

IZA DP No. 1460

## Optimal Redistributive Taxation in a Search Equilibrium Model

Mathias Hungerbühler  
Etienne Lehmann  
Alexis Parmentier  
Bruno Van der Linden

January 2005

# Optimal Redistributive Taxation in a Search Equilibrium Model

**Mathias Hungerbühler**

*IRES, Catholic University of Louvain  
and ERMES, University of Paris 2*

**Etienne Lehmann**

*ERMES, University of Paris 2,  
IRES, Catholic University of Louvain and IZA Bonn*

**Alexis Parmentier**

*EUREQua, University of Paris 1  
and ERMES, University of Paris 2*

**Bruno Van der Linden**

*IRES, Catholic University of Louvain, FNRS, Belgium,  
ERMES, University of Paris 2 and IZA Bonn*

Discussion Paper No. 1460  
January 2005

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
Email: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### Optimal Redistributive Taxation in a Search Equilibrium Model\*

This paper characterizes 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. Redistributive taxation distorts labor demand and wages. Compared to their efficient values, gross wages, unemployment and participation are lower. Average tax rates are increasing. Marginal tax rates are positive, even at the top. Finally, numerical simulations suggest that redistribution is much more important in our setting than in a comparable Mirrlees (1971) setting.

JEL Classification: D82, H21, H24, J64

Keywords: optimal income taxation, unemployment, wage bargaining, matching

Corresponding author:

Etienne Lehmann  
ERMES  
Université Paris 2  
12 Place du Panthéon  
75230 Paris Cedex 05  
France  
Email: [elehmann@u-paris2.fr](mailto:elehmann@u-paris2.fr)

---

\* We thank participants to seminars at EUREQua, the day for belgian labour economists, CREST, AFSE, Namur (PAI-IUAP), Evry, T2M 2004, ESEM 2004 and EALE 2004. We thank Antoine d'Autume, Robin Boadway, Pierre Cahuc, Bart Cockx, Per Engström, Jean Michel Grandmont, Jean Hindricks, Herwig Immervoll, Francis Kramarz, Guy Laroque, David Margolis, Dominique Meurs, Gwenaël Piasser, Thomas Piketty, Fabien Postel Vinay, Jean Charles Rochet, Emmanuel Saez, Etienne Wasmer, the editor Fabrizio Zilibotti and the two anonymous referees. The usual caveat applies. We also thank ERMES and IRES for their hospitality. This research is part of a program supported by the Belgian government (Poles d'Attraction inter-universitaires PAI P5/21).

# I Introduction

In this paper, we consider an economy where agents have different productivities and search frictions on the labor market create unemployment. We assume that this economy is efficient in the absence of taxes. This paper characterizes the nonlinear income tax schedule that optimizes the redistributive objective of the government. We put emphasis on the case where the government does not observe the productivity of the worker-firm pairs.

Following Mortensen and Pissarides (1999) and Pissarides (2000), frictions on the labor market generate rents for workers and firms who have matched. These rents are shared through a Nash bargain on wages, as typically assumed in the literature. This setting allows to deal with the effects of income taxation on wages and labor demand. To highlight the intuition of our mechanisms, we ignore the intensive margin. Higher levels of taxes reduce the surplus to be shared. This reduction is split between the worker and the firm through a decline in net wages and an increase in gross wages (or wage costs). In contrast, higher marginal tax rates imply that a given marginal increase in gross wages has a reduced impact on net wages and an unchanged one on profit levels. Therefore, a higher marginal tax rate makes it less rewarding for workers to bargain aggressively and hence the gross wage falls (see Malcomson and Sartor (1987) and Lockwood and Manning (1993)). Through their negative impact on wages, higher marginal tax rates or lower tax levels induce firms to raise the resources spent on posting vacancies. Put differently, labor demand increases. As the creation of vacancies absorbs resources, this increase in labor demand and eventually in gross output is not necessarily efficient. In this paper, thanks to an appropriate assumption about the workers' bargaining power (the so-called Hosios condition), an economy without taxes and benefits is efficient: Aggregate output minus vacancy costs is maximized. Hence, taxes are here not used to correct for inefficiencies. As in the seminal paper of Mirrlees (1971), they are used to redistribute income from more to less productive agents.

When the government observes the productivity of the individuals, there is, as expected, no trade-off between redistribution and efficiency. When productivity levels are not observed by the government, we prove five main analytical results.

First, for participants to the labor market, optimal wages are below their efficient levels. Thus, the optimal levels of employment are higher than their efficient values. This employment result contrasts with the standard literature where total working hours are below their efficient levels (Mirrlees (1971), Stiglitz (1982, 1987)). The extent of redistribution is constrained by the asymmetric information about productivity levels. Lowering gross wages below their efficient level (thereby increasing employment levels) is in our model a key ingredient to mitigate the effects of this constraint. Our result is important because it gives a new rationale for fiscal policies aiming at stimulating labor demand, such as the ones that have been recommended by e.g. Drèze and Malinvaud (1994) or Phelps (1997).

Second, average tax rates should be increasing through the whole distribution of productivities. Progressive taxation (in the sense of Musgrave and Musgrave (1976)) is therefore a main feature of the optimal tax schedule. It is not as detrimental as the standard optimal taxation literature suggests. There is clearly no equivalent property in the Mirrlees framework.

Third, marginal tax rates should be positive at the top of the distribution. This property holds even with a bounded distribution of productivities. Distorting employment

at the top generates no equity gain. The gross wage and the employment rate of the most productive workers are therefore at their efficient levels. However, the positive tax level at the top generates wage pressure that has to be compensated by a positive marginal tax rate. In the standard literature, marginal tax rates should be nil at the top of the distribution because there are no gains from distorting labor supply.

Fourth, our over-employment result is nevertheless tempered since the labor-force participation rate should be below its efficient level. Our model introduces a participation decision. Assuming that the value of inactivity is identical for all agents, every individual who is less productive than an endogenous threshold does not search for a job. Despite its efficiency cost, increasing this threshold (i.e. raising the number of welfare recipients) allows to reduce the informational rent that accrues to more productive workers. This explains our result concerning the participation rate.

Fifth, the least skilled employed workers may receive an in-work benefit. But in this case, it is always lower than the assistance benefit. This result is in line with the standard optimal income taxation literature. Saez (2002), Boone and Bovenberg (2004, 2005) Laroque (2002) and Choné and Laroque (2005) challenge this view by integrating a participation decision. They recommend higher in-work benefits than assistance benefits. This result does not come out in our framework. The introduction of a continuous extensive margin might change this fifth property.

These five results hold under the assumption of risk neutrality. Most of them are also true under risk aversion.

We then develop various numerical exercises <sup>1</sup>. They illustrate the properties of our optimal tax schedule. Although these exercises are very coarse, we hope they provide some insights on the characteristics of optimal tax schedules. Marginal tax rates appear to be high. Depending on the chosen parameters, they lie between 42% and 66%. For a given set of parameters, marginal tax rates fluctuate in a range that is smaller than 10 percentage points.

Finally, we check to what extent our setting and the Mirrlees one generate different optimal tax schedules. The latter model is calibrated to generate the same distribution of earnings and the same elasticity of earnings with respect to marginal tax rates as the former one. Our results turn out to be dramatically different since our recommended marginal tax rates are by and large twice higher.

Our main contribution is methodological since we build a model where the efficiency distortions induced by income taxation are due to matching and wage bargaining instead of the standard consumption-leisure trade-off. To what extent our assumptions are empirically relevant remains an open question. There is first some evidence concerning the wage moderating effect of tax progressiveness in wage bargaining models. The time series regressions by Malcomson and Sartor (1987) for Italy, Lockwood and Manning (1993) for the UK, or Holmlund and Kolm (1995) for Sweden lend support to this mechanism (see Sørensen (1997) for a survey). A second literature surveyed by Blundell and MaCurdy (1999) is concerned with labor supply responses in micro data. Elasticities of labor supply along the intensive margin are rather low for men, with a typical estimate between 0.1 and 0.5, but are higher for married women. However, only hours are observable by econometricians, not in-work effort. This suggests that labor supply responses should be better estimated by looking at gross earnings instead of working hours (see Feldstein (1995) and

---

<sup>1</sup>All programs are available on <http://www.u-paris2.fr/ermes/membres/lehmann/lehmann.htm>

Gruber and Saez (2002)). However, these estimates can be either interpreted as labor supply or wage bargaining responses. Both mechanisms predict that individual gross earnings are increasing with the level of taxes (when leisure is assumed to be a normal good in labor supply models) and decreasing with marginal tax rates. Predictions differ only as far as the effects on employment, working hours and hourly wages are concerned. Few papers have exploited such data. The time-series estimates by Hansen, Pedersen and Sløk (2000) conclude that higher tax progression decreases hourly wages for blue collar workers, suggesting that the wage bargaining mechanism dominates the labor supply effect for these workers. However, the reverse turns out to be true for white collar workers.

Labor demand is not absent in the standard literature. The typical model assumes an aggregate production function with perfect substitution between the different types of labor. Hence, hourly gross wages equal marginal products of labor and are independent of taxation. Stiglitz (1982) considers a two-skill model with imperfect substitution between high and low-skilled labor. The marginal tax rate should then be negative at the top. This increases high-skilled labor supply and employment and so reduces the hourly wage skill premium. Marceau and Boadway (1994) extend Stiglitz's (1982) model by introducing a minimum wage that generates unemployment for low-skilled workers. However, when marginal productivity is decreasing, the derivation of a non-trivial labor demand requires a finite number of skills. The matching model of Mortensen and Pissarides (1999) and Pissarides (2000) provides an interesting alternative since it allows to derive very easily a continuum of skill-specific labor demand functions. Engström (2002) extends the model of Stiglitz (1982) in a matching setting. His analytical results are developed for exogenous wages. They emphasize complementarities between the intensive margin and job creation. Our analytical framework is different since we consider a continuum of agents, fixed working hours but endogenous wages. His simulation results with endogenous wages are related to ours. In particular, he emphasizes that employment is a key feature of the redistributive scheme.

The presence of matching frictions raises new questions in welfare economics. The normative analyses developed in the matching framework have put strong emphasis on the conditions under which the allocation of resources is efficient. Hosios (1990) first established the condition under which the output generated by additional jobs is equal to the resources needed to create additional vacancies in an economy without taxes. When this condition is not fulfilled, labor-income taxation is one of the instruments that can be used to restore efficiency (see Boone et Bovenberg (2002) and Lehmann and Van der Linden (2002)). We do not further explore this research avenue. Taking the Hosios condition for granted, the labor-income tax schedule is chosen so as to maximize a social welfare function. The desire to redistribute income between skill-groups will as usual create an equity-efficiency trade-off but the latter will come out in a different theoretical setting <sup>2</sup>.

The paper is organized as follows. Section II presents the model and derives the analytical results. Section III is devoted to numerical simulations. In section IV, we compare the optimal schedule in our setting to the optimum in a Mirrlees-type setting. Section V checks the robustness of our analytical results under risk aversion. Section VI concludes.

---

<sup>2</sup>Hungerbühler (2004) extends the present model to the case where the Hosios condition is not fulfilled.

## II The model

We consider a static model<sup>3</sup> where jobs differ according to their exogenous productivity denoted by  $a \in [a_0, a_1]$  with  $0 \leq a_0 < a_1 \leq +\infty$ . The intensive margin is not taken into account here. Workers and firms are assumed to be risk neutral. Directed search is assumed for simplicity<sup>4</sup>. So, type- $a$  active workers search for type- $a$  jobs. Firms open type-specific vacancies. Each vacancy has to be filled with a single searching worker. Matching workers and vacancies is a time-consuming and costly activity. Following Mortensen and Pissarides (1999) and Pissarides (2000), we consider a well-behaved matching function that gives the number (measure) of type- $a$  jobs formed as a function of the number  $U_a$  of searching workers and the number  $V_a$  of vacancies. Employment in segment  $a$  is an increasing and constant-return-to-scale function  $H(U_a, V_a)$ . This matching function implicitly captures heterogeneities, frictions and information imperfections on the labor market.

The size of the population is normalized to 1. Workers' types are distributed according to a continuous density  $f(\cdot)$  and a c.d.f.  $F(\cdot)$ . These functions are common knowledge. Through costly screening, the productivity of a worker is observed by the firm. We assume the government has not this ability. The income tax schedule consists in a continuously differentiable non-linear tax function  $T(\cdot)$  and an untaxed assistance benefit  $b$ . Since the government observes gross wages but not productivity,  $T(\cdot)$  is only based on gross income. We assume that job search cannot be monitored by the government. The assistance benefit  $b$  is thus distributed to searching and non-searching jobless individuals. For each type  $a$ ,  $w_a$ ,  $L_a$  and  $x_a$  denote respectively the gross wage (or equivalently the wage cost), the employment rate and the workers' ex-post surplus in case of employment, with  $x_a \equiv w_a - T(w_a) - b$ . Hence, type- $a$  employed workers receive  $w_a - T(w_a) = x_a + b$ . Further, we define  $\Sigma_a \equiv x_a \cdot L_a$  as the workers' expected surplus and  $Y_a \equiv w_a \cdot L_a$  as the workers' expected gross income. Let  $d > 0$  be the value of inactivity. Irrespectively of their type, individuals who do not (respectively, do) search for a job receive  $b + d$  (resp.  $b$ ).

