Poverty Alleviation Programs: Monitoring vs. Workfare

Márcia Oliveira and Paulo Côrte-Real

Faculdade de Economia da Universidade Nova de Lisboa

2006

Poverty Alleviation Programs: Monitoring vs. Workfare*

Márcia Oliveira† and Paulo P. Côrte-Real‡

March 2006

Abstract

The role of Poverty Alleviation Programs (PAP) in fighting poverty and ensuring the satisfaction of basic economic needs is well known. However, informational asymmetries create the need for adequate instruments to prevent fraud.

This paper provides a static model of adverse selection where the government (principal) aims to minimize the costs of a PAP that ensures that all individuals have access to an exogenously defined minimum income level.

Agents may differ in their income-generating ability and disutility of labor. Under the different informational environments, we study the effectiveness of workfare (that involves unpaid and unproductive work in the public sector) as a screening device, based on the comparison with standard monitoring.

We find that when disutility of labor is the only unobservable variable, a workfare policy is inefficient because it "crowds out" private sector work and significantly increases the costs of the program. Under this informational context, monitoring may be the best instrument for preventing fraud. When income-generating ability is the only unobservable variable, choosing between workfare and monitoring depends not only on the cost function associated with the latter, but also on income distribution. The analysis of this case would suggest that a workfare policy might be inefficient in the context of undeveloped countries where income distribution exhibits strong inequalities, but appropriate in developed ones. These conclusions suggest that a mixed policy combining workfare and monitoring may be optimal when both income-generating ability and disutility of labor are unknown.

Keywords: Poverty Alleviation Programs, fraud, monitoring, workfare.

*This is a preliminary and incomplete version: please do not quote.

†Faculdade de Economia, Universidade Nova de Lisboa, 1099-032 Lisboa, Portugal. E-mail: marcia@esaelvas.pt.

‡Faculdade de Economia, Universidade Nova de Lisboa, 1099-032 Lisboa, Portugal. E-mail: ppc@fe.unl.pt.

1 Introduction

Risks are part of everyday life. But the impact on the poor and other vulnerable groups, such as the elderly and disabled, are often more immediate and more threatening than those faced by others in society. These risks can be associated with a household (i.e. illness, disability or death, and unemployment), with a community or region (i.e. floods, famine) or a country (i.e. drought, global financial crisis, shifts in terms of trade). The adverse effects of these risks will be far more damaging to the poor than to those better-off in terms of income, physical and mental well-being, and long-term human development. For poor people, lost income may force them to sell their land, their livestock or their tools, send their children to work rather than to school, or eat less. These drastic measures may help families survive from day to day, but they will make it that much harder for these families to escape poverty in the future.

Governments and international financial institutions can play an important role in helping households manage daily risks and cope with losses whenever they occur. The notion that Poverty Alleviation Programs (PAP) should be a permanent feature of social policy and not simply a temporary response to crisis is now increasingly embraced by the international development community. During economic downturns, problems are much more serious and immediate and the appeals for public action are impossible to deny. But even when a country prospers, some households will still face hard times.

Cross-country experience in design and implementation of programs confirms that appropriate program choice, program design, and delivery are key to reaching the poor successfully. Although controversial, the possibility of conditioning entitlement to social assistance on a work requirement for its beneficiaries (workfare) has become more and more accepted in the design of modern welfare systems in Europe and in the United States. The stated reasons for workfare vary. Broadly, there are three main arguments for introducing workfare. The first presents workfare as a screening tool when individuals differ with respect to valuations of leisure or earning capacity. The purpose of screening is, of course, improving the targeting of transfers. A second reason for considering workfare is the deterrent argument: if benefit claimants are required to substitute work for leisure, they may be more eager to leave unemployment. The third motive is directly connected to the first two: the screening and deterrent effects of workfare may also make income transfers to the poor more politically acceptable .

Although workfare itself is not a new concept, only recently has it been developed on a rigorous theoretical basis in public finance literature. Essentially two classes of social objectives are considered, one being conventional "welfarist" and the other being "non-welfarist". With welfarist objectives, the government is only concerned with individual welfare as given by the individuals' preferences. In the usual setup, this implies that the policy maker cares about individuals' consumption of goods as well as their demand for leisure. A utilitarian objective function is the prime example of a welfarist social objective. If the government has non-welfarist objectives, it may give zero weight to the value of leisure and focus exclusively on income.

Workfare models have in general been based on the contributions of Besley and Coate (1992,1995) which provide a detailed analysis of workfare as an income maintenance (or poverty alleviation) program. They introduce a static model of adverse selection in which the primary aim of the social planner is to minimize the costs of a PAP and also ensure that all individuals are above a minimum income level, when income-generating ability is unobservable. Workfare takes the form of a requirement to work in an unproductive public-sector job. They analyse two distinct incentive arguments to justify the use of work requirements in such programs: a screening argument suggesting it may serve as a means of giving transfers only to deserving individuals and a deterrent argument defending it is a device to encourage poverty-reducing investments such as acquiring additional education. In case workfare is implemented, it crowds out some private-sector output by reducing time spent in private-sector work. It is shown that workfare may be part of a cost-minimizing policy when the government is unable to observe wage rates and incomes and also when incomes (but not wage rates) are observable. In a two-class model (Besley and Coate, 1992), the high ability individuals are offered no benefits whereas those claiming to be of low ability are offered an income transfer in exchange for a work requirement. However, those of high ability have no incentive to pretend to be of low ability. In fact, the optimal work requirement is chosen so as to make high-ability individuals indifferent between claiming to be of low ability and receiving no benefit at all. The optimality of workfare is more likely to occur if there is a large wage differential between high and low ability workers and if the fraction of low-ability workers is small relative to the whole target population. When those conditions are met, the crowding-out effect from workfare is modest and the cost-saving from excluding high-ability workers from benefits dominates the crowding-out effect.

The optimality of workfare within a welfarist approach has been examined by Cuff (2000). The individuals in this model differ along two dimensions: ability and disutility of labor. It is shown that it may be part of the optimal package to impose a work requirement on low ability individuals who have a low disutility of labor. Cuff also shows that workfare is never optimal when all individuals have the same ability, unless they are more productive in workfare than in the private sector. Beaudry and Blackorby (1998) derive a similar result. They consider workfare in a utility maintenance program where the objective is to maximize a social welfare function. They also demonstrate that workfare may be welfare-improving if some individuals lack private-sector opportunities. In the latter case, introduction of workfare would not crowd out private-sector output due to reduced labor supply among low-ability individuals.

Tranaes and Thustrup Hansen (1999) adapt the workfare theory to unemployment insurance (UI). They consider the use of work requirements to address moral hazard problems arising from difficulties in monitoring job search efforts. They present a formal analysis of workfare in a model where individuals may have the same productivity but differ with respect to their preferences for leisure. There are two types of individuals referred to as workers (with low disutility of labor) and non-workers (with high disutility of labor). The government knows the distribution of individual characteristics but not the preferences of a particular individual. Job search effort and job acceptance decisions are also private information of the individuals. The paper examines whether workfare can be a Pareto improving welfare reform, i.e., whether it is possible to improve welfare for one type of individuals without worsening conditions for the other type. The answer is affirmative: workfare works as a welfare improving screening device if individuals are sufficiently heterogeneous with respect to their valuations of leisure. By introducing a work requirement, the government can induce non-workers to self-select out of the UI system, the reason being that they have a strong preference for leisure. At the margin, it is possible to simultaneously raise UI benefits and introduce a work requirement so as to make non-workers indifferent between claiming and not claiming UI benefits. The rise in UI benefits represents a strictly positive welfare improvement for workers.

In the contributions discussed above, standard monitoring does not take place as a screening device. The economics literature dealing with monitoring and sanctions in the context of UI is small and of recent origin, and it is non-existing in the context of PAP. There is however a growing literature on optimal law enforcement that is of relevance for the analysis of optimal UI design. Recent surveys of the literature are conducted by Garoupa (1997) and Polinsky and Shavell (2000).

The problem with monitoring is that it is not a free good. The theoretical literature on law enforcement has some fairly clear implications on this matter. The more costly monitoring is, the less should be spent on monitoring activities and the larger should the penalties for commiting fraud be. The crucial difficulty is to quantify the costs of monitoring. No empirical estimates appear to exist in this area.

