Welfare effects of deterrence-motivated activation policy

Martin Rasmussen

Welfare systems and policies
Working Paper 6:2004

Working Paper
The Danish National Institute of Social Research

Welfare effects of deterrence-motivated activation policy

Martin Rasmussen [1]

The Danish National Institute of Social Research, Herluf Trolles Gade 11, DK-1052 Copenhagen K, Denmark

Abstract

We investigate whether activation policy is part of optimal policy of a benevolent government, when the motivation for introducing activation is to deter some people from collecting benefits. The government offers a pure benefit programme and an activation programme, and individuals self-select into programmes. Individuals differ with respect to disutility and wage. Activation programmes are relatively costly and favour individuals who are relatively well off. Hence, for activation policy to used, labour supply effects have to be relatively small. We discuss how labour supply effects depend on the distribution of wage and disutility, and discuss previous literature in this light.

Theme: Labour market policy

Keywords: Workfare, Active labour market policy, Activation policy

JEL-Code: J65, J68

---

[1]Email address: mar@sfi.dk., Tel.: +45 33 48 09 10. Fax.: +45 33 48 08 33.

1. Introduction

We investigate theoretically whether use of 'activation policy' (i.e. the policy of making social benefits conditioned on work-like efforts – also phrased 'workfare' or 'active labour market policy') can be part of an optimal social policy of a benevolent government. We limit the discussion to the 'deterrence effect' (or 'motivation effect') of activation policy, i.e. the effect caused by some people choosing work rather than public benefits because activation programmes implies disutility in the same way as work does. We use a principal-agent approach with wage and disutility as individuals' private-knowledge parameters. The government designs a pure benefit programme and an activation programme, and individuals self-select into social programmes or into ordinary work.

Previous theoretical contributions differ somewhat in their conclusions about the social welfare consequences of the deterrence effects of activation. In this paper, we hope to increase the understanding of the consequences of especially the assumptions made on distributions of private-knowledge parameters to the conclusions about whether or not activation is part of optimal policy.

One way to motivate the use of activation policy to deter people from collecting social benefits might be as follows:

A work requirement imposes a utility loss to participants in activation programmes. If the work requirement is relatively harmless to the potential participants in activation programmes, but not to potential non-recipients, then the work requirement makes the social benefits less attractive especially to the second group. The government may therefore be able to raise benefits without unintended labour supply effects (e.g. Thustrup Kreiner and Tranæs (2003)) or to reduce the costs of the social programmes (Besley and Coate (1992, 1995)). Of importance to this deterrence argument is that the work requirement has different effect on different groups.

In the literature reviewed below, the models used differ especially with respect to the assumptions about the distributions of characteristics (i.e. the way that individuals differ from each other with respect to productivity and/or disutility of work) and with respect to assumptions about the government's criterion function.

Besley and Coate (1992) consider a government that seeks to minimize the costs of social benefit programmes that guarantees people an income above a certain level (income maintenance). Individuals are different with respect to productivity parameters but not with respect to disutility parameters. The government observes the income of the individuals but cannot observe the hours worked to earn the income. Individuals may simultaneously work at the ordinary labour market and in activation programmes. Work requirements may be used to deter people from collecting the social benefit rather than work at the ordinary labour market. This may be part of an optimal policy when the productivity parameters differ a lot between individuals. In this case, work in activation in 'exchange' for work at the ordinary market has a greater loss for high productivity workers than for low productivity workers, and therefore screening via work requirements is useful. Besley and Coate (1995) present a version of the model with a minimum level of utility (rather than income) as the sub-condition for the government. With that condition, the utility loss caused by the work requirement is included in the government's criterion function and is similar for individuals (since disutility is the same) and as a consequence, activation is no longer part of optimal policy. Brett (1998) obtains a similar result using a traditional criterion function of a benevolent government. (Assuming a useful product from work in activation, he finds that activation can be part of optimal policy.) Beaudry and Blackorby (1997), section 7, obtain a similar result using somewhat different assumptions.

In Thustrup Kreiner and Tranæs (2003) productivity and disutility parameters are both private knowledge. There are two types of individuals, since all individuals with high productivity have low disutility and all individuals with low productivity have high disutility. The government designs a social assistance programme intended for low productivity individuals (non-workers) and an unemployment insurance programme intended for high productivity individuals who become involuntarily unemployed with some probability. The government maximize utility of high productive individuals subject to the constraint that low productivity individuals obtain a minimum level of utility. In some cases, activation can be part of the optimal unemployment insurance programme. Activation deters voluntarily unemployed non-workers from the unemployment insurance programme (that has a high level of benefit). Screening is possible because participants in the activation programme have low disutility and hence is less hurt by activation than non-participants.

In Cuff (2001), the government's criterion function is in a sense the opposite of that in Thustrup Kreiner and Tranæs (2003), since the government maximizes utility for the individuals with the lowest level of utility (low productivity individuals). In one version of the model, the individuals with the lowest level of utility are also those with low disutility and in this version, activation deters high disutility individuals from obtaining benefit. (In that version, disutility equals lost value of leisure rather than pain from going to work.)

In this paper, we choose to use a traditional criterion function for a benevolent government and we do not assume any exogenous minimum income or utility levels in the social programmes. In this way, we hope to avoid that activation programmes are proved optimal only because other welfare programmes are designed non-optimally or because welfare for some individuals are left out of consideration. The government designs a *pure benefit* programme and an *activation* programme. The first consists simply of a benefit for the participants and the second of a (higher) benefit combined with a requirement of 'activation', i.e. some sort of effort carried out by the participants.

Individuals are different with respect to the disutility parameter and the wage rate (the productivity rate). The government knows the population distribution of the disutility parameter and the wage rate, but not the characteristics of each individual. The government knows whether an individual works. As is hopefully indicated in the literature review above, two dimensions of private knowledge are important for the analysis: social benefit programmes are typically for people with low productivity, and the different disutility parameters open for work requirements to be used to screen some people for programme participation. In the paper, we begin the analysis with a general two-dimensional continuous distribution function and discuss the relation between the shape of the distribution function and the conclusion about whether activation policy is part of optimal policy. Generally, the activation programme is expensive compared to pure the benefit programme (per participant) and is used by people who are relatively well off. Hence, for the activation programme to be part of optimal policy, the 'labour supply incentive effects' (i.e. the effect that a higher benefit level makes people transit form work to social programmes) has to be small for the activation programme compared to the pure benefit programme. For example, with a uniform distribution of characteristics, the activation programme turns out not part of an optimal policy.