Posting a type- $a$  vacancy costs  $\kappa_a$ . This parameter captures the cost of screening applicants and the investment cost of creating a workstation. A type- $a$  filled (respectively unfilled) vacancy yields a surplus of  $a - w_a - \kappa_a$  (resp.  $-\kappa_a$ ) to the firm-owner. In the literature, the vacancy cost is either taken as fixed or as proportional to productivity (see e.g. Pissarides (2000)). We therefore assume that<sup>5</sup>:

$$0 \leq \frac{\dot{\kappa}_a}{\kappa_a} \leq \frac{1}{a} \quad (1)$$

The timing of the model is:

1. The government commits to a taxation scheme  $T(\cdot)$  and a level of benefit  $b$ .
2. Firms open vacancies and workers decide whether or not they search for a job.
3. Matching occurs. The firm and the worker negotiate the wage once matched.
4. Transfers accrue to the agents.

---

<sup>3</sup>Our static model simplifies the dynamic version of the matching model but still captures its major mechanisms (see e.g. Boone and Bovenberg (2002)).

<sup>4</sup>Since we consider the whole distribution of productivity levels, directed search seems to be a less unrealistic assumption than undirected search.

<sup>5</sup>A dot over a variable denotes the total derivative with respect to type  $a$  (e.g.  $\dot{\kappa}_a = d\kappa_a/da$ ).

This section is organized as follows. First, we deal with the objective function and the budget constraint of the government. Second, we introduce frictions and the demand side of the labor market. Third, we characterize efficiency in the idealized case where the government has perfect information on productivities. Fourth, we deal with the wage bargain. Fifth, incentive and participation constraints are defined in the case where the government observes wages but not productivities. Finally, the properties of the second-best optimum are derived.

## II.1 The government

We first present the government's objective and its budget constraint. How productivity levels are allocated in the population is out of the scope of this article. People are simply not held responsible for their productivity. So, the government is ready to compensate for differences in productivity levels. We assume the following objective for the government:

$$\Omega = \int_{a_0}^{a_1} \{p_a \cdot \Phi [L_a (w_a - T (w_a)) + (1 - L_a) b] + (1 - p_a) \cdot \Phi (b + d)\} f (a) da \quad (2)$$

where  $\Phi'(\cdot) > 0$ ,  $\Phi''(\cdot) < 0$ ,  $p_a = 1$  if type- $a$  workers participate to the labor market and  $p_a = 0$  otherwise.

This objective expresses that the government cares about the distribution of expected utilities, namely  $L_a (w_a - T (w_a)) + (1 - L_a) b$  for those who are active and  $b + d$  for inactive people. It encompasses as limiting cases the maximin criterion (max  $b$ ) and the "pure" utilitarian criterion (whenever  $\Phi''(\cdot) = 0$ ). For expository reasons, we first neglect the issue of insurance against the unemployment risk. As will soon be shown, over-employment is optimal at the second best for it contributes to the fulfillment of the incentive compatibility constraint. Under risk aversion and when insurance is incomplete, overemployment comes out for a distinct reason, namely because it allows to better share risks. This is illustrated in Section V where risk aversion is introduced. There, the two motives for overemployment will be simultaneously present. Finally, objective (2) can be micro-founded if we assume that the economy is made of productivity-specific representative households that perfectly share consumption between their employed and unemployed members. From  $\Sigma_a = L_a (w_a - T (w_a) - b)$ , we rewrite this objective as:

$$\Omega = \int_{a_0}^{a_1} \{p_a \cdot \Phi [\Sigma_a + b] + (1 - p_a) \cdot \Phi (b + d)\} f (a) da \quad (3)$$

The government faces the budget constraint:

$$\int_{a_0}^{a_1} p_a \cdot T (w_a) \cdot L_a \cdot f (a) da = \left[ \int_{a_0}^{a_1} \{p_a (1 - L_a) + 1 - p_a\} f (a) da \right] b + E \quad (4)$$

where  $E \geq 0$  is an exogenous amount of public expenditures. Since  $Y_a - \Sigma_a = L_a [T (w_a) + b]$ , we rewrite the government's budget constraint (4) as:

$$\int_{a_0}^{a_1} p_a (Y_a - \Sigma_a) \cdot f (a) da = b + E \quad (5)$$



## II.2 The matching process

Following empirical studies (see Blanchard and Diamond (1989) or Petrongolo and Pissarides (2001)), we assume a Cobb-Douglas matching function. The number of type- $a$  matches is a function of the number of type- $a$  vacancies  $V_a = \theta_a \cdot f(a)$  and of the number of type- $a$  searching workers  $U_a$  according to:

$$H_a = A \cdot (U_a)^\gamma \cdot (V_a)^{1-\gamma} \quad \text{with} \quad \gamma \in (0, 1)$$

All type- $a$  individuals either search for a job or stay inactive. If they search, their number is  $U_a = f(a)$ . Their probability of finding a job (resp. the probability of filling a type- $a$  vacancy) is  $L_a = H_a/U_a = A \cdot \theta_a^{1-\gamma}$  (resp.  $H_a/V_a = A \cdot \theta_a^{-\gamma}$ ).

The expected return of posting a vacancy is  $A \cdot \theta_a^{-\gamma} \cdot (a - w_a) - \kappa_a$ . The higher the gross wage  $w_a$ , the lower this return. Firms enter freely the market and post vacancies as long as this return is positive. Therefore, in equilibrium, this return is nil (the so-called “free-entry condition”). One can then derive the type- $a$  probability of being employed (or the “labor demand”):

$$L_a = A^{\frac{1}{\gamma}} \cdot \left( \frac{a - w_a}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} \quad (6)$$

The free-entry condition implies that net output (i.e. total output net of search costs  $L_a \cdot a - \theta_a \cdot \kappa_a$ ) equals the workers’ expected gross income  $L_a \cdot w_a$ , so:

$$Y_a = w_a \cdot L_a = L_a \cdot a - \theta_a \cdot \kappa_a$$

Taking (6) into account, we abuse notations slightly by writing net output under the free-entry condition as:

$$Y_a(w_a) \equiv A^{\frac{1}{\gamma}} \cdot \left( \frac{a - w_a}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} \cdot w_a \quad (7)$$

## II.3 The first-best optimum

In this subsection only, we assume that the government perfectly observes productivities. The government then chooses the assistance benefit  $b$ , the wage  $w_a$ , the participation indicator  $p_a$  and workers’ expected surplus  $\Sigma_a$  to maximize the social objective (3) subject to the budget constraint (5) and the labor demand (6):

$$\begin{aligned} \max_{w_a, p_a, \Sigma_a, b} \quad & \int_{a_0}^{a_1} \{p_a \cdot \Phi[\Sigma_a + b] + (1 - p_a) \cdot \Phi(b + d)\} f(a) da \\ \text{s.t.} : \quad & \int_{a_0}^{a_1} p_a (Y_a(w_a) - \Sigma_a) \cdot f(a) da = b + E \end{aligned} \quad (8)$$

As shown in appendix A, the solution to this problem implies the following results. First,  $\Sigma_a^* = d$  for all participating types. Second, the first-best gross wage  $w_a^*$  maximizes net output. Therefore,  $Y_a$  is our measure of efficiency. Efficient values of gross wages, employment, and net output are:

$$\begin{aligned} w_a^* &= \gamma \cdot a \quad , \quad L_a^* = (1 - \gamma)^{\frac{1-\gamma}{\gamma}} \cdot A^{\frac{1}{\gamma}} \cdot \left( \frac{a}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} \\ Y_a^* &= Y_a(w_a^*) = \gamma (1 - \gamma)^{\frac{1-\gamma}{\gamma}} \cdot A^{\frac{1}{\gamma}} \cdot a \cdot \left( \frac{a}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} \end{aligned} \quad (9)$$

To increase employment above  $L_a^*$ , firms have to open more vacancies. The resources spent to create these vacancies are not offset by the increase in output. Matching frictions therefore imply that full employment is not optimal<sup>6</sup>. Equations (1) and (9) imply that the efficient level of employment is non-decreasing in  $a$ . Finally, let  $a_d^*$  be defined by:

$$Y_{a_d^*}^* \begin{matrix} = \\ \geq \end{matrix} d \quad \text{if} \quad a_d^* \begin{matrix} > \\ = \end{matrix} a_0 \quad (10)$$

The efficient participation condition implies that every type above (below)  $a_d^*$  participate (stay inactive).

## II.4 The wage bargain

We henceforth consider the case where the government does not observe jobs' productivities. Therefore, the tax schedule  $T(\cdot)$  is a differentiable function of gross wages only. Once a firm and a worker have matched, they bargain over the wage. In the absence of an agreement, nothing is produced and the worker gets the assistance benefit  $b$ . These outside options imply the existence of a positive rent  $a - T(w_a) - b$ . As it is standard in the literature (see Pissarides (2000)), this rent is shared by maximizing a Nash product. The wage  $w_a$  maximizes the Nash product depending on the worker's and the firm's surplus if they find an agreement:

$$\max_{w_a} [w_a - T(w_a) - b]^\beta \cdot [a - w_a]^{1-\beta}$$

taking  $b$  and  $T(\cdot)$  as given and where  $\beta \in (0, 1)$  denotes the worker's bargaining power. For convenience, we redefine the Nash product as:

$$\max_{w_a} [w_a - T(w_a) - b] \cdot A^{\frac{1}{\gamma}} \cdot [a - w_a]^{\frac{1-\beta}{\beta}}$$

From Equation (6), workers' expected surplus coincides with the (redefined) Nash product if, as we henceforth assume, the Hosios condition  $\beta = \gamma$  is fulfilled. This condition states that the relative weight of the firm's surplus in the Nash product  $(1 - \beta) / \beta$  is equal to the elasticity of labor demand with respect to the firm's surplus  $(1 - \gamma) / \gamma$ . Under the Hosios condition, efficiency is reached in the absence of taxes and benefits. The results of this paper would therefore also be obtained in the case of any other wage setting that maximizes workers' expected utility  $\Sigma_a + b$ , given the level of  $b$ , the labor demand function (6) and the tax schedule  $T(\cdot)$ . This is in particular the case with skill-specific monopoly unions. Henceforth,  $\Sigma_a$  will denote the workers' expected surplus evaluated at bargained wages:

$$\Sigma_a = \max_{w_a} [w_a - T(w_a) - b] \cdot A^{\frac{1}{\gamma}} \cdot \left[ \frac{a - w_a}{\kappa_a} \right]^{\frac{1-\gamma}{\gamma}} \quad (11)$$

The first-order condition leads to:

$$w_a = \frac{\gamma(1 - T'_a)a + (1 - \gamma)(T_a + b)}{\gamma(1 - T'_a) + 1 - \gamma} \quad (12)$$

---

<sup>6</sup>  $A$  and  $\kappa_a$  are assumed to be such that  $L_a^* < 1$  and  $a - w_a^* > \kappa_a$ .

where  $T'_a \equiv T'(w_a)$  denotes the marginal tax rate and  $T_a \equiv T(w_a)$  denotes the level of taxes for a type- $a$  worker. Furthermore, the first-order condition of (11) implies:

$$\frac{w_a}{a - w_a} = \frac{\gamma}{1 - \gamma} \cdot \frac{1 - T'_a}{1 - \frac{T_a + b}{w_a}} \quad (13)$$

Since neither the firm's nor the worker's ex-post surplus can be negative, we conclude that:

$$T'_a \leq 1$$

Therefore, the worker's ex-post surplus  $x_a = w_a - T(w_a) - b$  is necessarily an increasing function of the gross wage  $w_a$ .

In the rest of this subsection, we consider the level of tax  $T_a$  and the marginal tax rate  $T'_a$  as parameters. The tax schedule influences the labor market equilibrium in two ways. First, higher levels of taxes  $T_a$  (or benefits  $b$ ) reduce the workers' ex-post surplus. Therefore, workers claim higher wages. This positive effect on individual gross earnings is similar to what occurs in a labor supply framework when leisure is a normal good. There, a rise in the level of taxes at given marginal tax rates increases labor supply. However, here this upward pressure on wages reduces employment.

Second, keeping  $T_a$  constant, the wage is decreasing in the marginal tax rate  $T'_a$  because a unit rise in the gross wage increases net earnings at a rate of one minus the marginal tax rate. As the marginal tax rate rises, the worker earns less from each increase in gross wages while the effect on firms' profits remains unchanged. Therefore, workers have less incentives to claim higher wages (see Malcomson and Sartor (1987), Lockwood and Manning (1993), Holmlund and Kolm (1995), Pissarides (1998), Sørensen (1999) or Boone and Bovenberg (2002) among others). As in labor supply frameworks, higher marginal tax rates decrease individual gross earnings. However, in the labor supply literature, the channel is different because there gross earnings are decreasing due to lower working hours. Here, labor demand rises.