Our paper intends to provide a static model of adverse selection where the government aims to minimize the costs of a PAP that ensures that each individual in the economy has access to an exogenously defined minimum income level. We extend Besley and Coate's (1992) framework in order to research new issues regarding poverty relief. Under different informational environments, we study the effectiveness of workfare as a screening device based on a comparison with standard monitoring. Furthermore, since disutility of labor (possibly related to physical or intellectual handicaps, distance to the work place or even simple "laziness") represents an important factor in labor decisions, we allow agents to differ not only in their income-generating ability but also in their labor disutility. In what follows, we analyze the optimal PAP when disutility of labor or income-generating ability are unobservable and try to infer results for an environment where both characteristics are unobservable.

Our results suggest that the informational environment may be crucial for the determination of the optimal program. When disutility of labor is the only unobservable variable (for a given income-generating ability), a workfare policy is inefficient because it "crowds out" private sector work increasing the size of the poverty gap and the costs of the program. When only income-generating ability is unobservable, choosing between workfare and monitoring depends not only on the cost function associated with the latter but also on income distribution. The analysis of this case suggests that a workfare policy may be inefficient in the context of countries where most people exhibit low income levels. Under these results, a mixed policy that combines monitoring and workfare will generally be superior to welfare only solutions (giving the minimum income level to all agents in the economy).

The paper is organized as follows. Section 2 describes the model. Section 3 and 4 analyze the potential of, respectively, standard monitoring and workfare as instruments to prevent fraud whenever either disutility of labor or income-generating ability are unobservable. Section 5 concludes.

2 The Model

In what follows we present a multidimensional static model of adverse selection where the principal (government) is concerned with ensuring that each individual gets at least an exogenously defined level of income z, at minimum fiscal cost. Thus, only those who earn less than z should be assisted. Following Besley and Coate (1992), we set aside the revenue-raising implications of the budget required to finance government transfers and focus mostly on poverty-alleviation issues.

We consider an economy consisting of a set I of individuals. Letting the subscript i denote an individual agent in $I = \{1, ..., n\}$, individual agents may differ in their income-generating ability $w_i \in [\underline{w}, \overline{w}]$, and disutility of labor $k_i \in [\underline{k}, \overline{k}]$. Individual types are exogenous and observable to the individual but not to the government. Let $w = (w_i)_{i \in I}$ and $k = (k_i)_{i \in I}$. All variables are independent and $f(w, k)$ is the joint probability density function (pdf). The conditional pdfs are denoted by $f(w)$ and $f(k)$.

For each agent i, let x_i be consumption, $\bar{l}$ total time available for work, l_i total number of hours of work, and let t_i denote non-labor income (possibly a government transfer). We assume individual utilities to be quasilinear in income: $U(x_i, l_i) = x_i - k_i h(l_i)$, where $h(.)$ is strictly increasing and strictly convex. We further assume that $h_l^{-1}(\underline{w}/\overline{k}) = 0$ and that both $\underline{w} h_l^{-1}(\underline{w}/\underline{k}) > z$ and $\overline{w} h_l^{-1}(\overline{w}/\overline{k}) > z$ (i.e. an agent with either the lowest disutility of labor or the highest ability always receives an income that is above z).

Individual i solves the following problem:

$$\begin{aligned} &\max_{x_i, l_i} U(x_i, l_i) = x_i - k_i h(l_i) \\ &\text{s.t. } x_i \leq w_i l_i + t_i \\ &\qquad 0 \leq l_i \leq \bar{l} \end{aligned}$$

It is easy to show that the optimal x_i and l_i are respectively $x_i^* = w_i l_i^* + t_i$ and $l_i^* = h_l^{-1}(\frac{w_i}{k_i})$. Letting $l^*(w_i, k_i) \equiv h_l^{-1}(\frac{w_i}{k_i})$, $l_i^* = l^*(w_i, k_i)$ and we conclude that l^* is increasing in w_i and decreasing in k_i. Notice that $h(.)$ is common to all agents and so is l^*. Moreover, quasilinearity implies that there is no substitution effect and therefore l^* does not depend on t_i.

3 Monitoring

In this section the government aims to minimize the costs of the PAP, ensuring that all individuals in the economy are willing to participate and have access to at least z. Under different informational contexts, we study the effectiveness of standard monitoring used as the only instrument to prevent fraud, based on a comparison with the pure welfare case, where the government hands out z to all individuals in the economy.

The instruments available to the government are a monitoring degree q, where $0 \leq q \leq 1$ is interpreted as the probability of detecting individual fraud, penalties for committing fraud $p = (p_i)_{i \in I}$, and transfers $t = (t_i)_{i \in I}$. The Government does not have the ability to force participation on the PAP. Thus, population must be induced to participate.

The cost of monitoring is defined as $g(q)$. It is assumed that $g(0) = 0$, $g'(q) > 0$, $g''(q) > 0$ and $\lim_{q \to 1} g(q) = \lim_{q \to 1} g'(q) = \infty$. The latter assumption implicitly says that a PAP with $q = 1$ would not be cost-minimizing.

Using t, q and p, the government tries to guarantee that each agent reveals her true type rather than mimicking another one.

3.1 Finding the optimal PAP when k is unobservable

3.1.1 The Complete Information Problem

To illustrate the adverse selection problems arising under incomplete information we start by analyzing the complete information case. In this case, government aims to find the optimal PAP when k is known and w is fixed at $\underline{w}$. Thus, government knows both the vector of individual income-generating abilities (where all elements of the vector coincide) and that of individual parameters of labor disutility before designing the program.

Under this informational context, the problem is:

$$\begin{aligned}
&\min_{t,q,p} \textstyle\sum_i t_i + g(q) \\
&\text{s.t.} \\
&t \geq 0,\ 0 \leq q \leq 1,\ p \geq 0 \\
&(\text{PC}_k)\ wl_i^* + t_i - k_i h\left(l_i^*\right) \geq wl_i^* - k_i h\left(l_i^*\right) \text{ for all } i \Longleftrightarrow t_i \geq 0 \text{ for all } i \\
&(\text{MI}_k)\ wl_i^* + t_i \geq z \text{ for all } i
\end{aligned}$$

PC_k ensures that all individuals are willing to participate in the program and MI_k states that all individuals in the economy have a minimum income of z.

From the individual maximization problem we know that $l_i^* = h_l^{-1}(\frac{w}{k_i})$. With monitoring as the only instrument to prevent fraud, l_i^* remains the same as in the absence of any PAP, and equals the amount of labor for the private sector. The complete information solution is such that $t_i = \max\left\{z - wl_i^*, 0\right\}$. This result shows that the agents who benefit from the program are those with $k_i \in \left[w/h_l(\frac{z}{w}), \overline{k}\right]$ i.e. the agents whose income is below z. The poorest agents

- those with disutility of labor $\overline{k}$ - receive the highest transfer, that equals z. Transfers decrease to zero as k decreases.

It is easy to see that the first-order condition for q is $g\prime(q) = 0$. Therefore, there is no monitoring under complete information.

3.1.2 The Incomplete Information Problem

We now assume that the planner cannot observe k. However, w is common knowledge and it is the same for all agents (and equal to $\underline{w}$). In what follows we simplify the notation and write $l^*(k_i) = l^*(w, k_i)$.

Given that each k_i is private information, the principal needs to design a type-contingent contract: besides determining q, the contract must specify $t_i(k_i)$ and $p_i(k_i)$ for all i such that each agent will have an incentive to announce the true value of k_i in the associated direct revelation game. We therefore need additional constraints to the maximization problem of the previous section: IC_k is the incentive compatibility constraint that ensures truthful revelation and MIP_k guarantees that, even when agents are caught committing fraud, the penalty is such that they still have a final income of at least z[1].