Papers differ with respect to the tax-structure used to finance social programmes. In this paper, we assume a simple lump sum tax, while most other papers use a non-linear taxation model.

In regard to the public debate on activation policy, remark that we investigate whether activation policy is *part* of optimal policy, so that activation – if used – very likely is used along with a benefit programme without activation. In Denmark, activation policy is however mandatory in both the unemployment insurance programme and the social assistance programme. To be more precise, activation is in Denmark mandatory only after the individual have colleted the benefit for some time. This paper considers a static model and we cannot preclude that the ‘dynamics’ of real-world benefit policies affects the conclusions of the paper.

The paper is organized as follows. In section 2, we set up the model. In section 3, we derive a general necessary condition for activation policy to be part of optimal social policy. In subsection 3.1, we present three special cases, and in subsection 3.2, we give a short description of a similar type of analysis carried out in a companion paper (Rasmussen (2004)). In that paper, a particular disutility parameter is related to activation. The very short sections 4 and 5 discuss the criterion function and a useful variation of the model which is related to involuntary unemployment. Section 6 concludes.

2. The model

Population and private knowledge: Individuals are different with respect to the wage rate, w, and the disutility parameter, d. An individual is denoted (w,d). We assume that $w,d \geq 0$ and the joint density function for wage and disutility is denoted f. Each individual knows her own wage rate and disutility parameter. The government knows the distribution of characteristics.

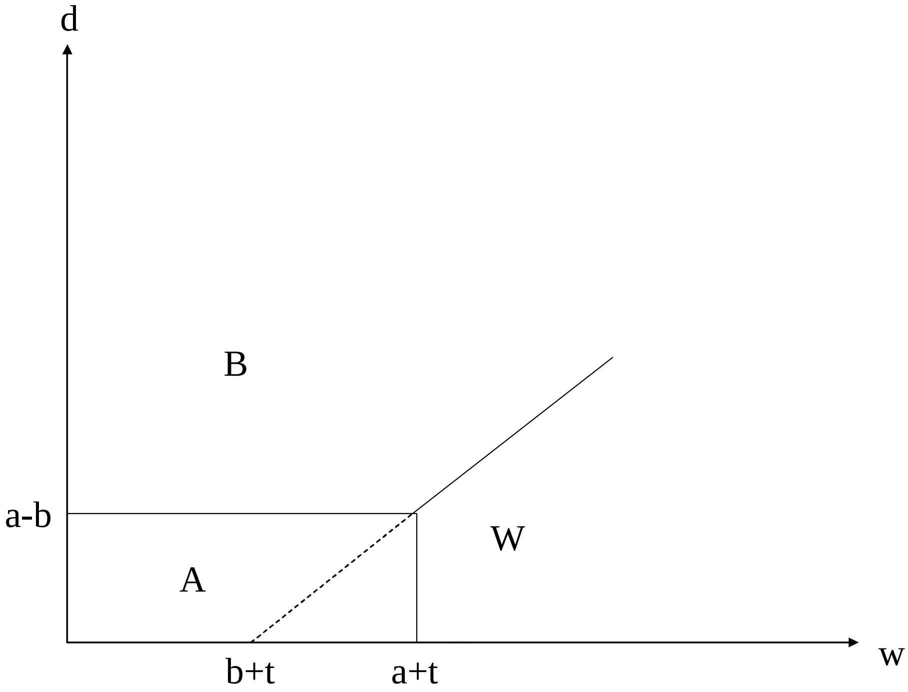
Individuals' utility in various states: An individual derives utility from income and disutility from work or activation. On the basis of the obtainable level of 'income minus disutility' in the three 'states', *work*, *pure benefit*, and *activation*, the individual chooses a state. If an individual works, she supplies one unit of labour. From this, she obtains wage, w, disutility, d, and pay the lump sum tax rate t. In the activation programme, disutility is as for regular work. The benefit rate net of tax is denoted a. In the pure benefit programme, there is no disutility and the benefit rate net of tax is denoted b.

We denote by $B, A,$ and W the sets of individuals who choose pure benefit, activation, or work. The population distribution across states as function of wage rates, disutility, and politically determined variables, is

$$\begin{aligned} W &= \{(w,d) \mid w-t-d \geq b \text{ and } w-t-d \geq a-d\} \\ A &= \{(w,d) \mid a-d > w-t-d \text{ and } a-d > b\} \\ B &= \{(w,d) \mid b > w-t-d \text{ and } b \geq a-d\} \end{aligned} \tag{1}$$

Figure 1 illustrates the sets.

Figure 1. Distribution of individuals across states as function of policy variables



In drawing figure 1, we implicitly assume $a, b, t \geq 0$ and $a > b$. Negative benefits are excluded because we implicitly assume existence of a fourth 'programme', namely a programme with no benefits and no work (e.g. home working wives), which is preferable to a programme with negative benefits. Hence, in optimum, negative benefits would never occur. Since taxes finance the costs of the programmes, the tax rate will also be positive. Furthermore, if $a < b$ no individual would choose activation. Finally, we like to include individuals who receive no public benefit and do not work in the set B: if $b>0$, all individuals in B receives the benefit with this broad definition. If $b=0$, we might say that individuals in B receives a benefit equal to zero. This modification of B is a technicality used in a proof below.

Government's problem: The government's criterion function depends on income minus disutility for each individual. Each individual contributes to the criterion function through the increasing, concave, and continuous function u, $u:[0,\infty)\to[0,\infty)$. The more 'curved' the u is, the greater weight the government puts on equality. The criterion function is

$$\begin{aligned}V=&\int_{(w,d)\in W} u(w-t-d)f(w,d)\,\mathrm{d}(w,d)+\int_{(w,d)\in A} u(a-d)f(w,d)\,\mathrm{d}(w,d)\\&+\int_{(w,d)\in B} u(b)f(w,d)\,\mathrm{d}(w,d)\\&=\int_{w=a+t}^{\infty}\int_{d=0}^{w-(b+t)} u(w-t-d)f(w,d)\,\mathrm{d}d\mathrm{d}w+\int_{w=0}^{a+t}\int_{d=0}^{a-b} u(a-d)f(w,d)\,\mathrm{d}d\mathrm{d}w\\&+\left[\int_{w=0}^{a+t}\int_{d=a-b}^{\infty} u(b)f(w,d)\,\mathrm{d}d\mathrm{d}w+\int_{w=a+t}^{\infty}\int_{d=w-(b+t)}^{\infty} u(b)f(w,d)\,\mathrm{d}d\mathrm{d}w\right]\end{aligned}\tag{2}$$

The government's budget constraint is

$$F=Wt-Aa-Bb=0\tag{3}$$

The government's problem is to maximize V subject to the budget constraint. The policy variables are the benefit rates and the tax rate.

