DISCUSSION PAPER SERIES

IZA DP No. 2957

On the Optimality of a Minimum Wage: New Insights from Optimal Tax Theory

Mathias Hungerbühler
Etienne Lehmann

July 2007

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

On the Optimality of a Minimum Wage: New Insights from Optimal Tax Theory

Mathias Hungerbühler
University of Namur

Etienne Lehmann
CREST and IZA

Discussion Paper No. 2957
July 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 2957
July 2007

ABSTRACT

On the Optimality of a Minimum Wage: New Insights from Optimal Tax Theory*

We build a theoretical model to study whether a minimum wage can be welfare-improving if it is implemented in conjunction with an optimized nonlinear income tax. We consider this issue in a framework where search frictions on the labor market generate unemployment. Workers differ in productivity. The government does not observe workers' productivity but only their wages. Hence, the redistributive policy solves an adverse selection problem. We show that a minimum wage is optimal if the bargaining power of the workers is relatively low. However, if the government controls the bargaining power, then it is preferable to set a sufficiently high bargaining power.

JEL Classification: D86, H21, H23, J64, J68

Keywords: optimal taxation, minimum wage, search-matching unemployment, bunching, wage bargaining

Corresponding author:

Etienne Lehmann
CREST-INSEE
Laboratoire de Macroéconomie
Timbre J360
15 Boulevard Gabriel Péri
92 245 Malakoff Cedex
France
E-mail: etienne.lehmann@ensae.fr

* We thank participants of seminars in ERMES - Paris 2, CREST and CORE and at the PET 2007 meeting in Nashville. In particular, we thank Pierre Cahuc and Guy Laroque for fruitful comments. A part of this research was realized while Hungerbühler was visiting ERMES-Paris 2 and another part while Lehmann was visiting the Department of Economics of the Université Catholique de Louvain.

I Introduction

The minimum wage is one of the most controversial economic policies. On the ground of equity considerations, a minimum wage aims to play a redistributive role by increasing income for the least skilled workers. One might however counter-argue that redistributive taxation can achieve this goal in a more efficient way. On the ground of efficiency considerations, the minimum wage is often blamed for its adverse effects on labor demand. This is true as long as labor markets are perfectly competitive. However, the minimum wage can be helpful to correct for noncompetitive wage setting (see e.g. Robinson 1933 or Stigler 1946). In our opinion, it is therefore necessary to include optimal taxation and labor market imperfections when one considers the normative issue of the minimum wage. In this paper, we propose a theoretical model to study whether a minimum wage can be welfare-improving if it is implemented in conjunction with an optimized nonlinear income tax *à la* Mirrlees (1971). To integrate explicitly the unemployment effects of a minimum wage, we consider this issue in a framework where search frictions on the labor market *à la* Mortensen and Pissarides (1999) generate endogenous "involuntary" unemployment (i.e. some workers are willing to work at the equilibrium wage, but fail to find a job).

In our model, workers differ with respect to productivity. They decide whether to search for a job, while firms search for workers to fill their job vacancies. If a worker and a firm are paired, they Nash-bargain the wage. The government observes wages, but not productivity. Hence, it faces an adverse selection problem. Since the productivity of a match is revealed through the wage, and since the negotiated wage is the one that maximizes the Nash Product, incentive constraints depend only on Nash Products. However, and contrary to the standard model in contract theory, workers' participation constraints depend on a different variable than incentive constraints. In our case, the participation depends on the workers' expected incomes while searching. We show that in such a context, bunching at the bottom of the wage distribution appears at the (second-best) optimum if the workers' bargaining power is relatively low. In our model where wages are negotiated, we interpret such a bunching as an argument in favor of a binding minimum wage.

This first result holds under the assumption that the government cannot influence the workers' bargaining power. We further show that, if the government can control the bargaining power, then it is desirable to increase a relatively low bargaining power, in which case our previous argument for the minimum wage disappears.