(IC_k) $(1-q)\left\{wl^*(k_i) + t_i(\widehat{k}) - k_i h\left[l^*(k_i)\right]\right\} + q\left\{wl^*(k_i) + t_i(k_i) - p_i(k_i) - k_i h\left[l^*(k_i)\right]\right\} \leq wl^*(k_i) + t_i(k_i) - k_i h\left[l^*(k_i)\right]$ for all $(i, k_i, \widehat{k})$

(MIP_k) $wl^*(k_i) + t_i(k_i) - p_i(k_i) \geq z$ for all (i, k_i)

The principal wants to minimize the expected cost of the program: $E_k(\sum_i t_i(k_i)) + g(q) = \sum_i E_{k_i}(t_i(k_i)) + g(q)$. Rewriting constraints PC_k and MI_k, the problem becomes[2]:

$$\min_{q,\{t_i(k_i),p_i(k_i)\}_{i\in I},} \sum_i E_{k_i}(t_i(k_i)) + g(q)$$

s.t. $t \geq 0$, $0 \leq q \leq 1$, $p \geq 0$

(PC_k) $wl^*(k_i) + t_i(k_i) - k_i h\left(l^*(k_i)\right) \geq wl^*(k_i) - k_i h\left(l^*(k_i)\right)$ for all (i, k_i) $\Longleftrightarrow$ $t_i(k_i) \geq 0$ for all (i, k_i)

(IC_k) $(1-q)\left\{wl^*(k_i) + t_i(\widehat{k}) - k_i h\left[l^*(k_i)\right]\right\} + q\left\{wl^*(k_i) + t_i(k_i) - p_i(k_i) - k_i h\left[l^*(k_i)\right]\right\} \leq wl^*(k_i) + t_i(k_i) - k_i h\left[l^*(k_i)\right]$ for all $(i, k_i, \widehat{k})$

---

[1] This constraint is imposed because the penalty threat must be credible. Otherwise, we could think of an infinite penalty that would ensure truthful revelation, but this penalty would contradict the primary goal of the program: that no individual income is below the minimum survival level z.

[2] If truthful revelation is ensured, the government could use all the information contained in the vector k to determine each t_i. Instead of $t_i(k_i)$,we could consider instead $T_i(k)$ and let $t_i(k_i) = E_{k_{-i}}(T_i(k))$. IC_k and PC_k would then be interim Bayesian constraints where each agent would take into account only $t_i(k_i)$. MI_k and MIP_k would instead be ex post requirements where $T_i(k)$ should be used instead. The versions of MI_k and MIP_k that we use would nevertheless be implied by those ex post requirements. Once we derive the optimal $t_i(k_i)$ in our simplified version of the problem, setting $T_i(k) = t_i(k_i)$ yields the exact same value for the objective function and satisfies all the constraints of the more general problem. We therefore focus only on this simplified version.

(MI_k) $wl^*(k_i) + t_i(k_i) \geq z$ for all (i, k_i)

(MIP_k) $wl^*(k_i) + t_i(k_i) - p_i(k_i) \geq z$ for all (i, k_i)

Given that all agents are equal ex ante, we look for a symmetric solution i.e. we look for functions $t(k_i)$ and $p(k_i)$ such that $t_i(k_i) = t(k_i)$ and $p_i(k_i) = p(k_i)$ for all i.

Before presenting a solution to the problem, notice that our assumption on $g(q)$ suffices to ensure $q < 1$. Notice also that MI_k is implied by MIP_k together with the non-negativity constraint imposed on the penalty: $wl^*(k_i) + t(k_i) \geq wl^*(k_i) + t(k_i) - p(k_i) \geq z$, and so we can ignore MI_k. PC_k is, in turn, equivalent to the non-negativity of t. Finally, notice that the inequality in IC_k may be rewritten as $t(k_i) + \frac{q}{1-q} p(k_i) \geq t(\widehat{k})$.

Lemma 1 *If disutility of labor is unobservable (for a fixed w), a cost-minimizing PAP is such that the maximum transfer is attributed to $k_i = \overline{k}$ and equals z i.e. $t(\overline{k}) = \max_{k_i} t(k_i) = z$.*

Proof. Let k^* be such that $t(k^*) = \max_{k_i} t(k_i)$ and $t(k^*) > z$. Then from IC_k, $t(k_i) + \frac{q}{1-q} p(k_i) \geq t(\widehat{k})$, for all $\widehat{k}$. Since $t(k^*) \geq t(k_i)$ for all k_i, we only need to ensure that $t(k_i) + \frac{q}{1-q} p(k_i) \geq t(k^*) > z$. It follows that $t(k_i) > z - \frac{q}{1-q} p(k_i)$ (1). From MIP_k, $p(k_i) \leq wl^*(k_i) + t(k_i) - z$. Manipulation of this condition yields $z - \frac{q}{1-q} p(k_i) \geq z - \frac{q}{1-q} \left[wl^*(k_i) + t(k_i) - z \right]$ (2). Combining (1) and (2), we have $t(k_i) > z - \frac{q}{1-q} \left[wl^*(k_i) + t(k_i) - z \right]$ and $t(k_i) > z - qwl^*(k_i)$ (for all k_i). Instead let $t(k_i) = \max \{ z - qwl^*(k_i), 0 \}$ and $p(k_i) = wl^*(k_i) + t(k_i) - z \geq 0$. It is straightforward to check that all the constraints are satisfied and for the same q total expected costs are smaller. Therefore, the solution must be such that $t(k_i) \leq z$ for all k_i. Also, from MI_k, $t(\overline{k}) \geq z$ and therefore $t_i(\overline{k}) = z = \max_{k_i} t(k_i)$. ∎

Lemma 2 *If disutility of labor is unobservable (for a fixed w), a cost-minimizing PAP is such that $t(k_i) = \max \{ z - qwl^*(k_i), 0 \}$ for all k_i.*

Proof. Since $t(\overline{k}) = \max_{k_i} t(k_i) = z$, we can rewrite IC_k as $t(k_i) + \frac{q}{1-q} p(k_i) \geq z$ for all k_i. Combined with MIP_k, this implies the requirement that $t(k_i) \geq z - qwl^*(k_i)$ for all k_i. For any given $q \in [0, 1)$, solving a simplified maximization problem subject only to the latter requirement and to the non-negativity constraints yields the result $t(k_i) = \max \{ z - qwl^*(k_i), 0 \}$ for all k_i. Finally, setting $p(k_i) = wl^*(k_i) + t(k_i) - z$ ensures MIP_k and non-negativity of the penalties in the original problem. ∎

Recall that $l^*(k_i) = h_l^{-1}(w/k_i)$. For $0 < q < 1$, let k_0 be the highest level of labor disutility that leads to a 0 transfer i.e. let k_0 be such that $k_0 = w/h_l\left(z/qw\right) < w/h_l\left(z/w\right)$ (and $\underline{k} < w/h_l\left(z/w\right)$ by assumption). From this result we can conclude that every agent i such that $w/h_l\left(z/w\right) > k_i > k_0$ has a private income that exceeds z but receives a transfer from the PAP. The principal is therefore paying informational rents to those agents. Moreover, the principal is paying informational rents to every agent i such that $k_0 < k_i < \overline{k}$ i.e. to all types getting positive transfers (except for $\overline{k}$) . As q goes to 1, k_0 goes to $w/h_l\left(z/w\right)$ and, although monitoring costs increase, the government saves on informational rents. As q goes to 0, so does k_0 and therefore $\underline{k} > k_0$ i.e. all agents with $k_i \in \left[\underline{k}, \overline{k}\right)$ receive an informational rent.

Lemma 3 *If disutility of labor is unobservable (for a fixed w), a cost minimizing PAP includes a positive degree of monitoring if and only if $g'(0) < |I| \, E(wl^*(k_i))$, in which case q is implicitly defined by $g'(q) = |I| \int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} wh_l^{-1}(\frac{w}{k_i})f(k_i)dk_i$.*

Proof. As $q \to 0$, $\frac{w}{h_l\left(\frac{z}{qw}\right)} \to 0 < \underline{k}$. By continuity, there is a $q_0 > 0$ such that $\underline{k} = \frac{w}{h_l\left(\frac{z}{q_0 w}\right)}$.

We can then substitute the expression for $t(k_i)$ in the principal's objective function and rewrite it as

$$\begin{array}{lcl} |I| \displaystyle\int_{\underline{k}}^{\overline{k}} \left[z - qwh_l^{-1}(\frac{w}{k_i})\right] f(k_i)dk_i + g(q) \qquad \text{for } q & \leq & q_0 \\ |I| \displaystyle\int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} \left[z - qwh_l^{-1}(\frac{w}{k_i})\right] f(k_i)dk_i + g(q) \qquad \text{for } q & > & q_0. \end{array}$$

The function is continuous and so is its first derivative

$$\begin{array}{lcl} -|I| \displaystyle\int_{\underline{k}}^{\overline{k}} wh_l^{-1}(\frac{w}{k_i})f(k_i)dk_i + g'(q) \qquad \text{for } q & \leq & q_0 \\ -|I| \displaystyle\int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} wh_l^{-1}(\frac{w}{k_i})f(k_i)dk_i + g'(q) \qquad \text{for } q & > & q_0 \end{array}$$

Notice that, since $f(k_i) = 0$ for $k_i \leq \underline{k}$, the derivative can be rewritten as $-|I| \int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} wh_l^{-1}(\frac{w}{k_i})f(k_i)dk_i + g'(q)$ for $q \in (0,1)$. We thus have an interior solution $q \in (0,1)$ if and only if

$$g'(q) = |I| \int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} wh_l^{-1}(\frac{w}{k_i})f(k_i)dk_i.$$

Given our assumptions, $q = 1$ is never optimal. By continuity, if $g'(0) < \lim_{q\to 0} |I| \int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} wh_l^{-1}(\frac{w}{k_i})f(k_i)dk_i$, then there exists an interior solution.

