Imperfect Tagging Revisited

Rehn, Eric

2007

Imperfect Tagging Revisited

Eric Rehn
eric.rehn@nek.lu.se
Department of Economics,
Lund University,
Box 7082,
220 07 Lund,
Sweden

February 19, 2007

Abstract

Revisiting Parsons' 1996 article about disability insurance with imperfect tagging in a two type-economy – individuals are either able or disabled. Here Parsons' analysis is extended in several directions. The model is generalized to allow for different utility functions over work status. The analysis extends to three different cases of a two-type economy. Finally Parsons' model is extended to three types: able, partially disabled and disabled - adapting the model to disability insurances allowing for more than two degrees of disability. The results are consistent with Parsons', but a complete ranking of the consumption allocations cannot be achieved in the general case.

Keywords: social insurance; imperfect tagging; partial disability
JEL classification: H21; H53

1 Introduction

Disability, be it partial or full, lowers individuals' ability to sustain a sufficient income. Disability insurance aims to provide income support to these individuals. Like all insurance policies a disability insurance is vulnerable to excessive use - moral hazard. In their seminal paper Diamond & Mirrlees (1978) define the basic moral hazard condition and show that in a two-type-economy, i.e. individuals are either able or disabled, an optimal disability insurance can be constructed taking moral hazard into account. This policy induces all able individuals to work if the income from working is sufficiently more generous than the benefit offered to non-workers. This implies a great strain on the targeted group for the disability insurance - the disabled, who cannot work. The benefits will generally be low implying a low replacement rate. To improve the efficiency (the replacement rate) of disability insurance tagging/classification of individuals is both used in practice, and suggested in theory.

Tagging is the practice of assigning a tag to individuals based on their characteristics, e.g. able and disabled. Akerlof (1978) show that differentiating between individuals through tagging may improve efficiency.[1] In his model the screening process (tagging process) never mistakes an able individual for a disabled (however not necessarily all disabled receive a tag) implying that the tagged disabled will be compensated fully for their loss of income. Empirical studies and casual observation suggest that mistakes are made and that the classification of individuals is less accurate than Akerlof envisions. The screening process produces two types of errors, often labeled type-I error and type-II error. Type-I errors arise when truly disabled are not tagged, while type-II errors are the opposite i.e. able individuals receiving a disability tag. Empirical studies show that the classification errors are substantial; Nagi (1969) concludes in an early study that both the type-I error and the type-II error is 20% in the American disability insurance (SSDI).[2] Benitez-Silva et al (2004) uses self-reported ability to evaluate the classification error and finds similar results for the type-II error and concludes that the type-I error might be as large as 60%.

Imperfect tagging in disability insurance in a theoretical setting has been studied by e.g. Diamond & Sheshinski (1995) and Parsons (1996). Both articles assume that some individuals are missclassified by the screening process. Diamond & Sheshinski's two-type-model recognizes the presence of both type-I and type-II errors but does not include incentives for the latter group in their optimal disability insurance. Parsons includes the incentives of the "false positives", the able with disability tag, in the optimal program. He shows in two-type-model, where individuals are either able or disabled, that it is optimal to provide work incentives for able individuals with disability tag. Salanié (2002) shows that Parsons results are very general as long as leisure is a normal good.

This paper extends Parsons' analysis in several directions, most profoundly the model is extended to three types - able, disabled and partially disabled. Several existing disability insurances allow for more than two degrees of disability. For example, the Swedish disability insurance has four degrees of disability; 25%, 50%, 75%, 100%. The analysis of the three-type-economy provides results that are applicable to all settings with multiple degrees of disability. Furthermore, this paper extends Parsons' two-type-model in two directions and compares social welfare over the different settings. Thus, the analysis stretches over three different cases of the two-type-model with imperfect tagging. The first case is Parsons' "four-price model". That is, the model is solved such that all able individuals work irrespective of tag-status. In the second case the model is solved while ignoring the type-II error, i.e. treating all tagged individuals as truly disabled. In the third case the model is solved such that all able individuals work irrespective of tag-status and that all workers have the same salary. All three models are illustrated by logarithmic examples. The three cases are meant to

[1] This in a optimal taxation setting, which is a similar setting to the one used here.
[2] Social Security Disability Insurance

represent that policy-makers may have other goals and priorities than income support to the disabled and that these change the structure of the optimal disability insurance. Notably, the same basic social welfare function is used in all three cases, but in each case modifications are made to accommodate for the different priorities of the policy-makers.

Policy-makers and insurance administration are for simplicity treated as one entity in the formal analysis (both in the two- and three-type-economy) such that the entire economy, in principle, is the disability insurance. This is of course limiting in the sense that disability insurance generally is a part of a bigger economy and there is, in any economy, also ample room for conflicts of interest between policy-makers and insurance administrations. However, the simplification makes the models tractable and is standard in the literature. Moreover the models may serve as examples of a welfare economy with low friction between politicians and administration.

Parsons' analysis is also extended by using different utility functions for workers and non-workers in all models.[3] The idea is that consumption is valued differently by workers compared to non-workers - consumption when not working is here assumed to render the individuals greater utility than consumption when working. This, of course, decreases the incentive to work, thus adding to the disutility of working. Furthermore the moral hazard condition introduced by Diamond & Mirrlees (1978) is imposed throughout the paper.

Section 2 presents the three cases of the two-type-model and the social welfare comparison. The analysis of the three-type-model is presented in section 3. Section 4 concludes and makes brief suggestions for future research.

2 The Two-Type Economy

Consider an economy with two types of individuals: able and disabled. The proportion of able individuals is ℓ^A, accordingly the proportion of disabled is $\ell^D = 1 - \ell^A$. Individuals in this model are either fully able to work (able) or not able to work at all (disabled). θ is the degree of disability which in the two-type-model is dichotomous, 0 (able) or 1 (disabled). All individuals in the economy are identical in all aspects but the degree of disability. All working individuals have the same marginal product and the marginal product is normalized to 1

2.1 Case 1: Imperfect Tagging

A disability insurance is designed to cover the loss of income (at least partially). A screening process decides whether an individual is eligible for disability benefits. The screening process is exogenous to the model and assigns a disability "tag" to an individual with probability p_θ. The tagging is imperfect in that it, with positive probability, fails to assigns disability tags to truly disabled individuals and, with positive probability, assign disability tags to able individuals.

[3] In this following Diamond & Sheshinski (1995)

2.1.1 The screening process

Able individuals (A) receive a disability tag (T) with probability p_0 and no tag (NT) with probability $\varphi_0 = 1 - p_0$. The disabled (D) however receive a tag with probability p_1 and no tag with probability $\varphi_1 = 1 - p_1$. The probability of getting a disability tag is greater for the truly disabled than for the able, that is $p_1 > p_0 > 0$ and thus $\varphi_0 > \varphi_1 > 0$. Thus, the type-1-error is $\varphi_1$and the type-2-error is p_0.

2.1.2 The utility functions

Able individuals who are working, f, have the following utility of consumption:

$$U_f^A = u\left(c\right) - D_\theta$$

where $D_\theta > 0$ is the disutility of working depending on the degree of disability, for the able we have $\theta = 0$, thus $U_f^A = u\left(c\right) - D_0$. Furthermore, the utility of consumption when not working, d, is given by:

$$U_d^i = v\left(c\right)$$

for $i = A, D$. $u(c)$ and $v(c)$ are concave and increasing in consumption, the marginal utilities go from ∞ to 0 as c goes from 0 to ∞ (c.f. Diamond & Sheshinski, 1995, Parsons, 1996). Work and consumption is preferable to no work and no consumption: $u(c) - D_0 > v(0)$. That is, able individuals will work in the absence of a disability insurance. Work is unpleasant such that; $u(c) - D_0 < v(c)$, *all* c. It is assumed that if $U_f^A(c) = U_d^A(\tilde{c}) \Longleftrightarrow u\left(c\right) - D_0 = v(\tilde{c})$ it follows that $u'\left(c\right) < v'(\tilde{c})$.[4] This is the moral hazard condition introduced by Diamond & Mirrlees (1978). The moral hazard condition states that equating the utilities between non-workers and workers will render a marginal utility higher for non-workers . This assumption is needed to characterize the relation between the utility functions $u\left(c\right)$ and $v\left(c\right)$.[5] Notably the moral hazard condition is satisfied in the models presented here and also for $u(c) = v(c)$.

2.1.3 Characterizing individuals and policy instruments

Given the two-sided classification error in the screening process the individuals are characterized by their work status, their ability and their tag status and can thus be divided in to 8 groups

$$\begin{aligned}[\mathit{ability}, \mathit{work}, \mathit{tag}] : [A, f, NT]\,, [A, f, T]\,, [A, d, NT]\,, [A, d, T]\,, [D, f, NT]\,, \\ [D, f, T]\,, [D, d, NT]\,, [D, d, T]\,.\end{aligned}$$

Since it is impossible for the disabled to work the two groups $[D, f, NT]$, $[D, f, T]$ are not realized in the model. The disability insurance administration,

[4] for utility functions that are unbounded from below.

[5] Note that Parsons (1996) has $U_f^A = u\left(c\right) + b, b < 0$ and $U_d^i = u\left(c\right)$,for $i = A, D$, if using the current notation.

performing the screening process and deciding on benefit levels, knows that there are six groups of individuals in the economy. However, it cannot fully distinguish between able and disabled individuals, i.e. disability is imperfectly observed. Work- and tag-status is observed by the insurance administration. Since the screening process is exogenous to the model the insurance administration's only available policy instrument is the benefit levels, here modelled as the different consumption allocations to the distinguishable groups in the economy: c_f, c_f^T, c_d and c_d^T, where the superscript is the tag status and the subscript is the work status.[6]

2.1.4 The optimization problem

In this case the model is solved such that all able individuals work (c.f. Parsons, 1996). The policy objective is to maximize the expected social welfare by choosing the benefit level (the consumption vector). Answering the question: which are the optimal consumption allocations given that all able individuals work? In this case the expected social welfare is given by:

$$\begin{aligned} SWF_1 &= (1-p_0)\ell^A(u(c_f) - D_0) + p_0\ell^A(u\left(c_f^T\right) - D_0) + \\ &+ (1-p_1)(1-\ell^A)v(c_d) + p_1(1-\ell^A)v\left(c_d^T\right) \end{aligned} \tag{1}$$

In maximizing expected social welfare the insurance administration is constrained by the resources available in the economy and by the work constraints. The resource constraint is:

$(1-p_0)\ell^A c_f + p_0\ell^A c_f^T + (1-p_1)(1-\ell^A)c_d + p_1(1-\ell^A)c_d^T \leq M$

where M may be all the resources in the economy or fraction of them. It is assumed that all working individuals have the same marginal product and this is normalized to 1, thus $M = \ell^A \times 1$ if M is all the resources in the economy. Note that individuals have no reason to forgo consumption in this model and the resource constraint therefore holds with equality (Parsons, 1996). The insurance administration is also constrained by the work constraints of the able individuals. If the able are to work irrespective of their tag, the following constraints need to be fulfilled:

$$u(c_f) - D_0 \geq v(c_d)$$

for no tag, and

$$u\left(c_f^T\right) - D_0 \geq v\left(c_d^T\right)$$

for those with disability tag.

[6] That is, the insurance adminstration uses the information it posesses, about work- and tag-status to construct the policy instrument.

Thus, the insurance administration solves the following maximization problem:

$$\max_{c_f, c_f^T, c_d, c_d^T} (1-p_0)\ell^A(u(c_f) - D_0) + p_0\ell^A(u\left(c_f^T\right) - D_0) + \tag{2}$$
$$+(1-p_1)(1-\ell^A)v(c_d) + p_1(1-\ell^A)v\left(c_d^T\right)$$

subject to the resource constraint:

$$(1-p_0)\ell^A c_f + p_0\ell^A c_f^T + (1-p_1)(1-\ell^A)c_d + p_1(1-\ell^A)c_d^T = M \tag{3}$$

and the work constraints:

$$u(c_f) - D_0 \geq v(c_d) \tag{4}$$
$$u\left(c_f^T\right) - D_0 \geq v\left(c_d^T\right) \tag{5}$$

The first order conditions, λ_j, for $j = i, ii, iii$ are the Lagrange multipliers:

$\frac{\partial L}{\partial C_f} = (1-p_0)\ell^A u'(c_f) - \lambda_i(1-p_0)\ell^A + \lambda_{ii}u'(c_f) = 0$
$\frac{\partial L}{\partial C_d} = (1-p_1)(1-\ell^A)v'(c_d) - \lambda_i(1-p_1)(1-\ell^A) - \lambda_{ii}v'(c_d) = 0$
$\frac{\partial L}{\partial C_f^d} = p_0\ell^A u'\left(c_f^T\right) - \lambda_i p_0\ell^A + \lambda_{iii}u'\left(c_f^T\right) = 0$
$\frac{\partial L}{\partial C_d^d} = p_1)(1-\ell^A)v'\left(c_d^T\right) - \lambda_i p_1(1-\ell^A) - \lambda_{iii}v'\left(c_d^T\right) = 0$

Since the resource constraint is binding we know that $\lambda_i > 0$, and it can be shown that $\lambda_{ii} > 0$ and $\lambda_{iii} > 0$ (see appendix), i.e the work constraints binding.

