CESifo Working Papers

www.cesifo.org/wp

Optimal Income Taxation with Endogenous Participation and Search Unemployment

Etienne Lehmann
Alexis Parmentier
Bruno Van Der Linden

CESIFO WORKING PAPER NO. 3324
CATEGORY 1: PUBLIC FINANCE
JANUARY 2011

CESifo *Working Paper No. 3324*

Optimal Income Taxation with Endogenous Participation and Search Unemployment

Abstract

We characterize optimal redistributive taxation when individuals are heterogeneous in 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. Average tax rates are increasing at the optimum. This shifts wages below their laissez faire value and distorts labor demand upwards. The marginal tax rate is positive at the top of the skill distribution even when the latter is bounded. These results are analytically shown under a Maximin objective when the elasticity of participation is decreasing in the skill level and are numerically confirmed under a more general objective.

JEL-Code: H210, H230, J640.

Keywords: non-linear taxation, redistribution, adverse selection, random participation, unemployment, labor market frictions.

Etienne Lehmann
CREST (INSEE)
France – 92245 Malakoff Cedex
etienne.lehmann@ensae.fr

Alexis Parmentier
EPEE – University of Evry Val d'Essonne
France – 91025 Evry Cedex
alexis.parmentier@univ-evry.fr

Bruno Van Der Linden
IRES, Catholic University of Louvain
Belgium – 1348 Louvain-la-Neuve
bruno.vanderlinden@uclouvain.be

3rd January 2011
We thank for their comments an anonymous Referee, the Editor, and participants at seminars at the Université Catholique de Louvain, CREST, Malaga, EPEE-Evry, Gains-Le Mans, CES-Paris 1, ERMES-Paris 2, Thema-Cergy Pontoise, Copenhagen, Uppsala, Rotterdam, Munich, the IZA-SOLE 2008 meeting, the LAGV 2008 meeting, the HetLab conference at Konstanz, The Economics of Labor Income Taxation workshop at IZA 2008, AFSE 2008, EALE 2008, T2M 2009, PET 2009 and the CESifo workshop on European Unemployment, with special thanks to Daron Acemoglu, Sören Blomquist, Pierre Cahuc, Helmuth Cremer, Bruno Decreuse, Christian Holzner, Laurence Jacquet, Bas Jacobs, Claus Thustrup Kreiner, Guy Laroque, Cecila Garcia-Peñalosa, Fabien Postel-Vinay and Laurent Simula. Mathias Hungerbühler was particularly helpful in providing suggestions and comments at various stages. Any errors are ours. This research was partly conducted while Bruno Van der Linden was visiting ERMES (Université Paris 2) and while Etienne Lehmann was visiting IRES (Université Catholique de Louvain). 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. Simulations programs are available at http://www.crest.fr/ck.nder/user.les/.les/Pageperso/elehmann/LPVsimulations.zip

I Introduction

In the literature on optimal redistributive taxation, the labor supply responses along the intensive (Mirrlees 1971) or extensive (Diamond 1980, Saez 2002) margins are the only sources of deadweight losses. However, in this literature, non-employment, if any, is synonymous with non-participation. According to Mirrlees (1999), a "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 the presence of (such involuntary) unemployment which is an important source of inequality, one should depart from the assumption of walrasian labor markets. We provide an optimal tax formula in a search-matching framework where wages, employment, (involuntary) unemployment and (voluntary) non participation are affected by taxation on labor incomes.

Our economy is made of a continuum of skill-specific labor markets. On each of them, we introduce matching frictions *à la* Diamond (1982) and Mortensen and Pissarides (1999) to generate unemployment. Taxes are distortive because the government can only condition them on endogenous wages. As in most labor market models, we assume that the equilibrium gross wage maximizes an objective that is increasing in the after-tax (net) wage and decreasing in the pre-tax (gross) wage. This is because the former increases employees' welfare, whereas the latter decreases employers' profit. When taxation becomes more progressive,[1] a higher pre-tax wage becomes less attractive to workers, so 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 the unemployment rate on each skill-specific labor market. We call this response the"*wage-cum-labor demand*" margin. To focus on redistribution, we abstract from standard inefficiencies arising from search frictions by imposing a wage-setting mechanism that maximizes total resources in the absence of taxes (i.e. in the *laissez faire*). To account for the extensive margin, we assume that whatever their skill level, individuals differ in their value of remaining out of the labor force.[3] A higher level of taxes reduces the returns to participation, thereby inducing some individuals to give up search.

[1] In this paper, a tax schedule is progressive when employment tax rates are increasing along the wage distribution. The employment tax rate divides the sum of the tax liability and the assistant benefit by the gross wage level. This corresponds to what Immervoll *et alii* (2007) among others call *participation* tax rate. In the presence of unemployment, it is more appropriate to use the term *employment* tax rate..

[2] As overviewed by Bovenberg (2006), this property holds in the monopoly union model (Hersoug, 1984), in the right-to manage union model (Lockwood and Manning, 1993), in the matching model (Pissarides, 1998), in the efficiency wage model (Pisauro, 1991) and also in the textbook competitive labor supply framework. In the latter, a higher (annual) pre-tax wage is obtained thanks to more effort in employment (the so-called intensive margin of the labor supply). For simplicity, we ignore labor supply responses along the intensive margin.

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

We derive an optimal tax formula in terms of behavioral elasticities by considering a tax perturbation approach in the spirit of Piketty (1997) and Saez (2001). Intuitively, given the redistributive objective and the participation response, the optimum typically requires a progressive income tax schedule. This generates a downward distortion along the wage-cum-labor demand margin. On the one hand, the latter distortion reduces resources available for redistribution. On the other hand, the rise in the employment probability enables to share more equally incomes among employed and unemployed workers of the same skill level. Simulations confirm our intuition that optimal taxation is progressive. Average tax rates are increasing along the wage distribution. Marginal tax rates are positive at the top of the income distribution, even when the skill distribution is bounded. Marginal tax rates can be negative at the bottom and transfers to low-skilled workers can be larger than transfers to the non-employed, i.e. an EITC can be desirable.

We obtain analytical results under a Maximin (Rawlsian) social objective. Optimal marginal tax rates are positive everywhere and optimal average tax rates are increasing when the elasticity of participation decreases along the distribution of skills. The reason is that a progressive income tax schedule is then optimal as it increases the level of tax at skill levels where participation reacts more strongly to the tax pressure. The optimal tax schedule thus reduces wages and increases labor demand to ease redistribution.

A number of studies are related to our work. In the optimal taxation literature that focuses on the intensive margin (Mirrlees 1971), the optimal marginal tax rate at the top is nil if the skill distribution is bounded (Sadka 1976, Seade 1977). This implies that the average tax has to be decreasing in the upper part of this distribution (Hindriks *et alii* 2006, Boadway and Jacquet 2008). Taking an unbounded (Pareto) distribution of skills, Diamond (1998) and Saez (2001) show that asymptotic marginal tax rates are positive. In our paper, the marginal tax rate is positive at the top even when the skill distribution is bounded.

Both the intensive labor supply and the wage-cum-labor demand margins account for the empirical fact that gross earnings decrease with marginal tax rates (Saez *et alii* 2010). Which of these two margins matters more remains an open empirical question. We believe that our wage-cum-labor demand margin might be crucial. Blundell and MacCurdy (1999) and Meghir and Phillips (2008) conclude that the elasticity of the intensive labor supply margin is likely very small. Manning (1993) finds a significantly negative effect of tax progressivity on the UK unemployment rate (see also Sørensen 1997 and Røed and Strøm 2002), which is consistent with the presence of a wage-cum-labor demand response to tax progressivity. The wage-cum-labor demand margin is also a plausible explanation for the result obtained by Blomquist and Selin (2010) according to which the hourly wage rate elasticity is similar to the taxable labor income elasticity with respect to the marginal tax rates for males in Sweden.

There is growing evidence that participation decisions matter a lot (Meghir et Philips, 2008). Diamond (1980), Saez (2002) and Choné and Laroque (2005, 2010) have thus studied optimal income taxation when individuals' decisions are limited to a binary choice between working or not, wages are exogenous and there is no unemployment. This "pure extensive" literature focuses on the rationale for an EITC and is silent on the shapes of optimal average and marginal tax rates. On the contrary, our paper provides results on the whole income tax profile. It can be shown in the pure extensive setting that average tax rates are increasing if the social objective is Maximin and the participation elasticity is decreasing along the skill distribution. We retrieve this analytical result in a more general model that does account for two important facts: the existence of gross incomes responses to marginal tax rates and the presence of involuntary unemployment, which is an important source of income inequality.

Saez (2002) has proposed a model of optimal taxation with both extensive and intensive labor supply margins. While he does not provide analytical results for the mixed case, his simulations show that the EITC is optimal when responses along the extensive margin are more important than responses along the intensive one and the social objective is not Maximin. We emphasize the role of the monotonicity of the elasticities of participation. Furthermore, his simulations consider only few points in the bottom half of the income distribution, while ours offer a much broader picture along the whole wage distribution. Moreover, in a "mixed" setting *à la* Saez (2002), the optimal marginal tax rate is nil at the top when the income distribution is bounded (Jacquet *et alii* 2010). Consequently, average tax rates cannot be increasing everywhere in such a setting.

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 cannot 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.

Hungerbühler *et alii* (2006), henceforth HLPV, propose an optimal income tax model with unobservable workers' skills and with wage-cum-labor demand responses in a matching framework. HLPV assumes that all individuals face the same cost of participation whatever their skill level. Consequently, every agent above (below) an endogenous threshold of skill participates (does not participate). Instead here, this cost varies both within and between skill levels. Our model thus integrates an extensive margin of the labor supply à la Diamond (1980), Saez (2002) and Choné and Laroque (2005, 2010). Contrary to HLPV, our model predicts that negative marginal tax rates and an EITC for the low-skilled can be optimal. In addition, the present paper differs from HLPV in the following aspects. First, HLPV does not express optimality conditions in terms of behavioral elasticities. Second, the social welfare function in HLPV does not take into account the issue of income redistribution between employed and unemployed individuals of the same skill level. Third, HLPV imposes a Cobb-Douglas matching function, while we do not. Finally, while HLPV calibrates the skill distribution by assuming a lognormal distribution, we here base the calibration of the model on the true earnings distribution in the US using a Kernel procedure.

The paper is organized as follows. The model and fiscal incidence are presented in the next section. Section III presents the optimality conditions under a general utilitarian criterion. Section IV characterizes the Maximin optimum. 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 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 with homogeneous tastes. Earnings differ in our model for three reasons. First, individuals are endowed with different levels of productivity (or skill) denoted by a. The density of skills $f(.)$ is continuous and positive on $[a_0, a_1]$, with $0 < a_0 < a_1 \leq +\infty$. Population size 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. χ represents 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). A worker of skill a produces a units of output if and only if she is employed in a type a job, otherwise her production is nil. This assumption of perfect-segmentation 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.

2. For each skill level a, firms choose the number of job vacancies they open. 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 and the wage level. An individual of type (a, χ) who chooses to participate renounces χ. All participants of skill a are alike during the matching process. We henceforth call these individuals participants of type a. Each employed worker supplies an exogenous amount of labor normalized to 1. So, earnings and (gross) wages are the same among equally skilled workers.

4. Each worker of skill a 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*. 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. An individual of type (a, χ) can 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 (and a utility level) $w_a - T(w_a)$ or she becomes unemployed and gets the assistance benefit b.[5]

II.1 Participation decisions

An individual of type (a, χ) participates if her expected income, $\ell_a (w_a - T(w_a)) + (1 - \ell_a) b$, is higher than in case of non participation, $b+\chi$. Let $\Sigma_a \equiv \ell_a (w_a - T(w_a) - b)$ denote the expected surplus of a participant of type a and $G(a, .)$ be the cumulative distribution of the value of non-market activities, conditional on the skill level. The participation rate among individuals of skill a equals $G(a, \Sigma_a)$ and the mass of participants of type a equals $U_a = G(a, \Sigma_a) f(a)$. We denote $g(a, \Sigma)$ the continuous conditional density of the value of non-market activities. It is supposed to be strictly positive on an interval whose lower bound is 0. Hence, at any skill level, there

[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. We also do not consider the possibility that redistribution could also be based on observable characteristics related to skills (see Akerlof, 1978).

[5] One can adopt an alternative assumption, namely that χ is a *cost* of searching. Then, nonparticipating individuals get utility b, the employed get $w_a - T(w_a) - \chi$ and the unemployed get $b - \chi$. The behavior of economic agents remains unchanged. However, the social objective is different and the Maximin case is much less tractable. Furthermore, our model can easily be extended to include a fixed cost of working, provided that this cost depends on the skill level a and not on χ.

is a positive mass of participants. Note that the characteristics a and χ can be independently distributed or not. We define

$$\pi_a \equiv \frac{\Sigma_a \; g(a, \Sigma_a)}{G(a, \Sigma_a)} \tag{1}$$

the elasticity of the participation rate with respect to Σ, at $\Sigma = \Sigma_a$. This elasticity is in general 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 suitable workers 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 given by the matching function $H(a, V_a, U_a)$, where V_a and U_a stand for the numbers of vacancies and job-seekers. As is usual in this literature, the matching function is twice-continuously differentiable, is increasing in both U_a and V_a, exhibits constant returns to scale in (U_a, V_a), verifies $H(a, V_a, 0) = H(a, 0, U_a) = 0$, and $H(a, V_a, U_a) < \min(V_a, U_a)$.

Define tightness θ_a as the ratio V_a/U_a. The probability that a vacancy is filled equals $q(a, \theta_a) \equiv H(a, 1, 1/\theta_a) = H(a, V_a, U_a)/V_a$. Due to congestion externalities in the matching process, 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 function of θ_a. Symmetrically, the probability that a job-seeker finds a job $\theta_a q(a, \theta_a) = H(a, \theta_a, 1) = H(a, V_a, U_a)/U_a$ is an increasing function of tightness. Firms and individuals being atomistic, they take tightness θ_a as given.

When a firm creates a vacancy of type a, it fills it with probability $q(a, \theta_a)$. Then, its profit at stage 4 equals $a - w_a$. Therefore, its expected profit at stage 2 equals $q(a, \theta_a)(a - w_a) - \kappa(a)$. Firms create vacancies until the free-entry condition $q(a, \theta_a)(a - w_a) = \kappa(a)$ is met. This pins down the value of tightness θ_a and in turn the probability of finding a job through[6]

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

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

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

From the assumptions made on the matching function, the labor demand function $L(.,.)$ is twice-continuously differentiable and admits values within $(0, 1)$. As the wage increases, firms

[6] $q^{-1}(a, .)$ denotes the reciprocal of $\theta \mapsto q(a, \theta)$, holding a constant.

get a lower profit on each filled vacancy, fewer vacancies are created and tightness decreases. Hence $\partial L/\partial w_a < 0$. Moreover, due to the constant-returns-to-scale assumption, the probability of being employed depends on skill and wage levels but 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

In the presence of search frictions, a match between an employer and a job-seeker creates a joint surplus that needs to be shared. When a vacancy is filled in segment a, a surplus $a - w$ accrues to the employer. A worker gains $w - T(w) - b$ from finding a job. This surplus does not depend on her value of nonmarket activities χ since she renounces χ when she decides to participate to the labor market.