If, however, $g'(0) \geq |I|\, E(wl^*(k_i)) = |I| \int_{\underline{k}}^{\overline{k}} w h_l^{-1}(\frac{w}{k_i}) f(k_i) dk_i$, then $q = 0$ is optimal and, for $q > 0$, we have $g'(q) > g'(0) \geq |I| \int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} w h_l^{-1}(\frac{w}{k_i}) f(k_i) dk_i$, ruling out an interior solution. ■

Starting from the benchmark case of no monitoring, an infinitesimal increase in q decreases costs, *via* the government transfers, by the expected value of the total income of the agents in the economy. Therefore, in case the economy is very poor, or if the infinitesimal increase in q represents a comparatively very large cost, monitoring will not be an optimal instrument. Naturally, if the former is true, the revenue-raising implications of the budget required to finance government transfers would also prove to be prohibitive. Having the marginal cost associated with achieving an infinitesimally small probability of detecting fraud be larger that the total expected income in the economy is also not very realistic. Therefore, the most interesting case is the one for which $g'(0) < |I|\, E(wl^*(k_i))$.

Under the additional assumption that $g'(0) < |I|\, E(wl^*(k_i))$, we can now combine the Lemmas to fully characterize the optimal PAP:

Proposition 4 *If disutility of labor is unobservable (for a fixed w), the cost-minimizing PAP includes:*

- *a monitoring degree q such that* $g'(q) = |I| \int_{\frac{w}{h_l\left(\frac{z}{qw}\right)}}^{\overline{k}} w h_l^{-1}(\frac{w}{k_i}) f(k_i) dk_i$,
- *and transfers* $t(k_i) = \max\{z - qwl^*(k_i), 0\}$ *for all* k_i.

3.2 Finding The Optimal PAP When W is Unobservable

3.2.1 The Complete Information Problem

In this case, government aims to find the optimal PAP when w is known and k is fixed at $\overline{k}$. Thus, government knows both the vector of individual parameters of labor disutility (where all elements of the vector coincide) and that of individual income-generating abilities before designing the program.

Under this informational context, the problem is:

$\min_{t,q,p} \sum_i t_i + g(q)$
s.t.
$t \geq 0$, $0 \leq q \leq 1$, $p \geq 0$
(PC_w) $w_i l_i^* + t_i - kh\left(l_i^*\right) \geq w_i l_i^* - kh\left(l_i^*\right)$ for all $i \Longleftrightarrow t_i \geq 0$ for all i
(MI_w) $w_i l_i^* + t_i \geq z$ for all i

PC_w again ensures that all individuals are willing to participate in the program and MI_w states that all individuals in the economy have a minimum income of z.

The complete information solution is such that $t_i = \max\{z - w_i l_i^*, 0\}$. Let $\widetilde{w}$ be such that $\widetilde{w} h_l\left(\frac{\widetilde{w}}{k}\right) = z$. Our result shows that the agents who benefit from the program are those with $w_i \in [\underline{w}, \widetilde{w}]$ i.e. the agents whose income is below z. The poorest agents - those with income-generating ability $\underline{w}$ - receive the highest transfer, that equals z. Transfers decrease to zero as w increases. Once more, since the first-order condition for q is $g\prime(q) = 0$, there is no monitoring under complete information.

3.2.2 The Incomplete Information Problem

We now assume that the planner cannot observe w. However, k is common knowledge and it is the same for all agents (and equal to $\overline{k}$). In what follows we simplify the notation and write $l^*(w_i) = l^*(w_i, k)$.

Given that each w_i is private information, the principal needs to design a type-contingent contract: besides determining q, the contract must specify $t_i(w_i)$ and $p_i(w_i)$ for all i such that each agent will have an incentive to announce the true value of w_i in the associated direct revelation game. We again need additional constraints to the maximization problem of the previous section: IC_w is the incentive compatibility constraint that ensures truthful revelation and MIP_w guarantees that, even when agents are caught committing fraud, the penalty is such that they still have a final income of at least z.

(IC_w) $(1-q)\{w_i l^*(w_i) + t_i(\widehat{w}) - kh[l^*(w_i)]\} + q\{w_i l^*(w_i) + t_i(w_i) - p_i(w_i) - kh[l^*(w_i)]\} \leq w_i l^*(w_i) + t_i(w_i) - kh[l^*(w_i)]$ for all $(i, w_i, \widehat{w})$

(MIP_w) $w_i l^*(w_i) + t_i(w_i) - p_i(w_i) \geq z$ for all (i, w_i)

The principal wants to minimize the expected cost of the program: $E_w(\sum_i t_i(w_i)) + g(q) = \sum_i E_{w_i}(t_i(w_i)) + g(q)$. Rewriting constraints PC_w and MI_w, the problem becomes:

$$\min_{q,\{t_i(w_i), p_i(w_i)\}_{i\in I},} \sum_i E_{w_i}(t_i(w_i)) + g(q)$$

s.t. $t \geq 0$, $0 \leq q \leq 1$, $p \geq 0$

(PC_w) $w_i l^*(w_i) + t_i(w_i) - kh(l^*(w_i)) \geq w_i l^*(w_i) - kh(l^*(w_i))$ for all (i, w_i) $\Longleftrightarrow$ $t_i(w_i) \geq 0$ for all (i, w_i)

(IC_w) $(1-q)\{w_i l^*(w_i) + t_i(\widehat{w}) - kh[l^*(w_i)]\} + q\{w_i l^*(w_i) + t_i(w_i) - p_i(w_i) - kh[l^*(w_i)]\} \leq w_i l^*(w_i) + t_i(w_i) - kh[l^*(w_i)]$ for all $(i, w_i, \widehat{w})$

(MI_w) $w_i l^*(w_i) + t_i(w_i) \geq z$ for all (i, w_i)

(MIP_w) $w_i l^*(w_i) + t_i(w_i) - p_i(w_i) \geq z$ for all (i, w_i)

The problem is similar to the one with the unobservable k. Given that all agents are equal ex ante, we again look for a symmetric solution i.e. we look for functions $t(w_i)$ and $p(w_i)$ such that $t_i(w_i) = t(w_i)$ and $p_i(w_i) = p(w_i)$ for all i.

Once more, our assumption on $g(q)$ suffices to ensure $q < 1$. Also, MI_w can be ignored because it is implied by MIP_w together with the non-negativity constraint imposed on the penalty. PC_w is again equivalent to the non-negativity of t. We state without proof the following results (the proofs are similar to those for the case where k is unobservable):

Lemma 5 *If individual income-generating ability is unobservable (for a fixed k), a cost-minimizing PAP is such that the maximum transfer is attributed to $w_i = \underline{w}$ and equals z i.e. $t(\underline{w}) = \max_{w_i} t(w_i) = z$.*

Lemma 6 *If individual income-generating ability is unobservable (for a fixed k), a cost-minimizing PAP is such that $t(w_i) = \max\{z - qw_i l^*(k), 0\}$ for all w_i.*

For $0 < q < 1$, let w_0 be such that $w_0 h_l^{-1}(w_0/k) = z/q$. Recalling that $\widetilde{w}$ is such that $\widetilde{w} h_l^{-1}(\widetilde{w}/k) = z$ then $w_0 > \widetilde{w}$ (and $\overline{w} > \widetilde{w}$ by assumption). From this result we can conclude that every agent i such that $\widetilde{w} < w_i < w_0$ has a private income that exceeds z but receive transfers from the PAP. The principal is therefore paying informational rents to those agents - and indeed to every agent i such that $\underline{w} < w_i < w_0$. As q goes to 1, w_0 goes to $\widetilde{w}$ and, although monitoring costs increase, the government saves on informational rents. As q goes to 0, w_0 becomes arbitrarily large and therefore $w_0 > \overline{w}$ i.e. all agents with $w_i \in (\underline{w}, \overline{w}]$ receive an informational rent.