The impact of a minimum wage in the case of a monopsony in the labor market has been studied among others by Robinson (1933) and Stigler (1946). Firms do not face competition on the labor market and thus distort wages downwards, thereby reducing labor supply and eventually employment. Therefore, a binding minimum wage can restore efficiency and increase employment (along the labor supply), provided its level is not above the equilibrium wage in a perfectly competitive labor market. Our contribution differs in several ways from Stigler's. First, we integrate taxation into the framework. As already noted by Stigler, tax measures might achieve the same result as the minimum wage, and possibly even in a more efficient way. We however show that a minimum wage is – under certain conditions – even optimal when tax measures are available. Second, while Stigler only analyses the efficiency problem, we analyze the impact of the minimum wage in a framework where the government wants to redistribute from low-income to high-income individuals, and thus faces an efficiency-equity trade-off[1]. Third, we introduce "involuntary" unemployment. In Stigler's simple monopsony model, every individual who is willing to work at the (monopsony) market wage is able to find a job. We assume search and matching frictions, which imply that not every individual who is willing to work at the equilibrium wage can find a job. Some people fail to find one and become "involuntarily" unemployed.

Many papers have already investigated whether the minimum wage can be useful in combination with an optimized redistributive tax (see Franklin Allen 1982, Stephen Allen 1987, Guesnerie and Roberts 1987, Drèze and Gollier 1993, Marceau and Boadway 1994, Boadway and Cuff 2001, among others). In particular, Stephen Allen (1987) considers a model with two types of imperfectly substitutable workers and endogenous hours of work *à la* Stiglitz (1982). He shows that a minimum wage is never optimal in conjunction with the optimized nonlinear income tax, because a rise in the minimum wage strengthens the relevant incentive constraint. We abandon this framework of two types of imperfect-substitute labors and we instead base our labor demand margin on matching frictions. In our model, search frictions drive a wedge between (marginal) productivity and the wage, while productivity and hours of work are exogenous. Furthermore, we consider a model with a continuum of productivity which is – as we will show – more relevant for considering the issue of the form of the optimal redistributive allocation.

Finally, we extend the model of optimal redistributive taxation in a search equilibrium framework developed by Hungerbühler, Lehmann, Parmentier and Van der Linden (2006, henceforth HLPV). Contrary to HLPV, we allow the bargaining power to be lower than the one prescribed by the so-called Hosios (1990) condition. In such a case, a rise in wage for a given level of the Nash Product increases workers' expected income. As we show, this effect opens the room for a welfare-improving role of the minimum wage.

The paper is organized as follows. In Section II, we present the basic model, including incentive and participation constraints. Next, in Section III, we solve the model for a given bargaining power. We show that a minimum wage is optimal if the bargaining power is sufficiently low. Section IV considers the case when the government can control

[1]For the case of the (non-)desirability of minimum wages in the context of a monopsony and redistributive taxation, see Cahuc and Laroque (2007).

the bargaining power. Finally, we conclude in Section V.

II The model

Our model follows the framework built in HLPV to deal with the optimal tax problem of Mirrlees (1971) within the equilibrium unemployment theory of Mortensen and Pissarides (1999) and Pissarides (2000). To keep things as simple as possible, we consider a static setting which has become standard in the models of search equilibrium with taxation (see also Boone and Bovenberg, 2002). There is a mass 1 of risk-neutral individuals. They can be either employed, unemployed or out of the labor force. Their preferences on consumption and leisure are assumed quasi-linear in consumption. Individuals differ with respect to their productivity $a \in [a_0, a_1]$ with $0 \leq a_0 < a_1 < \infty$ according to the continuous density function $f(a)$ and the cumulative density function $F(a)$[2]. These functions are common knowledge, while the productivity is private information to the worker. The timing of the model is as follows:

1. First, the government commits to its policy. The policy consists in a tax schedule $T(.)$, a welfare benefit b and a minimum wage $\underline{w}$. Since the government cannot observe the individuals' productivities but only their wages, this tax schedule is conditional on the gross wage w only. The government cannot observe whether a non-employed individual has searched for a job but not found one (hence, she is unemployed) or simply not searched for a job at all (hence, she is out of the labor force). Consequently, all non-employed individuals get the same welfare benefit b (whatever their productivity and their participation decision). In section IV only, we assume that the government has an additional policy instrument: It can control the workers' bargaining power β.