We focus on redistribution and consider a sharing rule of the total surplus such that the role of taxation is only to redistribute income (as in Mirrlees) and not to restore efficiency.[7] 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 the wage maximizes the Nash Product $(w - T(w) - b)^{\gamma} (a - w)^{1-\gamma}$ and that the workers' bargaining power γ equals the elasticity of the matching function with respect to unemployment (see Hosios 1990). The latter assumption is only meaningful if the elasticity of the matching function is constant and exogenous. With a Cobb-Douglas matching function $H(a, U_a, V_a) = A\ (U_a)^{\gamma} (V_a)^{1-\gamma}$, 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:

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

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

Various wage-setting mechanisms can provide alternative microfoundations for (4). The Competitive Search Equilibrium introduced by Moen (1997) and Shimer (1996) leads to this property when search is directed by wages and by skill levels (see Appendix A). Another possibility is to assume that a skill-specific utilitarian monopoly union selects the wage w_a after indi-

[7] Boone and Bovenberg (2002) studies how nonlinear taxation can restore efficiency when the Hosios condition is not fulfilled. Hungerbühler and Lehmann (2009) extends HLPV by relaxing the Hosios assumption.

viduals' participation decisions but before firms' decisions about vacancy creation (see Mortensen and Pissarides, 1999).

II.4 The equilibrium

The objective in (4) multiplies the employment probability by the difference between the net incomes in employment and in unemployment. We call this latter difference the *ex-post surplus* $x = w - T(w) - b$. It subtracts an "employment tax" $T(w) + b$ from the wage w.

For a given employment tax function $T(.)+b$, the equilibrium allocation is recursively defined. *i*) The wage-setting equations (4) determine wages w_a and in turn $x_a = w_a - T(w_a) - b$. *ii*) The labor demand functions (2) determine the skill-specific employment probabilities $\ell_a = L(a, w_a)$ and unemployment rates $1 - L(a, w_a)$. *iii*) From (3), expected surpluses equal $\Sigma_a = \ell_a x_a$. Participation rates are given by $G(a, \Sigma_a)$ and employment rates by $L(a, w_a) G(a, \Sigma_a)$. *iv*) For each additional worker of type a, the government collects taxes $T(w_a)$ and saves the assistance benefit b. Hence, each additional worker of skill a increases the government's revenue by an amount equal to the employment tax $T(w_a) + b$. Denoting $E \geq 0$ the exogenous amount of public expenditures, the government's budget constraint sets the level of b:

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

Each additional *participant* of skill a expects a gross wage $w_a\ L(a, w_a)$ and a net income $\Sigma_a + b$ (see (3)). Hence, the average employment tax $(T(w_a) + b)\ L(a, w_a)$ per participant of skill a equals $w_a\ L(a, w_a) - \Sigma_a$ and the budget constraint can be rewritten as:

$$b = \int_{a_0}^{a_1} [w_a\ L(a, w_a) - \Sigma_a]\ G(a, \Sigma_a)\ f(a)\, da - E \tag{5}$$

II.5 The *laissez-faire*

The *laissez-faire* is defined as the economy without tax and benefit. According to (4), the equilibrium level of wage maximizes $w\ L(a, w)$. To ensure that program (4) is well-behaved at the *laissez-faire*, we assume[8] that for any (a, w),

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

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

$$\frac{\partial^2 \log L}{\partial a \partial w}(a, w) > 0 \tag{7}$$

[8] We show in Appendix B how to recover the matching technology and the vacancy cost from the labor demand function $L(.,.)$. There is therefore no loss of generality in considering $L(.,.)$ as a primitive of the model.

These two properties are not restrictive. They are for instance met when the matching function takes a CES specification $H(a,U,V) = (U^{-\rho} + V^{-\rho})^{-\frac{1}{\rho}}$ with $\rho > 0$, and when the vacancy cost increases less than proportionally or decreases with the skill level a (so that $a\ \dot{\kappa}(a)/\kappa(a) \leq 1$). The latter is standard (Pissarides 2000). From (2) the labor demand function is $L(a,w) = \left[1 - \left(\frac{a-w}{\kappa(a)}\right)^{-\rho}\right]^{\frac{1}{\rho}}$ and one has in addition that $\partial L/\partial a > 0$.

II.6 Fiscal incidence

We now reintroduce the tax/benefit system and explain how tax reforms affect the equilibrium. The first-order condition[9] of (4) writes:

$$-\frac{\partial \log L}{\partial \log w}(a, w_a) = \eta(w_a) \tag{8}$$

where

$$\eta(w) \equiv \frac{1 - T'(w)}{1 - \frac{T(w)+b}{w}} = \frac{\partial \log(w - T(w) - b)}{\partial \log w} \tag{9}$$

When the wage increases by one percent, the left-hand side of (8) measures the relative decrease in the employment probability, while the right-hand side $\eta(w_a)$ denotes the wage elasticity of the ex-post surplus. In equilibrium, these two terms are equal.

Figure 1 illustrates condition (8). From (4), the tax and benefit system influences the wage through the shape of the ex-post surplus function $x(w) = w - T(w) - b$. Because of the multiplicative form of (4), we put the log of the wage along the horizontal axis and the log of the ex-post surplus along the vertical one. Hence in Figure 1, the slope of the ex-post surplus function equals the elasticity $\eta(w)$. Along the curves $\log x =$constant$- \log L(a,w)$, the wage-setting objective remains constant. From (2) and (6), these indifference curves are increasing and convex. The solution to Program (4) then consists in choosing the highest indifference curve tangent to the ex-post surplus function. The first-order condition (8) expresses this tangency condition.

Figure 1 here

Consider now a tax reform such that the ex-post surplus function becomes steeper in Figure 1, so η rises. 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. We denote ε_a the (substitution) elasticity of the wage with respect to η for workers

[9] 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.

of skill a. Moreover, we define α_a as the elasticity of the wage with respect to the skill level. Appendix C explains that these elasticities are endogenous and equal to:

$$\varepsilon_a = -\frac{\eta(w_a)}{w_a\left(\frac{\partial^2 \log L(a,w_a)}{\partial w\ \partial \log w} + \eta'(w_a)\right)} > 0 \tag{10a}$$

$$\alpha_a = -\frac{a}{\frac{\partial^2 \log L(a,w_a)}{\partial w\ \partial \log w} + \eta'(w_a)} \frac{\partial^2 \log L}{\partial a \partial w}(a, w_a) > 0 \tag{10b}$$

A rise in η is generated by either a rise in the assistance benefit ratio b/w_a, a rise in the average tax rate $T(w_a)/w_a$, or a decrease in the marginal tax rate $T'(w_a)$. For comparative statics purposes, consider for a while these three latter terms as parameters. So, $\eta(w_a)$ is provisionally a parameter, too. 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.[10] In particular, they are consistent with the empirical findings surveyed in Saez *et alii* (2010) according to which earnings increase when marginal tax rates decrease, especially for top-income earners. In our model, the link between marginal tax rates and wages comes through the surplus sharing process and not through the intensive margin of the labor supply.

We now determine the "efficient"level of η. We define efficiency by the maximization of government's revenue, holding constant the level of the wage-setting objective Σ_a (thereby the participation rates $G(a, \Sigma_a)$). Given the budget constraint (5), in the labor market of type a, the wage is efficient if it maximizes the average gross wage $w\ L(a, w)$ per participant. This occurs if the wage level reaches its *laissez-faire* value w_a^{LF} (see (4)). This corresponds to the case where $\eta = 1$ (see (8)). In this sense, non-distortive taxation is characterized by the equality between the marginal tax rate $T'(w_a)$ and the employment tax rate $\frac{T(w_a)+b}{w_a}$ (i.e. $\eta(w_a) = 1$). In the case where $\eta(.) < 1$ (resp. > 1), the wage is below (above) its *laissez-faire* value, so a rise in wage increases (decreases) resources available for redistribution.[11] From Equation (9), wages are distorted downwards (upwards) whenever the ex-post surplus increases less (more) than proportionally in wages, that is, whenever the employment tax rate $\frac{T(w)+b}{w}$ is increasing

[10] See Footnote 2 and empirical evidence in Sørensen (1997), Røed and Strøm (2002) and Manning (1993).

[11] This footnote compares the efficient marginal tax rates in our model and in a labor supply model that generates the same responses of gross earnings to taxation. For this purpose, we take b as a parameter and we consider a labor supply model where preferences over after-tax earnings $C = w - T(w)$ and before-tax earnings w are given by $U = (C - b) L(a, w)$. There, $L(., a)$ represents the disutility of labor for workers of skill a. In both models the gross wage w_a maximizes the same objective $(w - T(w) - b) L(a, w)$, so the wage responses are identical. However, efficient marginal tax rates are different. They are nil (i.e. lump-sum taxation) in the labor supply setting. Conversely, in our model, a rise in the marginal tax rate does not only reduce the gross wage, but it also increases the labor demand, thereby the number of taxpayers. This additional effect is beneficial (detrimental) for the government when the employment tax is positive (negative). This intuitively explains why the efficient marginal tax rate has the same sign as the employment tax in our model while it is nil in the labor supply model.

(decreasing) in the wage level.[12]

Assume now an upward shift of the ex-post surplus function in Figure 1, keeping the elasticity of the ex-post surplus η unchanged. Equation (8) indicates that the equilibrium wage remains unaffected. In this specific sense, there is no *income effect* on the wage w_a, nor on the employment probability $L(a, w_a)$. However, the reform reduces the employment tax $T(w_a) + b$. So, the surplus Σ_a an agent of type a can expect from participation increases. The participation rate $G(a, \Sigma_a)$ thus increases and the employment rate $L(a, w_a)\ G(a, \Sigma_a)$ as well. The magnitude of this behavioral response is captured by the elasticity of participation π_a defined in (1).

II.7 Government's objective and incentive constraints

The government cares about inequalities measured in terms of the net income that accrues to agents according to their position on the labor market. We consider the following Bergson-Samuelson social welfare function:

$$\int_{a_0}^{a_1} \left\{ \left[\ell_a\ \Phi\left(w_a - T\left(w_a\right)\right) + \left(1 - \ell_a\right)\Phi\left(b\right)\right]\ G\left(a, \Sigma_a\right) + \int_{\Sigma_a}^{+\infty} \Phi\left(b + \chi\right)\ g\left(a, \chi\right) d\chi \right\} f\left(a\right) da \quad (11)$$

where $\Phi'(.) > 0 > \Phi''(.)$. The stronger the concavity of $\Phi(.)$, the more averse to inequality is the government. The Benthamite objective corresponds to the case where $\Phi(.)$ is linear. The Maximin (Rawlsian) objective is the polar case and is considered in Section IV. There, the government values only the utility of the least well-off. Unemployed individuals are the least well-off because they get b, which is always lower than the workers' and non participants' utility levels, which are respectively equal to $w - T(w)$ and $b + \chi$.

The government does not observe the productivity of each job but only the wage negotiated by each worker-firm pair. So, it aims at maximizing its objective subject to the budget constraint (5) and the choices made by the agents. Since a worker-firm pair maximizes an objective $L(a, w)(w - T(w) - b)$ that is increasing in the ex-post surplus $x = w - T(w) - b$ and decreasing in gross wages, the government's self-selection problem can be viewed as one where worker-firms pairs of skill a are agents with an objective $L(a, w)\ x$. Therefore, 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:

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

This condition expresses that a firm-worker pair of type a chooses the bundle (w_a, x_a) rather than any other bundle $(w_{a'}, x_{a'})$ designed for firm-worker pairs of another type a'. From (7),

[12] Formally, $\partial \frac{T(w)+b}{w} / \partial w = \left(T'(w) - \frac{T(w)+b}{w}\right)/w = (1 - \eta(w))\left(1 - \frac{T(w)+b}{w}\right)/w$

the strict single-crossing condition holds. Hence, (12) is equivalent to the envelope condition associated to (4)

$$\dot{\Sigma}_a = \Sigma_a \frac{\partial \log L}{\partial a}(a, w_a) \tag{13}$$

and the monotonicity requirement that the wage w_a is a nondecreasing function of the skill level a. Following Mirrlees (1971), it is much more convenient to solve the government's problem in terms of allocations. Contrary to HLPV, and in the spirit of Saez (2001), we will express the optimality conditions in terms of behavioral elasticities that can be easily interpreted for applied purposes. We verify in Appendices D and E that the method in terms of allocations leads to the same optimality conditions.

III The general utilitarian case

Let $h_a = \ell_a\, G(a, \Sigma_a)\ f(a)$ denote the (endogenous) mass of workers of skill a. Under the social objective (11), Appendix D proves that:

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

$$\left(\frac{1-\eta(w_a)}{\eta(w_a)}\, w_a - \frac{\Phi(w_a - T(w_a)) - \Phi(b) - x_a\ \Phi'(w_a - T(w_a))}{\lambda}\right) \frac{\varepsilon_a}{\alpha_a}\ a\ h_a = Z_a \tag{14a}$$

$$Z_{a_0} = 0 \tag{14b}$$

$$\textit{where} \qquad Z_a = \int_a^{a_1} \left\{ \left(1 - \frac{\Phi'(w_t - T(w_t))}{\lambda}\right) x_t - \pi_t \left[T(w_t) + b + \Xi_t\right] \right\} h_t\ dt \tag{14c}$$

$$\textit{and} \qquad \Xi_t = \frac{\ell_t\ \Phi(w_t - T(w_t)) + (1-\ell_t)\,\Phi(b) - \Phi(b + \Sigma_t)}{\lambda\ \ell_t}, \tag{14d}$$

in which the positive Lagrange multiplier associated to the budget constraint (5), λ, verifies

$$\begin{aligned} \lambda \ &= \int_{a_0}^{a_1} \left\{ \ell_a\ G(a, \Sigma_a)\ \Phi'(w_a - T(w_a)) + (1-\ell_a)\ G(a, \Sigma_a)\ \Phi'(b) \right. \\ &\quad \left. + \int_{\Sigma_a}^{+\infty} \Phi'(b + \chi)\ g(a, \chi)\ d\chi \right\} f(a)\ da \end{aligned} \tag{15}$$

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). Moreover, w_a is determined by the wage-setting condition (8). This section provides an intuitive proof of this proposition and then studies the properties of the optimal tax schedule.