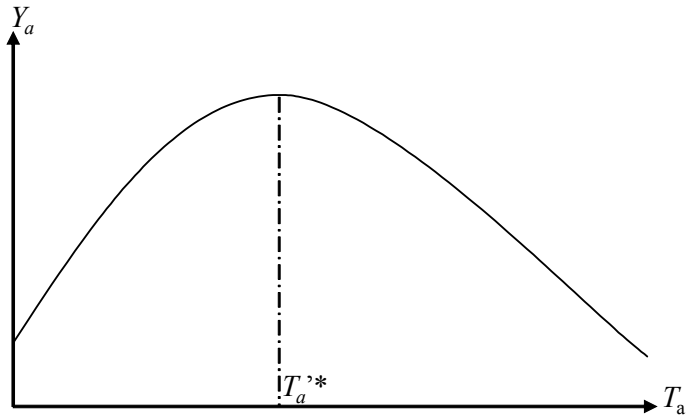


Figure 1: The impact of the marginal tax rate  $T'_a$  on efficiency at a given level of taxes  $T_a$ .

A rise in  $T'_a$  at a given tax level  $T_a$  has a non-monotonic effect on net output  $Y_a$ . The gross wage  $w_a$  decreases according to Equation (12). So, employment  $L_a$  and therefore gross output  $a \cdot L_a$  and total vacancy costs  $\kappa_a \cdot \theta_a$  increase. When the wage level is higher

(respectively lower) than its efficient value  $w_a^*$ , the effect on gross output (resp. on vacancy costs) dominates, so net output increases (resp. decreases). The relationship between net output and the marginal tax rate parameter is therefore hump shaped (see Figure 1). Let the “efficient marginal tax rate”  $T_a'^*$  be the one that maximizes net output for a given level of tax. From Equations (9) and (12):

$$T_a'^* = \frac{T_a + b}{\gamma \cdot a} \quad (14)$$

This equality establishes an upward relationship between the efficient marginal tax rate and the level of taxes.

## II.5 Incentive and participation constraints

The strategic interaction between the firm and the worker is modelled in a reduced form way. Given a tax schedule  $T(w)$  and an assistance benefit  $b$ , the worker-firm pair chooses the wage that maximizes the Nash product. We further assume that side-payments are not allowed<sup>7</sup>. The government therefore sets taxes as a function of wages subject to the following implementation constraint:

$$w_a = \arg \max_{w_{\hat{a}}} [w_{\hat{a}} - T(w_{\hat{a}}) - b] \cdot A^{\frac{1}{\gamma}} \cdot \left[ \frac{a - w_{\hat{a}}}{\kappa_a} \right]^{\frac{1-\gamma}{\gamma}} \quad (15)$$

This problem is equivalent to one in which worker-firm pairs participate in a direct mechanism where each worker-firm pair announces its productivity type and receives a worker surplus and a wage level from the government. Let  $w_a$  be the solution to (15). The outcome can then be described by the wage  $w(a) = w_a$  and the worker’s ex-post surplus  $x(a) = w_a - T(w_a) - b$ . Then, the mapping  $a \mapsto (w(a), x(a))$  defines a direct<sup>8</sup> truthful mechanism that is incentive compatible in the usual sense, i.e.

$$a = \arg \max_{\hat{a}} x(\hat{a}) \cdot A^{\frac{1}{\gamma}} \cdot \left[ \frac{a - w(\hat{a})}{\kappa_a} \right]^{\frac{1-\gamma}{\gamma}} \quad (16)$$

This equivalence result is commonly known as the “taxation principle” (see Hammond (1979) and Guesnerie (1995)). This rewriting shows the equivalence of the tax function to a class of direct mechanism where the worker and the firm are treated as a single agent with preferences described by the Nash product objective. However, the worker and the firm have diverging interests. While the firm wants to announce a type whose allocation gives a low wage to increase its profits, the worker prefers to announce a type that gives a higher worker’s surplus. In our model, the worker-firm pair can only send a single message<sup>9</sup>. Otherwise, the match would be interrupted and the firm and the worker would get their outside options. We assume that the worker and the firm negotiate the message to be sent to the government through a Nash bargain.

<sup>7</sup>Tax evasion is here neglected. Therefore, earnings are perfectly and costlessly observed.

<sup>8</sup>Since there is only one large principal and many small agents in our model, the revelation principle applies and we can concentrate on direct mechanisms without loss of generality.

<sup>9</sup>Cremer and McLean (1988) show that if agents who are ultimately informed about the other agents’ types have different interests and if coordination is not possible between them, the principal can achieve the first-best outcome by punishing the agents if they report different types. This possibility is not available in our context since we assume that the firm and the worker can coordinate their messages. In our model, this coordination actually takes place in a Nash bargain.

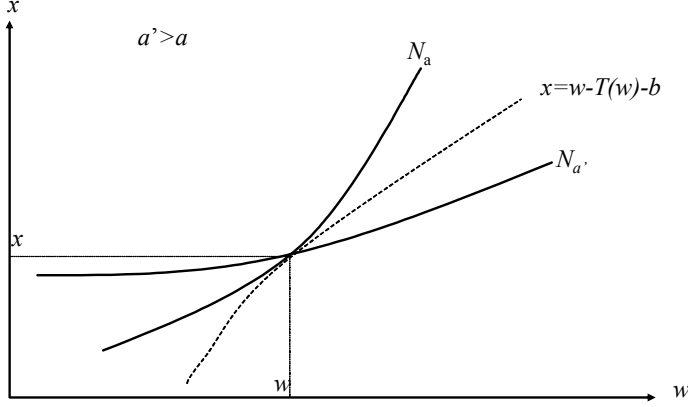


Figure 2: The single-crossing property

Standard principal-agent techniques then apply (see Salanié (1997) and Laffont and Martimort (2002)). The Nash product can be written as a function  $N(a, w, x)$  of the type  $a$ , the gross wage  $w$  and the worker's ex-post surplus  $x$  in the following way:

$$N(a, w, x) \equiv A^{\frac{1}{\gamma}} \left( \frac{a-w}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} x$$

The worker's surplus  $x$  has to increase when the gross wage  $w$  increases to keep the Nash product  $N(a, \cdot, \cdot)$  unchanged. For each pair  $(w, x)$  the marginal rate of substitution:

$$\left. \frac{\partial x}{\partial w} \right|_{N(a, \cdot, \cdot)} = \frac{1-\gamma}{\gamma} \frac{x}{a-w}$$

is a decreasing function of the type  $a$ . This single-crossing property is illustrated in Figure 2. This figure displays the indifference curves in terms of the Nash product  $N(a, \cdot, \cdot)$  for worker-firm combinations with different productivity levels  $a' > a$ . The higher the productivity of a match, the less elastic is the firm's surplus to the gross wage and the less sensitive is the Nash product to changes in the gross wage. Applying the envelope theorem to Equation (11), the incentive compatibility constraint (16) can be rewritten as the following first and second-order conditions (see appendix B.1)

$$\dot{\Sigma}_a = \frac{1-\gamma}{\gamma} \left( \frac{1}{a-w_a} - \frac{\dot{\kappa}_a}{\kappa_a} \right) \Sigma_a \quad (17)$$

$$\dot{w}_a \geq 0 \quad (18)$$

Equations (1) and (17) imply that the the maximized Nash product or equivalently the expected worker's surplus,  $\Sigma_a$ , is increasing with respect to the productivity<sup>10</sup>.

A worker decides to search as long as the expected utility when searching  $\Sigma_a + b$  is higher than the assistance benefit  $b$  plus the value of inactivity  $d$ . The participation constraint can then be written as:

$$\Sigma_a \geq d \quad (19)$$

<sup>10</sup>Given the single crossing property, first-order and second-order conditions are equivalent to the incentive compatibility constraints (15) (see Salanié (1997)). Throughout the paper, we consider only the first-order condition and verify ex-post in our simulations that  $\dot{w}_a > 0$ .

Since the worker's expected surplus  $\Sigma_a$  is increasing in  $a$  and since  $d$  is by assumption identical for all types, there exists a single threshold  $a_d \geq a_0$  such that workers endowed with  $a < a_d$  stay inactive whereas workers with  $a \geq a_d$  search for a job.

## II.6 The second-best optimum

Under asymmetric information, the government chooses the threshold  $a_d$ , the assistance benefit  $b$ , the wage and the workers' expected surplus functions  $(w_a, \Sigma_a)$  to maximize the government's objective (3) subject to the budget constraint (5), the labor demand (6), the incentive constraints (17) and the participation constraints (19):

$$\begin{aligned} & \max_{a_d, w_a, \Sigma_a, b} F(a_d) \Phi(b+d) + \int_{a_d}^{a_1} \Phi(\Sigma_a + b) f(a) da & (20) \\ \text{s.t. : } & \int_{a_d}^{a_1} (Y_a(w_a) - \Sigma_a) \cdot f(a) da = b + E \\ & \dot{\Sigma}_a = \frac{1-\gamma}{\gamma} \left( \frac{1}{a-w_a} - \frac{\dot{\kappa}_a}{\kappa_a} \right) \Sigma_a \quad \begin{cases} \Sigma_{a_d} = d & \text{if } a_d > a_0 \\ \Sigma_{a_d} \geq d & \text{if } a_d = a_0 \end{cases} \end{aligned}$$

For any value of  $a_d$  and  $b$ , this is a standard optimal control problem where workers' expected surplus  $\Sigma_a$  is the state variable and the gross wage  $w_a$  is the control variable. The first-order conditions lead to the following formulation of the equity-efficiency trade-off for all  $a \geq a_d$  (see Appendix B.2):

$$\lambda \cdot \frac{\partial Y_a}{\partial w_a} \cdot f(a) = \frac{1-\gamma}{\gamma(a-w_a)^2} \int_a^{a_1} \{\lambda - \Phi'_t\} \Sigma_t \cdot f(t) \cdot dt \quad (21)$$

where  $\lambda$  is the Lagrange multiplier associated with the budget constraint and  $\Phi'_a = \Phi'(\Sigma_a + b)$ . For  $a < a_d$ , we define  $\Phi'_a = \Phi'(b + d)$ .

Consider a marginal increase in the type- $a$  wage. The maximized Nash product achieved by type- $a$  worker-firm combinations,  $\Sigma_a$ , is fixed by the incentive constraints for less productive pairs. The rise in the wage decreases the employment level  $L_a$  and thereby gross output  $L_a \cdot a$  but also the resources spent on posting vacancies. The effect on net output  $Y_a$  is therefore ambiguous. If  $w_a < w_a^*$  (resp.  $>$ ), the total effect is positive (resp. negative). Multiplying this by the number of type- $a$  agents  $f(a)$  and the shadow cost of public funds  $\lambda$ , the left-hand side of (21) measures the social value of the net marginal change in output. This captures the efficiency side of the trade-off.

The right-hand side of (21) represents the equity cost of a higher gross wage for type- $a$  worker-firm combinations. As  $w_a$  rises at fixed  $\Sigma_a$ , more productive worker-firm pairs find it more attractive to mimic type- $a$  combinations. To prevent this, the Nash product accruing to the former has to grow. Looking at Equation (17), the term in front of the integral measures by how much the rate of change of the maximized Nash product  $\dot{\Sigma}_a/\Sigma_a$  has to grow when  $w_a$  marginally increases. The incentive compatibility constraints will remain satisfied if all combinations with a productivity higher than  $a$  benefit from an equivalent *relative* increase in their Nash product. For any type  $t$  above  $a$ , this relative increase times  $\Sigma_t$  gives the rise in the Nash product. Each unit of the latter generates an increase in the social welfare measured by  $\Phi'_t$  and implies a budgetary cost equal to  $\lambda$ .

The proof of the following normative properties is left to appendices.

**Proposition 1** *The levels of the gross wage and of employment should be efficient at the top of the distribution.*

The right-hand side of (21) indicates that changing the gross wage at the top has no distributional effect. The government can therefore set this gross wage at the level  $w_{a_1}^*$  which maximizes net output.

**Proposition 2** *For all worker-firm pairs with productivity  $a_d \leq a < a_1$ , the gross wage should be below and the employment should be above their efficient levels.*

Consider that the tax schedule has been optimized for all workers up to type  $a$ . The maximized Nash product of type- $a$  worker-firm combinations,  $\Sigma_a$ , is predetermined by the incentive compatibility constraints. This level of the Nash product is depicted in Figure 3 by the curve denoted  $\Sigma_a$ . Let  $w_a$  decrease below its efficient value  $w_a^*$ . This only has a second-order effect on efficiency. But there is a first-order effect on the Nash products and thus on workers' expected surplus  $\Sigma_{\tilde{a}}$  for types  $\tilde{a} > a$ . By reducing  $w_a$ , the government can reduce the latter (see Equation (17) and the downward shift of the curve labeled  $\Sigma_{\tilde{a}}$  in Figure 3). The government can extract more tax revenues from these types above  $a$ . This gain in resources is valued at the marginal cost of public funds  $\lambda$ . The loss in the workers' expected surplus for these types is valued at the marginal social welfare  $\Phi'_{\tilde{a}}$ ,  $\tilde{a} > a$ . Following Equation (21), these two effects are integrated over all types above  $a$ . Since the assistance benefit is optimally chosen, the property  $\int_{a_0}^{a_1} (\Phi'_t - \lambda) f(t) dt = 0$  holds (see Equation (34) in appendix B.2). As  $\Sigma_a$  is increasing in  $a$  and  $\Phi'_a$  is therefore decreasing in  $a$ , the right-hand side of (21) is always positive. In other words, the additional tax revenues are more valued than the loss in utility above  $a$ . Therefore the optimal value of  $w_a$  should be below its efficient level.

