When redistribution leads to regressive taxation

Cyril HARITON\textsuperscript{1}  
Gwenaël PIASER\textsuperscript{2}

March 2004

Abstract

We introduce labor contracts, in a framework of optimal redistribution: firms have some local market power and try to discriminate among heterogeneous workers. In this setting we show that if the firms have perfect information, i.e., they perfectly discriminate against workers and take all the surplus, the best tax function is flat. If the firms have imperfect information, i.e., if they offer incentive contracts, then (under some assumptions) the best redistributive taxation is regressive.

Key words: Income Taxation, Redistribution, Labor market, Multi-principals, Adverse selection, Mechanism design.

JEL Classification: D21, D82, H21, L14

\textsuperscript{1}Groupe ESC Toulouse.  
\textsuperscript{2}CORE, Université catholique de Louvain.

This paper is a substantial revision of Chapter 3 of the second author’s unpublished Ph.D. dissertation at the University of Toulouse. We would like to thank Maurice Marchant, Alessandro Pavan, Pierre Pestieau, Pierre Picard, Jean-Charles Rochet, Motohiro Sato, Eric Strobl and Wilfried Zantman for helpful discussions and also seminar audiences in Copenhagen, Lausanne, Liège, Louvain-la-Neuve and Paris. All errors remain, of course, ours.

This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister’s Office, Science Policy Programming. The scientific responsibility is assumed by the authors.
1 Introduction

There have been numerous papers on optimal taxation since the seminal papers by Mirrlees (1971) and Stiglitz (1982). One of the main contribution of these papers is to set down a unified theoretical framework, contract theory, to analyze the interaction between the government and tax payers who are also workers. Workers are employed in a perfectly competitive labor market and, thus, are paid their marginal productivity. The government faces an asymmetry of information about the taxpayers because it does not know their marginal productivity. Therefore, it has to build a direct taxation contract that incitates them to reveal their true private information.

We depart from this literature by introducing an imperfect labor market. By “imperfect markets” we mean mainly two things. First, we consider a monopolistic labor market. Second, we allow firms to discriminate among workers: wages are nonlinear. In the real world labor markets are not always informed and perfectly competitive. Indeed, the opposite assumption, i.e, an imperfect labor market, is consistent with empirical evidence.\textsuperscript{1} Specifically, we focus on a specific form of imperfect markets, which is an extreme of imperfect markets. By doing this we want to question the role of competition in driving the main results derived by the literature on the properties of the marginal direct tax rate and of the optimal direct tax. We keep in mind that real labor markets are neither perfectly competitive nor monopolistic, but we think that considering a different polar case rather than the traditional one helps to understand how taxation is modified when markets are not fully competitive.\textsuperscript{2} The main objective of this paper is to illustrate difficulties for a government in redistributing wealth among worker when its taxation schemes modify the compensation schemes of the firms and vice versa.

The market failure in our model comes from two particular assumptions: market power and lack of information. The firms are price maker and the labor market is segmented in several independent local markets, each firm being a monopsony on one local market. In the most general version of our model, firms do not know the productivity of each workers. As any monopolist constrained by incomplete information, firms must offer revealing contracts to prevent opportunistic behaviors, i.e, they must reward productive workers. By consequence, contracts offered by the firms constrain the government’s taxation: given a taxation schedule, the firms adapt their offers to keep a high wage differential between more and less productive workers. This interaction between income tax and labor contracts is the driving force behind our results.

A paper by Hungerbüelher, Lehmann, Parmentier, and Van der Linden (2003) address the same questions. The authors characterize optimal non-linear income taxation in an economy with a continuum of unobservable productivity levels and endogenous involuntary unemployment due to frictions in the labor markets. They show that redistributive

\textsuperscript{1}See Boal and Ransom (1997) for a survey of the topic.

\textsuperscript{2}In a different framework, Strobl and Walsh (2003) show that if firms are monopsonists in labor markets, a minimum wage may raise hours which are already too high, but has that this ambiguous effects on the number of employees and utility.
taxation distorts labor demand. Moreover, compared to the laissez-faire, gross wages, unemployment, and participation are lower. The results are quite intuitive and common in this field. However, they provide some other meaningful findings: in their model average tax rates are increasing and marginal tax rates are always positive, even at the top.

We depart from this work by considering a different type of market failure. In the latter paper, the labor market is characterized by a matching problem: firms and workers are not equivalent. Workers differ by their type and search for a job corresponding to their type. Firms open type-specific vacancies and each vacancy has to be filled with a single searching worker. Matching workers and vacancies is a time-consuming and costly activity. The “laissez-faire” equilibrium is not efficient and exhibits unemployment. In our model, we do not consider matching problems nor competitive firms. Information revelation occurs thru endogeneous labor contracts. Interestingly, ours results are quite different, which shows how important the assumptions concerning the labor market are.

From a theoretical point of view our model is derived from the standard one. Workers are characterized by private information and their marginal productivity, with a continuous and closed support. Their utility is quasi-linear with a heterogenous parameter on the labor part. On one hand, the labor market is organized by a local monopolist which faces asymmetric information about the workers and maximizes its profits. It proposes labor contracts and produces a good in a competitive final market. On the other hand, the government observes the income from each worker’s contract and imposes a direct tax system in order to maximize a social welfare function. The timing is the following. First, workers discover their marginal productivity. Second, the government imposes a direct taxation scheme. Then, the monopolist proposes a set of contracts. Finally, workers choose their contracts and payments take place.

This work analyses a sequential game with multi-principals where the realization of the contracts occurs at the end of the game. The closest paper in this direction is Martimort (1999), but deals with a pure regulation problem and uses discrete types. Our analysis is made in several steps. In a first part, we keep the assumption of a perfectly competitive labor market in order to find the standard results in our framework. Then, in a second part, we introduce an informed monopolist and exhibit the effect of introducing a monopolist instead of perfect competition in the labor market. Finally, we relax the assumption of perfect information by the monopolist and analyze the game with an uninformed monopolist.

Our main finding is that in this modified setting we can get regressive taxation. By regressive, we mean that the tax paid is decreasing with the income. We call progressive taxation any increasing income tax function. More specifically, we show that if the firms have perfect information, i.e, they perfectly discriminate against workers and take all the surplus, the best tax function is flat. Any progressive taxation would be useless. The firms take all the surplus and each worker receives his reservation utility. As all workers have the same outside option, they get the same level of utility. Redistribution among workers is not needed.

If the firms have imperfect information, i.e, if they offer incentive contracts, then the best redistributive taxation is regressive. It is regressive basically because any progressive
taxation is inefficient. Obviously, a worker has less incentive to work if part of his salary is taxed. Incomplete information gives him the opportunity to work less by misreporting his type. In order to prevent that, the firms will react to any progressive taxation by increasing the range of its salaries and consequently, by paying less unskilled workers. As taxation and compensation policies involve inefficiencies, the resulting after tax incomes are lower than resulting incomes if there were no income tax. On the contrary, and for the same reason, a regressive income tax increases incomes.

The next section introduces the model and the notations. Section 3 presents two benchmarks. Then, the game is solved by backward induction (section 4). Section 5 briefly provides clues on possible extensions and concludes.

2 The model

The model has four different players and two markets. The players are two kinds of households (with one parameter of heterogeneity), the firm and the state. There are two goods: labor and a homogeneous consumption good (with a fixed price $P = 1$), and there is a market for each.