III.1 Intuitive proof of Proposition 1

To determine the marginal social value of public funds λ, consider a unit increase in public expenditures E financed by a unit decrease in b and a unit increase in tax liabilities $T(w)$. The employment tax rates are unaffected, so there is neither a response along the wage-cum-labor demand nor along the participation margins. The right-hand side of (15) captures the impact of such policy change on the social welfare function. Hence, λ is a weighted average of the social marginal utilities of consumption in the economy.

We represent in Figure 2 a tax reform that leads to (14a). Given that it is the ex-post surplus (and not the after-tax wage) that enters the wage-setting objective (4), and given the multiplicative form of this objective, we again represent the log of wages in the horizontal axis and the log of the ex-post surplus in the vertical axis. Starting from the optimal tax schedule, the wage-elasticity η of the ex-post surplus is marginally decreased by $\Delta\eta < 0$ for wages in the small interval $[w_a - \delta w, w_a]$.[13] This reform affects the government's problem along three effects: a *mechanical* effect, a *participation* effect for individuals of skill t above a, and a *wage* effect for participants whose wage before the reform lies in the $[w_a - \delta w, w_a]$ interval. These three effects have an impact not only on the government's revenue but also on the social objective (11). We divide these impacts on the social objective by the marginal social value of public funds λ to express them in monetary units.

Figure 2 here

The mechanical effect

Consider first the participants of any skill level t above a. As $\eta(.)$ is unchanged around w_t, the tax reform keeps the equilibrium wage and hence the employment probability ℓ_t unchanged. Conversely, the tax reform decreases the ex-post surplus by (see (9)):

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

The decrease in x_t corresponds to a rise in the employment tax level $T(w_t)+b$ by $\Delta(T(w_t)+b) = x_t \ (-\Delta\eta) \ (\delta w/w)$. This increases the resources of the government, but decreases the social welfare of the type t workers, which the government values at a rate $\Phi'(w_t - T(w_t))/\lambda$. Adding this welfare loss to the gain in tax receipts, and multiplying by the mass of employed workers, the *mechanical* effect at skill level t equals

$$\left(1 - \frac{\Phi'(w_t - T(w_t))}{\lambda}\right) \ x_t \ h_t \ (-\Delta\eta) \ \frac{\delta w}{w} \tag{16}$$

The sign of this effect is given by the difference between one and the marginal social welfare weight (expressed in terms of the value of public funds) $\frac{\Phi'(w_t - T(w_t))}{\lambda}$.

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

The participation effect

Consider next the participation decisions of individuals of skill t above a. From (3), since their employment probability is unchanged, their expected surplus decreases by the same relative amount $\Delta\Sigma_t/\Sigma_t = \Delta\eta \ (\delta w/w)$ as their ex-post surplus x_t does. According to (1) the mass of employed individuals of type t thus decreases by

$$\Delta h_t = \pi_t \ h_t \ \Delta\eta \ (\delta w/w)$$

For each of these additional non-employed individuals, the government loses $T(w_t) + b$ employment taxes.

On the labor market of skill t, a pivotal individual who is indifferent between participation or nonparticipation is characterized by $\chi = \Sigma_t$. The social welfare function decreases by $\ell_t \ \Phi(w_t - T(w_t)) + (1 - \ell_t)\,\Phi(b)$ when she exits the labor market but increases by $\Phi(b + \Sigma_t)$ when she enters the pool of non-participants. From (14d), the net effect equals $-\Xi_t \ \Delta h_t$ and is positively valued by an inequality-averse government as $\Xi_t < 0$. The *participation* effect at skill level t equals:

$$-\pi_t \ \left(T(w_t) + b + \Xi_t\right) \ h_t \ (-\Delta\eta) \ \frac{\delta w}{w} \tag{17}$$

In comparison with a setup with full employment, the welfare gain of additional participants is here less important because joblessness for some of them rises inequalities (an effect captured by the negative Ξ_t term).

Wage response effect

This effect concerns participants whose wage lies in the interval $[w_a - \delta w, w_a]$ if they are employed in the pre-reform economy. The interval $[w_a - \delta w, w_a]$ of the wage distribution corresponds to the interval $[a - \delta a, a]$ of the skill distribution, where according to (10b):

$$\delta a = \frac{a}{\alpha_a} \ \frac{\delta w}{w} \tag{18}$$

The number of participants concerned by this effect is $(a/\alpha_a)\,G(\Sigma_a, a)\,f(a)\,(\delta w/w)$.

Due to the small tax reform, those employed face a more increasing employment tax rate schedule. The tax reform thus induces a wage reduction Δw_a that substitutes employment probability for ex-post surplus. From (10a), one has

$$\frac{\Delta w_a}{w_a} = \frac{\varepsilon_a}{\eta(w_a)} \ \Delta\eta \tag{19}$$

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 effect can be decomposed into its impact on government's revenue and its impact on the social objective.

Recall from the government's budget constraint (5) that the average additional employment tax per participant of skill a equals the difference between the expected gross wage $w_a\ L(a, w_a)$ and the expected surplus Σ_a. Only the former term is affected by the wage response effect. According to (8), the expected gross wage is changed by

$$\Delta\left(w_a L\left(a, w_a\right)\right) = \left(1 - \eta\left(w_a\right)\right)\ L\left(a, w_a\right)\ \Delta w_a$$

Using (18) and (19), the government's revenue is changed by

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

This term is negative (positive) whenever $\eta(w_a)$ is below (above) 1, that is whenever the wage level is inefficiently low (high), i.e. the employment tax rates are increasing (decreasing) around the wage w_a.

The wage response effect affects the social welfare objective (11) through the change in the average social utility $\ell_a \Phi\left(w_a - T\left(w_a\right)\right) + \left(1 - \ell_a\right) \Phi(b)$ per participant of skill a. Unemployment and the after-tax wage decrease, while the average income $\Sigma_a + b$ of participants of skill a remains unchanged. Each additional employed individual increases social welfare by $\Phi\left(w_a - T\left(w_a\right)\right) - \Phi(b)$. However, the reduction in after-tax wage lowers social welfare by $\left(w_a - T\left(w_a\right) - b\right)\ \Phi'\left(w_a - T\left(w_a\right)\right)\ \Delta\ell_a$.[14] The same average income is shared more equally among participants, so, by concavity of $\Phi(.)$, the net effect is positive. Multiplying this by the number of participants of skill a in the $[a - \delta a, a]$ interval, and taking (8), (18) and (19) into account, the impact of the wage effect on the social objective (expressed in terms of public funds) equals:

$$\frac{\Phi\left(w_a - T\left(w_a\right)\right) - \Phi(b) - \left(w_a - T\left(w_a\right) - b\right)\ \Phi'\left(w_a - T\left(w_a\right)\right)}{\lambda}\ \frac{\varepsilon_a}{\alpha_a}\ a\ h_a\ \left(-\Delta\eta\right)\ \frac{\delta w}{w} \tag{21}$$

Combining (20) and (21), the total *wage response* effect on the interval $[w_a - \delta w, w_a]$ is the left-hand side of (14a) times $\Delta\eta\ \frac{\delta w}{w}$. Adding the mechanical (16) to the participation (17) effects for all skill levels t above a gives $-\Delta\eta\ (\delta w/w)\ Z_a$. At the optimum, the sum of these effects should be nil. This yields (14a) in Proposition 1.

To obtain $Z_{a_0} = 0$ in (14b), consider a tax reform that rises $\log\left(w - T(w) - b\right)$ by a constant amount for all w. This reform does not change $\eta(w)$, so there is no wage response effect. However, it induces a mechanical and a participation effect whose sum is proportional to Z_{a_0}. At the optimum, such a marginal reform should not have a first-order impact on the government's objective, i.e. $Z_{a_0} = 0$.

[14] Using (9), one has $\Delta\left(w_a - T\left(w_a\right)\right) = \left(w_a - T\left(w_a\right) - b\right)\ \eta\left(w_a\right)\ (\Delta w/w)$. Given (8), $\Delta\ell_a = -\eta\left(w_a\right)\ (\Delta w/w) > 0$. The decrease in the after-tax wage decreases social welfare by

$$\ell_a\ \Phi'\left(w_a - T\left(w_a\right)\right)\ \Delta\left(w_a - T\left(w_a\right)\right) = -\left(w_a - T\left(w_a\right) - b\right)\ \Phi'\left(w_a - T\left(w_a\right)\right)\ \Delta\ell_a$$

III.2 Properties of the optimal tax schedule

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, employment probabilities ℓ_a do not react to taxation. Nevertheless, wages 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, 2010). There is however one difference: as in Boone and Bovenberg (2004, 2006), participants face a positive but exogenous probability of being unemployed. In the absence of a wage response effect, the sum of the mechanical (16) and participation effect (17) should be nil for each skill level, so:

$$\frac{T(w_t)+b}{w_t-T(w_t)-b} = \frac{1-\frac{\Phi'(w_t-T(w_t))}{\lambda}}{\pi_t} - \frac{\Xi_t}{w_t-T(w_t)-b} \tag{22}$$

If there were no unemployment, the term Ξ_t would be nil (see (14d) and (3)). Then the participation effect would be proportional to the employment tax $T(w_t) + b$ and Formula (22) would be identical to Expression (4) in Saez (2002). The sign of the employment tax would be given by $1 - \frac{\Phi'(w_t-T(w_t))}{\lambda}$. If for sufficiently low skill levels, the marginal social welfare weight $\Phi'(w_t - T(w_t))/\lambda$ is larger than one, the employment tax $T(w_t) + b$ would be negative. Then, low income workers would reveive higher transfers than those jobless. So an EITC would prevail.

When there is unemployment, the term Ξ_t becomes negative (see (14d)). The sign of the employment tax is then no longer the sign of one minus the marginal social welfare weight. In particular for a level of skill such that the marginal social welfare weight equals one, the employment tax has to be positive. The new term Ξ_t in the optimal tax formula is inherently due to the implication of unemployment on income inequalities. C*eteris paribus* the larger the unemployment, the larger this mechanism, the less likely an EITC is optimal. This conclusion contradicts the expectation of Immervoll *et alii* (2007, p.36) that the presence of unemployment should reinforce the desirability of EITC.

Intuitively, the higher the skill level, the higher the pre-tax and the post-tax wage. Therefore, given the concavity of $\Phi(.)$, $1 - \frac{\Phi'(w_t-T(w_t))}{\lambda}$ is increasing in the skill level. Moreover, empirical evidence suggests that the elasticity of participation π_t is lower for more skilled workers (Juhn *et alii* 1991, Immervoll *et alii* 2007, Meghir and Phillips 2008). Therefore, the first term in the right-hand side of (22) is increasing in the skill level (provided that the skill level is not too small). Consequently, employment tax rate tend to be increasing in the wage. However how Ξ_t evolves with skill is ambiguous. Numerical simulations are thus required. The section (V.2) suggests that optimal employment tax rates are increasing along the wage and skill distributions.

We now reintroduce the wage response effect $\varepsilon_a > 0$. Equation (14a) describes two reasons why it may be optimal to distort wages away from their *laissez faire* value.

First, a tax reform that reduces the wage elasticity of the ex-post surplus reduces wages and increases the probability of being employed without any first-order effect on the expected income of participants $\Sigma_a + b$. A given average income is then shared more equally among employed and unemployed individuals of the same skill level. This mechanism provides a first argument for distorting wages downwards by shifting $\eta(w_a)$ below one.

Second, by making the ex-post surplus function flatter in Figure 2, this reform makes the employment tax rate schedule more increasing. This shift of the tax burden on high skilled workers is socially desirable only if the sum of mechanical and participation effects over skill levels above a is positive. This condition is precisely summarized by $Z_a > 0$.[15] The case where $Z_a > 0$ for all $a \in (a_0, a_1)$ is highly plausible, as shown analytically in the Maximin case (Section IV) and numerically in the general case (Section V). This desire to make employment tax rates increasing in the wage provides a second argument for distorting wages downwards.

At the two extremities of the skill distribution one has $Z_{a_0} = Z_{a_1} = 0$ (see (14b) and (14c)). Only the first argument for distorting wages downwards remains and wages have to be distorted downwards. So, $\eta(w_{a_0}) < 1$ and $\eta(w_{a_1}) < 1$ in the absence of bunching, which are the only analytical results with a general objective function. Moreover, in the plausible case where $Z_a > 0$ everywhere at the optimum, the two abovementioned arguments imply that $\eta(w_a)$ has to be below 1 at all wages. Therefore, the optimal employment tax rate has to be increasing and so does the average tax rate $T(w)/w$. In addition, the employment tax has to be positive for top-income earners, otherwise the budget constraint of the government (5) could not clear. The property $\eta(w_{a_1}) < 1$ then implies that the marginal tax rate has to be positive at the top. Notice that this result is achieved even with a bounded skill distribution.

The result of positive marginal tax rate at the top does not hold in the model with intensive margin and a bounded distribution (Sadka 1976, Seade 1977). As average tax rates have to be higher than marginal tax rates for top income earners, optimal average tax rates cannot be increasing along the whole income distribution in such a model. The same conclusion holds in a model with labor supply responses along both margins. Diamond (1998) and Saez (2001) need to approximate the skill distribution by an unbounded skill distribution to get positive asymptotic marginal tax rates. Such an approximation can then also generate increasing average tax rates (Hindriks *et alii*, 2006 and Boadway and Jacquet, 2008).

IV The Maximin case

We now provide an analytical characterization of the optimum by restricting to the case of a Maximin social objective. The mechanical, participation and wage response effects then matter

[15] Given (14b) and (14c), Z_a is also equal to the aggregation of mechanical and participation effects of a tax reform that generates a uniform increase in ex-post surpluses over all skill levels t below a.

only through their impact on government's revenue to finance the assistance benefit b. The optimality conditions of Proposition 1 become (see Appendix E):

Proposition 2 *For any skill level* $a \in [a_0, a_1]$*, the Maximin-optimal tax schedule verifies:*

$$\frac{1-\eta(w_a)}{\eta(w_a)} \frac{\varepsilon_a}{\alpha_a} w_a \; a \; h_a = Z_a \qquad \textit{and} \qquad Z_{a_0} = 0 \tag{23}$$

$$Z_a = \int_a^{a_1} [w_t - (1+\pi_t)(T(w_t)+b)] \; h_t \; dt, \tag{24}$$

IV.1 No wage response effect

We again provisionally assume the absence of wage response effects ($\varepsilon_a = 0$), so employment probabilities ℓ_a are exogenous. Given the simplifications under the Maximin, Equation (22) becomes:

$$\frac{T(w_a)+b}{w_a} = \frac{1}{1+\pi_a} \tag{25}$$

The optimal employment tax rate decreases with the elasticity of participation. In accordance with Saez (2002) and Choné and Laroque (2005), the employment tax rate is positive, i.e. there is no EITC.

