WORK REQUIREMENTS AND LONG-TERM POVERTY

FRED SCHROYEN
Norwegian School of Economics & Business Administration

GAUTE TORSVIK
University of Bergen

Abstract

We study how work requirements can be used to target transfers to the long-term poor. Without commitment, time consistency requires all screening measures to be concentrated in the first phase of the program. We show that this increases the effectiveness of workfare; it is optimal to use work requirements for a wider range of prior beliefs about the size of the poor population, and work requirements are used more intensively. We compare these results with the optimal policy under commitment.

1. Income Transfers and Incentive Problems

To contain the cost of poverty relief programs, it is important to channel resources to those in real need of them. Otherwise, unnecessarily large outlays will take place in the form of transfers flowing to people not in need of support, and this may undermine the political legitimacy of such programs. We analyze how effective work requirements are in target when poverty is persistent.

We are not the first to evaluate work requirements in the light of these considerations. Most notably, this issue has been addressed in a formal model by Besley and Coate (1992). The novelty of our study is the focus on *long-term poverty*. We let individuals' income opportunities be correlated over time, which means that a welfare administrator can collect information about these income opportunities as time passes. Potential welfare claimants might understand this, and adjust their behavior accordingly.

Fred Schroyen, Department of Economics Norwegian School of Economics & Business Administration, Helleveien 30, N-5045 Bergen, Norway (fred.schroyen@nhh.no). Gaute Torsvik, Department of Economics, University of Bergen, Fosswinckelgate 6, N-5007 Bergen, Norway and CMI (gaute.torsvik@econ.uib.no).

We thank two referees and an associate editor for their useful comments.
Received April 2002; Accepted November 2003.

Journal of Public Economic Theory, 7 (3), 2005, pp. 427–448.

To get a rough idea of how this long-term perspective influences the cost–benefit analysis of work requirements, consider the following problem. Let there be two groups of individuals in a society, one with a low income potential and one with a high income potential—we call them L- and H-individuals, respectively. The government wants to guarantee everyone a minimum income z, which is higher than the income L earns in the market, but lower than the income H earns. H-individuals may nevertheless claim benefits intended for the poor, since the welfare administrator cannot observe a person's income opportunities. Work requirements, or workfare, can be used to prevent such dissembling behavior. Requiring welfare recipients to work c hours in the public sector to qualify for transfers makes it costly for those with a relatively high earning capacity to join the program. Every hour spent in a public sector job could alternatively be used in the private sector, and since an H has a relatively high income potential this loss is relatively high. The negative effect of a work requirement is that it crowdes out L's market income and thus necessitates larger transfers to the poor in order to guarantee them an income above the poverty line.

Ignore for a moment the learning aspect associated with long-term poverty. Assume that there is no correlation between a person's present and future earning capacity—i.e., there exists only short-term poverty. Let the proportion of genuinely poor be low. There are, in other words, a lot of potential dissemblers around and it is important to deter them from joining the poverty program. Let c^s be the minimum level of public work that scares them off. As we have constructed the problem, the government minimizes costs by imposing a workfare program that requires the poor to work c^s hours in exchange of their benefits.

Assume now that individual's earning capacities are correlated over time. This means that the welfare administrator can learn about peoples' income potential by keeping a record of their past behavior. In fact, since a work requirement of c^s separated the two groups, she correctly infers that those who participated in the workfare program are genuinely poor. If she is free to change policy later on, she will certainly not make individuals work for their benefits in future periods. Now that the screening is done, it is only costly to use workfare. But, and this is the crux of the argument, if H-individuals perceive that welfare will be provided unconditionally later on, they will not be discouraged from participating in a poverty program that requires individuals to work c^s hours in the initial period.

As this example indicates, in a multi-period framework it becomes essential to specify whether or not policy makers can commit to the design of future poverty alleviation policies. We evaluate the effectiveness of different policy programs both with and without commitment.

1.1. Optimal Policy

We find that work requirements should in general be concentrated in the first phase of the program. Compared with the cost-efficient policy for eliminating short-term poverty, we find that workfare—as opposed to universal welfare—becomes a more efficient policy in containing the overall cost when poverty is long term. In some cases though—which we specify in detail later—the concentrated use of work requirements will scare away the poor from the program. To avoid that, the welfare administrator should allocate work requirements more evenly in time, even though this implies that fewer non-poor people separate. Finally, we analyze the optimal program if the welfare administrator *can* commit herself and find that in many cases the optimal commitment policy coincides with the equilibrium policy under non-commitment.

1.2. Methodology and Related Literature

In addition to the light that our model sheds on an important policy issue, we believe it has some methodological interest. Formally, we study the design of a dynamic Bayesian game. Our problem is therefore closely related to the literature on dynamic principal—agent relationships which emphasize the role that asymmetric information and long-term commitment plays in governance. Our problem of alleviating long-term poverty resembles the basic structure of, for example, a dynamic regulation problem. Still, the results we derive differ sharply from those obtained there. A central result in optimal regulation is that a regulator who is able to commit herself to a multi-period contract ought to repeat the optimal static policy in every period; and that this policy is not time consistent: The regulator will not follow the plan if she is free to re-optimize later on (cf. Laffont and Tirole 1990). Lack of commitment is therefore detrimental in a standard dynamic regulation problem.[1] In poverty alleviation, it is *not* always optimal to repeat the static program in each period, and, as a consequence of this, lack of intertemporal commitment is *not* always a problem. Another notable feature of our model is that if a semi-separating equilibrium exists, it involves randomization from both the agents (welfare recipients) *and* the principal (the welfare administrator).

Before we present the details of our arguments, we should say something about the scope of our perspective, and how it relates to the existing literature. The literature on how policy instruments can be used to target transfers to the poor is extensive (see Lipton and Ravallion (1995) for a discussion and references). Although the possibility of using work requirements to screen the needy from the not-so-needy had been discussed before, Besley and Coate (1992) were the first to give a formal analysis of the argument.[2] It is their

---

[1]Weitzman (1980) was the first to use a principal agent framework to point out the negative effects lack of intertemporal commitment has on the agent's behavior. Freixas et al. (1985) developed the first game theoretic analysis of a dynamic principal–agent relationship governed by linear incentive schemes. For other references and for a general discussion of this topic, see chapters 9 and 10 in Laffont and Tirole (1993). Dillén and Lundholm (1996) use the framework developed by Freixas et al. to discuss optimal income taxation in a dynamic model.

[2]See also Besely and Coate (1995).

model we extend to a dynamic environment. We think this is an important extension, both because there is virtually no theoretical work on the dynamics of poverty alleviation programs, and because long-term poverty is a serious problem: A substantial share of those who live below the poverty line do so persistently.[3]

Admittingly, the "cost efficiency perspective" on poverty alleviation and the effects of workfare that we borrow from Besley and Coate are narrow. One limitation is that it considers work requirements solely as a stick that scares the non-poor from claiming benefits. This is obviously not the whole story. Having a job can also be seen as an essential aspect of life, something that provides people with social recognition and self-esteem. Another important point is that making welfare claimants work for their benefits may prevent a deterioration of their working morale and human capital. Furthermore, it is not obvious that individuals are poor—as we assume—because they are endowed with an insufficient earning capacity. Alternatively, one may argue that it is the lack of well-functioning economic institutions to deal with property rights, information problems, etc. which is the main reason why so many people live in poverty (see Hoff 1996). We also ignore the political legitimacy of different poverty alleviation programs (see Besley 1996). We are not saying that these arguments are unimportant, only that they are irrelevant for the incentive problem we focus on.