2.1 Workers and tax payers

Households consist of both workers and tax payers. They are heterogeneous with respect to a variable $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$, distributed according to a cumulative distribution function $F(.)$ with density $f(.)$. Each household is characterized by this parameter $\theta$. One can interpret $\theta$ as an arbitrary index, it may be, for example, a name if the name summarizes all the relevant information. The underlying assumption is rather strong: all the differences of the workers can be comprehensively described by a real variable.

More specifically, a worker $\theta$ has two characteristics: a constant disutility of labor $a(\theta)$ (with $a(\theta) > 0$) and a constant marginal productivity $\omega(\theta)$, (with $\omega(\theta) > 0$). To be even more specific, we will always assume that $\dot{a}(\theta) < 0$ and $\dot{\omega}(\theta) > 0$. Thus we do not have two kinds of heterogeneity but only one, $\theta$. We impose such a link on $\omega$ and $a$ in order to avoid the effect arising from “counterveilling incentives”. There is here a clear ranking over types, everybody would like to have a higher $\theta$ if it were possible. As types are randomly distributed, there is in the economy a motivation for redistribution. Note that we consider that the disutility of labor is given for workers and that they are not responsible for this disutility.

If there is no clear ranking of types, i.e. if for some $\theta$ we have $\dot{a}(\theta) > 0$ or $\dot{\omega}(\theta) < 0$, incentives schemes may present counterintuitive properties even in a simple setting. As we want to stress the particular outcomes arising when firms offer contracts, we have to keep the model as simple as possible. Therefore, we will refer to type $\bar{\theta}$ as the “good” agent (high marginal productivity, low disutility of labor) and $\underline{\theta}$ as the “bad” agent (low marginal productivity, high disutility of labor). In order to avoid useless technical difficulties, we
assume that \( f \) is strictly positive and finite over \([\theta, \bar{\theta}]\). It is also helpful to assume that \( a(\theta) > -\infty \), which is quite natural.

The disutility of effort \( a \) as well as the marginal productivity \( \omega \) are private information only observed by each worker \( \theta \) and the repartition function \( F \) is common knowledge.

Each agent \( \theta \) takes two decisions. First, he (or she) can sell a quantity \( L(\theta) \) of labor force in exchange for income \( I(\theta) \) or decide not to work. The compensation scheme \( I(L) \) is taken as given. Second, it pays an income tax \( T(I) \) that yields net income \( R(\theta) = I(\theta) - T[I(\theta)] \). No agent can avoid the payment of this income tax. Finally, it buys a quantity \( Q(\theta) \) of the homogeneous consumption good. From this it derives a utility that we assume to be quasi-linear:

\[
U(\theta, R, L) = V(R) - a(\theta) L,
\]

and faces a budget constraint:

\[
PQ \leq R
\]

with \( V \) increasing and concave.

This class of utility functions is the one considered for example by Lollivier and Rochet (1983), Weymark (1987), Rochet (1991) and Boadway, Cuff, and Marchand (2000). For a given individual the quasi-linearity assumption implies that the marginal rate of substitution is independent of income. This independence allows us to derive some explicit solutions and makes our model tractable.

### 2.2 Firm

The imperfect production sector is modeled through a monopolist in the labor market, with perfect competition in the final good market yielding a price \( P \) normalized to 1. This is in a way an extreme departure from the perfectly competitive labor market, but one should think of this assumption as a situation in which several firms would perfectly compete on the final good market while being locally a monopsony on the labor market. Moreover, it is assumed that the firm is not owned by any household but is an agent “on its own”. Let us say that owners of the firms are agents that do not participate in the labor market but benefit from the profits of the firm.

The firm experiences a constant return-to-scale with respect to labor input \( L(\theta) \) of agent \( \theta \), which produces an output \( Q(\theta) = \omega(\theta) L(\theta) \) for a cost corresponding to the salary paid \( I(\theta) \) (there are no other costly inputs). Thus, the total profit the firms maximizes is:

\[
\Pi(L, I) = \int_{\theta} [\omega(\theta) L(\theta) - I(\theta)] dF(\theta) = \int_{\theta} \pi(\theta) dF(\theta).
\]

We assume that the firm can write contracts on both \( L(\theta) \) and \( I(\theta) \) but not on \( \pi(\theta) \). A first interpretation would be that the firms observe \( L(\theta) \) and \( I(\theta) \), but not \( \pi(\theta) \). This kind of assumption does not seem realistic. We argue that the firm cannot write a contract on \( \pi(\theta) \), not because it cannot be observed by a firm, but because a third party (as a judge)
can observe only the whole profit of the firm and the contribution of the individual \( \theta \). Then contracts based on \( \pi(\theta) \) are not enforcable\(^3\). In the general case (presented in section 4), the firm neither observes the marginal productivity nor the desutility of labor of agent \( \theta \).

The firm’s strategy is the compensation scheme \( I(L) \) as the decision to hire a particular worker. It is analytically convenient to consider the function \( L(I) \) rather than the function \( I(L) \). By doing this, we restrain the firm’s choices: we impose \( I(L) \) to be bijective. This may problematic if one wants to be as general as possible. As it is a quite natural restriction, we do not think that this technical requirement can be affect our results.

### 2.3 Government

The government maximizes social welfare, and to achieve this it has two tools. First, it insures a minimal revenue, in the sense that if an agent does not work, the state provides him with a minimal revenue \( R \) (net of income tax) that yields an utility \( U = V(R) \). Second, it designs an income tax that cannot be escaped by households. We do not consider other tax means (e.g. on the consumption good) because we want to focus on the effect of the imperfection of the labor market on the specific income tax. When designing its taxation contract, the government maximizes a weighted sum of consumers surplus and firm profits:

\[
SW = \int_{\theta}^{\bar{\theta}} U(\theta) \lambda(\theta) dF(\theta) + \int_{\theta}^{\bar{\theta}} \pi(\theta) dF(\theta),
\]

under the budget constraint

\[
\int_{\theta}^{\bar{\theta}} T(I(\theta)) dF(\theta) \geq 0.
\]

By choosing such a welfare function we implicitly assume that the government can take part or all of the profit of the firm by using a undistortive lump-sum taxation. As the lump-sum taxation is not distortive, and as all our results are derived using first the order condition, we do not explicitely introduce this taxation. Our results are unaffected by this simplification, which makes the model a little bit simpler.

The government does not know the disutility of labor and the current income of agent \( \theta \), but it observes the contracts proposed by the firm to the agents, and it can verify whether an agent works or not. To simplify the notation, in the following, we will denote

\[
\lambda(\theta) f(\theta) = g(\theta),
\]

and

\[
G(\theta) = \int_{\theta}^{\bar{\theta}} g(t) dt.
\]

The repartition \( G \) describe the preferences of the government. For example, if \( G \) first order dominates \( F \), this means that the government puts much more weight on the poor than a pure utilitarist government.

\(^3\)An alternative model would be to introduce moral hazard.
The government’s strategy is a couple \((T(\cdot), U)\). We will assume that the function \(T(\cdot)\) is twice differentiable. It allows us to use infinitesimal calculus and standard optimization methods. This is a technical restriction and we do not think that the conclusions of the model would not be affected if one allows for more general taxation schemes.

### 2.4 Timing

The timing is the following. At \(t = 0\), agents discover their private information \(\theta\). At \(t = 1\), the government proposes its taxation contract \(T(I)\) to agents. At \(t = 2\), the firm proposes its labor contract \(L(I)\) to agents. At \(t = 2.5\), agents choose both contracts and at \(t = 3\), all contracts are realized\(^4\) (labor, income, income tax, consumption).