2. In a second step, individuals and firms decide to participate in the labor market. The individuals have the binary choice to invest all their leisure time in search for a job or to stay out of the labor force and enjoy utility from leisure, $d > 0$, and the welfare benefit b. They do decide to participate in the market if their expected income is above the utility of staying out of the labor force. Firms can decide to open job vacancies. To do so, they first have to invest capital to build up the workstation. This capital investment is assumed irreversible and productivity-specific, i.e. to be able to hire a worker of ability a, the firm has to buy κ_a units of capital[3]. Conversely,

[2] Our results still hold if $a_1 \to \infty$ as long as $\int_{a_0}^{a_1} a dF(a)$ is finite.

[3] A firm might consist of a large number of different jobs. The government observes in a firm's account only the sum of the total investment costs κ_a and total profits $a - w$, but cannot disentangle the costs and profits of each job. For this reason, we assume that no tax scheemes based on investment costs κ_a or profits $a - w_a$ are available to the government.

an individual of ability a can only work at a firm that has made the appropriate investment in equipment, i.e. that has invested the exact amount κ_a of capital. We assume that[4]

$$\frac{\dot{\kappa}_a}{\kappa_a} \leq \frac{1}{a} \tag{1}$$

Firms decide to enter into the labor market as long as their expected profit is positive.

3. Next, individuals and firms match on their skill-specific labor markets. To model the search frictions, we use a standard matching function. The number of matches is a function of the number of individuals searching for a job and of firms searching for a worker. Capital investment and labor markets are productivity-specific. In this sense, we assume directed search[5]. The outcome of this matching process determines the unemployment rate in our model. During this matching process, the firm observes the productivity of the worker. Unmatched firms loose their invested capital.

4. The worker and the firm bargain the gross wage that is paid to the worker.

5. Production and transfers occur.

II.1 The matching process, participation decisions and employment

The matching function $H(U_a, V_a)$ in labor market a gives the number (density) H_a of employed individuals of type a as a function of the number U_a of workers searching for a job, and the number V_a of job vacancies. This matching function is assumed to represent heterogeneities and frictions that we do not model explicitly. It is usually assumed that the matching function $H(.,.)$ is increasing in both its arguments, concave and homogeneous of degree 1. Empirical studies have found that a Cobb-Douglas approximation of the matching function fits the data well (Petrongolo and Pissarides, 2001). We therefore assume that the matching function is given by

$$H(U_a, V_a) = A\,(U_a)^{\gamma}\,(V_a)^{1-\gamma} \qquad \text{with} \qquad \gamma \in (0,1) \tag{2}$$

where $A > 0$ is a scale parameter of the matching function. If individuals of type a search for a job, then $U_a = f(a)$, otherwise $U_a = 0$. The probability for a firm to hire a

[4]This assumption ensures that in an economy without government intervention, the unemployment rate among the high-skilled individuals is not higher than the unemployment among the low-skilled individuals.

[5]While the government does not observe a worker's productivity, we assume that firms observe it after a match. Hence, if a type-a worker searches on a type-$t \neq a$ labor market and finds a job, the match becomes unproductive and the worker is fired. Since search is costly for workers, a type-a worker has no incentive to search on another labor market.

worker H_a/V_a is a decreasing function of the number of vacancies, while the probability $L_a = H_a/U_a$ for a searching worker to find a job increases in the number of vacancies.

A firm has to invest κ_a units of capital to open a type-a vacancy. When the firm finds a suitable worker for this vacancy, this match produces a units of goods. Note that the investment takes place before the matching to the worker. Since there are matching frictions on the labor market, some firms do not find a worker. In that case, the loss of the firm is equal to the investment cost κ_a. If the firm finds a worker for its vacancy, then they have to bargain on the gross wage w_a and the firm's profit writes $a - w_a - \kappa_a$. Since the probability that a firm finds a worker of type a is equal to H_a/V_a, the expected profit from posting a vacancy can be written as $(H_a/V_a)(a - w_a) - \kappa_a$. Firms enter the market as long as these expected profits are positive. Hence, this expected profit is nil at equilibrium, and therefore

$$\frac{H_a}{V_a}(a - w_a) = \kappa_a \tag{3}$$

Given (2), one has $H_a/V_a = A \cdot (V_a/U_a)^{-\gamma}$. The free-entry condition (3) therefore determines the ratio V_a/U_a, and thereby the probability $L_a = H_a/U_a = A \cdot (V_a/U_a)^{1-\gamma}$ for a type-a searching worker to find a job, which is given by:

$$L_a = A^{\frac{1}{\gamma}} \left(\frac{a - w_a}{\kappa_a}\right)^{\frac{1-\gamma}{\gamma}} \tag{4}$$