Having pointed out the limits of our scope, we should, however, hasten to add that we believe the problem we point at warrants attention. Our arguments should be mentioned in a general debate about how one ought to provide assistance to the long-term poor, which is an important debate, both in developing countries and in more modern welfare states.

This paper is organized as follows. The next section presents a formal model of the costs and benefits of using workfare in targeting the poor. In Section 3, we characterize the cost-minimizing program in a static framework. In Section 4, which is the heart of the paper, we introduce dynamics and study how workfare can be used to minimize the cost of providing transfers to the long-term poor. In Section 5, we compare these results with the case where the welfare administrator can commit herself. Section 6 concludes the paper.

[3]For example, Heady et al. (1994) find that 10% of the population in Germany are frequently poor or near-poor. Rodgers and Rodgers (1993) conclude that about one-thirds of measured poverty in the United States as of 1987 can be regarded as "chronic," and that over the period they studied, "poverty not only increased, it became more chronic and less transitory in nature" (p. 51). Adams and Duncan (1988), in a study of US urban poverty, estimated that of the 13.4% of urban people that where poor in 1979, 34.6% were poor in at least 1 year between 1974 and 1983, and 5.2% were "persistently poor"—defined as poor in 8 out of 10 years.

In poor underdeveloped countries, the problem of chronic poverty is even more pronounced, Gibson (2001) uses data from a recent household survey in Papua New Guinea to conclude that close to half of those classified as poor has a chronic poverty problem.

2. A Formal Model of the Costs and Benefits of Using Workfare to Target Benefits to the Poor

As a prerequisite to the dynamic analysis, we analyze poverty alleviation in a static (one period) model. We follow Besley and Coate (1992) and assume that a welfare administrator, hereafter referred to as the WA, faces a target population of a size normalized to 1. A fraction γ has a very low productivity a_L and a fraction $(1-\gamma)$ is endowed with a higher productivity a_H. The latter are also "low class," but not as destitute as the former. All people have the same strictly concave utility function defined as over-disposable income (x) and leisure (ℓ), $u(x,\ell)$, and a time endowment normalized to unity. People choose the level of private labor earnings which maximizes their utility level. Without any program, the L-people (and only them) earn a disposable income below the poverty line z. The WA faces the task of designing a cost minimizing welfare program that guarantees everybody at least the minimal income z.[4]

A transfer program consists of a menu $\{(b_L, c_L), (b_H, c_H)\}$, where b is a money transfer and c the number of hours of public work an applicant is required to carry out in order to qualify for the transfer.[5] The menu must guarantee that (i) all people voluntarily participate in the program, (ii) everybody at least enjoys a disposable income z, (iii) nobody has an incentive to apply for the package intended for somebody with a different productivity, and (iv) the total cost of the program, $\gamma b_L + (1-\gamma) b_H$, is kept at a minimum (because it will be financed by distortionary taxation on the other people in the economy).

2.1. Individual Behavior

An individual with ability a, receiving the package (b, c) decides how much income (y) to earn:

$$\max_{y\geq 0} u\left(b+y, 1-c-\frac{y}{a}\right).$$

Let us denote the solution by $y(b,c,a)$. Normality of consumption and leisure means that as long as $y(b,c,a) > 0$, the derivatives w.r.t. c and b are negative.[6]

---

[4]Poverty is thus defined exclusively in terms of income, an attitude that is ubiquitous in public debate. Still, our main results would go through if the WA's aim is to guarantee a minimal *living standard*, including the value of leisure. For an analysis of the dynamics of redistribution in a utilitarian setting, see Dillén and Lundholm (1996).

[5]As Besley and Coate, we will assume that public sector work is unproductive. We discuss the impact of this assumption in footnote 14.

[6]Regarding $|\frac{\partial y}{\partial b}|$, Moffitt (1992) reports on a value of 0.37 for females, while Sawhill (1988, p. 1103) reports on values in the range [0.16, 0.71].

The corresponding maximal utility level is written as $v(b, c, a)$. Note that if the transfer b and/or the work requirement c are very high, it may be optimal to refrain from working privately altogether—the utility level then reduces to $u(b, 1 - c)$. Also note that our concavity assumption on $u(\cdot)$ implies $v_{bb} < 0$.

2.2. The Costs of Workfare

For a given work requirement c_L, let $b_L(c_L)$ be the lowest transfer that guarantees L-people a disposable income of at least z:

$$b_L(c_L) + y(b_L(c_L), c_L, a_L) \equiv z.$$

Implicit derivation shows that $\frac{\mathrm{d}b_L(c_L)}{\mathrm{d}c_L} = a_L$: A higher work requirement crowds out private sector earnings with a_L, and thus requires an extra a_L Euro to top up disposable income to the poverty line. Imposing a work requirement is thus costly because it necessitates larger transfers to needy people.

We define c^{co} as the work requirement that *crowds out* private sector earnings completely:

$$c^{co} \stackrel{\text{def}}{=} \max\{c : y(b_L(c), c, a_L) \geq 0\}.$$

The necessary transfer $b_L(c)$ thus satisfies

$$\begin{aligned} b_L(c) &= b_L(0) + a_L c \quad \text{if} \quad && c \leq c^{co}, \\ &= z && c \geq c^{co}, \end{aligned}$$

and is clearly concave in c.

Another important value is the work requirement that brings L down to his reservation utility level:

$$c^{\max} \stackrel{\text{def}}{=} \max\{c : v(b_L(c), c, a_L) \geq v(0, 0, a_L)\}.$$

Clearly, $c^{\max}$ puts an upper bound on the WA's selection of work requirements.

2.3. The Benefits of Workfare

The WA has to offer appropriate incentives to prevent H-individuals from joining the program. Pretending to be poor can be easy or difficult, depending on what the WA observes. One possibility is that the WA observes no personal characteristics; applying for a welfare package is then a sufficient condition for getting it. Another possibility is that the WA observes private sector earnings, and that welfare applicants qualify for transfers only when their earnings do not exceed a certain limit. In this paper, we limit ourselves to the first case.[7]

The maximum utility H gets if he receives a transfer b_H in exchange of a work requirement c_H is thus $v(b_H, c_H, a_H)$. On the other hand, when H pretends to be of type L, he attains a welfare level $v(b_L(c_L), c_L, a_H)$. The screening, or no mimicking constraint can thus be written as

$$v(b_H, c_H, a_H) \geq v(b_L(c_L), c_L, a_H).$$

Obviously, it is optimal to choose $c_H = 0$. Supplementing b_H with a positive work requirement implies a higher transfer to H, which increases the total cost of the program. To ease exposition, we drop the subscript on the work requirement since this policy is only relevant for the package intended for the poor.

