and workers' motivation in the public sector, *Economic Journal*, 118, 171–191.
- Diamond, P. and E. Sheshinski, 1995. Economic aspects of optimal disability benefits, *Journal of Public Economics*, 57, 1–23.
- Dixit, A., 2002. Incentives and Organizations in the Public Sector: An Interpretative Review, *Journal of Human Resources*, 38(4), 696–727.
- Duclos, J.-Y., 1995. Modelling the take-up of state support, *Journal of Public Economics*, 58, 391–415.
- Francois, P. and M. Vlassopoulos, 2008. Pro-social Motivation and the Delivery of Social Services, *CESifo Economic Studies*, doi:10.1093/cesifo/ifn002.
- Frey, B.S. and R. Jegen, 2001. Motivation Crowding Theory, *Journal of Economic Surveys*, 15(5), 589–611.
- Gruber, J. and J.D. Kubik, 1997. Disability insurance rejection rates and the labor supply of older workers, *Journal of Public Economics*, 64, 1–23.
- Kleven, H.J. and W. Kopczuk, 2011. Transfer Program Complexity and the Take-Up of Social Benefits, *American Economic Journal: Economic Policy*, 3, 54–90.
- Laffont, J.-J. and J. Tirole, 1993. A Theory of Incentives in Procurement and Regulation. MIT Press.
- Naegelen, F. and M. Mougeot, 2011. Power of incentives with motivated agents in public organizations, *Journal of Public Economic Theory*, 13(3), 391–416.
- OECD, 2003. Transforming Disability into Ability: Policies to Promote Work and Income Security for Disabled People. Directorate for Employment, Labor and Social affairs, Section on Sickness and Disability Policies.

- Parsons, D.O., 1991. The health and earnings of rejected Disability Insurance claimants: Comment, *American Economic Review*, 81, 1419–1426.
- Parsons, D.O., 1996. Imperfect 'tagging' in social insurance programs, *Journal of Public Economics*, 62, 183–207.
- Perry, J.L., Hondeghem, A. and L.R. Wise, 2010. Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future, *Public Administration Review*, September- October, 681–690.
- Prendergast, C., 2007. The Motivation and Bias of Bureaucrats, *American Economic Review*, 97(1), 180–196.
- Prendergast, C., 2008. Intrinsic Motivation and Incentives, *American Economic Review: Papers & Proceedings*, 98(2), 201–205.
- Tirole, J., 1992. Collusion and the Theory of Organizations, *In Advances in Economic Theory: Proceedings of the Sixth World Congress of the Econometric Society*, ed. J.-J. Laffont, 2, 151–206.
- von Wachter, T., Song, J. and J. Manchester, 2011. Trends in Employment and Earnings of Allowed and Rejected Applicants to the Social Security Disability Insurance Program, *American Economic Review*, 101, 3308–3329.