If the wage w_a decreases, then the firm's surplus $a - w_a$ increases relative to the vacancy cost κ_a, and hence, firms create more vacancies and the probability of finding a job increases for type-a searching workers. An additional firm that enters the market increases employment and therefore *gross* output. But it also increases the resources spent for capital investments. The impact of an additional vacancy on *net* output (net of investment costs) is then ambiguous and depends on the number or vacant jobs that are already on the labor market. If the wage is sufficiently low, the firm has incentives to enter the market, even though this might not be optimal from a social point of view, because too many resources might then be spent for capital investments. The gross output generated by workers of type a is equal to $aH_a = aL_a f(a)$. Let

$$Y_a \equiv aL_a - \frac{V_a \kappa_a}{f(a)}$$

Output net of investment costs on the type-a labor market can then be written as $aH_a - V_a\kappa_a = Y_a f(a)$. Multiplying Equation (3) by V_a, one gets

$$Y_a(w_a) \equiv w_a L_a = A^{\frac{1}{\gamma}} \left(\frac{a - w_a}{\kappa_a}\right)^{\frac{1-\gamma}{\gamma}} w_a \tag{5}$$

Because of free entry, firms' expected profits equal 0, so their total surplus $(a - w_a) L_a f(a)$ equals total investment costs $V_a\kappa_a$. As a consequence, net output consists only in the total gross wages of workers, $w_a L_a f(a)$. Net output is an inverse-U shaped function of gross wage.

A type-a participating worker finds a job with probability L_a. In this case, she gets the wage w_a and has to pay income taxes $T(w_a)$. If she doesn't find a job, her income consists of the welfare benefit b. The (ex-ante) expected income of a searching individual equals $L_a [w_a - T(w_a)] + (1 - L_a) b$. If the individual decides to stay out of the labor force, then she gets the welfare benefit b and enjoys her leisure time which gives her utility d. Hence, the individual participates in the labor market as long as $L_a [w_a - T(w_a)] + (1 - L_a) b \geq b + d$. By defining worker's *ex-post* surplus as $x_a \equiv w_a - T(w_a) - b$ and worker's *expected* surplus from participation as $\Sigma_a \equiv x_a L_a$, one gets:

$$\Sigma_a = A^{\frac{1}{\gamma}} \left(\frac{a - w_a}{\kappa_a} \right)^{\frac{1-\gamma}{\gamma}} (w_a - T(w_a) - b) \tag{6}$$

Then, the participation constraint for type-a workers simplifies to

$$\Sigma_a \geq d \tag{7}$$

II.2 The wage bargain

At this stage of the game, the entry costs are sunk. If there is an agreement between the firm and the worker on the wage w, the output is produced and the firm pays the worker the negotiated wage. In the absence of an agreement, nothing is produced, and the worker only gets the welfare benefit b. The *ex-post* surplus of the worker is therefore equal to $x = w - T(w) - b$, whereas the surplus of the firm equals $a - w$. As it is standard in the literature, we assume that the wage negotiation amounts to maximize the Nash Product defined by

$$\mathcal{N}(w, x, a) \equiv (a - w)^{\frac{1-\beta}{\beta}} \; x \tag{8}$$

where $\beta \in (0, \gamma]$ denotes the worker's relative bargaining power. In this paper, we are only interested by the case where workers' bargaining power β is lower than the elasticity of the matching function η, that is

$$\beta \leq \gamma \tag{9}$$

As we will show later, the minimum wage appears at the optimum as soon as $\beta < \gamma$. If the wage that maximizes the Nash product $\mathcal{N}(w, x, a)$, lies below the minimum wage $\underline{w}$, then the minimum wage must be paid to the worker. This leads to:

$$w_a \equiv \arg\max_{w \geq \underline{w}} \quad \mathcal{N}(w, x, a) = \max \left[\underline{w}, \frac{\beta [1 - T'(w_a)] a + (1 - \beta)(T(w_a) + b)}{1 - \beta \cdot T'(w_a)} \right] \tag{10}$$