Let $b_H^s(c)$ be the minimum transfer H must receive in order not to register as poor (superscript s for "static"). This is an information rent–resources H receives because the WA cannot observe his earning capacity. Its magnitude is implicitly defined by

$$v\big(b_H^s(c), 0, a_H\big) \equiv v\big(b_L(c), c, a_H\big). \tag{1}$$

Requiring the poor to work for their benefits makes it less attractive for H to mimic L and thus the minimum transfer b_H^s can be reduced. The following lemma informs about the shape of $b_H^s(c)$ (proven in the Appendix[8]).

LEMMA 1: *The transfer function $b_H^s(c)$ has the following first and second derivatives:*

$$\begin{aligned}\frac{db_H^s(c)}{dc} &= -(a_H - a_L) \quad \textit{if } c \leq c^{co},\\ &= -a^H \qquad\qquad c^{co} \leq c \leq c^{\max},\\ \frac{d^2 b_H^s(c)}{dc^2} &= 0.\end{aligned}$$

Moreover $b_H^s(0) = b_L(0)$.

By the last property, *universal welfare* is equivalent to $c = 0$. Since the transfer function is decreasing and concave in c, there exists a critical value

[7] The income observable case is discussed in Besley and Coate (1992) for short-term poverty alleviation and in Schroyen and Torsvik (1999) for long-term poverty alleviation. Allowing for means-testing will in general reduce the need for work requirements, although Besley and Coate (1995) have shown that even with a nonlinear income transfers (including earnings subsidies), workfare remains useful, as long as one is concerned with *income* maintenance. If the objective is *utility* maintenance, work requirements loose their role once means-testing scheme becomes flexible enough.

[8] None of the proofs are included in this paper. They are relegated to the appendix of the working paper, which is available upon request or may be downloaded from http://www.nhh.no/sam/res&publ/2001/dp34.pdf.

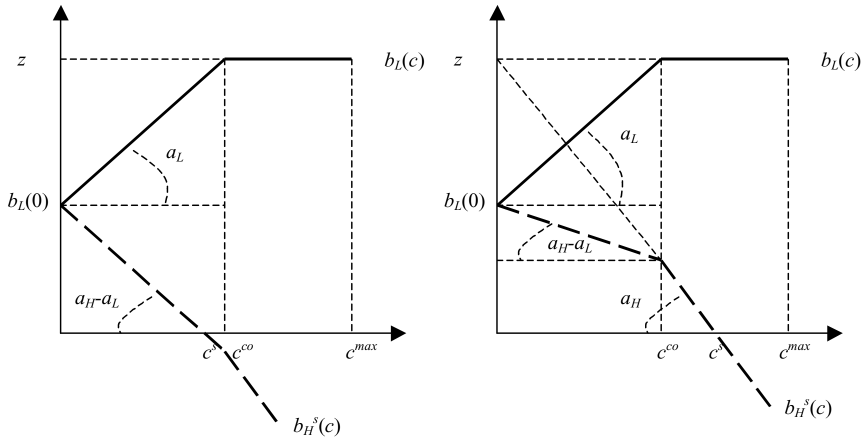

Figure 1: $b_L(\cdot)$ and $b_H^s(\cdot)$ when $c^s < c^{co}$ (left) and $c^s > c^{co}$ (right)

for the work requirement on L-persons, c^s, for which the transfer b_H can be set to 0 and still secure self-selection, i.e., $b_H^s(c^s) \equiv 0$. It is easy to see that $c^s < c^{\max}$. Figure 1 displays $b_L(c)$ and $b_H^s(c)$.

3. The Cost Minimizing Static Program

We can now construct the function which maps the work requirement c into the total cost of the program,

$$K^s(c) \stackrel{\text{def}}{=} \gamma b_L(c) + (1-\gamma) b_H^s(c).$$

By definition, this function gives—for any arbitrary work requirement—the minimal pair of transfer payments which satisfy both the poverty alleviation and the incentive compatibility constraints. As H-persons always have the option to stay away from the program, they cannot be imposed any taxes. This is equivalent to requiring that $b_H(c) \geq 0$ or $c \leq c^s$. The WA's problem can thus be stated as

$$\min_{c \leq c^s} K^s(c).$$

Since both transfer functions are piecewise linear but concave in c, there are two possible solutions: either c^s or 0. Workfare is either used so extensively that H-people do not sign up for poverty transfers, or workfare will not be used at all and poverty is alleviated through universal welfare. In the first case, the costs of alleviating poverty are $\gamma b_L(c^s)$; in the second, they amount to $b_L(0)$.

It is easy to understand that the choice between a welfare or a workfare program depends on how large the population of the poor is relative to the number of potential mimickers. The fewer potential mimickers there are in the population, the lower is the cost of paying them the rent which prevents them from applying for the package meant for the really needy. In the limit, as γ approaches 1, (almost) all individuals are of the L-type and it would be wasteful to distort the behavior of (almost) the whole population in order to eliminate a cost (the rent to the H-people) that is negligible.

Let γ^s be the value of γ for which the administrator is indifferent between universal welfare and workfare. It is then easy to check that

$$\gamma^s \stackrel{\text{def}}{=} \frac{b_L(0)}{b_L(c^s)} = 1 - \frac{a_L}{a_H}\frac{\min\{c^s, c^{co}\}}{c^s}. \tag{2}$$

Thus, the WA will prefer a workfare policy iff $\gamma < \gamma^s$.

To understand what comes later, it is important to keep in mind that the transfer which H-agents receive is a discontinuous function of γ. It is defined as

$$\beta_H(\gamma) \equiv \begin{matrix} b_H^s(0) > 0 & \text{if } \gamma > \gamma^s, \\ 0 & \text{if } \gamma \leq \gamma^s. \end{matrix} \tag{3}$$

This model contains many interesting insights that we cannot elaborate here (but see Besley and Coate 1992). We just mention that the discontinuity of the rent function (3)—due to the concavity of the cost function—gives the problem a particular feature which is absent in standard dynamic agency problems (such as regulatory problems), as will be seen in the next section.

4. Dynamics and the Problem of Targeting the Poor

So far, we have followed Besley and Coate (1992) and taken it for granted that the information people reveal by opting for a particular poverty program cannot be utilized by the WA later on. Suppose know that the poverty program runs over several periods, and that the WA can learn something about people's earning capacity as time passes. This assumption adds a new dimension to the poverty alleviation problem: The fact that the welfare administrator can collect information about peoples' income opportunities as time passes will be anticipated by potential welfare claimants who will adjust their behavior.

We start by describing the classes of equilibria that exists when the *WA* is unable to commit herself to a particular poverty alleviation program in the future. Next, we discuss the optimality of the different equilibria. In Section 5, we compare the non-commitment case with optimal policy under commitment.

Preferences are taken to be additive across periods, with a 0 rate of discount. Also, the WA uses a zero discount rate to compute intertemporal costs. This choice of discount rate is not crucial to our results, but considerably facilitates the exposition of the arguments. A prerequisite for our analysis is that poverty is to some extent persistent. To simplify, we make the extreme assumption that individual's earning capacities are perfectly correlated over time. We do not allow individuals to save or borrow, for several reasons. First, we want to limit the connection between periods to one stock variable (information). Second, once saving and borrowing is allowed, the definition of the poverty line becomes more fuzzy. Third, it can be regarded as a stylized representation of the poors' imperfect access to capital markets.