Lemma 7 *If individual income-generating ability is unobservable (for a fixed k), a cost-minimizing PAP includes a positive degree of monitoring if and only if $g'(0) < |I|\, E(w_i l^*(w_i))$, in which case q is implicitly defined by $g'(q) = |I| \int_{\underline{w}}^{w_0} w_i h_l^{-1}(\frac{w_i}{k}) f(w_i) dw_i$.*

Again, the most interesting case is the one for which $g'(0) < |I|\, E(w_i l^*(w_i))$ and under that additional assumption, we can now fully characterize the optimal PAP:

Proposition 8 *If individual income-generating ability is unobservable (for a fixed k), the cost-minimizing PAP includes:*

- *a monitoring degree q such that $g'(q) = |I| \int_{\underline{w}}^{w_0} w_i h_l^{-1}(\frac{w_i}{k}) f(w_i) dw_i$, where w_0 is implicitly defined by $w_0 h_l^{-1}(w_0/k) = z/q$;*
- *and transfers $t(w_i) = \max\{z - qw_i l^*(w_i), 0\}$ for all w_i.*

4 Workfare

In this section the government aims to minimize the costs of the PAP, ensuring that all individuals in the economy are willing to participate and have access to at least z. Under different informational contexts, we study the effectiveness of a workfare program used as the only instrument to prevent fraud, based on a comparison with the pure welfare case, where the government hands out z to all individuals in the economy. We therefore consider $q = 0$.

Under a worfare system, agent i receives transfer t_i but is required to put in an amount of work r_i in the public sector and therefore $\max\{l_i - r_i, 0\}$ is the time spent working in the private sector. Work in the public sector is assumed to be non-productive i.e. a mere screening instrument that produces net profits of 0. The instruments available to the government are therefore transfers $t = (t_i)_{i\in I}$ and workfare $r = (r_i)_{i\in I}$. Using t and r, the government tries to guarantee that each agents reveals her true type rather than mimicking another one.

4.1 Finding The Optimal PAP When K Is Unobservable

4.1.1 The Complete Information Problem

To illustrate the adverse selection problems arising under incomplete information we again start by analyzing the complete information case. In this case, government aims to find the optimal PAP when k is known and w is fixed at $\underline{w}$. Thus, government knows both the vector of individual income-generating abilities (where all elements of the vector coincide) and that of individual parameters of labor disutility before designing the program.

Under this informational context, the problem is:

$\min\limits_{t,r} \sum_i t_i$
s.t.
$t \geq 0$, $r \geq 0$
(PC_k) $w\max\{l_i^* - r_i, 0\} + t_i - k_i h(\max\{l_i^*, r_i\}) \geq w l_i^* - k_i h(l_i^*)$ for all i
(MI_k) $w\max\{l_i^* - r_i, 0\} + t_i \geq z$ for all i

PC_k ensures that all individuals are willing to participate in the program and MI_k states that all individuals in the economy have a minimum income of z.

From the individual maximization problem we know that $l_i^* = h_l^{-1}(\frac{w}{k_i})$. Workfare does not affect l_i^*. The complete information solution is such that $r_i = 0$ and $t_i = \max\{z - wl_i^*, 0\}$ for all i. This result shows that the agents who benefit from the program are those with $k_i \in \left[w/h_l(\frac{z}{w}), \overline{k}\right]$ i.e. the agents whose income is below z. The poorest agents - those with disutility of labor $\overline{k}$ - receive the highest transfer, that equals z. Transfers decrease to zero as k decreases.

4.1.2 The Incomplete Information Problem

We now assume that the planner cannot observe k. However, w is common knowledge and it is the same for all agents (and equal to $\underline{w}$). In what follows we simplify the notation and write $l^*(k_i) = l^*(w, k_i)$.

Given that each k_i is private information, the principal needs to design a type-contingent contract: the contract must specify $t_i(k_i)$ and $r_i(k_i)$ for all i such that each agent will have an incentive to announce the true value of k_i in the associated direct revelation game. We therefore need an additional constraint to the maximization problem of the previous section: IC_k is the incentive compatibility constraint that ensures truthful revelation.

(IC_k) $w \max\{l^*(k_i) - r_i(k_i), 0\} + t_i(k_i) - k_i h\left(\max\{l_i^*(k_i), r_i(k_i)\}\right) \geq w \max\left\{l^*(k_i) - r_i(\widehat{k}_i), 0\right\} + t_i(\widehat{k}_i) - k_i h\left(\max\left\{l_i^*(k_i), r_i(\widehat{k}_i)\right\}\right)$ for all $(i, k_i, \widehat{k}_i)$

The principal wants to minimize the expected cost of the program: $E_k(\sum_i t_i(k_i)) = \sum_i E_{k_i}(t_i(k_i))$. Rewriting constraints PC_k and MI_k, the problem becomes:

$\min_{\{t_i(k_i), r_i(k_i)\}_{i \in I},} \sum_i E_{k_i}(t_i(k_i))$

s.t. $t \geq 0$, $r \geq 0$

(PC_k) $w \max\{l^*(k_i) - r_i(k_i), 0\} + t_i(k_i) - k_i h\left(\max\{l_i^*(k_i), r_i(k_i)\}\right) \geq w l^*(k_i) - k_i h\left(l^*(k_i)\right)$ for all (i, k_i)

(IC_k) $w \max\{l^*(k_i) - r_i(k_i), 0\} + t_i(k_i) - k_i h\left(\max\{l_i^*(k_i), r_i(k_i)\}\right) \geq w \max\left\{l^*(k_i) - r_i(\widehat{k}_i), 0\right\} + t_i(\widehat{k}_i) - k_i h\left(\max\left\{l_i^*(k_i), r_i(\widehat{k}_i)\right\}\right)$ for all $(i, k_i, \widehat{k}_i)$

(MI_k) $w \max\{l^*(k_i) - r_i(k_i), 0\} + t_i(k_i) \geq z$ for all (i, k_i)

Given that all agents are equal ex ante, we look for a symmetric solution i.e. we look for functions $t(k_i)$ and $r(k_i)$ such that $t_i(k_i) = t(k_i)$ and $r_i(k_i) = r(k_i)$ for all i.

We present a set of lemmas that lead us to our proposition. The first one is a direct consequence of MI_k applied to $\overline{k}$, IC_k with $\widehat{k}_i = \overline{k}$ for all k_i, and cost-minimization.

Lemma 9 *If $r(\overline{k}) = 0$, the optimal PAP attributes z to all individuals and $r(k_i) = 0$ for all k_i.*

Lemma 10 *If the optimal PAP is such that $r(\overline{k}) > 0 = l^*(\overline{k})$, then $r(k_i) \geq r(\overline{k})$ for all $k_i < \overline{k}$.*

Proof. Assume not. Let k_i be such that $r(k_i) < r(\overline{k})$.There are three possible cases:

1. $l^*(k_i) \geq r(\overline{k}) > r(k_i)$

   Then, IC_k for k_i and $\widehat{k}_i = \overline{k}$ and IC_k for $\overline{k}$ and $\widehat{k}_i = k_i$ respectively yield:

$$t(k_i) \geq t(\overline{k}) + w(r(k_i) - r(\overline{k})) \tag{1}$$

$$t(k_i) \leq t(\overline{k}) + \overline{k}(h(r(k_i)) - h(r(\overline{k}))) \tag{2}$$

   Then $wr(\overline{k}) - \overline{k}h(r(\overline{k})) \geq wr(k_i) - \overline{k}h(r(k_i))$. Since $r(\overline{k}) > r(k_i) \geq l^*(\overline{k}) = 0$, and the agent's utility function is concave, we reach a contradiction

2. $r(\overline{k}) > r(k_i) \geq l^*(k_i)$

   Then, IC_k for k_i and $\widehat{k}_i = \overline{k}$ and IC_k for $\overline{k}$ and $\widehat{k}_i = k_i$ respectively yield:

$$t(k_i) \geq t(\overline{k}) + k_i(h(r(k_i)) - h(r(\overline{k}))) \tag{3}$$

$$t(k_i) \leq t(\overline{k}) + \overline{k}(h(r(k_i)) - h(r(\overline{k}))) \tag{4}$$

   But then $(\overline{k} - k_i)(h(r(k_i)) - h(r(\overline{k}))) \geq 0$ and we reach a contradiction.

