When Redistribution Leads to Regressive Taxation

Cyril Hariton
Gwenaël Piaser

Department of Economics
Ca' Foscari University of Venice
When Redistribution Leads to Regressive Taxation

Cyril Hariton
Toulouse Business School

Gwenaël Piaser
University of Venice

Abstract
We introduce labour 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 firms have imperfect information, i.e., if they offer incentive contracts, then (under some assumptions) the best redistributive taxation is regressive.

Keywords
Income Taxation, Redistribution, Labor market, Multi-principals, Adverse selection, Mechanism design

JEL Codes
D21, D82, H21, L14
1 Introduction

There have been numerous papers published on optimal taxation since the seminal works of Mirrlees (1971) and Stiglitz (1982). One of the main contributions of these papers is to set down a unified theoretical framework, contract theory, to analyse the interaction between the government and taxpayers who are also workers. Workers are employed in a perfectly competitive labour market and, thus, are paid according to 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 leads them to reveal their true private information.

We depart from this literature by introducing an imperfect labour market. By “imperfect markets” we mean mainly two things. First, we consider a monopsonistic labour market. Second, we allow firms to discriminate among workers: wages are nonlinear. Indeed, the opposite assumption, i.e., an imperfect labour market, is consistent with empirical evidence.\(^1\) Specifically, we focus on a particular form of imperfect markets, which is an extreme one. 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 labour markets are neither perfectly competitive nor monopsonistic, but we think that considering a different polar case rather than the traditional one helps to explain how taxation is modified when markets are not fully competitive.\(^2\) The main objective of this paper is to illustrate the difficulties faced by a government in redistributing wealth among workers 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: existence of market power and lack of information. The firms are price makers and the labour market is segmented into several independent local markets, with 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 worker. As any monopsonist constrained by incomplete information, firms must offer revealing contracts to prevent opportunistic behaviours, i.e., they must reward productive workers. As a consequence, contracts offered by the firms constrain the government’s taxation: given a taxation schedule, the firms adapt their offers to keep

---

\(^1\)See Boal & Ransom (1997) for a survey of the literature on this topic.

\(^2\)In a different framework, Strobl & Walsh (2003) show that if firms are monopsonists in labour markets, a minimum wage may raise the number of hours which are already too high, but this has ambiguous effects on the number of employees and utility.
a high wage differential between more and less productive workers. This interaction between income tax and labour contracts is the driving force behind our results.

A paper by Hungerbüelher et al. (2006) addresses the same questions. The authors characterise optimal non-linear income taxation in an economy with a continuum of unobservable productivity levels and endogenous involuntary unemployment due to frictions in the labour markets. They show that redistributive taxation distorts labour demand. Moreover, compared to the “laissez-faire” equilibrium, 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 labour market is characterised by a matching problem: firms and workers are not equivalent. Workers differ by their type and they 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 results in unemployment. In our model, we do not consider matching problems nor competitive firms. Information revelation occurs through endogenous labour contracts. Interestingly, our results are quite different, thus showing the importance of the assumptions concerning the labour market.

From a theoretical point of view, our model is derived from the standard one. Workers are characterised by private information and their marginal productivity, with a continuous and closed support. Their utility is quasi-linear with a heterogeneous parameter on the labour part. On one hand, the labour market is organized by a local monopsonist which faces asymmetric information about the workers as it seeks to maximise its profits. It proposes labour 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 maximise 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 monopsonist 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 realisation of the contracts occurs at the end of the game. The closest analysis in this direction is Martimort (1999), but this paper deals with a pure regulation problem and uses discrete types. Our analysis is made in several steps. In the first part, we retain the assumption of a
perfectly competitive labour market in order to find the standard results in our framework. Then, in the second part, we introduce an informed monopsonist and observe the effect of introducing a monopsonist instead of perfect competition in the labour market. Finally, we relax the assumption of perfect information by the monopsonist and analyse the game with an uninformed monopsonist.

Our main finding is that in this modified setting we can observe 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 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 necessary.

On the other hand, if 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 to 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 in section 4. Section 5 briefly provides clues on possible extensions and makes conclusions.

