Optimal income taxation with endogenous participation and search unemployment

E. Lehmann, A. Parmentier and B. Van der Linden

Discussion Paper 2008-36

Département des Sciences Économiques de l'Université catholique de Louvain

UCL

Optimal income taxation with endogenous participation and search unemployment*

Etienne LEHMANN†
CREST, IZA, IDEP and
Université Catholique de Louvain

Alexis PARMENTIER‡
EPEE-TEPP - Université d'Evry and
Université Catholique de Louvain

Bruno VAN DER LINDEN§
IRES - Department of Economics - Université Catholique de Louvain,
FNRS, ERMES - Université Paris 2 and IZA

October 29, 2008

Abstract

This paper characterizes the optimal redistributive taxation when individuals are heterogeneous in two exogenous dimensions: their skills and their values of non-market activities. Search-matching frictions on the labor markets create unemployment. Wages, labor demand and participation are endogenous. The government only observes wage levels. Under a Maximin objective, if the elasticity of participation decreases along the distribution of skills, at the optimum, the average tax rate is increasing, marginal tax rates are positive everywhere, while wages, unemployment rates and participation rates are distorted downwards compared to their *laissez-faire* values. A simulation exercise confirms some of these properties under a general utilitarian objective. Taking account of the wage-cum-labor demand margin deeply changes the equity-efficiency trade-off.

Keywords: Non-linear taxation; redistribution; adverse selection; random participation; unemployment; labor market frictions.

JEL codes: D82; H21; J64

*We thank for their comments participants at seminars at the Université Catholique de Louvain, CREST, Malaga, EPEE-Evry, Gains-Le Mans, CES-Paris 1, ERMES-Paris 2, the IZA-SOLE 2008 Transatlantic meeting, the 7th Journées Louis-André Gérard-Varet, the HetLab conference at the University of Konstanz, The Economics of Labor Income taxation workshop at IZA, Thema-Cergy Pontoise, AFSE 2008, EALE 2008, with a particular mention to Pierre Cahuc, Helmuth Cremer, Laurence Jacquet, Guy Laroque, Cecila Garcia-Penlosa, Fabien Postel-Vinay. Mathias Hungerbühler was particularly helpful in provinding suggestions and comments at various stages. Any errors are ours. This research has been funded by the Belgian Program on Interuniversity Poles of Attraction (P6/07 Economic Policy and Finance in the Global Economy: Equilibrium Analysis and Social Evaluation) initiated by the Belgian State, Prime Minister's Office, Science Policy Programming.

†Address: CREST-INSEE, Timbre J360, 15 boulevard Gabriel Péri, 92245, Malakoff Cedex, France. Email: etienne.lehmann@ensae.fr.

‡Address: EPEE - Université d'Evry Val d'Essonne, 4 boulevard François Mitterand, 91025, Evry Cedex, France. Email: alexis.parmentier@univ-evry.fr

§Address: IRES - Département d'économie, Université Catholique de Louvain, Place Montesquieu 3, B1348, Louvain-la-Neuve, Belgium. Email: bruno.vanderlinden@uclouvain.be

I Introduction

In the literature on optimal redistributive taxation initiated by Mirrlees (1971), non-employment, if any, is synonymous with non-participation. The importance of participation decisions is not debatable. However, according to Mirrlees (1999), "another desire is to have a model in which unemployment [in our words, "non-employment"] can arise and persist for reasons other than a preference for leisure". Along this view, it is important to recognize that some people remain jobless despite they do search for a job at the market wage. To account for this fact, one should depart from the assumption of walrasian labor markets. Our paper characterizes the optimal redistribution policy in a framework where wages, employment, (involuntary) unemployment and (voluntary) non participation are endogenously affected by taxation on labor income.

As it is standard in the optimal tax literature, we assume that the government is only able to condition taxation on wages. Our economy is made of a continuum of skill-specific labor markets. On each of them, we introduce matching frictions à la Mortensen and Pissarides (1999). This setting is particularly attractive because both labor supply (along the participation/extensive margin) and labor demand determine the equilibrium levels of employment. In our model, taxes are distorsive via the participation margin and the wage-cum-labor demand margin. Concerning participation, we assume that whatever their skill level, individuals differ in their value of remaining out of the labor force.[1] A higher level of taxes reduces the skill-specific value of participation, thereby inducing some individuals to leave the labor force. Labor demand is affected by taxation through wage formation. In various wage-setting models, the equilibrium gross wage maximizes an objective that is increasing in the after-tax (net) wage and decreasing in the pre-tax (gross) wage. A higher pre-tax wage increases worker's consumption but, for instance, in a monopoly union model, it reduces the labor demand while it reduces firms' profit in Nash bargaining models. Since a compensated increase in tax progressivity renders a higher pre-tax wage less attractive to workers, a lower pre-tax wage is substituted for a lower after-tax wage.[2] This wage moderation effect of tax progressivity stimulates labor demand and reduces unemployment. In order to be as general as possible, we deal with wage-formation in a reduced form way that is consistent with those properties.

As it is standard in the optimal taxation literature, we stick to the welfarist view according to which the government's objective depends on utility levels. Moreover, in order to focus sharply on redistributive issues, we assume that the economy without taxes (*laissez faire*) is efficient

[1] Because of this additional unobserved heterogeneity, the government has to solve an adverse selection problem with « random participation » *à la* Rochet and Stole (2002).

[2] It is worth noting that this mechanism also holds in the textbook competitive labor supply framework. There, the after-tax wage equals consumption and a higher pre-tax wage is obtained thanks to more effort. Hence, solving the consumption/leisure tradeoff amounts to maximize an objective that is increasing in the net wage and decreasing in the gross wage. For simplicity, we ignore labor supply responses along the intensive margin in our model.

(in the Benthamite sense). When the government has a Maximin (Rawlsian) objective, it only values the utility level of the unemployed. The optimal tax schedule has then clear analytical properties if the elasticities of participation verify a monotonicity assumption. In the most plausible case, these elasticities are decreasing along the skill distribution, so the participation rates are more elastic for low skilled labor markets than for high skilled ones. Then, we prove that optimal marginal tax rates are positive everywhere and optimal average tax rates are increasing. The reason is that a more progressive tax schedule increases the level of tax at the top of the skill distribution where participation decisions are less elastic and decreases the level of tax where participation reacts more strongly to the tax pressure. Since redistribution lowers participants' expected surplus, the participation rate is lower at the optimum than at the *laissez faire*. However, a more progressive tax schedule distorts wages and unemployment rates downwards. Thus, at the optimum, the efficiency losses generated by the wage and employment distortions equalize the net gains due to redistribution.

We also derive the optimal tax formula under a general utilitarian criterion. As in the Maximin case, we provide a formula that expresses the optimal tax as a function of behavioral elasticities and the skill distribution. We also interpret this formula by considering a marginal tax reform around the optimum. Unemployment has now two effects on social welfare that cannot be recognized if the wage-cum-labor demand margin is ignored. First, since income net of taxes and transfers has to be higher in employment than in non-employment (to induce participation), unemployment *per se* causes a loss in social welfare. Second, because some participants to the labor market are eventually unemployed, enhancing participation increases earnings inequalities, which has a detrimental effect on social welfare. Compared to the Maximin case, these channels call for further downward distortions of wages and push down the optimal unemployment rates.

To illustrate how our optimal tax formulas could be used for applied purposes, we calibrate our model for the US economy. Our numerical results illustrate the properties found in the Maximin case and cast light on the more complex mechanisms at work in the general utilitarian case. In the Maximin case, it turns out that the optimal tax profile is well approximated by an assistance benefit tapered away at a high and nearly constant rate. If the government maximizes a Bergson-Samuelson social welfare function, the optimal tax profile is different with hump-shaped marginal tax rates. Moreover, an EITC can be optimal.

A number of studies are related to our work. In the optimal taxation literature that follows Mirrlees (1971), the intensive margin (i.e. work effort) is the only source of deadweight losses. In this competitive labor market model, tax progressivity induces a downward distortion of work effort and thus of pre-tax wages. In our non-competitive model, tax progressivity reduces pre-tax wages and increases labor demand. Thus, the equity-efficiency trade-off in our non walrasian labor market framework is dramatically different from the one appearing in the Mirrleesian literature. Both mechanisms can account for the empirical fact that gross incomes decrease with marginal tax rates (Feldstein, 1995, Gruber and Saez, 2002). Whether this wage moderating effect of tax progressivity is due to a labor supply response along the intensive margin or to a non competitive wage formation remains an open empirical question. However, we believe that the mechanism on which our model is based might be crucial. On the one hand, Blundell and MacCurdy (1999) and Meghir and Phillips (2008) conclude that the labor supply responses along the intensive margin are empirically very small. On the other hand, Manning (1993) finds a significantly negative effect of tax progressivity on the UK unemployment rate (see also Sørensen 1997 and Røed and Strøm 2002).

There is now growing evidence that the extensive margin (i.e. participation decisions) matters a lot. Diamond (1980) and Choné and Laroque (2005) have studied optimal income taxation when individuals' decisions are limited to a dichotomic choice about whether to work or not. The optimum trades off the equity gain of a higher level of tax against the efficiency loss of a lower level of participation. However, gross incomes are not distorted in these models because of a competitive labor market and exogenous productivity levels. Saez (2002) has proposed a model of optimal taxation where both extensive and intensive margins of the labor supply are present. He shows that the optimal tax schedule heavily depends on the comparison between the elasticities of participation decisions with respect to tax levels and of earnings with respect to marginal tax rates. Our model emphasizes that the monotonicity of the elasticities of participation is also important.

Some papers have made a distinction between unemployment and non-participation. Boadway *et alii* (2003) study redistribution when unemployment is endogenous and generated by matching frictions or efficiency wages. The government's information set is different from ours because they assume that it observes productivities and can distinguish among the various types of non-employed. Boone and Bovenberg (2004) depart from the standard model of nonlinear income taxation *à la* Mirrlees (1971) by adding a job-search margin that is the single determinant of the unemployment risk. As in our model, the government can not verify job search. However, in their model, the cost of participation is homogeneous in the population and the unemployment risk does not depend on wages nor on taxation. In Boone and Bovenberg (2006), the framework is similar but since the government observes employed workers' skill, taxation is skill-specific. Their focus is on the respective roles of the assistance benefit and of in-work benefits in redistributing income while ours is on redistributive taxation when the government observes only wages.

Closely related to the current paper, Hungerbühler *et alii* (2006), henceforth HLPV, have proposed an optimal income tax model with unobservable worker's ability and where unemployment is endogenous due to matching frictions. The present paper differs from HLPV in four important respects. First, the cost of participation takes a unique value in HLPV. Hence, every agent above (below) an endogenous threshold of skill participates (does not participate). In the present paper, we allow the opportunity cost of participation to vary within and between skill levels. This leads to a more general and to us much more realistic treatment of participation decisions. In this sense, HLPV is a particular case of the present paper where the elasticity of participation is infinite at the threshold, and zero above. Second, following Saez (2001), the present paper expresses our optimality conditions in terms of behavioral elasticities and interprets them in the light of marginal tax reforms. This emphasizes the economics behind the optimal tax formulas. Moreover, in the Maximin case, it reveals the critical role played by the shape of the elasticities of participation along the skill distribution. Third, the social welfare function in HLPV does not take into account the issue of income redistribution between employed and non-employed individuals of the same skill level. On the contrary, our paper deals with this issue, a point that appears important when workers are involuntarily unemployed. Finally, in both papers wage-setting implies that the *laissez-faire* allocation is efficient but our paper is compatible with a wider class of matching functions. As a matter of fact, by focusing on Nash bargaining over wages under the so-called Hosios (1990), HLPV restrict to Cobb-Douglas matching functions.

The paper is organized as follows. The model and fiscal incidence are presented in the next section. Section III characterizes the Maximin optimum. Section IV presents the optimality conditions under the general utilitarian criterion. Section V explains how we calibrate the model and presents numerical simulations of optimal tax schedules. Finally, Section VI concludes.

II The model

As usual in the optimal non linear tax literature that follows Mirrlees (1971), we consider a static framework where the government is averse to inequality. For simplicity we assume risk-neutral agents. In our model, the sources of differences in earnings are threefold. First, individuals are endowed with different levels of productivity (or skill) denoted by a. The distribution of skills admits a continuous density function $f(.)$ on a support $[a_0, a_1]$, with $0 \leq a_0 < a_1 \leq +\infty$. The size of the population is normalized to 1. Second, whatever their skill, some people choose to stay out of the labor force while some others do participate to the labor market. To account for this fact, we assume that individuals of a given skill differ in their individual-specific gain χ of remaining out of the labor force. We call χ the value of non-market activities. Third, among those who participate to the labor market, some fail to be recruited and become unemployed. This "involuntary" unemployment is due to matching frictions *à la* Mortensen and Pissarides (1999) and Pissarides (2000). Labor markets are perfectly segmented by skill. This assumption is made for tractability and seems more realistic than the polar one of a unique labor market for all skill levels. The timing of events is the following:

1. The government commits to an untaxed assistance benefit b and a tax function $T(.)$ that only depends on the (gross) wage w.[3]

2. For each skill level a, firms decide how many vacancies to create. Creating a vacancy of type a costs $\kappa(a)$. Individuals of type (a, χ) decide whether they participate to the labor market of type a.

3. On each labor market, the matching process determines the number of filled jobs. Since an individual of type (a, χ) who chooses to participate renounces χ, all participants of skill a are alike. We henceforth call these individuals participants of type a for short. Each participant supplies an exogenous amount of labor normalized to 1. So, earnings and (gross) wages are equal.

4. Each worker of skill a produces a units of goods, receives a wage $w = w_a$ and pays taxes. Taxes finance the assistance benefit b and an exogenous amount of public expenditures $E \geq 0$. Agents consume.

We assume that the government does neither observe individuals' types (a, χ) nor the job-search and matching processes.[4] It only observes workers' gross wages w_a and is unable to distinguish among the *non-employed* individuals those who have searched for a job but failed to find one (the *unemployed*) from the *non participants*.[5] Moreover, as our model is static, the government is unable to infer the type of a jobless individual from her past earnings. Therefore, the government is constrained to give the same level of assistance benefit b to all non-employed individuals, whatever their type (a, χ) or their participation decisions.[6] An individual of type (a, χ) can decide to remain out of the labor force, in which case her utility equals $b+\chi$. Otherwise, she finds a job with an endogenous probability ℓ_a and gets a net-of-tax wage $w_a - T(w_a)$ or she becomes unemployed with probability $1 - \ell_a$ and gets the assistance benefit b.[7]

[3] If the income tax and the assistance schemes were administered by different authorities, new issues would arise that we do not consider here.

[4] The government is therefore unable to infer the skill of workers from the screening of job applicants made by firms. So, the tax schedule cannot be skill-specific. Moreover, we do not consider the possibility that redistribution could also be based on observable characteristics related to skills (see Akerlof, 1978).

[5] However, the government is able to compute the skill-specific unemployment and participation rates. It also knows the density $f(.)$ and the boundaries of the support of a.

[6] Similarly, in Boone and Bovenberg (2004, 2006), the welfare benefit does not depend on the ability of the jobless individual.

[7] Our model can easily be extended to include a skill-specific fixed cost of working.