The second equality holds only for values of w_a where the function $w \mapsto T(.)$ admits a derivative. To simplify notations in what follows, we define the *maximized* Nash product

N_a as:

$$N_a = (a - w_a)^{\frac{1-\beta}{\beta}} (w_a - T(w_a) - b) \tag{11}$$

so $N_a = \mathcal{N}(w_a, x_a, a)$. Given this definition, we can rewrite the worker's expected surplus from participation as

$$\Sigma_a = A^{\frac{1}{\gamma}} (\kappa_a)^{\frac{\gamma-1}{\gamma}} \; N_a \; (a - w_a)^{\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}} \tag{12}$$

II.3 The government's problem

II.3.1 Incentive constraints

Since the government only observes the income w of employed individuals but not their productivity a, it faces an adverse selection problem. Therefore, the government has to choose a menu of bundles (w_a, x_a) that leads agents to reveal their ability. By determining the wage w_a, agents send a "message" about their productivity. The particularity of this problem in our context is that the message is here jointly determined by the worker and the firm of a match. However, since the wage maximizes the Nash product $\mathcal{N}(w, x, a)$, we[6] treat this problem as if a single agent chooses the wage that maximizes $\mathcal{N}(w, x, a)$. Therefore we can apply standard techniques (see e.g. Salanié 2002) and express incentive constraints in terms of Nash products. Using the taxation principle, it is equivalent to design a tax function $T(w)$ or to let the firm-worker pair choose among the menu of proposed bundles (w_a, x_a). To be optimal, the allocations must induce the individual matches to truthfully reveal their type, which is the case for a firm-worker pair of type a if and only if

$$\text{for all } a' \neq a \qquad \mathcal{N}(w_a, x_a, a) \geq \mathcal{N}(w_{a'}, x_{a'}, a) \tag{13}$$

In other words, a worker-firm pair of type a prefers the wage w_a designed for it (which induces the workers' ex-post surplus to be $x_a = w_a - T(w_a) - b$), instead of wage $w_{a'}$ designed for any other type a' (which induces the workers' ex-post surplus to be equal to $x_{a'} = w_{a'} - T(w_{a'}) - b$). Since at a constant value for the Nash product, we have that the marginal rate of substitution between the wage w and the worker's ex-post surplus x

$$\left.\frac{\partial x}{\partial w}\right|_{\mathcal{N}(.,.,a)} = \frac{1-\beta}{\beta} \frac{x}{a-w} \tag{14}$$

is decreasing in a, the single-crossing property is fulfilled. This allows a full description of incentive compatible allocations.

[6] For further discussions, see HLPV. In particular, since the government observes only wages, and since we rule out side-payments or tax evasion, the firm and the worker of a match cannot send separate messages.

Lemma 1 *If an allocation $a \mapsto (w_a, x_a, \Sigma_a, N_a)$ is incentive compatible then, $a \mapsto w_a$ is a nondecreasing function, $a \mapsto \Sigma_a$ is an increasing function, there exists a single type a_d such that $\Sigma_{a_d} = d$, all types with $a \geq a_d$ do participate and*

$$N_a = N_{a_d} \cdot \exp\left[\frac{1-\beta}{\beta}\int_{a_d}^{a}\frac{dt}{t-w_t}\right] \tag{15}$$

$$\Sigma_a = d \cdot \left(\frac{a-w_a}{a_d-w_{a_d}}\right)^{\frac{1-\gamma}{\gamma}-\frac{1-\beta}{\beta}} \cdot \left(\frac{\kappa_a}{\kappa_{a_d}}\right)^{\frac{\gamma-1}{\gamma}} \cdot \exp\left[\frac{1-\beta}{\beta}\int_{a_d}^{a}\frac{dt}{t-w_t}\right] \tag{16}$$

$$x_a = A^{\frac{-1}{\gamma}} \cdot \left(\frac{a-w_a}{\kappa_a}\right)^{-\frac{1-\gamma}{\gamma}} \cdot \Sigma_a = N_a \cdot (a-w_a)^{-\frac{1-\beta}{\beta}} \tag{17}$$

Conversely, for any real $a_d \in [a_0, a_1]$ and any nondecreasing function $a \mapsto w_a$, the allocation $a \mapsto (w_a, x_a, \Sigma_a, N_a)$ defined by Equations (15), (16) and (17) is incentive compatible and workers of type a participate if and only if $a \geq a_d$.