4.1. Equilibria: Types and Existence

The simplest framework to discuss long-term poverty alleviation is a game with two periods and four stages. The structure of this game is as follows:

Period 1

Stage 1: The WA designs a first period poverty program $[(b_L^1, c_L^1), (b_H^1, c_H^1)]$.

Stage 2: Individuals decide which package they want to sign up for.

Period 2

Stage 3: The WA is not committed to any prior announcements. Given her updated information on the basis of what she observed in Stage 2, she designs the cost minimizing poverty program $[(b_L^2, c_L^2), (b_H^2, c_H^2)]$.

Stage 4: Individuals decide which package they want to sign up for.

Let γ^2 be the WA's updated belief that an agent who opted for bundle (b_L^1, c^1) in the first period is of type L. We can simplify the game in several respects. First, note that the second period game is just like the static problem but now for a belief γ^2. Second, because the WA has to alleviate poverty also in the first period, she will also set b_L^1 equal to $b_L(c_L^1)$. Third, we claim that if the first period transfers given to H-persons are not too high, L will never want to choose the package intended for H and therefore first period transfers to H will not be made conditional on a work requirement: $c_H^1 = 0$. In the appendix, we give sufficient conditions for this to be verified by the optimal policy. Thus, again, we drop the subscript L on c without any risk of confusion. (see Figure 2).

If H applies in the first period for the bundle $(b_L^1(c^1), c^1)$, he gets $(\beta_H(\gamma^2), 0)$ in the second. On the other hand, should he not register as poor he gets $(b_H^1, 0)$ in the first period and $(0, 0)$ in the second. The values of these two options are $v(b_L(c^1), c^1, a_H) + v(\beta_H(\gamma^2), 0, a_H)$ and $v(b_H^1, 0, a_H) + v(0, 0, a_H)$, respectively. Depending on the magnitude of the

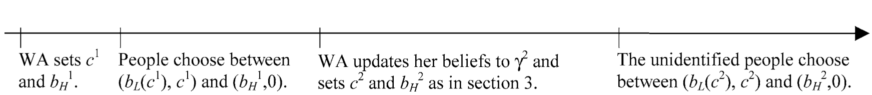

Figure 2: The time sequence in the simplified game

transfers, and the work required, there exists three kinds of equilibria.[9] A *separating equilibrium* in which H-people do not register as poor. To implement such an equilibrium, the WA must either impose extensive work requirements on those who claim poverty transfers, or she must give generous transfers to the non-poor. On the other hand, with very low work requirements associated with poverty transfers and very low transfers to the non-poor, these non-poor clearly prefer to mimic the poor and we have a *pooling equilibrium.* For intermediate values for the two instruments, we may have a *semi-separating equilibrium* in which the non-poor randomize between registering as poor or not.

4.1.1. Separating Equilibrium

We have a separating equilibrium when H prefers not to register as poor even if the WA knows this and is thus convinced that all who do register are genuinely poor (i.e., sets $\gamma^2 = 1$). That is, if

$$v\left(b_H^1, 0, a_H\right) + v\left(0, 0, a_H\right) \geq v\left(b_L(c^1), c^1, a_H\right) + v\left(b_L(0), 0, a_H\right).$$

Separation can be induced either by a welfare policy or by a workfare policy. The lower boundary of (b_H^1, c^1)-values giving rise to a separating equilibrium is found by letting the inequality above bind. Let $b_H^d(c^1)$ be defined as the minimum transfer that induces separating for a first period work requirement c^1, then

$$v\big(b_H^d(c^1), 0, a_H\big) + v(0, 0, a_H) \equiv v\big(b_L(c^1), c^1, a_H\big) + v(b_L(0), 0, a_H). \quad (4)$$

The following Lemma informs about the shape of $b_H^d(c)$ (proven in the Appendix).

LEMMA 2: *The transfer function $b_H^d(c)$ has the following first and second derivatives:*

---

[9]The proper equilibrium concept for this game is perfect Bayesian equilibrium. This means that (P1) the agents make an optimal choice in Period 2 among the packages made available to them by the *WA*; (P2) the *WA*'s design of the second period's program should be optimal, given her updated beliefs; (P3) the choice of the agents in Stage 1 should be optimal given the packages made available by the *WA* in Stage 1 and taking into account the fact that the second period program that is made available to them will depend on the *WA*'s updated beliefs, and therefore on their first period choice; (P4) the *WA*'s choice of program in the first period is optimal given the strategies of the agents and of her own 2nd period strategies; and (B) the *WA* updates her beliefs by observing the participants' first period behavior, thus $\gamma^2 = \text{Prob}(\text{agent is of type } L \mid \text{agent chose in Period 1 the package } [b_L(c^1), c^1])$. In this subsection, we look at *continuation equilibria,* i.e., strategies of the agent in both periods, and of the *WA* in Period 2, and an updating rule, that satisfy P1–P3 and B. See Laffont and Tirole (1993, pp. 380–381). In Section 4.2, we inquire about the optimal choice for the *WA* in Period 1, i.e., impose P4.

$$\frac{db_H^d(c)}{dc} = \frac{v_b^s}{v_b^d}\frac{db_H^s(c)}{dc} < 0$$

$$\frac{d^2 b_H^d(c)}{dc^2} = \frac{(v_b^s)^2}{v_b^d}\left[\frac{v_{bb}^s}{(v_b^s)^2} - \frac{v_{bb}^d}{(v_b^d)^2}\right]\left(\frac{db_H^s(c)}{dc}\right)^2,$$

where v_b^s and v_b^d are shorthands for $v_b(b_H^s(c), 0, a_H)$ and $v_b(b_H^d(c), 0, a_H)$, respectively, and likewise for the second order income derivatives v_{bb}^s and v_{bb}^d.

Concavity of $b_H^d(c)$ is no longer guaranteed by the assumptions we have invoked so far, but can be established with some mild conditions on the risk aversion coefficients. In the sequel, we therefore assume concavity of this transfer function.[10]

With a transfer function that is decreasing and concave in c there again exists a critical value for the work requirement on L-persons, c^d, for which the transfer b_H^d can be reduced to zero while still securing self-selection, i.e., $b_H^d(c^d) \equiv 0$. It is an empirical issue whether c^d exceeds $c^{\max}$ or not. If it does, c^d is not implementable, since that would scare away L-people and make the program meaningless. Then, the best the WA can do is replace it by $c^{\max}$ and leave a positive information rent $b_H^d(c^{\max})$ to H-people.