2.1.5 The consumption allocations

How are the different consumption allocations related to each other? First, in the optimal program both work constraints are binding and thus:

$u(c_f) - v(c_d) = D_0$
$u\left(c_f^T\right) - v\left(c_d^T\right) = D_0$
$\Rightarrow u(c_f) - v(c_d) = u\left(c_f^T\right) - v\left(c_d^T\right)$

Moreover the first order conditions yields the following redistribution principle (c.f. Parsons, 1996):

$$\omega\frac{1}{u'(c_f)} + (1-\omega)\frac{1}{v'(c_d)} = \omega^D\frac{1}{u'(c_f^T)} + (1-\omega^D)\frac{1}{v'(c_d^T)} \tag{6}$$

where $\omega = \frac{(1-p_0)\ell^A}{(1-p_1)(1-\ell^A)+(1-p_0)\ell^A}, \omega^D = \frac{p_0\ell^A}{p_1(1-\ell^A)+p_0\ell^A}$. Both these weights are positive (by the screening mechanism) and less than or equal to one. To rank the consumption levels the ranking within the NT-state and the D-state are first determined; it turns out that $c_f > c_d$ and $c_f^T > c_d^T$, since $v(\tilde{c}) > u(c) + D_0$ when $\tilde{c} = c$, and $v(c_d) = u(c_f) + D_0 \Rightarrow u'(c_d) < v'(c_d)$, $v\left(c_d^T\right) = u\left(c_f^T\right) +$

$D_0 \Rightarrow u'\left(c_d^f\right) < v'(c_d^T)$. That is, $v(c)$ is always above $u(c) - D_0$ for all positive consumptions and $v(c)$ is steeper than $u(c)$ when the utility levels are equalized, thus consumption needs to be higher when working than when not working for the utilities to be equal, an intuitive result since there is a disutility of working in the model. Now, how are the consumption allocations in different tag-states related to each other? Unlike in Parsons (1996) the consumption allocations cannot be fully ranked in this setting.

$$c_f^T > c_f \lesseqqgtr c_d^T > c_d$$

Proof. $p_1 > p_0$ implies that $\omega > \omega^D$. Knowing this, now assume that $c_f = c_f^T$, which implies that $c_d = c_d^T$ since the work constraints are binding. Thus $\frac{1}{u'(c_f)} = \frac{1}{u'(c_f^T)}$ and $\frac{1}{v'(c_d)} = \frac{1}{v'(c_d^T)}$, and the left-hand side of equation 6 is greater than the right-hand side. That is $\omega\frac{1}{u'(c_f)} + (1-\omega)\frac{1}{v'(c_d)} > \omega^D\frac{1}{u'(c_f^T)} + (1-\omega^D)\frac{1}{v'(c_d^T)}$. However in the optimal program the redistribution principle is satisfied with equality, thus the redistribution principle requires that, since LHS is an increasing function of c_f and RHS an increasing function of c_f^T, c_f is reduced relative to c_f^T. Thus $c_f < c_f^T$ in the optimal program implying that $c_d < c_d^T$, since the work constraints are binding and both $v(c)$ and $u(c)$ are increasing functions. However, the ranking between c_d^T and c_f cannot be established with the current assumptions about utility functions. ■

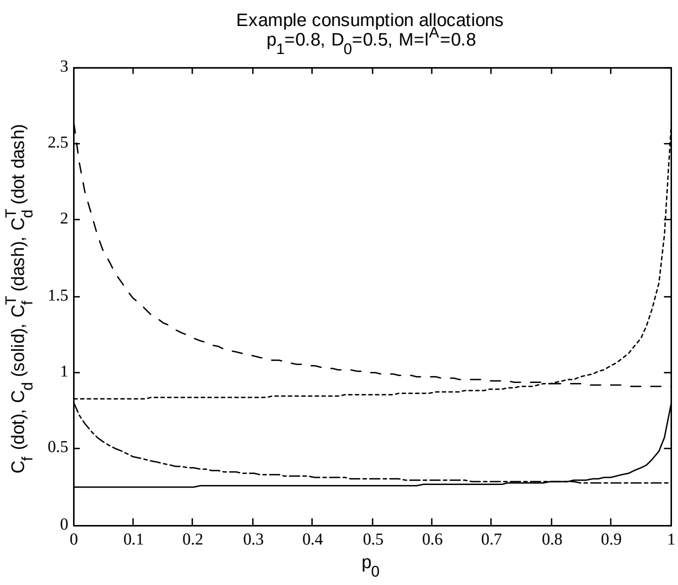
Concerning the ranking of c_d^T and c_f it can be concluded that $u'(c_f) < v'(c_d^T)$ in the optimal program, since the smallest element in a weighted average cannot exceed or equal the greatest element of another weighted average if the averages are equal and have positive weights, that is $\frac{1}{u'(c_f)} > \frac{1}{v'(c_d^T)}$. It can also be established that $u(c_f) - D_0 < v\left(c_d^T\right)$ since it is known that $c_d^T > c_d$. Assuming $v'(c) = u'(c)$ all c would ensure that the consumption allocations could be fully ranked ($c_d^T < c_f$), and this is implictily assumed in the examples below, but not assumed in the general model.

It is not obvious that tagged non-workers should have a lower income (consumption) than untagged workers in this model, as it is in Parsons model. It might be optimal, from a utilitarian point of view, to increase the consumption of tagged non-workers above the level of the untagged workers if the marginal utility of consumption for the former group is high enough. However, this setup follows Parsons' results in providing a premium for being tagged and working, and also in keeping income for untagged non-workers low (e.g. low social assistance level).

2.1.6 Logarithmic example

To illustrate the model, as is done by Diamond & Mirrlees (1978) and, of course, Parsons (1996), a logarithmic example is constructed. This example serve as intuition and sheds some light on policy issues such as the adoption of consumption allocations to increasing type-I and type-II errors and increasing work disutility for the able. Assume that $U_f^A = u(c) - D_0 = \ln c - D_0$ and $U_d^i = v(c) = \ln 2c$ for $i = A, D$. This specification of the model is solved for c_f, c_f^T, c_d and c_d^T, which are then plotted in figure 1-3 below.

Figure 1 shows the consumption allocations assigned to the different groups in the optimal insurance program when the probability of able getting a disability tag (p_0) varies. This is done under the following assumptions: $p_1 = 0.8$, $D_0 = 0.5$ and $M = \ell^A = 0.8$. These assumptions follow the assumptions made by Parsons (1996) and figure 1 thus resembles Parsons' figure 2[7]. The difference between figures stems from the specification of the utility functions, that is the introduction of the moral hazard condition and the higher utility of consumption for non-workers. Note that the required difference in consumption between working and not working is greater than in Parsons' example. For $p_0 = 0.2$, the consumption allocations are $\left[c_f, c_d, c_f^T, c_d^T\right] = [0.83, 0.25, 1.23, 0.37]$ in this example compared to $[0.819, 0.497, 0.996, 0.604]$ in Parsons' example. Thus this model requires greater reward for workers, as expected. Note that when $p_0 > 0.8$ this violates the assumption of the model, i.e. $p_1 > p_0$, explaining the "exponential" growth of the consumption allocations for the non-tagged groups. As in Parsons, when there are no type-I and type-II errors, i.e. $p_0 = 0, p_1 = 1$, then both workers and non-workers get the consumption allocation 0.8 - the disabled are fully insured.

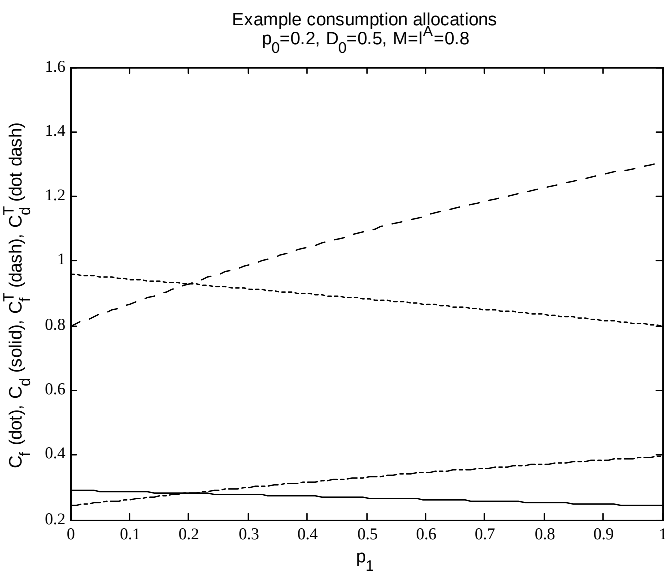


(Figure 1)

Figure 2 illustrates the optimal insurance program when the probability that disabled will receive a disability tag (p_1) varies. Besides $p_0 = 0.2$ the assumptions are the same as for figure 1. For $p_1 < 0.2$, the assumptions of the model

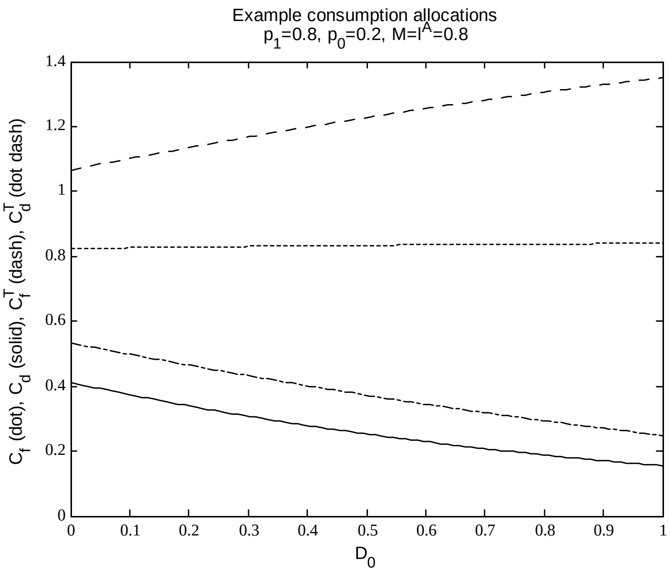
[7] see page 197

are once again violated, explaining the deviation from the results in the general model for this range of p_1. The difference in consumption allocation between the working untagged and the non-working tagged is shrinking as the type-I error becomes smaller (increasing probability that disabled receive a disability tag). That is, the coverage of the income loss for the targeted group, the disabled, will be improved as the screening process is improved. The relative proportion of able and tagged individuals will fall with increasing p_1 and this in combination with the increasing consumption allocation for the tagged non-workers will drive up the consumption allocation for the tagged workers. Since fewer of the disabled will be untagged non-workers (with increasing p_1), the consumption for this group will go down to and the consumption allocation for the untagged workers follow this movement (keeping the optimal difference between them).



(Figure 2)

Figure 3 plots the consumption allocations over the disutility of working and replicates the familiar pattern from the figures above. Notably the consumption allocation for untagged and working individuals is only modestly rising (virtually constant) over the range of disutilities. The small rise in c_f, is compensated by the fall c_d, the consumption allocation for the untagged non-workers. Also for tagged non-workers the allocated consumption declines. The decline in consumption for non-workers is coupled with a increase in consumption for workers (although marginally for untagged workers). This is needed to ensure that working is an attractive option for the individuals in the economy.