![Figure 1: Timing](image)

This sequential contracting procedure with late contract realization allows us to apply the Revelation Principle on both contracts\(^5\) as there is no decision in between the two contract proposals, basically the revelation principal is valid for the contract designed at time 1. Moreover, as taxation rules are more rigid than labor contracts (the change of a tax by a government is a complicated process in comparison to a change in a labor contract by a firm), the Stackelberg role to the government is natural.

### 3 Benchmarks

#### 3.1 Competitive and informed labor market

In this section we do not allow the firms to be local monopolists, but we assume they are competitive. We compute the optimal taxation in our framework and consider this situation as a benchmark. Except the utility function, the model is close to the one used by Mirrlees (1971). As long as the labor markets are informed and competitive, wages are equal to marginal productivity and profits are equal to zero. Then, gross income is given by:

\[
I(\theta) = \omega(\theta)L(\theta),
\]

The utility function (for any worker \(\theta\)) using the equality between wage and marginal productivity can be written as:

\[
V[R(\theta)] - \frac{I(\theta)}{\omega(\theta)/a(\theta)}.
\]

\(^4\)This setting follows the one used in Martimort (1999).

\(^5\)Calzolari and Pavan (2000) discuss extensively the benefits of a sequential contract of this type.
Note that given our assumptions on utility functions, the ratio \( \omega(\theta)/a(\theta) \) can be interpreted as an 'adjusted productivity'. Workers differ by their marginal productivity \( \omega \) and have different disutilities of labor \( a \). The quasi linearity allows us to use a natural index which is summerized perfectly this two variables.

The government’s budget constraint is denoted by:

\[
\int_{\theta}^{q} [I(\theta) - R(\theta)]dF = 0.
\]

This budget constraint can be rewritten as:

\[
B(\theta) = \int_{\theta}^{q} [I(\theta) - R(\theta)]dF, \\
B(q) = 0.
\]

The optimal taxation is computed using the so-called “optimal revelation mecanism”. The government asks to worker their real type (i.e, their \( \theta \)), and according to their answer, the government gives them an allocation \( (I,C) \). To be optimal, the allocation must induce truthful revelation of types (each worker reports his or her real type), must be a feasible function and must maximize the government’s utility function. A worker \( \theta \) reveals his true type if and only if:

\[
\forall \theta, \tilde{\theta}, V[R(\tilde{\theta})] - \frac{a(\theta)}{\omega(\theta)}I(\tilde{\theta}) \leq V[R(\theta)] - \frac{a(\theta)}{\omega(\theta)}I(\theta),
\]

which with standard manipulations yields the following first order condition:

\[
\dot{U}(\theta) = -\frac{d}{d\theta} \left[ \frac{a(\theta)}{\omega(\theta)} \right] I(\theta).
\]

This necessary condition becomes sufficient if we add a second order condition, which can be rewritten \(^6\) as:

\[
-\frac{d}{d\theta} \left[ \frac{a(\theta)}{\omega(\theta)} \right] I'(\theta) \geq 0.
\]

As we want to focus on the relationship between the compensation and the optimal taxation, we will keep the taxation problem as simple as possible and will assume that the first order condition is sufficient. Since \( \omega(\cdot) \) is an increasing function and \( a(\cdot) \) a decreasing function, this condition is equivalent to the following: 'the function \( I(\theta) \) is non-decreasing’. Then, as long as the contract is incentive compatible, that condition is equivalent to the condition ‘\( R(\theta) \) is non-decreasing’. In the following we will assume that the solution is such that this condition is satified.\(^7\) Thus (as \( \Pi = 0 \)) the program of the government is:

---

\(^6\)See Laffont and Tirole (1993) for details on computations which lead to this second order condition.

\(^7\)This implies that there is no “bunching”, or any workers with different productivities end up with a different after tax income. Lollivier and Rochet (1983) and Boadway, Cuff, and Marchand (2000) provide some conditions under which this approach is fully valid.
\[
\max_{I(.),U(.),B} \left\{ \int_{\theta}^\theta U(\theta) \, dG(\theta) \right\}
\]

subject to:
\[
\begin{align*}
 U(\theta) &= -d\left[ \frac{a(\theta)}{\omega(\theta)} \right] \frac{\omega(\theta)}{d\theta} I(\theta), \quad (\mu) \\
 B(\theta) &= [I(\theta) - R(\theta)] f(\theta), \quad (\eta) \\
 B(\theta) &= B(\bar{\theta}) = 0.
\end{align*}
\]

Lemma 1 The optimal allocation is characterized by:
\[
\begin{align*}
 \forall \theta & \quad \eta(\theta) = \eta \\
 & \quad -d\left[ \frac{a(\theta)}{\omega(\theta)} \right] V'[R(\theta)] \mu(\theta) + \frac{\omega(\theta)}{a(\theta)} V'[R(\theta)] f(\theta) \eta - f(\theta) \eta = 0, \\
 & \quad -\mu(\theta) = g(\theta) - \frac{\eta f(\theta)}{V'[R(\theta)]} - d\left[ \frac{a(\theta)}{\omega(\theta)} \right] \frac{1}{V'[R(\theta)]}, \\
 & \quad \mu(\bar{\theta}) = \mu(\bar{\theta}) = 0.
\end{align*}
\]

From this proposition, we deduce a general property of this class of model. As long as \( f(\theta) \) and \( f(\bar{\theta}) \) are strictly positive, both workers \( \theta \) and \( \bar{\theta} \) receive their first best allocation, i.e, their marginal utility of consumption is equal to their marginal disutility of labor:
\[
V'[R(\theta)] = \frac{a(\theta)}{\omega(\theta)},
\]
\[
V'[R(\bar{\theta})] = \frac{a(\bar{\theta})}{\omega(\bar{\theta})}.
\]

In the following, we assume that these optimal allocations are implemented by an indirect differentiable mechanism \( T(I) \). Then a worker gets the following utility
\[
U(\theta) = V[I - T(I)] - I(\theta) \frac{a(\theta)}{\omega(\theta)}.
\]

The worker \( \theta \) chooses the income \( I(\theta) \) which maximizes his utility. The first order condition can be written as:
\[
[1 - T'(I(\theta))] \omega(\theta) V'[I(\theta) - T(I(\theta))] = a(\theta),
\]
which is the equality between the marginal utility of labor and the marginal disutility of labor. From this, we can deduce the optimal marginal tax and its shape.

Proposition 1 If \( F(\theta) \leq G(\theta) \), the marginal tax rate is non-negative.
The interpretation of the condition $F(\theta) \leq G(\theta)$ is quite straightforward. The government puts more weight on unproductive workers than the real distribution. In some sense we assume that the government has a distributive objective function. The proposition remains true if we assume that the government maximizes the well-known utilitarian criterium: $g \equiv f$.

This proposition is not surprising in this framework. The government has an incentives to redistribute wealth from 'lucky' people to 'unlucky' people. The only difference with a more traditional framework is the new heterogeneity $a(\theta)$. Since $a(.)$ is a decreasing function of $\theta$ and $\omega(.)$ an increasing function of $\theta$, this new variable does not interfere with the standard analysis.

Compared to the 'laissez-faire', the after-tax income remains unchanged for the most productive workers ($\Theta$ workers) and for the less productive workers ($\check{\Theta}$ workers).

![Figure 2: 'Laissez-faire' and Second-Best income distribution](image)

For these two kinds of workers redistribution occurs thru labor: the less productive agents work less compared to the laissez-faire and the more productive agents work more. The levels of consumption remains optimal given their productivity.