II.1 Participation decisions

To participate, an individual of type (a, χ) should expect an income, $\ell_a (w_a - T(w_a)) + (1 - \ell_a) b$, higher than in case of non participation, $b + \chi$. Let

$$\Sigma_a \stackrel{\text{def}}{\equiv} \ell_a (w_a - T(w_a) - b)$$

denote the expected surplus of a participant of type a. Let $G(a, .)$ be the cumulative distribution of the value of non-market activities, conditional on the skill level, that is

$$G(a, \Sigma) \stackrel{\text{def}}{\equiv} \Pr[\chi \leq \Sigma | a]$$

Then, the participation rate among individuals of skill a equals $G(a, \Sigma_a)$ and hence the number of participants of type a equals $U_a = G(a, \Sigma_a) f(a)$. We denote the continuous conditional density of the value of non-market activities by $g(a, \Sigma)$. The support of $g(a, .)$ is an interval whose lower bound is 0. Note that the characteristics a and χ can be independent or not. We define

$$\pi_a \stackrel{\text{def}}{\equiv} \frac{\Sigma_a \cdot g(a, \Sigma_a)}{G(a, \Sigma_a)} \tag{1}$$

the elasticity of the participation rate with respect to Σ, at $\Sigma = \Sigma_a$. This elasticity is in general both endogenous and skill-dependent. Note that π_a also equals the elasticity of the participation rate of agents of skill a with respect to $w_a - T(w_a) - b$ when ℓ_a is fixed. The empirical literature typically estimates the latter elasticity.

II.2 Labor demand

On the labor market of skill a, creating a vacancy costs $\kappa(a) > 0$. This cost includes the investment in equipment and the screening of applicants. Only a fraction of vacancies finds a suitable worker to recruit. Following the matching literature (Mortensen and Pissarides 1999, Pissarides 2000 and Rogerson *et alii* 2005), we assume that the number of filled positions is a function $H(a, V_a, U_a)$ of the numbers V_a of vacancies and U_a of job-seekers. The matching function $H(a, ., .)$ on the labor market of skill a is increasing in both arguments and exhibits constant returns to scale.[8] Moreover, $H(a, V_a, 0) = H(a, 0, U_a) = 0$, and for all V_a and U_a, one has $H(a, V_a, U_a) < \min(V_a, U_a)$. Finally, $H(., ., .)$ is twice-continuously differentiable on $[a_0, a_1] \times \mathbb{R}_+^2$.

Define tightness θ_a as the ratio V_a/U_a. The probability that a vacancy is filled equals $q(a, \theta_a) \stackrel{\text{def}}{\equiv} H(a, 1, 1/\theta_a) = H(a, V_a, U_a)/V_a$. Due to search-matching externalities, the job-filling probability decreases with the number of vacancies and increases with the number of job-seekers. Because of constant returns to scale, only tightness matters and $q(a, \theta_a)$ is a decreasing

[8] See Petrongolo and Pissarides (2001) for microfoundations and empirical evidence about the matching function.

function of θ_a. Symmetrically, the probability that a job-seeker finds a job is an increasing function of tightness $p(a,\theta_a) \stackrel{\text{def}}{\equiv} H(a,\theta_a,1) = H(a,V_a,U_a)/U_a$. Firms and individuals being atomistic, they take tightness θ_a as given.

When a firm creates a vacancy of type a, she fills it with probability $q(a,\theta_a)$. Then, her profit at stage 4 equals $a - w_a$. Therefore, her expected profit at stage 2 equals $q(a,\theta_a)(a-w_a) - \kappa(a)$. Firms create vacancies until the free-entry condition $q(\theta_a)(a-w_a) = \kappa(a)$ is met. This pins down the value of tightness θ_a and in turn the probability of finding a job through[9]

$$L(a,w_a) \stackrel{\text{def}}{\equiv} p\left(a, q^{-1}\left(a, \frac{\kappa(a)}{a-w_a}\right)\right) \tag{2}$$

At the equilibrium, one has $\ell_a = L(a,w_a)$ and

$$\Sigma_a = L(a,w_a)(w_a - T(w_a) - b) \tag{3}$$

The $L(.,.)$ function is a reduced form that captures everything we need on the labor demand side. From the assumptions made on the matching function, $L(.,.)$ is twice-continuously differentiable and admits values within $(0,1)$. As the wage increases, firms get lower profit on each filled vacancy, fewer vacancies are created and tightness decreases. This explains why $\partial L/\partial w_a < 0$. Moreover, due to the constant-returns-to-scale assumption, the probability of being employed depends only on skill and wage levels and not on the number of participants. If for a given wage, there are twice more participants, the free-entry condition leads to twice more vacancies, so the level of employment is twice higher and the employment probability is unaffected. This property is in accordance with the empirical evidence that the size of the labor force has no lasting effect on group-specific unemployment rates. Finally, because labor markets are perfectly segmented by skill, the probability that a participant of type a finds a job depends only on the wage level w_a and not on wages in other segments of the labor market.

II.3 The wage setting

As the literature dealing with optimal redistribution in a competitive framework (Mirrlees 1971 and followers), we focus on the redistribution issue and abstract from the standard inefficiency arising from matching frictions. In other words, we consider a setting such that the role of taxation is only to redistribute income (as in Mirrlees) and not to restore efficiency.[10] For this purpose, we consider a wage-setting mechanism that maximizes the sum of utility levels in the absence of taxes and benefits. To obtain this property, the matching literature typically assumes that wages are the outcome of a Nash bargaining and that the workers' bargaining power equals the elasticity of the matching function with respect to unemployment (see Hosios

[9] Where $q^{-1}(a,.)$ denotes the inverse function of $\theta \mapsto q(a,\theta)$, holding a constant.

[10] Boone and Bovenberg (2002) studies how nonlinear taxation can restore efficiency in a matching model where the Hosios condition is not fulfilled.

1990). This assumption is only meaningful if the elasticity of the matching function is constant and exogenous. When the matching function is of the Cobb-Douglas form $H(a, U_a, V_a) = A (U_a)^{\gamma} (V_a)^{1-\gamma}$, Equation (2) implies that $L(a, w) = A^{1/\gamma} ((a-w)/\kappa(a))^{((1-\gamma)/\gamma)}$. Then, Nash bargaining under the Hosios condition leads to a wage level that solves (see HLPV):[11]

$$w_a = \arg\max_{w} \quad L(a, w) \cdot (w - T(w) - b) \tag{4}$$

When the matching function is not of the Cobb-Douglas form, we assume that (4) still holds. So, $\Sigma_a = \max_{w} \quad L(a, w) \cdot (w - T(w) - b)$ and the equilibrium wage maximizes the participation rate given the tax/benefit system.

Different wage-setting mechanisms can provide microfoundations for (4). The Competitive Search Equilibrium introduced by Moen (1997) and Shimer (1996) leads to this property. Another possibility is to assume that a skill-specific utilitarian monopoly union selects the wage w_a after individuals' participation decisions but before firms' decisions about vacancy creation (see Mortensen and Pissarides, 1999).

II.4 The *laissez faire*

The *laissez faire* is defined as the economy without tax and benefit. According to (4), the equilibrium level of wage maximizes $L(a, w) \cdot w$. A wage increase has a direct positive effect on $L(a, w) \cdot w$ and a negative effect through the employment probability. To ensure that program (4) is well-behaved at the *laissez faire*, we assume that for any (a, w),

$$\frac{\partial^2 \log L}{\partial w \cdot \partial \log w}(a, w) < 0 \tag{5}$$

We henceforth denote w_a^{LF} the wage at the *laissez faire*. To guarantee that w_a^{LF} increases with the level of skill, we further assume that for any (a, w):

$$\frac{\partial^2 \log L}{\partial a \partial w}(a, w) > 0 \tag{6}$$

Appendix A verifies that, when the exogenous amount of public expenditures E is nil, the *laissez-faire* economy maximizes the Benthamite objective, which equals the sum of utility levels. Because of our wage-setting mechanism (4), wages at the *laissez faire* maximize "efficiency" (i.e. maximize the Benthamite criterion). Note that participation decisions are then also efficient.

II.5 Fiscal incidence

We now reintroduce the tax/benefit system and explain how tax reforms affect the equilibrium. Starting with the wage, notice that the objective in (4) multiplies the employment probability

[11] If different wage levels solve (4), then we make the tie-breaking assumption that the wage level preferred by the government will be selected. See also the discussion in Mirrlees (1971, footnotes 2 and 3 pages 177).

by the difference between the net incomes in employment and in unemployment. We call this difference the *ex-post surplus* and denote it $x \overset{\text{def}}{\equiv} w - T(w) - b$. It subtracts an "employment tax", $T(w) + b$, from the earnings w. In our setting, the influence of the tax and benefit system comes through the profile of the relationship between the ex-post surplus x and earnings w. Because of the multiplicative form of (4), what actually matters is how $\log x$ varies with $\log w$. When $T(.)$ is differentiable, the first-order condition[12] of Program (4) writes:

$$-\frac{\partial \log L}{\partial \log w}(a, w_a) = \eta(w_a) \tag{7}$$

where[13]

$$\eta(w) \overset{\text{def}}{\equiv} \frac{1 - T'(w)}{1 - \frac{T(w)+b}{w}} = \frac{\partial \log(w - T(w) - b)}{\partial \log w} \tag{8}$$

When the wage increases by one percent, the term $\partial \log L / \partial \log w\,(a, w)$ measures the relative decrease in the employment probability, while $\eta(w)$ measures the relative increase in the ex-post surplus. At equilibrium, Equation (7) requires that these two relative changes sum to zero. Notice that in our setting the profile of $\eta(w)$ gathers all the information about the profile of the tax/benefit system needed to fix the equilibrium wage. Figure 1 displays indifference expected surplus curves. The equation of these indifference curves can be written as $\log x =$ constant $- \log L(a, w)$. From (2) and (5), these curves are increasing and convex. The solution to Program (4) then consists in choosing the highest indifference curve taking the relationship between $\log x$ and $\log w$ into account. In case of differentiability, this amounts to choosing the highest indifference curve tangent to the $\log w \mapsto \log x = \log(w - T(w) - b)$ schedule. The first-order condition (7) combined with (8) expresses this tangency condition.

For comparative static purposes, consider for a while the average tax rate $T(w_a)/w_a$, the assistance benefit ratio b/w_a and the marginal tax rate $T'(w_a)$ as parameters. So, $\eta(w_a)$ is provisionally a parameter, too. Under Condition (5), Equations (7) and (8) imply that the equilibrium wage w_a (thereby the unemployment rate $1 - L(a, w_a)$) increases with the average tax rate and the assistance benefit ratio and decreases with the marginal tax rate. These properties are standard in the equilibrium unemployment literature. They hold under monopoly unions (Hersoug 1984), right-to-manage bargaining (Lockwood and Manning 1993), efficiency wages with continuous effort (Pisauro 1991) or matching models with Nash bargaining (Pissarides 1998). Sørensen (1997) and Røed and Strøm (2002) provide some empirical evidence in favor

[12] The solution to (4), if any, necessarily lies in $(-\infty, a - \kappa(a))$. Since $L(a, a - \kappa(a)) = 0$, $w = a - \kappa(a)$ does not solve (4). From a theoretical viewpoint, the wage can be negative whenever $T(.)$ is negative enough to keep some agents of type a participating to the labor market (i.e. $w - T(w) > b$). Hence the solution to (4) is necessarily interior. In the rest of the paper, we focus on positive wage levels.
Since $\partial \log L / \partial w < 0$, $\eta(w)$ has to be positive. As the expected surplus is positive, so is $w - T(w) - b$. Hence, the marginal tax rate $T'(w)$ has to be lower than 1.

[13] $\eta(w)$ is reminiscent of the Coefficient of Residual Income Progression which measures the wage elasticity of net earnings (Musgrave and Musgrave 1976). $\eta(w)$ is actually the Coefficient of Residual Income Progression divided by one minus the net replacement ratio $b/(w - T(w))$.

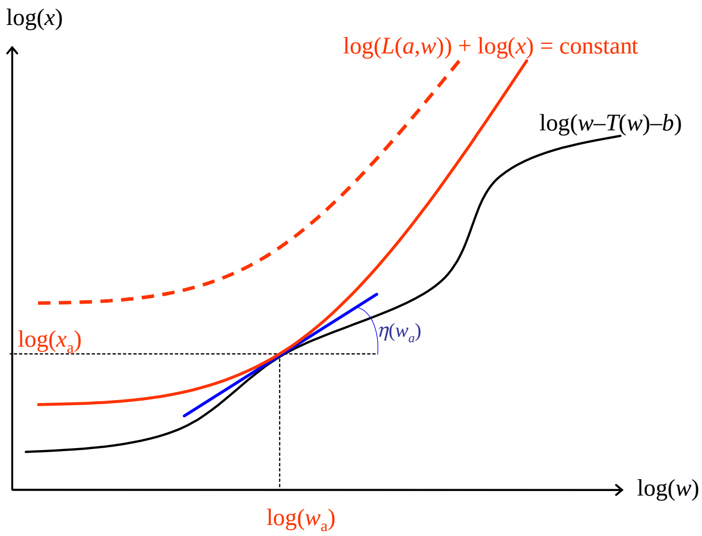

Figure 1: The choice of the wage for a type a match.

of the wage-moderating effect of higher marginal tax rates. In addition, Manning (1993) finds that higher marginal tax rates lower unemployment in the UK.

From Equation (7), the average tax rate, the assistance benefit ratio and the marginal tax rate affect the equilibrium wage only through changes in the slope η of the $\log w \mapsto \log x$ function. To see why, imagine a tax reform such that participants of type a face a steeper $\log w \mapsto \log x$ tax function. A relative rise in the wage induces now a higher relative gain in the ex-post surplus x. Still, the relative loss in the employment probability is unchanged. Consequently, the rise in η induces an increase in the equilibrium wage w_a that substitutes ex-post surplus for employment probability. This is reminiscent of the *substitution effect* in a competitive framework with adjustments along the intensive margin. There, a lower marginal tax rate raises the net hourly wage and leads to a substitution toward consumption and away from leisure time. Returning to our setting, Equation (7) indicates that for a given slope of the $\log w \mapsto \log x$ function, the level of this function does not affect the equilibrium wage. In this specific sense, there is no *income effect* of the tax schedule on wages in our model.

In the general case where η is a function of the wage, a change in this slope produces a direct change in wage levels. This in turn creates a second change in η which produces a further change in the wage. To clarify this circular process[14] and to prepare the analysis developed in Sections III and IV in terms of a small tax reform, imagine that the slope $\eta(w)$ of the $\log w \mapsto \log x$ relationship exogenously increases by a small amount $\tilde{\eta}$. Let us rewrite the first-order condition

[14]Which is also present in the optimal non-linear taxation literature with competitive labor markets and labor supply decisions (see Saez 2001).