3. A necessary condition for activation policies to be optimal

Let a, b and t be a solution to the government's problem. Proposition 1 gives a necessary condition for $a > b$, i.e. for an activation programme to be part of an optimal social policy. Remember, if $a \leq b$ no individuals would prefer activation to pure benefits. Hence, if $a > b$, the government actually sets up an activation programme, and if $a \leq b$, it does not.

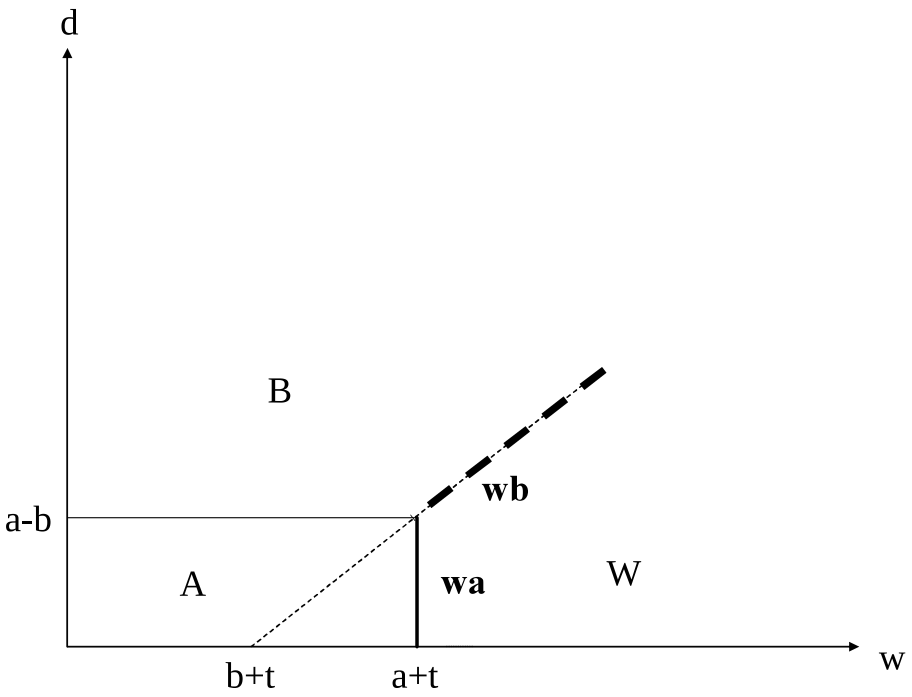
Proposition 1. Assume that the density function f is continuous, and let a, b be a solution to the government's problem. A necessary condition for $a > b$ is that

$$0 < \frac{\partial W / \partial a}{A} < \frac{\partial W / \partial b}{B} \tag{4}$$

is true at the optimal values. •

Condition (4) says that the number of individuals who move out of work in response to a benefit increase relative to the number of 'initial' beneficiaries, should be low for the activation programme compared to the pure benefit programme. In figure 2 below, $\frac{\partial W}{\partial a}$ is the bold line segment denoted $\boldsymbol{wa}$ and $\frac{\partial W}{\partial b}$ is the dotted line segment denoted $\boldsymbol{wb}$ (see the appendix).

Figure 2. Transition from work to social programmes



Technically, (4) is an assumption about the distribution of individual characteristics at the optimum values and hence not an assumption directly on the exogenous elements of the theory. Of course, if we assume (4) to hold for all values for all values of $a > b \geq 0$, then (4) is an assumption about the exogenous elements of the theory.

The proof consists of a very simple derivation and inspection of the first order conditions of the Lagrange function. It is lengthy, but parts of the proof are used to give an intuitive understanding of the government's problem.

Proof.

The Lagrange function is $L(a,b,t)=V-\lambda F$. Before we write the first order conditions, we explain three ways to simplify the notation.

First, we denote the marginal values of u for pure benefit recipients and the averages of the marginal values for participants in the activation programme and for individuals who work as

$$u'(b)=\partial u(b)/\partial b$$

$$\bar{u}'(a-d)=\frac{1}{A}\int_{(w,d)\in A} u'(a-d)f(w,d)\mathrm{d}(w,d) \tag{5}$$

$$\bar{u}'(w-t-d)=\frac{1}{W}\int_{(w,d)\in W} u'(w-t-d)f(w,d)\mathrm{d}(w,d)$$

Second, as regards the partial derivatives of V, consider an increase of b. The value of V is affected because individuals obtaining the pure benefit gain. Furthermore, some individuals who work and some individuals who participate in the activation programme move to the pure benefit programme. It turns out however that these 'movers' do not affect V. This is because individuals on the border of two states are indifferent between the two states. Hence, below we omit the effect from 'movers' when evaluation e.g. $\frac{\partial V}{\partial b}$. Result 1 establishes this.

Result 1.

$$\frac{\partial V}{\partial b}=u'(b),\frac{\partial V}{\partial a}=\bar{u}'(a-d),\frac{\partial V}{\partial t}=-\bar{u}'(w-t-d). \tag{6}$$

Proof.

$$\frac{\partial V}{\partial t} = \int_{w=a+t}^{\infty} \int_{d=0}^{w-(b+t)} -u'(w-t-d) f(w,d) \mathrm{d}d\mathrm{d}w - \int_{w=a+t}^{\infty} u(b) f(w, w-(b+t))\mathrm{d}w$$

$$- \int_{d=0}^{a-b} u(a-d) f(a+t,d)\mathrm{d}d$$

$$+ \int_{d=0}^{a-b} u(a-d) f(a+t,d)\mathrm{d}d$$

$$+ \int_{d=a-b}^{\infty} u(b) f(a+t,d)\mathrm{d}d - \int_{d=a-b}^{\infty} u(b) f(a+t,d)\mathrm{d}d + \int_{w=a+t}^{\infty} u(b) f(w, w-(b+t))\mathrm{d}w$$

$$= -W\overline{u}'(w-t-d)$$

This proves the result for a change of the tax rate. The remaining parts of the results are proved in a similar way. •

Third, the partial derivatives of the government budget are (we use $\frac{\partial B}{\partial b} = -\frac{\partial W}{\partial b} - \frac{\partial A}{\partial b}$)

$$F_b = \frac{\partial F}{\partial b} = -B + t\frac{\partial W}{\partial b} - b\frac{\partial B}{\partial b} - a\frac{\partial A}{\partial b} = -B + (b+t)\frac{\partial W}{\partial b} - (a-b)\frac{\partial A}{\partial b}$$

(and analogously for a and t).