2 The model

The model has three different players and two markets. The players are the households, the firm and the state. There are two goods: labour 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 are heterogeneous with respect to a variable $\theta \in \Theta = [\underline{\theta}, \overline{\theta}]$, distributed according to a cumulative distribution function $F(.)$ with density $f(.)$. Each household is characterised 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 $\theta$ is characterised by a constant disutility of labour $a(\theta)$ (with $a(\theta) > 0$) and a constant marginal productivity $\omega(\theta)$, (with $\omega(\theta) > 0$).

As we want to stress the particular outcomes arising when firms offer contracts, we will always assume that $\dot{a}(\theta) < 0$ and $\dot{\omega}(\theta) > 0$. We impose such a link on $\omega$ and $a$ in order to get clear ranking over types: in a laissez-faire economy, 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 labour is given for workers and that they are not responsible for this disutility.

Therefore, we will refer to type $\overline{\theta}$ as the “good” agent (high marginal productivity, low disutility of labour) and $\underline{\theta}$ as the “bad” agent (low marginal productivity, high disutility of labour). In order to avoid useless technical difficulties, we assume that $f$ is strictly positive and finite over $[\underline{\theta}, \overline{\theta}]$. It is also helpful to assume that $\dot{a}(\theta) > -\infty$, which is quite natural.

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

Each agent $\theta$ can sell a quantity $L(\theta)$ of labour force in exchange for income $I(\theta)$. 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 & Rochet (1983), Weymark (1987), Rochet (1991) and Boadway et al. (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 monopsonist in the labour market, with perfect competition in the final good market yielding a price $P$ normalized to 1. In one respect, this is an extreme departure from the perfectly competitive labour 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 labour 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 labour market but benefit from the profits of the firm.

The firm experiences a constant return-to-scale with respect to labour 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 which the firms maximises is:

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

(3)

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 enforceable\(^3\). In the general case (presented in section 4), the firm neither observes the marginal productivity nor the disutility of labour 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 affect our results.

\(^3\)An alternative model would be to introduce moral hazard.
2.3 Government

The government maximises social welfare, and to achieve this goal, it has three 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 a utility $U = V(R)$. Second, it designs an income tax that cannot be escaped by households. Finally, the government can take part or all of the profit of the firm by using a nondistortive lump-sum taxation denoted $\tau$. 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 labour market on the specific income tax. When designing its taxation contract, the government maximises a weighted sum of consumer surplus and firm profits:

$$SW = \int_{\theta} U(\theta) \lambda(\theta) dF(\theta) + \int_{\theta} \pi(\theta) dF(\theta) - \tau,$$

under the budget constraint

$$\int_{\theta} T(I(\theta)) dF(\theta) + \tau \geq 0.$$  

(4)

As the lump-sum taxation is not distortive, and as all our results are derived first using the order condition, we will assume that the government does not tax the firm, i.e., $\tau = 0$. Our results are unaffected by this simplification, which makes the model a little bit simpler.

The government does not know the disutility of labour 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

$$\lambda(\theta) f(\theta) = g(\theta),$$

(6)

and

$$G(\theta) = \int_{0}^{\theta} g(t) dt.$$  

(7)

The cumulative distribution $G$ 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 utilitarian government.

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 opti-
misation methods. This is a technical restriction and we do not think that the conclusions of the model would 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 labour contract \( L(I) \) to agents. At \( t = 2.5 \), agents choose both contracts and at \( t = 3 \), all contracts are realised\(^4\) (labour, income, income tax, consumption).

```
<table>
<thead>
<tr>
<th>t=0</th>
<th>t=1</th>
<th>t=2</th>
<th>t=2.5</th>
<th>t=3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agents discover their type</td>
<td>Government announces T(I)</td>
<td>The firms propose labor contracts</td>
<td>Agents choose contracts</td>
<td>Contracts are executed</td>
</tr>
</tbody>
</table>
```

Figure 1: Timing

This sequential contracting procedure with late contract realisation 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 principle is valid for the contract designed at time 1. Moreover, as taxation rules are more rigid than labour contracts (the change of a tax by a government is a complicated process in comparison to a change in a labour contract by a firm), the Stackelberg role attributed to the government is natural.