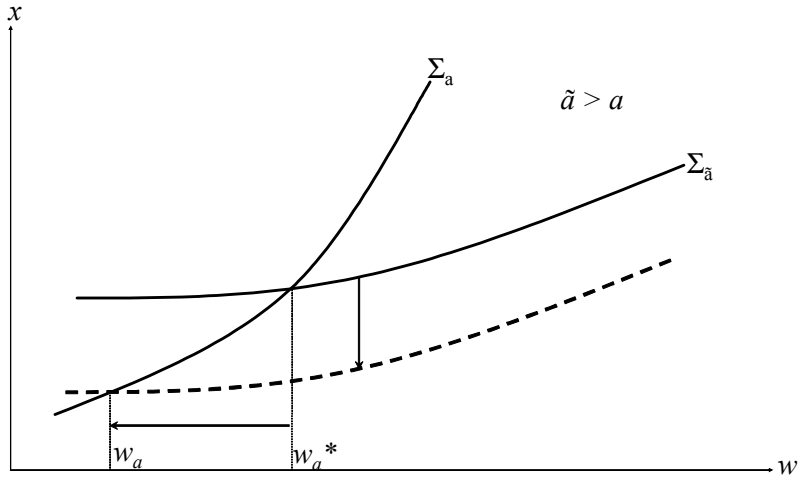


Figure 3: The equity-efficiency trade-off.

Another intuition for Proposition 2 is given by Figure 1. Keeping the level of taxes unchanged up to  $T_a$ , a rise in the marginal tax  $T'_a$  creates an equity gain since it allows to tax richer workers more heavily. At the optimum, this equity gain has to be compensated by a loss in efficiency. According to Figure 1, the marginal tax rate is then necessarily higher than the efficient one. Consequently, the wage rate is below its efficient level and therefore there is overemployment. So, as in the Mirrlees model, incentive constraints lead to a decline in gross earnings. However, this reduction here implies a rise in labor demand whereas in the standard literature it follows from a reduction in labor supply.

Proposition 2 recommends overemployment for all participating types (except at the top of the productivity distribution). Many are convinced that there actually is underemployment, in particular at medium and low productivity levels. Adopting this view, Proposition 2 has to be considered in a normative way. As the underlying model is highly stylized, this result should be considered with care. The main message is that endogenizing wages and labor demand leads to recommend a fiscal stimulation of labor demands for participating types. The next proposition deals with the adverse effect of this stimulation on the optimal participation rate.

**Proposition 3** *The participation rate should be lower than or equal to its efficient value.*

Aggregate net output and hence efficiency increase if individuals of types  $a \in [a_d^*, a_d)$  participate. But their participation also gives to worker-firm pairs with productivity above  $a_d$  the possibility to mimic them. To avoid this mimicking, the government has to give an additional informational rent to these more productive matches. If this equity cost is higher than the efficiency gain, the government prefers that individuals of types  $a \in [a_0, a_d)$  do not participate.

Combining Propositions 2 and 3, the efficient and the second-best optimal levels of total employment  $L = \int_{a_0}^{a_1} L_a f(a) da$  cannot be ranked. On the one hand, less people should participate to the labor market. On the other hand, more participants should be employed.

**Proposition 4** *In-work benefits (if any) should be lower than assistance benefits.*

In-work benefits that are higher than assistance benefits increase participation. However, the previous proposition shows that the government chooses not to increase participation. This proposition implies that an EITC would not be optimal at the second best.

**Proposition 5** *Average tax rates should increase with the wage level. Marginal tax rates should be positive everywhere.*

The first part of this proposition states that the tax schedule  $T(\cdot)$  has to be progressive in the sense of Musgrave and Musgrave (1976). This means that the coefficient of residual income progression is below 1 everywhere. The importance of this conclusion should be stressed since standard optimal income taxation models do not yield precise analytical results about the shape of average tax rates. With a bounded distribution of productivity, this literature has shown that the marginal tax rate should equal zero at the top. Therefore, one only knows that the average tax rate should necessarily be decreasing close to the top of the distribution.

Since the value of inactivity is unique, the second part of Proposition 5 is in accordance with common wisdom<sup>11</sup>, except at the top of the distribution. The reason why the marginal tax rate should be positive at the top is easily understood. As the government wants to redistribute income in favor of less productive agents, the level of taxes is positive at the top. This distorts the gross wage upwards. To restore an efficient level of wage (Proposition 1), a positive marginal tax rate is therefore needed at the top (see Equation (14)).

---

<sup>11</sup>See Saez (2002) or Boone and Bovenberg (2004) for a critical appraisal of this wisdom.



There are no analytical results about the profile of marginal tax rates. If the level of employment was efficient everywhere, following the previous argument, the upward profile of tax levels would require a rising profile of marginal tax rates according to Equation (14). However, Proposition 2 implies that there is over-employment everywhere, except at the top. Marginal taxes are therefore higher than their efficient values. From Propositions 1 and 2, one knows that the positive difference between optimal and efficient employment levels has to decline with productivity levels in the neighborhood of the top of the distribution. We do not know more. So, possible changes in the intensity of the over-employment effect could lead to a non-monotonic relationship between marginal tax rates and wages.

### III Simulations

#### III.1 Calibration

The parameters are chosen to roughly represent key figures for France. For the productivity distribution, we use a truncated log-normal density:

$$f(a) = \frac{K}{a} \exp\left(\frac{\log a - \log(\mu \cdot a_1 + (1 - \mu) a_0)}{2 \cdot \xi^2}\right)$$

This form used to be typical in the literature (Mirrlees (1971), Tuomola (1990) and Boadway et al. (2000); see Diamond (1998) and Saez (2001) for a critique).  $K$  is a scale parameter. We select the values of  $\mu$  and  $\xi$  from data on wages found in the French Labor Force Survey 2002 (*Enquête Emploi*). We consider monthly earnings of all full time working individuals. To keep things simple, as Saez (2002 page 1072), we assume that the distribution of the log of productivities is obtained through a translation of the distribution of the log of wages. A sensitivity analysis with respect to the productivity distribution will be conducted in Subsection III.3.4. The standard deviation of the log of wages is 0.42, so we take  $\xi = 0.4$ . Finally, we find a mean for the log of wages equal to 7.23. This corresponds to a weight of 0.21, so we take  $\mu = 0.2$ . We truncate the distribution at  $a_1 = 20\,000$  Euros. Finally  $a_0 = 0$ . The distribution of  $a$  in our benchmark case is displayed in solid lines in Figure 4.

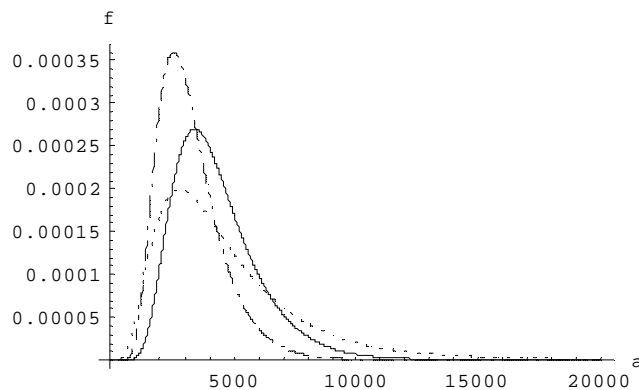


Figure 4: Density Functions  $f(a)$ . The benchmark case is in solid line. Dotted line is used for  $(\mu, \xi) = (0.2; 0.6)$ . Dashed line for  $(\mu, \xi) = (0.15; 0.4)$

$\sigma$	$\gamma$	$d$	$\mu$	$\xi$	$E/Y_1$
1	0.5	500	0.2	0.4	0

Table 1: Parameter values in the benchmark case

The elasticity of the matching function,  $\gamma$ , is set at 0.5. This corresponds to the average estimate in the empirical literature (see Petrongolo and Pissarides (2001)). Since the Hosios condition is assumed, the worker and the firm have equal bargaining power. There is no micro-evidence available for the values of  $\kappa_a$  and  $A$ . Actually, only the value of  $\kappa_a \cdot A^{\frac{1}{\gamma-1}}$  matters. We assume an iso-elastic shape so that  $\kappa_a \cdot A^{\frac{1}{\gamma-1}} = \kappa_0 \cdot a^l$ . We select the values of  $l$  and  $\kappa_0$  so that the unemployment rate equals 2% at the top and 25% at the bottom in the economy without taxes. This leads to  $\kappa_0 = 0.62$  and  $l = 0.88$ . In the benchmark case, we assume that the value of leisure  $d$  equals 500 euros. The government's expenditures  $E$  are set equal to zero. Finally, the government's utility function is assumed to be a CES function of the expected surplus. We have,  $\Phi(\Omega) = \Omega^{1-\sigma}/(1-\sigma)$ . In the benchmark case, we take  $\sigma$  equal to 1, so  $\Phi(\Omega) = \log(\Omega)$ . This corresponds to the basic parameterization in Saez (2002). Table 1 summarizes the values of the parameters.

### III.2 The benchmark

Figure 5 illustrates the propositions in section II. The upper-left panel displays the employment rate  $L$  as a function of the productivity level  $a$ . The upper-right panel displays the after-tax income level  $C$  as a function of the gross wage  $w$ . The panels at the bottom show the levels of taxes  $T$  and of marginal tax rates  $T_m$  as functions of gross wages. Dotted lines correspond to the economy without taxes. Table 2 displays the main features of the optimum.  $WT$  denotes the economy without taxes,  $SB$  the second best optimum and  $\Delta$  the relative differences of second-best values compared to those in the equilibrium without taxes.  $L = \int_{a_0}^{a_1} L_a f(a) da$  denotes the total employment rate in the economy.

Employment should be above its efficient level except at the top. Over-employment should be more pronounced for low productivity levels. This property is not surprising. As in the Mirrlees setting, our model predicts that the government should highly distort the outcomes of the low productivity types because there are few of them. The corresponding marginal tax rates should therefore be high. However, introducing a continuous extensive labor supply margin might remove this property. In Saez (2002), marginal tax rates can become negative if participation elasticities are sufficiently high. Similarly, Boone and Bovenberg (2004, 2005) show that the introduction of search costs might also lead to negative marginal tax rates. At the bottom of the distribution, such decreases in marginal tax rates could attenuate the over-employment property of our optimum.

The participation rate should equal 94%. Optimal total employment is nevertheless higher than in the economy without taxes (see Table 2). Due to these employment distortions, total output net of vacancy cost is 2.7% lower than in the equilibrium without taxes. However, redistribution increases the certainty equivalent of social welfare,  $\Phi^{-1}(\Omega)$  by around 7%. In addition, the level of assistance benefit is quite high. The tax function turns out to be close to linear. Marginal tax rates are first decreasing and then slightly hump-shaped. They lie between 59% and 52%.

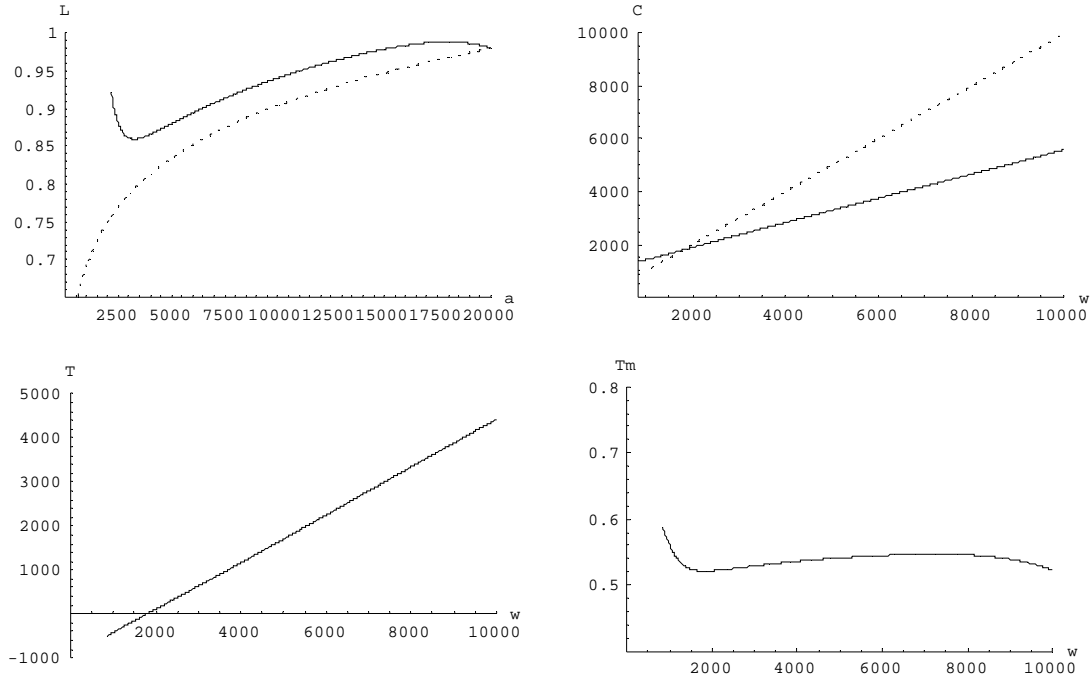


Figure 5: Benchmark Case. Values in the economy without taxes in dotted lines.

### III.3 Sensitivity analysis

The following section indicates to what extent the optimum is sensitive to changes in the main parameters.