Now first, suppose a,b,t is an interior solution to the government's problem, so that $a > b > 0$ and $A, B > 0$. The Lagrange conditions for an interior solution are

$$\begin{aligned}
L_b &= u'(b)B - \lambda\left(-B + (b+t)\frac{\partial W}{\partial b} - (a-b)\frac{\partial A}{\partial b}\right) = 0 \\
L_a &= \overline{u}'(a-d)A - \lambda\left(-A + (b+t)\frac{\partial W}{\partial a} - (a-b)\frac{\partial A}{\partial a}\right) = 0 \\
L_t &= -\overline{u}'(w-t-d)W - \lambda\left(W + (b+t)\frac{\partial W}{\partial t} - (a-b)\frac{\partial A}{\partial t}\right) = 0 \\
F &= Wt - Aa - Bb = 0
\end{aligned} \quad (7)$$

Note that $\lambda < 0$ in an optimum: to see this, note the term in the bracket is unambiguously negative in the equation for L_a (it is easy to show that $\frac{\partial W}{\partial a} < 0, \frac{\partial A}{\partial a} > 0$, see the appendix).

We rewrite the first two equations

$$\begin{aligned} &\frac{L_b}{B} = 0 \\ &\quad \Leftrightarrow\ u'(b) - \lambda\left(-1-(a-b)\frac{\partial A/\partial b}{B}\right) = \lambda(b+t)\frac{\partial W/\partial b}{B} \\ &\frac{L_a}{A} = 0 \\ &\quad \Leftrightarrow\ \bar{u}'(a-d) - \lambda\left(-1-(a-b)\frac{\partial A/\partial a}{A}\right) = \lambda(b+t)\frac{\partial W/\partial a}{A} \end{aligned} \tag{8}$$

It is easy to show that $\frac{\partial A}{\partial b} < 0$, see the appendix. Also, since individuals the choose state themselves, and since all individuals obtain the same b if they choose the pure benefit programme, and finally, since u is concave, then $u(a-d) > u(b)$ for all individuals who choose the activation programme rather than the pure benefit programme, and hence $\bar{u}(a-d) > u(b)$, and $\bar{u}'(a-g) < u'(b)$. Therefore, for (8) to be fulfilled, (4) must hold.

Finally, suppose $a > b = 0$. This is a border solution where $A > 0$ and $L_b \leq 0$. However, the arguments above remain exactly the same in this case, and (4) is a necessary condition for $L_a = 0 \geq L_b$ to be possible. The only technical modification is that B is interpreted as those who receives no benefits (or a benefit b equal to 0). This is used to ensure $B > 0$ which is used above. •

To compare the welfare effects of marginal increases in the benefit rates, we might use the first order conditions (8) in the proof of the proposition to distinguish between three effects.

- A 'targeting effect' arising from $u'(b) > \overline{u}'(a-d)$. The inequality says that pure benefits have the advantage of targeting people with low utility better than benefits in activation programmes.
- A direct budget effect, namely $-1-(a-b)\frac{\partial A/\partial b}{B} > -1-(a-b)\frac{\partial A/\partial a}{A}$. The inequality says that pure benefits have the advantage of burdening the public budget less than activation benefits.
- An 'adverse labour supply effect', namely $\frac{\partial W/\partial b}{B}$ compared to $\frac{\partial W/\partial a}{A}$. This is the effect that arises because people move from work to welfare programmes in response to benefit increases.

The first two bullets are – loosely speaking – arguments for using pure benefit programmes rather than activation programmes. The third bullet however may be an argument for using either of the programmes. With a criterion function as the one in this paper, the labour supply effect in general prevents the government from redistributing income as much as it would otherwise like to. Activation policies might be useful to reduce this problem, if the labour supply effect is small for the individuals who will actually choose activation (i.e. individuals with low disutility in activation and low wage rate).

Finally, notice that the three remaining possible optimal social policies are either no real programmes open (i.e. $a = b = 0$), or a policy with only the pure benefit programme open (i.e. $b > 0, b \geq a$), or a policy with only the activation programme open (i.e. $b = 0, a > 0$). The third of these options is a special case of proposition 1. As concerns the case $a = b = 0$, inspection of the Lagrange conditions (7) using these values (and $t = 0$) leads to a condition $u'(0) < \bar{u}'(w - d)$ which is not true because u is concave.[2] The case $b > 0, b \geq a$ is optimal if (4) is not fulfilled.

3.1. Special cases

In this subsection we discuss whether the condition (4) is fulfilled in three special cases. The first two cases concern continuous distributions of w and d, and the third case is a verbal discussion of a discrete distribution of characteristics.

[2] However, the use of $t = 0$ is critical in this argument. In the model, $a = b = 0$ implies $t = 0$, because the government uses the tax revenue only to cover the costs of the two programmes. With benefit rates and tax rates equal to zero, people moving out of work does not affect the public budget. If the government has expenditures for other purposes, we cannot infer $t = 0$ from $a = b = 0$. In this case, it is possible that $a = b = 0$ could be an optimal social policy.