(Figure 3)

Notably when the assumption that $p_0 > p_1$ is violated, being tagged and not being tagged switches roles. This can be seen in figure 1 and 2 where not having a tag is a better signal of disability than having a tag when $p_0 > p_1$and thus the consumption allocations are greater for the untagged groups in the economy in this range. Clearly, when p_1 is high and p_0 low, i.e. the screening process is very reliable, the model approaches full insurance and it is optimal to compensate the few able and tagged individuals substantially if they choose to work in spite of the tag. Can it be socially optimal to ignore a sufficiently low type-II error (low p_0) and accept that some able receive disability benefits (avoiding the extra cost of compensating them)? The next section deals with this case.

2.2 Case 2: Imperfect Tagging - ignoring the type-II error[8]

Parsons (1996) compares his imperfect tagging model with a model with "incomplete tagging" or as he also calls it "the Akerlof model"[9]. The incomplete tagging model features a type-I error ($1 - p_1 > 0$) but assumes that there is no type-II error ($p_0 = 0$). As discussed in the introduction empirical studies suggest that both errors exist, therefore an alternative approach is used in this section. Here it is assumed that the type-II error is positive but ignored by the insurance administration. A reason for this is could be that the insurance administration believes that the screening process produces very small type-II errors and thus

[8] Diamond & Sheshinski (1995) analyses a similar case

[9] See the introduction for discussion concerning Akerlof (1978).

concludes that it is not worth the effort to construct an insurance with four different consumption allocations - assumes that it is optimal to treat all tagged individuals as truly disabled. Able individuals will either be untagged and work or tagged and not working. The insurance administration recognizes that the type-I error might be substantial and thus constructs a disability insurance with three different consumption allocations; one for the untagged workers c_f, one for untagged non-workers c_d, and one for tagged non-workers c_d^T. The consumption allocations are chosen to maximize social welfare. The social welfare function in this setup is:

$$\begin{aligned} SWF_2 &= (1-p_0)\ell^A(u(c_f)-D_0)+p_0\ell^A v\left(c_d^T\right)+ \\ &\quad +(1-p_1)(1-\ell^A)v(c_d)+p_1(1-\ell^A)v\left(c_d^T\right) \end{aligned} \tag{7}$$

2.2.1 The optimization problem

The work constraint for the able without tag is, once again: $u(c_f)-D_0 \geq v(c_d)$. Since the insurance administration is only interested in ensuring that the untagged and able individuals work this is the only work constraint. It believes that very few or no able individuals are tagged. Furthermore the insurance administration faces a binding resource constraint:

$$\begin{aligned} &(1-p_0)\ell^A c_f + p_0\ell^A c_d^T + (1-p_1)(1-\ell^A)c_d + p_1(1-\ell^A)c_d^T = \\ &= (1-p_0)\ell^A c_f + (1-p_1)(1-\ell^A)c_d + \left(p_0\ell^A + p_1(1-\ell^A)\right)c_d^T = M \end{aligned}$$

Note that M in this case depend on the size of the type-II error, e.g. if M is the working population times their marginal product (normalized to one), then the working population is decreasing with the type-II error since the able and tagged individuals have no incentive to work. Thus M could be $M=\ell^A(1-p_0)*1$, this specification of M is used in the examples below.

Thus the insurance administration solves the following maximization problem:

$$\begin{aligned} &\max_{c_f,c_d,c_d^T}(1-p_0)\ell^A(u(c_f)-D_0)+p_0\ell^A v\left(c_d^T\right)+ \\ &+(1-p_1)(1-\ell^A)v(c_d)+p_1(1-\ell^A)v\left(c_d^T\right) \end{aligned} \tag{8}$$

subject to the resource constraint:

$$(1-p_0)\ell^A c_f + (1-p_1)(1-\ell^A)c_d + \left(p_0\ell^A + p_1(1-\ell^A)\right)c_d^T = M \tag{9}$$

and the work constraint:

$$u(c_f)-D_0 \geq v(c_d) \tag{10}$$

The first order conditions, where λ_j, for $j=i,ii$ are the Lagrange multipliers:

$\frac{\partial L}{\partial C_f} = (1-p_0)\ell^A u'(c_f) - \lambda_i(1-p_0)\ell^A + \lambda_{ii}u'(c_f) = 0$
$\frac{\partial L}{\partial C_d} = (1-p_1)(1-\ell^A)v'(c_d) - \lambda_i(1-p_1)(1-\ell^A) - \lambda_{ii}v'(c_d) = 0$
$\frac{\partial L}{\partial C_d^d} = p_0\ell^A v'\left(c_d^T\right) + p_1(1-\ell^A)v'\left(c_d^T\right) - \lambda_i p_0\ell^A - \lambda_i p_1(1-\ell^A) = 0$

The first two FOC:s are unchanged from the previous problem thus the work constraint is binding and the resource constraint is binding by assumption.

2.2.2 The consumption allocations

The FOC:s for untagged individuals are the same as in the original problem and thus will the left hand side of the redistribution principle be the same:

$\omega\frac{1}{u'(c_f)} + (1-\omega)\frac{1}{v'(c_d)} = \frac{1}{\lambda_i}$, where $\omega = \frac{(1-p_0)\ell^A}{(1-p_1)(1-\ell^A)+(1-p_0)\ell^A}$ and it is obvious from the third FOC that $\frac{1}{\lambda_i} = \frac{1}{v'(c_d^T)}$. Accordingly the redistribution principle in this case is[10]:

$$\omega\frac{1}{u'(c_f)} + (1-\omega)\frac{1}{v'(c_d)} = \frac{1}{v'(c_d^T)} \tag{11}$$

The binding work constraint ensures that $c_f > c_d$ in the optimal program and the consumption of tagged non-workers (c_d^T) is greater than c_d. Again the ranking between c_f and c_d^T cannot be established, thus $c_f \lesseqgtr c_d^T > c_d$.

Proof. Assume that $c_d^T = c_d \Rightarrow LHS > RHS$ since the weighted average of the number ($\frac{1}{v'(c_d^T)}$) and a greater number ($\frac{1}{u'(c_f)}$) is greater than the first number ($\frac{1}{v'(c_d^T)}$), with positive weights. LHS is an increasing function of c_d implying that c_d has to be decreased relative to c_d^T to fulfill the optimal redistribution principle, thus $c_d^T > c_d$. ∎

Note once again that $\frac{1}{v'(c_d^T)}$ cannot exceed $\frac{1}{u'(c_f)}$ (the greatest element of the weighted average) if the equality is to hold (with positive weights) thus $u'(c_f) < v'\left(c_d^T\right)$.

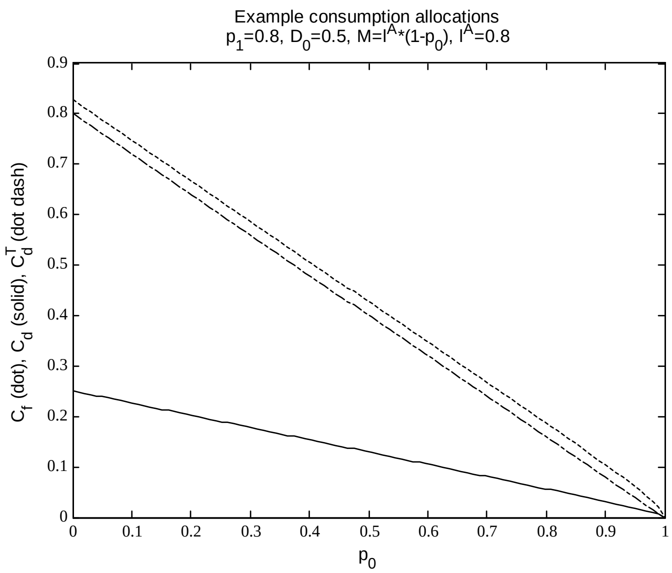
2.2.3 Logarithmic example

Assume that $u_f^A = u(c) - D_0 = \ln c - D_0$ and $u_d^i = v(c) = \ln 2c$ for $i = A, D$, and now solve the model for c_f, c_d and c_d^T.

In figure 4 it is assumed that $p_1 = 0.8$, $D_0 = 0.5$ and the proportion of able individuals in the economy is $\ell^A = 0.8$ and $M = \ell^A(1-p_0)$ (see discussion above). The figure plots the consumption allocations for the different groups over p_0 i.e. the probability that able individuals get a disability tag. The consumption allocations follow the general result in the model $c_f > c_d^T > c_d$. Figure 4 also shows that the viability of the disability insurance is falling with the size of the type-II error when the latter is ignored in the design of the insurance. With high type-II errors it is obvious that all groups are worse of compared to the imperfect tagging model. High p_0 obviously puts a strain on the disability insurance since it lowers the resources available to the insurance

[10] C.f. equation 13, page 192 in Parsons (1996)

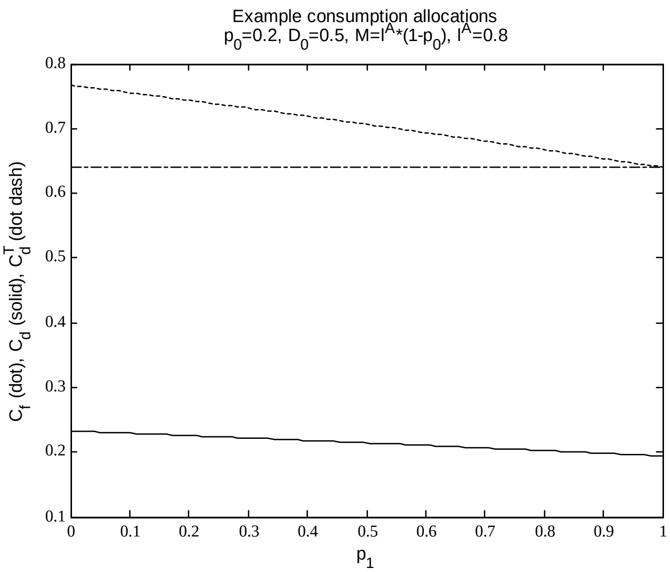
administration. Adding a type-II error of 0.2, as is assumed above, gives the following consumption allocations $\left[c_f, c_d^T, c_d\right] = [0.67, 0.64, 0.20]$ implying that the targeted group (the disabled) are better of in this model IF tagged, but worse off IF untagged compared to the imperfect tagging model. The untagged working population is worse off (0.67 compared to 0.83).



(Figure 4)

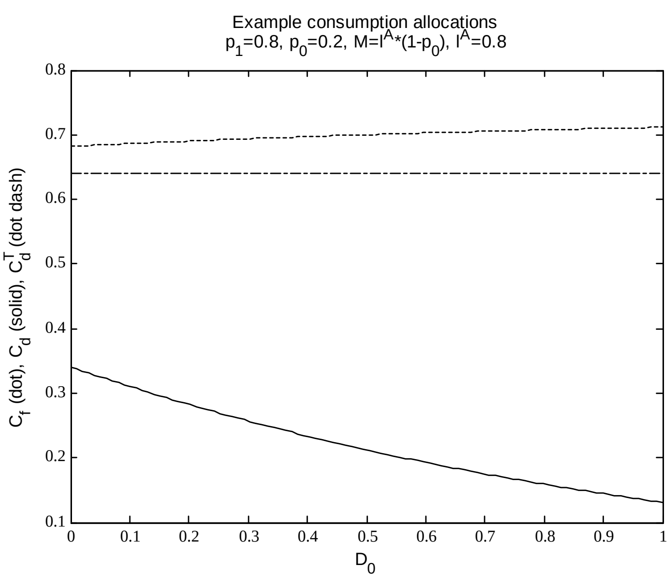
Figure 5[11] plots the consumptions over the type-I error, p_1, for $p_0 = 0.2$ and otherwise the same assumptions as in figure 4 apply. Here the restrictive effect of the positive type-II error is less visible, it only shows in the fact that the tagged non-workers does not reach the full insurance consumption 0.8 as they would be if $p_0 = 0$ (c.f. Parsons 1996). Instead they reach a "constrained full insurance level" which is lower (0.64). Constrained by the fact that the resources available in the economy is limited by the positive type-II error and full insurance in the sense that untagged workers and tagged non-workers receive the same consumption when $p_1 = 1$. That is, as more individuals with disability receive a disability tag, better screening to avoid type-I errors, the difference between c_f and c_d^T diminishes. The untagged non-workers will "bear an increasing share of the required work incentive differential.." (Parsons, 1996:194). The number untagged non-workers (untagged disabled) diminish as p_1 increases (to zero when $p_1 = 1$).