Here is a brief sketch of the Proof. Applying the envelope theorem to the wage bargaining program (11) gives the derivative of $\log N_a$. Integrating this derivative between a_d and a gives (15). To obtain (16), we first use Equation (12) at $a = a_d$ together with the participation constraint $\Sigma_{a_d} = d$ to obtain N_{a_d}. We then use (15) to get N_a and finally, we use (12) to obtain Σ_a. Equation (17) is obtained directly thanks to (6). The proof in Appendix A in particular takes care that N_a may not admit a derivative everywhere. We call equations (15) and (16) the first-order incentive constraints. We call the second-order incentive constraint the requirement that:

$$a \mapsto w_a \text{ is a non-decreasing function} \tag{18}$$

If this constraint is binding at the bottom of the wage distribution, we interpret this result as a minimum wage[7].

II.3.2 The government's objective and budget constraint

As in HLPV, we assume that the government cares only about the distribution of expected utilities, i.e. $\Sigma_a + b$ for the participating types and $b + d$ for the non participating ones. We assume the following objective to the government:

$$\Omega = F(a_d)\,\Phi(b+d) + \int_{a_d}^{a_1} \Phi\left[L_a\left(w_a - T(w_a)\right) + (1-L_a)\,b\right] f(a)\,da$$

[7] The minimum wage as a legally binding instrument might not be necessary to implement an optimum with bunching at the bottom of the wage distribution (See Cahuc and Laroque 2007 for a discussion). The tax system might be sufficient. In particular, one way to implement the bunching at the bottom is to set $T(w)$ at a very high level for $w < \underline{w}$. This is obiously equivalent to imposing a high fine for all matches that choose a wage below the minimum wage.

where $\Phi(.)$ is a twice continuously differentiable, increasing and concave function. This formulation implies that the government compensates individuals for their innate productivity a, but not for their labor market status[8]. It admits as a limiting case the *maximin* criterion. Using the definition of the worker's expected surplus (6), one can rewrite this objective as

$$\Omega = F(a_d)\Phi(b+d) + \int_{a_d}^{a_1} \Phi(b+\Sigma_a) f(a)\, da \tag{19}$$

The budget constraint can be written as

$$\int_{a_d}^{a_1} T(w_a) L_a f(a)\, da = \left[F(a_d) + \int_{a_d}^{a_1} (1-L_a) f(a)\, da\right] b + E$$

where $E \geq 0$ denotes exogenous public expenditures. Since firms' profits net of vacancy costs are nil and only labor incomes are taxed, the Walras's law implies that the government's budget constraint can be replaced by the resource constraint

$$\int_{a_d}^{a_1} Y_a(w_a) f(a)\, da = \int_{a_d}^{a_1} \Sigma_a f(a)\, da + E + b \tag{20}$$

The left-hand side of (20) denotes total production net of vacancy costs and the right-hand side the distribution of these available resources to all individuals.

Hence, the government maximizes its objective (19), subject to the budget constraint (20), the first-order incentive constraint (16), the second-order incentive constraint (18) and the participation constraint (7). Therefore, the government's problem is reduced to choosing a threshold value a_d and a nondecreasing function $a \mapsto w_a$. Given this choice, the participation constraint $\Sigma_{a_d} = d$ and Equation (12) give N_{a_d}. The first-order incentive constraint (15) gives N_a for all $a > a_d$. For all $a \geq a_d$, we get Σ_a either from N_a and equation (12) or directly through (16). The budget constraint (20) yields the value of the welfare benefit b. Finally we get the workers' ex-post surplus $x_a = w_a - T(w_a) - b$, thereby the level of tax from (6).