3.1.1. w and d independent and uniformly distributed

Proposition 2. Suppose that $w, d \in [0,1]$ and that the parameters are independent and uniformly distributed. Then condition (4) is not fulfilled. •

Proof. We prove that $a > b$ and condition (4) cannot be fulfilled simultaneously by proving that for $a > b$

$$-\frac{\partial W / \partial a}{A} > -\frac{\partial W / \partial b}{B} \tag{9}$$

Relevant sets and derivatives are (see figure 2, plot in lines for $w = 1$ and $d = 1$)

$$W = 0.5\left(1-(b+t)\right)^2 - 0.5(a-b)^2$$
$$A = (a-b)(a+t)$$
$$B = 1 - W - A$$

and

$$\frac{\partial W}{\partial b} = (a+t) - 1$$
$$\frac{\partial W}{\partial a} = -(a-b)$$

Note that $(a-b), (a+t), (1-(a+t)) \in (0,1)$ for sensible values of a, b, t. Note that $B > 1-(a-b)(a+t)$. Hence, if the first inequality in

$$-\frac{\partial W / \partial a}{A} = \frac{a-b}{(a-b)(a+t)} > \frac{1-(a+t)}{(1-(a-b)(a+t))} > \frac{1-(a+t)}{B} = -\frac{\partial W / \partial b}{B} \tag{10}$$

holds, then (9) is proved. Rearrange the first inequality

$$(a-b)(1-(a-b)(a+t)) > (a-b)(a+t)(1-(a+t))$$
$$\Leftrightarrow$$
$$(a-b)-(a-b)^2(a+t) > (a-b)-(a-b)(a+t)$$
$$=(a-b)(1-(a+t)) > (a-b)(a+t)(1-(a+t))$$

where the first inequality is true because $(a-b)\in(0,1)$, and the second inequality is true because $(a+t)\in(0,1)$. •

In figure 2, $\frac{\boldsymbol{wb}}{B}$ is small relative to $\frac{\boldsymbol{wa}}{A}$. This verifies the proposition graphically.

3.1.2. A simple correlation between w and d

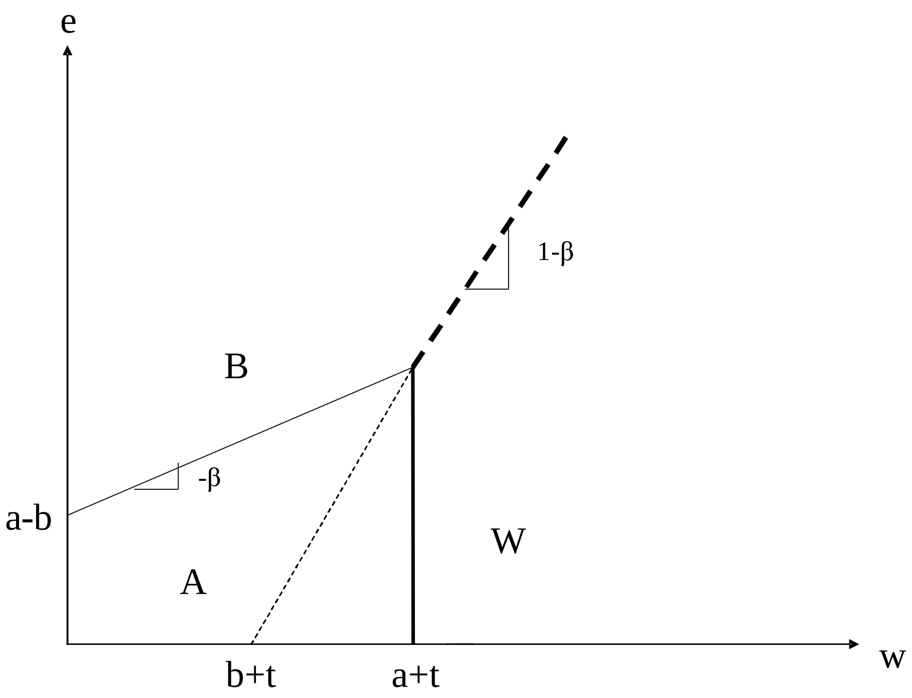
Assume that $w \geq 0$ is continuous and distributed according to the continuous density function h. Assume that $d = \beta w + e$, where $e \geq 0$ is continuous and distributed according to the continuous density function g, and assume that $\beta \in (-1,0)$. Assume that w and e are independent. The corresponding distribution functions are H and G. A negative β means that individuals who are lucky in one dimension (have a high wage) are also lucky in the other dimension (are not hurt much by working). This appears very likely.

The sets of individuals in the three states are

$$A = \int_{w=0}^{a+t}\int_{d=0}^{a-b} f(w,d)\,\mathrm{d}d\mathrm{d}w = \int_{w=0}^{a+t}\int_{e=0}^{a-b-\beta w} h(w)g(e)\,\mathrm{d}e\mathrm{d}w$$
$$W = \int_{w=a+t}^{\infty}\int_{d=0}^{w-(b+t)} f(w,d)\,\mathrm{d}d\mathrm{d}w = \int_{w=a+t}^{\infty}\int_{e=0}^{(1-\beta)w-(b+t)} h(w)g(e)\,\mathrm{d}e\mathrm{d}w$$
$$B = 1 - A - W$$

Figure 3 shows the sets

Figure 3. Distribution of individuals across states with simple correlation of characteristics



First, define $m \geq a+t$ such that

$$\int_{w=a+t}^{\infty} h(m)g((1-\beta)w-(b+t))\mathrm{d}w = \int_{w=a+t}^{\infty} h(w)g((1-\beta)w-(b+t))\mathrm{d}w \qquad (11)$$

(such an m clearly exists) The density $h(m)$ represents average density of the wage on the border illustrated by the bold dotted line in figure 3.

Proposition 3. Let m be defined as in (11). If

$$h(m)\frac{1}{1-\beta} < h(a+t) \tag{12}$$

then condition (4) is not fulfilled. •

Hence, the more closely the two parameters are correlated (the greater the numerical value of β), the less likely it is that (4) is fulfilled.

Proof.