(7) as $\mathcal{W}(w_a, a, 0) = 0$, where:

$$\mathcal{W}(w, a, \tilde{\eta}) \overset{\text{def}}{\equiv} \frac{\partial \log L}{\partial \log w}(a, w) + \eta(w) + \tilde{\eta} \tag{9}$$

The second-order condition of (4) writes $\mathcal{W}'_w(w_a, a, 0) \leq 0$ where

$$\mathcal{W}'_w(w_a, a, \tilde{\eta}) = \frac{\partial^2 \log L(a, w_a)}{\partial w \cdot \partial \log w} + \eta'(w_a)$$

This second-order condition states that at the equilibrium wage w_a, the $\log w \mapsto \log x$ relationship depicted in Figure 1 has to be either concave or less convex than the indifference expected surplus curves.[15]

Consider now how the equilibrium wage w_a is influenced by small changes in the parameter $\tilde{\eta}$ and in the type a. Whenever the second-order condition of (4) is a strict inequality, we can apply the implicit function theorem on $\mathcal{W}(w_a, a, \tilde{\eta}) = 0$. We then obtain the elasticity ε_a of the equilibrium wage w_a with respect to a small local change in $\tilde{\eta}$ around a given $\log w \mapsto \log x$ function. We also obtain the elasticity α_a of the equilibrium wage w_a with respect to the skill level a along the same $\log w \mapsto \log x$ function:

$$\varepsilon_a \overset{\text{def}}{\equiv} \frac{\eta(w_a)}{w_a} \cdot \frac{\partial w_a}{\partial \tilde{\eta}} = -\frac{\eta(w_a)}{w_a \cdot \mathcal{W}'_w(w_a, a, 0)} > 0 \tag{10a}$$

$$\alpha_a \overset{\text{def}}{\equiv} \frac{a}{w_a}\frac{\partial w_a}{\partial a} = -\frac{a}{w_a \cdot \mathcal{W}'_w(w_a, a, 0)} \cdot \frac{\partial^2 \log L}{\partial a \partial w}(a, w_a) > 0 \tag{10b}$$

These elasticities are in general endogenous and in particular they depend on the curvature term $\eta'(w_a)$ in $\mathcal{W}'_w$. This is because a change in wage Δw_a, that is either caused by a change in $\tilde{\eta}$ or in a, induces a change in $\eta(w_a)$ that equals $\eta'(w_a)\Delta w_a$ and a further change in the wage. This is at the origin of a circular process captured by the term $\eta'(w_a)$ in $\mathcal{W}'_w$. However, as will be clear in Sections III and IV, only the ratio ε_a/α_a enters the optimality conditions and this ratio does not depend on $\eta'(w_a)$ but only on a and w_a. The positive signs of ε_a and α_a follow from the strict second-order condition $\mathcal{W}'_w < 0$ and from (6).

In addition to its effect on wage and unemployment through $\eta(.)$, taxation also influences participation decisions. To isolate this effect, consider a tax reform that rises $\log(w - T(w) - b)$ by a constant amount for all w so that $\eta(w)$ is kept unchanged. Such a tax reform does neither change the wage level, nor the employment probability. However, the employment tax $T(w_a) + b$ is reduced and hence the surplus Σ_a an agent of type a can expect from participation increases. Therefore, such a reform increases the participation rate $G(a, \Sigma_a)$, thereby the employment rate $L(a, w_a) \cdot G(a, \Sigma_a)$. The magnitude of this behavioral response is captured by the elasticity π_a defined in (1). In sum, the income effect affects the participation margin and not the wage-cum-labor demand margin.

[15] When this condition is not verified over an interval, the earnings function $a \mapsto w_a$ is discontinuous.

II.6 The equilibrium

For a given function $\log w \mapsto \log x$, the equilibrium allocation can be found recursively. The wage-setting equations (4) determine wages w_a and in turn $x_a = w_a - T(w_a) - b$. The labor demand functions (2) determine the skill-specific unemployment rates $1 - L(a, w_a)$. Then, from (3), the participation rates are given by $G(a, \Sigma_a)$ and the employment rates equal $L(a, w_a)\, G(a, \Sigma_a)$. For each additional worker of type a, the government collects taxes $T(w_a)$ and saves the assistance benefit b. Since $E \geq 0$ is the exogenous amount of public expenditures, the government's budget constraint defines the level of b:

$$b = \int_{a_0}^{a_1} (T(w_a) + b) \cdot L(a, w_a) \cdot G(a, \Sigma_a) \cdot f(a)\, da - E \tag{11}$$

III The Maximin case

Under the Maximin (Rawlsian) objective, the government only values the utility of the least well-off. Unemployed individuals get b, which is always lower than the workers' and non participants' utility levels, respectively $w - T(w)$ and $b + \chi$. Therefore, a Maximin government aims at maximizing b subject to the budget constraint (11) and incentive compatibility constraints. The latter state that, for each skill level, the selected wage w_a maximizes the expected surplus $L(a, w)(w - T(w) - b)$. According to the *taxation principle* (Hammond 1979, Rochet 1985 and Guesnerie 1995), the set of allocations induced by a tax/benefit system $\{T(.), b\}$ through the wage-setting equations (4) corresponds to the set of incentive-compatible allocations $\left(b, \{w_a, x_a, \Sigma_a\}_{a \in [a_0, a_1]}\right)$ that verify:

$$\forall \left(a, a'\right) \in [a_0, a_1]^2 \qquad \Sigma_a \equiv L(a, w_a)\ x_a \geq L(a, w_{a'})\ x_{a'} \tag{12}$$

From (6), the strict single-crossing condition holds. Hence (12) is equivalent to the differential equation $\dot{\Sigma}_a = \Sigma_a \cdot \partial \log L / \partial a\, (a, w_a)$ and the monotonicity requirement that the wage w_a is a nondecreasing function of the skill level a (see Appendix B).

Following Mirrlees (1971), it is much more convenient to solve the government's problem in terms of allocations.[16] In Appendix C, we follow this approach to derive our optimal tax formula. Let $h_a = L(a, w_a)\ G(a, \Sigma_a)\ f(a)$ denote the (endogenous) mass of workers of skill a. We obtain:

[16] We assume the existence of an optimal allocation $a \mapsto (w_a, x_a)$ that is continuous, differentiable and increasing. Existence and continuity are usual regularity assumptions (see e.g. Mirrlees 1971, 1976 or Guesnerie and Laffont 1984). The monotonicity assumption means that we rule out bunching. We verify in the simulations that the monotonicity requirement is verified along the optimum. The differentiability assumption is made only for convenience. It implies that the tax schedule $T(.)$ is almost everywhere differentiable in the wage.

Proposition 1 *For any skill level $a \in [a_0, a_1]$, the maximin-optimal tax schedule verifies:*

$$\frac{1-\eta(w_a)}{\eta(w_a)} \cdot \frac{\varepsilon_a}{\alpha_a} \cdot w_a \cdot a \cdot h_a = Z_a \qquad \textit{and} \qquad Z_{a_0} = 0 \tag{13a}$$

$$Z_a = \int_a^{a_1} [x_t - \pi_t (T(w_t) + b)] h_t \cdot dt, \tag{13b}$$

where $T(w_t) + b = w_t - x_t$ and since $\eta(w) = \partial \log(w - T(w) - b) / \partial \log w$, x_t verifies:

$$\forall t, u \qquad \log x_t = \log x_u + \int_{w_u}^{w_t} \eta(w)\, d \log w$$

In Proposition 1, the elasticities π_a of the participation rate, ε_a of the wage with respect to η and α_a of the wage with respect to the skill level a are respectively given by (1), (10a) and (10b) along the optimal allocation. Moreover, w_a is determined by the wage-setting condition (7).

III.1 Intuitive proof of Proposition 1

The resolution in terms of incentive-compatible allocations enables a rigorous derivation. However, this method does not provide much economic intuition. So, we propose here an intuitive proof in the spirit of Saez (2001). Recall that in our model, it is much more convenient to think of the tax schedule as a function that associates the log of the ex-post surplus to the log of the wage. We consider the effect of the following small tax reform around the optimum depicted in Figure 2. The slope $\eta(w)$ of $\log w \mapsto \log x$ is marginally increased by $\tilde{\eta} = \Delta\eta$ for wages in the small interval $[w_a - \delta w, w_a]$.[17] We take $\Delta\eta$ sufficiently small compared to δw, so that bunching or gaps in the wage distribution around $w_a - \delta w$ or w_a induced by the tax reform can be neglected. This reform has two effects on the government's objective (11). There is first a *tax level* effect that concerns individuals of skill t above a. Those of them who are employed thus receive a wage w_t above w_a. Second, there is *a wage response* effect. It takes place for those whose wages lie in the $[w_a - \delta w, w_a]$ interval.

The tax level effect

Consider skill levels t above a. Since $\eta(.)$ is unchanged around w_t, the equilibrium wage w_t is unaffected by the tax reform, and so is the employment probability $L(t, w_t)$. From (8), the tax reform increases the ex-post surplus $x_t = w_t - T(w_t) - b$ by

$$\frac{\Delta x_t}{x_t} = \Delta\eta \cdot \frac{\delta w}{w}$$

[17] The reasoning below will be entirely developed in terms of this local change in η. For the reader interested by the implementation of such a reform, the small local increase $\Delta\eta$ would be the result of a small decline in the marginal tax rate, the level of the average employment tax being kept locally constant. Above w_a, the induced reduction in the employment tax should be compensated for by an appropriate reduction of the marginal tax rate to keep η unchanged.

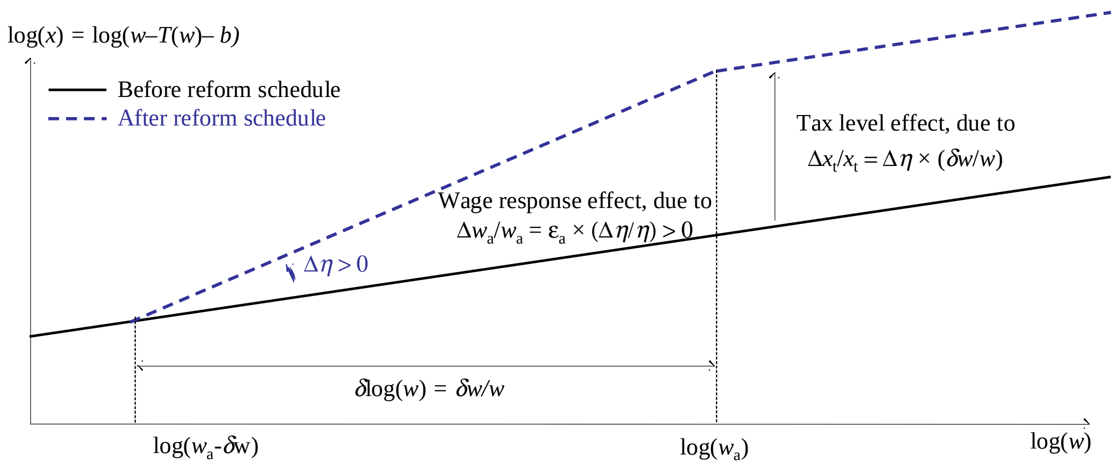

Figure 2: The tax reform

Due to:	**Wage response effect** $\Delta w_a / w_a = \varepsilon_a \ (\Delta\eta/\eta)$	**Tax Level effect** $\Delta x_t / x_t = \Delta\eta \ (\delta w / w)$
Mechanical Component	$T(w_a)$ increases by $T'(w_a)\ \Delta w_a$	$T(w_t)$ decreases by $\Delta T(w_t) = -\Delta x_t$
Behavioral component	Labor demand is reduced by $\Delta \ell_a / \ell_a = -\eta(w_a)(\Delta w_a / w_a)$	Participation rates increase by: $\Delta G_t / G_t = \pi_t \ (\Delta x_t / x_t)$

Table 1: Summary of the different components of the effects of the tax reform

(see Figure 2). The consequence of this rise of (the log of) the ex-post surplus can be decomposed into a *mechanical component* and a *behavioral component* through a change in the participation decisions (see Table 1).

The rise in x_t corresponds to a reduction in the employment tax level $T(w_t) + b$ such that $\Delta(T(w_t) + b) = -x_t \cdot \Delta\eta \cdot (\delta w / w)$. Since there are h_t workers of type t, the mechanical component of the tax level effect at skill level t equals:

$$-x_t \cdot h_t \cdot \Delta\eta \cdot \frac{\delta w}{w} \tag{14}$$

Consider now the participation decisions of individuals of skill t above a. From (3), since their employment probability is unchanged, their expected surplus increases by the same relative amount $\Delta\Sigma_t / \Sigma_t = \Delta\eta \cdot (\delta w / w)$ as their ex-post surplus x_t. According to (1) the number of employed individuals of type t thus increases by $\pi_t \cdot h_t \cdot \Delta\eta \cdot (\delta w / w)$. For each of these additional employed individuals, the government receives $T(w_t) + b$ additional employment taxes. Hence, the behavioral component of the tax level effect at skill level t equals:

$$\pi_t \cdot (T(w_t) + b) \cdot h_t \cdot \Delta\eta \cdot \frac{\delta w}{w} \tag{15}$$

From (13b), the sum of the mechanical and behavioral components over all skill levels t above a gives *the tax level effect*. It equals $-Z_a \cdot \Delta\eta \cdot (\delta w / w)$.

The wage response effect

This effect concerns individuals whose skill level is such that their wage in case of employment lies in the interval $[w_a - \delta w, w_a]$. Let $[a - \delta a, a]$ be the corresponding interval of the skill distribution. From (10b), one has

$$\delta a = \frac{a}{\alpha_a} \cdot \frac{\delta w}{w} \tag{16}$$

Therefore, the number of agents concerned by this effect is $(a/\alpha_a)\, f(a)\, (\delta w/w)$.

Due to the small tax reform, those employed face a more increasing $\log w \mapsto \log x$ tax schedule. The tax reform thus induces a wage increase Δw_a that substitutes ex-post surplus for employment probability. From (10a), one has

$$\frac{\Delta w_a}{w_a} = \frac{\varepsilon_a}{\eta\left(w_a\right)} \cdot \Delta\eta \tag{17}$$

Since the equilibrium wage maximizes participants' ex-post surplus Σ_a, the tax reform has only a second-order effect on Σ_a and thereby on the participation rate of these individuals. The wage response effect can be decomposed into a *mechanical component* and a *behavioral component* through a change in the labor demand decisions (see Table 1).

The wage increase Δw_a changes the employment tax paid by $T'\left(w_a\right) \cdot \Delta w_a$. From (8), one gets $1 - T'\left(w_a\right) = x_a \cdot \eta(w_a)/w_a$, so

$$\Delta\left(T\left(w_a\right) + b\right) = T'\left(w_a\right) \cdot \Delta w_a = \left[\left(1 - \eta\left(w_a\right)\right) w_a + \eta\left(w_a\right)\left(T\left(w_a\right) + b\right)\right] \frac{\Delta w_a}{w_a} \tag{18}$$

Multiplying the last term by the number of employed individuals h_a gives the mechanical component of the wage response effect.