[11] C.f. figure 1 page 193 in Parsons (1996)



(Figure 5)

Figure 6 depicts the consumption allocations as a function of the disutility of working. Obviously, c_d^T is independent of the disutility since no one in this group is working. For the untagged workers the consumption is modestly increasing for increasing disutility. The untagged non-workers i.e. disabled without disability tag will again carry the large share of the required work incentive differential.



(Figure 6)

Assuming that it is optimal to ignore the type-II error is obviously a bad idea if this error is substantial. It would put a strain on the whole economy lowering the optimal consumption allocations for all groups in the economy (c.f. figure 4). For a modest type-II error the tagged and disabled is better off than in the original model (while the few able and tagged are worse off). The untagged individuals who does not work are generally treated more harshly in this setup (compared to the original model).

2.3 Case 3: Imperfect tagging - equalizing salaries

Assume that the insurance administration is reluctant to reward tagged workers in the way that is done in the original model, but still want all able individuals work. Equal salary could be a normative objective for the insurance administration (policy-makers), much like the Swedish policy objective of low variation in the income distribution. Thus the consumption allocation assigned to workers will be the same across tag-status in this specification of the model and the allocation must be chosen such that it is optimal to work for the able. The screening process is still producing a two-sided-classification error and, given the assumption above, the policy instruments available to the insurance administration are:c_f, c_d and c_d^T. Social welfare in this setting is represented by:

$$\begin{aligned} SWF_3 \quad &= \quad (1-p_0)\ell^A(u(c_f) - D_0) + p_0\ell^A(u(c_f) - D_0) + \\ &\quad +(1-p_1)(1-\ell^A)v(c_d) + p_1(1-\ell^A)v\left(c_d^T\right) \end{aligned} \tag{12}$$

2.3.1 The optimization problem

The insurance administration is once again constrained in the choice of consumption allocation by a resource constraint and two work constraints. The resource constraint is:

$$(1-p_0)\ell^A c_f + p_0\ell^A c_f + (1-p_1)(1-\ell^A)c_d + p_1(1-\ell^A)c_d^T = M \tag{13}$$

and the work constraints:

$$u(c_f) - D_0 \quad \geq \quad v(c_d) \tag{14}$$

$$u(c_f) - D_0 \quad \geq \quad v\left(c_d^T\right) \tag{15}$$

The first order conditions, where λ_j, for $j = i, ii, iii$ are the Lagrange multipliers:

$\frac{\partial L}{\partial C_f} = \ell^A u'(c_f) - \lambda_i \ell^A + \lambda_{ii} u'(c_f) + \lambda_{iii} u'(c_f) = 0$

$\frac{\partial L}{\partial C_d} = (1-p_1)(1-\ell^A)v'(c_d) - \lambda_i(1-p_1)(1-\ell^A) - \lambda_{ii}v'(c_d) = 0$

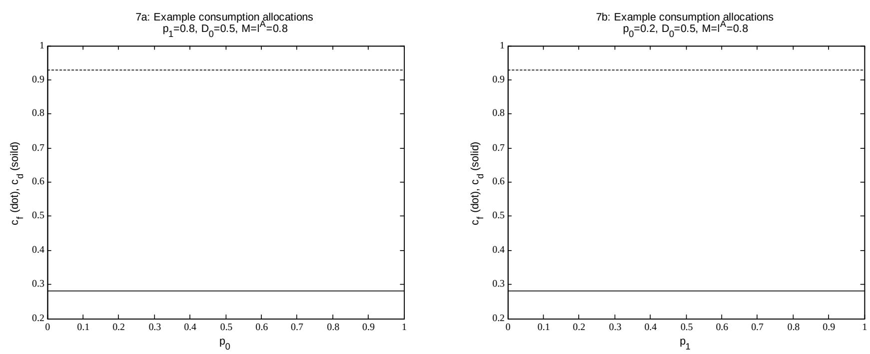
$\frac{\partial L}{\partial C_d^d} = p_1(1-\ell^A)v'\left(c_d^T\right) - \lambda_i p_1(1-\ell^A) - \lambda_{iii}v'\left(c_d^T\right) = 0$

It can be shown that the optimization has a solution when both work constraints are binding (see appendix) implying that $c_d = c_d^T < c_f$ through the specification of the utility functions. Thus it is optimal to have only one social insurance benefit (one consumption allocation for all non-workers) when it is a policy goal that all workers have the same salary irrespective of tag-status. The information given by the screening process is ignored and the model replicates the results in Diamond & Mirrlees (1978) - all able individuals will work if the salary is sufficiently more generous than the disability benefit. This is easily seen in the examples below where the consumption allocations are constant over p_0 and p_1 and the difference $c_f - c_d$ is increasing in D_0.

This solution to the imperfect tagging model suggests that if all workers are to receive the same salary then it is optimal to give equal consumption allocation (benefit) to all non-workers. The salary need to outweigh the disutility of working and is therefore strictly greater than the benefit given to non-workers.

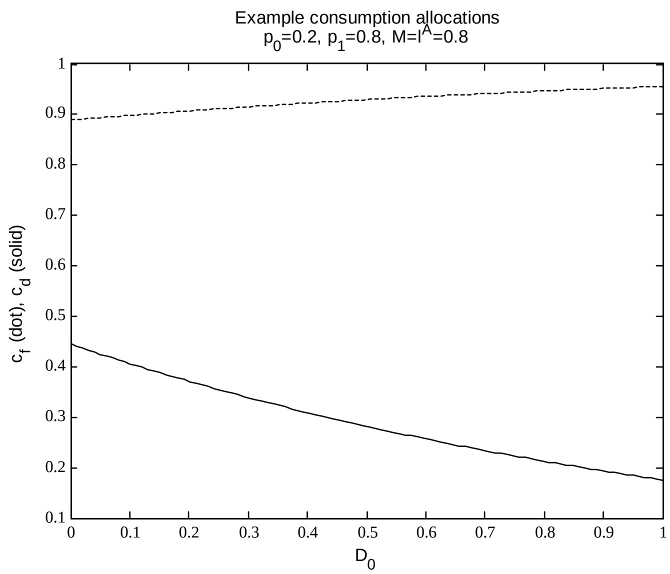
2.3.2 Logarithmic example

Assume that $u_f^A = u(c) - D_0 = \ln c - D_0$ and $u_d^i = v(c) = \ln 2c$ for $i = A, D$, and now solve the model for c_f, $c_d = c_d^T$. Figure 7a and 7b show that the consumption allocations are independent of the screening process (remember that its is exogenous to the model and thus without cost).



(Figure 7)

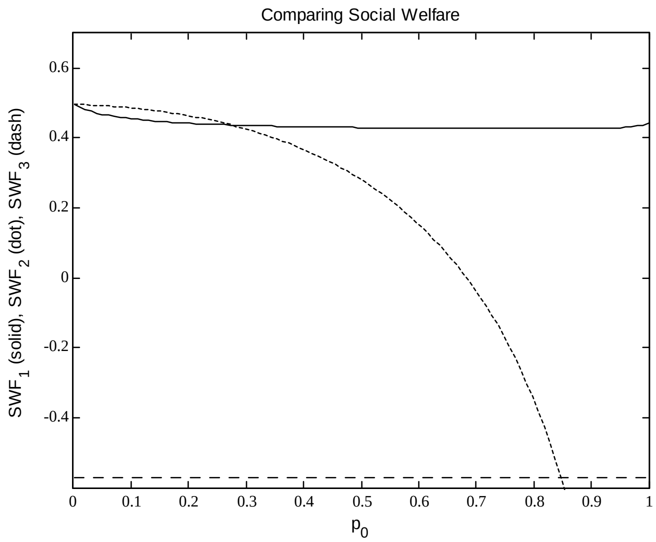
Figure 8 shows that the difference between the optimal salary and the optimal disability benefit is increasing in the disutility of working. Underlining the similarity with the result in Diamond & Mirrlees (1978).



(Figure 8)

2.4 Logarithmic example - comparing social welfare.

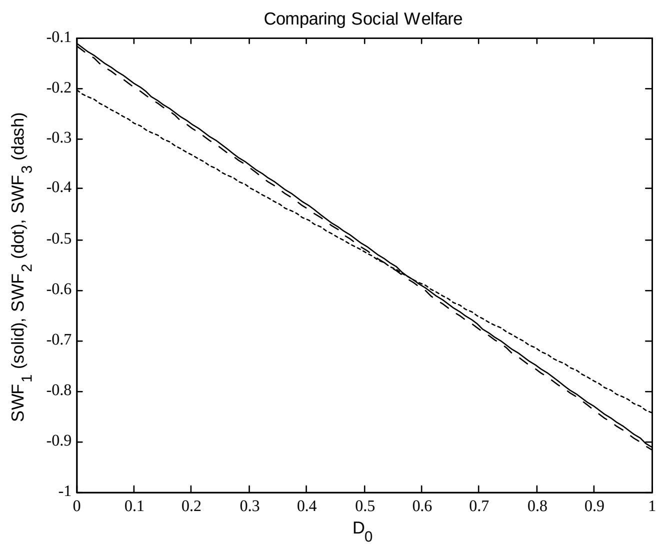
This section compares the resulting social welfare from the three models. Figure 9 plots social welfare in the three settings over the type-II error (with $D_0 = 0.5$). Social welfare is greater in case 1 and case 2 compare to case 3 for almost the whole range of p_0. Notably for small type-II errors case 2 produces the greatest social welfare while it is the worst model for very big type-II errors. The explanation is that for small p_0 it is optimal to ignore its existence, while big p_0 puts a strain on the whole economy - as discussed earlier.



(Figure 9)

In figure 10 social welfare is plotted against the disutility of working[12] for a given precision in the screening process. The screening process is characterized by $p_1 = 0.8$, $p_0 = 0.2$. Social welfare in model 3 is always worse than social welfare in case 1 (the original model), but better than case 2 for substantial range of work disutility. For work disutilities greater than 0.6 the ignoring the type-II error (case 2) produces greater social welfare than both model 1 and model 3.

[12] Following Parsons (1996)



(Figure 10)

For low type-II errors combined with high disutility of working it is obviously superior to ignore the type-II error. In all other circumstances the original model (case 1) yields the greatest social welfare. Thus, with high disutility of working it is optimal, in terms of social welfare, to allow all tagged individuals to receive disability benefits - as long as the type-II error is modest, that is as long as few able individuals are tagged as disabled. Furthermore it is intuitive that when the screening process produces big type-II errors it is harmful to ignore this imperfection and case 2 is the worst solution to the imperfect tagging model. Notably, ignoring all available information, as in case 3, is never the best alternative either case 1 or case 2 (or both) is better in the dimensions presented here. Reflecting the low replacement rate for the targeted groups in case 3, the disabled will not be replaced for their income loss in an optimal manner when the tagging information is ignored[13].

3 The Three-Type Economy

Now consider an economy with three types of individuals: able, partially disabled and disabled. The population weights for each type is ℓ^A, ℓ^P, and $\ell^D = 1 - \left(\ell^A + \ell^P\right)$ respectively. These population weights are exogenous to the model. The three types differ in their disutility of working, which is increasing in the degree of disability, $0 \leq \theta \leq 1$. All individuals in the economy are identical in all aspects but the degree of disability. All working individuals have the same marginal product and the marginal product is normalized to 1. The disabled are unable to work. Furthermore, all degrees of partial disability are,

[13] in line with the results presented by Akerlof (1978)

for simplicity, treated as one type. The partially disabled may work part-time or not work at all. The able, however, may work full time, part-time or not at all. Able individuals have $\theta = 0$, partially disabled $\theta = \bar{\theta}$ ($\bar{\theta} \in (0,1)$), and disabled have $\theta = 1$. Apart from the the addition of a third type is this model similar to the two-type-model - with appropriate modifications to accommodate for the third type.

3.1 Imperfect tagging - extended to three types

Again the aim is to design a disability insurance given an imperfect screening process. Being partially disabled and disabled imposes a loss of income on the individuals, the disability insurance aims to alleviate some of this loss.