Relevant derivative of the sets of individuals are

$$\frac{\partial W}{\partial b} = -\int_{w=a+t}^{\infty} h(w)g((1-\beta)w-(b+t))\mathrm{d}w$$
$$\frac{\partial W}{\partial a} = -h(a+t)G((1-\beta)(a+t)-(b+t))$$

Note from figure (3) that

$$A < H(a+t)G(a-b-\beta(a+t))$$

$$B > H(a+t)(1-G(a-b-\beta(a+t)))\,.$$

Then write

$$\begin{aligned} \frac{\partial W/\partial a}{A} &> \frac{h(a+t)G(a-b-\beta(a+t))}{H(a+t)G(a-b-\beta(a+t))} = \frac{h(a+t)}{H(a+t)} \\ \frac{\partial W/\partial b}{B} &< \frac{\int_{w=a+t}^{\infty} h(w)g((1-\beta)w-(b+t))\mathrm{d}w}{H(a+t)(1-G(a-b-\beta(a+t)))} \end{aligned} \tag{13}$$

Use the substitution $k = w(1-\beta) - (b+t)$ and m as defined in (11) such that

$$\int_{w=a+t}^{\infty} h(w) g((1-\beta)w - (b+t))\mathrm{d}w = h(m) \int_{w=a+t}^{\infty} g((1-\beta)w - (b+t))\mathrm{d}w$$

$$= h(m) \int_{k=a-b-\beta(a+t)}^{\infty} g(k) \frac{1}{1-\beta} \mathrm{d}k \tag{14}$$

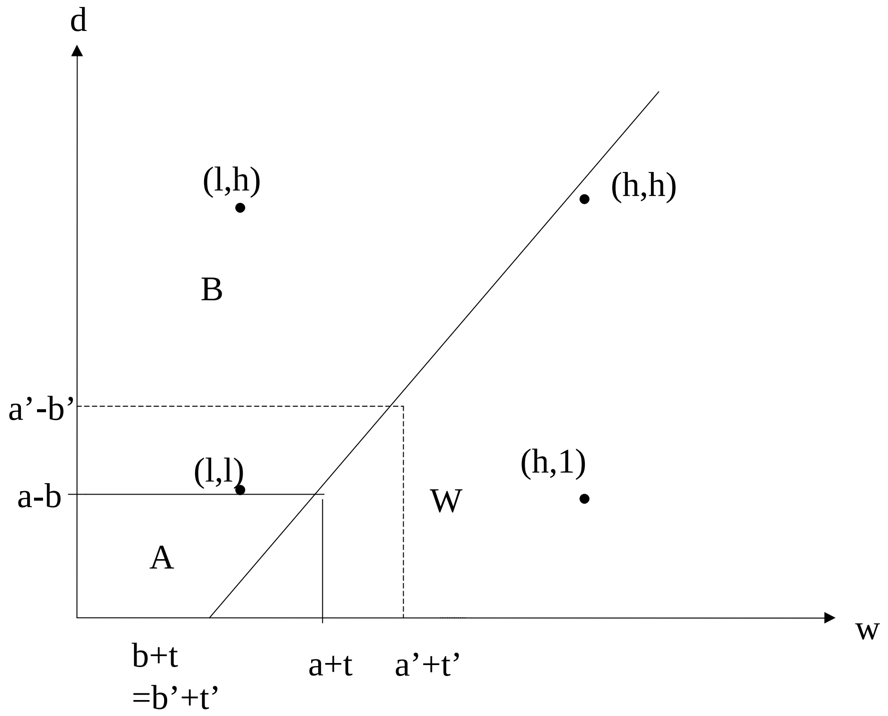
$$= h(m) \frac{1}{1-\beta} \left[1 - G(a-b-\beta(a+t))\right]$$

Inserting (14) into (13) shows that if (12) is fulfilled, then (4) is not fulfilled. •

3.1.3. A four-point discrete distribution

As showed above, for activation policy to be used, the 'adverse labour supply effect' has to be small for the activation programme compared to the pure benefit programme (condition (4)). In figure 2, this is the case if the border between A and W is 'thin' relative to the number of participants in activation (A) and relative to the same measure for pure benefits. It is easy to illustrate how condition (4) can be fulfilled if we use a discrete distribution of characteristics. In figure 4, we consider a four point distribution (high and low values of wages and disutility). The lines separating individuals in various states are similar to those in figure 1-3. In this section, we verbally discuss this distribution.

Figure 4. Distribution of individuals across states with a four point discrete distribution of characteristics



In figure 4, two candidates for a solution to the governments problem are considered, namely alternative 1) with a, b, t and alternative 2) with a', b', t' made so that $b + t = b' + t'$. In both of the alternatives, types (h,l) and (h,h) work and type (l,h) are collecting pure benefits, while type (l,l) is collecting pure benefits in alternative 1) but collecting activation benefits in alternative 2).

Now consider a change from 1) to 2). There is no labour supply effect related to this change. The change makes type (l,l) better off, while type (l,h) becomes worse off. Taxpayers also become worse off. If the number of people of type (l,h) is small compared to the number of people of type (l,l), the change might be gainful for the government. In this case, activation is a way to deter types (l, h) and (h, h) from obtaining the high activation benefit. The low level of the pure benefit makes type (h, h) to prefer work rather than the pure benefit. Finally, the motivation for having a pure benefit programme is to give people of type (l, l) a relatively high level of utility. Hence, the activation programme is not the only programme open (so activation is not mandatory in figure 4 as it is in many real-world cases).

3.2. A variation: distinct disutility in activation

One way to vary the model is to assume a disutility parameter related to activation different from disutility related to ordinary work. It appears realistic that such a difference exists. A model with this feature might ideally also be useful in the practical design of activation programmes, because the way that activation programmes expose participants to disutility may be manipulated in the design of the programmes. In a companion paper (Rasmussen (2004)), we analyse such cases. The models lead to conclusions much in line with the analysis above. In this section, we therefore very briefly illustrate and discuss two models. In both models, the disutility parameter of activation is private knowledge. This parameter is denoted g. In the first model, the second private-knowledge parameter is disutility in work (d) while the wage rate (w) is constant across individuals. The roles of w and d are reversed in the second model.