IV.2 Constant elasticity of participation

We now investigate under which condition the tax schedule described by Equation (25) is optimal when wages are responsive to taxation ($\varepsilon_a > 0$). This tax schedule induces that the sum Z_a of the mechanical and of the participation effects for all skill levels above a equals 0 (See Equation 24). Therefore, the wage response effect has to be nil. So, according to (23), the elasticity η of the ex-post surplus has to equal 1 everywhere and the employment tax rate has to be constant. This is consistent with (25) 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 at a value π and consider the tax policy defined by an employment tax $T(w) + b$ equal to $w/(1+\pi)$ for all wage levels. In this case, the mechanical (16) and participation (17) effects sum to 0 at each skill level. Moreover, from (9), 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 2.

IV.3 Decreasing elasticity of participation

The assumption of a constant elasticity of participation is not plausible. Empirical evidence suggests a decreasing profile with a (see Juhn *et alii*, 1991, Immervoll *et alii*, 2007 or Meghir and Phillips, 2008). Of course, the profile of π_a at the optimum can be different from the one observed in the current economy. Still, the two following examples suggest that the slopes of π_a in the current economy and at the optimum might have the same signs.

The first example specifies

$$G(a,\Sigma) = \min\left[A(a)\ \Sigma^{\pi_a}, 1\right] \qquad \text{with} \qquad A(a) > 0 \text{ and } \pi_a > 0 \tag{26}$$

Then, provided that $\Sigma_a \leq (A(a))^{-1/\pi_a}$, the participation rates remain within $(0,1)$ and the elasticity of participation is exogenous and equals π_a.

The second example is based on the following assumptions. Firstly, the value of non market activities χ is distributed independently of the skill level a. Secondly, for a given wage level, the employment probability increases in the skill level (see the example given in Section II.5). Hence $\dot{\Sigma}_a > 0$ from (13). Then, along any allocation, the elasticity of participation decreases with skill whenever $\Sigma g(\Sigma)/G(\Sigma)$ is a decreasing function of Σ. This is for instance the case if χ follows the exponential distribution $G(\Sigma) = 1 - \exp(-\sigma_1\Sigma)$ or the Pareto distribution $G(\Sigma) = 1 - \sigma_0\Sigma^{-\sigma_1}$ with $\sigma_0, \sigma_1 > 0$.

We then get (See Appendix F):

Proposition 3 *If everywhere along the Maximin optimum one has $\dot{\pi}_a < 0$, then*

***i*)** *The average tax rate $T(w)/w$ is an increasing function of the wage. Marginal tax rates are higher than employment tax rates, which are positive (no EITC).*

ii) $w_a < w_a^{LF}$ *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* $L(a_1, w_{a_1}) = L\left(a, w_{a_1}^{LF}\right)$.

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

Figure 3 depicts the employment tax rate as a function of the level of skill. In the absence of wage responses, optimal employment tax rates are decreasing in the elasticity of participation (see (25)). The dashed increasing curve $1/(1+\pi_a)$ in Figure 3 illustrates this profile in the current context where $\dot{\pi}_a < 0$. Employment tax rates increase then more than proportionally in the wage w_a, so $\eta(w_a) < 1$. When wage are responsive to taxation, this policy distorts wages downwards. Hence, to limit wage distortions, the optimal tax function depicted by the solid curve varies less with the wage than the optimal curve without wage responses. It has also to be as close as possible to the optimal curve without wage response to limit the departures from the optimal trade-off between the participation and mechanical effects. Employment tax rates are therefore increasing, but less so compared to the case without wage response. This implies that wages and unemployment rates are lower than their efficient levels, except at the two extremes of the distribution because of the transversality conditions (Results *ii*)).

Figure 3 here

A comparison of the magnitude of the expected surpluses at the optimum Σ_a and at the *laissez-faire* $w_a^{\mathrm{LF}} L\left(a, w_a^{\mathrm{LF}}\right)$ determines whether and to what extent participation rates are distorted. Let us write Σ_a as $w_a L\left(a, w_a\right)\ \left(1-\left(\left(T\left(w_a\right)+b\right)/w_a\right)\right)$. First, since wages are distorted, $w_a L\left(a, w_a\right)$ is lower at the optimum than at the *laissez-faire*. Second, as illustrated by Figure 3 the employment tax rate reaches its lowest value at the bottom. Moreover, the employment tax rate at the bottom is larger at the optimum with wage response effect than at the optimum without wage response effect. These two features hold because the profile of optimal employment tax rates is flatter than at the optimum without wage responses. Finally, along the optimum without wage response effect, the employment tax rate at the bottom is positive because the government has a Maximin objective (see 25). Hence, employment tax rates are everywhere positive along the optimum with wage response effect. So, participation rates are distorted downwards (Result *iii*)). Results *ii*) and *iii*) imply that the net effect on aggregate employment is ambiguous.

Since the employment tax rate is increasing in wages, the average tax rate is increasing in wages as well and the marginal tax rate is higher than the employment tax rate. Finally, since $\left(T(w)+b\right)/w$ is positive everywhere, the marginal tax rate is also positive everywhere, including at the boundaries of the skill distribution (Result *i)* of the Proposition).

The literature on optimal income taxation with an extensive margin only is, to the best of our knowledge, silent on the shape of optimal average tax rates. However, formula (25) applies under Maximin. Therefore, there is no EITC and the average employment tax rate is increasing when the participation elasticity is decreasing along the skill distribution. Consequently, average tax rates are increasing and marginal tax rates are above employment tax rates, thereby they are positive. Proposition 3 shows that these analytical results also hold in our more general model that does account for the empirical evidence that gross incomes are responsive to marginal tax rates and for the existence of unemployment.

V Simulations

To illustrate how our optimal tax formulas could be used for applied purposes and because of the lack of broad analytical properties in the general utilitarian case, we now turn to a simulation exercise.

V.1 Calibration

To avoid the complexity of interrelated participation decisions within families, we only consider single adults in the US.[16] In addition to the function $\Phi(.)$ for the Bergson-Samuelson criterion, we need to specify the density function of skills $f(a)$, the cumulated density function of non-market activities $G(a,\chi)$ and the labor demand function $L(.,.)$. We take $\log L(a,w) = B(a) - \varepsilon \left(\frac{w}{c\ a}\right)^{\frac{1}{\varepsilon}}$. Under this specification, the first-order condition (8) for the wage-setting program implies:

$$w_a = c\ a\ \ (\eta(w_a))^{\varepsilon} \tag{27}$$

Next, we roughly approximate the tax system that is applied to single adults without children by a linear function $T(w) = \tau\ w + \tau_0$ with $\tau = 25\%$ and $\tau_0 = -3000$ dollars per year. 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 in the current economy. 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. So we take $b = -\tau_0 = 3000$. 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 (see 10a). This elasticity also equals the *compensated* elasticity of wage with respect to $1 - T'$.[17] Following Saez *et alii* (2010), estimates of the latter elasticity would lie between 0.12 and 0.4. We take a conservative value $\varepsilon = 0.1$ in the benchmark calibration and conduct a sensitivity analysis. We set c to 2/3, so that in the current economy, total wage income represents two third of total production. Finally, we use (27) and the distribution of weekly earnings of the Current Population Survey of May 2007 to approximate the skill distribution among employed workers. Reexpressing variables in annual terms, the range of skills is [\$3,900; \$218,400].[18] 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 dotted line in Figure 4.

We adopt the simplest specification of the cumulative distribution of non-market activities, namely (26). So, the elasticity of participation varies exogenously with the level of skill. 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

[16] 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.

[17] For any *compensated* change in marginal tax rates (i.e. with $\Delta T = 0$), one has $\Delta\eta = \frac{\Delta(1-T')}{1-T'}\frac{1-T'}{1-(T+b)/w} = \frac{\Delta(1-T')}{1-T'} \cdot \eta$.

[18] 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.

$\pi_a = (\pi_{a_0} - \pi_{a_1}) \left(\frac{a_1 - a}{a_1 - a_0}\right)^3 + \pi_{a_1}$. 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 the scale parameters $B(a)$ of the labor demand function and $A(a)$ of (26) to generate some realistic properties of skill-specific unemployment and participation rates along our approximation of the current economy. The profile of unemployment (resp. participation) rates in the current economy is calibrated by a decreasing (increasing) function of a:[19]

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

In our approximation of the current economy, the mean unemployment rate is 5.1%, 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 (solid line) $f(a)$,[20] the distribution of skill among participants in the current economy $\mathcal{G}_a f(a)$ (dashed line) and the distribution of skills among employed individuals (dotted line) $\ell_a \mathcal{G}_a f(a)$.

Figure 4 here

We compute the level of exogenous public expenditures E from the government's budget constraint (5). 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 finance a public good that generates social utility that is considered as a perfect substitute to private consumption under this specification.

V.2 The results in the benchmark

Figure 5 (Maximin optimum) here

The Maximin case is depicted in Figure 5 (Maximin optimum). Confirming Figure 3, in comparison with the case where wages are exogenously fixed, employment tax rates vary less when wages are responsive to taxation. As emphasized in Proposition 3, marginal tax rates are always higher than employment tax rates, except at both extremes of the distribution. This illustrates to what extent the wage-cum labor demand is distorted. Under the Maximin, redistribution takes the form of a Negative Income Tax (NIT) in the following sense: An annual assistance benefit of $14,198 is taxed away at a high, and in this case nearly constant, marginal

[19] By approximating a by the educational attainment, our unemployment and participation rates for the whole population, for a_0 and a_1 are in line with the CPS data in June 2007.

[20] $f(a)$ is deducted from the above approximation of $\ell_a \mathcal{G}_a f(a)$.and the calibrations of functions ℓ_a and $\mathcal{G}_a$.

tax rate close to 80%. This figure appears to be very high. However, it is not unusual under Maximin (see e.g. Saez 2001).

Figure 6 (Bergson Samuelson optimum) here

Under the Bergson Samuelson objective, the employment tax rate without wage response but with unemployment is based on Equation (22) that extends Equation (4) of Saez (2002) to the impact of unemployment on inequalities among participants of a given skill. The latter effect does not preclude the presence of an EITC at the bottom as illustrated by the dashed curve of Figure 6. The employment tax rate is upward sloping. When wages are responsive to taxation, this tax profile (solid curve) distorts wages downwards. This is detrimental to efficiency. However, the optimal tax profile does not only trade off efficiency and distortions to participation but takes also the inequalities between the employed and the unemployed of the same skill into account (see Equation (21) and the corresponding term in (14a)). Distorting wages downwards reduces unemployment and therefore these inequalities. In Figure 6, to limit distortions along the wage-cum-labor demand margin, the optimal employment tax rate (solid) curve with wage responses varies less than the employment tax rate (dashed) curve without wage responses. The gap between these two curves is noteworthy given the low elasticity ε of wage in the calibration. In the presence of wage responses, marginal tax rates (dotted curve) are everywhere larger than employment tax rates. At the extremes of the distribution, because of the transversality conditions, only the wage-cum labor demand effect on welfare (21) exists, which explains why wages are distorted downwards. Marginal tax rates are thus higher than the employment tax rates at the extremes, which is in contrast with the Maximin case. At the top, the marginal tax rate equals 40% and the employment tax rate equals 34%.

Let us now briefly compare the Bergson Samuelson and the Maximin optima when wages respond to taxation. The well-being of workers, in particular the low-paid ones, enters the scene under the more general objective. The employment tax rate at the bottom becomes negative and the tax schedule has now the basic features of an EITC. 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 $T(w_{a_0}) = -\$3,167$. At the bottom of the wage distribution, the marginal tax rate is negative as well and then it sharply increases.

The hump-shaped profile of the marginal tax rates contrasts with the U-shaped profile found by Saez (2001) in a"Mirrlees-type model" whatever the social objective criterion. 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. Our numerical simulations are thus in line with Saez (2002) on this point and confirm an important difference with HLPV who treated the participation decisions in a crude way.

Figure 7 (Unemployment rates) here

To illustrate Part *ii*) of Proposition 3, let us compare the actual profile of unemployment rates and the optimal ones under the Maximin and Bergson-Samuelson criteria (Figure 7). The actual unemployment rate turns out to be too high from a Maximin perspective. However, the actual unemployment rate is optimal at both ends of the skill distribution. This result is in contrast to HLPV where the unemployment distortions were maximal for the lowest participating type. Again, the result of no-distortion at the bottom is the consequence of the additional heterogeneity introduced in the present paper. Here, the highest unemployment distortions appear in the middle of the skill distribution. This explains why the optimal unemployment rate is U-shaped. From the general utilitarian viewpoint, the unemployment rate should even decrease further with respect to the actual one, confirming the importance of the wage-cum labor demand effect on welfare (21). This impact is stronger for the less skilled because their welfare is highly valued by the government. Their unemployment rate is thus strongly distorted downwards.

Figure 8 (Participation) here

As an illustration of Part *iii*) of Proposition 3, Figure 8 shows that a Maximin government would accept a sharp decline in participation rates. In order to finance an important assistance benefit, the government has to set high employment tax rates. Under the more general utilitarian objective, optimal participation rates are higher for low-skilled workers and lower for high-skilled workers. In order to reduce the unemployment rates of the low-skilled, the government distorts highly their wages downwards. However, it compensates them by fixing a negative employment tax. Their expected surplus is thus increased, thereby pushing up their participation rates. 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.

V.3 Sensitivity analysis

Figure 9 (Sensitivity ε) here

The wage response effects are reinforced if the sensitivity of wages to taxation is raised from $\varepsilon = 0.1$ to $\varepsilon = 0.2$. When ε becomes higher, the employment tax rate schedule at the Maximin becomes less increasing in the wage, so as to prevent too important distortions along the wage-cum-labor demand margin. The marginal tax schedule becomes closer to a linear one. The simulations displayed in Figure 9 show that this also arises along the Bergson-Samuelson optimum.

Figure 10 (Sensitivity lower π) here

Next, we decrease by a constant amount of 0.05 all the shape of π_a. In the Maximin case without wage response, Equation (25) implies that the government would choose higher employment tax rates as participation responds less, so the dashed curve in Figure 3 is shifted upwards. Consequently, in the presence of wage responses, the solid curve shifts upwards too. Hence the Maximin optimum implements higher participation tax rates and therefore higher marginal tax rates. Figure 10 quantifies this mechanism. Once again, the Bergson-Samuelson optimum is affected in a similar way compared to the Maximin optimum.