The following two observations indicate a potential advantage of work requirements to separate to two groups:

1. $b_H^d(0) > 2b_H^s(0)$: If the WA decides to fight first period poverty by using welfare, she must offer H-people *more than twice the amount* she needed to give them in the static case. The reason is that v_{bb} is negative.[11]
2. $c^d < 2c^s$: If she decides to use workfare to scare fraudulent H-people off, she has to impose a higher work requirement than in the static case, but the number of hours that are sufficient to drive H's rent to zero is *less than twice the amount* needed in the static case. The reason is again that v_{bb} is negative.[12]

---

[10]The RHS of (4) can be rewritten as $v(b_H^s(c^1), 0, a_H) + v(b_H^s(0), 0, a_H)$. Since $b_H^s(c^1)$ is concave in c^1, 1st period (and thus intertemporal) utility when mimicking is strictly concave in c^1. At the same time, 1st period (and thus intertemporal) utility when being honest is strictly concave as well in b_H^1. However, if the first mentioned concavity is "strong" compared with the second one, the term $[\frac{v_{bb}^s}{(v_b^s)^2} - \frac{v_{bb}^d}{(v_b^d)^2}]$ will be negative. In the appendix to the working paper, we showed that the sign of this term is given by the sign of $\frac{\mathrm{d}\log R_a}{\mathrm{d}\log m} + R_r$, where R_a and R_r are the coefficients of absolute and relative risk aversion for uncertainty regarding full income m. Decreasing absolute risk aversion and a not too large R_r is thus sufficient for concavity of $b_H^d(c)$.

[11]Evaluating (4) at $c^1 = 0$, and noting that $b_L(0) = b_H^s(0)$, we get that $v(b_H^d(0), 0, a_H) + v(0, 0, a_H) = 2v(b_H^s(0), 0, a_H)$.

[12]Evaluating (4) at $c^1 = c^d$, noting that $v(0, 0, a_H) = v(b_H^s(c^s), 0, a_H)$ and using the alternative formulation for the RHS, we get that $2v(b_H^s(c^s), 0, a_H) = v(b_H^s(c^d), 0, a_H) + v(b_H^s(0), 0, a_H)$.

Since $b_H^s(c)$ is decreasing and concave in c, and $v(b, 0, a_H)$ increasing and strictly concave in b, it follows that $c^d < 2c^s$.

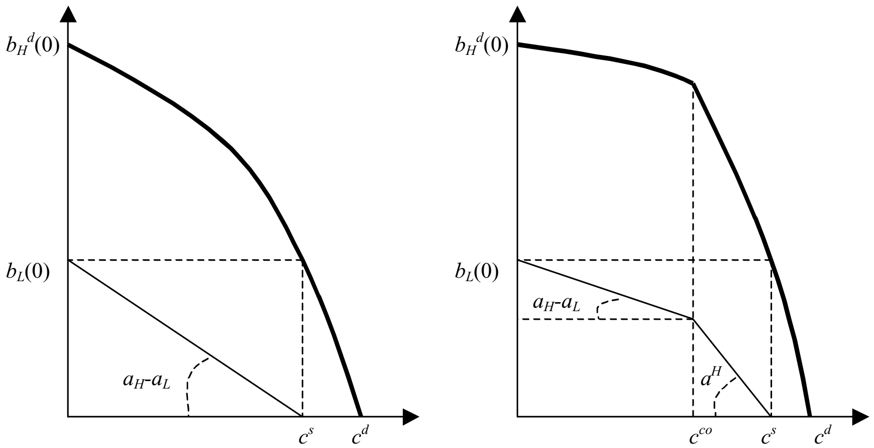

Figure 3: $b_H^d(\cdot)$ (bold) and $b_H^s(\cdot)$ when $c^s < c^{co}$ (left) and $c^s > c^{co}$ (right)

Moreover, $b_H^d(c^s) = b_H^s(0)$, implying that $b_H^d(c)$ everywhere lies above $b_H^s(c)$. Figure 3 shows the relation of $b_H^d(c)$ to $b_H^s(c)$.

With the two groups successfully separated in the first period, the second period policy reduces to the first best type contingent policy: A cash transfer $b_L(0)$ is offered the poor while H-people receive nothing.

4.1.2. Pooling Equilibrium

Clearly, if b_H^1 and c^1 are sufficiently low, an H-person may prefer to mimic the poor even though the WA knows this and therefore set γ^2 equal to γ^1. The condition for a pooling equilibrium is given by the inequality

$$v\big(b_L(c^1), c^1, a_H\big) + v\big(\beta_H(\gamma^1), 0, a_H\big) \geq v\big(b_H^1, 0, a_H\big) + v(0, 0, a_H).$$

The upper boundary for pooling depends on the value γ^1 takes. If $\gamma^1 \geq \gamma^s$, mimicking in the first period generates a welfare policy in the second period and a monetary rent $\beta_H(\gamma^1) = b_H^s(0)$. In this case, we can easily see that the upper boundary of the pooling equilibrium coincides with the lower boundary of the separating equilibrium (since by definition $v(b_H^s(0), 0, a_H) = v(b_L(0), 0, a_H))$. If on the other hand $\gamma^1 < \gamma^s$, we know that pooling in the first period implies workfare in the second period and no second period rent for the non-poor even if they pose as poor in the first period. In that case, pooling occurs when $v(b_L(c^1), c^1, a_H) \geq v(b_H^1, 0, a_H)$, which with equality is the equation for separation in the static model—Equation (1). Hence, when $\gamma^1 < \gamma^s$ there will be an area of (c^1, b_H^1) values that generate neither pooling nor full separation. It is for these values that a semi-separating equilibrium will occur. See the left-hand panel of Figure 4.

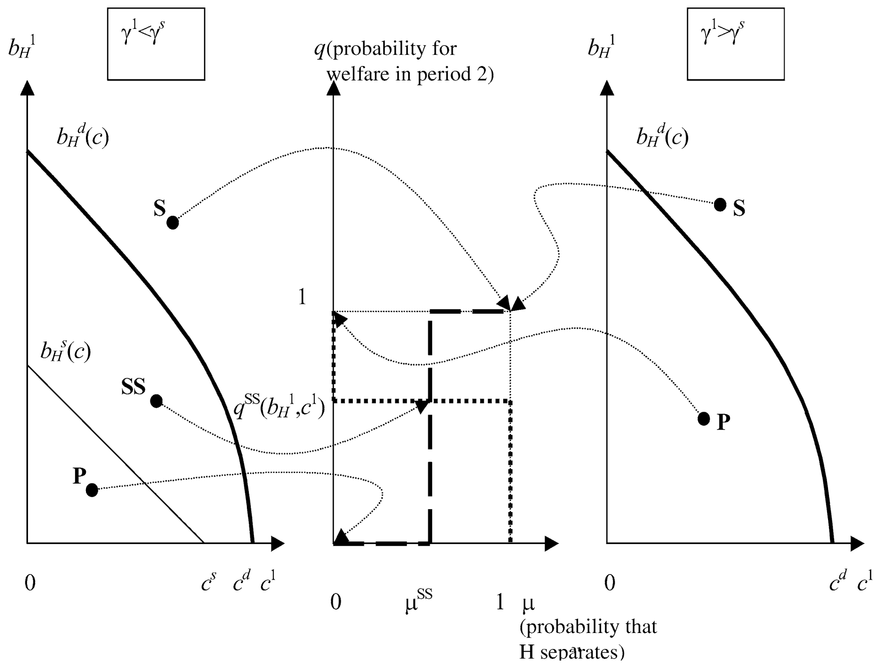

Figure 4: The different continuation equilibria (left, right) and the reaction curves (middle) of the WA (dashed) and *H* (dotted) for the semi-separating equilibrium