3.2.1 Disutility parameters in activation, g, and work, d, as private knowledge

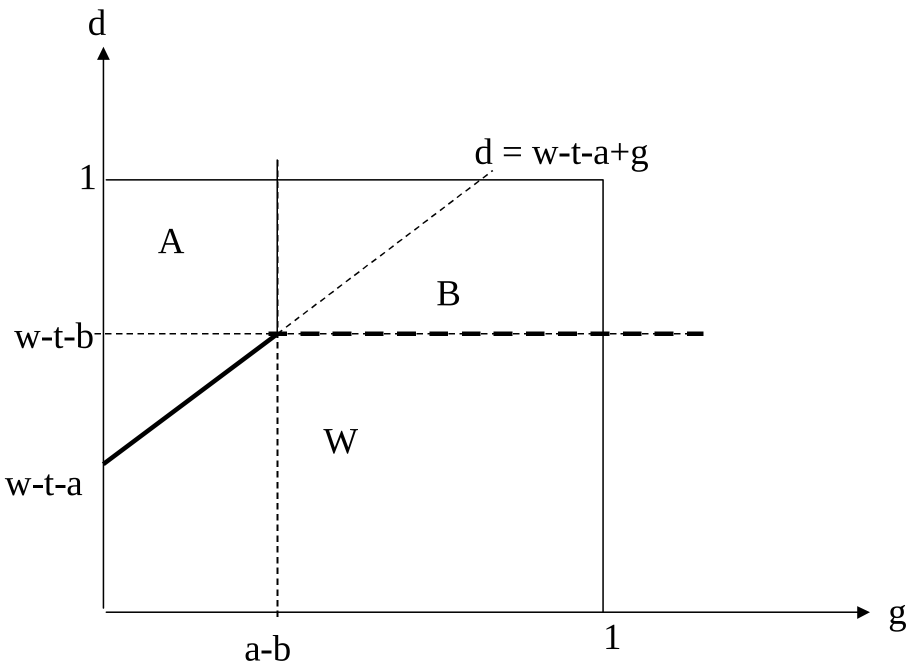
The disutility parameters in activation, g, and in work, d, are private knowledge, and the parameter w is observable and constant. A participant in activation gains $a-g$. We assume $g, d \geq 0$. The notation is otherwise as above.

The sets of individuals in the three states are

$$
\begin{aligned}
W &= \{(g,d) \mid w-t-d \geq b \text{ and } w-t-d \geq a-g\} \\
A &= \{(g,d) \mid a-g > w-t-d \text{ and } a-g > b\} \\
B &= \{(g,d) \mid b > w-t-d \text{ and } b \geq a-g\}
\end{aligned}
$$

Figure 5 illustrates the sets (we also assume $g, d \in [0,1]$ in the figure).

Figure 5. Distribution of individuals across states with g and d as private knowledge



For the labour supply effect to be small in this figure, the line segment ***wa*** should be small relative to the number of activation participants A and relative to ***wb*** relative to B. This *might* be the case with a strong negative correlation between g and d, for example if $d = 1 - g$.

The government cannot affect disutility parameters g of the individuals, but it might be able to set up different types of activation programmes with different types of work requirements, and each of such programmes might affect people differently. The government might therefore design the activation programme such that condition (4) is fulfilled. Such a design demands that individuals who are only lightly affected by the disutility in activation are very strongly affected by the disutility in work (so that the density in the neighbourhood of $(g,d)=(0,1)$ is high compared to the density near ***wa***), and, reversely, those with disutility characteristics making them approximately indifferent between work and pure benefits are more strongly affected by the disutility in activation (so that the density near ***wb*** is high compared to the density near $g>a-b, d=1$). It appears nevertheless difficult to imagine and suggest real-world examples of activation characteristics with these features.

3.2.2 Disutility parameter in activation, g, and wage rate, w, as private knowledge

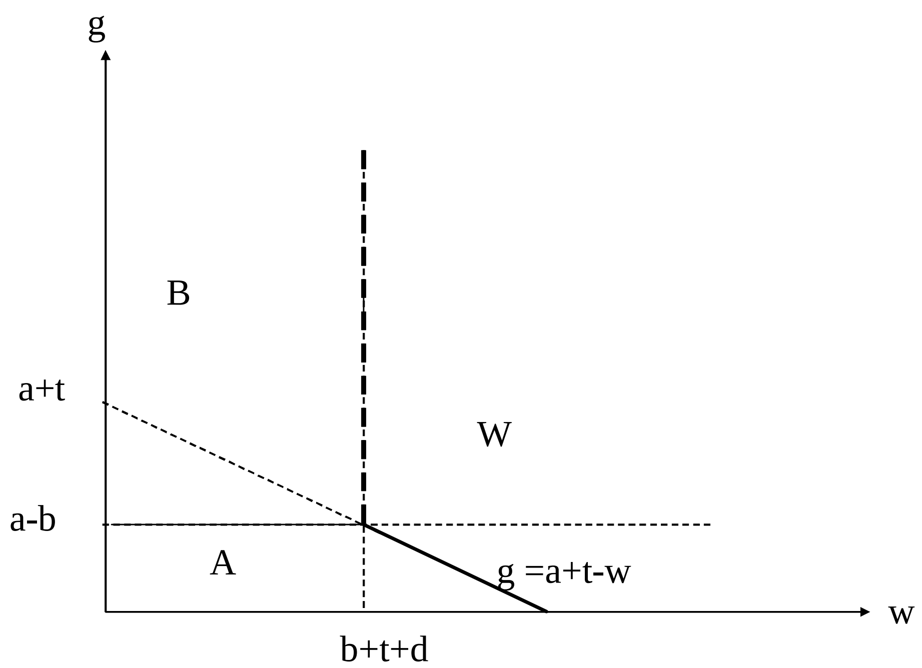
The disutility parameter in activation, g, and the wage rate, w, are private knowledge, and the parameter d is observable and constant. We assume $w,g\geq 0$. The notation is otherwise as above.

The sets of individuals in the three states are

$$
\begin{aligned}
W &= \{(w,g)\,|\,w-t-d\geq b \text{ and } w-t-d\geq a-g\}\\
A &= \{(w,g)\,|\,a-g>w-t-d \text{ and } a-g>b\}\\
B &= \{(w,g)\,|\,b>w-t-d \text{ and } b\geq a-g\}
\end{aligned}
$$

Figure 6 displays the sets.

Figure 6. Distribution of individuals across states with g and w as private knowledge



For activation to be optimal, the activation programme should be designed so that individuals with a wage rate much below the benefit rates have little disutility in the activation programmes, while individuals with a wage rate approximately equal to the pure benefit rate (or higher) have high disutility in activation. As in the previous section, it is not easy to find good realistic ideas to such activation characteristics.

4. The criterion function

In some studies (Besley and Coate (1992, 1995) and Thustrup Kreiner and Tranæs (2003)) a minimum level of income or utility is considered fixed (whereas we above consider the pure benefit level as endogenous). To compare with this approach, suppose the pure benefit rate b is exogenous.