3.1.1 The screening process

Able individuals receive a disability tag (T) with probability p_0, partial disability tag (P) with probability π_0 and no tag (NT) with probability $\varphi_0 = 1 - \pi_0 - p_0$. The disabled however receive a disability tag with probability p_1, partial disability tag with probability π_1 and no tag with probability $\varphi_1 = 1 - \pi_1 - p_1$. For the partially disabled are these probabilities given by $p_{\bar{\theta}}, \pi_{\bar{\theta}}$ and $\varphi_{\bar{\theta}} = 1 - \pi_{\bar{\theta}} - p_{\bar{\theta}}$ respectively. The probability of getting a disability tag is greater for the truly disabled than for the partially disabled which in turn is greater than the probability for the able, that is $p_1 > p_{\bar{\theta}} > p_0 > 0$. Furthermore; $\pi_{\bar{\theta}} > \pi_1 > \pi_0 > 0$. These two conditions states that the screening mechanism has easier to separate the able from the partially disabled and disabled than the latter two from each other - this since the able have the smallest probabilities of being tagged in both instances. Finally it is assumed that $0 < \varphi_1 < \varphi_{\bar{\theta}} < \varphi_0$ which implies that $p_1 - p_{\bar{\theta}} > \pi_{\bar{\theta}} - \pi_1$, which may be interpreted as follows; the screening mechanism will relatively more often tag a disabled as a partially disabled than tag a partially disabled as disabled. Finally it is assumed that probabilities within each type is ranked as follows:

$$
\begin{aligned}
p_1 &> \pi_1 > \varphi_1 \text{ for the disabled} \\
\pi_{\bar{\theta}} &> p_{\bar{\theta}} > \varphi_{\bar{\theta}} \text{ for the partially disabled} \\
\varphi_0 &> \pi_0 > p_0 \text{ for the able}
\end{aligned}
$$

Thus the probability for each type to get the "right" tag is greatest followed by the probability of getting the tag closest to the individuals actual status. The assumption for the partially disabled implicitly says that the partially disabled are closer to being disabled than able.

The imperfect tagging is more complex in the three-type-economy than in the two-type-economy. The number possible of erroneous judgements in the screening process is increased from two to six: able getting T-tag (p_0), able getting P-tag (π_0), partially disabled getting T-tag ($p_{\bar{\theta}}$), partially disabled getting no tag ($\varphi_{\bar{\theta}}$), disabled getting P-tag (π_1), and disabled getting no tag (φ_1).

3.1.2 The utility functions

The individuals utility depend on their work-status, ability and of course the consumption they are allocated. Able and partially disabled have different disutility of working, partially disabled has a greater disutility of working $D_{\bar{\theta}}$ than able individuals, i.e. $D_{\bar{\theta}} > D_0$. Remember, the disutility of working is increasing in degree of disability. The disutility of working, for both types, is reduced by working part time.

Able individuals who are working full time, f, the have the following utility of consumption:

$$U_f^A = u\,(c) - D_0$$

Able individuals may also choose to work part-time, p, instead of full time. If they do so the utility of consumption is given by:

$$U_p^A = u\,(c) - kD_0$$

where k measures the reduction in disutility from working part-time, $0 < k \leq 1$. k can be interpreted as the extent of the part-time work, for example working 50 %.Thus it is assumed that the disutility of working part-time may equal the disutility of working full time, but not equal zero, for able individuals. Utility of consumption for the partially disabled working part-time is:

$$U_p^P = u\,(c) - kD_{\bar{\theta}}$$

For all types the utility of consumption when not working, n, is given by:

$$U_d^i = v\,(c)$$

where $i = A, P, D$. As in the two-type-economy $u(c)$ and $v(c)$ are both concave and increasing. $u'(c)$ goes from ∞ to 0 as c goes from 0 to ∞. This also holds for $v'(c)$. For all $c > 0$ and $\theta < 1$ it is assumed that $u(c) - D_\theta > v(0)$ implying that $u(c) - kD_\theta > v(0)$, that is work and consumption is preferred to no work and no consumption for both the able and the partially disabled. Work is unpleasant, it entails a disutility, thus $u(c) - kD_\theta < v(c)$ implying that $u(c) - D_\theta < v(c)$ for all c. The moral hazard condition from the two-type-model is extended to suite the three-type-model:

$u(\hat{c}) - kD_\theta = v(\tilde{c}) \Rightarrow u'(\hat{c}) < v'(\tilde{c})$[14]

That is, when the utility of working (full time as able or part-time as partially disabled) is equal to the utility of not working and receiving disability benefits then the marginal utility of extra consumption is higher for not working.

[14]implying $u(c) - D_\theta = v(\tilde{c}) \Rightarrow u'(c) < v'(\tilde{c})$ through the concavity of $u(\bullet)$

3.1.3 Characterizing individuals and policy instruments

The individuals are characterized by their work-status, their ability and their tag-status and can thus be divided in to 27 groups $[ability, work, tag]$:

$$[A, f, NT], [A, f, P], [A, f, T], [A, p, NT], [A, p, P], [A, p, T], [A, d, NT],$$
$$[A, d, P], [A, d, T];$$
$$[P, f, NT], [P, f, P], [P, f, T], [P, p, NT], [P, p, P], [P, p, T], [P, d, NT],$$
$$[P, d, P], [P, d, T];$$
$$[D, f, NT], [D, f, P], [D, f, T], [D, p, NT], [D, p, P], [D, p, T], [D, d, NT],$$
$$[D, d, P], [D, d, T]$$

As mentioned above partially disabled cannot work full-time and disabled individuals cannot work at all. Therefore can 9 groups be eliminated from the problem, namely: $[P, f, NT]$, $[P, f, P]$, $[P, f, T]$, $[D, f, NT]$, $[D, f, P]$, $[D, f, T]$, $[D, p, NT]$, $[D, p, P]$ and $[D, p, T]$.

The number of policy instruments are extended to match the increase of possible states in the three-type-economy compared to the two-type economy. The insurance policy now consists of nine consumption allocations:

$$c_f, c_p, c_d, c_f^P, c_p^P, c_d^P, c_f^T, c_p^T, c_d^T$$

where the superscript is the tag-status and the subscript is the work-status.

3.1.4 Optimality and the optimization problem

The insurance administration uses the policy instruments to maximize social welfare. It is assumed that all able individuals work full time and all partially disabled individuals work part-time in the optimal program. The insurance administration optimizes the following social welfare function:

$$\begin{aligned} SWF_4 = \varphi_0 \ell^A (u(c_f) - D_0) + \pi_0 \ell^A (u(c_f^P) - D_0) + p_0 \ell^A (u(c_f^T) - D_0) + \\ + \varphi_{\bar{\theta}} \ell^P (u(c_p) - kD_{\bar{\theta}}) + \pi_{\bar{\theta}} \ell^P (u(c_p^P) - kD_{\bar{\theta}}) + p_{\bar{\theta}} \ell^P (u(c_p^T) - kD_{\bar{\theta}}) + \\ + \varphi_1 \ell^D v(c_d) + \pi_1 \ell^D v\left(c_d^P\right) + p_1 \ell^D v\left(c_d^T\right) \end{aligned} \tag{16}$$

The insurance administration is constrained in its choice of consumption allocations by a resource constraint:

$$\begin{aligned} \varphi_0 \ell^A c_f + \pi_0 \ell^A c_f^P + p_0 \ell^A c_f^T + \varphi_{\bar{\theta}} \ell^P c_p + \pi_{\bar{\theta}} \ell^P c_p^P + p_{\bar{\theta}} \ell^P c_p^T + \\ + \varphi_1 \ell^D c_d + \pi_1 \ell^D c_d^P + p_1 \ell^D c_d^T = M \end{aligned} \tag{17}$$

Where M could be equal to the production in the economy (remember that each able individual has a marginal product equal to one and assume that marginal product of partially disabled is k), i.e. $M = \ell^A * 1 + \ell^D * k$. As in the two-type-economy the resource constraint is assumed to be binding, no reason for the individuals to forgo consumption in the model. The insurance administration also has to ensure that able individuals will work full time and that partially disabled will work part-time in the optimal program. As discussed earlier the able have three options concerning work status irrespective of tag status. To induce them to work full-time in all tag states the following constraints need to be fulfilled:

For able with no tag:

$$\begin{aligned} u(c_f) - D_0 &\geq u(c_p) - kD_0 & (18) \\ u(c_f) - D_0 &\geq v(c_d) & (19) \end{aligned}$$

For able with a P-tag:

$$\begin{aligned} u(c_f^P) - D_0 &\geq u(c_p^P) - kD_0 & (20) \\ u(c_f^P) - D_0 &\geq v(c_d^P) & (21) \end{aligned}$$

For able with a T-tag:

$$\begin{aligned} u(c_f^T) - D_0 &\geq u(c_p^T) - kD_0 & (22) \\ u(c_f^T) - D_0 &\geq v(c_d^T) & (23) \end{aligned}$$

These work constraints ensure that able individuals will choose to work full-time over working part-time and not working at all. The options open to the partially disabled are limited since they cannot work full-time. For the partially disabled the (part-time) work constraints are:

For partially disabled with no tag:

$$u(c_p) - kD_{\bar{\theta}} \geq v(c_d) \quad (24)$$

For partially disabled with a P-tag:

$$u(c_p^P) - kD_{\bar{\theta}} \geq v(c_d^P) \quad (25)$$

For partially disabled with a T-tag:

$$u(c_p^T) - kD_{\bar{\theta}} \geq v(c_d^T) \quad (26)$$

The part-time work constraints ensures that the partially disabled will choose to work part-time over not working at all, irrespective of tag status. Disabled individuals cannot work and thus there is no need for constraints on their behavior. Before setting up the optimization problem it will be helpful to simplify notation, let: $a = \varphi_0 \ell^A, b = \pi_0 \ell^A, c = p_0 \ell^A, d = \varphi_{\bar{\theta}} \ell^P, e = \pi_{\bar{\theta}} \ell^P, f = p_{\bar{\theta}} \ell^P, g = \varphi_1 \ell^D, h = \pi_1 \ell^D, i = p_1 \ell^D$. The optimization problem thus becomes:

$$\max_{c_f, c_p, c_d, c_f^P, c_p^P, c_d^P, c_f^T, c_p^T, c_d^T} SWF_4$$

subject to

$$ac_f + bc_f^P + cc_f^T + d\ell^D c_p + ec_p^P + fc_p^T + gc_d + hc_d^P + ic_d^T = M \tag{r1}$$

$$\begin{aligned} u(c_f) - D_0 &\geq u(c_p) - kD_0 && \text{(r2)} \\ u(c_f) - D_0 &\geq v(c_d) && \text{(r3)} \end{aligned}$$

$$\begin{aligned} u(c_f^P) - D_0 &\geq u(c_p^P) - kD_0 && \text{(r4)} \\ u(c_f^P) - D_0 &\geq v(c_d^P) && \text{(r5)} \end{aligned}$$

$$\begin{aligned} u(c_f^T) - D_0 &\geq u(c_p^T) - kD_0 && \text{(r6)} \\ u(c_f^T) - D_0 &\geq v(c_d^T) && \text{(r7)} \end{aligned}$$

$$u(c_p) - kD_{\bar{\theta}} \geq v(c_d) \tag{r8}$$

$$u(c_p^P) - kD_{\bar{\theta}} \geq v(c_d^P) \tag{r9}$$

$$u(c_p^T) - kD_{\bar{\theta}} \geq v(c_d^T) \tag{r10}$$

It is easily seen that all work constraints are not slack or binding at the same time. Moreover, it is obvious, by implication, that constraint r3, r5 and r7 are fulfilled with strict inequality if the other constraints are satisfied with equality. It can also be shown that the last case is a solution candidate to the optimization problem (see appendix for discussion of the complementary slackness conditions). Thus r3, r5 and r7 are redundant and other constraints are binding in the optimal program.