3. $r(\overline{k}) > l^*(k_i) > r(k_i)$

   Let $k_0 = \sup\left\{k_i : r(\overline{k}) > l^*(k_i) > r(k_i)\right\}$.

   If $k_0 = \overline{k}$, then from MI_k applied to $\overline{k}$, $t(\overline{k}) \geq z$. Let k_i be arbitrarily close to $\overline{k}$. Since $l^*(k_i) > r(k_i)$ and $l^*(k_i)$ is arbitrarily close to 0, so is $r(k_i)$. Also, MI_k applied to k_i implies $t(k_i) \geq z - \varepsilon$, with ε arbitrarily close to 0. But then IC_k for $\overline{k}$ and $\widehat{k}_i = k_i$ is violated: $t(\overline{k}) - \overline{k}h(r(\overline{k})) < t(k_i) - \overline{k}h(r(k_i)) \Leftrightarrow \varepsilon < \overline{k}(h(r(\overline{k})) - h(r(k_i)))$.

   Therefore, $k_0 < \overline{k}$. From (2), we know that $r(k_j) \geq r(\overline{k})$ for $k_j > k_0$ (and for $k_j = k_0$ whenever $k_0 \notin \left\{k_i : r(\overline{k}) > l^*(k_i) > r(k_i)\right\}$. Therefore, in any $\varepsilon-$neighborhood of k_0, we can find k_i and k_j such that $k_i < k_j$ and $r(k_i) < l^*(k_j) < l^*(k_i) < r(k_j)$. Then IC_k for k_i and $\widehat{k}_i = k_j$ and IC_k for k_j and $\widehat{k}_j = k_i$ respectively yield:

$$t(k_i) \geq t(k_j) + k_i(h(l^*(k_i)) - h(r(k_j))) - w(l^*(k_i) - r(k_i)) \tag{5}$$

$$t(k_i) \leq t(k_j) + k_j(h(l^*(k_j)) - h(r(k_j))) - w(l^*(k_j) - r(k_i)) \tag{6}$$

   Then $wl^*(k_i) - k_i h(l^*(k_i)) - (wl^*(k_j) - k_j h(l^*(k_j))) \geq (k_j - k_i)h(r(k_j))$. Letting $v(k_i)$ denote the value function associated with the utility maximization problem, $v(k_i) = wl^*(k_i) - k_i h(l^*(k_i))$ and, from the envelope theorem, $dv/dk_i = -h(.)$. Applying a first-order Taylor expansion of $v(k_i)$ around k_j, we can approximate $v(k_i) - v(k_j)$ by $-h(l^*(k_j))(k_i - k_j) = (k_j - k_i)h(l^*(k_j))$. Since $r(k_j) > l^*(k_j)$, we reach a contradiction.

■

Using the previous lemmas, we can now characterize the optimal PAP:

Proposition 11 *If labor disutility is unobservable (for a fixed w), the cost-minimizing PAP is the pure welfare program:*
- $r(k_i) = 0$ and $t(k_i) = z$ for all i.

Proof. If the optimal PAP is such that $r(\overline{k}) > 0 = l^*(\overline{k})$, then $r(k_i) \geq r(\overline{k})$ for all $k_i < \overline{k}$. From MI_k applied to $\overline{k}$, $t(\overline{k}) \geq z$. But then incentive compatibility implies that $t(k_i) \geq t(\overline{k}) \geq z$ for all $k_i < \overline{k}$ and the pure welfare program would be the cost-minimizing PAP. ■

This analysis suggests that a workfare policy might be inefficient if labor disutility is the prominent cause of low incomes, because it "crows-out" private sector work, increasing the size of the poverty gap and the costs of the program.

4.2 Finding The Optimal PAP When W Is Unobservable

4.2.1 The Complete Information Problem

To illustrate the adverse selection problems arising under incomplete information we again start by analyzing the complete information case. In this case, government aims to find the optimal PAP when w is known and k is fixed at $\overline{k}$. Thus, government knows both the vector of individual parameters of labor disutility (where all elements of the vector coincide) and that of individual income-generating abilities before designing the program.

Under this informational context, the problem is:

$\min\limits_{t,r} \sum_i t_i$
s.t.
$t \geq 0$, $r \geq 0$
(PC_w) $w_i \max\{l_i^* - r_i, 0\} + t_i - kh(\max\{l_i^*, r_i\}) \geq w_i l_i^* - kh(l_i^*)$ for all i
(MI_w) $w_i \max\{l_i^* - r_i, 0\} + t_i \geq z$ for all i

PC_w ensures that all individuals are willing to participate in the program and MI_w states that all individuals in the economy have a minimum income of z.

From the individual maximization problem we know that $l_i^* = h_l^{-1}(\frac{w_i}{k})$. Workfare does not affect l_i^*. The complete information solution is such that $r_i = 0$ and $t_i = \max\{z - w_i l_i^*, 0\}$ for all i. Let $\widetilde{w}$ be such that $\widetilde{w} h_l\left(\frac{\widetilde{w}}{k}\right) = z$. Our result shows that the agents who benefit from the program are those with $w_i \in [\underline{w}, \widetilde{w}]$ i.e. the agents whose income is below z. The poorest agents - those with income-generating ability $\underline{w}$ - receive the highest transfer, that equals z. Transfers decrease to zero as w increases.

4.2.2 The Incomplete Information Problem

We now assume that the planner cannot observe w. However, k is common knowledge and it is the same for all agents (and equal to $\overline{k}$).

Given that each w_i is private information, the principal needs to design a type-contingent contract that must specify $t_i(w_i)$ and $r_i(w_i)$ for all i such that each agent will have the incentive to announce the true value of w_i in the associated direct revelation game. We therefore need an additional constraint to the maximization problem of the previous section: IC_w is the incentive compatibility constraint that ensures truthful revelation.

(IC_w) $w_i \max[l^*(w_i) - r_i(w_i), 0] + t_i(w_i) - kh\max[l^*(w_i), r_i(w_i)] \geq w_i \max[l^*(w_i) - r_i(\widehat{w}_i), 0] + t_i(\widehat{w}_i) - kh\max[l^*(w_i), r_i(\widehat{w}_i)]$ $\quad for \quad all \quad (i, w_i, \widehat{w}_i)$

Rewriting constraints PC and MI, the problem becomes:

$\min\limits_{t_i(w_i), r_i(w_i)} \sum_i E_{w_i} t_i(w_i)$

s.t. $t \geq 0$, $r \geq 0$

(PC$_w$) $w_i \max[l^*(w_i) - r_i(w_i), 0] + t_i(w_i) - kh\max[l^*(w_i), r_i(w_i)] \geq w_i l^*(w_i) - kh[l^*(w_i)]$ *for all* (i, w_i)

(IC$_w$) $w_i \max[l^*(w_i) - r_i(w_i), 0] + t_i(w_i) - kh\max[l^*(w_i), r_i(w_i)] \geq w_i \max[l^*(w_i) - r_i(\widehat{w}_i), 0] + t_i(\widehat{w}_i) - kh\max[l^*(w_i), r_i(\widehat{w}_i)]$ *for all* $(i, w_i, \widehat{w}_i)$

(MI$_w$)$w_i \max(l^*(w_i) - r_i(w_i), 0) + t_i(w_i) \geq z$ *for all* (i, w_i)

We again look for a symmetric solution i.e. functions $t(w_i)$ and $r(w_i)$ such that $t_i(w_i) = t(w_i)$ and $r_i(w_i) = r(w_i)$ for all i.

We present a set of lemmas that lead us to our proposition. The first one is a direct consequence of MI$_w$ applied to $\underline{w}$, IC$_w$ with $\widehat{w}_i = \underline{w}$ for all w_i, and cost-minimization.

Lemma 12 *If* $r(\underline{w}) = 0$, *the optimal PAP attributes* z *to all individuals and* $r(w_i) = 0$ *for all* w_i.

Lemma 13 *If the optimal PAP is such that* $r(\underline{w}) > 0 = l^*(\underline{w})$, *then there must be a* w_i *s.t.* $r(w_i) < l^*(w_i)$.