The wage increase Δw_a also induces a reduction in the employment probability $L\left(a, w_a\right)$. Given (7), the fraction of employed among participants is decreased by:

$$\Delta L\left(a, w_a\right) = -\eta\left(w_a\right) \frac{\Delta w_a}{w_a} L\left(a, w_a\right) \tag{19}$$

When an additional participant of type a finds a job, the government levies additional taxes $T\left(w_a\right)$ and saves b. Multiplying the employment tax $T\left(w_a\right) + b$ by $\Delta \ell_a$ times the number of participants $G\left(a, \Sigma_a\right) f\left(a\right) \delta a$ gives the behavioral component of the wage response effect. The sum of these two components equals

$$\Delta\left[\left(T\left(w_a\right) + b\right) \cdot L\left(a, w_a\right)\right] \cdot G\left(a, \Sigma_a\right) \cdot f\left(a\right) \cdot \delta a = \left(1 - \eta\left(w_a\right)\right) w_a \cdot h_a \cdot \frac{\Delta w_a}{w_a} \cdot \delta a$$

Given (16), (17) and the last expression, the total *wage response* effect on the interval $[w_a - \delta w, w_a]$ equals

$$\frac{1 - \eta\left(w_a\right)}{\eta\left(w_a\right)} \cdot \frac{\varepsilon_a}{\alpha_a} \cdot a \cdot w_a \cdot h_a \cdot \Delta\eta \cdot \frac{\delta w}{w} \tag{20}$$

The wage response effect can be either positive or negative. From Subsection II.4, recall that the *laissez-faire* value of the wage is efficient. If $\eta\left(w_a\right) < 1$, (resp. $\eta\left(w_a\right) > 1$) the wage is below (above) its *laissez-faire* value, hence it is inefficiently low (high). Adding the wage response and the tax level effects gives (13a) in Proposition 1.

To obtain $Z_{a_0} = 0$ in (13a), consider a tax reform that rises $\log(w - T(w) - b)$ by a constant amount for all w, so that $\eta(w)$ is kept unchanged. This reform is implemented by increasing the level of Σ_{a_0} and thus the level of Σ_a for all a (the rise in Σ_a has to be a proportional rise since $\dot{\Sigma}_a = \Sigma_a \cdot \partial \log L / \partial a\,(a, w_a)$). Such a tax reform induces an effect that is proportional to Z_{a_0} and no wage response effect. At the optimum, such a marginal reform has to have no-first-order impact on the government's objective. This implies that the sum of all mechanical and behavioral effects has to be nil i.e. that $Z_{a_0} = 0$.

III.2 Instructive cases

To better understand the implications of our optimal tax formula, we now consider its implications when additional restrictions are imposed. Given the literature, a natural starting point is the case where wages are exogenously fixed ($\varepsilon_a = 0$). Then, we return to the case where wages are endogenous but impose some constraints on the elasticities of participation.

No wage response effect

We provisionally assume that marginal tax reforms do not change the employment probabilities ℓ_a. However, wages still increase exogenously with the skill (i.e. α_a remains positive). This case corresponds to the model with only extensive margin responses of labor supply considered by Diamond (1980), Saez (2002) and Choné and Laroque (2005).[18] Intuitively, as wages do not react to changes in taxes, the solution is given by putting to zero the sum of the mechanical (14) and behavioral (15) components of the tax level effect. This has to be true for all levels of skill. Consequently, $x_a - \pi_a (T(w_a) + b) = 0$ whatever the skill a.[19] Therefore, at the optimum, the employment tax (respectively, the employment surplus) verify:

$$\frac{T(w_a) + b}{w_a} = \frac{1}{1 + \pi_a} \qquad \Leftrightarrow \qquad \frac{x_a}{w_a} = \frac{\pi_a}{1 + \pi_a} \tag{21}$$

These relationships are implicit ones when π_a depends on the expected surplus. The optimal employment tax rate only depends on the behavioral response (through π_a) and not on the distribution of skills. In Figure 1, the optimal allocation $\log x_a$ is necessarily below the 45 degree line at a distance given by $|\ \log(\pi_a/(1+\pi_a))\ |$. In accordance with Saez (2002) in the Maximin case, the employment tax $T(w_a) + b$ is positive i.e. there is no EITC. Where the participation rate is more elastic, the behavioral component matters more. Therefore, the

[18] However here, as in Boone and Bovenberg (2004, 2006), participants face a positive but exogenous probability to be "involuntarily" unemployed.

[19] Formally, from (13a) as $\varepsilon_a = 0$ for all a, $Z_a = 0$ everywhere. So, from (13b), $x_a - \pi_a (T(w_a) + b) = 0$ everywhere, too.

optimal ex-post surplus has to be higher to induce participation (a necessary condition to collect taxes to finance b).

Constant elasticity of participation

We now investigate under which condition the tax schedule described by Equation (21) is optimal when wages are responsive to taxation ($\varepsilon_a > 0$). This tax schedule induces that the aggregate tax level effect Z_a equals 0 everywhere along the skill distribution (See Equation 13b). Therefore, the wage response effect has to be nil everywhere. So, according to (13a), the slope η of the $\log w \mapsto \log x$ function has to equal 1 everywhere. Therefore, from (8), the ratio x_a/w_a has to be constant. This is consistent with (21) only when the elasticity of participation π_a is the same for all skill levels at the optimum.

Reciprocally, assume that the elasticity of participation is constant and consider the tax policy defined by an employment tax $T(w) + b$ that equals $w/(1+\pi)$ for all wage levels w. In this case, the mechanical (14) and behavioral (15) components of the tax level effect sum to 0 at each skill level. Moreover, from (8), this policy induces $\eta(w)$ to be constant and equal to 1, so wages are not distorted and the wage response effect is nil everywhere. Therefore, this policy satisfies the conditions in Proposition 1.

Decreasing elasticity of participation

The assumption of a constant elasticity of participation is convenient but not plausible. Empirical studies suggest that participation decisions are more elastic at the bottom of the skill distribution (see the empirical evidence surveyed by Immervoll *et alii*, 2007, and Meghir and Phillips, 2008). This elasticity is in general a function of the expected surplus (see (1)), hence it is endogenous. Therefore, the profile of π_a at the optimum may be different from the corresponding profile in the current economy. It seems nevertheless reasonable to assume that the elasticity of participation is decreasing in skill levels along the optimum.[20] In this case, we get:

Proposition 2 *If everywhere along the Maximin optimum one has $\dot{\pi}_a < 0$, then*

i) $w_a < w_a^{LF}$ *and* $L(a, w_a) > L\left(a, w_a^{LF}\right)$ *for all* a *in* (a_0, a_1)*, while* $w_{a_0} = w_{a_0}^{LF}$, $L(a_0, w_{a_0}) = L\left(a_0, w_{a_0}^{LF}\right)$, $w_{a_1} = w_{a_1}^{LF}$ *and* $L(a_1, w_{a_1}) = L\left(a, w_{a_1}^{LF}\right)$.

ii) *Compared to the laissez faire, the participation rates are distorted downwards.*

iii) *The average tax rate $T(w)/w$ is an increasing function of the wage and the marginal tax rates $T'(w)$ are positive everywhere. The in-work benefit (if any) at the bottom-end of the distribution is lower than the assistance benefit $-T(w_{a_0}) < b$.*

[20] The polar assumption where π_a is increasing in a leads to symetric analytical results. We do not present them here since this case very implausible.

This Proposition is proved in Appendix D. Its intuition is illustrated in Figure 3. This Figure depicts the ratio of the ex-post surplus over the wage, x_a/w_a, as a function of the level of skill. In the absence of wage responses, as we have seen above, the optimum implements a policy such that x_a/w_a is equal to $\pi_a/(1+\pi_a)$ and hence the tax level effect is nil. The dashed decreasing curve $\pi_a/(1+\pi_a)$ in Figure 3 illustrates this profile in the current context where $\dot{\pi}_a < 0$. However, when wages are responsive to taxation (i.e. when $\varepsilon_a > 0$), implementing this policy means that $x_a = w_a - T(w_a) - b$ increases less than proportionally in the wage w_a, so $\eta(w_a) < 1$. Hence, wages are distorted downwards. The optimum trades off the distortions along the wage response effect and along the tax level effect. Since an optimization along the wage response effect corresponds to a flat curve, the optimal policy corresponds to the one illustrated by the solid curve in Figure 3. Thus, the solid curve remains decreasing, which induces that wages and unemployment are distorted downwards for all interior skill levels (point *i)* of the Proposition).

Whether participation rates are distorted upwards or downwards compared to the *laissez faire,* depend on whether the expected surplus is higher along the optimum Σ_a or along the *laissez faire* $w_a^{\mathrm{LF}} L\left(a, w_a^{\mathrm{LF}}\right)$. Let us write Σ_a as $w_a L(a, w_a) \cdot (x_a/w_a)$. First, since wages are distorted, $w_a L(a, w_a)$ is lower at the optimum compared to the *laissez faire*. Second, as illustrated by Figure 3 the ratio x_a/w_a reaches its highest value along the optimum for the lowest skill level. Moreover, x_{a_0}/w_{a_0} is lower at the optimum with wage response effect compared to the optimum without wage response effect. These two features hold because the optimum with wage response effect trades off distortions along the tax level effect and along the wage response effect. Finally, along the optimum without wage response effect, x_{a_0}/w_{a_0} is lower than 1 because the government has a Maximin objective (see 21). Hence, $x_a/w_a < 1$, which finally gives point *ii*) of the Proposition.

Moreover, as $x_a/w_a < 1$, one has $T(w) + b > 0$ for all wage levels. So, transfers for (low income) workers are never higher than for the jobless: There is no EITC in the words of Saez (2002, p. 1055). Furthermore, since x/w is decreasing, $(T(w) + b)/w$ is increasing in wages, hence average tax rates are increasing in wages, too. Finally, since $(T(w) + b)/w$ is positive everywhere and marginal tax rates are higher than this ratio (because $\eta < 1$), marginal tax rates are positive everywhere, including at the boundaries of the skill distribution (Point *iii)* of the Proposition).

Point *i*) of the Proposition 2 is in contrast to the literature initiated by Mirrlees (1971). There, optimal marginal tax rates are positive whenever the government values redistribution (see e.g. the discussion in Choné and Laroque 2007). Therefore, labor supply, thereby the volume of labor used, are distorted downwards, while in our case the volume of labor among participants is distorted upwards. However, Point *ii)* reduces this contrast. In our model,

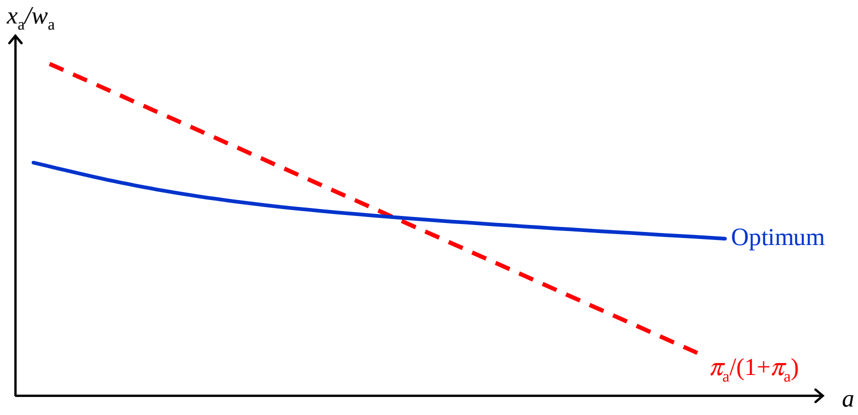

Figure 3: Intuition of Proposition 2

participation is distorted downwards. Consequently, the net effect on aggregate employment is ambiguous. Proposition 2 generalizes HLPV. There, the value χ of non market activities is identical for all types. Therefore, a unique threshold level of skill separates nonparticipants from participants. The elasticity of participation is thus infinite at the threshold and then nil, which is a very specific decreasing $a \mapsto \pi_a$ relationship. Finally, the property according to which employment tax rates are always positive is also obtained in the models of Saez (2002) and Choné and Laroque (2005) where participation margins are central. Saez (2002) however emphasizes that this result only holds under a Maximin criterion. With a more general objective, he finds that the optimal income tax schedule is typically characterized by a negative employment tax at the bottom provided that labor supply responses along the extensive margin are high enough compared to responses along the intensive margin.

IV The general utilitarian case

In this section, we derive the optimal tax formula when the government has the following Bergson-Samuelson social welfare function:

$$\begin{aligned}\Omega \;=\; & \int_{a_0}^{a_1} \{L(a,w_a)\,G(a,\Sigma_a)\,\Phi(w_a - T(w_a)) + (1 - L(a,w_a))\,G(a,\Sigma_a)\,\Phi(b) \\ & + \int_{\Sigma_a}^{+\infty} \Phi(b+\chi)\,g(a,\chi)\,d\chi\} f(a)\,da\end{aligned} \tag{22}$$

where $\Phi'(.) > 0 > \Phi''(.)$. The "pure" (Benthamite) utilitarian case sums the utility levels of all individuals and corresponds to the case where $\Phi(.)$ is linear. The stronger the concavity of $\Phi(.)$, the more averse to inequality is the government. Under this objective, Appendix E shows that the optimum verifies (recall that $\ell_a = L(a,w_a)$ and $h_a = \ell_a\; G(a,\Sigma_a)\; f(a)$):

Proposition 3 *For any skill level $a \in [a_0, a_1]$, the optimal tax schedule verifies:*

$$\left(\frac{1-\eta(w_a)}{\eta(w_a)} \cdot w_a - \frac{\Phi(w_a - T(w_a)) - \Phi(b) - x_a \cdot \Phi'(w_a - T(w_a))}{\lambda}\right) \cdot \frac{\varepsilon_a}{\alpha_a} \cdot a \cdot h_a = Z_a \quad (23a)$$

$$Z_{a_0} = 0 \quad (23b)$$

$$\textit{where} \qquad Z_a = \int_a^{a_1} \left\{ \left(1 - \frac{\Phi'(w_t - T(w_t))}{\lambda}\right) x_t - \pi_t \left[T(w_t) + b + \Xi_t\right] \right\} h_t \cdot dt \quad (23c)$$

$$\textit{and} \qquad \Xi_t = \frac{\ell_t \cdot \Phi(w_t - T(w_t)) + (1-\ell_t)\Phi(b) - \Phi(b + \Sigma_t)}{\lambda \cdot \ell_t}, \quad (23d)$$

in which the positive Lagrange multiplier associated to the budget constraint (11), λ, verifies

$$\lambda = \int_{a_0}^{a_1} \left\{ \ell_a G(.)\, \Phi'(w_a - T(w_a)) + (1-\ell_a) G(.)\, \Phi'(b) + \int_{\Sigma_a}^{+\infty} \Phi'(b+\chi)\, g(a,\chi)\, d\chi \right\} f(a)\, da \quad (24)$$

We now explain how to extend the intuitive proof of Section III. Equation (24) defines the marginal social value of public funds, λ. It is obtained by a unit increase in E financed by a unit decrease in b holding $w \mapsto w - T(w) - b$ constant. Next, we consider again the small tax reform depicted in Figure 2. This tax reform has a tax level effect and a wage response effect, each of them being decomposed into mechanical and behavioral components (see Table 1). In the Maximin case, these components only capture the impact on the least well-off (i.e. on additional tax receipts to finance the assistance benefit b). Now, the government also values how the utility levels of all other economic agents are affected by the tax reform. To make the formula comparable, we divide these additional impacts by λ, so as to express them in terms of the value of public funds. For each component, we now examine how the various components are changed.