3 Labour Contract

In order to solve the game backwards, the problem of the firm has to be solved for a given tax schedule \( T(I) \). It is assumed that the firm 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 firm:

\[
U(\theta, \tilde{\theta}) = V[I(\tilde{\theta}) - T(I(\tilde{\theta}))] - a(\theta)L(\tilde{\theta}). \tag{8}
\]

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

\(^5\)Calzolari & Pavan (2000) discuss extensively the benefits of a sequential contract game of this type.
Thus, the program of the firm is:

\[
\max_{L(.),I(.)} \left\{ \int_{\theta}^{\hat{\theta}} \left[ \omega(\theta) L(\theta) - I(\theta) \right] dF(\theta) \right\}
\]

subject to the incentive compatibility and individual rationality constraints

\[
\begin{align*}
V [I(\theta) - T(I(\theta))] - a(\theta) L(\theta) & \geq [I(\hat{\theta}) - T(I(\hat{\theta}))] - a(\theta) L(\hat{\theta}), \\
V [I(\theta) - T(I(\theta))] & \geq U.
\end{align*}
\]

(9)

(10)

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

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

(11)

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

\[
\dot{U}(\theta) = -\dot{a}(\theta) L(\theta),
\]

(12)

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 \(\dot{L}(\theta) \geq 0\) or equivalently \(\dot{I}(\theta) \geq 0\). In the appendix we argue that if the quantity denoted by \(K(.)\) and defined in the following, is increasing, then \(I(.)\) is also increasing.

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

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

(13)

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 \(\theta\). Thus, the program of the firm can be rewritten as a dynamic optimisation problem with \(I\) as control variables and \(U\) as the state variable

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

(14)

\[\text{This explains the assumption on the structure of the utility function. If the quasi-linearity has been taken with respect to the income -which would be more satisfactory in a sense- then the labour becomes an implicit function of income and utility. This makes the analysis more difficult at the second step and, as a consequence, makes the overall study of the optimal taxation more complicated.}\]
subject to the constraints

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

\[
U(\theta) = U_0
\]

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

**Lemma 1** Let \( I(.) \) solve, for a given tax schedule \( T(I) \), the following problem

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

where

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

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

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

Note that the labour contract implemented by the firm depends on the tax the government is to implement. This reflects the fact that the 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:

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

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: it would be the optimal compensation scheme if the firm were perfectly informed, i.e., if \( H(\theta) \) were equal to zero.

The second term, \( 1 + \frac{\dot{a}(\theta)}{a(\theta)} \frac{H(\theta)}{\omega(\theta) f(\theta)} \) corresponds to the cost induced by truthful revelation by the agent. Then we can interpret \( K(\theta) \) as the “apparent productivity” of the agent \( \theta \) for the firm. If \( K(\theta) = \omega(\theta)/a(\theta) \), i.e., if \( \frac{\dot{a}(\theta)}{a(\theta)} \frac{H(\theta)}{\omega(\theta) f(\theta)} = 0 \), it means that the asymmetric
information has no impact on firm behaviour, 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 \( \frac{\dot{a}(\theta)}{a(\theta)} \frac{H(\theta)}{f(\theta)} = -1 \), the information is so costly that the worker \( \theta \) is “apparently unproductive”.

To keep the analysis meaningful, we assume that the “apparent productivity” is always greater than zero and smaller than the “adjusted productivity”, formally:

\[
\forall \theta \in \Omega, \quad \frac{\omega(\theta)}{a(\theta)} > K(\theta) > 0, \quad (20)
\]

and by construction

\[
\frac{\omega(\theta)}{a(\theta)} = K(\theta) > 0. \quad (21)
\]

Given their definitions, \( \dot{a}(\cdot) \) is negative and \( H(\cdot) \) is positive. The inequality \( \frac{\omega(\theta)}{a(\theta)} > K(\theta) \) is implied by our definitions. By assuming \( K(\theta) > 0 \), we assume that every worker remains productive when we introduce asymmetric information.\(^7\)

Compared to the “laissez-faire” approach, the discriminating monopsonist enlarges the range of incomes. To see that, let us consider the optimal compensation scheme under asymmetric information \( I^{**} \) characterised by the following first order condition

\[
K(\theta) V'[I^{**}(\theta) - T(I^{**}(\theta))] \left[1 - T'(I^{**}(\theta))\right] = 1. \quad (22)
\]

If we consider the optimal compensation scheme under full information \( I^* \),

\[
\frac{\omega(\theta)}{a(\theta)} V'[I^*(\theta) - T(I^*(\theta))] \left[1 - T'(I^*(\theta))\right] = 1. \quad (23)
\]

It implies for all \( \theta \in [\underline{\theta}, \bar{\theta}] \), we get \( I^*(\theta) < I^{**}(\theta) \) and \( I^*(\theta) \neq I^{**}(\bar{\theta}) \).