With b fixed, we consider levels of a's such that $a > b$ (so some individuals prefers activation) and $L_a = 0$ (assuming such levels of a's exists), and ask whether the exogenous level of b is in-optimally high ($L_b < 0$) or in-optimally low ($L_b > 0$). (Using this argument, we implicitly assume second order partial derivatives to be negative on the entire domain).

If condition (4) is not fulfilled, proposition 1 implies that $L_a < L_b$. Hence, $L_a = 0$ implies that pure benefit level is in-optimally high. If condition (4) is fulfilled the level of b might be in-optimally high or in-optimally low.

An exogenous level of b generally allows activation policies to be considered optimal in more cases than if b is not considered exogenous (because pure benefit programmes and activation programmes are 'competing' policies).

5. Involuntary unemployment

So far, we assume that people choose for themselves whether to be working or to collect public benefits. Hence, the model's explanation for non-employment solely relies on 'supply side' factors. It is however very simple to include 'demand side' factors in a simple way by assuming that each individual receives a job offer with a certain (private knowledge) probability p. We will argue that the model is very similar to the one described above.

If an individual does not supply labour she chooses activation if $a-d>b$, and pure benefits if the reverse is true.

If an individual supplies labour, she have to choose the type of benefit she prefers if she is involuntary unemployed. Again, $a-d$ is compared to b.

Finally, the individual has to decide whether or not to supply labour. Suppose $a-d>b$ so activation is preferred to pure benefits if the person is out of work (voluntarily or involuntarily). For supply of labour to be preferred to activation

$$p(w-t-d)+(1-p)(a-d)>a-d$$
$$\Leftrightarrow$$
$$w-t-d>a-d$$

This condition is exactly as in the previous sections. An analogous argument holds if $a-d\leq b$. Therefore, individuals preferred state is not affected by the introduction of a probability of being the involuntary unemployed.

6. Conclusion and discussion

We show how assumptions on the distribution of individuals' private-knowledge characteristics affect whether a rational benevolent government should use activation programmes as part of social policy. We only consider one aspect of activation policy, namely the 'deterrence effect' that induces some individuals to work rather than to collect social benefit because of the disutility related to activation programmes. We derive a simple, intuitive and necessary condition for activation programmes to be used in optimum, namely that ‘adverse labour supply effects’ related to activation benefit should be small compared to the adverse effects related to pure benefits. Even though we have not found general classes of distributions such that this condition is fulfilled or violated, the examples appear to indicate that activation programmes are optimal only with rather special distributional assumptions. In this way, the analysis confirms the result by Besley and Coate (1995) and Brett (1998).

On the other hand, the analysis only considers the deterrence effect from activation programmes. Activation programmes in the real world might produce valuable outcomes such as goods or improvements of participants’ education or labour market ability. But real-world costs of activation programmes are also excluded from the analysis.

Acknowledgements

Financial assistance from the Danish Social Science Research Council is gratefully acknowledged. Thanks to Torben Tranæs for helpful discussions.

Appendix

We write the distribution of individuals across states and derivatives of the sets of individuals. The sets of individuals in the three states are

$$A = \int_{w=0}^{a+t} \int_{d=0}^{a-b} f(w,d)\,\mathrm{d}d\,\mathrm{d}w$$

$$W = \int_{w=a+t}^{\infty} \int_{d=0}^{b+t-w} f(w,d)\,\mathrm{d}d\,\mathrm{d}w$$

$$B = 1 - A - W$$

The derivatives are

$$\frac{\partial A}{\partial b} = -\int_{w=0}^{a+t} f(w, a-b)\,\mathrm{d}w \qquad [<0,\text{ used in proposition 1}]$$

$$\frac{\partial W}{\partial b} = -\int_{w=a+t}^{\infty} f(w, w-(b+t))\,\mathrm{d}w \qquad [= \boldsymbol{wb},\text{ used in figure 2}]$$

$$\frac{\partial B}{\partial b} = -\frac{\partial W}{\partial b} - \frac{\partial A}{\partial b} = \int_{w=0}^{a+t} f(w, a-d)\,\mathrm{d}w + \int_{w=a+t}^{\infty} f(w, w-(b+t))\,\mathrm{d}w$$

$$\frac{\partial A}{\partial t} = \int_{d=0}^{a-b} f(a+t, d)\,\mathrm{d}d$$

$$\frac{\partial W}{\partial t} = -\int_{d=0}^{a-b} f(a+t, d)\,\mathrm{d}d - \int_{w=a+t}^{\infty} f(w, w-(b+t))\,\mathrm{d}w$$

$$\frac{\partial B}{\partial t} = \int_{w=a+t}^{\infty} f(w, w-(b+t))\,\mathrm{d}w$$

$$\frac{\partial A}{\partial a} = \int_{d=0}^{a-b} f(a+t, d)\,\mathrm{d}d + \int_{w=0}^{a+t} f(w, a-b)\,\mathrm{d}w \qquad [>0,\text{ used in proposition 1}]$$

$$\frac{\partial W}{\partial a} = -\int_{d=0}^{a-b} f(a+t, d)\,\mathrm{d}d \qquad [= \boldsymbol{wa},\text{ used in figure 2}]$$

$$\frac{\partial B}{\partial a} = \int_{w=0}^{a+t} f(w, a-b)\,\mathrm{d}w \qquad [<0,\text{ used in proposition 1}]$$

References

Besley, T., Coate, S., 1992. Workfare versus Welfare: Incentive Requirements in Poverty-Alleviation Programs. The American Economic Review, 82, 249-261.

Besley, T., Coate, S., 1995. The Design of Income Maintenance Programmes. Review of Economic Studies, 62, no. 1. p. 187-221.

Brett, C., 1998. Who should be on workfare? The use of work requirement as part of an optimal tax mix. Oxford Economic Papers, 50, 607-622.

Beaudry, P., Blackorby, C., 1997. Taxes and Employment Subsidies in Optimal Redistribution Programs. The University of British Columbia, Department of Economics, Discussion Paper no. 97-21.

Cuff, K., 2000. Optimality of workfare with heterogeneous preferences. Canadian Journal of Economics, 33, 149-174.

Rasmussen, M., 2004. Welfare effects of deterrence-motivated activation policy: the case of distinct activation-policy. Manuscript.

Thustrup Kreiner, C., Tranæs, T., 2003. Optimal Workfare with Voluntary and Involuntary Unemployment. Manuscript.