### 3.2 Informed Monopolist

We consider that enterprises are local monopsonists in the labor market, but they perfectly know the marginal productivity of each worker. As all perfectly informed monopolists, a firm has no reason to leave any rent to its employees. Thus it chooses $I(\theta)$ and $L(\theta)$ such that:

$$V (I(\theta) - T (I(\theta))) - a(\theta) L(\theta) = U.$$

---

8See Seade (1982) or Myles (1995) for the same result with a more general utility function and worker who differ only with respect to their marginal productivity.
The quasi-linearity gives us a simple relationship between $I(\theta)$ and $L(\theta)$:

$$L(\theta) = \frac{V[I(\theta) - T(I(\theta))] - U}{a(\theta)}.$$  

We can deduce from this the firm’s profit as a function of $I(\theta)$ only:

$$\Pi[I(.)] = \int_{\theta} V\left[\frac{V[I(\theta) - T(I(\theta))] - U}{a(\theta)}\omega(\theta) - I(\theta)\right] dF(\theta).$$

From this expression one can deduce the optimal before-tax income function given the income tax schedule $T(.)$. This optimal income depends on functions $a(.)$ and $\omega(.)$.

To characterize the optimal tax function, we may now proceed to characterize allocations that are constrained by the self-selection constraint and the budget constraint. The self-selection constraint in this setting is derived from the the profit of the firm. As long as the firm knows the $\theta$ of its employees, a worker cannot lie about his type. The only informational problem arises between the firm and the government. The government uses a mechanism which consists of a specific before-tax income $I(\theta)$ and after-tax income $R(\theta)$. These before and after-tax incomes make the firm reveal the type of its employees.

Now, let us suppose that the firm misrepert to the government the type of its employees following the rule $\tilde{\theta}(\theta)$, i.e, it reports to the government $\tilde{\theta}(\theta)$ when it hires a worker who is of the type $\theta$. The profit of the firm is:

$$\Pi[I(.), R(.), \tilde{\theta}(.)] = \int_{\theta} V\left[\frac{V[R(\tilde{\theta}(\theta))] - U}{a(\theta)}\omega(\theta) - I(\tilde{\theta}(\theta))\right] dF(\theta).$$

Then the firm tells the truth if:

$$\forall \theta \in \Theta \quad V'[R(\theta)] \frac{\omega(\theta)}{a(\theta)} R'(\theta) - I'(\theta) = 0$$

This condition is equivalent to the following:

$$\pi(\theta) = \frac{d}{d\theta} \left[\frac{\omega(\theta)}{a(\theta)} (V[R(\theta)] - U)\right],$$

where:

$$\pi(\theta) = \frac{V[R(\theta)] - U}{a(\theta)}\omega(\theta) - I(\theta).$$

The second order condition imposes that $R(\theta)$ is a non-decreasing function. As in the previous section, we neglect this condition and assume that our solution fulfills this condition.\(^9\)

\(^9\)Using classical methods, one can show that we should add the constraint $R'(\theta) \geq 0$ to get a sufficient condition. Moreover the optimal taxation given by Proposition 3 leads to a net income function which follows this condition (as shown in appendix).
The government maximizes:
\[
\max_{\theta \in \Theta, R(\cdot), U(\cdot)} \int_{\theta}^{\bar{\theta}} U dG(\theta) + \int_{\theta}^{\bar{\theta}} \pi(\theta) dF(\theta),
\]
under the constraints:
\[
\begin{align*}
\pi(\theta) &= \frac{d}{d\theta} \left[ \frac{\omega(\theta)}{\pi(\theta)} \right] (V[R(\theta)] - U), \\
\pi(\theta) &\geq 0, \\
B(\theta) &= [I(\theta) - R(\theta)] f(\theta), \\
B(\bar{\theta}) &= B(\theta) = 0,
\end{align*}
\]
where \( \mu \) and \( \eta \) are lagrangian multipliers. In the following we will distinguish between two kinds of solutions: the solution such that \( \pi(\theta) > 0 \), and the solution such that \( \pi(\theta) = 0 \).

Secondly, if we write down the Hamiltonian of this program,
\[
\Lambda_1 = U g(\theta) + \pi(\theta) f(\theta) + \mu(\theta) \frac{d}{d\theta} \left[ \frac{\omega(\theta)}{\pi(\theta)} \right] (V[R(\theta)] - U) + \eta(\theta) [I(\theta) - R(\theta)] f(\theta).
\]
From the definition of \( \pi(\theta) \), we can deduce:
\[
I(\theta) = \frac{V[R(\theta)] - U}{\omega(\theta)} \omega(\theta) - \pi(\theta).
\]

**Lemma 2** The optimal allocation must satisfy the following condition:
\[
\begin{align*}
-\eta(\theta) &= 0, \\
-\mu(\theta) &= f(\theta) - \eta(\theta) f(\theta), \\
\mu(\bar{\theta}) &= \mu(\theta) = 0, \\
\mu(\theta) \frac{d}{d\theta} \left[ \frac{\omega(\theta)}{\pi(\theta)} \right] V'[R(\theta)] &= -\eta(\theta) \left[ V'[R(\theta)] \frac{\omega(\theta)}{\pi(\theta)} - 1 \right] f(\theta).
\end{align*}
\]
From this first order conditions we can deduce the shape of the income taxation.

**Proposition 2** The government can achieve the optimal social welfare with a uniform taxation.

Any progressive taxation would be useless in this setting. The monopsony takes all the surplus: whatever his type and the taxes, a worker receives reservation utility \( U \). Thus the only feasible income taxation must be flat and redistribution among workers is not needed.

We have made the assumption that the firm hires all the workers. This assumption is necessary to ensure that the function \( T(\cdot) \) (or the function \( R(\cdot) \)) is differentiable, and write down the incentive constraint of the firm. If it is not the case, the problem is more technical.
4 Uninformed Monopolist

4.1 Labor Contract

In order to solve the game backwards, the problem of the rm has to be solved for a given tax schedule $T(I)$. It is assumed that the rm has an incentive to propose a contract to all types of workers. Moreover, the Revelation Principle holds at $t = 2$ because $T(I)$ is fixed and there is no strategic use of information after the offer of the firm.

Define $U(\theta, \tilde{\theta})$ as the utility that agent $\theta$ gets when he (or she) is of type $\theta$ but announces $\tilde{\theta}$ to the rm:

$$U(\theta, \tilde{\theta}) = V[I(\tilde{\theta}) - T(I(\tilde{\theta}))] - a(\theta)L(\tilde{\theta}).$$

Thus, the program of the rm is:

$$\max_{L(\cdot), I(\cdot)} \left\{ \int_\theta \omega(\theta)L(\theta) - I(\theta) \right\}$$

subject to the incentive compatibility and individual rationality constraints

$$V[I(\theta) - T(I(\theta))] - a(\theta)L(\theta) \geq [I(\tilde{\theta}) - T(I(\tilde{\theta}))] - a(\theta)L(\tilde{\theta}),$$

$$V[I(\theta) - T(I(\theta))] \geq U.$$

One can rewrite the incentive compatibility constraint. This condition has to be verified for every $\theta$ and $\tilde{\theta}$, which yield the following first order condition.

$$V[I(\theta) - T(I(\theta))] \left[ 1 - T'(I(\theta)) \right] - a(\theta)L'(\theta) = 0.$$

Using the total differential of $U$ with respect to $\theta$, the first order equation is equivalent to

$$\hat{U}(\theta) = -\hat{a}(\theta)L(\theta),$$

which means that the utility the firm should give to an agent is increasing in its type. The second order condition is assumed to be verified. Using classical methods, as in the previous section, it is sufficient to add the constraint $\hat{L}(\theta) \geq 0$ or equivalently $I(\theta) \geq 0$. In the appendix we argue that if the quantity denoted by $K(\cdot)$ and defined in the following, is increasing, then $I(\cdot)$ is also increasing.