The resulting shape of income is a consequence of the lack of information and exhibits the usual properties of a second best mechanism.

The less productive workers are paid less, their total income dramatically decreases and their “apparent productivity” is quite low compared to their real adjusted productivity. On the contrary the compensation of the most productive worker is not distorted. The

\(^7\)We assume that all the productivities are strictly positive, which is technically restrictive, but allowing the “apparent productivities” to be negative would introduce difficulties which are outside the remit of this paper.
monopsonist firm wants to discourage productive workers from mimicking unproductive
workers.

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

**Lemma 2** Let $I^* (\theta)$ be the optimal compensation scheme for a given tax schedule $T (.)$, 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

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

(24)

The “laissez-faire” best contract offer is not feasible 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 $\theta$:

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

(25)

and the first order conditions must be equivalent as well,

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

(26)

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 progressive taxation: by definition a progressive taxation system
reduces income inequalities. So, any progressive taxation decrease makes labour more
expensive for the firm. Generally 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 taxa-
tion by increasing the range of its salaries and, consequently, by paying less to unskilled
workers. The preceding lemma shows that 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” equilibrium.

Our intuitive argument relies on 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 system) a government allows the firm (or encourages it) to increase inequalities.

4 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 labour contract to all types of agents.

As stated by Martimort (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 the agent’s type to the government. 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. $\hat{\theta}(\theta) = \theta$ if: $\Pi'(\theta) = 0$, which can be written as:

$$K(\theta) V' [R(\theta)] R'(\theta) - I'(\theta) = 0.$$  \hspace{1cm} (27)

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

$$\dot{\pi}(\theta) = K(\theta) V [R(\theta)] - \frac{\omega(\theta)}{a(\theta)} U,$$  \hspace{1cm} (28)

where:

$$\pi(\theta) = K(\theta) V [R(\theta)] - I(\theta) - \frac{\omega(\theta)}{a(\theta)} U.$$  \hspace{1cm} (29)

Moreover, the budget constraint can be rewritten, using $B(\theta) = \int_0^\theta [I(\theta) - R(I(\theta))] dF(\theta)$, as $\dot{B}(\theta) = [I(\theta) - R(I(\theta))] f(\theta)$, with $B(\theta) = B(\hat{\theta}) = 0$. 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:

$$\max_{\pi(\cdot), R(\cdot), B(\cdot)} \int_0^{\theta} U(\theta) dG(\theta) + \int_0^{\theta} \pi(\theta) dF(\theta),$$  \hspace{1cm} (30)
under the constraints:

\[ \forall \theta \in \Theta \quad \dot{\pi}(\theta) = K(\theta) V(R(\theta)) - \frac{\omega(\theta)}{a(\theta)} U, \]
\[ \forall \theta \in \Theta \quad \dot{B}(\theta) = [I(\theta) - R(I(\theta))] f(\theta), \]
\[ B(\theta) = B(\bar{\theta}) = 0, \]
\[ \forall \theta \in \Theta \quad \pi(\theta) \geq 0. \]  

(31)

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

\[ \dot{K}(\theta) \geq \frac{\dot{\omega}(\theta)}{a(\theta)}. \]  

(32)

The “apparent productivity” must sufficiently increase with \( \theta \). In order to get a monotonic compensation rule, i.e., high skilled workers are paid much more than unskilled workers, the adjusted productivity must be increasing. But this is not enough: rather, it must be “sufficiently increasing”. The firm, given the taxation, will offer individually rational contracts, then we can anticipate that \( V(R(\theta)) \geq U \). We can deduce from that \( \dot{\pi}(\theta) > 0 \). Thus the only relevant participation constraint is:

\[ \pi(\bar{\theta}) \geq 0. \]  

(33)

As in the previous section the program has two important features: the optimal taxation may not be 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 1** There exists a tax schedule \( \bar{T}(\cdot) \) which leads to the optimal social welfare and which gives to household \( \theta \) an allocation such that:

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

(34)

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 characterised 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 \( V'[R(\bar{\theta})] = \frac{a'(\theta)}{\omega(\theta)} \), i.e., the marginal benefit of consumption is equal to the marginal disutility of labour. 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

$$\tilde{T}'(I(\theta)) = -\frac{\dot{a}(\theta) J(\theta)}{a(\theta) f(\theta) K(\theta)}.$$  (35)

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

If $a(.)$ is a constant over a segment, then the firm has no means 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 government can only ask for the same income tax.