Figure 11 (Sensitivity steeper π) here

Finally, we change the elasticities of participation so that the profile of π_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 11, it is again convenient to come back to Figure 3. In the Maximin optimum without wage response, the government wishes to implement a more increasing employment tax rate curve, 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 9).

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 wage-cum labor demand effect on welfare (21). 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 but are higher for lower skill workers. Average tax rates are always increasing at the Bergson-Samuelson optimum.

VI Conclusion

We derive an optimal tax formula in a model where wages, unemployment rates and participation decisions are endogenous. In general, the optimal income tax schedule is characterized by increasing average tax rates. This shifts the burden of taxation away from low skilled workers whose utility are the most valued and whose participation decisions are the most responsive. Due to the reactions of wages and hence of labor demand to taxation, the progressive tax schedule shifts wages below their *laissez faire* value, so labor demand is distorted upwards. Moreover, we obtain increasing average tax rates and a positive marginal tax rate at the top of the skill distribution, even when the skill distribution is bounded.

Most of these properties are obtained numerically under a general objective. We are able to derive analytically some of these properties under a Maximin objective when the elasticity of participation is decreasing in the skill level. We then show that the qualitative profile of the average employment tax rate turns out to be the same whether wages and unemployment are exogenous or endogenous. However, the theoretical setting with exogenous wages contradicts the empirical evidence that gross earnings are responsive to taxation. This paper has instead developed a theoretical framework compatible with this evidence and with the existence of unemployment. Moreover, according to our simulation exercise, there is a substantial différence between the optimal employment tax rates in the two settings, even with a conservative calibration of the wage response effect.

Our paper also contributes to the literature of self-selection models with random participation. The seminal paper of Rochet and Stole (2002) is developed in a context of nonlinear pricing. Our paper proposes a new method for signing the distortions in a different framework. This method considers in a first step the tax function that minimizes distortions along the extensive margin by removing wage responses. In a second step, we show that the "full" optimum with wage responses is the result of a trade-off between this tax function and a linear function for which wages and unemployment are not distorted. Jacquet *et alii* (2010) use a similar method in a model with perfectly competitive labor markets and labor supply responses along both the intensive and the extensive margins. However, they obtain results only on the sign of marginal tax rates.

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 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. Third, in the real world, labor supply decisions are typically taken at the household level, not at the individual one (see Kleven *et alii* 2009). Finally, it would be interesting to address the issue of top managers' income using an alternative bargaining framework without unemployment.

A Competitive Search Equilibrium

In this Appendix, we derive the Competitive Search Equilibrium (CSE) when search is directed by skill and wage levels. There is one potential submarket for each skill level a *and* each wage level w. In the CSE setting, at stage 2 of our timing of events, firms and individuals of skill a not only decide whether or not to enter the labor market, but also on which (single) submarket (a, w) they enter. This defines a non-cooperative game between firms (deciding how many vacancy to create on each submarket) and individuals (deciding whether or not to participate, and if they do, on which submarket to enter) whose Nash equilibrium(a) define(s) the so-called CSE(s) (Moen 1997).

Let $U_{a,w}$, $V_{a,w}$ and $\theta_{a,w} = V_{a,w}/U_{a,w}$ denote the mass of job-seekers, vacancies and tightness on submarket (a, w). Firms (respectively job-seekers) being atomistic, they take $\theta_{a,w}$ as given, thereby their probabilities of matching $q(\theta_{a,w})$ (resp. $\theta_{a,w} q(\theta_{a,w})$) as well.

We first show that the free-entry condition $q(\theta_{a,w})(a-w) - \kappa(a) = 0$ is verified on each submarket. If $q(\theta_{a,w})(a-w) - \kappa(a) > 0$ (resp. $<$), firms find profitable to create more (less) jobs of skill a at the posted wage w. This increases (decreases) the mass of vacancies $V_{a,w}$ on this submarket and reduces (increases) the probability of filling each of these vacancy $q(\theta_{a,w})$. Following (2), job-seekers find a job on submarket (a, w) with probability $\theta_{a,w} q(\theta_{a,w}) = L(a, w)$.

Second, let (a, w) and (a, w') such that $L(a, w)(w - T(w) - b) > L(a, w')(w' - T(w') - b)$. An individual of type (a, χ) gets $b + \chi$ if she does not participate, expects $L(a, w)(w - T(w)) + (1 - L(a, w)) b$ if she enters submarket (a, w) and expects $L(a, w')(w' - T(w') - b) + b$ is she enters submarket (a, w'). Entering submarket (a, w') is therefore not profitable for any individual of skill a. So, an individual of type (a, χ) finds profitable to enter submarket (a, w) only if w maximizes $L(a, \omega)(\omega - T(\omega) - b)$ in ω and only if $\chi < L(a, w)(w - T(w) - b)$. In particular, if individuals (a, χ) and (a, χ') participate, they choose to enter the same submarket. Hence, they are employed at the same probability and at the same wage.

B Labor demand as a structural primitive

Lemma 1 *Let $L(a, w_a)$ be a twice-continuously differentiable labor demand function such that $\partial L/\partial w_a < 0$ and, for each a, there exists $\overline{w}_a$ such that $L(a, w_a) > 0$ if and only if $w_a < \overline{w}_a$. Then, there exists a unique matching technology $H(a, ., .)$ and a vacancy cost function $\kappa(a)$ that generates $L(., .)$ through (2). Moreover, $H(., ., .)$ verifies all the assumptions we made about the matching function.*

From the free-entry condition, tightness on the labor market is positive if and only if $a - w > \kappa(a)$. Hence, one must have $\kappa(a) = a - \overline{w}_a$. Let $\xi = \kappa(a) / (a - w_a)$, so that $w_a = a - (\kappa(a)/\xi)$. As w_a increases from $-\infty$ to $\overline{w}_a$, ξ increases from 0 to 1. Then, (2) can be rewritten $L\left(a, a - \frac{\kappa(a)}{\xi}\right) = \xi \; q^{-1}(a, \xi)$ and so one must have

$$q^{-1}(a, \xi) = \frac{L\left(a, a - \frac{\kappa(a)}{\xi}\right)}{\xi}$$

for each ξ in $(0, 1]$. For each a, since $L\left(a, a - \frac{\kappa(a)}{\xi}\right)$ is bounded, one has $\lim_{\xi \mapsto 0} q^{-1}(a, \xi) = +\infty$. Hence, $q^{-1}(a, .)$ defines a decreasing function from $[0, 1]$ onto $[0, +\infty]$. Inverting the last function, one retrieves $q(a, \theta)$ defined over $[a_0, a_1] \times \mathbb{R}_+$ onto $[0, 1]$. Then, the matching function $H(a, V, U) \equiv V \; q(a, V/U)$ is well defined over $[a_0, a_1] \times \mathbb{R}_+^2$ and exhibits constant returns to scale. Since $q(., .)$ is bounded above by 1, one obtains that $H(a, 0, U) = H(a, V, 0) = 0$ and $H(a, V, U) < V$. Moreover, since the elasticity in ξ of $L\left(a, a - \frac{\kappa(a)}{\xi}\right)$ is negative, the elasticity with respect to ξ of q^{-1} is lower than -1. Consequently, the elasticity of q with respect to θ must lie in between -1 and 0. Therefore $H(a, ., .)$ is increasing in both arguments. To finally show that $H(a, V, U) < U$, let us define $\theta = V/U$ and w by $(a - w) q(a, \theta) = \kappa(a)$. Then, one has that $H(a, V, U)/U = \theta q(a, \theta) = L(a, w) < 1$.

C Elasticities of wages

In this appendix, we compute how the equilibrium wage that solves (4) is modified when i) the slope of the ex-post surplus function is uniformly increased by an amount $\Delta\eta$ ii) when the skill level increases by δa. Let us rewrite the first-order condition (8) as $\mathcal{W}(w_a, a, 0) = 0$, where:

$$\mathcal{W}(w, a, \tilde{\eta}) \equiv \frac{\partial \log L}{\partial \log w}(a, w) + \eta(w) + \tilde{\eta} \tag{28}$$

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

$$\mathcal{W}'_w(w_a, a, \tilde{\eta}) = \frac{\partial^2 \log L(a, w_a)}{\partial w\ \partial \log w} + \eta'(w_a) \tag{29}$$

This second-order condition states that in Figure 1, the ex-post surplus function is either locally concave (i.e. $\eta'(w_a) < 0$) or less convex than the indifference expected surplus curves.[21] Consider now how the equilibrium wage w_a is influenced by small changes in η or in the skill level a. Whenever the second-order condition of (4) is a strict inequality and the maximum of (4) is globally unique, we can apply the implicit function theorem on $\mathcal{W}(w_a, a, \tilde{\eta}) = 0$. We then obtain the elasticity ε_a (resp. α_a) of the equilibrium wage w_a with respect to a small local change $\Delta\eta$ in the elasticity of the ex-post surplus (resp. in the skill level a):

$$\varepsilon_a = -\frac{\eta(w_a)}{w_a}\frac{\mathcal{W}'_{\Delta\eta}}{\mathcal{W}'_w} \qquad \alpha_a = -\frac{a}{w_a}\frac{\mathcal{W}'_a}{\mathcal{W}'_w}$$

which gives (10a) (resp. (10b)) given (29) and that $\mathcal{W}'_{\Delta\eta} = 1$ (resp. $\mathcal{W}'_a = w_a \frac{\partial^2 \log L}{\partial a\ \partial w}(a, w_a)$). 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 $\mathcal{W}'_w$. The positive signs of ε_a and α_a follow from the strict second-order condition $\mathcal{W}'_w < 0$ and from (7).

D Proof of Proposition 1

Let $\sigma_a = \log \Sigma_a$. We use optimal control by considering σ_a as the state variable and w_a as the control. Incentive constraint (13) implies $\dot{\sigma}_a = \partial \log L(a, w)/\partial a$. Let λ be the Lagrange multiplier associated to the budget constraint (5) and q the co-state variable. The Hamiltonian writes:

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

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

We assume that a maximum exists where w_a is a continuous function of a.[22] Then, there exists a continuously differentiable function $a \mapsto q_a$, such that the following first-order conditions are verified:

$$0 = \frac{1}{\lambda}\frac{\partial \mathcal{H}}{\partial w} = \left[\frac{\partial L(a,w_a)}{\partial w_a} \frac{\Phi\left(\frac{\Sigma_a}{L(a,w_a)}+b\right) - \Phi(b) - \frac{\Sigma_a}{L(a,w_a)}\Phi'\left(\frac{\Sigma_a}{L(a,w_a)}+b\right)}{\lambda} + \frac{\partial\left(w_a\ L(a,w_a)\right)}{\partial w_a}\right] G(a,\Sigma_a)\ f(a) + \frac{q_a}{\lambda}\frac{\partial^2 \log L}{\partial a \partial w}(a,w_a) \tag{30a}$$

$$-\frac{\dot{q}_a}{\lambda} = \frac{1}{\lambda}\frac{\partial \mathcal{H}}{\partial \sigma} = \left\{\left(\frac{\Phi'\left(\frac{\Sigma_a}{L(a,w_a)}+b\right)}{\lambda} - 1\right) G(a,\Sigma_a) + \left[w_a\ L(a,w_a) - \Sigma_a\right]\ g(a,\Sigma_a) \right. \tag{30b}$$

$$\left. \frac{L(a,w_a)\,\Phi\left(\frac{\Sigma_a}{L(a,w_a)}+b\right) + (1-L(a,w_a))\,\Phi(b) - \Phi(b+\Sigma_a)}{\lambda} g(a,\Sigma_a)\right\} \Sigma_a\ f(a)$$

together with the transversality conditions $q_{a_0} = q_{a_1} = 0$. Let us define $Z_a = -q_a/\lambda$, Using $q_{a_1} = 0$, one has $Z_a = \int_a^{a_1} \frac{\dot{q}_t}{\lambda}\, dt$. Hence, (30b) with (1) gives (14c). The transversality condition $q_{a_0} = 0$ gives $Z_{a_0} = 0$ in (14b). From (8), one has

$$\frac{\partial L}{\partial w}(a,w_a) = -\eta(w_a)\frac{L(a,w_a)}{w_a} \qquad \frac{\partial\left(w\ L(a,w)\right)}{\partial w}(a,w_a) = (1-\eta(w_a))\ L(a,w_a) \tag{31}$$

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

$$\frac{\partial^2 \log L}{\partial a \partial w}(a,w_a) = \frac{\alpha_a}{\varepsilon_a}\ \frac{\eta(w_a)}{a}\ \frac{1}{w_a} \tag{32}$$

Introducing these last expressions into (30a) gives the first equality in (14a).

E Proof of Proposition 2

The government's problem under Maximin is:

$$\max_{w_a,\sigma_a} \int_{a_0}^{a_1} \left[L(a,w_a)\ w_a - \exp\sigma_a\right]\ G(a,\exp\sigma_a)\ f(a)\, da \qquad s.t: \dot{\sigma}_a = \frac{\partial \log L}{\partial a}(a,w_a)$$

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

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

Equations (30a) and (30b) are simplified:

$$0 = \frac{\partial \mathcal{H}}{\partial w} = \frac{\partial\left(w_a\ L(a,w_a)\right)}{\partial w_a} G(a,\Sigma_a)\ f(a) + Z_a\ \frac{\partial^2 \log L}{\partial a \partial w}(a,w_a) \tag{33a}$$

$$-\dot{Z}_a = \frac{\partial \mathcal{H}}{\partial \sigma} = \left\{-G(a,\Sigma_a) + \left[w_a\ L(a,w_a) - \Sigma_a\right]\ g(a,\Sigma_a)\right\}\ \Sigma_a\ f(a) \tag{33b}$$

[22] 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 follow this first-order approach only to save on space. We check 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.

These two conditions with the transversality conditions $q_{a_0} = q_{a_1} = 0$, (31) and (32) give (23) and (24).

F Proof of Proposition 3

We first show that Z is positive on (a_0, a_1). From (24), one has

$$\dot{Z}_a = \left(\frac{\pi_a}{1+\pi_a} - \frac{x_a}{w_a}\right)(1+\pi_a)\ w_a\ h_a \tag{34}$$

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 $Z_a \leq 0$ 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 (34) this implies:
$$\frac{\pi_{\underline{a}}}{1+\pi_{\underline{a}}} \leq \frac{x_{\underline{a}}}{w_{\underline{a}}}$$
- Since $Z_a \leq 0$, one has from (23) that $\eta(w_a) \geq 1$ for all $a \in [\underline{a}, \overline{a}]$. Given (9), 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 (34) 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.[23]

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 (34)

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

However since $Z_{a_2} = 0$, one has $\eta(w_{a_2}) = 1$ from (23). Hence, from (9) 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 (34), 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, $(\pi_{a_4}/(1+\pi_{a_4})) \simeq$

[23] In the presence of bunching, (30b) and the transversality conditions $q_{a_0} = q_{a_1} = 0$ continue to hold, while only the integration of (30a) over the skill interval where bunching occurs holds. Therefore, the present proof is still valid since on a bunching interval, w_a and x_a being constant x_a/w_a also remains constant.