Proof. If $r(w_i) \geq l^*(w_i)$, for all w_i, then from MI$_w$, $t(w_i) \geq z$ for all w_i, and for all w_i such that $w_i l^*(w_i) > z$, PC$_w$ implies that $t(w_i) > z$. Therefore, welfare would result in a smaller total cost. ∎

Lemma 14 *If* $l^*(w_i) > r(w_i)$, *then for all* $w_i' > w_i$, *we have* $r(w_i') \leq r(w_i)$.

Proof. Assume not. Let $w_i' > w_i$ and $r(w_i') > r(w_i)$. If $l^*(w_i') > l^*(w_i) \geq r(w_i') > r(w_i)$, then IC$_w$ for w_i and $\widehat{w}_i = w_i'$ and IC$_w$ for w_i' and $\widehat{w}_i = w_i$ respectively yield:

$$t(w_i) \geq t(w_i') + w_i(r(w_i) - r(w_i')) \tag{7}$$

$$t(w_i) \leq t(w_i') + w_i'(r(w_i) - r(w_i')) \tag{8}$$

But then $(w_i' - w_i)(r(w_i) - r(w_i')) \geq 0$ and therefore $r(w_i) \geq r(w_i')$.

Now if $w_i' > w_i$ and $r(w_i') > r(w_i)$ where $l^*(w_i') > r(w_i') > l^*(w_i) > r(w_i)$, if r is continuous in $[w_i, w_i']$ then there exists a w_i'' such that $l^*(w_i'') > l^*(w_i) \geq r(w_i'') > r(w_i)$ and we reach a contradiction. If there is a positive discrete jump in r at some w_i'' i.e. $l^*(w_i'' - \varepsilon) > l^*(w_i) \geq r(w_i) \geq r(w_i' - \varepsilon)$ for all $\varepsilon \in (0, l^*(w_i'') - l^*(w_i))$ but $l^*(w_i'') > r(w_i'') > l^*(w_i) > r(w_i)$, then for some ε sufficiently small, $l^*(w_i'') > l^*(w_i'' - \varepsilon) \geq r(w_i'') > r(w_i) \geq r(w_i' - \varepsilon)$ and we again reach a contradiction.

Finally if $w_i' > w_i$ and $r(w_i') > r(w_i)$ where $r(w_i') \geq l^*(w_i') > l^*(w_i) > r(w_i)$, again continuity of r would suffice for a contradiction. If there is a positive discrete jump in r at some w_i'' i.e. $l^*(w_i'' - \varepsilon) > l^*(w_i) \geq r(w_i) \geq r(w_i' - \varepsilon)$ for all $\varepsilon \in (0, l^*(w_i'') - l^*(w_i))$ but $r(w_i'') > l^*(w_i'') > l^*(w_i) > r(w_i)$, then IC$_w$ for w_i and $\widehat{w}_i = w_i''$ and IC$_w$ for w_i'' and $\widehat{w}_i = w_i$ respectively yield:

$$t(w_i) + w_i(l^*(w_i) - r(w_i)) - kh(l^*(w_i)) \geq t(w_i'') - kh(r(w_i'')) \tag{9}$$

$$t(w_i'') - kh(r(w_i'')) \geq t(w_i) + w_i''(l^*(w_i'') - r(w_i)) - kh(l^*(w_i'')) \qquad (10)$$

But then $t(w_i)+w_i(l^*(w_i)-r(w_i))-kh(l^*(w_i)) \geq t(w_i)+w_i''(l^*(w_i'')-r(w_i))-kh(l^*(w_i''))$ and therefore $w_i l^*(w_i) - kh(l^*(w_i)) \geq w_i'' l^*(w_i'') - kh(l^*(w_i''))$. But this contradicts the fact that $l^*(\widetilde{w}_i)$ results from the maximization of $\widetilde{w}_i l - kh(l)$ and that, from the envelope theorem, $\frac{d(\widetilde{w}_i l^*(\widetilde{w}_i) - kh(l^*(\widetilde{w}_i)))}{d\widetilde{w}_i} = l^*(\widetilde{w}_i) > 0$ for all $\widetilde{w}_i > \underline{w}$. ■

Therefore, r is non-increasing for all $w_i > w_0 = \inf \{w_i : l^*(w_i) > r(w_i)\}$. Therefore, for $w_i < w_0$, we have $r(w_i) \geq l^*(w_i)$. Set $r(w_0) \geq l^*(w_0)$ as well (which does not affect our results given that there is a continuum of agents).

Lemma 15 *For all $w_i < w_0$, we have $r(w_i) \geq l^*(w_0)$.*

Proof. We know that $r(w_i) \geq l^*(w_i)$. If $r(w_0) \geq l^*(w_0) > r(w_i) \geq l^*(w_i)$, then IC_w for w_0 and $\widehat{w}_0 = w_i$ and IC_w for w_i' and $\widehat{w}_i = w_0$ respectively yield:

$$t(w_0) - kh(r(w_0)) \geq t(w_i) + w_0(l^*(w_0) - r(w_i)) - kh(l^*(w_0)) \qquad (11)$$

$$t(w_i) - kh(r(w_i)) \geq t(w_0) - kh(r(w_0)) \qquad (12)$$

But then $t(w_i) - kh(r(w_i)) \geq t(w_i) + w_0(l^*(w_0) - r(w_i)) - kh(l^*(w_0))$ and $w_0 r(w_i) - kh(r(w_i)) \geq w_0 l^*(w_0) - kh(l^*(w_0))$, which is only possible for $r(w_i) = l^*(w_0)$ and we reach a contradiction. ■

Again, given that we have a continuum of agents, we set $r(w_0) = l^*(w_0)$ as well (which does not affect our results).

Lemma 16 *In a cost-minimizing PAP all agents such that $w_i \leq w_0$ have the same level of utility.*

Proof. For any $w_i \leq w_0$ and $w_j \leq w_0$, IC_w for w_i and $\widehat{w}_i = w_j$ and IC_w for w_j and $\widehat{w}_j = w_i$ combined yield $t(w_i) - kh[r(w_i)] = t(w_j) - kh[r(w_j)]$. We can then set $t(w_i) - kh[r(w_i)] = c$, where c is a constant, for all $w_i \leq w_0$. ■

Lemma 17 *In a cost-minimizing PAP all agents such that $w_i \leq w_0$ have $r(w_i) = r(w_0)$ and $t(w_i) = t(w_0)$.*

Proof. From PC_w for all $w_i \leq w_0$, $c + kh[r(w_i)] - kh[r(w_i)] \geq w_i l^*(w_i) - kh[l^*(w_i)]$ i.e. $c \geq w_i l^*(w_i) - kh[l^*(w_i)]$ and provided this inequality is met for $w_0 = \arg\max_{\underline{w} \leq w_i \leq w_0} w_i l^*(w_i) - kh[l^*(w_i)]$, it will hold for all $w_i \leq w_0$. Also, it follows that $c \geq 0$.

From MI_w, we just need to ensure that $t(w_i) \geq z$ for all $w_i \leq w_0$.

From our previous result stating that $t(w_i) - kh[r(w_i)] = c$ for all $w_i \leq w_0$, it is clear that IC_w for any $w_i \leq w_0$ and for any $\widehat{w}_i = w_j > w_0$ does not depend on $r(w_i)$. Also, for any $w_j > w_0$ and $w_i \leq w_0$, IC_w for w_j and $\widehat{w}_j = w_i$ yields:

$$w_j [l^*(w_j) - r(w_j)] + t(w_j) - kh[l^*(w_j)] \geq w_j \max[l^*(w_j) - r(w_i), 0] + t(w_i) - kh[\max(l^*(w_j), r(w_i)]$$

Letting $t(w_i) = c + kh(r(w_i))$, if $l^*(w_j) \leq r(w_i)$, this reduces to $w_j\left[l^*(w_j) - r(w_j)\right] + t(w_j) - kh\left[l^*(w_j)\right] \geq c$. If $l^*(w_j) > r(w_i)$, this reduces to $w_j\left[l^*(w_j) - r(w_j)\right] + t(w_j) - kh\left[l^*(w_j)\right] \geq w_j\left[l^*(w_j) - r(w_i)\right] + c + kh(r(w_i)) - kh\left(l^*(w_j)\right)$. The right hand side of this inequality is decreasing in $r(w_i)$: its first derivative is $-w_j + kh'(r(w_i)) < 0$ since $l^*(w_j) > r(w_i)$ and $h'' > 0$. The constraint may only be binding for w_i such that $r(w_i)$ is minimal i.e. for $w_i = w_0$.