Tax level effect

The rise in the ex-post surplus x_t increases the social welfare of the corresponding workers by $\Phi'(w_t - T(w_t))/\lambda$. Adding this welfare gain to the loss in tax receipts, the mechanical component of the tax level effect at skill level t equals

$$-\left(1 - \frac{\Phi'(w_t - T(w_t))}{\lambda}\right) \cdot x_t \cdot h_t \cdot \Delta\eta \cdot \frac{\delta w}{w} \quad (25)$$

instead of (14). The integral of relation (25) over the skill distribution above a corresponds to the "*between-skill*" motive of redistribution. Since λ averages marginal social welfare over the whole population and Φ is concave, the term in parentheses is positive for most workers. This means that the rise in x_t is in general detrimental to the government's objective. This might however not be true for workers with sufficiently low earnings. In this case, the government would increase the ex-post surplus with respect to the laissez faire for these workers. In opposition to the case where the government has a Maximin objective, this would imply a rise in the participation rate of the less skilled workers.

As far as the behavioral component is concerned, consider individuals of type t who are induced to participate by the tax reform. Their expected utility levels only change by a second-order amount. However, this change in participation decisions increases inequalities because participants' income is different whether they get a job or not. The inequality-averse government values this by $(\ell_t \cdot \Phi(w_t - T(w_t)) + (1-\ell_t)\Phi(b) - \Phi(b+\Sigma_t))/\lambda$, which equals $\ell_t \cdot \Xi_t$ (by Definition (23d)) and is negative. So, the behavioral component of the tax level effect at skill level t equals

$$\pi_t \{T(w_t) + b + \Xi_t\} \cdot h_t \cdot \Delta\eta \cdot \frac{\delta w}{w} \tag{26}$$

instead of (15). From (23c), the sum of the mechanical and behavioral components over all skill levels t above a equals $-\Delta\eta \cdot (\delta w/w) \cdot Z_a$. It is hard to draw clear conclusions about the value of Z_a. Still, two opposite effects are specific to the general utilitarian case. Compared to the Maximin, raising the ex-post surplus for skills above a is now less detrimental for the social welfare in terms of the mechanical component but the welfare gain of additional participants is less important because of the negative induced impact of increased inequalities on social welfare (the negative Ξ_t term).

Wage response effect

In addition to its impact on b through the tax receipts (described in (18) and (19)), the wage response effect has also a direct influence on social welfare through a change in the expected social welfare of participants of type a, $\ell_a\Phi(w_a - T(w_a)) + (1-\ell_a)\Phi(b)$. Holding b constant, a mechanical and a behavioral component should again be distinguished.

The wage increase Δw_a rises $\Phi(w_a - T(w_a))$ by the marginal social welfare $\Phi'(w_a - T(w_a))$ times the small increase in the post-tax wage $(1 - T'(w_a))\Delta w_a$. Using (8), the additional mechanical component expressed in terms of the value of public funds equals:

$$\frac{\Phi'(w_a - T(w_a))}{\lambda} \cdot x_a \cdot \eta(w_a) \cdot h_a \cdot \frac{\Delta w_a}{w_a} \cdot \delta a$$

This component has a positive effect on social welfare. However, the rise in the wage lowers the employment probability ℓ_a by $\Delta\ell_a = -\eta(w_a) \cdot (\Delta w_a/w_a) \cdot \ell_a$. Each additional unemployed individual decreases social welfare by $\Phi(w_a - T(w_a)) - \Phi(b)$. Hence, using (7), the additional behavioral component equals

$$-\frac{\Phi(w_a - T(w_a)) - \Phi(b)}{\lambda} \cdot \eta(w_a) \cdot h_a \cdot \frac{\Delta w_a}{w_a} \cdot \delta a$$

Adding these two components, then using (16) and (17), we get the welfare consequence of the wage response effect

$$-\frac{\Phi\left(w_a-T\left(w_a\right)\right)-\Phi\left(b\right)-x_a\cdot\Phi'\left(w_a-T\left(w_a\right)\right)}{\lambda}\cdot\frac{\varepsilon_a}{\alpha_a}\cdot a\cdot h_a\cdot\delta a \tag{27}$$

The welfare consequence of the wage response effect is negative because it increases inequalities among participants. This is first due to the fact that for a given number of unemployed, the ex-post surplus of each employee increases. Secondly, for a given employee surplus, the number of unemployed increases. The wage response effect implies a "*within-skill*" motive of redistribution that attenuates the will of the government to mitigate the between-skill inequalities. Thus, this effect pushes optimal wages downwards to reduce inequalities among participants and to lower unemployment.

By adding (27) to the impact (20) of the wage response effect on the level of the assistance benefit b, one obtains the left-hand side of (23a) times $\Delta\eta\cdot(\delta w/w)$.

This intuitive proof of Proposition 3 has highlighted that (search) unemployment has two effects on social welfare that cannot be recognized if the wage-cum-labor demand margin is ignored. First, unemployment *per se* is a source of loss in social welfare which calls for downward wage distortions. This is captured by the negative sign of (27). Second, because the fate of participants is not employment for sure, policies that enhance participation have a detrimental induced effect on inequality. To see the implication of this second effect, consider the particular case where wages are not responsive to taxation ($\varepsilon_a=0$ everywhere). Then, the tax level effect has to be nil everywhere at the optimum. From (23c), whatever the skill t, the employment tax should verify:

$$\frac{T(w_t)+b}{w_t-T(w_t)-b}=\frac{1}{\pi_t}\left(1-\frac{\Phi'\left(w_t-T\left(w_t\right)\right)}{\lambda}\right)-\frac{\Xi_t}{w_t-T(w_t)-b} \tag{28}$$

If Ξ_t was zero, Formula (28) would be identical to Expression (4) in Saez (2002). Then, if the welfare of low skill workers is highly valued by the government, i.e. if their ability and thus wage is sufficiently low (i.e. such that $\Phi'\left(w_t-T(w_t)\right)/\lambda>1$), the employment tax $T(w_t)+b$ should be negative, meaning that transfers for low income workers, $-T(w_t)$, are higher than for the jobless. Now because of unemployment, inequalities between the agents induced to participate by this policy are increased (since Ξ_t is negative). This reduces the willingness of the government to redistribute to low income workers.

When wages are responsive to taxation, the only analytical result in the general utilitarian case concerns wage distortions at both extremes of the skill distribution. There, as in the Maximin case, the tax level effect is nil. Nevertheless, there is a reason to choose an inefficient wage level. This is because unemployment reduces social welfare. To mitigate this effect, it is worth distorting wages downwards at both extremes of the skill distribution.

Concerning the robustness of Proposition 2 obtained under a Maximin objective, we cannot say whether nor when the two new terms in (25) and (26) change the sign of the tax level effect. We can nevertheless make the following conjectures in line with this proposition. As far as point *i*) is concerned, the government has now an additional incentive to reduce wages and stimulate labor demand since the welfare impact of the wage response effect (27) is negative. However, pushing wages downwards obviously reduces social welfare, and the more so as one moves towards the low-end of the wage distribution. Therefore, compensating transfers for low-skilled workers are expected. Numerical simulations are needed to throw some light on these conjectures.

V Simulations

To illustrate how our optimal tax formulae could be used for applied purposes, this Section proposes a calibration of our model based on the US economy. This enables us to compute optimal income tax schedules that provide some numerical feel of the policy implications of our analysis. As the underlying model remains stylized in several dimensions the following simulation results should only be considered as illustrative.

V.1 Calibration

To avoid the complexity of interrelated participation decisions within families, we only consider single adults in the US.[21] We need to specify the labor demand function $L(.,.)$ and the distribution of types (a,χ) through functions $G(.,.)$ and $f(.)$. In choosing functional specifications of $L(.,.)$ and $G(.,.)$, we want to control the behavioral responses ε_a, α_a and π_a defined respectively by Equations (10a), (10b) and (1). We take

$$\log L(a,w) = B(a) - \varepsilon \left(\frac{w}{c \cdot a}\right)^{\frac{1}{\varepsilon}}$$

Under this specification, the first-order condition (7) for the wage-setting program implies:

$$w_a = c \cdot a \cdot (\eta(w_a))^{\varepsilon} \tag{29}$$

Next, we roughly approximate the tax system that is applied to single adults without children by a linear function $T(w) = \tau \cdot w + \tau_0$ with $\tau = 25\%$ and $\tau_0 = -3000$. The selection of a value of b for the current economy determines whether $\eta(w)$ is lower or larger than 1, and, consequently, whether wages (and thus unemployment) are distorted upwards or downwards. As a benchmark and to be consistent with our theoretical analysis where taxes are used only to redistribute income, we assume that wages are efficient in the current economy, so we take $b = -\tau_0 = 3000$.

[21] These are "primary individuals", i.e. persons without children living alone or in households with adults who are not their relatives. They are older than 16 and younger than 66.

Since η is then constant, the elasticity α_a of the wage with respect to the skill equals 1 in the current economy (see 10b), as it would be the case in a perfectly competitive economy. Moreover ε equals the elasticity of the wage with respect to η in the current economy (10a). This elasticity also equals the *compensated* elasticity of wage with respect to $1-T'$.[22] Following Gruber and Saez (2002), estimates of the latter elasticity would lie between 0.2 and 0.4. We take a conservative value $\varepsilon = 0.1$ in the benchmark calibration and conduct a sensitivity analysis where $\varepsilon = 0.2$. We set c to 2/3, so that in the current economy, total wage income represents two third of the total production.[23] Finally, we use (29) and the distribution of weakly earnings of the Current Population Survey of May 2007 to approximate a distribution of skills among employed workers. Reexpressing variables in annual terms, the range of skills is $[\$3,900; \$218,400]$.[24] Using a quadratic Kernel with a bandwidth of \$63,800 we get an approximation of $L(a)\,G(a,\Sigma_a)\,f(a)$ in the current economy which is depicted by the lowest curve in Figure 4.

We then assume that the elasticity of participation varies exogenously with the level of skill. More specifically, we assume the following cumulative distribution of non-market activities $\Pr[\chi \leq \Sigma \mid a]$:[25]

$$G(a,\Sigma) = A(a) \cdot \Sigma^{\pi_a} \qquad \text{where } A(a) > 0 \text{ and } \pi_a > 0 \tag{30}$$

Because, to our knowledge, the empirical literature does not provide any information about the concavity of the function $a \mapsto \pi_a$, we assume the following simple declining profile $\pi_a = (\pi_{a_0} - \pi_{a_1})\left(\frac{a_1-a}{a_1-a_0}\right)^3 + \pi_{a_1}$. We set the elasticity at the bottom, π_{a_0}, to 0.4 and the elasticity at the top, π_{a_1}, to 0.2 in the benchmark calibration and conduct sensitivity analysis. These elasticities are in line with the evidence summarized by Immervoll *et alii* 2007 and Meghir and Phillips (2008).

We adjust scales parameters $B(a)$ and $A(a)$ to get realistic profiles of skill-specific unemployment rates and participation rates. The profile of unemployment (resp. participation) rates is approximated by a decreasing (increasing) function of a:

$$1 - \ell_a = 0.035 + \left(\frac{a_1 - a}{a_1 - a_0}\right)^4 0.045 \qquad \text{and} \qquad G(a,\Sigma_a) = 0.31\left(1 - \left(\frac{a_1 - a}{a_1 - a_0}\right)^6\right) + 0.58$$

In our approximation of the current economy, the mean unemployment rate is 5.06%, the mean participation rate equals 80.3% and the mean elasticity of the participation rate equals 0.29. Figure 4 depicts the calibrated skill distribution $f(a)$, the distribution of skill among participants in the current economy $G(a,\Sigma_a)\,f(a)$ and the distribution of skills among employed individuals

[22] For any *compensated* change in marginal tax rates, one has $\Delta\eta = \frac{\Delta(1-T')}{1-T'} \frac{1-T'}{1-(T+b)/w} = \frac{\Delta(1-T')}{1-T'} \cdot \eta$.

[23] In the equilibrium matching approach, workers receive less than their marginal product because firms have to recoup their initial investment ($\kappa(a)$ in the theory developed above).

[24] The data are collected for wage and salary workers. We ignore weekly earnings below 50\$, which corresponds to the lowest 1.2% of the earnings distribution.

[25] When we adopt this specification, we implicitly assume that $A(a)$ is such that one always has $\Sigma_a \leq [A(a)]^{-1/\pi_a}$. Otherwise, the participation rate equals one and becomes inelastic.

$L(a, w_a)\, G(a, \Sigma_a)\, f(a)$. We compute the level of exogenous public expenditures E from the

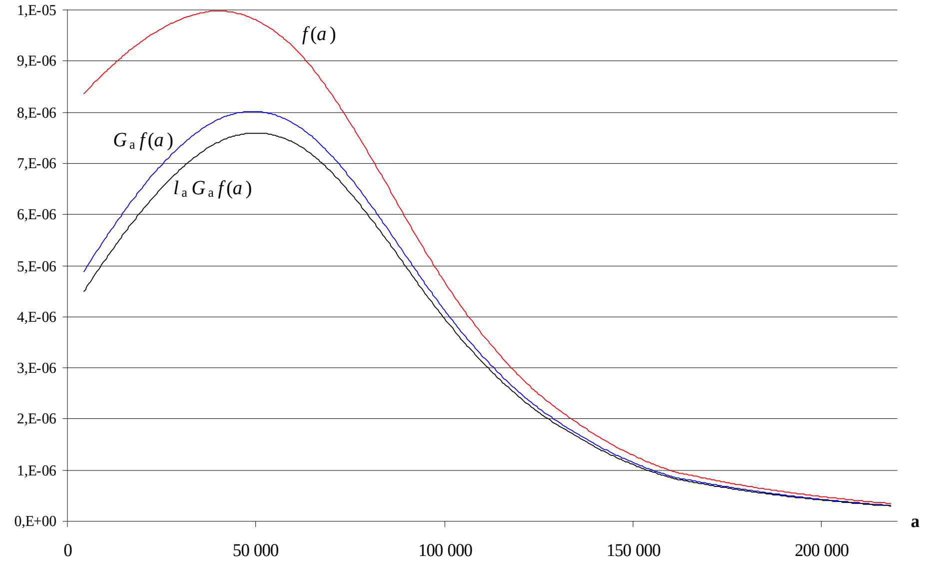

Figure 4: Densities $f(a)$, $G(a, \Sigma_a)\, f(a)$ and $L(a, w_a)\, G(a, \Sigma_a)\, f(a)$ in the current economy.

government's budget constraint (11). This leads to an amount $E = \$5,636$ per capita. In the Bergson-Samuelson utilitarian case, we take $\Phi(y) = (y + E)^{1-\sigma}/(1-\sigma)$, with $\sigma = 0.2$ in the benchmark. The exogenous public expenditures finances a public good that generates social utility that is considered as a perfect substitute to private consumption under this specification.