$(\pi_{a_2}/(1+\pi_{a_2})) + \left(\dot{\pi}_{a_2}/(1+\pi_{a_2})^2\right)(a_4 - a_2)$ at a first-order approximation. Hence, since $\dot{\pi}_{a_2} < 0$, 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 (23), one has $\eta(w_a) < 1$ for any $a \in (a_0, a_1)$, which has different implications. First, for any $a \in (a_0, a_1)$, one has $\partial \log L/\partial w(a, w_a) > -1$ from (8). Moreover, at the *laissez-faire*, $\partial \log L/\partial w\left(a, w_a^{\text{LF}}\right) = -1$ from (8) and (9). Hence, from (6) $w_a < w_a^{\text{LF}}$ which means that optimal wages are distorted downwards. Furthermore, since $\partial L/\partial w(a, .) < 0$, one has $1 - L(a, w_a) < 1 - L\left(a, w_a^{\text{LF}}\right)$ and unemployment rates are distorted downwards. Finally, $Z_{a_0} = Z_{a_1} = 0$ induces $w_{a_0} = w_{a_0}^{\text{LF}}$, $L(a_0, w_{a_0}) = L\left(a_0, w_{a_0}^{\text{LF}}\right)$, $w_{a_1} = w_{a_1}^{\text{LF}}$ and $L(a_1, w_{a_1}) = L\left(a_1, w_{a_1}^{\text{LF}}\right)$. This proves *ii*).

Second, as $\eta(w_a) < 1$, 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, $x_{a_0}/w_{a_0} \leq \pi_{a_0}/(1+\pi_{a_0}) < 1$. Hence for all a, $x_a < w_a$ and participation rates are distorted downwards. This proves *iii*).

Lastly, as $x < w$ for all w, employment tax rates $(T(w) + b)/w$ are always positive. Moreover, it is nondecreasing since $\eta(w) < 1$. So, the average tax rate $T(w)/w$ is increasing in wage w. Finally (9) and $\eta(w) \leq 1$ induces $T'(w) \geq (T(w) + b)/w$, so marginal tax rate are positive everywhere. This proves *i*).

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] Bovenberg, L (2006), *Tax Policy and Labor Market Performance*, in J. Agell and P. Birch Sorensen (eds.), Tax Policy and Labor Market Performance, MIT Press

[3] 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.

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

[5] Boadway, R. and L. Jacquet, 2008, Optimal Marginal and Average Income Taxation under Maximin, *Journal of Economic Theory*, **143**(1), 425-41.

[6] Blomquist, S. and H. Selin, 2010, Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates, *Journal of Public Economics*, **94**(11-12), 878-889.

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

[8] 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.

[9] 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.

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

[11] Choné, P. and G. Laroque, 2010, Optimal Taxation in the extensive model, *Journal of Economic Theory*, Forthcoming.

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

[13] Diamond, P., 1982, Wage determination and efficiency in search equilibrium, *Review of Economic Studies*, **49**, 217-227.

[14] Diamond, P., 1998, Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates, *American Economic Review*, **88**(1), 83-95.

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

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

[17] 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.

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

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

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

[21] Hindriks, J., E. Lehmann and A. Parmentier, 2006, Optimal Income Taxation and the shape of average tax rates, *Economics Bulletin*, **8**(2), pp. 1-6.

[22] Hungerbühler, M., E. Lehmann, 2009, On the Optimality of a Minimum Wage: New Insights from Optimal Tax Theory, *Journal of Public Economics*, **93**, 464-481.

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

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

[25] Jacquet, L., E. Lehmann and B. Van der Linden, 2010, Optimal Redistributive Taxation with both Extensive and Intensive Responses, *IRES Discussion Paper* 2010-33.

[26] Juhn, C., Murphy K. M.,Topel R. H., Yellen J. L. and Baily M. N., 1991, Why has the Natural Rate of Unemployment Increased over Time?, *Brookings Papers on Economic Activity*, 75-142.

[27] Kleven, H. J., C. T. Kreiner, E. Saez, 2009, The Optimal Income Taxation of Couples, *Econometrica*, **77**(2), 537-60.

[28] 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.

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

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

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

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

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

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

[35] 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.

[36] Piketty, T., 1997, La redistribution fiscale face au chômage, *Revue Française d'Economie*, **12**(1), 157-201

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

[38] 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.

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

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

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

[42] 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.

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

[44] Sadka, E., 1976, On income distribution, incentive effects and optimal income taxation, *Review of Economic Studies*, **43**(2), 261-67

[45] Saez, E., 2000, Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses, *NBER Working Paper* 7708.

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

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

[48] Saez, E., J. Slemrod and S. Giertz, 2010, The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review, *Journal of Economic Literature*, forthcoming.

[49] Seade, J., 1977, On the shape of optimal tax schedules, *Journal of Public Economics*, **7**(2), 203-35.

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

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

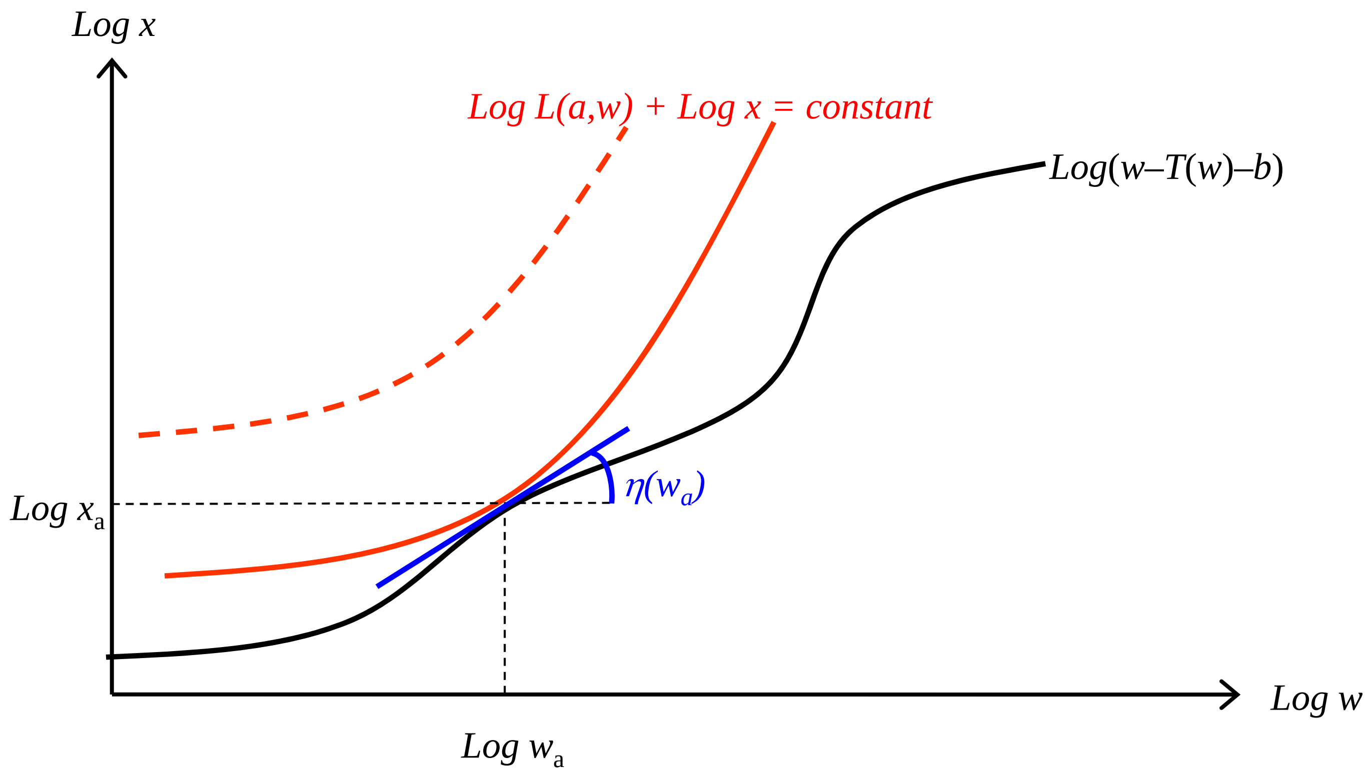

Figure 1: The wage determination for a match of type *a*.

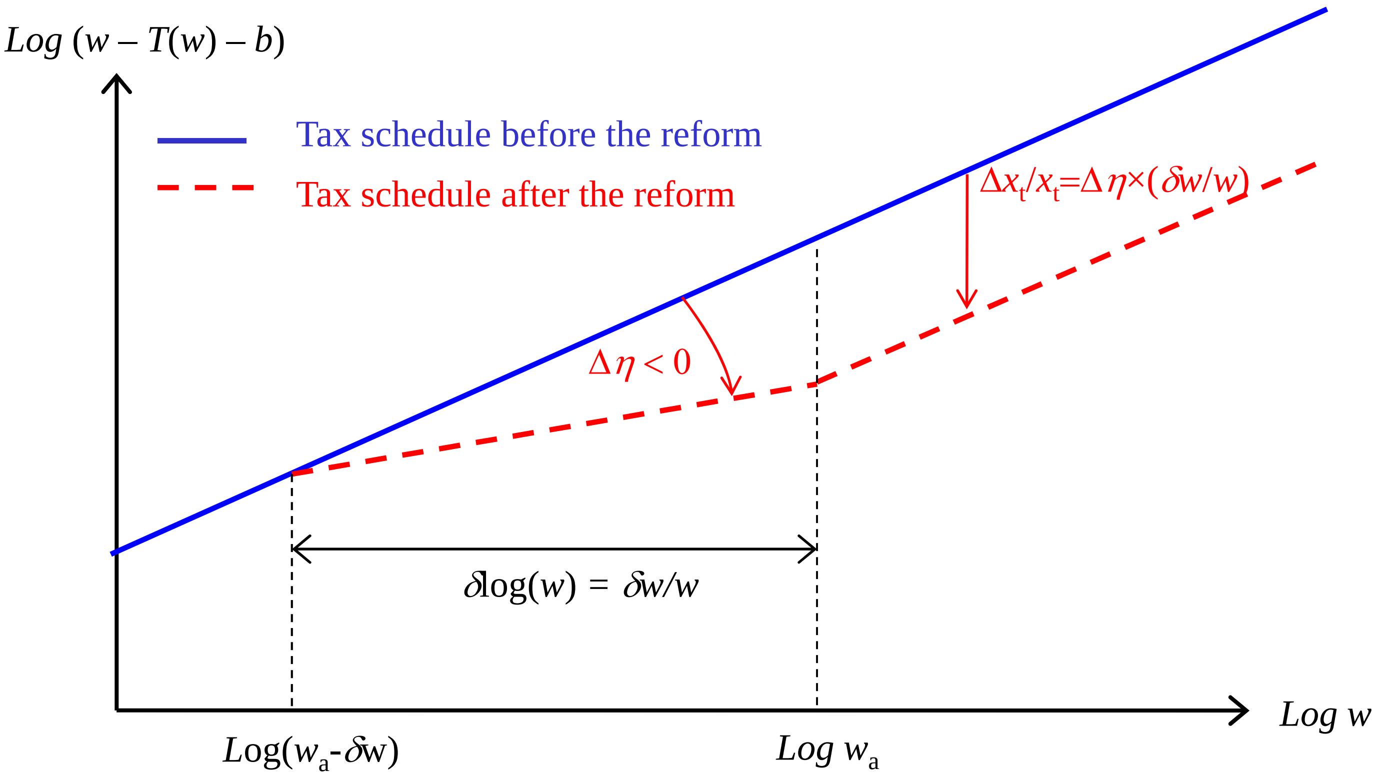

Figure 2: The Tax reform.