Moreover, one can rewrite labor as a function of utility and income

$$L(\theta) = \frac{1}{a(\theta)}[V[I(\theta) - T(I(\theta))] - U(\theta)].$$

Finally, as giving extra utility to agents is costly for the firm, it will make the individual rationality constraint binding by giving $U$ to the inefficient agent $\tilde{\theta}$. Thus, the program of

\footnote{This explains the assumption on the structure of the utility function. If the quasi-linearity has been taken with respect to the income -which would be more satisfactory in a sense- then the labor becomes an implicit function of income and utility. This complicates the analysis at the second step and, as a consequence, makes the overall study of the optimal taxation more complicated.}
the firm can be rewritten as a dynamic optimization problem with \( I \) as control variables and \( U \) as the state variable

\[
\max_{U(\cdot), I(\cdot)} \left\{ \int_{\Omega} \left[ \frac{\omega(\theta)}{a(\theta)} \left[ V \left( I(\theta) - T(I(\theta)) \right) - U(\theta) \right] - I(\theta) \right] dF(\theta) \right\},
\]

subject to the constraints

\[
\dot{U}(\theta) = -\frac{\hat{a}(\theta)}{a(\theta)} \left[ V \left( I(\theta) - T(I(\theta)) \right) - U(\theta) \right] \quad \forall \theta
\]

\[
U(\theta) = U.
\]

The following proposition describes the optimal compensation policy of the firm given the taxation rules.

**Lemma 3** Let \( I(\cdot) \) solve, for a given tax schedule \( T(I) \), the following problem

\[
\max_{I(\cdot)} \left\{ \frac{\omega(\theta)}{a(\theta)} \left[ 1 + \frac{\hat{a}(\theta)}{a(\theta)} \frac{H(\theta)}{\omega(\theta)} f(\theta) \right] V \left( I(\theta) - T(I(\theta)) \right) - I(\theta) - \frac{\omega(\theta)}{a(\theta)} U \right\}
\]

where

\[
H(\theta) = \int_{\theta}^{q} \omega(\theta) f(\theta) d\theta
\]

and \( L(\theta) \) is given by

\[
L(\theta) = \frac{1}{a(\theta)} \left[ V \left( I(\theta) - T(I(\theta)) \right) - U(\theta) \right].
\]

Note that the labor contract implemented by the firm depends on the tax the government is to implement. This reflects the fact that this latter retains ultimate control on the final outcome by moving first. In the appendix we discuss the validity of the second order conditions of this problem and their links with the optimal taxation.

In order to simplify the discussion we need to introduce some notation:

\[
K(\theta) = \frac{\omega(\theta)}{a(\theta)} \left[ 1 + \frac{\hat{a}(\theta)}{a(\theta)} \frac{H(\theta)}{\omega(\theta)} f(\theta) \right].
\]

To induce revelation, the firm must offer a contract that satisfies the incentive constraint. This induces a decrease in the marginal benefit the firm can enjoy from an agent \( \theta \): the term \( K(\theta) \) can be divided into parts. First, \( \omega(\theta)/a(\theta) \) corresponds to “adjusted productivity” of the agent \( \theta \), which has a clear interpretation. The second term, \( \frac{\hat{a}(\theta)}{a(\theta)} \frac{H(\theta)}{\omega(\theta)} f(\theta) \), corresponds to the cost induced by truthful revelation by the agent. Then we can interpret \( K(\theta) \) as the ‘apparented productivity’ of the agent \( \theta \) for the firm. If \( K(\theta) = \frac{\omega(\theta)}{a(\theta)} \), i.e., if \( \frac{\hat{a}(\theta)}{a(\theta)} \frac{H(\theta)}{\omega(\theta)} f(\theta) = 0 \), it means that the asymmetric information has no impact on firm behavior, information revelation is costless, and for a given tax function \( T(\cdot) \), the compensation is optimal, and the redistribution analysis is similar to the one presented in
the previous section. On the contrary, if $K(\theta)$ is equal to 0, i.e., if \( \frac{\hat{a}(\theta)}{a(\theta)} \cdot \frac{H(\theta)}{\omega(\theta) f(\theta)} = -1 \), the information is so costly that the the worker $\theta$ is 'apparently unproductive'.

To keep the analysis meaningfull, we assume that the 'apparented productivity' is always greater than zero and smaller than the 'adjusted productivity', formally:

$$\forall \theta \quad \frac{\omega(\theta)}{a(\theta)} \geq K(\theta) > 0.$$  

Given their definitions, $\hat{a}(\cdot)$ is negative and $H(\cdot)$ is positive. The first inequality is implied by our definitions. We also assume that every worker remains productive when we introduce asymmetric information.\(^{11}\)

Compared to the competitive case, the discriminating monopolist enlarges the range of incomes. The less productive workers are paid less (and work less), their total income dramatically decreases and their 'apparented productivity' is quite low compared to their real adjusted productivity. The monopolist wants to discourage productive workers to mimick unproductive workers. Thus, the resulting shape of income is a consequence of the lack of information and exibits the usual properties of a second best mechanism.

![Figure 3: 'Laissez-faire', Second-Best, and monopolistic income distribution](image)

**Figure 3:** 'Laissez-faire', Second-Best, and monopolistic income distribution

The next lemma describes the effect of a progressive tax on the compensation scheme, and gives the intuition of the next proposition.

**Lemma 4** Let $I^*(\theta)$ be the optimal compensation scheme for a given tax schedule $T(\cdot)$, and let $\tilde{I}^*(\theta)$ be the optimal compensation scheme if there is no income tax (i.e., if $T \equiv 0$). If $T$ is an increasing function, then

$$\forall \theta \in \Theta, \quad I^*(\theta) - T[I^*(\theta)] < \tilde{I}^*(\theta).$$

\(^{11}\)We assume that all the productivities are strictly positive, which is technically restrictive, but allow the 'apparented productivities' to be negative would introduce difficulties which are out of the reach of this paper.
The 'laissez-faire' best contract offer is not feasible anymore when the government introduces income taxation. The income taxation modifies the level of income, but also the shape of the compensation scheme. The firm is then unable to reproduce the former optimal scheme. To do that, it would be necessary to satisfy the following conditions for every $q$:

$$I^*(\theta) - T[I^*(\theta)] = \tilde{I}^*(\theta),$$

and the first order conditions must be equivalent as well,

$$(1 - T'[I^*(\theta)]) V'(I^*(\theta) - T[I^*(\theta)]) = V'(\tilde{I}^*(\theta)).$$

If $T'$ is different from 0, these two conditions are incompatible. The firm can choose a compensation scheme such that the after-tax incomes are unmodified, but this scheme for the firm itself would not be equivalent to the former compensation scheme.

If a government wants to reduce the inequalities induced by such a compensation scheme, it may want to increase the total income of the less productive workers. In order to do that, it cannot use a progressive taxation: by definition a progressive taxation reduces income inequalities. So, any progressive taxation decrease makes labor more expensive for the firm. Roughly speaking, a worker has less incentive to work if part of his salary is taxed. Incomplete information gives him the opportunity to work less by misreporting his type. In order to prevent that, the firm will react to any progressive taxation by increasing the range of its salaries and, consequently, by paying less unskilled workers. The preceeding lemma shows that the after-tax incomes become lower.

If the government 'taxes the poors', roughly speaking, it becomes less difficult to make workers work. As a consequence, workers are paid more compared to the 'laissez-faire'.