Now, let λ_j for $j = 1, ..., 10$ be the Lagrange multipliers of the optimization problem and in the optimal program is shown that $\lambda_j = 0$ for $j = 3, 5, 7$ and $\lambda_i > 0$, for $i = 1, 2, 4, 6, 8, 9, 10$. Then the first order conditions are the following:

$\frac{\partial \mathcal{L}}{c_f} = au'(c_f) - \lambda_1 a + \lambda_2 u'(c_f) = 0$

$\frac{\partial \mathcal{L}}{c_p} = du'(c_p) - \lambda_1 d - \lambda_2 u'(c_p) + \lambda_8 u'(c_p) = 0$

$\frac{\partial \mathcal{L}}{c_d} = gv'(c_d) - \lambda_1 g - \lambda_8 v'(c_d) = 0$

$\frac{\partial \mathcal{L}}{c_f^P} = bu'(c_f^P) - \lambda_1 b + \lambda_4 u'(c_f^P) = 0$

$\frac{\partial \mathcal{L}}{c_p^P} = eu'(c_p^P) - \lambda_1 e - \lambda_4 u'(c_p^P) + \lambda_9 u'(c_p^P) = 0$

$\frac{\partial \mathcal{L}}{c_d^P} = hv'(c_d^P) - \lambda_1 h - \lambda_9 v'(c_d^P) = 0$

$\frac{\partial \mathcal{L}}{c_f^T} = cu'(c_f^T) - \lambda_1 c + \lambda_6 u'(c_f^T) = 0$

$\frac{\partial \mathcal{L}}{c_p^T} = fu'(c_p^T) - \lambda_1 f - \lambda_6 u'(c_p^T) + \lambda_{10} u'(c_p^T) = 0$

$\frac{\partial \mathcal{L}}{c_d^T} = iv'(c_d^T) - \lambda_1 i - \lambda_{10} v'(c_d^T) = 0$

3.1.5 Redistribution principle and consumption allocations

The optimization problem can be solved when r3, r5, r7 are slack. In this case an inspection of the work constraints gives the following ranking of consumption allocations within each tag-status group:

$c_f > c_p > c_d$, for not tagged
$c_f^P > c_p^P > c_d^P$, for P-tagged
$c_f^T > c_p^T > c_d^T$, for T-tagged

This is not surprising given the setup of the model. However, the relation between consumption allocations across tag-status is less obvious and thus more interesting . As in the two-type-economy a redistribution principle can be elaborated from the first order conditions. The redistribution principle describes the redistribution between untagged, P-tagged and T-tagged individuals, and states that the weighted average of the inverse marginal utilities is equalized across tag-status.

$$\begin{aligned}
\omega_a \frac{1}{u'(c_f)} + \omega_g \frac{1}{v'(c_d)} + (1-\omega_a-\omega_g)\frac{1}{u'(c_p)} &= \\
= \omega_b \frac{1}{u'(c_f^P)} + \omega_h \frac{1}{v'(c_d^P)} + (1-\omega_b-\omega_h)\frac{1}{u'(c_p^P)} &= \\
= \omega_c \frac{1}{u'(c_f^T)} + \omega_i \frac{1}{v'(c_d^T)} + (1-\omega_c-\omega_i)\frac{1}{u'(c_p^T)} &
\end{aligned}$$

where

$$\begin{aligned}
\omega_a &= \frac{\varphi_0 \ell^A}{\varphi_0 \ell^A + \varphi_{\bar{\theta}} \ell^P + \varphi_1 \ell^D}, \omega_g = \frac{\varphi_1 \ell^D}{\varphi_0 \ell^A + \varphi_{\bar{\theta}} \ell^P + \varphi_1 \ell^D}, \\
\omega_b &= \frac{\pi_0 \ell^A}{\pi_0 \ell^A + \pi_{\bar{\theta}} \ell^P + \pi_1 \ell^D}, \omega_h = \frac{\pi_1 \ell^D}{\pi_0 \ell^A + \pi_{\bar{\theta}} \ell^P + \pi_1 \ell^D}, \\
\omega_c &= \frac{p_0 \ell^A}{p_0 \ell^A + p_{\bar{\theta}} \ell^P + p_1 \ell^D}, \omega_i = \frac{p_1 \ell^D}{p_0 \ell^A + p_{\bar{\theta}} \ell^P + p_1 \ell^D}
\end{aligned}$$

Ranking of consumption allocations across tag-status Could it be optimal to ignore the individuals tag-status and give all workers the same consumption allocation? Formally this is represented by setting $c_f = c_f^P = c_f^T$, implying, through the work constraints, that $c_p = c_p^P = c_p^T$ this in turn implies $c_d = c_d^P = c_d^T$. Under these conditions the redistribution principle becomes[15]:

$$\omega_a \frac{1}{u'(c_f)} + \omega_g \frac{1}{v'(c_d)} + (1-\omega_a-\omega_g)\frac{1}{u'(c_p)} > \tag{A}$$

$$\omega_b \frac{1}{u'(c_f^P)} + \omega_h \frac{1}{v'(c_d^P)} + (1-\omega_b-\omega_h)\frac{1}{u'(c_p^P)} > \tag{B}$$

$$\omega_c \frac{1}{u'(c_f^T)} + \omega_i \frac{1}{v'(c_d^T)} + (1-\omega_c-\omega_i)\frac{1}{u'(c_p^T)} \tag{C}$$

[15] see appendix

However, the redistribution principle is fulfilled with equality in the optimal program, thus it cannot be optimal to ignore the information that the tag-status gives. Instead note that A is a increasing function in c_f since $\frac{1}{u'(c_f)}$ is increasing in c_f , and that B (C) is increasing in c_f^P (c_f^T) by the same reasoning. Therefore c_f needs to lowered compared to both c_f^P and c_f^T, and c_f^P needs lowered compared to c_f^T to achieve equality, implying that $c_f < c_f^P < c_f^T$ in the optimal program. Now the binding work constraints yield that $c_p < c_p^P$ $< c_p^T$ and $c_d < c_d^P < c_d^T$. Notably the information contained in the screening process is used in optimum. That is, individuals with the same work-status but different tag-status will receive different consumption allocations. Thus, the ranking between tag-status groups for individuals with the same work status is:

$c_f^T > c_f^P > c_f$, for full-time workers
$c_p^T > c_p^P > c_p$, for part-time workers
$c_d^T > c_d^P > c_d$, for non-workers

Combining this with the ranking within each tag-status groups a partial ranking can be achieved. Obviously c_f^T is the greatest consumption allocation and c_d the smallest, in line with the results from the two-type model, *table 1* presents the partial ranking.

Table 1: The partial ranking

	c_f	c_p	c_d	c_f^P	c_p^P	c_d^P	c_f^T	c_p^T	c_d^T
c_f	$=$	$>$	$>$	$<$	$\lesseqgtr$	$\lesseqgtr$	$<$	$\lesseqgtr$	$\lesseqgtr$
c_p	$<$	$=$	$>$	$<$	$<$	$\lesseqgtr$	$<$	$<$	$\lesseqgtr$
c_d	$<$	$<$	$=$	$<$	$<$	$<$	$<$	$<$	$<$
c_f^P	$>$	$>$	$>$	$=$	$>$	$>$	$<$	$\lesseqgtr$	$\lesseqgtr$
c_p^P	$\lesseqgtr$	$>$	$>$	$<$	$=$	$>$	$<$	$<$	$\lesseqgtr$
c_d^P	$\lesseqgtr$	$\lesseqgtr$	$>$	$<$	$<$	$=$	$<$	$<$	$<$
c_f^T	$>$	$>$	$>$	$>$	$>$	$>$	$=$	$>$	$>$
c_p^T	$\lesseqgtr$	$>$	$>$	$\lesseqgtr$	$>$	$>$	$<$	$=$	$>$
c_d^T	$\lesseqgtr$	$\lesseqgtr$	$>$	$\lesseqgtr$	$\lesseqgtr$	$>$	$<$	$<$	$=$

Interpretation: The inequality signs show how the cons. allocations in the vertical column is related to the cons. allocation in the top row.

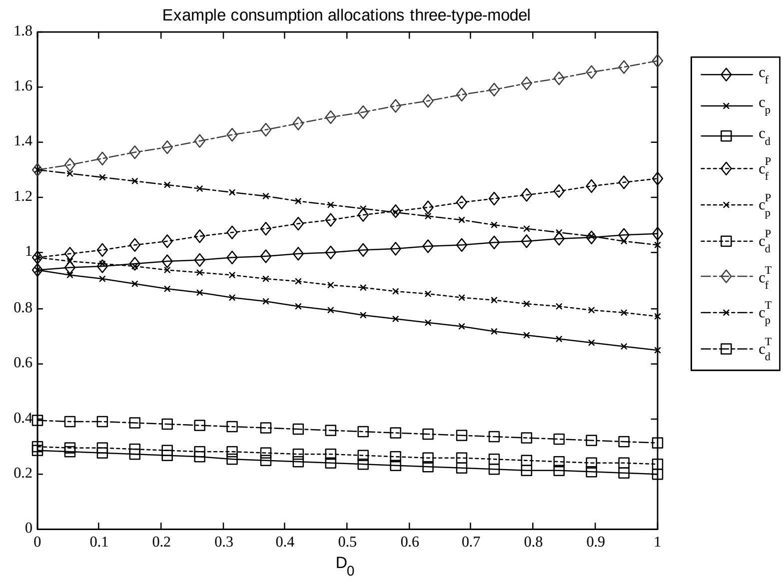
The assumption $v'(c) = u'(c)$ all c gives us a complete ranking in the two-type case, imposing the same assumption for the three-type-model ensures that $c_f > c_d^T$, $c_f^P > c_d^T$ and that $c_f > c_f^P$. This follows from the fact that the smallest element of a weighted average cannot exceed or equal the greatest element of another weighted average if equality is to hold.[16] Thus the assumption improves

[16] For positive weights and as in this case $0 < \omega_j < 1$, $j = a, b, c, d, e, f, g, h, i$

the ranking but does not make it complete for any weights between zero and one. The ranking of consumption allocations given to part-time workers compared to full-time workers of different tag status is undetermined in this case and will depend on the disutility of working for the groups as can be seen in the logarithmic example below. The ranking between part-time workers and non-workers of different tag-status is also undetermined and will depend (similar to the two-type model) on marginal utility of consumption - if non-workers have a very high marginal utility of consumption it might be socially optimal to give them a higher consumption than part-time workers (from a utilitarian point of view).

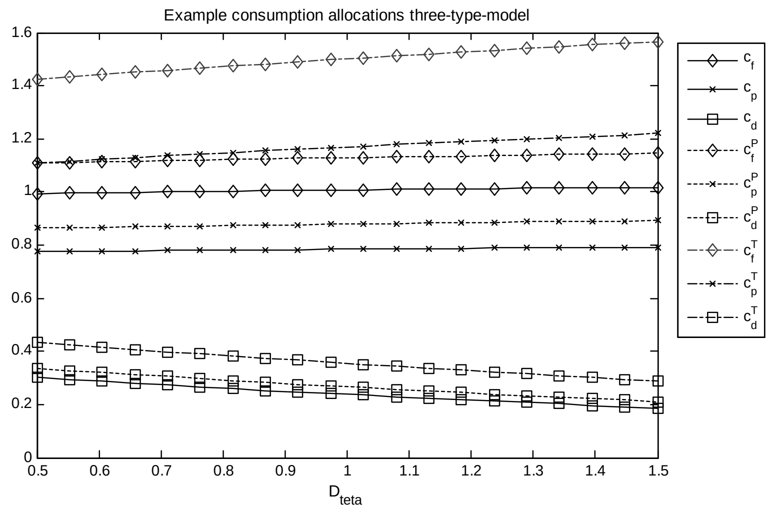
3.2 Logarithmic example

To illustrate the three-type model it is once again assumed that $u(c) = \ln c$ and $v(c) = \ln 2c$. The somewhat rigorous assumptions about the screening process limits the opportunities for comparisons but some tendencies can be found in the examples below. Especially the effect of work disutility on the consumption allocations can be easily depicted. Figure 11 and figure 12 plots the consumption allocations over D_0 and D_θ respectively. Both under the following assumptions concerning the screening process and the population weights: $p_0 = 0.1$, $p_1 = 0.7$, $p_\theta = 0.2$, $\pi_0 = 0.2$, $\pi_1 = 0.2$, $\pi_\theta = 0.7$, $\ell^P = 0.1$, $\ell^D = 0.1$, $\ell^A = 0.8$. Furthermore it is assumed that $k = 0.5$, the part-time workers work 50 percent. Much can be said about these assumption, but most interesting for the examples presented here is the assumption about k. k affects the consumption allocations for part-time workers and full-time worker since it affects the work disutility for part-time workers. A low k implies that the income for full-time workers (e.g. c_f) will be greater than the income for part-time workers (e.g. c_p^P) for lower work disutilities than would be the case if k was high. That is, when k is low part-time workers will need less compensation for their work disutility and the difference, in consumption, compared to full-time workers have to increase faster with growing disutility for full-time workers - to ensure that full-time work is a attractive alternative. Moreover, as k approaches 1 for given work disutilities the difference in income between full-time and part-time workers naturally approaches zero.