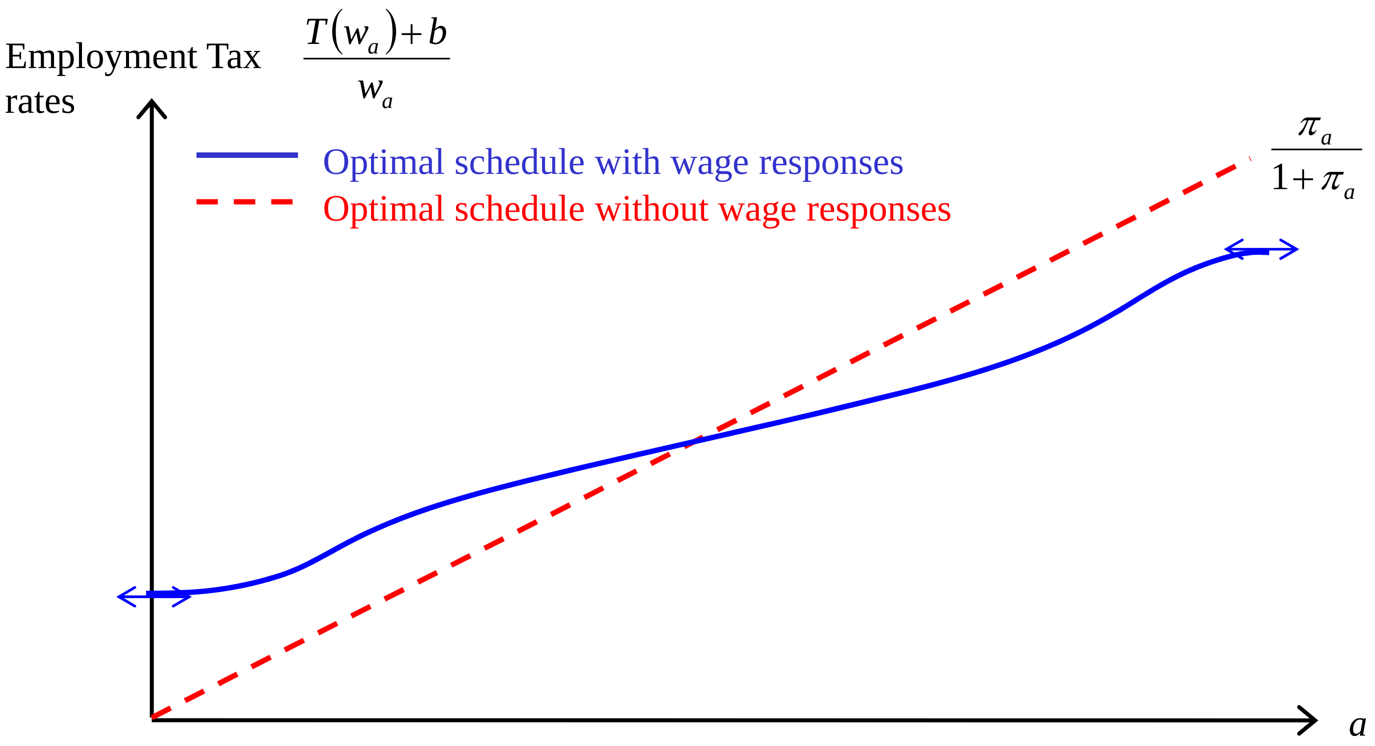

Figure 3: Intuition for Proposition 3

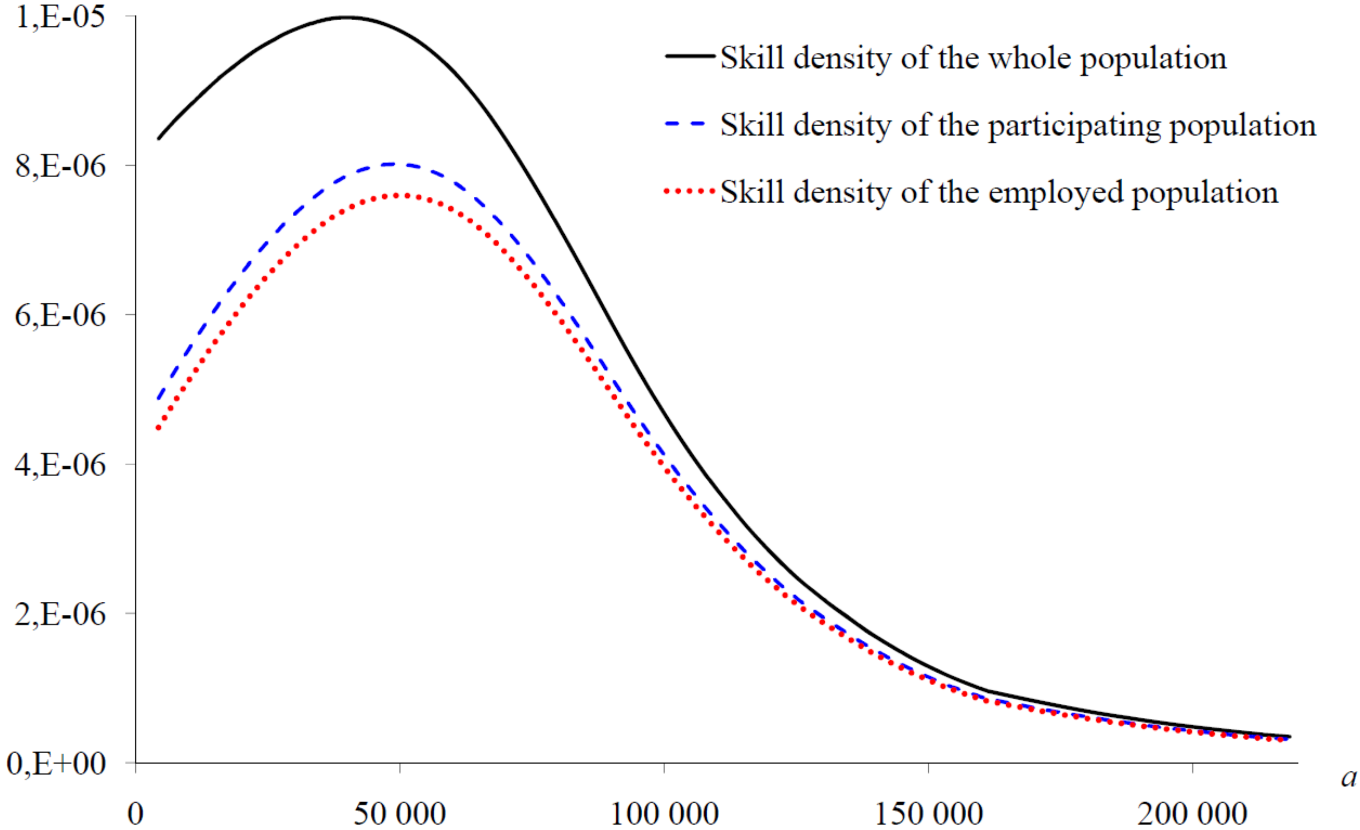

Figure 4: Densities in the current economy

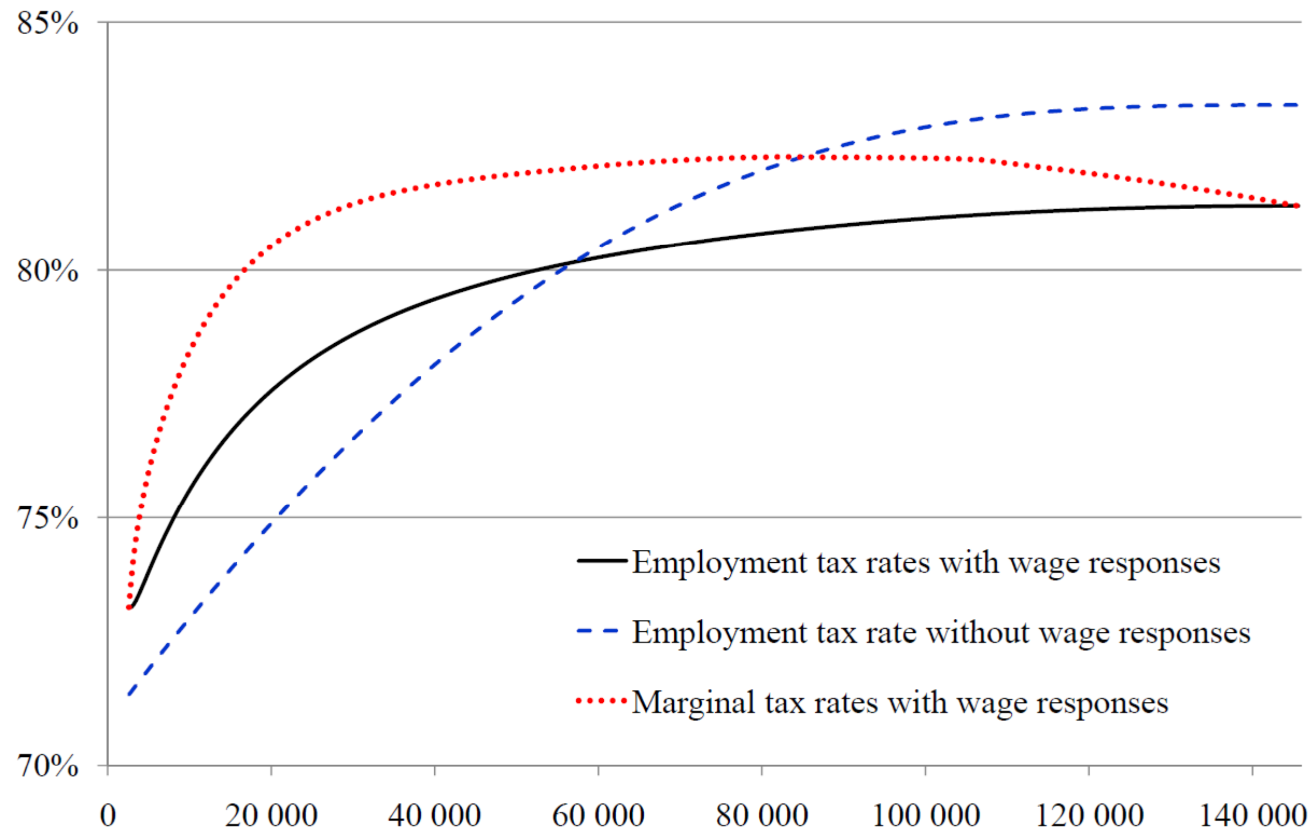

Figure 5: The Optimal Tax schedules under Maximin

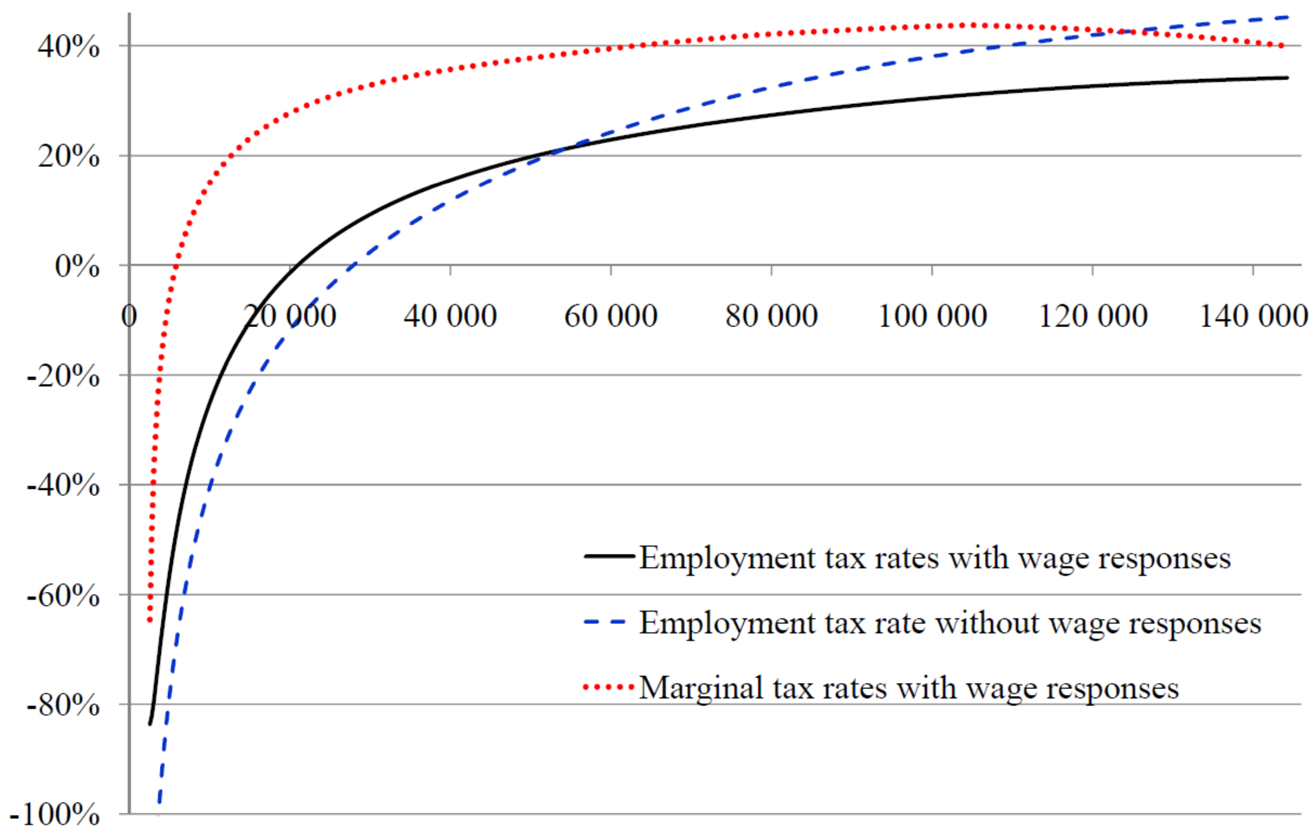

Figure 6: The Optimal Tax schedules under Bergson-Samuelson

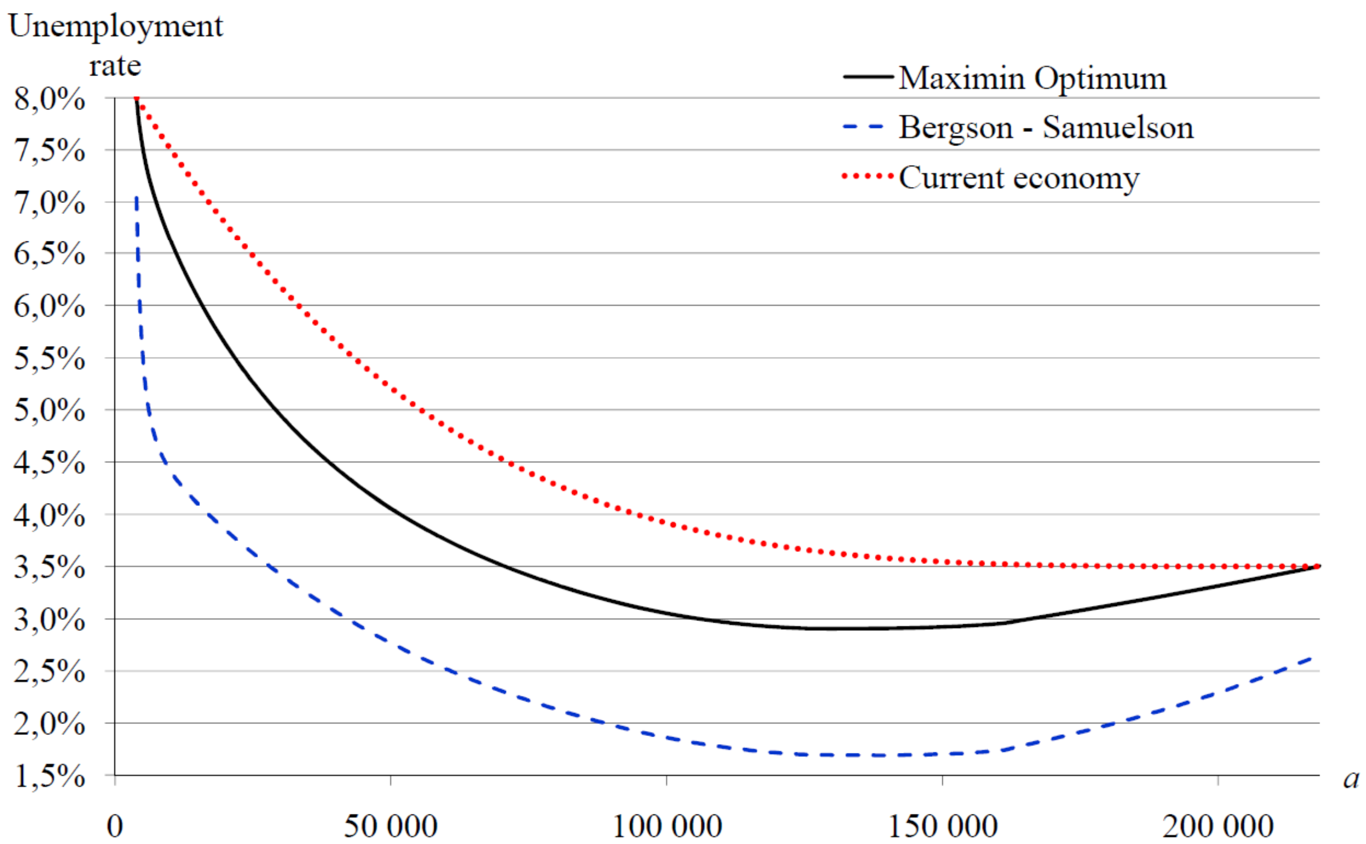