#### III.3.1 The aversion to inequality

When the aversion to inequality  $\sigma$  increases, the government is ready to distort more the allocation of resources. When the aversion to inequality is doubled, marginal tax rates shift upwards by about 5 percentage points (see Figure 6), and the assistance benefit varies slightly. These rather small changes might be explained by the fact that the benchmark already leads to a high degree of redistribution.

#### III.3.2 The value of inactivity

When the value of inactivity  $d$  increases, the participation constraint becomes more stringent. Therefore, the expected surplus of marginal participants increases. Keeping the type of marginal participants unchanged, the incentive constraints imply a higher expected surplus for all participants. The government is therefore forced to redistribute less. To mitigate this effect, the government lets the participation rate decline. As shown in Figure 7 and in Table 2, marginal tax rates shift downwards and assistance benefits decrease substantially. Hence, from a quantitative viewpoint, the optimal taxation is sensitive to this parameter.

	Employment $L$		Net output $Y$			Welfare $\Phi^{-1}(\Omega)$			$b$	Participation		
	SB	$\Delta$	WT	SB	$\Delta$	WT	SB	$\Delta$		WT	SB	$\Delta$
Benchmark	82.6	+1.8%	1794	1747	-2.7%	1627	1738	+6.8%	809	99.6	94.0	-5.6
$\sigma = 0.5$	82.7	+1.9%	1794	1766	-1.6%	1710	1760	+3.0%	722	99.6	95.9	-3.7
$\sigma = 2.0$	84.8	+4.6%	1794	1735	-3.3%	1475	1708	+15.8%	853	99.6	93.7	-5.9
$d = 300$	84.6	+3.9%	1796	1782	-0.8%	1626	1763	+8.4%	1070	100	98.3	-1.7
$d = 700$	80.1	+1.0%	1779	1702	-4.3%	1634	1725	+5.5%	575	97.1	89.4	-7.7
$\gamma = 0.4$	82.0	+2.2%	1430	1378	-3.6%	1304	1381	+5.9%	511	98.4	91.5	-6.9
$\gamma = 0.6$	84.1	+3.4%	2155	2122	-1.5%	1952	2100	+7.6%	1090	99.9	96.8	-3.1
$\xi = 0.6$	79.5	+1.1%	1982	1890	-4.6%	1631	1900	+16.4%	960	96.1	85.6	-10.5
$\mu = 0.15$	77.1	+0.5%	1292	1234	-4.5%	1185	1251	+5.6%	430	97.3	89.2	-8.1
$\frac{E}{Y^*} = 0.2$	0.85	+4.3%	1794	1751	-2.4%	—	1371	—	445	99.6	95.0	-4.6

Table 2: Numerical results WT for the economy without taxes and SB for second best

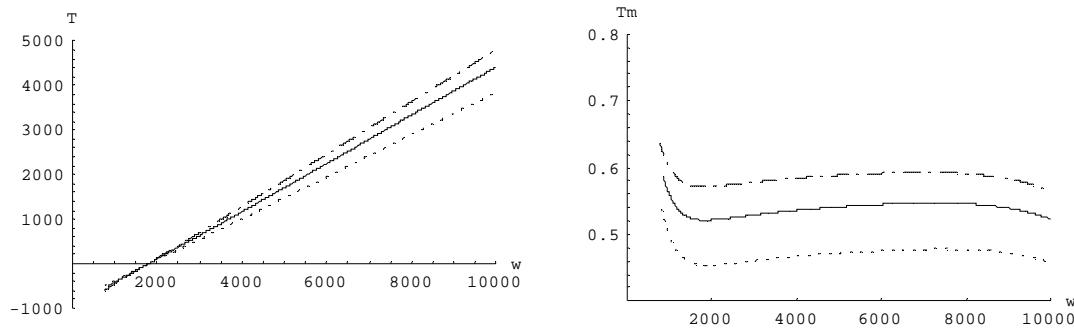


Figure 6: Dotted, solid and dashed lines respectively for  $\sigma$  equal to 0.5, 1 and 2.

### III.3.3 The labor demand elasticity and the bargaining power

We only consider variations that simultaneously change the elasticity of the matching function  $\gamma$  and the bargaining power so that the Hosios' condition remains satisfied. The vacancy costs are adjusted so as to keep the extreme values of the unemployment rate unchanged in the economy without taxes. The higher the level of  $\gamma$ , the higher is the matching effectiveness of searching workers. Hence, the parameter  $\gamma$  is an efficiency parameter. As this parameter increases, labor demand becomes less elastic to the firm's expected surplus. Put differently, taxation becomes less distortive. Consequently, assistance benefits increase and marginal tax rates shift upwards (see Figure 8 and Table 2).

### III.3.4 The form of the distribution

In Figure 9, we report the effect of a rise in  $\xi$  (dotted lines) and of a decrease in  $\mu$  (dashed lines) (see Figure 4). A rise in  $\xi$  means a higher standard deviation for the log of productivity. Therefore, the government wishes to redistribute more, with more distortions and higher marginal tax rates and assistance benefit. A decrease in  $\mu$  basically implies a shift of the productivity distribution to the left. Given the participation constraint, this

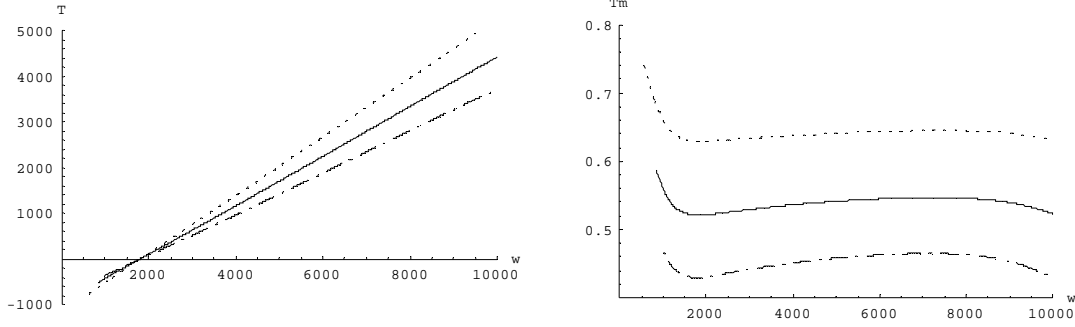


Figure 7: Doted, solid and dashed lines respectively for  $d$  equal to 300, 500 and 700 Euros.

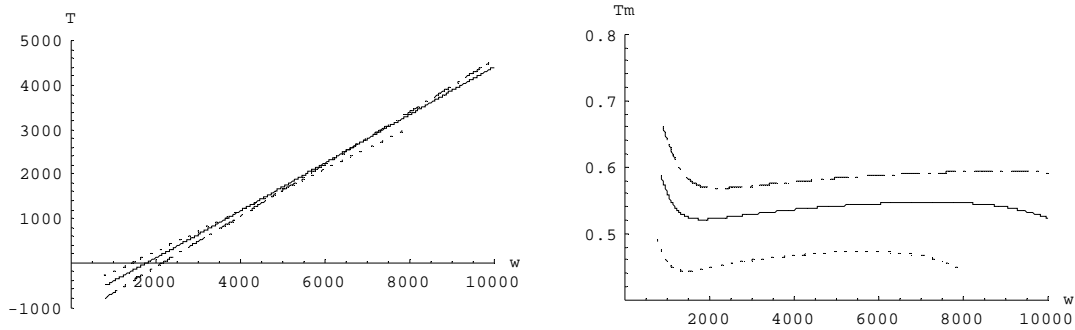


Figure 8: Dotted, solid and dashed lines respectively for  $\gamma$  equal to 0.4, 0.5 and 0.6.

increases the density of participating workers at the bottom  $f(a_d)$ . Hence, redistribution becomes more costly, and in particular, participation becomes much more elastic. As a consequence, the government redistributes less with lower marginal tax rates.

### III.3.5 Public expenditures

With positive public expenditures  $E$ , the marginal tax profile is almost the same. The main difference is a reduction in the assistance benefit (see Table 2). Comparing levels of welfare is meaningless here, because by definition  $E = 0$  in the economy without taxes.

## IV Comparing the Mirrlees approach to ours

Our main contribution is methodological since we build a model where the efficiency distortions induced by income taxation are due to matching frictions and wage bargaining instead of the standard consumption-leisure trade-off. Economists do observe the distribution of wages and the elasticity of earnings with respect to the marginal tax rate. Both the labor supply-Mirrlees approach and the bargaining-matching model are consistent with this observed elasticity. To what extent is the optimal tax schedule sensitive to the micro-foundation of this elasticity is the question we now address.

To compare our model to a framework that only incorporates the labor supply choice of the individuals, we build a model à la Mirrlees (1971) that generates the same distribution

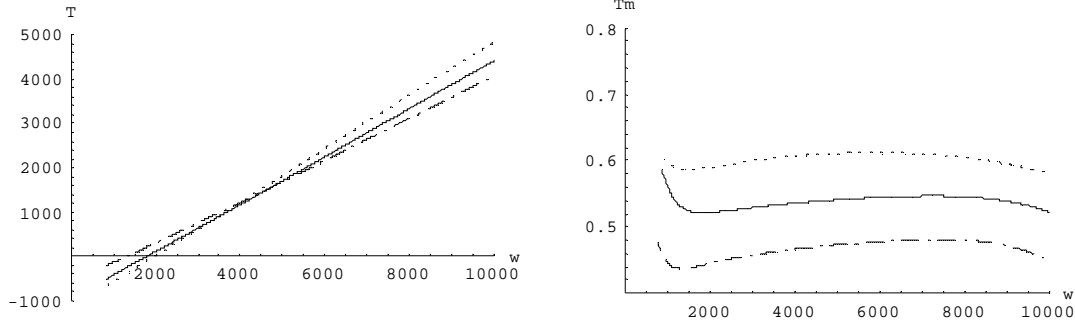


Figure 9: Dotted, solid and dashed lines respectively for  $(\mu; \xi) = (0.2; 0.6)$ ,  $= (0.2; 0.4)$  and  $= (0.15; 0.4)$

of wages and the same elasticity of income with respect to the marginal tax rate at the equilibrium without taxes. As e.g. Diamond (1998), we consider utility functions that are quasi-linear in consumption. The worker's surplus is now  $x_a \equiv w_a - T(w_a) - v(h_a) - b$  with  $v(h_a)$ , the disutility of work,  $v(h_a)' > 0$  and  $v''(h_a) > 0$ . Let  $\eta$  denote the elasticity of labor supply, so  $1/\eta = h \cdot v''(h) / v'(h)$ . Employment rates are assumed equal to 1. Net output is given by  $Y_a = a \cdot h_a - v(h_a)$ . Gross wages  $w_a$  are here equal to  $a \cdot h_a$ . Labor supply  $h_a$  solves:

$$v'(h_a) = a \cdot (1 - T'(a \cdot h_a)) \quad (22)$$

$\Sigma_a$  now corresponds to the workers' surplus for an optimal labor supply:

$$\Sigma_a = x_a = a \cdot h_a - v(h_a) - T(a \cdot h_a) - b$$

$\Sigma_a$  evolves according to:

$$\dot{\Sigma}_a = h_a (1 - T'(w_a)) = h_a \frac{v'(h_a)}{a} \quad (23)$$

In this setting *à la* Mirrlees, the optimal income taxation solves:

$$\begin{aligned} & \max_{a_d, h_a, \Sigma_a, b} F(a_d) \Phi(b+d) + \int_{a_d}^{a_1} \Phi(\Sigma_a + b) f(a) da \\ & s.t. : \int_{a_d}^{a_1} \{a \cdot h_a - v(h_a) - \Sigma_a\} \cdot f(a) da = b + E \\ & \dot{\Sigma}_a = h_a \frac{v'(h_a)}{a} \quad \begin{cases} \Sigma_{a_d} = d & \text{if } a_d > a_0 \\ \Sigma_{a_d} \geq d & \text{if } a_d = a_0 \end{cases} \end{aligned}$$

Optimal taxation verifies the following condition:

$$\frac{T'(w_a)}{1 - T'(w_a)} = \frac{\int_{a_d}^{a_1} \left(1 - \frac{\Phi'_t}{\lambda}\right) f(t) dt}{a \cdot f(a)} \left(1 + \frac{1}{\eta}\right)$$

which coincides with the classic formula provided by the standard literature when utility is quasi-linear in consumption (see e.g. Diamond (1998)).

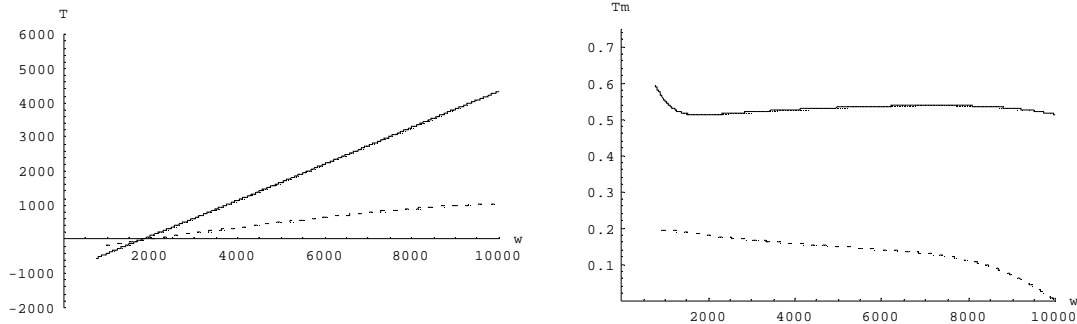


Figure 10: Benchmark comparison of the two models. Dotted lines for the Mirrlees model.