(Figure 11)

Given the assumptions above (e.g. $k = 0.5$) and the additional assumption of $D_\theta = 1$ figure 11 plots the consumption allocations over the work disutility for able individuals (i.e full-time workers). Notably the consumption allocations for non-workers are significantly lower than for both part-time and full-time workers and thus has the specification of the example improved the ranking of consumption allocations. In the general case and also after the additional assumption that $v'(c) = u'(c)$ all c, the ranking between part-time workers and non-workers of different tag-status could not be determined. However, the ranking between part-time workers and full-workers of different tag-status is not absolute even in this example. The income for T-tagged part-time workers (c_p^T) is, for example, greater than income for both P-tagged and untagged full-time workers over a wide range of D_0. However, as D_0 grows difference in income between part-time and full-time workers within the same tag-status grows - to ensure that full-time work is an attractive alternative. This effect can be seen for all tagging-groups in figure 11.



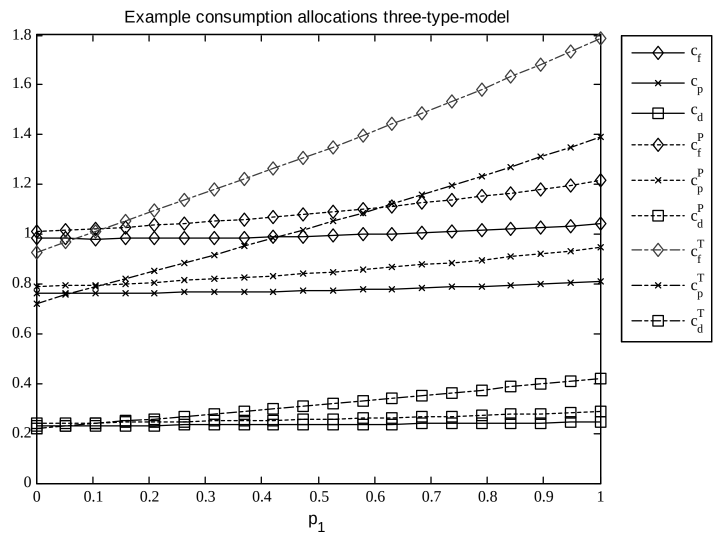
(Figure 12)

Figure 12 plots the consumption allocations, given the assumptions above and $D_0 = 0.5$, over $0.5 \leq D_\theta \leq 1.5$ since it is assumed that $D_0 < D_\theta$ in the three-type-model. Interestingly the income for part-time workers is kept quite constant over the range of D_θ and, similar to the two-type-model[17], the consumption allocations for non-workers is suppressed to make working part-time an attractive alternative. Moreover this also implies that the income for full-time workers is quite constant over the disutility of working for partially disabled.

These two examples provides some intuition to three-type-model, but their main propose is to show how the ranking of consumption allocations behave in different settings. The assumptions about screening process limits amount of reasonable examples, however figure 13 plots the consumption allocation over p_1 and produces a similar picture to figure 2 (p. 9). That is, the coverage of the income loss for the disabled is improved as the probability that disabled receive a disability tag becomes higher. A similar effect can be found for partially disabled as the probability for assigning a P-tag to partially disabled increases

[17] e.g figure 3

(not presented here).



(Figure 13)

It is apparent that similar forces are a work in the three-type-model as in the two-type-model, but the results are less straight forward and obvious - a natural effect of introducing a third type. The three illustrations above underlines this finding.

4 Concluding remarks

This paper takes Parsons' 1996 model for disability insurance under imperfect tagging and extends in several directions. The most substantial extension is the introduction of a third type: the partially disabled. This is done to investigate and illustrate a optimal disability insurance allowing for several degrees of disability under imperfect tagging. The Swedish disability insurance is a real life example of a disability insurance with several degrees of disability. The Swedish disability insurance has four degrees of disability (thus five types of individuals in the economy): 25, 50, 75 and 100 percent. The analysis of a three-type-model lends intuition to a discussion of a five type economy - introducing additional types to the formal analysis is straight forward but clouds the analysis more than it clarifies the intuition. So what conclusion can be drawn from the three type model? They are essentially the same as the ones in the two-type-model; with imperfect tagging it is optimal to reward individuals working in line with their ability and that this leaves room for a improved replacement rates for the targeted groups. Moreover, both the three-type and the two-type-model underlines the need to improve the screening processes, when the classification errors become smaller the disability insurance become more efficient (in terms of replacement rates). This realization provides a good foundation for further research in the area, e.g. an analysis of how the screening probabilities are affected by stricter or more lenient rules and how this is manifested in the optimal program. Notably, in most countries rules are more common as policy instruments than monetary incentives.

Besides looking at the three-type-economy this paper look at three different solutions to the two-type-model. This is done since policy-makers may have other goals and priorities besides disability insurance that affect the optimal solution. Policy-makers may have wowed to equalize salaries for all workers (e.g. within the same industry) or be reluctant to admit to type-II errors in the screening process. The comparison of social welfare in the different cases show that it is never in the best interest of society, *ceteris paribus*, to ignore the information given in the screening process by equalizing salaries for workers. The targeted groups in the disability insurance will have a low replacement rate in this setting. The comparison also show, quite intuitively, that when the type-II error is small it is optimal to ignore it and act as if it did not exist. However, if the screening process deteriorates and the type-II error thereby increases this will put a strain on the whole economy and eventually will the disability insurance collapse.

The two-type-analysis contributes to Parsons' analysis in two ways; first the analysis is generalized to a setting allowing for different utility over work status - in this case such that non-workers have greater utility from consumption than workers. Second, it provides additional intuition about the optimality of the four-price-model (here called the original model or case 1). It is obvious that four-price model is optimal when the e.g. type-II errors are substantial and empirical research points in that direction. However, much empirical research remains to be done and, as is underlined by Parsons (1996), the model might be difficult to implement. Future research might, besides analyzing the effect of stricter rules, analyze whether it is better to have a disability insurance allowing for several degrees of disability or a dichotomous disability insurance.

References

[1] Akerlof G A, 1978, "The Economics of Tagging as Applied to Optimal Income Tax, Welfare Programs, and Manpower Planning, American Economic Review 68, 8-19

[2] Benitez-Silva H, Buchinsky M & Rust J, 2004, "How large are the classification errors in the social security disability award process", NBER Working paper 10219, http://www.nber.org/papers/w10219

[3] Diamond P A & Mirrlees J A, 1978, "A Model of Social Insurance with Variable Retirement", Journal of Public Economics 10, 295-336

[4] Diamond P A & Sheshinski E, 1995, "Economic Aspects of Optimal Disability Benefits", Journal of Public Economics 57, 1-23

[5] Nagi S Z, 1969, “Disability and rehabilitation: Legal, clinical and self-concepts and measurement”, Ohio State University Press

[6] Parsons D O, 1996, “Imperfect ‘tagging’ in social insurance programs”, Journal of Public Economics 62, 183-207

[7] Salanié B, 2002, "Optimal Demogrants with Imperfect Tagging", Economic Letters 75, 319-324

Appendix

Complementary slackness two-type-model

The complementary slackness conditions requires that;

$\lambda_{ii} \geq 0$, $\lambda_{ii}\left[v\left(c_d\right)-u\left(c_f\right)+D_0\right]=0$ that is $\lambda_{ii}=0$ if $v\left(c_d\right)-u\left(c_f\right)+D_0<0$

$\lambda_{iii} \geq 0$, $\lambda_{iii}\left[v\left(c_d^T\right)-u\left(c_f^T\right)+D_0\right]=0$ that is $\lambda_{iii}=0$ if $v\left(c_d^T\right)-u\left(c_f^T\right)+D_0<0$

Imperfect tagging - original model

1. $\lambda_{ii}=\lambda_{iii}=0$, if remaining lagrange multipliers (after concluding that $\lambda_i>0$) both equal zero then the work incentive constraints are fulfilled with inequality: $v\left(c_d\right)-u\left(c_f\right)+D_0<0$, $v\left(c_d^T\right)-u\left(c_f^T\right)+D_0<0$. Since $u\left(c\right)-D_0=v(\tilde{c}) \Rightarrow u'\left(c\right)<v'(\tilde{c})$ and $u(c)-D_0<v(c)$, *all c* these inequalities imply that $u'\left(c_f\right)<v'(c_d)$, $u'\left(c_f^T\right)<v'(c_d^T)$ when $\lambda_{ii}=\lambda_{iii}=0$. However an investigation of the first order conditions reveals that $u'\left(c_f\right)=v'(c_d)=u'\left(c_f^T\right)=v'(c_d^T)$ when $\lambda_{ii}=\lambda_{iii}=0$, thus there is a contradiction in this case. No solution candidate to the maximization problem in this case.

2. $\lambda_{ii}>0, \lambda_{iii}=0$ implies that $v\left(c_d\right)-u\left(c_f\right)+D_0=0$, $v\left(c_d^T\right)-u\left(c_f^T\right)+D_0<0$. Both constraints imply that the marginal utility when working is lower than the marginal utility when not working, i.e. $u'\left(c_f\right)<v'(c_d)$, $u'\left(c_f^T\right)<v'(c_d^T)$. The first order conditions, however, imply that $u'\left(c_f\right)<v'(c_d), u'\left(c_f^T\right)=v'(c_d^T)$, again there is a contradiction in this case for the tagged individuals. No solution candidate to the maximization problem in this case.

3. $\lambda_{ii}=0, \lambda_{iii}>0$ implies that $v\left(c_d\right)-u\left(c_f\right)+D_0<0$, $v\left(c_d^T\right)-u\left(c_f^T\right)+D_0=0$. Once again this implies $u'\left(c_f\right)<v'(c_d)$, $u'\left(c_f^T\right)<v'(c_d^T)$. In this case the first order conditions imply that $u'\left(c_f\right)=v'(c_d), u'\left(c_f^T\right)<v'(c_d^T)$, thus we have a contradiction in this case as well. No solution candidate to the maximization problem in this case.

4. $\lambda_{ii}>0, \lambda_{iii}>0$ implies that $v\left(c_d\right)-u\left(c_f\right)+D_0=0$, $v\left(c_d^T\right)-u\left(c_f^T\right)+D_0=0$. This in turn implies that $u'\left(c_f\right)<v'(c_d)$, $u'\left(c_f^T\right)<v'(c_d^T)$ and in this case the first order conditions also result in these inequalities, thus there is solution candidate to the maximization problem in this case.

Imperfect tagging - equalizing salaries

1. $\lambda_{ii} = \lambda_{iii} = 0$ and $\lambda_i > 0 \Rightarrow u'(c_f) = v'(c_d) = v'(c_d^T)$ but this leads to the contradiction with similar reasoning as in case 1 in the original model (above). Marginal utilities cannot be equalized when the utility when working is strictly greater than the utility when not working.

2. $\lambda_{ii} > 0, \lambda_{iii} = 0$ and $\lambda_i > 0$, give the following (from the first order conditions):
$\ell^A u'(c_f) - \lambda_i \ell^A + \lambda_{ii} u'(c_f) = 0,$
$(1-p_1)(1-\ell^A)v'(c_d) - \lambda_i(1-p_1)(1-\ell^A) - \lambda_{ii} v'(c_d) = 0 \Rightarrow v'(c_d)\left(1 - \frac{\lambda_{ii}}{(1-p_1)(1-\ell^A)}\right) = \lambda_i$
$p_1(1-\ell^A)v'\left(c_d^T\right) - \lambda_i p_1(1-\ell^A) = 0 \Rightarrow v'\left(c_d^T\right) = \lambda_i$ and $v'(c_d)\left(1 - \frac{\lambda_{ii}}{(1-p_1)(1-\ell^A)}\right) = v'\left(c_d^T\right)$ implies that
$v'(c_d) > v'\left(c_d^T\right) \Rightarrow c_d < c_d^T$. However the work constraints in this case are $u(c_f) - D_0 > v\left(c_d^T\right)$ and $u(c_f) - D_0 = v(c_d)$ implying that $c_d > c_d^T$ - a contradiction.