Figure 7: Unemployment rates

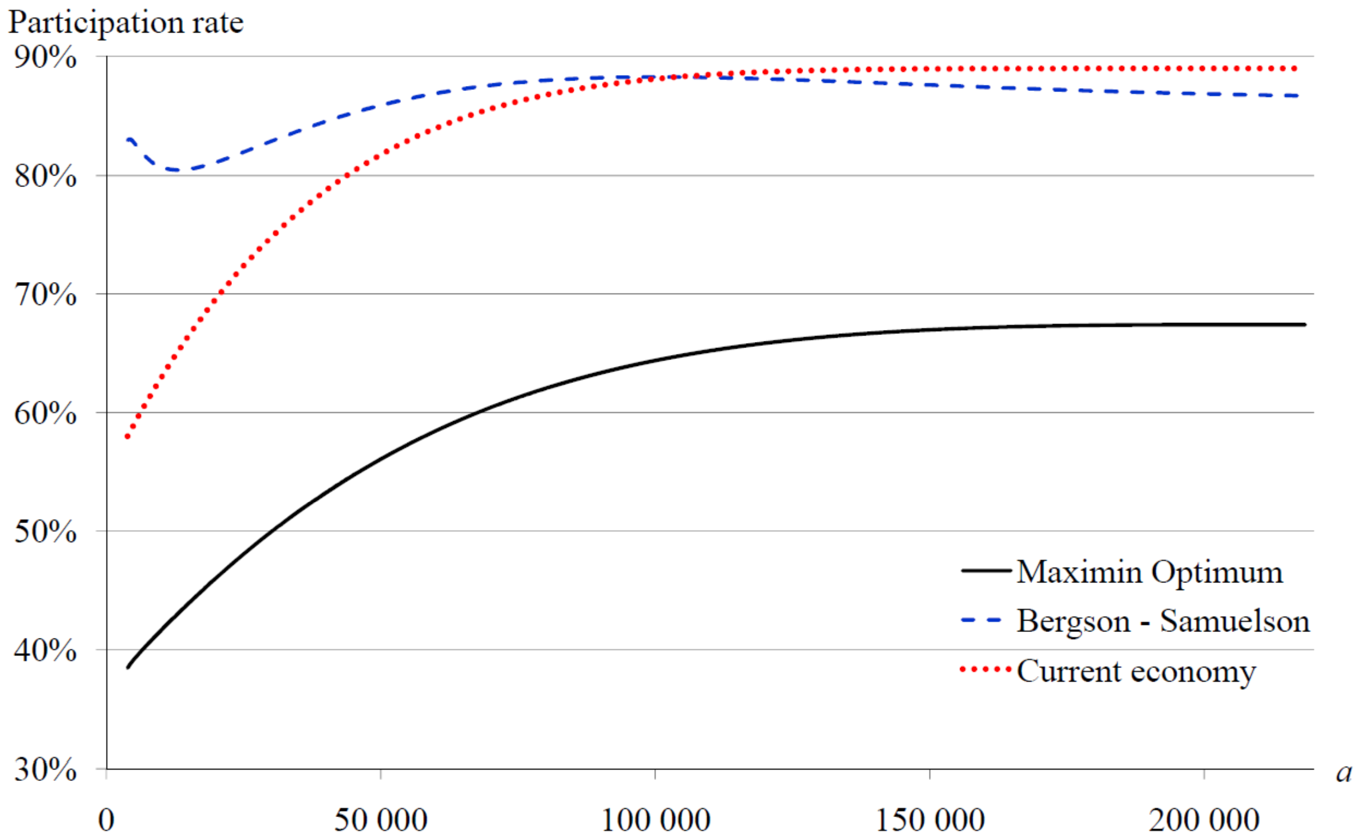

Figure 8: Participation rates

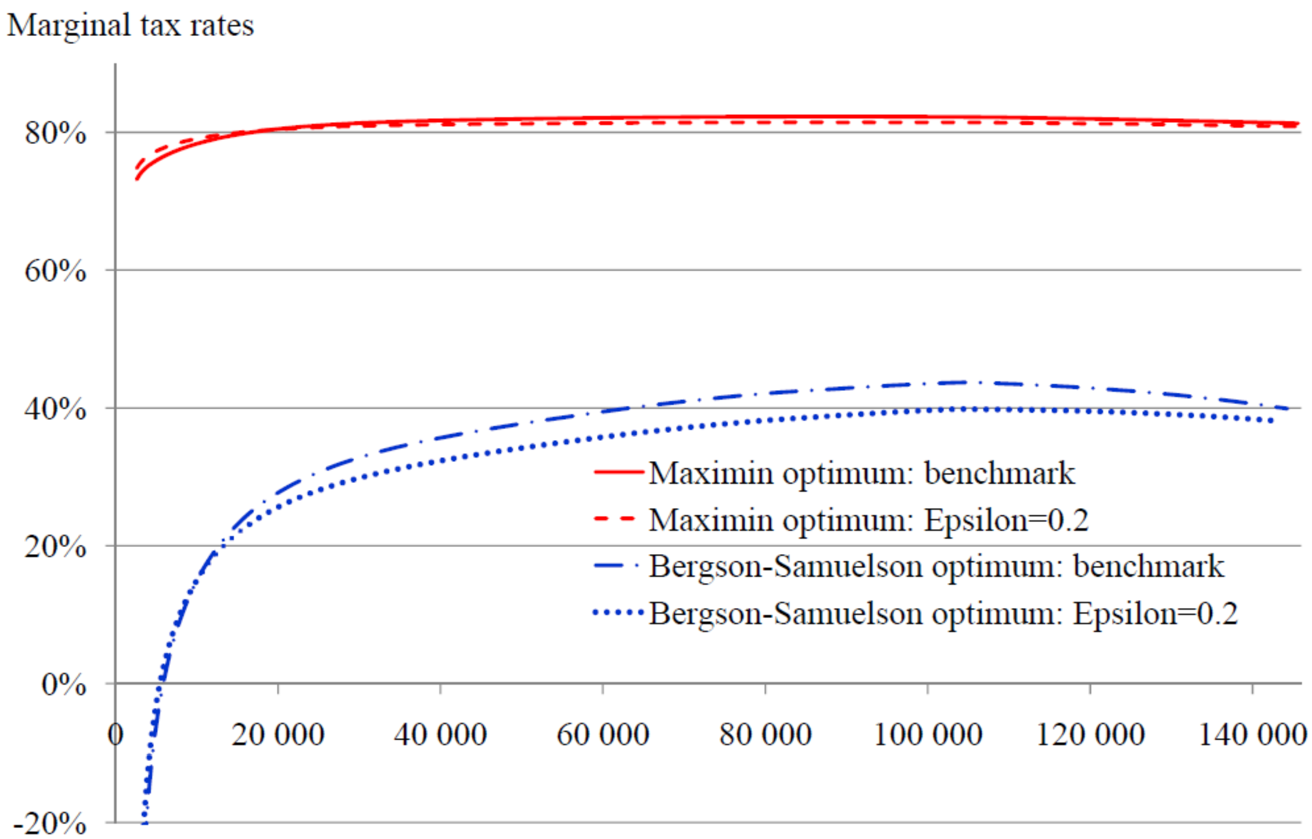

Figure 9: The Optimal Tax schedules when ε_a varies

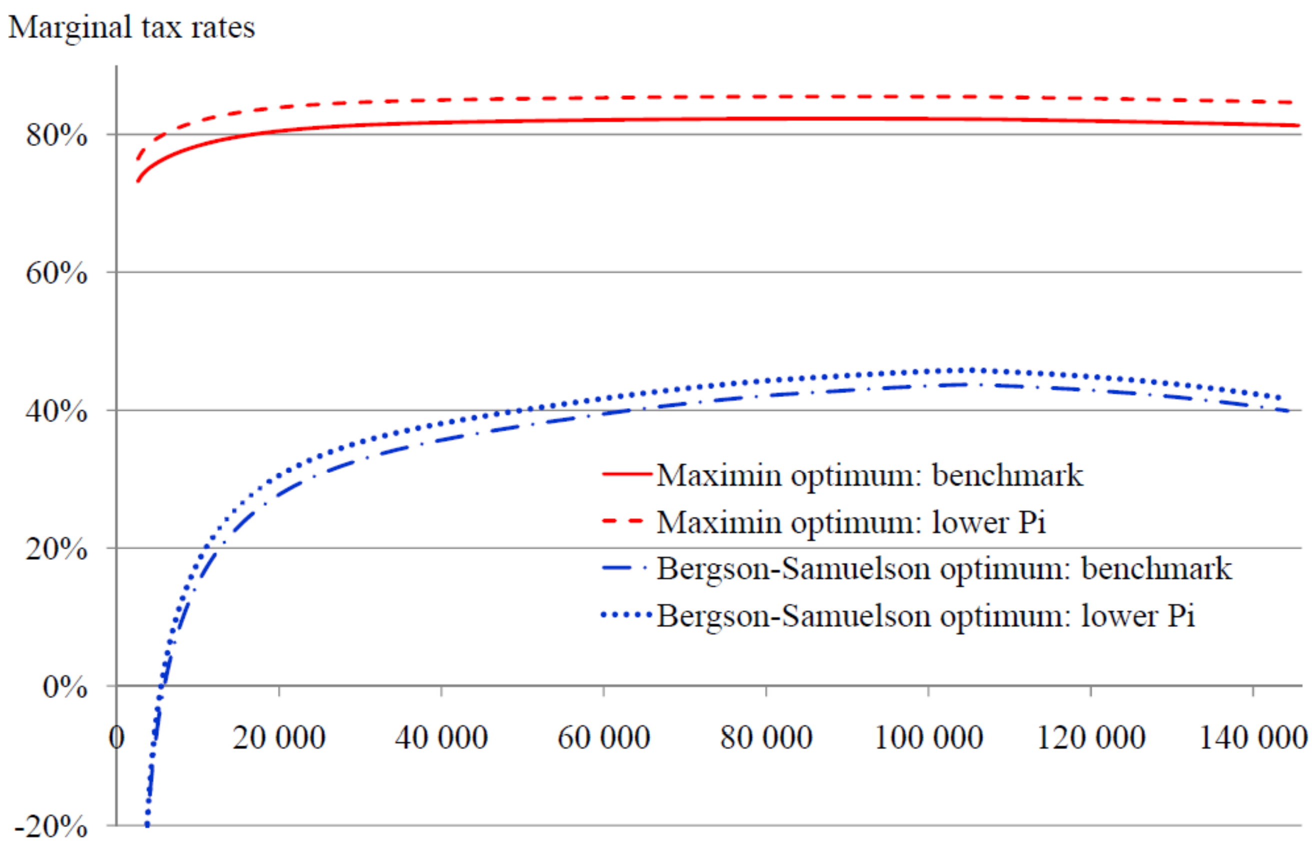

Figure 10: The Optimal Tax schedules when π_a is lower

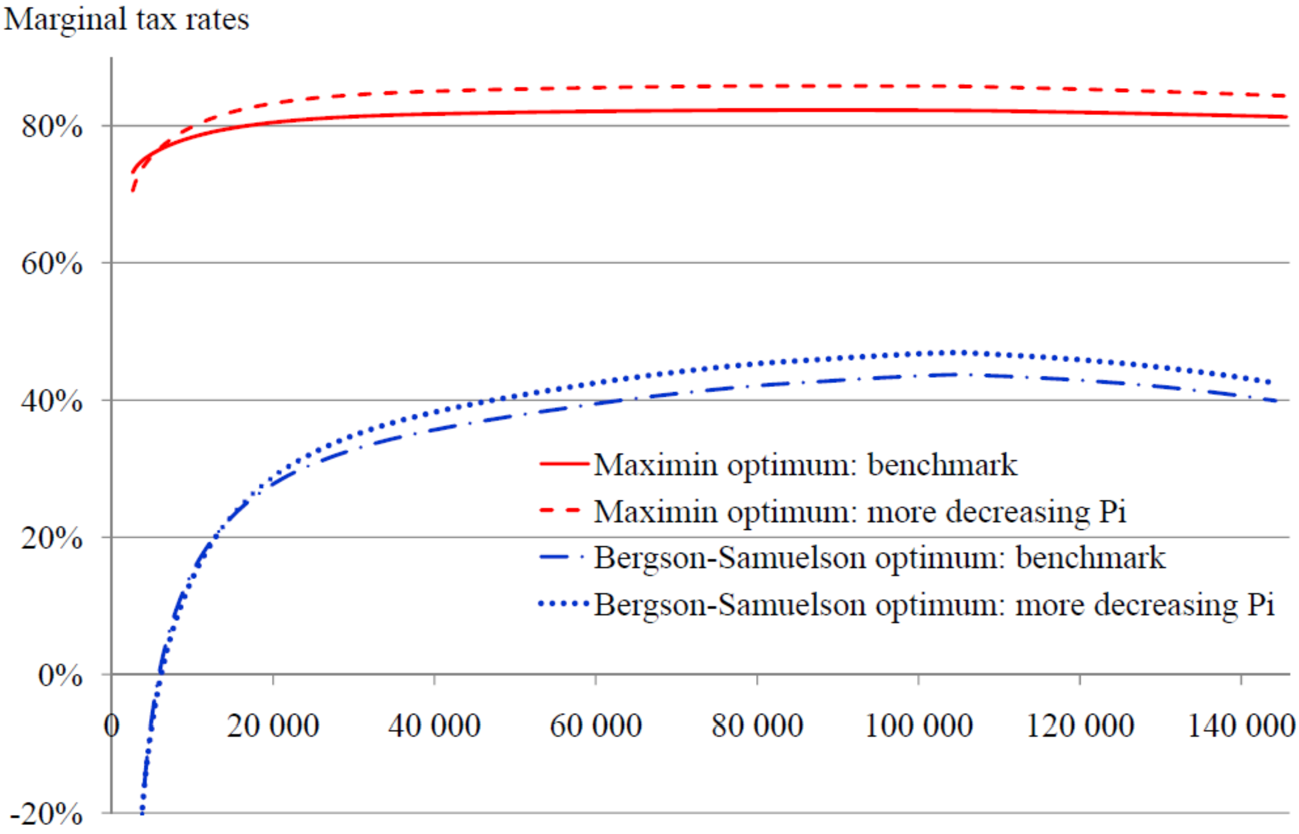

Figure 11: Steeper profile of π_a