III The optimal policy for a given bargaining power

This section describes the consequences of policy changes on welfare when the government cannot influence the workers' bargaining power. To understand the underlying economic mechanisms at work, we need to describe the optimality conditions for all variables. We

[8] Adding this latter motive leads to an objective function of the form

$$F(a_d)\Phi(b+d) + \int_{a_d}^{a_1} \left[L_a \Phi(w_a - T(w_a)) + (1-L_a)\Phi(b)\right] f(a)\, da$$

This alternative formulation does not change the mechanisms that lead to our main results, but makes the model less tractable.

first compute the shadow cost of public funds (III.1), and then the optimality condition for the wage in the absence of bunching (III.2). After this preliminary, we derive the main Proposition of this paper (III.3), that is the optimality condition for the minimum wage level. We finally turn to the condition with respect to the participation threshold (III.4). The mathematical derivations of the optimality conditions are in Appendix B.

III.1 The shadow cost of public funds

Consider a rise in the level of public expenditures E, holding the threshold a_d and the wage distribution $a \mapsto w_a$ constant. Then, neither $Y_a(w_a)$ nor Σ_a is affected by this change. Therefore, the only change in Equation (20) is that b decreases one-to-one when E rises. The social utility of non-participating individuals decreases by $\Phi'(b+d)$ while the social expected utility of participating individuals of type t decreases by $\Phi'(b+\Sigma_t)$. Let λ be the shadow cost of public funds. We hence get that:

$$\lambda = \Phi'(b+d)\,F(a_d) + \int_{a_d}^{a_1} \Phi'(\Sigma_t + b)\,f(t)\,dt \tag{21}$$

III.2 Optimal negotiated wages w_a

We now describe the optimality condition for the wage w_a at a point of the skill distribution where there is no bunching (and at a point where $a \mapsto w_a$ is continuous). We consider a marginal translation δw of the function $a \mapsto w_a$ on an infinitesimal interval[9] $[a, a+\delta a]$. We get from Appendix B.1:

$$\begin{aligned} 0 \;=\; & \underbrace{\lambda \frac{\partial Y_a}{\partial w_a}(w_a)\,f(a)}_{\text{Efficiency effect}} - \underbrace{\frac{1-\beta}{\beta\,(a-w_a)^2} Z_a}_{\text{Informationnal rents effect}} \\ & \underbrace{+\left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right)\frac{1}{a-w_a}\left[\Phi'(\Sigma_a+b)-\lambda\right]\Sigma_a f(a)}_{\text{Bargaining power effect}} \end{aligned} \tag{22}$$

where

$$Z_a = \int_a^{a_1} \left(\lambda - \Phi'(\Sigma_t + b)\right)\Sigma_t f(t)\,dt \tag{23}$$

Consider the optimization problem for agents of type a. Given the participation constraint (7), Equation (12) and the first-order incentive constraint (15), the Nash product for type a is predetermined and not affected by the change in the wage w_a.

[9] Since we are considering the optimal wage w_a at at point where there is no bunching, this optimal wage has also to solve a "locally relaxed problem" that is identical to the government's problem, except that the second-order incentive constraint (18) is not imposed in a neighborhood of a. It is the necessary condition of this locally relaxed problem that we are actually deriving.

The first term in equation (22) stands for the *efficiency* part of the trade-off. An increase in the wage rate w_a implies that less vacancies are created, which has two consequences. First, it decreases employment and therefore gross output. But it also decreases the resources used for investments in capital to build workstations. The effect on output net of investment costs equals $\partial Y_a / \partial w_a$ and is therefore ambiguous (see equation (5)). Multiplying this by the shadow cost λ of public funds and the density $f(a)$ of type-a workers gives the *efficiency* term in (22).

The second term in equation (22) represents the impact on *informational rents* of a higher gross wage for type-a workers. When firm-worker matches endowed with productivity a have a higher gross wages (while keeping the Nash product N_a fixed), more productive firm-worker matches find it more attractive to mimic them. In other words, a type-$t > a$ worker-firm pair finds it profitable (in terms of Nash Product 8) to choose the wage w_a designed for type-a jobs instead of the wage w_t designed for them.