4.1.3. Semi-separating Equilibrium (When $\gamma^1 < \gamma^s$)

The third kind of equilibrium requires the following set of inequalities to be fulfilled:

$$\begin{aligned} v\big(b_L(c^1), c^1, a_H\big) + v\big(b_H^s(0), 0, a_H\big) &> v\big(b_H^1, 0, a_H\big) + v(0, 0, a_H) \\ &> v\big(b_L(c^1), c^1, a_H\big) + v\big(\beta_H(\gamma^1), 0, a_H\big). \end{aligned}$$

The LHS is H's utility when mimicking as L when the WA believes everybody is of type L ($\gamma^2 = 1$), while the RHS is utility under mimicking when the WA sets $\gamma^2 = \gamma^1$. Then we claim that there exists a semi-separating equilibrium in which an H-person chooses the bundle intended for him (does not register as poor) with probability

$$\mu^{SS} \stackrel{\text{def}}{=} \frac{\gamma^s - \gamma^1}{(1-\gamma^1)\gamma^s}, \tag{5}$$

and the WA chooses a zero work requirement in the second period (i.e., $c^2 = 0$) with probability

$$q^{SS}\big(b_H^1, c^1\big) \stackrel{\text{def}}{=} \frac{\big[v\big(b_H^1, 0, a_H\big) - v\big(b_L(c^1), c^1, a_H\big)\big]}{\big[v\big(b_H^s(0), 0, a_H\big) - v\big(0, 0, a_H\big)\big]}. \tag{6}$$

To understand this claim, note that if H mimics with probability μ^{SS}, a Bayesian updating WA will believe that among those who opted for poverty transfers in the first period exactly a fraction γ^s are genuinely poor. With such a belief, the WA is indifferent between a workfare and a welfare program in the second period, and therefore willing to randomize between these two policies.[13] A simple computation shows that she must randomize with probability $q^{SS}(b_H^1, c^1)$ in order to make H indifferent between pooling with L-individuals and separating.[14] The semi-separation equilibrium is depicted in the middle part of Figure 4 below.

Let us summarize the facts we have established so far.

PROPOSITION 1: *Depending on the value of* γ^1*, the following equilibria exist: For* $\gamma^1 < \gamma^s$*:*

(i) ***separating equilibrium.*** *H and L are separated in the first period, and a type contingent welfare policy is implemented in the second period;* (b_H^1, c^1) *satisfy* $b_H^1 \geq b_H^d(c^1), 0 \leq c^1 \leq \min\{c^d, c^{\max}\}$*;*

(ii) ***semi-separating equilibrium.*** *H and L are partly separated in the first period, and WA chooses randomly between welfare and workfare in the second period;* (b_H^1, c^1) *satisfy* $b_H^s(c^1) \leq b_H^1 < b_H^d(c^1), 0 \leq c^1 \leq \min\{c^d, c^{\max}\}$*; and*

(iii) ***pooling equilibrium.*** *H and L are not separated in the first period, and a separating workfare program is offered in the second period;* (b_H^1, c^1) *satisfy* $0 \leq b_H^1 \leq b_H^s(c^1), 0 \leq c^1 \leq c^s$*.*

For $\gamma^1 \geq \gamma^s$*:*

(i) ***separating equilibrium.*** *H and L are separated in the first period, and a type contingent welfare policy is implemented in the second period;* (b_H^1, c^1) *satisfy* $b_H^1 \geq b_H^d(c^1), 0 \leq c^1 \leq \min\{c^d, c^{\max}\}$*; and*

(ii) ***pooling equilibrium.*** *H and L are not separated, and universal welfare is offered in the second period;* (b_H^1, c^1) *satisfies* $0 \leq b_H^1 < b_H^d(c^1), 0 \leq c^1 \leq \min\{c^d, c^{\max}\}$.

These different equilibria are depicted in Figure 4 (for the case where $c^d < c^{co}$).

---

[13]That the WA plays a mixed strategy is due to the discontinuity of the rent function (3). In the standard regulation problem, the rent to the efficient firm is continuous in the regulator's belief. Her updated belief in the semi-separating regime is then uniquely given by equating the second period rent to the opportunity cost that the efficient firm has when pooling (cf. Laffont and Tirole 1993, p. 429).

[14]H's utility when pooling and separating are $v(b_L(c^1), c^1, a_H) + (1-q)v(b_L(c^s), c^s, a_H) + qv(b_L(0), 0, a_H)$ and $v(b_H^1, 0, a_H) + v(0, 0, a_H)$, respectively. Since $v(b_L(c^s), c^s, a_H) = v(0, 0, a_H)$ and $v(b_L(0), 0, a_H) = v(b_H^s(0), 0, a_H)$, (6) follows.

4.2. Optimal Poverty Alleviation Programs

Now that we have outlined the continuation equilibrium for an arbitrary first period program (b_H^1, c^1), we have enough information to identify the cost minimizing first period program. The first period policy is made up of two instruments: c^1 hours of work requirement on L, and a cash transfer b_H^1 to H. Both the instruments are costly, but an appropriate use of them can make it more efficient to target transfers to the long-term poor and to economize on second period transfers. When H-persons separate in the first period with probability μ, the cost of the program in that period is

$$K^1\left(c^1, b_H^1, \mu; \gamma^1\right) \stackrel{\text{def}}{=} [\gamma^1 + (1-\gamma^1)(1-\mu)]\, b_L(c^1) + (1-\gamma^1)\mu b_H^1. \quad (7)$$

The first square bracket's term denotes the number of persons displaying type L behavior: The really needy and the fraction of H-persons pretending to be needy. The second term gives the amount of transfers handed over to those H-persons who reveal themselves as non-needy. Since both instruments c^1 and b_H^1 give rise to first period costs, it will be efficient to select them on the lower boundary of each regime. Thus, if separation ($\mu = 1$) is aimed at, the WA should set $b_H^1 = b_H^d(c^1)$ and $c^1 \le \min\{c^d, c^{\max}\}$. An efficient semi-separation policy requires that $b_H^1 = b_H^s(c^1)$. And efficient pooling is obtained when $b_H^1 = 0$ and $c^1 = 0$. Note that an efficient semi-separation policy involves no randomization on the part of the WA since $q^{SS}(b_H^s(c^1), c^1) = 0$ (identically in c^1).

We now turn to second period costs. If the WA randomizes and chooses a welfare policy with probability q in the second period, expected costs are given by

$$\begin{aligned} \mathrm{E}[K^2(\mu, q; \gamma^1)] \stackrel{\text{def}}{=} {} & \gamma^1\left[(1-q)\, b_L(c^s) + q\, b_L(0)\right] \\ & + (1-\gamma^1)(1-\mu)\left[(1-q)\cdot 0 + q\, b_H^s(0)\right], \end{aligned} \quad (8)$$

where (μ, q) take on the values $(1, 1)$ under separation and type-contingent welfare policy, $(\mu^{SS}, 0)$ under (efficient) semi-separation, $(0, 0)$ under pooling and workfare (if $\gamma^1 < \gamma^s$), and $(0, 1)$ under pooling and welfare (if $\gamma^1 \ge \gamma^s$). In this expression, the first square bracket term is the expected transfer which will be handed over to L-persons, while the second square bracket term is the expected amount of money that will be transferred to every H-person that pooled in the first period with the L-types (those H-persons who revealed themselves in the first period—a fraction $(1-\gamma^1)\mu$—receive no transfer at all).