Proposition 2  For all relevant incomes, the tax function $\tilde{T}(I)$ is decreasing.

As long as the firm plays after the government in the sequential game, any progressive income taxation would be mitigated and even canceled 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 labour market and for the adverse selection problems arising between households and both the firm and the government. We characterise the optimal income taxation, as well as the corresponding labour contracts from the firm.

We show that, in our setting, there always exists an optimal regressive taxation, even if the government has a utilitarian social welfare function. Any progressive taxation has negative effects on the firm’s incentives. These results give some interesting insights. Imperfect labour 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 effect should be taken into account by competition authorities when they assess the merits of a proposed merger.

In this article, the firm has a strong bargaining power. This is a strong and sharp assumption made for technical reasons rather than a realistic description of the labour market. Our argument is based on the competition between the government and basically
one firm. If we introduce more firms in each labour market, our argument should continue to apply as long as each firm has a power market and is allowed to offer contracts 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 emphasises 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, by adapting its 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., assuming the firm plays before the government, as one could to check the robustness of the traditional models in a different way.

6 Acknowledgments

We would like to thank Maurice Marchant, Alessandro Pavan, Pierre Pestieau, Pierre Picard, 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, of course, remain our own.
A Appendix

Proof of Lemma 1. The program of the firm is, given $T(\cdot)$, to maximise

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

(36)

subject to the constraints

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

(37)

$$U(\theta) = U_0.$$  

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

$$U(\theta) = -a(\theta) \int_\Theta \frac{\dot{a}(s)}{a(s)^2} V(R(s)) ds + \frac{a(\theta)}{a(\theta)} U_0,$$

(38)

which yields

$$\int_\Theta \frac{\omega(\theta)}{a(\theta)^2} U(\theta) dF(\theta) = -\int_\Theta \omega(\theta) \left[ \frac{\dot{a}(s)}{a(s)^2} V(R(s)) \right] dF(\theta) + \int_\Theta \frac{\omega(\theta)}{a(\theta)^2} U dF(\theta),$$

$$= -\int_\Theta \int_\Theta \frac{\dot{a}(s)}{a(s)^2} V(R(s)) \omega(\theta) f(\theta) d\theta ds + \int_\Theta \frac{\omega(\theta)}{a(\theta)^2} UdF(\theta),$$

$$= -\int_\Theta \frac{\omega(\theta)}{a(\theta)^2} V(R(\theta)) \frac{H(\theta)}{f(\theta)} dF(\theta) + \int_\Theta \frac{\omega(\theta)}{a(\theta)^2} UdF(\theta).$$

(39)

where $H(\theta) = \int \omega(\theta) dF(\theta)$ such that $H(\Theta) = 0$. Then, one can rewrite the objective function of the firm, which ends up at a point-by-point maximisation of the member under the integral. ■

Proof of Lemma 2. The optimal compensation for a worker $\theta$ is given by the following equation:

$$V'[I(\theta) - T(I(\theta))] = \frac{1}{\omega(\theta) \left[ 1 + \frac{a(\theta)}{\omega(\theta)f(\theta)} \right] \left[ 1 - V'(I(\theta)) \right]}.$$  

(40)

The proof follows directly. ■

Proof of Proposition 1. From lemma 1, we have

$$U(\theta) = -a(\theta) \int_\Theta \frac{\dot{a}(s)}{a(s)^2} V(R(s)) ds + \frac{a(\theta)}{a(\theta)} U.$$  