We had that for all i such that $w_i \leq w_0$, $r(w_i) \geq r(w_0) = l^*(w_0)$. For cost minimization, we can now set $r(w_i) = r(w_0)$ for all i such that $w_i \leq w_0$ and still meet all the constraints, letting $t(w_i) = c + kh(r(k_0)) = t(w_0)$. ■

Using the previous lemmas, we can now characterize the optimal PAP:

Proposition 18 *Let w_0 be such that $w_0 l^*(w_0) = z$. If income-generating ability is unobservable (for a fixed k), the cost-minimizing PAP consists of:*
- $r(w_i) = l^*(w_0)$ and $t(w_i) = z$ for all i such that $w_i \leq w_0$;
- $r(w_i) = 0$ and $t(w_i) = 0$ for all i such that $w_i > w_0$.

Proof. From IC_w for all $w_i > w_0$ we know $w_i = \arg\max_{\widehat{w}_i > w_0} t(\widehat{w}_i) - w_i r(\widehat{w}_i)$. We can define the maximum function $U(w_i) = \max_{\widehat{w}_i > w_0} t(\widehat{w}_i) - w_i r(\widehat{w}_i)$ and applying the envelope theorem, $\frac{dU}{dw_i} = -r(w_i)$. Then we can rewrite $U(w_i) = U(\overline{w}) + \int_{w_i}^{\overline{w}} r(\widetilde{w}_i) d\widetilde{w}_i$ and $t(w_i) - w_i r(w_i) = U(\overline{w}) + \int_{w_i}^{\overline{w}} r(\widetilde{w}_i) d\widetilde{w}_i$.

$$t(w_i) = U(\overline{w}) + w_i r(w_i) + \int_{w_i}^{\overline{w}} r(\widetilde{w}_i) d\widetilde{w}_i.$$

Given our previous results, the principal's objective function can be rewritten as $|I|\left\{F(w_0)t(w_0) + \int_{w_0}^{\overline{w}} t(w_i) f(w_i) dw_i\right\}$. Replacing $t(w_i)$ by the expression above, and after some manipulation, we obtain

$$|I|\left\{F(w_0)t(w_0) + \left[1 - F(w_0)\right]U(\overline{w}) + \int_{w_0}^{\overline{w}} r(w_i)\left[w_i + \frac{F(w_i) - F(w_0)}{f(w_i)}\right] f(w_i) dw_i\right\}.$$

From PC_w for $w_i \leq w_0$, we know that the only potentially binding condition is $t(w_0) - kh(r(w_0)) \geq w_0 r(w_0) - kh(r(w_0)) \Leftrightarrow t(w_0) \geq w_0 r(w_0)$.

From PC_w for $w_i > w_0$,.we have that $U(\overline{w}) \geq 0$ (it is easy to prove that this constraint is binding and we incorporate it into the problem).

Finally, from MI_w for $w_i \leq w_0$, we have that $t(w_0) \geq z$.

We have already incorporated all the relevant IC_w constraints into the problem and we ignore MI_w for $w_i > w_0$ for now.

The problem thus becomes:

$\min |I|\left\{F(w_0)t(w_0) + \int_{w_0}^{\overline{w}} r(w_i)\left[w_i + \frac{F(w_i) - F(w_0)}{f(w_i)}\right] f(w_i) dw_i\right\}$
s.t. $t(w_0) \geq w_0 r(w_0)$
$t(w_0) \geq z$

Regardless of the value of w_0, since $\left[w_i + \frac{F(w_i) - F(w_0)}{f(w_i)}\right] f(w_i) > 0$ for all $w_i > w_0$, it is optimal to set $r(w_i) = 0$ and therefore $t(w_i) = 0$ for all $w_i > w_0$.

Therefore, $t(w_0) = \max\{z, w_0 r(w_0)\}$. We only need to determine w_0. If $z > w_0 r(w_0)$, it would be optimal to revert to the welfare case. For $z \leq w_0 r(w_0)$, $t(w_0)$ is minimized for w_0 such that $z = w_0 r(w_0) = w_0 l^*(w_0)$ and the total cost of the program is $|I|\, F(w_0)z = |I|\, F(w_0) w_0 r(w_0) < |I|\, z$. Therefore, the optimal w_0 is such that $z = w_0 r(w_0)$. ■

This analysis suggests that, given an exogenous cost of designing and implementing the program, a workfare policy might be inefficient in the context of underdeveloped countries where income distribution exhibits strong inequalities and large number of agents exhibit a low income-generating ability; it may however be appropriate for countries with a fair distribution of income, where there are larger savings due to the introduction of workfare.

5 Conclusion

We explore the potential of workfare as a screening instrument based on a comparison with standard monitoring. The cost of using workfare in this model is that it reduces the amount of labor in the private-sector, increasing the poverty gap and, consequently, the costs of the program. However, it may reduce transfers to the non-poor and this benefit may exceed the cost stemming from the reduced private-sector earnings.

We show that the informational environment is crucial for the determination of the optimal program. When only labor disutility is unobservable, a cost-minimizing PAP does not include workfare. Thus, monitoring may be the best policy, depending on its cost function. However, in the case of very poor economies with limited monitoring ability, it is not an optimal policy. In those situations, the cost-minimizing PAP would then be the pure welfare program. If only income-generating ability is unobservable, the optimal program may use monitoring or workfare. The results depend not only on the monitoring cost function but also on the distribution of income among people in the economy. Our analysis suggests that a workfare policy may be efficient in the context of economies with a reduced number of poor agents, where the savings on informational rents exceed the costs due to lower private-sector earnings.

Our analysis also suggest that the results on workfare achieved in a two-type model by Besley and Coate (1992) are very similar to those of our continuum of types model.

Future research may shed additional light on this comparison for the two-dimensional problem, where both income-generating ability and labor disutility are unobservable. The results we have achieved suggest that a mixed policy combining monitoring and workfare will generally be superior to welfare only solutions (giving the minimum income level to all agents in the economy), justifying the need to further the study of these two instruments.

6 References

Barnes, M (1999). New Deal: Welfare-to-Work in the U.K. University of Essex Department ef Economics, Msc Economics Dissertation

Beaudry P, Blackorby C (1998) Taxes and Employment Subsidies in Optimal Redistribution Program. NBER Working PaperNo. 6355

Besley T, Coate S (1992) Workfare vs Welfare: Incentive Arguments For Work Requirements In Poverty-Alleviation Programs. American Economic Review 82: 249-261

Besley T, Coate S. (1995) The Design of Income Maintenance Programmes. Review of Economic Studies 62: 187-221

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

Cardoso A, Ramos G (2002) Integrated approaches to active welfare and employment policies - Portugal. European Foundation for the Improvement of Living Condition

Cuff K (2000) Optimality of Workfare with Heterogeneous Preferences. Canadian Journal of Economics 33,149-174

Garoupa N (1997) The Theory of Optimal Law Enforcement, Journal of Economic Surveys 11, 267-295

LeBlanc G (1998) Optimal Income Maintenance and The 'Unemployable. Concordia University (Canada), Department of Economics, Working Paper DP9809

Polinsky AM, Shavell S (2000) The Economics Theory of Public Enforcement Law. Journal of Economic Literature 38, 45-76

Ravallion M (1999) The World Bank Research Observer, Vol.14, No.1 (February)

Rodrigues CF (2000) Anti-poverty effectiveness and efficiency of the Guaranteed Minimum Income Programme in Portugal. ISEG/ Universidade Técnica de Lisboa, Working Paper

Schroyen F, Torsvik G (1999) Work Requirements and Long Term Poverty. University of Bergen, Department of Economics, Working Paper 0899

Subbarao K (1997) Public Works as an Anty-Poverty Program: An Overview of Cross-Country Experience. American Journal of Agricultural Economics 79 (May): 78:83

Tranaes T, Hansen CT (1999) OPtimal Workfare in a Society of Workers and Non workers. Manuscript EPRU, University of Copenhagen

Tranaes T, Hansen CT (1999) Optimal workfare in unemployment insurance, Institute of Economics, University of Copenhagen, Working Paper

Wilson S, Fretwell D (1996) Public Service Employment: A Review of Programmes in Selected OECD Countries and Transition Economies. Working Paper 6 on Regional Development Policies. OECD, Paris