![Figure 4: Effect of a regressive taxation on monopolistic income distribution](image-url)
Our intuitive argument relies on the income distribution and it is not welfare based. In the next subsection we show that this argument fully applies even if we consider a utilitarian government. The main reason is that by doing nothing (or by using a progressive taxation) a government allows the firm (or encourages it) to increase inequalities.

### 4.2 Optimal taxation

We assume that the firm can decide to shut down in case it can not make positive profits with an incentive contract. This is to ensure that the government is not intended to indirectly tax the firm by its income taxation contract, making it hard for the firm to continue to propose a labor contract to all types of agents.

As stated by Martinort (1999), there is no restriction to use a direct revelation mechanism at \( t = 1 \), provided that a "coalition" between the firm and the agents wants to truthfully reveal to the government the agent’s type. The government anticipates the contract offered by the firm and acts as if the firm was an agent who has private information.

The optimal lie is to tell the truth, i.e \( \tilde{\theta}(\theta) = \theta \) if: \( \Pi'(\theta) = 0 \), which can be written as:

\[
K(\theta) V'[R(\theta)] R'(\theta) - I'(\theta) = 0.
\]

As usual, this constraint can be rewritten through a first order equation:

\[
\hat{\pi}(\theta) = K(\theta) V [R(\theta)] - \frac{\omega(\theta)}{a(\theta)} U.
\]

Where:

\[
\pi(\theta) = K(\theta) V [R(\theta)] - I(\theta) - \frac{\omega(\theta)}{a(\theta)} U.
\]

Moreover, the budget constraint can be rewritten, using \( B(\theta) = \int_{\theta}^{\tilde{\theta}} [I(\theta) - R(I(\theta))] dF(\theta) \), as \( B(\theta) = [I(\theta) - R(I(\theta))] f(\theta) \), with \( B(\theta) = B(\tilde{\theta}) = 0 \). To simplify to computation and the notation, we assume that the government has a pure utilitarian social welfare function: \( G \equiv F \). The program of the government becomes:

\[
\max_{\pi(\cdot), R(\cdot), B(\cdot)} \int_{\theta}^{\tilde{\theta}} U(\theta) dG(\theta) + \int_{\theta}^{\tilde{\theta}} \pi(\theta) dF(\theta),
\]

under the constraints:

\[
\forall \theta \quad \pi(\theta) = K(\theta) V [R(\theta)] - \frac{\omega(\theta)}{a(\theta)} U,
\]

\[
\forall \theta \quad B(\theta) = [I(\theta) - R(I(\theta))] f(\theta),
\]

\[
B(\theta) = B(\tilde{\theta}) = 0,
\]

\[
\forall \theta \quad \pi(\theta) \geq 0.
\]

In order to solve this program, we need \( \pi \geq 0 \), and then we assume another property of \( a(\cdot), \omega(\cdot), f(\cdot) \) and \( g(\cdot) \):

\[
K(\theta) \geq \frac{\omega(\theta)}{a(\theta)}.
\]
The ‘apparented productivity’ must sufficiently increase with $\theta$. In order to get a mono-
tonic compensation rule (high skilled workers are much more paid than unskilled work-
ers), the adjusted productivity must by increasing. But this is not enough, it must be
’sufficiently increasing’. The firm (given the taxation) will offer individually rational
contracts, then we can anticipated that $V(R(\theta)) \geq U$. We can deduce from that $\pi(\theta) > 0$.
Thus the only relevant participation constraint is:

$$\pi(\theta) \geq 0.$$ 

As in the previous section the program has two important features: the optimal taxa-
tion may not unique and the Hamiltonian is not strictly convex in $\pi$, $B$, and $R$. Moreover,
since $\alpha > 0$, we cannot rule out a solution in which the participation constraint is not
binding.

**Proposition 3** There exists a tax schedule $\tilde{T}(\cdot)$ which leads to the optimal social welfare
and which gives to household $\theta$ an allocation such that:

$$V'(R(\theta)) = \frac{f(\theta)}{K(\theta) f(\theta) + \frac{a(\theta)}{a(\theta)} J(\theta)},$$

where $H(\theta)$ is a primitive of $\omega(\theta)f(\theta)$ such that $H(\bar{\theta}) = 0$ and $J(\theta)$ is a primitive of
$a(\theta)f(\theta)$ such that $J(\bar{\theta}) = 0$.

This optimal tax schedule is thus characterized by the absence of distortion ‘at the top’.
The most productive households are allocated with the optimal allocation. Indeed, for
$\theta = \bar{\theta}$, $H(\bar{\theta}) = J(\bar{\theta}) = 0$ and $V'[R(\bar{\theta})] = \frac{a(\bar{\theta})}{a(\bar{\theta})}$, i.e, the marginal benefit of consumption
is equal to the marginal disutility of labor. From this proposition, one can deduce the two
following results.

**Corollary 1** The tax schedule $\tilde{T}(I)$ that implements the optimal social welfare is such
that

$$\tilde{T}'(I(\theta)) = \frac{a(\theta) J(\theta)}{a(\theta) f(\theta) K(\theta)}.$$

**Corollary 2** Whenever $a(\cdot)$ is a constant, $\tilde{T}(\cdot)$ is also a constant.

If $a(\cdot)$ is a constant over a segment, then the firm has no mean to distinguish between
workers in this segment and proposes the same contract. Therefore, all these households
obtain the same utility, even if their marginal productivities are different. Thus, the gov-
ernment can only ask for the same income tax.

**Proposition 4** For all relevant incomes, the tax function $\tilde{T}(I)$ is decreasing.

As long as the firm plays after the government, any progressive income taxation would
be mitigated and even cancelled by the compensation scheme of the firm. Rather than
introduce more efficiencies, the government prefers to help the firm to discriminate, which
has a positive effect on incomes.
5 Conclusion

This article proposes a departure from the standard income taxation literature in allowing for an imperfect labor market and for the adverse selection problems arising between households and both the firm and the government. We characterize the optimal income taxation, as well as the corresponding labor contracts from the firm.

We show that, in our setting, there always exists an optimal regressive taxation, even if the government has an utilitarian social welfare function. Any progressive taxation has bad effects on the firm’s incentives. This results give some interesting insights. Imperfect labor markets are inefficient per se, however, they also reduce the possibility of redistribution from the productive workers to the less productive workers. The only efficient income taxation is regressive. This should be taken into account by competition authorities when they have to consider cases of in merging.

In this article, the firm has in the labor market both monopolist and bargaining power. This is a strong and sharp assumption made for technical reasons rather than a realistic description of the labor market. Our argument is based on the competition between the government and basically one firm. If we introduce more firms in each labor market, our argument should continue to apply as long as each firm as a power market and is allowed to offer contract rather than linear wages. We think that one can weaken our monopsony assumption and find qualitatively similar results.

We have assumed that both the firm and government used continuous and differentiable mechanisms. Technically, this assumption is very restrictive. However, we do not think that our results are driven by this restriction. Government and firms have different objectives, allowing them to use more general tools would probably not help to solve the coordination failure.

The utility functions chosen are also quite restrictive. We think that solving the model for very general utility function is technically impossible. But, since the main result is driven by the conflict between the government and the firm, we do not think that different utility functions would give qualitatively different results.

Finally, the timing chosen for the games emphases the role of the government. This assumption is crucial and considering a different timing would radically change our conclusions. The firm observes the taxation, and reacts to it. As the firm is the second player, adapting is compensation scheme it can cancel all the redistributive properties of any progressive taxation. If the firm is the first player, such a strategy is not possible anymore. Thus, it would be interesting to study the alternate timing, i.e, the firm plays before the government, check the robustness of the traditional models in a different way.
A Appendix

Proof of lemma 1.