With generic cost functions given by (7) and (8), we can inquire about the kind of equilibrium that ought to be established in the first period, and

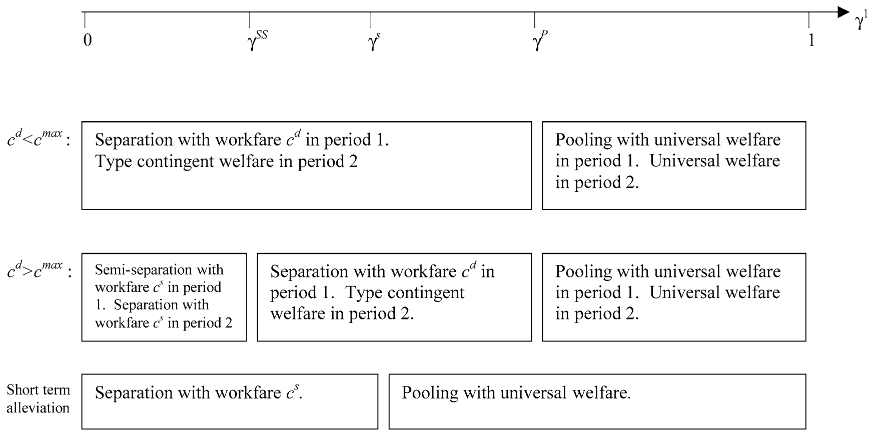

Figure 5: Optimal policy rules for long- and short-term poverty alleviation

how that equilibrium should be implemented. We first define two critical values for γ^1:

1. γ^{SS} makes the WA indifferent between a separation policy with work requirement $\min\{c^d, c^{\max}\}$ and a semi-separation policy with work requirement c^s; and
2. γ^P makes the WA indifferent between a separation policy with work requirement $\min\{c^d, c^{\max}\}$ and a pooling policy with universal transfer $b_L(0)$.

These critical values are given by

$$\gamma^{SS} \stackrel{\text{def}}{=} \frac{b_H^d(\min\{c^d, c^{\max}\})}{b_H^d(\min\{c^d, c^{\max}\}) + \left(1 + \dfrac{1}{\gamma^s}\right) b_L(c^s) + z + b_L(0)}, \tag{9}$$

$$\gamma^P \stackrel{\text{def}}{=} \frac{2b_L(0) - b_H^d(\min\{c^d, c^{\max}\})}{b_L(\min\{c^d, c^{\max}\}) + b_L(0) - b_H^d(\min\{c^d, c^{\max}\})}. \tag{10}$$

We can now formulate the WA's optimal policy rule (illustrated in Figure 5 and proven in the Appendix).

PROPOSITION 2:

(i) The critical γ^1-values can be ranked as follows:

$$0 \leq \gamma^{SS} < \gamma^s < \gamma^P < 1,$$

with $\gamma^{SS} = 0$ if $c^d = \min\{c^d, c^{\max}\}$.

(ii) If $\gamma^1 > \gamma^P$ the most efficient policy is universal welfare inducing pooling. If $\gamma^1 < \gamma^P$ and $c^d < c^{\max}$, the most efficient policy is workfare c^d inducing

separation. However, if $c^d > c^{\max}$, then for a small range of a priori *beliefs $\gamma^1 \in [0, \gamma^{SS}]$ the most efficient policy is semi-separation with workfare c^s.*

Proposition 2 highlights that workfare should be used for a larger range of prior beliefs in the first period of a long-term poverty alleviation program than under short-term poverty alleviation program. This policy, however, is *non-stationary:* once people have been screened, workfare has no longer any role to play and second period transfers are made categorical (a cash transfer to the identified *L*-persons, nothing to the others). The other alternative, which then is used "less often," is a universal welfare policy: A welfare grant $b_L(0)$ is handed out unconditionally to any person who applies for it. In a short-term poverty problem, this is the optimal policy for $\gamma^1 > \gamma^s$. In the long-term problem, γ^1 must exceed γ^P for this to be the efficient policy. As the WA does not learn anything about applicants' types in this case, she enters the second period as uninformed as she was in the first. Because $\gamma^P > \gamma^s$, she continues in the second period to hand out a welfare grant $b_L(0)$ to anybody who asks for it. Put differently, for $\gamma^1 > \gamma^P$ universal welfare is a *stationary* optimal policy.[15] Finally, there is the possibility that the voluntary participation condition on the poor prevents using a high work requirement $(c^d > c^{\max})$. In that case, separation with workfare requires leaving some rent $b^d_H(c^{\max})$ to the *H*-people because voluntary participation of the *L*-people prevents the use of a work requirement c^d. If there are many non-poor around (if γ^1 is very low), the dominant concern is rent reduction. And this can be achieved by a semi-separation policy where a work requirement c^s is imposed in both periods. To see this, note that if exactly $(1 - \gamma^1)\mu^{SS}$ of the non-poor separate in the first period, the WA agrees to impose a work requirement c^s in the second period, and a first period work requirement c^s is sufficient to make the non-poor indifferent between separation and mimicking. Though this policy imposes a higher total work requirement $(c^s + c^s)$ on the poor, it leaves no rents to the non-poor, of which there are many around. In this case, we thus have a stationary policy with a work requirement c^s in each period.

5. Optimal Policy under Commitment

So far, we have analyzed the costs of different transfer programs assuming that the WA cannot commit to a future program. We have assumed she implements

[15]In Schroyen and Torsvik (1999), we showed that when income is observable and means-testing possible, it may happen that for high γ^1 values the pooling policy is dominated by separation without work requirement. With pooling, the WA learns nothing and, if γ^1 is high, will want to separate in the second period without workfare. *H*-people then receive $b_L(0) + b^s_H(0)$. When separating with welfare in the first period, *H*-people receive $b^d_H(0) + 0$. If this amount is less than the former, it pays to separate with welfare in the 1st period. (If income is unobservable, this will never be the case since $b^s_H(0) = b_L(0)$. But with means-testing, it may be the case because $b^s_H(0) < b_L(0)$ as it is more costly for *H* to mimick *L*.)

the second period policy that minimizes costs, given the information she has at that stage. In this section, we characterize the optimal commitment policy and verify how it differs from the time consistent policy when the WA cannot commit.

The “no commitment” assumption prevents a separating policy program from specifying any work requirements or transfers to H-individuals in the second period. Formally, separation and sequential rationality imply $c^2 = 0$ and $b_H^2 = 0$. Repeating the static program is therefore impossible for a WA who operates a program that runs over two periods. Does this constraint increase the overall costs of poverty alleviation? Based on what we know about dynamic screening problems in general, we might expect lack of commitment to be a burden—see, e.g., Laffont and Tirole (1990) for a discussion of commitment problems in a regulation context, and Dillén and Lundberg (1996) for a discussion of the welfare consequences due to lack of commitment in optimal income taxation.