Model	Distribution	Net output $Y$			Welfare $\Phi^{-1}(\Omega)$			$b$	Participation SB
		WT	SB	$\Delta$	WT	SB	$\Delta$		
Benchmark	Benchmark	1794	1747	-2.7%	1627	1738	6.8%	809	94.0%
Mirrlees	Benchmark	2162	1957	-9.5%	1335	1365	2.3%	262	98.8%
Benchmark	With Pareto	2378	2305	-3.1%	1905	2258	18.5%	1254	93.7%
Mirrlees	With Pareto	2748	2340	-14.7%	1538	1652	7.5%	422	99.0%

Table 3: Numerical results WT for the economy without taxes and SB for second best

To compare it with the benchmark model of Sections II and III, we calibrate the Mirrlees model in such a way that in the economy without taxes both models have the same distributions of wages<sup>12</sup> and the same elasticity of gross wages with respect to one minus the marginal tax rate. From Equation (12), this elasticity equals  $1 - \gamma$  in our model<sup>13</sup>, whereas it equals  $\eta$  in the Mirrlees model. Hence, we take  $\eta = 0.5$ .

As Figure 10 and Table 3 show, the differences between the Mirrlees setting and ours are quantitatively very important. The optimum is much more redistributive when wages are bargained over and the intensive margin is neglected. Marginal tax rates are more than twice higher. Assistance benefits are almost three times greater. The gain in welfare is considerably higher and the loss in net output is much lower. Furthermore, the profile of marginal tax rates is substantially different.

Two major mechanisms are at work. First, the profiles of efficient marginal tax rates differ. In the Mirrlees model, efficient marginal tax rates are nil (since lump-sum transfers are the only way to redistribute income without distorting labor supply). Conversely, in our model, efficient marginal tax rates are positive according to Proposition 5. Furthermore, according to unreported simulations, efficient marginal tax rates are increasing with type  $a$ .

Second, in both models, marginal tax rates are above their efficient values, except at the top of the distribution. This prevents more productive workers from mimicking. As we move to the left of the distribution, the fraction of workers potentially involved in mimicking others increases. This generates a greater and greater upward pressure on marginal tax rates. In our model, the incentive compatibility constraint is expressed in

<sup>12</sup>This implies that the distribution of abilities in the “Mirrlees setting” has been appropriately reparametrized.

<sup>13</sup>We derive  $\frac{\partial w_a}{\partial(1-T'_a)} \frac{1-T'_a}{w_a}$  and evaluate this expression at  $T'_a = T_a = b = 0$ .

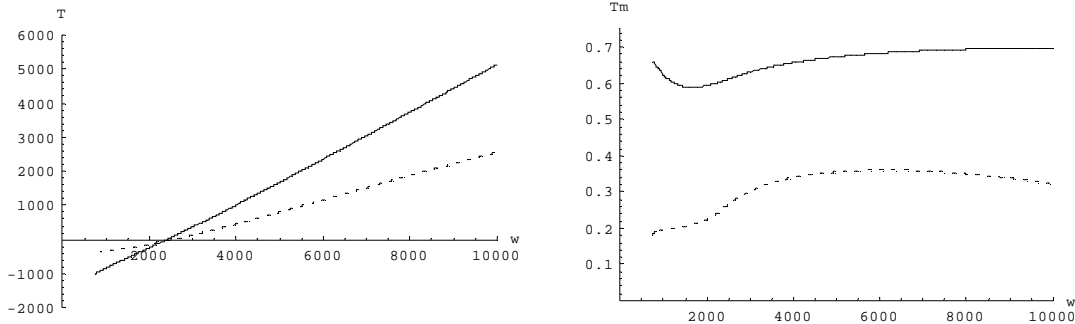


Figure 11: Comparison of the two models with a Pareto function at the top of the distribution. Dotted lines for the Mirrlees model.

terms of growth rates of workers' expected surplus (see Equation 17). In the Mirrlees version, the incentive constraint is formulated in terms of absolute changes (see Equation 23). Hence, in our model the upward pressure on marginal tax rates is stronger at the low end of the distribution.

Figure 10 assumes a lognormal distribution. However, within the Mirrlees' approach, it is well known that the tax schedule is highly sensitive to the skill distribution chosen. In particular, several recent authors (Diamond (1998) and Saez (2001)) have moved away from the lognormal assumption. They argue that the upper part of the productivity distribution is well approximated by an unbounded Pareto function. Under this assumption, they show that optimal marginal tax rates are much higher at the top of the distribution. In Figure 11, we conduct a sensitivity analysis. We take the same lognormal distribution as before for the lower-part of the distribution but we assume a Pareto distribution with density  $f_{par}(a) = K_p/a^{1+\pi}$  for the upper-part. Following Saez (2001), we take  $\pi = 2$ . The boundary between the two densities and  $K_p$  are chosen in such a way that the entire distribution is continuously differentiable. This leads to a boundary productivity level of  $a_{lim} = 5510$ . For this distribution, Figure 11 compares the optimal tax schedules in our model and in the Mirrlees one<sup>14</sup>. As expected, in the latter model, marginal tax rates significantly increase in the upper part of the wage distribution. However, marginal tax rates remain below our optimal ones. Our optimal marginal tax rates also shift upwards but to a lesser extent.

## V Introducing risk aversion

To check the robustness of our analytical results, this section introduces risk averse workers and a pure utilitarian criterion (keeping hours of work exogenous). In this setting, the rationale for public policy is twofold. On the one hand, the government wants to compensate individuals with different productivities. On the other hand, it wants to insure workers against the unemployment risk.

Let  $u(\cdot)$  be the workers' utility function with  $u'(\cdot) > 0$  and  $u''(\cdot) < 0$ . Let  $\Xi(\cdot)$  be its

<sup>14</sup>To approximate an unbounded distribution, we have now truncated the Pareto distribution at  $a_1 = 40\,000$ .



inverse function (for any  $x \geq 0$ ,  $\Xi(u(x)) \equiv x$ ), with:

$$\Xi'(x) = \frac{1}{u'(\Xi(x))} > 0 \quad \text{and} \quad \Xi''(\cdot) > 0$$

The government's objective becomes

$$\Psi = F(a_d) \cdot [u(b) + d] + \int_{a_d}^{a_1} \{L_a \cdot u(w_a - T(w_a)) + (1 - L_a) u(b)\} f(a) da$$

Labor demand and net output functions remain obviously unchanged. The negotiated wage solves now:

$$\Sigma_a \equiv \max_{w_a} A^{\frac{1}{\gamma}} \cdot \left( \frac{a - w_a}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} (u(w_a - T(w_a)) - u(b)) \quad (24)$$

Hence, as before,  $\Sigma_a$  corresponds to the worker's expected utility once the wage has been negotiated. Therefore, the government's objective can be rewritten as:

$$\Psi = F(a_d) \cdot [u(b) + d] + \int_{a_d}^{a_1} (\Sigma_a + u(b)) f(a) da \quad (25)$$

The incentive compatibility constraints remain (17) and  $\dot{w}_a > 0$ . The participation constraints are  $\Sigma_a \geq d$ . For the budget constraint, one has to note that  $u(w_a - T(w_a)) = \frac{\Sigma_a}{L_a} + u(b)$ , so:

$$w_a - T(w_a) - b = \Xi \left( \frac{\Sigma_a}{L_a} + u(b) \right) - b \quad (26)$$

Hence, the budget constraint (4) can be rewritten as:

$$\int_{a_d}^{a_1} \left\{ Y_a(w_a) - L_a(w_a) \left[ \Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - b \right] \right\} f(a) da = b + E \quad (27)$$

The government's problem now becomes the maximization of (25) subject to the budget constraint (27), the incentive constraints (17) and the participation constraints.

In Appendix C, we prove the following results. For all participating types, including the one at the top of the distribution, employment is higher than its efficient value<sup>15</sup>. Furthermore, if an in-work benefit is given to low-skilled employed workers, this in-work benefit is lower than the assistance benefit  $b$ . Finally, marginal tax rates are positive, including at the top of the distribution.

Compared to the risk-neutral case, the same mechanisms are here at work. In particular, adverse selection constraints tend to decrease gross wages and the participation rate. However, an additional mechanism reinforces the overemployment result. This mechanism comes from the inability to perfectly insure risk-averse workers against unemployment. Here, perfect insurance means that the level of income is the same whether the individual is employed or not. The participation constraint rules out this possibility<sup>16</sup>. For a given level of workers' expected utility (and so, for a given level of the Nash product  $\Sigma_a$ ), an additional decrease in the gross wage has now a first-order positive effect on efficiency.

<sup>15</sup>Here and in the rest of this section, efficiency has still the meaning introduced in Section II.3

<sup>16</sup>This would still be true if the assistance benefit was skill-specific.

It decreases workers' ex-post surplus  $x_a$  and it increases employment  $L_a$ . Therefore, it reduces the income risk supported by workers. Hence, a given level of expected utility can be guaranteed with less resources.

There is no proof that average tax rates should be increasing. Compared to risk neutral workers, risk averse ones bargain less aggressively over wages. Under the Hosios condition, it is then easily shown that wages lie below their efficient levels in an economy without taxes. In presence of a government with a redistributive objective and a concern for insurance, it is therefore no longer necessary to have marginal tax rates higher than average tax rates to decentralize the optimum.

## VI Conclusion

The optimal income taxation literature has essentially focused on distortions created through the consumption-leisure trade-off. This trade off is however not the unique way of explaining that earnings are affected by the profile of taxes. We have adopted an alternative setting where frictions on the labor market generate involuntary unemployment and rents to be shared by employers and employees. In this framework with exogenous working hours, the optimal income taxation has properties that strongly differ from those found in the Mirrlees competitive setting. Employment is higher than at the equilibrium without taxes, average tax rates are increasing in wages and marginal tax rates are strictly positive including at the top of the wage distribution. Compared to the prescriptions of a comparable Mirrlees setting, our numerical simulations show that assistance benefits are always by and large twice higher.

In sum, estimating the elasticity of gross earnings with respect to taxes is not sufficient to derive clear policy recommendations about the optimal tax schedule. One needs in addition to clarify which theoretical setting is empirically the most relevant. We left this for further research.

This paper also points to many interesting theoretical extensions. First, the assumption that employment is efficient in the no tax equilibrium could be relaxed. Second, the modeling of the extensive margin could be enriched. Finally, our contribution has been essentially methodological. Numerical simulations have therefore not tried to exploit rich datasets. All these extensions are also left for further research.

## Appendix

### A The first-best optimum

The first-order conditions of problem (8) are for each  $a$ :

$$0 = \{\Phi'(\Sigma_a + b) - \lambda\} p_a \cdot f(a) \quad (\Sigma_a)$$

$$0 = \int_{a_0}^{a_1} \{p_a \cdot \Phi'(\Sigma_a + b) + (1 - p_a) \Phi'(b + d)\} f(a) da - \lambda \quad (b)$$

$$0 = \lambda \cdot \frac{\partial Y_a}{\partial w_a} \cdot p_a \cdot f(a) \quad (w_a)$$

$$p_a = \begin{cases} 1 & \text{if } 0 \leq \\ 0 & \text{if } > \end{cases} \{\Phi(\Sigma_a + b) - \Phi(b + d) + \lambda[Y_a(w_a) - \Sigma_a]\} f(a) \quad (p_a)$$

where  $\lambda$  is the Lagrange Multiplier of the budget constraint.

Hence, at the first-best optimum, every participating type at the first best receives the same expected surplus  $\Sigma_a^*$  whose value is defined by  $\Phi'(b + \Sigma_a^*) = \lambda$ . Taking the condition on  $b$  into account, one further gets  $\lambda = \Phi'(\Sigma_a^* + b) = \Phi'(d + b)$  so  $\Sigma_a^* = d$ . The condition on  $w_a$  means that the wage level is fixed at its efficient value,  $w_a^*$ . From Equation (7), we then have:

$$\frac{\partial Y_a}{\partial w_a}(w_a) = \frac{\gamma \cdot a - w_a}{\gamma} (a - w_a)^{\frac{1}{\gamma}-2} \cdot \kappa_a^{\frac{\gamma-1}{\gamma}} \quad (28)$$

Therefore, the efficient wage  $w_a^*$  equals  $\gamma \cdot a$ .