V.2 Results

To illustrate Part i) of Proposition 1, let us compare the actual profile of unemployment rates and the optimal ones under the Maximin and Bergson-Samuelson criteria (Figure 5). The actual unemployment rate turns out to be too high from a Maximin perspective (except at the extremes of the skill distribution). From the general utilitarian viewpoint, it should even decrease further, confirming the importance of the welfare impact of the wage response effect (27). As an illustration of Part ii) of Proposition 1, Figure 6 shows that a Maximin government would accept a sharp decline in participation rates. Under the more general utilitarian objective, optimal participation rates are higher for low skilled workers and lower for high skilled workers. Since unemployment rates are lower and participation rates are higher at the bottom of the skill distribution, the tax-schedule is designed to boost low-skill employment.

Marginal tax rates are drawn in Figure 7. Under the Maximin, redistribution takes the form of a Negative Income Tax (NIT) in the following sense: An assistance benefit close to $\$14,198$ is taxed away at a high, and in this case nearly constant, marginal tax rate close to 80%. With the more general utilitarian criterion, the well-being of workers, in particular the

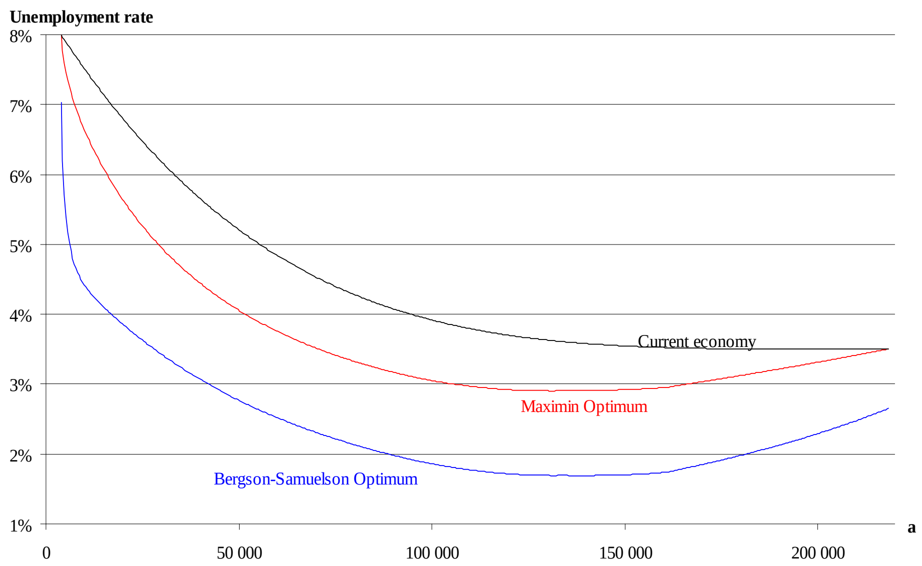

Figure 5: Unemployment under the benchmark calibration

low-paid ones, enters the scene. This changes dramatically the form and the level of marginal tax rates. At the bottom of the skill distribution, the marginal tax rate is negative and then sharply increases to about 40%. The tax schedule has now the basic features of an EITC-type taxation. In particular, the level of b equals $\$1,015$ per year, while there is an in-work benefit at the bottom whose level is substantially higher since $T(w_{a_0}) = -\$3,167$. In order to reduce the unemployment of the less skilled, the government strongly distorts their wages downwards.

In Figure 7, we have also depicted the optimal tax schedule if the reaction of wages to taxation is ignored ($\varepsilon = 0$). Compared to our benchmark where $\varepsilon = 0.1$, the optimal profile is notably different. In particular, the marginal tax rates are lower at the low-end of the wage distribution since, by assumption, there is no adjustment in wages and hence in unemployment. So, this Figure illustrates that taking into account or ignoring the wage-cum-labor demand margin has substantial quantitative implications. Still, the assistance benefit and the tax reimbursement at the bottom are close to those just mentioned (so that the property $T(w_{a_0}) + b < 0$ still holds).

If the sensitivity of wages to taxation is raised from $\varepsilon = 0.1$ towards $\varepsilon = 0.2$, the wage response effects are reinforced. The Maximin optimum therefore implements a tax schedule where the function $w \mapsto x(w)/w$ vary less (i.e. the solid curve of Figure 3 becomes flatter) so as to prevent too important distortions along the wage-cum-labor demand margin. The tax schedule becomes closer to a linear one, marginal tax rates vary less. The simulations displayed in Figure 8 show that this also happens along the Bergson-Samuelson optimum.

The other sensitivity analyses we conduct concern the calibration of the elasticity of participation π_a. First, we decrease by a constant amount of 0.05 all the shape of $a \mapsto \pi_a$. In the

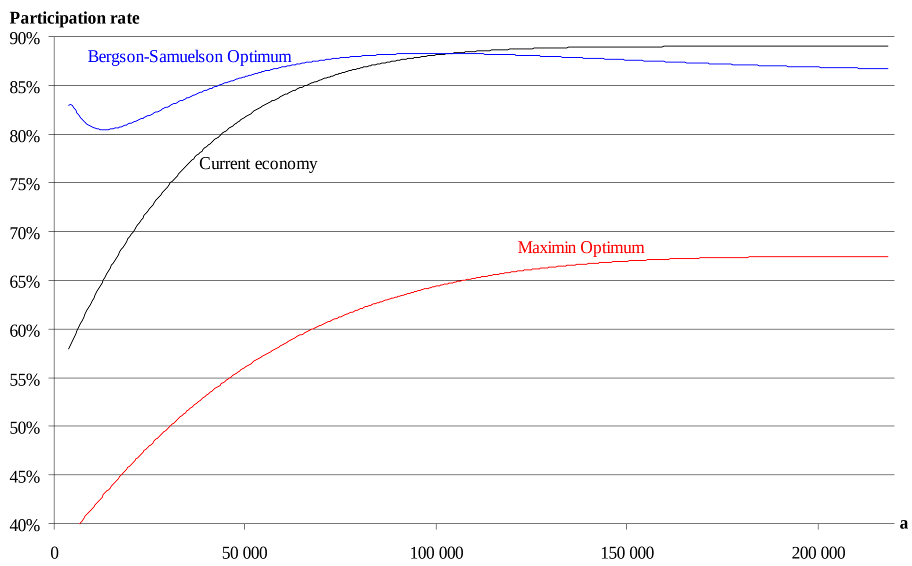

Figure 6: Participation rates under the benchmark calibration.

Maximin case without wage response, Equation (21) implies that the government would choose higher tax levels as participation responds less, so the dashed curve in Figure 3 is shifted downwards. Consequently, in the presence of wage response, the solid curve shifts downwards too. Hence the Maximin optimum implements higher levels of $(T(w) + b)/w$ and therefore higher marginal tax rates. Figure 9 quantifies this mechanism. Once again, The Bergson-Samuelson optimum is affected in a similar way compared to the Maximin optimum.

Last, we change the elasticities of participation so that the relationship $a \mapsto \pi_a$ is steeper while keeping the average elasticity in the current economy almost constant. For that purpose, we take $(\pi_{a_0}, \pi_{a_1}) = (0.48; 0.13)$ instead of $(0.4; 0.2)$. To understand the rise in marginal tax rates displayed by Figure 10, it is again convenient to come back to Figure 3. In the Maximin optimum without wage response, the government whishes to implement a tax schedule with a more decreasing $a \mapsto x_a/w_a$ function, so the dashed curve of Figure 3 becomes stepper. Hence, in the presence of wage responses, the distortions along the wage cum labor demand are reinforced and the solid curve of Figure 3 becomes stepper too. As a consequence, $\eta(w_a)$ are decreased and marginal tax rates are raised (see 8).

In all the simulation exercises, unemployment rates are even lower at the Bergson-Samuelson optimum than at the Maximin one. This confirms the importance of the welfare impact of the wage response effect (27). Participation rates are always higher at the Bergson-Samuelson optimum compared to the Maximin one. They remain lower than the current ones for high skill workers and higher for lower skill workers. Average tax rates are always increasing at the Bergson-Samuelson optimum.

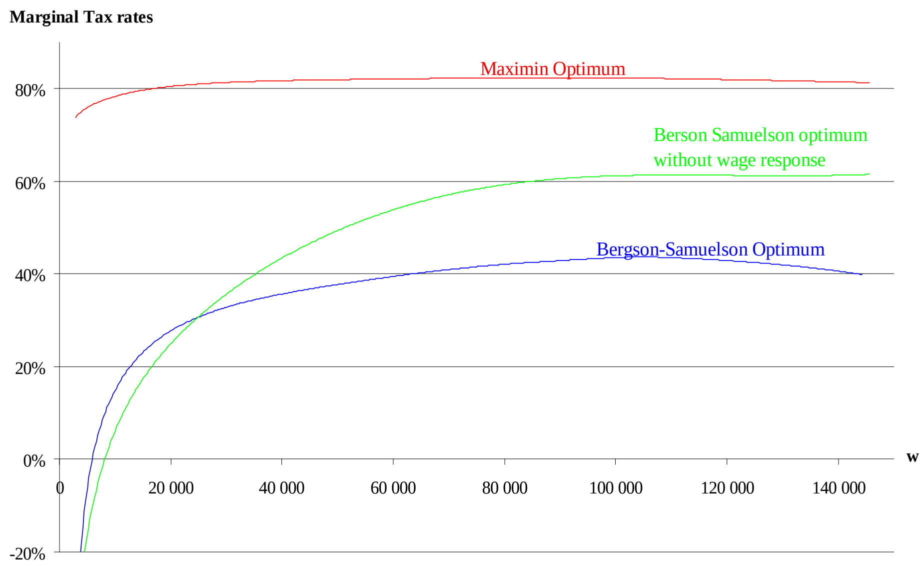

Figure 7: Marginal Tax Rates under the benchmark calibration

Saez (2001) has simulated optimal marginal tax rates using the empirical distribution of income to compute the underlying distribution of skills, as we do for our model. He has found that optimal marginal tax rates are U-shaped whereas we find a hump-shaped profile in all our simulations.

One analytical result in HLPV was that an EITC is never optimal. However, as we have already point in the introduction, participation decisions were treated in a crude way. This was an important limitation. In particular, Saez (2002) has proposed simulations of optimal tax rates at the bottom of the distribution with labor supply responses along both the extensive and the intensive margins. He has showed that an EITC can emerge if the government is not Maximin. In the present paper, we treat participation decisions more carefully than in HLPV. Our numerical simulations are then in line with Saez (2002) on this point, and not with HLPV.

VI Conclusions

According to authors such as Immervoll et al (2007), optimal income taxation can be studied in a competitive framework and the introduction of imperfections would not deeply modify the equity-efficiency trade-off. By modelling jointly participation decisions, wage formation and labor demand in a frictional economy, we show on the contrary that this trade off is deeply modified. In the Maximin case, a set of clear-cut analytical properties are shown if the elasticity of participation decreases with the level of skill. Then at the optimum, the average tax rate is increasing, marginal tax rates are positive everywhere, while wages, unemployment rates

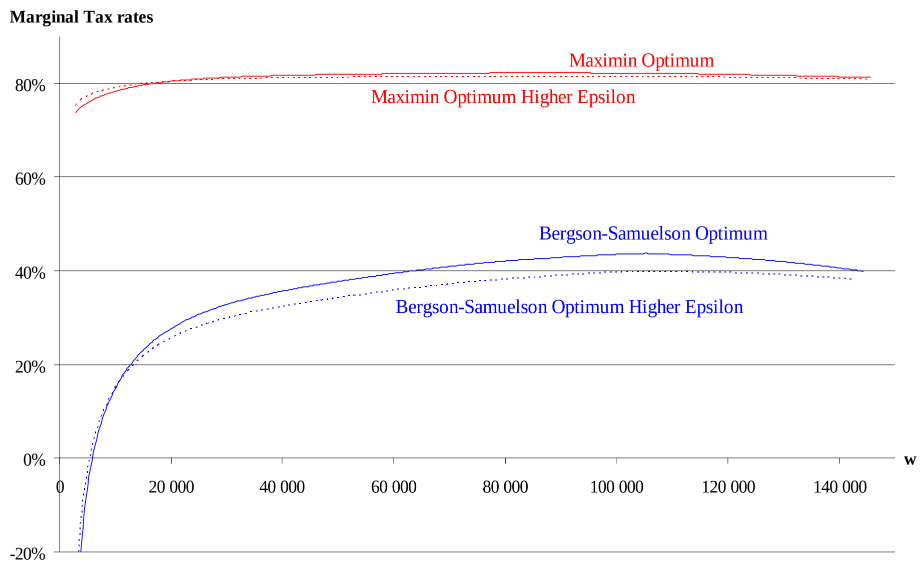

Figure 8: Dotted curves: ε equals 0.2 instead of 0.1 (solid curves).

and participation rates are distorted downwards compared to their laissez-faire values. These precise recommendations contrast with the small number of analytical properties derived in the literature following Mirrlees (1971).

When the government has a general utilitarian social welfare function, the equity-efficiency trade-off is more deeply affected by the wage-cum-labor demand margin. To induce participation, the net income of workers should be higher than the one of the non-employed. This creates an inequality that matters from a utilitarian perspective. Taxation should then promote wage moderation to reduce the detrimental effect of unemployment on social welfare. Moreover, the role of taxation on participation is more complex because some participants will not find a job. Therefore, stimulating participation through lower tax levels raises inequalities. Our numerical exercise shows that optimal unemployment rates are substantially distorted downwards and that an EITC can be optimal.

The present model could be extended in different directions. First, a dynamic model would enable to introduce earning-related unemployment insurance. Hence, one can expect that a "dynamic optimal taxation" version (*à la* Golosov *et alii* (2003)) of our model would deliver interesting insights about the optimal combination of unemployment insurance and taxation to redistribute income. Second, we abstract from any response of the labor supply along the intensive margin. Although we are confident that responses along the extensive margin are much more important, enriching our framework to include hours of work, in-work effort or educational effort belongs to our research agenda. Finally, in the real world, labor supply decisions are typically taken at the household level, not at the individual one. All these extensions are left

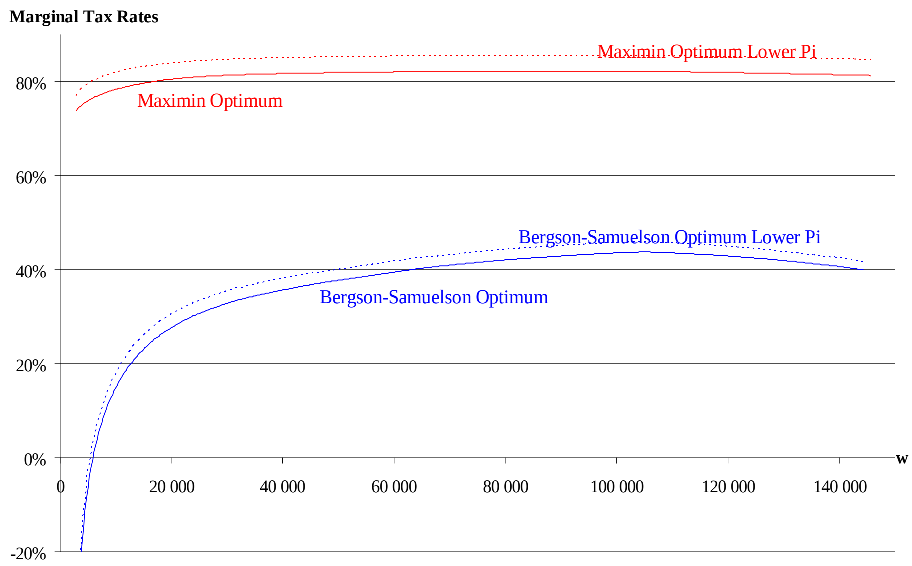

Figure 9: Dashed curves: (π_{a_0}, π_{a_1}) equals $(0.35; 0.15)$ instead of $(0.4; 0.2)$ (solid curves)

for future research.