3. $\lambda_{ii} = 0, \lambda_{iii} > 0$ and $\lambda_i > 0$, the mirror image of the case 2 above, and thus leading to a contradiction. The work constraints imply that $c_d < c_d^T$ while the first order conditions imply that $c_d > c_d^T$ - an obvious contradiction.

4. $\lambda_{ii} > 0, \lambda_{iii} > 0$ and $\lambda_i > 0$, giving the work constraints $u(c_f) - D_0 = v(c_d) = v\left(c_d^T\right)$ implying that $u'(c_f) < v'(c_d) = v'\left(c_d^T\right)$ since the moral hazard condition is satisfied. The first order conditions becomes:
$u'(c_f)\left(1 + \frac{\lambda_{ii}+\lambda_{iii}}{\ell^A}\right) = \lambda_i$
$v'(c_d)\left(1 - \frac{\lambda_{ii}}{(1-p_1)(1-\ell^A)}\right) = \lambda_i$
$v'\left(c_d^T\right)\left(1 - \frac{\lambda_{iii}}{p_1(1-\ell^A)}\right) = \lambda_i$
Given the reasoning in 1-3, i.e. no solution candidates in these cases, a solution can be found when $\frac{\lambda_{iii}}{p_1(1-\ell^A)} = \frac{\lambda_{ii}}{(1-p_1)(1-\ell^A)}$ implying that $u'(c_f) < v'(c_d) = v'\left(c_d^T\right)$. Thus, there is no contradiction in this case. Case 4 where all constraints are binding is a solution candidate.

Complementary slackness three-type-model

1. Assume that all work constraints are slack i.e. $\lambda_i = 0$ for $i = 2, 3, ..., 10$, this e.g. implies that $u'(c_f) = u'(c_p)$ and $u(c_f) - D_0 > u(c_p) - kD_0$. It is known that $D_0 \geq kD_0$ for all $D_0 > 0$ implying that $u(c_f) > u(c_p)$ which in turn implies that $c_f > c_p$ and thus $u'(c_f) < u'(c_p)$ - contradicting the assumption that $\lambda_i = 0$ for $i = 2, 3, ..., 10$.

2. Assume that all work constraints are binding in the optimal program, i.e. $\lambda_i > 0$ for $i = 2, 3, ..., 10$. Constraint r2, r3 and r8 cannot be fulfilled with equality at the same time. If this were the case $u(c_f) - D_0 = u(c_p) - kD_0$, $u(c_f) - D_0 = v(c_d)$, $u(c_p) - kD_{\bar{\theta}} = v(c_d)$ implying that $u(c_p) - kD_{\bar{\theta}} =$

$u(c_p) - kD_0$. Since its assumed that $D_{\bar{\theta}} > D_0$ it's obvious that $kD_{\bar{\theta}} > kD_0$ implying that $u(c_p) - kD_{\bar{\theta}} < u(c_p) - kD_0$ contradicting the assumption that constraints r2,r3 and r8 are fulfilled with equality at the same time. The equivalent reasoning holds for constraint r4, r5 and r9 and constraint r6, r7 and r10.

3. Assume that constraint r3, r5 and r7 are slack and the other constraints are binding, i.e.$\lambda_i > 0$, for $i = 1, 2, 4, 6, 8, 9, 10$ and $\lambda_j = 0$ for $j = 3, 5, 7$. Why would these constraints be slack? Consider constraint r3. If $u(c_p) - kD_{\bar{\theta}} \geq v(c_d)$ (r8) then $u(c_f) - D_0 > v(c_d)$ (r3) since it's required that $u(c_f) - D_0 \geq u(c_p) - kD_0$ and it is known that $u(c_p) - kD_{\bar{\theta}} < u(c_p) - kD_0$. With the equivalent reasoning it is obvious that constraints r5 and r7 also are slack. Does the other constraints bind under these conditions? Assume that $\lambda_2 = 0$ this implies (through the FOC:s) that $u'(c_f) > u'(c_p)$ and thus that $u(c_f) < u(c_p) \Leftrightarrow u(c_f) - D_0 < u(c_p) - kD_0$ since $D_0 \geq kD_0$ but $\lambda_2 = 0$ implies that $u(c_f) - D_0 > u(c_p) - kD_0$ - a contradiction arises. Thus it must be the case that $\lambda_2 > 0$, i.e. constraint r2 binds. The same reasoning can be applied to constraints r4 and r6. Now assume that $\lambda_8 = 0$ then the FOC:s give that $u'(c_p) > v'(c_d)$ when $u(c_p) - kD_{\bar{\theta}} > v(c_d)$, but the moral hazard condition gives that $u'(c_p) < v'(c_d)$ when $u(c_p) - kD_{\bar{\theta}} = v(c_d)$, and $v(c) > u(c)$ for all c, thus it must be the case that $u'(c_p) < v'(c_d)$ when $u(c_p) - kD_{\bar{\theta}} > v(c_d)$ - once again a contradiction is reached. Thus $\lambda_8 > 0$, the same reasoning holds for r9 and r10. Thus, a solution candidate for the optimization problem is found.

Redistribution principle and consumption allocations, three-type-model

For increased comparability assume that $\ell^A > \ell^P = \ell^D = \ell$. First consider the weights in the redistribution principle: $\omega_a, \omega_b, \omega_c, \omega_g, \omega_h, \omega_i, \omega_d = (1 - \omega_a - \omega_g)$, $\omega_e = (1 - \omega_b - \omega_h)$, $\omega_f = (1 - \omega_c - \omega_i)$.

$$
\begin{aligned}
\omega_a &= \frac{\varphi_0 \ell^A}{\varphi_0 \ell^A + \varphi_{\bar{\theta}} \ell + \varphi_1 \ell}, \omega_g = \frac{\varphi_1 \ell}{\varphi_0 \ell^A + \varphi_{\bar{\theta}} \ell + \varphi_1 \ell}, \\
\omega_b &= \frac{\pi_0 \ell^A}{\pi_0 \ell^A + \pi_{\bar{\theta}} \ell + \pi_1 \ell}, \omega_h = \frac{\pi_1 \ell}{\pi_0 \ell^A + \pi_{\bar{\theta}} \ell + \pi_1 \ell}, \\
\omega_c &= \frac{p_0 \ell^A}{p_0 \ell^A + p_{\bar{\theta}} \ell + p_1 \ell}, \omega_i = \frac{p_1 \ell}{p_0 \ell^A + p_{\bar{\theta}} \ell + p_1 \ell}
\end{aligned}
$$

All weights are positive and can be partially ranked as follows: $\omega_a > \omega_b > \omega_c$, $\omega_i > \omega_h > \omega_g$, $\omega_e > \omega_f > \omega_d$, and furthermore it can be shown that $\omega_a > \omega_d > \omega_g$, $\omega_e > \omega_h$, $\omega_i > \omega_f$. This given the assumption about the screening process.

It is known that $c_f > c_p > c_d$, $c_f^P > c_p^P > c_d^P$, $c_f^T > c_p^T > c_d^T$. Now assume that $c_f = c_f^P = c_f^T$ implying that $c_p = c_p^P = c_p^T$ which in turn implies that $c_d = c_d^P = c_d^T$. Thus$\frac{1}{u'(c_f)} = \frac{1}{u'(c_f^P)} = \frac{1}{u'(c_f^T)} \equiv \alpha$, $\frac{1}{u'(c_p)} = \frac{1}{u'(c_p^P)} = \frac{1}{u'(c_p^T)} \equiv \beta$

and $\frac{1}{v'(c_d)} = \frac{1}{v'(c_d^P)} = \frac{1}{v'(c_d^T)} \equiv \gamma$. The redistribution principle can now be written (note that the relation between the groups is not specified):

$$\omega_a\alpha + \omega_g\gamma + \omega_d\beta \lesseqqgtr \omega_b\alpha + \omega_h\gamma + \omega_e\beta \lesseqqgtr \omega_c\alpha + \omega_i\gamma + \omega_f\beta$$

In the optimal program the redistribution principle is fulfilled with equality, is this the case when the consumption allocations are equalized over work-status? First note that $\alpha > \beta > \gamma$, and that the left-hand side can be rewritten as $\omega_a\alpha+\omega_g\gamma+\omega_d\beta = (1-\omega_g-\omega_d)\,\alpha+\omega_g\gamma+\omega_d\beta = \alpha-\alpha\omega_g-\alpha\omega_d+\omega_g\gamma+\omega_d\beta = \alpha+\omega_g\,(\gamma-\alpha)+\omega_d\,(\beta-\alpha)$. The middle expression of the redistribution principle is $\omega_b\alpha+\omega_h\gamma+\omega_e\beta = (1-\omega_h-\omega_e)\,\alpha+\omega_h\gamma+\omega_e\beta = \alpha-\alpha\omega_h-\alpha\omega_e+\omega_h\gamma+\omega_e\beta = \alpha+\omega_h(\gamma-\alpha)+\omega_e\,(\beta-\alpha)$. Finally the right-hand side is $\omega_c\alpha+\omega_i\gamma+\omega_f\beta = (1-\omega_i-\omega_f)\,\alpha+\omega_i\gamma+\omega_f\beta = \alpha-\alpha\omega_i-\alpha\omega_f+\omega_i\gamma+\omega_f\beta = \alpha+\omega_i\,(\gamma-\alpha)+\omega_f\,(\beta-\alpha)$. Thus; $\alpha+\omega_g\,(\gamma-\alpha)+\omega_d\,(\beta-\alpha) \lesseqqgtr \alpha+\omega_h(\gamma-\alpha)+\omega_e\,(\beta-\alpha) \lesseqqgtr \alpha+\omega_i\,(\gamma-\alpha)+\omega_f\,(\beta-\alpha)$, implying $\omega_g\,(\gamma-\alpha)+\omega_d\,(\beta-\alpha) \lesseqqgtr \omega_h(\gamma-\alpha)+\omega_e\,(\beta-\alpha) \lesseqqgtr \omega_i\,(\gamma-\alpha)+\omega_f\,(\beta-\alpha)$ and $\omega_i > \omega_h > \omega_g \Rightarrow \omega_g\,(\gamma-\alpha) > \omega_h(\gamma-\alpha) > \omega_i\,(\gamma-\alpha)$ since $(\gamma-\alpha)$ is a negative number, furthermore $\omega_e > \omega_f > \omega_d \Rightarrow \omega_d\,(\beta-\alpha) > \omega_f\,(\beta-\alpha) > \omega_e\,(\beta-\alpha)$ since $(\beta-\alpha)$ is a negative number. Thus it can be concluded that $\omega_a\alpha+\omega_g\gamma+\omega_d\beta > \omega_b\alpha+\omega_h\gamma+\omega_e\beta$ and $\omega_a\alpha+\omega_g\gamma+\omega_d\beta > \omega_c\alpha+\omega_i\gamma+\omega_f\beta$, what about relation between the middle expression and the right-hand side expression?

Rewrite $\omega_b\alpha+\omega_h\gamma+\omega_e\beta$ as $\omega_b\alpha+\omega_h\gamma+(1-\omega_b-\omega_h)\,\beta = \beta+\omega_b(\alpha-\beta)+\omega_h\,(\gamma-\beta)$ and $\omega_c\alpha+\omega_i\gamma+\omega_f\beta$ as $\omega_c\alpha+\omega_i\gamma+(1-\omega_c-\omega_i)\,\beta = \beta+\omega_c(\alpha-\beta)+\omega_i\,(\gamma-\beta)$. Thus comparing the middle expression and the right-hand side it is found that $\omega_b(\alpha-\beta)+\omega_h\,(\gamma-\beta) > \omega_c(\alpha-\beta)+\omega_i\,(\gamma-\beta)$ since $\omega_b(\alpha-\beta) > \omega_c(\alpha-\beta)$ ($\omega_b > \omega_c$ and $(\alpha-\beta)$ is positive) and $\omega_h\,(\gamma-\beta) > \omega_i\,(\gamma-\beta)$ ($\omega_i > \omega_h$ and $(\gamma-\beta)$ is negative). Thus

$$\omega_a\alpha + \omega_g\gamma + \omega_d\beta > \omega_b\alpha + \omega_h\gamma + \omega_e\beta > \omega_c\alpha + \omega_i\gamma + \omega_f\beta$$

when the consumption allocations are the same for the same work status. This situation is not optimal since the redistribution principle is not fulfilled with equality.