Finally, the condition on participation can be simplified as:

$$p_a = \begin{cases} 1 & \text{if } d \leq Y_a(w_a^*) \\ 0 & \text{if } d > Y_a(w_a^*) \end{cases}$$

Hence, every type above (below)  $a_d^*$  should participate (be inactive) at the first best, where  $a_d^*$  is the unique solution in  $a$  to:

$$Y_a(w_a^*) = d$$

## B The second best

### B.1 The incentive compatibility constraints

This section follows Salanié (1997) very closely. Let  $\mathcal{N}(a, t)$  be the logarithm of the Nash product for a type- $a$  job when the negotiated wage is the one designed for type  $t$ -jobs. So

$$\mathcal{N}(a, t) \equiv \log N(a, w_t, x_t) = \frac{1-\gamma}{\gamma} \log\left(\frac{a-w_t}{\kappa_a}\right) + \log(w_t - T(w_t) - b) + \frac{1}{\gamma} \log A$$

and <sup>17</sup>:

$$\frac{\partial \mathcal{N}}{\partial a}(a, t) = \frac{1-\gamma}{\gamma} \left( \frac{1}{a-w_t} - \frac{\dot{\kappa}_a}{\kappa_a} \right) \quad \frac{\partial^2 \mathcal{N}}{\partial a \partial t}(a, t) = \frac{1-\gamma}{\gamma} \frac{\dot{w}_t}{(a-w_t)^2} \quad (29)$$

Equation (16) means that the function  $t \mapsto \mathcal{N}(a, t)$  reaches a maximum for  $t = a$ . So,  $\log \Sigma_a = \mathcal{N}(a, a)$ . The first-order condition can be written as  $\frac{\partial \mathcal{N}}{\partial t}(a, a) = 0$ . So, for any  $a$

$$\frac{\dot{\Sigma}_a}{\Sigma_a} = \frac{\partial \mathcal{N}}{\partial a}(a, a) + \frac{\partial \mathcal{N}}{\partial t}(a, a) = \frac{\partial \mathcal{N}}{\partial a}(a, a)$$

which, combined with (29) leads to (17). Furthermore, since  $\frac{\partial \mathcal{N}}{\partial t}(a, a) = 0$  for all  $a$ , one has  $\frac{\partial^2 \mathcal{N}}{\partial t \partial a}(a, a) + \frac{\partial^2 \mathcal{N}}{\partial t \partial t}(a, a) = 0$ . So, the second-order condition  $\frac{\partial^2 \mathcal{N}}{\partial t \partial t}(a, a) < 0$  is equivalent to  $0 < \frac{\partial^2 \mathcal{N}}{\partial t \partial a}(a, a)$  for all  $a$ . From (29) the second-order condition requires that for all  $a$ :

$$\dot{w}_a > 0$$

Finally, one has to verify that these local conditions are sufficient for (15). For any  $a$  and any  $t \neq a$  there exists  $\theta \in (0, 1)$  such that for  $\hat{t} = \theta a + (1 - \theta)t$

$$\mathcal{N}(a, a) - \mathcal{N}(a, t) = \frac{\partial \mathcal{N}}{\partial t}(a, \hat{t}) \cdot (a - t)$$

---

<sup>17</sup>It is here assumed that the mechanism  $a \rightarrow (w(a), x(a))$  is differentiable so  $\dot{w}_a$  exists.

Provided that for all  $t$ ,  $\dot{w}_t > 0$ , one has  $\frac{\partial^2 \mathcal{N}}{\partial a \partial t}(a, t) > 0$  and therefore  $\frac{\partial \mathcal{N}}{\partial t}(a, t)$  is increasing in  $a$ . Since  $\frac{\partial \mathcal{N}}{\partial t}(\hat{t}, \hat{t}) = 0$ , this implies that  $\frac{\partial \mathcal{N}}{\partial t}(a, \hat{t}) \geq 0$  if  $a \geq \hat{t}$ , that is if  $a \geq t$ . Hence  $\frac{\partial \mathcal{N}}{\partial t}(a, \hat{t}) \cdot (a - t) > 0$  and  $t = a$  is a global maximum for  $t \mapsto \mathcal{N}(a, t)$ .

## B.2 The first-order conditions of the optimization problem

We solve problem (20) in two steps. First, we solve for given values of  $b$  and  $a_d$ . Second, we characterize the optimal values of  $b$  and  $a_d$ . Given  $b$  and  $a_d$ , we define the Hamiltonian for  $a \in [a_d, a_1]$  as:

$$\mathbb{H}_a = \{\Phi(\Sigma_a + b) + \lambda \cdot Y_a(w_a) - \lambda \cdot \Sigma_a\} f(a) + q_a \cdot \frac{1 - \gamma}{\gamma} \left( \frac{1}{a - w_a} - \frac{\dot{\kappa}_a}{\kappa_a} \right) \Sigma_a \quad (30)$$

where  $\lambda$  is the Lagrange multiplier associated with the budget constraint and  $q$  is the co-state variable. The necessary conditions are:

$$\lambda \cdot \frac{\partial Y_a}{\partial w_a} \cdot f(a) + q_a \cdot \Sigma_a \cdot \frac{1 - \gamma}{\gamma (a - w_a)^2} = 0 \quad (w_a)$$

The co-state variable evolves according to:

$$-\dot{q}_a = \{\Phi'_a - \lambda\} f(a) + q_a \frac{\dot{\Sigma}_a}{\Sigma_a} \quad (\Sigma_a)$$

and the transversality conditions are:

$$q_{a_d} \cdot [\Sigma_{a_d} - d] = 0 \quad q_{a_1} = 0$$

As usual,  $q_a$  is the shadow cost of a marginal increase in  $\Sigma_a$ . Define  $Z_a = q_a \cdot \Sigma_a$ . The condition for  $\Sigma_a$  implies:

$$-\dot{Z}_a = \{\Phi'_a - \lambda\} \Sigma_a f(a) \quad (31)$$

So, together with the transversality condition:

$$Z_a = \int_a^{a_1} \{\Phi'_t - \lambda\} \Sigma_t \cdot f(t) \cdot dt \quad (32)$$

$Z_a$  corresponds to the opposite of the integral on the right hand side of Equation (21). Since  $Z_a \cdot \frac{d\Sigma_a}{\Sigma_a} = q_a \cdot d\Sigma_a$ ,  $Z_a$  stands for the shadow cost of a relative marginal increase in  $\Sigma_a$ .

The first order condition w.r.t.  $w_a$  can be written as

$$\lambda \cdot \frac{\partial Y_a}{\partial w_a} \cdot f(a) = -Z_a \frac{1 - \gamma}{\gamma (a - w_a)^2} \quad (33)$$

which, together with the expression for  $Z_a$  gives (21). The conditions with respect to  $b$  and  $a_d$  are (see Leonard and Van Long (1992)):

$$\int_{a_0}^{a_1} (\Phi'_a - \lambda) f(a) da = 0 \quad (34)$$

$$\Phi(b + d) f(a_d) - \mathbb{H}_{a_d} \leq 0 \quad \text{with } = \text{ if } a_d > a_0 \quad (35)$$

### B.2.1 Proof of Propositions 1 and 2

For  $a = a_1$ . The transversality condition  $q_{a_1} = 0$  implies that the integral in the right hand side of Equation (21) is nil for  $a = a_1$ , so  $w_{a_1} = w_{a_1}^*$  and  $L_{a_1} = L_{a_1}^*$ .

For  $a \in [a_d, a_1)$ . Since  $\Sigma_a$  is increasing in  $a$  by Equations (1) and (17),  $\Phi'_a$  is decreasing in  $a$ . Equation (34) implies that there exists a unique  $\hat{a}$  such that  $\Phi'_{\hat{a}} = \lambda$ . For  $t < \hat{a}$ , we get  $\Phi'_t - \lambda > 0$  and  $\Sigma_t < \Sigma_{\hat{a}}$  and for  $t > \hat{a}$ , we get  $\Phi'_t - \lambda < 0$  and  $\Sigma_t > \Sigma_{\hat{a}}$ . Therefore, for any  $t \neq \hat{a}$ , we have  $(\Phi'_t - \lambda) \Sigma_t < (\Phi'_t - \lambda) \Sigma_{\hat{a}}$ . Using this inequality and Equations (34) and (32), we obtain

$$\begin{aligned} Z_a &= \int_a^{a_1} (\Phi'_t - \lambda) \Sigma_t \cdot f(t) \cdot dt < \int_a^{a_1} (\Phi'_t - \lambda) \Sigma_{\hat{a}} \cdot f(t) \cdot dt \\ &< \Sigma_{\hat{a}} \left[ \int_a^{a_1} \Phi'_t \cdot f(t) \cdot dt - \lambda(1 - F(a)) \right] = \Sigma_{\hat{a}} \cdot (1 - F(a)) \cdot \{ \mathbb{E}_f [\Phi'_t | t \geq a] - \lambda \} \end{aligned}$$

where

$$\mathbb{E}_f [\Phi'_t | t \geq a] = \frac{\int_a^{a_1} \Phi'_t \cdot f(t) dt}{1 - F(a)}$$

Hence

$$Z_a < \Sigma_{\hat{a}} \cdot (1 - F(a)) \cdot \{ \mathbb{E}_f [\Phi'_t | t \geq a] - \mathbb{E}_f [\Phi'_t | t \geq a_0] \}$$

by Equation (34). Therefore,  $Z_a$  is negative for all  $a < a_1$  because  $\Phi'_t$  is decreasing with respect to the productivity. From (33), we obtain  $\partial Y_a / \partial w_a > 0$ , so from (28), we have over-employment for all types  $a < a_1$ .

### B.2.2 Proof of Propositions 3 and 4

We will prove that  $a_d \geq a_d^*$ . From the first order condition on  $a_d$ , we have:

$$0 \geq f(a_d) \Phi(b+d) - \{ \Phi(\Sigma_{a_d} + b) + \lambda \cdot (Y_{a_d} - \Sigma_{a_d}) \} f(a_d) - Z_{a_d} \frac{1-\gamma}{\gamma} \left[ \frac{1}{(a_d - w_{a_d})} - \frac{\dot{\kappa}_{a_d}}{\kappa_{a_d}} \right]$$

Since  $Z_a$  is always negative for  $a < a_1$ , the transversality condition on  $a_d$  implies that  $\Sigma_{a_d} = d$ . Rearranging the terms, we get:

$$Y_{a_d} - \Sigma_{a_d} \geq -Z_{a_d} \cdot \frac{\dot{\Sigma}_{a_d}}{\Sigma_{a_d}} \cdot \frac{1}{\lambda f(a_d)} > 0 \quad (36)$$

So,  $d = \Sigma_{a_d} < Y_{a_d}$ . Furthermore, we get  $Y_{a_d} < Y_{a_d}^*$  since  $Y_{a_d}^*$  corresponds to the efficient value at  $a_d$ . Hence, we get

$$Y_{a_d}^* > d$$

So, given Equation (10) we have two possible cases. First,  $a_d = a_0 = a_d^*$ . Second,  $a_d > a_0$  with  $Y_{a_d}^* > Y_{a_d}^* = d$ , which implies that  $a_d > a_d^*$ .

$Y_{a_d} > \Sigma_{a_d}$  implies  $w_{a_d} > x_{a_d}$  so  $T_{a_d} + b > 0$  and  $b > -T_{a_d}$ .  $-T_{a_d}$  should be understood as the in-work benefits for the least skilled workers who participate to the labor market.

### B.2.3 Proof of Proposition 5

From Equations (9) and (13) the gross wage is below its efficient level if and only if

$$T'_a > \frac{T_a + b}{w_a}$$

Since

$$\frac{\partial \left( \frac{T(w_a)}{w_a} \right)}{\partial w_a} = \frac{T'_a - \frac{T(w_a)}{w_a}}{w_a}$$

Proposition 2 implies that the average tax rate is increasing in the wage. Finally, according to Proposition 4, one has  $T'(w_a) > (T(w_a) + b)/w_a \geq (T(w_{a_d}) + b)/w_{a_d} > 0$ .

## C The case under risk aversion

For  $a \geq a_d$  the Hamiltonian writes:

$$\begin{aligned} \mathbb{H}_a = & \left\{ \Sigma_a + u(b) + \lambda \left[ Y_a(w_a) - L_a(w_a) \left( \Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - b \right) \right] \right\} f(a) \\ & + q_a \frac{1-\gamma}{\gamma} \left( \frac{1}{a-w_a} - \frac{\dot{\kappa}_a}{\kappa_a} \right) \Sigma_a \end{aligned}$$

Let  $Z_a = q_a \Sigma_a$  and  $\Xi'_a = \Xi' \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right)$ , so

$$\Xi'_a = \frac{1}{u'(w_a - T(w_a))}$$

The first-order conditions with respect to respectively  $\Sigma_a$ ,  $w_a$ ,  $b$  and  $a_d$  are:

$$-\dot{Z}_a = (1 - \lambda \cdot \Xi'_a) \Sigma_a \cdot f(a) \quad (37)$$

$$\begin{aligned} 0 = & \lambda \left\{ \frac{\partial Y_a}{\partial w_a} - \frac{\partial L_a}{\partial w_a} \left[ \Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - b - \frac{\Sigma_a}{L_a} \cdot \Xi'_a \right] \right\} f(a) \\ & + \frac{1-\gamma}{\gamma} \cdot \frac{Z_a}{(a-w_a)^2} \end{aligned} \quad (38)$$

$$\lambda = \int_{a_0}^{a_1} \{ u'(b) (1 - \lambda \cdot L_a \cdot \Xi'_a) + \lambda \cdot L_a \} f(a) da \quad (39)$$

$$0 \geq f(a_d) [u(b) + d] - \mathbb{H}_{a_d} \quad (40)$$

together with the transversality conditions:  $Z_{a_1} = 0$  and  $q_{a_d} [\Sigma_{a_d} - d] = 0$ .