Appendices

A Benthamite efficiency of the *laissez-faire* allocation

Let $\mathcal{U}$ be the Benthamite objective. Consider an equilibrium allocation. There are $G(a, \Sigma_a) f(a)$ participants of type a whose net income is $w_a - T(w_a)$ if they are employed and b otherwise, while non participants obtain $b + \chi$. So, the Benthamite objective writes:

$$\begin{aligned}\mathcal{U} &= \int_{a_0}^{a_1} \left\{ (L(a, w_a)(w_a - T(.)) + (1 - L(a, w_a))b) \cdot G(a, \Sigma_a) + \int_{\Sigma_a}^{+\infty} (b + \chi) \cdot g(a, \chi) \cdot d\chi \right\} f(a) \cdot da \\ &= \int_{a_0}^{a_1} \left\{ (\Sigma_a + b) \cdot G(a, \Sigma_a) + \int_{\Sigma_a}^{+\infty} (b + \chi) \cdot g(a, \chi) \cdot d\chi \right\} f(a) \cdot da\end{aligned}$$

where the second equality uses (3). Given the government's budget constraint (11), this objective can be rewritten when $E = 0$ as:

$$\mathcal{U} = \int_{a_0}^{a_1} \left\{ L(a, w_a) \cdot w_a \cdot G(a, \Sigma_a) + \int_{\Sigma_a}^{+\infty} \chi \cdot g(a, \chi) \cdot d\chi \right\} f(a) \cdot da$$

The Benthamite objective aggregates average earnings plus the value of non-market activities over the whole population, no matter how they are distributed. In this sense, the Benthamite criterion is an extreme case.

For each a and Y, the function $\Sigma \mapsto L \cdot w \cdot G(a, \Sigma) + \int_{\Sigma}^{+\infty} \delta \cdot g(a, \delta) \cdot d\delta$ reaches a unique maximum for $\Sigma = L \cdot w$. Therefore, when we compare any allocation $a \mapsto (w_a, \Sigma_a)$ to the

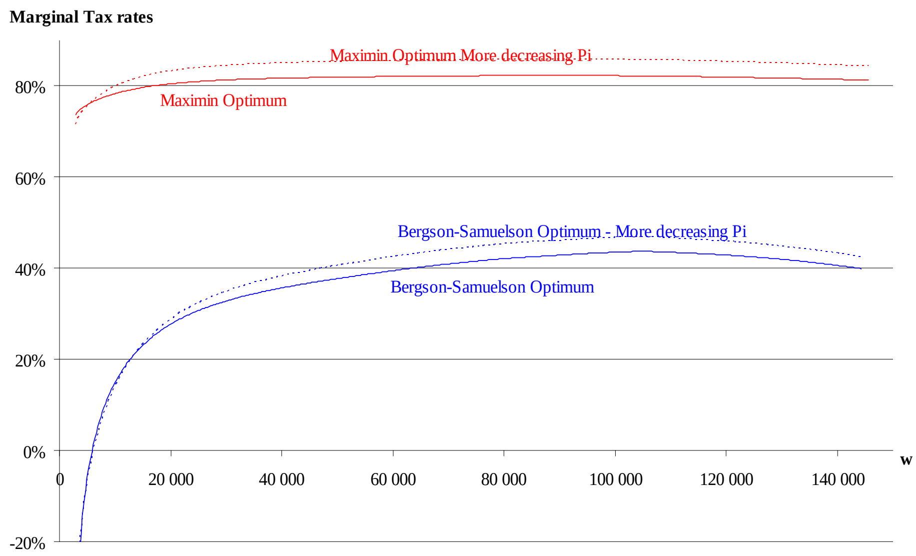

Figure 10: Dashed curves: (π_{a_0}, π_{a_1}) equals $(0.48; 0.13)$ instead of $(0.4; 0.2)$ (solid curves)

laissez-faire one, we get:

$$\begin{aligned}
\mathcal{U}^{\mathrm{LF}} &= \int_{a_0}^{a_1} \left\{ L\left(a, w_a^{\mathrm{LF}}\right) \cdot w_a^{\mathrm{LF}} \cdot G\left(a, \Sigma_a^{\mathrm{LF}}\right) + \int_{\Sigma_a^{\mathrm{LF}}}^{+\infty} \chi \cdot g\left(a, \chi\right) \cdot d\chi \right\} f\left(a\right) \cdot da \\
&\geq \int_{a_0}^{a_1} \left\{ L\left(a, w_a^{\mathrm{LF}}\right) \cdot w_a^{\mathrm{LF}} \cdot G\left(a, \Sigma_a\right) + \int_{\Sigma_a}^{+\infty} \chi \cdot g\left(a, \chi\right) \cdot d\chi \right\} f\left(a\right) \cdot da \\
&\geq \int_{a_0}^{a_1} \left\{ L\left(a, w_a\right) \cdot w_a \cdot G\left(a, \Sigma_a\right) + \int_{\Sigma_a}^{+\infty} \chi \cdot g\left(a, \chi\right) \cdot d\chi \right\} f\left(a\right) \cdot da = \mathcal{U}
\end{aligned}$$

The first inequality holds because $\Sigma_a^{\mathrm{LF}} = L\left(a, w_a^{\mathrm{LF}}\right) \cdot w_a^{\mathrm{LF}}$ at the *laissez faire*, according to (3). The second inequality holds because w_a^{LF} maximizes $w \mapsto L\left(a, w\right) \cdot w$

B Incentive Compatible allocations

Let $\mathcal{K}$ be the set of types (a, χ), $\mathcal{K}_P$ being the subset of participating types and $\mathcal{K}_0 = \mathcal{K} - \mathcal{K}_P$. An incentive-compatible allocation is given by a real number b and a mapping that associates to any element (a, χ) of $\mathcal{K}_P$ a bundle of wage $w_{a\chi}$ and ex-post surplus $x_{a\chi} = w_{a\chi} - T\left(w_{a\chi}\right) - b$, such that

$$\text{For any } \left(\left(a, \chi\right), \left(a', \chi'\right)\right) \in \left(\mathcal{K}_P\right)^2 : L\left(a, w_{a\chi}\right) \cdot x_{a\chi} \geq L\left(a, w_{a'\chi'}\right) \cdot x_{a'\chi'} \tag{31a}$$

$$\text{For any } \left(a, \chi\right) \in \mathcal{K}_P : b + L\left(a, w_{a\chi}\right) x_{a\chi} \geq b + \chi \tag{31b}$$

$$\text{For any } \left(a, \chi\right) \in \mathcal{K}_0 \text{ and any } \left(a', \chi'\right) \in \mathcal{K}_P : b + \chi \geq b + L\left(a, w_{a'\chi'}\right) x_{a'\chi'} \tag{31c}$$

and b clears the budget constraint

$$b + E = \iint_{\mathcal{K}_P} L\left(a, w_{a\chi}\right) \left(w_{a\chi} - x_{a\chi}\right) dG\left(a, \chi\right) f\left(a\right) da$$

Inequality (31a) ensures that the wage-setting process described by equation (4) induces for an employed worker of type (a,χ) the wage $w_{a\chi}$ and the associated ex-post surplus $x_{a\chi}$ designed for her type instead of the wage $w_{a'\chi'}$ and ex-post surplus $x_{a'\chi'}$ designed for any other participating type (a',χ'). Inequality (31b) ensures that participating types get a higher expected utility if they enter the labor force, while condition (31c) ensures that non-participating types are better off out of the labor force. It is worth noting that the value of the assistance benefit b has no impact on conditions (31a) to (31c).

We first consider $((a,\chi),(a,\chi')) \in (\mathcal{K}_P)^2$. From (31a), one obtains

$$L(a,w_{a\chi})\cdot x_{a\chi} \geq L\left(a,w_{a\chi'}\right)\cdot x_{a\chi'} \geq L(a,w_{a\chi})\cdot x_{a\chi}$$

The first inequality is obtained by replacing a' by a in the right-hand side of (31a). The second by inverting the roles of (a,χ) and (a,χ'). They together imply that $L(a,w_{a\chi})\cdot x_{a\chi} = L\left(a,w_{a\chi'}\right)\cdot x_{a\chi'}$. In other words, the government cannot distinguish between participants of the same skill level, but with different values of χ. This is because the χ characteristic is irrelevant for labor demand and wage-setting decisions and only matters for determining the participation decisions. Hence, although there is a bidimensional heterogeneity, the screening problem under random participation *à la* Rochet and Stole (2002) can be treated as a unidimensional screening problem by considering (12) instead of (31a) and the allocation can be indexed with respect to skill a only, as we do in the main text and henceforth in the Appendices.

Let $a \mapsto (w_a, x_a, \Sigma_a)$ be an allocation such that for all a, $\Sigma_a = L(a,w_a)\cdot x_a$ and for all a and a' (12) is verified. Condition (12) can be rewritten as:

$$\log \Sigma_{a'} - \log \Sigma_a \leq \log L\left(a',w_{a'}\right) - \log L(a,w_{a'})$$

Using the symmetric inequality where a and a' are inverted gives:

$$\log L\left(a',w_a\right) - \log L(a,w_a) \leq \log \Sigma_{a'} - \log \Sigma_a \leq \log L\left(a',w_{a'}\right) - \log L(a,w_{a'}) \tag{32}$$

Assume $a' > a$ and consider the two extreme parts of (32). They implies that

$$0 \leq \int_a^{a'} \left\{ \frac{\partial \log L}{\partial a}(t,w_{a'}) - \frac{\partial \log L}{\partial a}(t,w_a) \right\} dt$$

Since $a' > a$, and $\partial^2 \log L(a,w)/\partial a \partial w > 0$, this last inequality requires $w_{a'} \geq w_a$. Hence $a \mapsto w_a$ has to be nondecreasing. It is thus almost everywhere continuous. Take $a' > a$. Then from (32) we get

$$\frac{\log L(a',w_a) - \log(a,w_a)}{a'-a} \leq \frac{\log \Sigma_{a'} - \log \Sigma_a}{a'-a} \leq \frac{\log L(a',w_{a'}) - \log(a,w_{a'})}{a'-a}$$

As a' tends to a, the left-hand side of this condition tends to $\partial \log L(a,w_a)/\partial a$. The right-hand side tends to $\partial \log L(a,w_a)/\partial a$ as well, for any a where $a \mapsto w_a$ is continuous. Hence, $t \mapsto \Sigma_t$ admits a right-derivative for such $t = a$, which equals to $\partial \log L(a,w_a)/\partial a$. Redoing the same reasoning for $a' < a$ implies:

$$\frac{\dot{\Sigma}_a}{\Sigma_a} = \frac{\partial \log L}{\partial a}(a,w_a) \qquad \text{almost everywhere} \tag{33}$$

To show the reciprocal, let $a \mapsto (w_a, x_a, \Sigma_a)$ be an allocation such that for all a, $\Sigma_a = L(a,w_a)\cdot x_a$, $a \mapsto w_a$ is non-decreasing and (33) holds. We have to show that (12) holds for all

$a' \neq a$. Assume that $a' < a$ (respectively $a' > a$). Then we have for all $t \in [a', a]$ (resp. for all $t \in [a, a']$), that $w_t \geq w_{a'}$ (respectively $w_t \leq w_{a'}$). Since $\partial^2 \log L(a,w)/\partial a \partial w > 0$ this implies that:

$$\int_{a'}^{a} \left\{ \frac{\partial \log L}{\partial a}(t, w_t) - \frac{\partial \log L}{\partial a}(t, w_{a'}) \right\} dt \geq 0$$

which induces

$$\int_{a'}^{a} \frac{\partial \log L}{\partial a}(t, w_t)\, dt \geq \log L(a, w_{a'}) - \log L\left(a', w_{a'}\right)$$

Integrating (33) between a' and a, we see that the left-hand side of the last inequality equals to $\log \Sigma_a - \log \Sigma_{a'}$. Therefore, one has

$$\log \Sigma_a \geq \log \Sigma_{a'} + \log L(a, w_{a'}) - \log L\left(a', w_{a'}\right)$$

which is equivalent to (12).

C Proof of Proposition 1

From (3), one gets that $(T(w_a) + b)\, L(a, w_a)$ equals $L(a, w_a) \cdot w_a - \Sigma_a$, so the budget constraint (11) can be rewritten as

$$b = \int_{a_0}^{a_1} [L(a, w_a) \cdot w_a - \Sigma_a] \cdot G(a, \Sigma_a) \cdot f(a) \cdot da - E$$

Let $\sigma_a = \log \Sigma_a$. We use optimal control by considering σ_a as the state variable and w_a as the control.

$$\begin{aligned} \max_{w_a, \sigma_a} \quad & \int_{a_0}^{a_1} [L(a, w_a) \cdot w_a - \exp \sigma_a] \cdot G(a, \exp \sigma_a) \cdot f(a)\, da \\ s.t \quad : \quad & \dot{\sigma}_a = \frac{\partial \log L}{\partial a}(a, w_a) \end{aligned}$$

Let q_a be the multiplier associated to the equations of motion of σ_a and let $Z_a = -q_a$. The Hamiltonian writes

$$\mathcal{H}(w, \sigma, q, a) \stackrel{\text{def}}{\equiv} [L(a, w) \cdot w - \exp \sigma] \cdot G(a, \exp \sigma) \cdot f(a) + q \cdot \frac{\partial \log L}{\partial a}(a, w)$$

Since we assume that a maximum exists where w_a is a continuous function of a (see footnote 16), there exists a continuously differentiable function $a \mapsto q_a$, such that the following first-order condition are verified:

$$0 = \frac{\partial \mathcal{H}}{\partial w} = \frac{\partial (L(a, w) \cdot w)}{\partial w}(a, w_a) \cdot G(a, \Sigma_a) \cdot f(a) + q_a \cdot \frac{\partial^2 \log L}{\partial a \partial w}(a, w_a) \tag{34a}$$

$$-\dot{q}_a = \frac{\partial \mathcal{H}}{\partial \sigma} = -\left\{G(a, \Sigma_a) - [L(a, w_a) \cdot w_a - \Sigma_a] \cdot g(a, \Sigma_a)\right\} \cdot \Sigma_a \cdot f(a) \tag{34b}$$

together with the transversality conditions $q_{a_0} = q_{a_1} = \mu_{a_0} = \mu_{a_1} = 0$. Using $q_{a_1} = 0$, $Z_a = -q_a$, one has $Z_a = \int_a^{a_1} \dot{q}_t \cdot dt$. Hence, (34b) with (1) gives (13b). The transversality condition $q_{a_0} = 0$ gives $Z_{a_0} = 0$ in (13a). From (7), one has

$$\frac{\partial (L(a, w) \cdot w)}{\partial w}(a, w_a) = (1 - \eta(w_a)) \cdot L(a, w_a) \cdot w_a \tag{35}$$

From (10a) and (10b) one obtains

$$\frac{\partial^2 \log L}{\partial a \partial w}(a, w_a) = \frac{\alpha_a}{\varepsilon_a} \cdot \frac{\eta(w_a)}{a} \tag{36}$$

Introducing these two last expressions into (34b) gives the first equality in (13a).