The program of the government is:

$$\max_{R(.)U(.)} \left\{ \int_\Theta U(\Theta) dG(\Theta) \right\}$$

subject to:

$$\dot{U}(\Theta) = -\frac{d[\omega(\Theta)]}{d\Theta} I(\Theta), \quad (\mu)$$

$$\dot{B}(\Theta) = [I(\Theta) - R(\Theta)] f(\Theta). \quad (\eta)$$

The first order conditions for this program are:

$$\begin{align*}
-d[\omega(\Theta)] & \mu(\Theta) \frac{\partial I(\Theta)}{\partial R(\Theta)} - f(\Theta) \eta + \frac{\partial I(\Theta)}{\partial R(\Theta)} f(\Theta) \eta = 0, \\
-\dot{\mu}(\Theta) & = g(\Theta) + \eta f(\Theta) \frac{\partial I(\Theta)}{\partial U(\Theta)}, \\
-\dot{\eta}(\Theta) & = 0, \\
\mu(\Theta) & = \mu(\Theta) = 0.
\end{align*}$$

From the utility function, we have the equality

$$V[R(\Theta)] = \frac{a(\Theta)}{\omega(\Theta)} f(\Theta) + U(\Theta).$$

Replacing $I(\Theta)$ by its value gives us the conditions in the proposition.

Proof of Proposition 1.

$$\frac{dV'[R(\Theta)]}{d\Theta} = V''[R(\Theta)] R'(\Theta).$$

From the second order condition we have $R'(\Theta) \geq 0$, and from the assumptions $V'' \leq 0$. Then we can conclude that $V'[R(\Theta)]$ is a non-increasing function of $\Theta$.

Solving the incentive constraint gives:

$$\mu(\Theta) = \int_{\Theta} \bar{\theta} \left[ 1 - \eta \frac{f(s)}{g(s)} \frac{1}{V''[R(s)]} \right] g(s) ds.$$

The transversality condition requires $\mu(\Theta) = \mu(\bar{\Theta}) = 0$ and we have $V'[R(\Theta)] \geq 0$, it follows that $\eta \geq 0$. Combining these observations and the fact that $\frac{f(\Theta)}{g(\Theta)}$ is a non-decreasing...
function, \( 1 - \eta \frac{f(\theta)}{g(\theta)} - \frac{1}{g(\theta)} \) cannot be always negative or always positive. In fact, it must be negative for low values of \( \theta \) and positive for high values. This implies that \( \mu(\theta) \) is decreasing for all \( \theta \) less than some \( \hat{\theta} \), and increasing for \( \theta \) greater than \( \hat{\theta} \). Since \( \mu(\hat{\theta}) = 0 \), \( \mu(\theta) \) must be non-positive.

From proposition 1, the expression is:

\[
- \int \left[ \frac{a(\theta)}{\omega(\theta)} \mu(\theta) + \frac{\omega(\theta)}{a(\theta)} f(\theta) \eta \right] V'[R(\theta)] = f(\theta) \eta.
\]

If the optimal allocation is implemented by the indirect mechanism \( T(\cdot) \), the marginal tax rate \( T' \) must satisfy:

\[
[1 - T'(I(\theta))] V'[R(\theta)] = \frac{a(\theta)}{\omega(\theta)}.
\]

Then:

\[
T'(I(\theta)) = \frac{d}{d\theta} \left[ \frac{a(\theta)}{\omega(\theta)} \mu(\theta) \right] \frac{a(\theta)}{\omega(\theta)} f(\theta) \eta.
\]

Since \( \frac{\mu(\theta)}{\eta} \) is non-positive, \( T' \) is non-negative. ■

Proof of Lemma 2. The government maximizes:

\[
\max_{R(\cdot), \Pi(\cdot), \bar{U}} \int_0^\theta U dG(\theta) + \int_\theta^{\bar{\theta}} \pi(\theta) dF(\theta),
\]

subject to:

\[
\begin{align*}
\pi(\theta) &= \frac{d}{d\theta} \left[ \frac{\omega(\theta)}{\pi(\theta)} \right] (V[R(\theta)] - \bar{U}), \quad (\mu) \\
\pi(\theta) &\geq 0, \\
B(\theta) &= [I(\theta) - R(\theta)] f(\theta), \quad (\eta) \\
B'(\theta) &= B'(\theta) = 0.
\end{align*}
\]

The first order conditions are:

\[
\begin{align*}
-\eta(\theta) &= 0, \\
-\mu(\theta) &= f(\theta) + \frac{\partial I(\theta)}{\partial R(\theta)} \eta f(\theta), \\
\mu(\theta) &= 0, \\
\mu(\theta) \frac{d}{d\theta} \left[ \frac{\omega(\theta)}{\pi(\theta)} \right] V'[R(\theta)] &= -\eta \left[ \frac{\partial I(\theta)}{\partial R(\theta)} - 1 \right] f(\theta)
\end{align*}
\]

From the expression of the profit \( \pi \), we have:

\[
\begin{align*}
\frac{\partial I(\theta)}{\partial R(\theta)} &= V'[R(\theta)] \frac{\omega(\theta)}{a(\theta)}, \\
\frac{\partial I(\theta)}{\partial \pi(\theta)} &= -1.
\end{align*}
\]
These equalities and the first order conditions give us the conditions in the lemma.

Proof of Proposition 2. If the constraint \( p(q) \) is binding, then the relevant transversality condition is \( \mu(\theta) = 0 \). This condition implies:

\[
\int_{\theta}^{\bar{\theta}} f(t) dt - \eta \int_{\theta}^{\bar{\theta}} f(t) dt = 0.
\]

\[
\mu = 0,
\]

and so:

\[
V'[R(\theta)] = \frac{a(\theta)}{\omega(\theta)}.
\]

We suppose that the optimal allocation is implemented by the following indirect mechanism \( T(I) \). The firm chooses the salary which maximizes its profits (for a given individual \( \theta \)):

\[
[1 - T'(I)] \frac{V'[I - T(I)]}{a(\theta)} \omega(\theta) = 1,
\]

and then, for all \( \theta, T' = 0 \). All the workers pay the same taxes (whatever their before-tax income).

If the condition \( \Pi(\theta) \) is not binding, then the relevant transversality condition is \( \mu(\bar{\theta}) = \mu(\theta) = 0 \). Then the optimal allocation is unique and such that \( \eta = 1 \). There is no relevant informational problems and then the government can reach the first best.

One can remark that as long as \( a(\theta)/\omega(\theta) \) and \( V' \) are decreasing, the net income function \( R(\cdot) \) is increasing. Then it satisfies the so-called “second order condition”.

Proof of lemma 3.