First, we show that for any  $a \in [a_d, a_1)$ ,  $Z_a < 0$ . From (37) and the transversality condition  $Z_{a_1} = 0$ , one has:

$$Z_a = \int_a^{a_1} (1 - \lambda \cdot \Xi'_t) \Sigma_t \cdot f(t) dt \quad (41)$$

Some manipulations of (39) leads to:

$$\int_{a_0}^{a_1} (1 - \lambda \cdot \Xi'_a) \cdot f(a) da = \frac{\lambda}{u'(b)} \int_{a_0}^{a_1} (1 - L_a) (1 - u'(b) \cdot \Xi'_a) \cdot f(a) da$$

but

$$\Xi'_a \cdot u'(b) = \frac{u'(b)}{u'(w_a - T(w_a))} > 1$$

so

$$\mathbb{E}_f [1 - \lambda \cdot \Xi'_t] = \int_{a_0}^{a_1} (1 - \lambda \cdot \Xi'_a) \cdot f(a) da < 0$$

Finally, function  $a \mapsto \Sigma_a + b$  is increasing by (17) and (1), so function  $a \mapsto 1 - \lambda \cdot \Xi'_a$  is decreasing. Hence,  $1 - \lambda \cdot \Xi'_{a_1} < 0$ . Two cases are then possible.

1. If  $1 - \lambda \cdot \Xi'_{a_d} \leq 0$ , then for all  $t \in (a_d, a_1)$ , one has  $(1 - \lambda \cdot \Xi'_t) \Sigma_t < 0$ , so  $Z_a < 0$  by (41).
2. If  $1 - \lambda \cdot \Xi'_{a_d} > 0$ , then there exists a single  $\hat{t}$ , such that  $1 - \lambda \cdot \Xi'_{a_{\hat{t}}} = 0$ . But then for any  $t \in [a_d, a_{\hat{t}}) \cup (a_{\hat{t}}, a_1]$ , one has:

$$(1 - \lambda \cdot \Xi'_t) \Sigma_t < (1 - \lambda \cdot \Xi'_{\hat{t}}) \Sigma_{\hat{t}}$$

so

$$\begin{aligned} Z_a &< \Sigma_{\hat{t}} \int_a^{a_1} (1 - \lambda \cdot \Xi'_t) f(t) dt = \Sigma_{\hat{t}} (1 - F(a)) \cdot \mathbb{E}_f [1 - \lambda \cdot \Xi'_t | t \geq a] \\ &< \Sigma_{\hat{t}} (1 - F(a)) \cdot \mathbb{E}_f [1 - \lambda \cdot \Xi'_t] < 0 \end{aligned}$$

Second, (38) can be rewritten as:

$$\lambda \cdot \frac{\partial Y_a}{\partial w_a} \cdot f(a) = -\frac{1-\gamma}{\gamma} \cdot \frac{Z_a}{(a-w_a)^2} + \lambda \cdot \frac{\partial L_a}{\partial w_a} \left[ \Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - b - \frac{\Sigma_a}{L_a} \cdot \Xi'_a \right] \cdot f(a)$$

However, by the strict convexity of  $\Xi$ , one has:

$$\Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - \Xi(u(b)) < \frac{\Sigma_a}{L_a(w_a)} \cdot \Xi'_a$$

so  $\Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - b - \frac{\Sigma_a}{L_a} \cdot \Xi'_a < 0$ , therefore:

$$\frac{\partial L_a}{\partial w_a} \left[ \Xi \left( \frac{\Sigma_a}{L_a(w_a)} + u(b) \right) - b - \frac{\Sigma_a}{L_a} \cdot \Xi'_a \right] > 0$$

and finally,  $\frac{\partial Y_a}{\partial w_a} > 0$  for any  $a \in [a_d, a_1]$ . Hence there is overemployment over the whole distribution of productivity, including at  $a = a_1$ .

Third, since  $Z_{a_d} < 0$ , the transversality condition implies  $\Sigma_{a_d} = d$ . So, from (40), one gets:

$$0 \leq \lambda \left\{ Y_{a_d} - L_{a_d} \left[ \Xi \left( \frac{\Sigma_{a_d}}{L_{a_d}(w_{a_d})} + u(b) \right) - b \right] \right\} + Z_{a_d} \frac{1-\gamma}{\gamma} \left( \frac{1}{a_d - w_{a_d}} - \frac{\dot{\kappa}_{a_d}}{\kappa_{a_d}} \right)$$

Since  $Z_{a_d} < 0$  and  $\frac{1}{a_d - w_{a_d}} > \frac{1}{a_d} > \frac{\kappa_{a_d}}{\kappa_{a_d}}$  (by (1)), this implies that:

$$L_{a_d} \left[ \Xi \left( \frac{\Sigma_{a_d}}{L_{a_d}(w_{a_d})} + u(b) \right) - b \right] < Y_{a_d} = w_{a_d} L_{a_d}$$

so, together with (26), one has  $b + T(w_{a_d}) > 0$ .

Fourth, the first-order condition of the optimization program with respect to the wage (38) and the negativity of  $Z_a$  lead to:

$$\frac{\partial Y_a}{\partial w_a} - \frac{\partial L_a}{\partial w_a} \left[ w_a - T(w_a) - b - \frac{u(w_a - T(w_a)) - u(b)}{u'(w_a - T(w_a))} \right] > 0$$

Since,

$$\frac{\partial L_a}{\partial w_a} = -\frac{1-\gamma}{\gamma(a-w_a)} L_a \quad \frac{\partial Y_a}{\partial w_a} = \left( 1 - \frac{(1-\gamma)w_a}{\gamma(a-w_a)} \right) L_a = \frac{\gamma \cdot a - w_a}{\gamma(a-w_a)} L_a$$

we get:

$$\frac{u(w_a - T(w_a)) - u(b)}{u'(w_a - T(w_a))} < \frac{\gamma \cdot a - w_a}{1-\gamma} + w_a - T(w_a) - b$$

Furthermore, the first-order condition of the wage bargaining program implies

$$\frac{u(w_a - T(w_a)) - u(b)}{u'(w_a - T(w_a))} = \frac{\gamma(1 - T'_a)}{1-\gamma} (a - w_a)$$

Putting these two Equations together implies

$$\frac{\gamma}{1-\gamma} (1 - T'_a) (a - w_a) < \frac{\gamma}{1-\gamma} (a - w_a) - T(w_a) - b$$

so

$$\frac{1-\gamma}{\gamma} \cdot \frac{T(w_a) + b}{a - w_a} < T'_a$$

Since  $T(w_{a_d}) + b > 0$ , this implies that  $T(w_a) + b$  is positive and increasing over the whole distribution. It is therefore positive everywhere, so marginal tax rates are positive, including at the top of the distribution.

## References

- [1] Blanchard, O. J. and Diamond, P., 1989, "The Beveridge Curve", *Brookings Papers on Economic Activity*, 1, 1-76.
- [2] Blundell, R. and MaCurdy, T., 1999, "Labor Supply: A Review of alternative approaches", in O. Ashelfelter and D. Card (eds), *Handbook of Labor Economics*, North-Holland.
- [3] Boadway, R., Cuff, K. and Marchand, M., 2000, "Optimal Income Taxation with Quasi-Linear Preferences Revisited", *Journal of Public Economic Theory*, 2, 435-60.



- [4] Boone, J. and Bovenberg, L., 2002, “Optimal Labor Taxation and Search”, *Journal of Public Economics*, 85, 53-97.
- [5] Boone, J. and Bovenberg, L., 2004, “The optimal taxation of unskilled labor with job search and social assistance”, *Journal of Public Economics*, 88(11), 2227-2258.
- [6] Boone, J. and Bovenberg L., 2005, “Optimal welfare and in-work benefits with search unemployment and observable abilities”, *Journal of Economic Theory*, Forthcoming.
- [7] Choné, P. and Laroque, G., 2005, “Optimal Incentives for Labor Force Participation” *CREST-INSEE Discussion paper* 25, Forthcoming in *Journal of Public Economics*.
- [8] Cremer, J. and McLean, R., 1988, “Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions”, *Econometrica*, 56(6), 1247-1258.
- [9] Diamond, P. 1998, “Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates”, *American Economic Review*, 88, 83-95.
- [10] Drèze, J. and Malinvaud, E., 1994, “Growth and employment: The scope of a European Initiative”, *European Economic Review*, 38(3-4), 489-504.
- [11] Engström, P., 2002, “Optimal Non-Linear Taxation in Search Equilibrium”, Working Paper 2002:18, Uppsala University.
- [12] Feldstein, M., 1995, “The Effect of Marginal Tax Rates on Taxable Income: A Panel Study on the 1986 Tax Reform Act”, *Journal of Political Economy*, 103 (3), 551-572.
- [13] Guesnerie, R., 1995, “*A contribution to the pure theory of taxation*”, Cambridge University Press.
- [14] Gruber, J. and Saez, E., 2002, “The elasticity of taxable income: evidence and implications”, *Journal of Public Economics*, 84(1), 1-32.
- [15] Hammond, P., 1979, Straightforward Individual Incentive Compatibility in Large Economies, *Review of Economic Studies*, 46(2), 263-282.
- [16] Hansen, C. T., Pedersen L. H. and Sløk., T., 2000, “Ambiguous effects of the marginal tax rate”, *Labour Economics*, 7, 335-347.
- [17] Holmlund, B and Kolm, A. S 1995, “Progressive Taxation, Wage setting, and Unemployment : Theory and Swedish evidence”, *Swedish Economic Policy Review* 2, 423-460.
- [18] Hosios, A.,1990, “On the Efficiency of Matching and Related Models of Search and Unemployment”, *Review of Economic Studies*, 57, 279-298.
- [19] Hungerbühler, M., 2004, “The Impact of Union Power on the Optimal Income Tax Schedule”, *IREES Discussion Paper* n°2004 – 34, Université Catholique de Louvain, Louvain La Neuve, Belgium.
- [20] Laffont, J. J and Martimort, D., 2002, *The theory of incentives: the principal agent model*, Princeton University Press.
- [21] Laroque, G., 2002, “Income Maintenance and Labor Force Participation”, *Presidential Address to the Econometric Society*.

- [22] Lehmann, E. and Van der Linden, B., 2002, “On the optimality of search matching equilibrium when workers are risk averse”, *IREES Discussion paper* n° 2002-23.
- [23] Leonard, D. and Van Long N., 1992, *Optimal control theory and static optimization in economics*, Cambridge University Press.
- [24] Lockwood B. and Manning A. , 1993, “Wage Setting and the Tax System: Theory and evidences for the United Kingdom”, *Journal of Public Economics*, 52, 1-29.
- [25] Malcomson, J. M. and Sartor, N., 1987, “Tax Push Inflation in a Unionized Labour Market”, *European Economic Review*, 31, 1581-1596.
- [26] Marceau, N. and Boadway, R., 1994, “Minimum Wage Legislation and Unemployment Insurance as Instruments for Redistribution”, *Scandinavian Journal of Economics*, 96(1), 67-81.
- [27] Mirrlees, J., 1971, “An Exploration in the Theory of Optimum Income Taxation”, *Review of Economic Studies*, 38, 175-208.
- [28] Mortensen, D. and Pissarides, C., 1999, “New developments in models of search in the Labor Market”, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol 3 (North-Holland, Amsterdam)
- [29] Musgrave, R. and Musgrave, P., 1976, *Public Finance in Theory and Practice*, Second Edition, McGraw-Hill, New York.
- [30] Phelps, E., 1997, *Rewarding work: How to restore participation and self-support to free enterprise*, Harvard University Press, Cambridge, US.
- [31] Petrongolo, B. and Pissarides, C. A. , 2001, “Looking into the Black Box: A Survey of the Matching Function”, *Journal of Economic Literature*, 38, 390-431.
- [32] Pissarides, C. A., 1998, “The impact of employment tax cuts on unemployment and wages: the role of unemployment benefits and tax structure”, *European Economic Review*, 42, 155-183.
- [33] Pissarides, C. A., 2000, *Equilibrium Unemployment Theory*, Second Edition, MIT Press, Cambridge, US (First edition published in 1990 by Basil Blackwell).
- [34] Salanié, B., 1997, *The Economics of Contracts: a Primer*, MIT Press.
- [35] Saez, E., 2001, “Using Elasticities to Derive Optimal Income Tax Rates”, *Review of Economic Studies*, 68, 205-229.
- [36] Saez, E., 2002, “Optimal Income Transfers Programs: Intensive versus Extensive Margin Labor Supply”, *Quarterly Journal of Economics*, 117, 1039-1073.
- [37] Sørensen, P. B., 1997, “Public Finance solutions to the european unemployment problem”, *Economic Policy*, 25, 223-264.
- [38] Sørensen, P. B., 1999 , “Optimal Tax Progressivity in imperfect Labour markets”, *Labour Economics*, 6, 435-452.
- [39] Stiglitz, J. E., 1982, “Self-Selection and Pareto Efficient Taxation”, *Journal of Public Economics*, 17, 213-240.

- [40] Stiglitz, J. E., 1987, "Pareto Efficient and Optimal Taxation and the New Welfare Economics", Chap 15, *Handbook of Public Economics*, vol II, Auerbach, A.J. and Feldstein, M. (eds), Elsevier Science Publishers B.V (North-Holland).
- [41] Tuomala, M., 1990, *Optimal Income Tax and Redistribution*, Oxford: Clarendon Press.