D Proof of Proposition 2

We first show that Z is positive on (a_0, a_1). From (13b), one has

$$\dot{Z}_a = \left(\frac{\pi_a}{1+\pi_a} - \frac{x_a}{w_a}\right)(1+\pi_a)\cdot w_a \cdot h_a \tag{37}$$

Assume by contradiction that Z is negative at some point. Since $a \mapsto Z_a$ is continuous, there exists an interval where Z remains negative. Given that $Z_{a_0} = Z_{a_1} = 0$, this implies the existence of an interval $[\underline{a}, \overline{a}]$ such that $Z_{\underline{a}} = Z_{\overline{a}} = 0$ and such that $Z_a \leq 0$ for all $a \in [\underline{a}, \overline{a}]$.

- Since $Z_{\underline{a}} = 0$ and Z_a is negative in the neighborhood on the right of $\underline{a}$, one has $\dot{Z}_{\underline{a}} \leq 0$. Given (37) this implies that:
$$\frac{\pi_{\underline{a}}}{1+\pi_{\underline{a}}} \leq \frac{x_{\underline{a}}}{w_{\underline{a}}}$$
- Since $Z_a \leq 0$, one has from (13a) that $\eta(w_a) \geq 1$ for all $a \in [\underline{a}, \overline{a}]$. Given (8), this implies that x_a/w_a is nondecreasing, so
$$\frac{x_{\underline{a}}}{w_{\underline{a}}} \leq \frac{x_{\overline{a}}}{w_{\overline{a}}}$$
- Since $Z_{\overline{a}} = 0$ and Z_a is negative in the neighborhood on the left of $\overline{a}$, one has $\dot{Z}_{\overline{a}} \geq 0$. Given (37) this implies that
$$\frac{x_{\overline{a}}}{w_{\overline{a}}} \leq \frac{\pi_{\overline{a}}}{1+\pi_{\overline{a}}}$$

These three inequalities leads to $\pi_{\overline{a}} \geq \pi_{\underline{a}}$, so one must have $\underline{a} = \overline{a}$ since $a \to \pi_a$ is decreasing. Hence, Z_a is nonnegative on (a_0, a_1) and can only be nil pointwise.

Next, assume by contradiction that there exists $a_2 \in (a_0, a_1)$ such that $Z_{a_2} = 0$. Since Z_a is everywhere nonnegative, a_2 is an interior minimum of Z_a, so $\dot{Z}_{a_2} = 0$, and from (37)

$$\frac{\pi_{a_2}}{1+\pi_{a_2}} = \frac{x_{a_2}}{w_{a_2}}$$

However since $Z_{a_2} = 0$, one has $\eta(w_{a_2}) = 1$ from (13a). Hence, from (8) and the differentiability of $a \mapsto w_a$, x_a/w_a admits a derivative with respect to a that is nil. Since Z_a can only be nil pointwise within (a_0, a_1), there exists a real a_3 in the neighborhood of a_2 such that $a_3 > a_2$ and $Z_{a_3} > 0$. According to the mean value theorem, there exists $a_4 \in (a_2, a_3)$ such that $\dot{Z}_{a_4} = (Z_{a_3} - Z_{a_2})/(a_3 - a_2) > 0$. From (37), one obtains

$$\frac{\pi_{a_4}}{1+\pi_{a_4}} > \frac{x_{a_4}}{w_{a_4}}$$

Since a_4 is in the neighborhood of a_2 and $a \mapsto x_a/w_a$ has a zero derivative at a_2, then one has $(x_{a_4}/w_{a_4}) \simeq (x_{a_2}/w_{a_2})$ at a first-order approximation. However, $(\pi_{a_4}/(1+\pi_{a_4})) \simeq (\pi_{a_2}/(1+\pi_{a_2})) + \left(\dot{\pi}_{a_2}/(1+\pi_{a_2})^2\right)(a_4 - a_2)$ at a first-order approximation. Hence, since since $\dot{\pi}_{a_2} < 0$, one must have

$$\frac{\pi_{a_4}}{1+\pi_{a_4}} < \frac{\pi_{a_2}}{1+\pi_{a_2}} = \frac{x_{a_2}}{w_{a_2}} \simeq \frac{x_{a_4}}{w_{a_4}}$$

which leads to the contradiction. Therefore, Z_a is positive everywhere within (a_0, a_1).

From (13a), one has $\eta(w_a) < 1$ for any $a \in (a_0, a_1)$, which has different implications.

***i*)** For any $a \in (a_0, a_1)$, one has $\partial \log L/\partial w\,(a, w_a) > -1$ from (7). Moreover, at the *Laissez faire,* $\partial \log L/\partial w\left(a, w_a^{\mathrm{LF}}\right) = -1$ from (7) and (8). Hence, from (5) $w_a < w_a^{\mathrm{LF}}$ which means that optimal wages are distorted downwards. Furthermore, since $\partial L/\partial w\,(a, .) < 0$, one has $1 - L\,(a, w_a) < 1 - L\left(a, w_a^{\mathrm{LF}}\right)$ and unemployment rates are distorted downwards. Finally, $Z_{a_0} = Z_{a_1} = 0$ induces $w_{a_0} = w_{a_0}^{\mathrm{LF}}$, $L\,(a_0, w_{a_0}) = L\left(a_0, w_{a_0}^{\mathrm{LF}}\right)$, $w_{a_1} = w_{a_1}^{\mathrm{LF}}$ and $L\,(a_1, w_{a_1}) = L\left(a_1, w_{a_1}^{\mathrm{LF}}\right)$.

***ii*)** Since $\eta\,(w_a) < 1$, x_a/w_a is nonincreasing in a, so it is maximized at a_0. Since $Z_{a_0} = 0$ and $Z_a > 0$ on (a_0, a_1), one must have $\dot{Z}_{a_0} \geq 0$. Therefore, $x_{a_0}/w_{a_0} \leq \pi_{a_0}/\left(1 + \pi_{a_0}\right) < 1$. Hence for all a, $x_a < w_a$ and participation rates are distorted downwards.

***iii*)** $x < w$ for all w implies that the employment tax rate $\left(T\,(w) + b\right)/w$ is always positive. Moreover, it is nondecreasing since $\eta\,(w) < 1$. So, the average tax rate $T\,(w)\,/w$ is increasing in wage w. Finally (8) and $\eta\,(w) \leq 1$ induces $T'\,(w) \geq \left(T\,(w) + b\right)/w$, so marginal tax rate are positive everywhere.

E Proof of Proposition 3

The proof of Proposition 3 extends the one of Proposition 1 in Appendix C. Let λ be the multiplier associated to the budget constraint. From (3), $w_a - T\,(w_a) = \left(\Sigma_a/L\,(a, w_a)\right) + b$, so the Hamiltonian becomes:

$$\mathcal{H}\,(w, \sigma, q, a, b, \lambda) \stackrel{\mathrm{def}}{\equiv} \left[L\,(a, w)\,\Phi\left(\frac{\exp\sigma}{L\,(a, w)} + b\right) + \left(1 - L\,(a, w)\,\Phi\,(b)\right)\right] G\,(a, \exp\sigma) \cdot f\,(a)$$
$$+ \int_{\exp\sigma}^{+\infty} \Phi\,(b + \chi)\, g\,(a, \chi)\, f\,(a)\, d\chi + \lambda\left[L\,(a, w) \cdot w - \exp\sigma\right] \cdot G\,(a, \exp\sigma) \cdot f\,(a) + q \cdot \frac{\partial \log L}{\partial a}\,(a, w)$$

The first-order conditions now becomes, where we define $Z_a = -q_a/\lambda$

$$\begin{aligned} 0 &= \frac{1}{\lambda}\frac{\partial \mathcal{H}}{\partial w} = [\frac{\partial L\,(a, w)}{\partial w}\,(a, w_a)\,\frac{\Phi\left(\frac{\Sigma_a}{L(a,w_a)} + b\right) - \Phi\,(b) - \frac{\Sigma_a}{L(a,w_a)}\Phi'\left(\frac{\Sigma_a}{L(a,w_a)} + b\right)}{\lambda} \\ &\quad + \frac{\partial\left(L\,(a, w) \cdot w\right)}{\partial w}\,(a, w_a)] \cdot G\,(a, \Sigma_a) \cdot f\,(a) - Z_a \cdot \frac{\partial^2 \log L}{\partial a \partial w}\,(a, w_a) \end{aligned}$$

$$\begin{aligned} \dot{Z}_a &= \frac{1}{\lambda}\frac{\partial \mathcal{H}}{\partial \sigma} = \left\{\frac{\Phi'\left(\frac{\Sigma_a}{L(a,w_a)} + b\right)}{\lambda} - G\,(a, \Sigma_a) + \left[L\,(a, w_a) \cdot w_a - \Sigma_a\right] \cdot g\,(a, \Sigma_a)\right. \\ &\quad \left.\frac{L\,(a, w_a)\,\Phi\left(\frac{\Sigma_a}{L(a,w_a)} + b\right) + \left(1 - L\,(a, w_a)\right)\Phi\,(b) - \Phi\,(b + \Sigma_a)}{\lambda} \cdot g\,(a, \Sigma_a)\right\} \cdot \Sigma_a \cdot f\,(a) \end{aligned}$$

These two conditions with the transversality conditions $Z_{a_0} = Z_{a_1} = 0$, (35) and (36) give (23a) to (23d). Finally, the condition with respect to b is exactly (24).

References

[1] Akerlof, G, 1978, The Economics of Tagging as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Training, *American Economic Review*, **68**, 8-19.

[2] Blundell, R. and T. MaCurdy, 1999, Labor Supply: A Review of alternative approaches, in O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, North-Holland.

[3] Boadway, R., K. Cuff and N. Marceau, 2003, Redistribution and employment policies with endogenous unemployment *Journal of Public Economics*, **87**, 2407-2430.

[4] Boone, J. and L. Bovenberg, 2002, Optimal Labour Taxation and Search , *Journal of Public Economics*, **85**, 53-97.

[5] Boone, J. and L. Bovenberg, 2004, The Optimal Taxation of Unskilled Labor with Job Search and Social Assistance, *Journal of Public Economics*, **88**, 2227-2258.

[6] Boone, J. and L. Bovenberg, 2006, The Optimal Welfare and In-work benefits with search unemployment and observable abilities, *Journal of Economic Theory*, **126**, 165-193.

[7] Choné, P. and G. Laroque, 2005, Optimal Incentives for Labor force participation, *Journal of Public Economics*, **89**, 395-425.

[8] Choné, P. and G. Laroque, 2007, Negative marginal tax rates and heterogeneity, *mimeo* CREST.

[9] Diamond, P., 1980, Income Taxation with fixed hours of work, *Journal of Public Economics*, **13**, 101-110.

[10] Feldstein, M., 1996, The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act, *Journal of Political Economy*, **103**, 551-72.

[11] Golosov, M., N. Kocherlakota and A. Tsyvinski, 2003, Optimal Indirect and Capital Taxation, *Review of Economic Studies*, **70**, 569-547.

[12] Gruber, J. and E. Saez, 2002, The elasticity of taxable income: evidence and implications, *Journal of Public Economics*, **84**, 1-32.

[13] Guesnerie, R., 1995, *A Contribution to the Pure Theory of Taxation*, Cambridge University Press.

[14] Guesnerie, R. and Laffont, J-J, 1984, A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm, *Journal of Public Economics*, **25**, 329-369.

[15] Hammond, P., 1979, Straightforward Individual Incentive Compatibility in Large Economies, *Review of Economic Studies*, **46**, 263-282.

[16] Hersoug, T., 1984, Union Wage Responses to Tax Changes, *Oxford Economic Papers*, **36**, 37-51.

[17] Hosios, A.,1990, On the Efficiency of Matching and Related Models of Search and Unemployment, *Review of Economic Studies*, **57**, 279-298.

[18] Hungerbühler, M., E. Lehmann, A. Parmentier, B. Van der Linden, 2006, Optimal Redistributive Taxation in a Search Equilibrium Model, *Review of Economic Studies*, **73**, 743-767.

[19] Immervoll, H., Kleven, H., Kreiner, C. T. and Saez, E., 2007, Welfare reforms in European countries: a Microsimulation analysis, *Economic Journal*, **117**, 1-44.

[20] 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.

[21] Manning, A., 1993, Wage Bargaining and the Phillips Curve: The Identification and Specification of Aggregate Wage Equations, *Economic Journal*, **103**, 98-118.

[22] Meghir C. and D. Phillips, 2008, Labour Supply and Taxes, *IZA discussion paper* 3405.

[23] Mirrlees, J., 1971, An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies*, **38**(2), 175-208.

[24] Mirrlees, J., 1976, Optimal Tax Theory: A Synthesis, *Journal of Public Economics*, **6**(3), 327-358.

[25] Mirrlees, J., 1999, Optimal marginal tax rates at low incomes, *mimeo*, Faculty of Economics, Cambridge, England.

[26] Moen, E., 1997, Competitive Search Equilibrium, *Journal of Political Economy*, **105**(2), 385-411.

[27] 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 B, North-Holland, Amsterdam.

[28] Musgrave, R. A. and. Musgrave P.B, 1976, *Public Finance in Theory and Practice*, Second Edition, McGraw-Hill, New-York.

[29] 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.

[30] Pisauro, G., 1991, The effect of taxes on labour in efficiency wage models, *Journal of Public Economics*, **46**(3), 329-345.

[31] 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.

[32] Pissarides, C. A., 2000, *Equilibrium Unemployment Theory*, Second Edition, MIT Press, Cambridge, USA.

[33] Rochet, J-C, 1985, The taxation principle and multi-time Hamilton-Jacobi equations, *Journal of Mathematical Economics*, **14**, 113-128.

[34] Rochet, J-C and L. Stole, 2002, NonLinear Pricing with Random Participation, *Review of Economic Studies*, **69**, 277-311.

[35] Røed, K., S. Strøm, 2002, Progressive Taxes and the Labour Market: Is the Tradeoff Between Equality and Efficiency Inevitable? , *Journal of Economic Surveys*, **16**, 77-110.

[36] Rogerson, R., R. Shimer and R. Wright, 2005, Search-Theoretic Models of the Labor Market: A Survey, *Journal of Economic Literature*, **43**, 959-988.

[37] Saez, E, 2001, Using Elasticities to Derive Optimal Income Tax Rates, *Review of Economics Studies*, **68**, 205-229.

[38] Saez, E., 2002, Optimal Income Transfer Programs:Intensive Versus Extensive Labor Supply Responses, *Quarterly Journal of Economics*, **117**, 1039-1073.

[39] Shimer, R., 1996, Contracts in frictional labor markets, MIT mimeo.

[40] Sørensen, P. B., 1997, Public finance solutions to the European unemployment problem?, *Economic Policy*, **25**, 223-264.

Département des Sciences Économiques
de l'Université catholique de Louvain
Institut de Recherches Économiques et Sociales

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2008/3082/036