The program of the firm is, given \( T(\cdot) \), to maximize

\[
\max_{U(\cdot),T(\cdot)} \left\{ \int_{\theta}^{\bar{\theta}} \left[ \frac{\omega(\theta)}{a(\theta)} [V(I(\theta) - T(I(\theta))) - U(\theta)] - I(\theta) \right] dF(\theta) \right\}
\]

subject to the constraints

\[
\forall \theta \quad U(\theta) = -\frac{a(\theta)}{a(\theta)} [V(R(\theta)) - U(\theta)],
\]

\[
U(\theta) = U.
\]

The two constraints have the following solution in \( U(\cdot) \)

\[
U(\theta) = -a(\theta) \int_{\theta}^{\bar{\theta}} \frac{\dot{a}(s)}{a(s)^2} V(R(s)) ds + \frac{a(\theta)}{a(\theta)} U.
\]
which yields
\[
\int_\theta^\bar{\theta} \frac{\omega(\theta)}{a(\theta)} U(\theta) dF(\theta) = - \int_\theta^\bar{\theta} \omega(\theta) \left[ \int_\theta^s \frac{\dot{a}(s)}{a(s)^2} V(R(s)) ds \right] dF(\theta) + \int_\theta^\bar{\theta} \frac{\omega(\theta)}{a(\theta)} U dF(\theta),
\]
\[
= - \int_\theta^\bar{\theta} \int_s^\bar{\theta} \frac{\dot{a}(s)}{a(s)} V(R(s)) \omega(\theta) f(\theta) d\theta \right] ds + \int_\theta^\bar{\theta} \frac{\omega(\theta)}{a(\theta)} U dF(\theta),
\]
\[
= - \int_\theta^\bar{\theta} \frac{\dot{a}(\theta)}{a(\theta)} V(R(\theta)) \frac{H(\theta)}{f(\theta)} \right] dF(\theta) + \int_\theta^\bar{\theta} \frac{\omega(\theta)}{a(\theta)} U dF(\theta).
\]

where \( H(\theta) = \int \omega(\theta) dF(\theta) \) such that \( H(\bar{\theta}) = 0 \).

Then, one can rewrite the objective function of the firm, which ends up at a point-by-point maximization of the member under the integral.

Proof of lemma 4.
The optimal compensation for a worker \( \theta \) is given by the following equation:
\[
V'[I(\theta) - T(I(\theta))] = \frac{1}{\frac{\omega(\theta)}{a(\theta)} \left[ 1 + \frac{a(\theta)}{a(\theta)} \frac{H(\theta)}{f(\theta)} \right]} \left[ 1 - T'(I(\theta)) \right].
\]
The proof follows directly.

Proof of Proposition 3.
Recalling from Proposition 2 that the utility is equal to
\[
U(\theta) = -a(\theta) \int_\theta^\bar{\theta} \frac{\dot{a}(s)}{a(s)^2} V(R(s)) ds + \frac{a(\theta)}{a(\theta)} U,
\]
one can compute, using the same tool as before,
\[
\int_\theta^\bar{\theta} U(\theta) dF(\theta) = \int_\theta^\bar{\theta} \left[ \frac{\dot{a}(\theta)}{a(\theta)^2} V(R(\theta)) \frac{J(\theta)}{g(\theta)} \right] dG(\theta) + \int_\theta^\bar{\theta} \frac{a(\theta)}{a(\theta)} dG(\theta) U,
\]
where
\[
J(\theta) = \int_\theta^\bar{\theta} a(\theta) g(\theta).
\]
Thus, the program of the government becomes
\[
\max_{\pi(\cdot),R(\cdot),B(\cdot)} \int_\theta^\bar{\theta} \left[ \frac{\dot{a}(\theta)}{a(\theta)^2} V(R(\theta)) \frac{J(\theta)}{g(\theta)} + \frac{a(\theta)}{a(\theta)} U \right] dG(\theta) + \int_\theta^\bar{\theta} \pi(\theta) dF(\theta),
\]

22
under the constraints
\[
\begin{align*}
(\mu) \quad & \dot{\pi}(\theta) = \dot{K}(\theta) V(R(\theta)) + \frac{\dot{a}(\theta)}{a(\theta)} U, \\
(\eta) \quad & \dot{B}(\theta) = [I(\theta) - R(\theta)] f(\theta), \\
& B(\theta) = B(\bar{\theta}) = 0, \\
& \pi(\theta) \geq 0.
\end{align*}
\]

The associated Hamiltonian is
\[
\Lambda_2 = \frac{\dot{a}(\theta)}{a(\theta)^2} V(R(\theta)) J(\theta) + \frac{\dot{a}(\theta)}{a(\theta)^2} g(\theta) U + \pi(\theta) f(\theta)
\]
\[
+ \mu(\theta) \left[ \dot{K}(\theta) V(R(\theta)) + \frac{\dot{a}(\theta)}{a(\theta)} U \right]
\]
\[
+ \eta(\theta) [I(\theta) - R(I(\theta))] f(\theta).
\]

The first order conditions are
\[
\begin{align*}
-\dot{\eta}(\theta) &= 0, \\
-\dot{\mu}(\theta) &= f(\theta) + \eta(\theta) f(\theta) \frac{\dot{a}(\theta)}{a(\theta)}, \\
\mu(\theta) &= \mu(\bar{\theta}) = 0,
\end{align*}
\]
\[
\eta(\theta) \left[ 1 - \frac{\partial f(\theta)}{\partial R(\theta)} \right] f(\theta) = \frac{f(\theta)}{\partial a(\theta)^2} J(\theta) + \dot{K}(\theta) V'(R(\theta)) \mu(\theta).
\]

Using the expression of $\pi(\theta)$ yields
\[
\begin{align*}
\frac{\partial f(\theta)}{\partial a(\theta)} &= -1, \\
\frac{\partial f(\theta)}{\partial R(\theta)} &= K(\theta) V'(R(\theta)).
\end{align*}
\]

Using in turn these equalities, the first order equations become
\[
\begin{align*}
-\dot{\eta}(\theta) &= 0, \\
-\dot{\mu}(\theta) &= f(\theta) - \eta f(\theta), \\
\mu(\theta) &= \mu(\bar{\theta}) = 0,
\end{align*}
\]
\[
\left[ 1 - K(\theta) V'(R(\theta)) \right] f(\theta) = \frac{f(\theta)}{\partial a(\theta)^2} J(\theta).
\]

From the last equation, the expression of $V'(\cdot)$ is
\[
V'(R(\theta)) = \frac{f(\theta)}{K(\theta) f(\theta) + \frac{\partial a(\theta)}{a(\theta)^2} J(\theta)}.
\]
In order to get $R(\theta)$ increasing, we need two assumptions. First we need $K(\theta)$ increasing with $\theta$ and $\frac{f(\theta)a^2(\theta)}{a(\theta)J(\theta)}$ decreasing with $\theta$. The first assumption has already been discussed in the paper. The second assumption, is linked to the usual condition on the hazard rate, recall that $J(\cdot)$ is a primitive of $a(\cdot)g(\cdot)$.

Proof of corollary 2.

Using the first order equation of the firm’s program, one finds that the firm chooses $I(\theta)$ such that
\[ \left[ 1 - T'(I(\theta)) \right] K(\theta) V'(I(\theta) - T(I(\theta))) = 1, \]
which, using the optimal allocation given by proposition 4, yields
\[ \left[ 1 - T'(I(\theta)) \right] = \frac{K(\theta) f(\theta)}{K(\theta) f(\theta) + \frac{a(\theta)}{a(\theta)^2} J(\theta)} = 1, \]
or
\[ T'[I(\theta)] = - \frac{\dot{a}(\theta)}{a(\theta)^2} J(\theta) K(\theta). \]

Proof of Proposition 4.

By assumption we have $K(\theta) > 0$. By construction, $\dot{a}(\cdot)$ and $J(\cdot)$ are negative, $f(\cdot)$ is a positive function.

Given the expression of $T'$:
\[ T'(I(\theta)) = - \frac{\dot{a}(\theta)}{a(\theta)^2} J(\theta) K(\theta), \]
we can easily conclude that $T'$ is negative.

Given that we have an expression of $T'$ we can put conditions on our exogenous functions such that $T''$ positive and compatible with all the previous assumption made. It justifies the maximization of profit given by proposition 4.
References