(41)
One can compute, using the same tool as before,

\[
\int_\bar{\theta}^\theta U(\theta) dF(\theta) = \int_\bar{\theta}^\theta \left[ \frac{\dot{a}(\theta)}{a(\theta)^2} V(R(\theta)) \frac{J(\theta)}{g(\theta)} \right] dG(\theta) + \int_\bar{\theta}^\theta \frac{a(\theta)}{a(\theta)} dG(\theta) U, \tag{42}
\]

where

\[
J(\theta) = \int_0^\theta \dot{a}(\theta) g(\theta). \tag{43}
\]

Thus, the program of the government becomes

\[
\max_{\pi(\cdot), R(\cdot), B(\cdot)} \int_\bar{\theta}^\theta \left[ \frac{\dot{a}(\theta)}{a(\theta)^2} V(R(\theta)) \frac{J(\theta)}{g(\theta)} + \frac{a(\theta)}{a(\theta)} U \right] dG(\theta) + \int_\bar{\theta}^\theta \pi(\theta) dF(\theta), \tag{44}
\]

under the constraints

\[
\begin{align*}
(\mu) \quad \dot{\pi}(\theta) &= \dot{K}(\theta) V(R(\theta)) + \frac{a(\theta)}{a(\theta)} U, \\
(\eta) \quad \dot{B}(\theta) &= [I(\theta) - R(\theta)] f(\theta),
\end{align*} \tag{45}
\]

\[
B(\theta) = B(\bar{\theta}) = 0, \quad \pi(\theta) \geq 0.
\]

The associated Hamiltonian is

\[
\Lambda_2 = \frac{\dot{a}(\theta)}{a(\theta)^2} V(R(\theta)) J(\theta) + \frac{a(\theta)}{a(\theta)} g(\theta) U + \pi(\theta) f(\theta) + \mu(\theta) \left[ \dot{K}(\theta) V(R(\theta)) + \frac{a(\theta)}{a(\theta)} \right] U \\
+ \eta(\theta) [I(\theta) - R(I(\theta))] f(\theta). \tag{46}
\]
The first order conditions are
\[ -\dot{\eta}(\theta) = 0, \]
\[ -\dot{\mu}(\theta) = f(\theta) + \eta(\theta) f'(\theta) \frac{\partial l(\theta)}{\partial \pi(\theta)}, \]  
\[ \mu(\theta) = \mu(\overline{\theta}) = 0, \]  
\[ \eta(\theta) \left[ 1 - \frac{\partial l(\theta)}{\partial R(\theta)} \right] f(\theta) = V'(R(\theta)) \frac{\partial l(\theta)}{a(\theta)^2} J(\theta) + K(\theta) V'(R(\theta)) \mu(\theta). \]

Using the expression of \( \pi(\theta) \) yields
\[ \frac{\partial l(\theta)}{\partial \pi(\theta)} = -1, \]
\[ \frac{\partial l(\theta)}{\partial R(\theta)} = K(\theta) V'(R(\theta)). \]

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

From the last equation, the expression of \( V'(\cdot) \) is
\[ V'(R(\theta)) = \frac{f(\theta)}{K(\theta) f(\theta) + \frac{\partial l(\theta)}{a(\theta)^2} J(\theta)}. \]

In order to get \( R(\theta) \) increasing, we need two assumptions. First we need \( K(\theta) \) increasing with \( \theta \) and \( K(\theta) \) 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 the function \( a(\cdot) g(\cdot) \).

**Proof of Corollary 1.** Using the first order equation of the firm’s program, one finds that the firm chooses \( I(\theta) \) such that
\[ \left[ 1 - T'(I(\theta)) \right] K(\theta) V'(I(\theta) - T(I(\theta))) = 1, \]  

18
which, using the optimal allocation given by proposition 1, yields

\[
[1 - T'(I(\theta))] = \frac{K(\theta) f(\theta)}{K(\theta) f(\theta) + \frac{\dot{a}(\theta)}{a(\theta)^2} J(\theta)} = 1, \quad (52)
\]

or

\[
T'(I(\theta)) = -\frac{\dot{a}(\theta)}{a(\theta)^2} \frac{J(\theta)}{f(\theta) K(\theta)}. \quad (53)
\]

**Proof of Proposition 2.** 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 \(T'\):

\[
T'(I(\theta)) = -\frac{\dot{a}(\theta)}{a(\theta)^2} \frac{J(\theta)}{f(\theta) K(\theta)} \quad (54)
\]

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 maximisation of profit given by proposition 1. ■

**References**