The fact is, however, that lack of commitment causes no additional screening costs *as long as separation by workfare is the cost minimizing policy and* $c^d < c^{\max}$. If the WA imposes a work requirement c^d in the first period and a zero requirement in the second, she is able to separate the two types at a total cost of $\gamma^1[b_L(c^d) + b_L(0)]$. On the other hand, if she implements twice the optimal static workfare policy, she is also able to separate, but now at a total cost of $\gamma^1[b_L(c^s) + b_L(c^s)]$. We know that $c^d < 2c^s$ and since $b_L(c)$ is concave in c, it is optimal to impose work requirements only in one period. Hence, even if the WA can commit to a future policy, and therefore choose $c^2 > 0$, she is better off choosing $c^1 = c^d$ and $c^2 = 0$.[16]

On the other hand, lack of commitment is a potential problem if $c^d > c^{\max}$. To see this, suppose γ^1 is low but still bigger than γ^{SS}. In this case, it is clearly optimal to use work requirements as much as possible, to constrain the rent of the non-poor. The problem is that even a maximal work requirement in the first period implies some rents to the non-poor. If the WA can commit to a second period program, she is better off implementing a program with work requirement c^s in each period, and achieve complete separation without handing out any transfers to the non-poor.

Why are these results opposite to those in the regulation framework? The manager of a regulated firm has a disutility of effort function that is convex. The regulator, who needs to compensate for this disutility of effort out of costly public funds, would therefore like to smooth out the distortion of effort over time. Time consistency, however, forces her to take the entire distortion in

[16]Our assumption of constant productivity (normalized to zero) in the public sector partly drives this result. Indeed, with a decreasing marginal productivity of public work, there would be an argument for smoothing total work requirement over time. If this effect is strong enough, it could counterweigh the concavity of the transfer function and make lack of commitment costly.

the first period. On the other hand, the manager's marginal utility of rent income is assumed constant, and the intertemporal distribution of this rent is thus immaterial. In our model, the compensation of the L-type is concave in the distortion, while the rent seeking type (H) has a decreasing marginal utility of income. Not smoothing out the rent is costly; not smoothing out the distortion is cost effective.[17]

6. Concluding Remarks

We have analyzed how work requirements can be used as a device for targeting transfers to the poor in an environment in which individuals' earning capacities are persistent over time. The welfare administrator can make it less tempting for the non-poor to pose as poor in two different ways. She can increase their utility if they do not join the program by giving them a transfer or she can reduce their utility when joining the program by imposing a work requirement on applicants. A central feature in a dynamic model is that, unless the administrator can commit to a future policy, separation requires type contingent transfers in the second period. Hence all policy measures used to separate the poor from the non-poor must be concentrated early on in the program. We have shown that this increases the effectiveness of workfare as a screening instrument. There is, however, one proviso to this result: In some cases the concentrated use work requirements in the first period exceeds what the poor can bear. In order not to scare them away from the program, the use of work requirements should be spread over time and at the same time the non-poor should be presented with a modest transfer. Though this will no longer result in full separation, it is the best that can be achieved when the number of initially poor is "small." It is only in this latter case, when some information rent must be given to the non-poor, that the welfare administrator would achieve a better result if she could commit to a long-term workfare program.

In this paper, we let individual earnings capacities be fixed over time, and thus ruled out the possibility for poor people to escape poverty in the future by investing today in human capital. In a follow-up paper, we investigate how work requirements may act as sticks and carrots in solving this moral hazard problem, and how the latter interacts with the screening problem studied here. Preliminary results are reported in Schroyen and Torsvik (2001).

[17]For this reason, the WA of a static poverty alleviation program with work requirement c^s could do better by introducing a random work requirement: 0 and c^d, each with probability $\frac{1}{2}$. A similar observation was made by Brito et al. (1991): The desirable effects of randomizing the income tax schedule can be reaped in an intertemporal model by committing to a non-stationary income tax policy.

References

ADAMS, T., and G. DUNCAN (1988) The persistence of urban poverty and its demographic and behavioral correlates, University of Michigan, Survey Research Center, Institute of Social Research.

BESLEY, T. (1996) Political economy of alleviating poverty: Theory and institutions, in *Annual World Bank Conference on Development Economics*, M. Bruno and B. Pleskovic, eds. Washington DC: The Bank, 122–135.

BESLEY, T., and S. COATE (1992) Workfare versus welfare: Incentive arguments for work requirements in poverty-alleviation programs, *American Economic Review* **82**, 249–261.

BESLEY, T., and S. COATE (1995) The design of income maintenance programmes, *Review of Economic Studies* **62**, 187–221.

BRITO, D., J. HAMILTON, S. SLUTSKY, and J. STIGLITZ (1991) Dynamic optimal income taxation with government commitment, *Journal of Public Economics* **44**, 15–35.

DILLÉN, M., and M. LUNDHOLM (1996) Dynamic income taxation, redistribution and the ratchet effect, *Journal of Public Economics* **59**, 69–93.

FREIXAS, X., R. GUESNERIE, and J. TIROLE (1985) Planning under incomplete information and the ratchet effect, *Review of Economic Studies* **52**, 173–191.

GIBSON, J. (2001) Measuring chronic poverty without a panel, *Journal of Development Economics* **2**, 243–266.

HEADY, B., P. KRAUSE, and R. HABICH (1994) Long and short-term poverty—is Germany a 2/3 society?, *Social Indicators Research* **31**, 1–25.

HOFF, K. (1996) Comment on "Political economy of alleviating poverty: Theory and institutions", in *Annual World Bank Conference on Development Economics*, M. Bruno and B. Pleskovic, eds. Washington DC: The Bank, 135–140.

LAFFONT, J.-J., and J. TIROLE (1990) Adverse selection and renegotiation in procurement, *Review of Economic Studies* **57**, 597–625.

LAFFONT, J.-J., and J. TIROLE (1993) *A Theory of Incentives and Procurement in Regulation*. Cambridge, MA: MIT Press.

LIPTON, M., and M. RAVALLION (1995) Poverty and policy, in *Handbook of Development Economics*, Vol. 3B, J. Behrman and T. N. Srinivasan, eds. New York: North-Holland.

MOFFITT, R. (1992) Incentive effects of the US welfare system: A review, *Journal of Economic Literature* **30**, 1–61.

RODGERS, J., and J. RODGERS (1993) Chronic poverty in the United States, *Journal of Human Resources* **28**, 25–54.

SAWHILL, I. (1988) Poverty in the US: Why is it so persistent?, *Journal of Economic Literature* **26**, 1073–1119.

SCHROYEN, F., and G. TORSVIK (1999) Work requirements and long term poverty, Norwegian School of Economics and Business Administration, Department of Economics, Discussion paper 9/99.

SCHROYEN, F., and G. TORSVIK (2001) Sticks and carrots for the alleviation of long term poverty, Norwegian School of Economics and Business Administration, Department of Economics, Discussion paper 34/01.

WEITZMAN, M. (1980) The ratchet principle and performance incentives, *Bell Journal of Economics* **11**, 302–308.