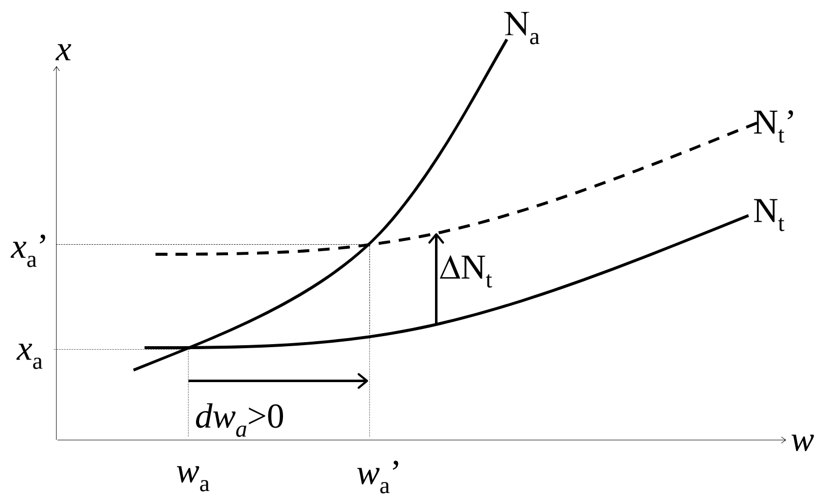

Figure 1: The informational rent effect

Figure 1 illustrates this point. It displays the iso-Nash curves for types a and t with $t > a$. Because productivity is higher, the elasticity of the firm's surplus with respect to the gross wage is smaller for worker-firm pairs of type t than for pairs of type a. Consequently, the iso-Nash curves corresponding to type a are steeper than iso-Nash curves corresponding to type t (see Equation 14). For a given level of Nash product accruing to type a, the government can propose different bundles, for instance a low wage combined with a low ex-post surplus (w_a, x_a) or a higher wage combined with a higher ex-post surplus (w'_a, x'_a). If the government proposes (w_a, x_a), it has to give a Nash product at least equal to N_t to prevent type-t worker-firm pairs to mimic type-a pairs. If the government however proposes the allocation (w'_a, x'_a), then, it has to give a Nash product N'_t, which is strictly higher than N_t. Hence, the higher the wage w_a, the higher the Nash product for types t above a, which means that a higher wage increases informational rents. In the "informational rent" term of equation (22), the term in front of Z_a measures the rate at which the growth rate of the worker's maximized Nash product has to increase to prevent more productive matches from mimicking the type a match.

From the first-order incentive constraint (15), the Nash product designed for all types t above $a + da$ has to increase by the same proportion to prevent mimicking. Since wages designed for all these types above $a+da$ are not changed, we get that the expected surplus for these types increases in the same proportion, so $\Delta\Sigma_t/\Sigma_t = \Delta N_t/N_t = \Delta N_{a+da}/N_{a+da}$. The increase in the expected surplus obtained by type-t worker-firm pairs increases the social welfare by $\Phi'(\Sigma_t)$, but implies a budgetary cost equal to λ. Integrating these two terms over all types t above $a + da$ gives the shadow cost Z_a of a *relative* increase of the type-a Nash product (see equation (23))[10]. Taking the limit when da tends to 0 leads to the *informational rent* effect in (22). We get the following Lemma, which is proved in Appendix B.2:

Lemma 2 *The shadow cost Z_a of a relative increase in the Nash product is positive for all $a < a_1$.*

In other words, the government desires to avoid informational rents.

A third effect appears in the present model due to the fact that workers' bargaining power β is below the elasticity of the matching function, γ. In our model, the expected surplus Σ_a that the government focuses on does not coincide with the Nash product N_a that the firm and the worker maximize when they negotiate the wage. For a given maximized Nash product N_a (that is predetermined by the incentive constraints), a change in the wage w_a has also an impact on the expected surplus Σ_a as described in equation (12).

The intuition is depicted in Figure 2. Both Nash product N_a and workers' expected surplus Σ_a are increasing functions of workers' ex-post surplus x_a and of firms' surplus $a - w_a$ (thereby a decreasing function of the gross wage w_a). However, if the Hosios condition $\beta = \gamma$ is not fulfilled, they put different relative weights on these two components. Hence, the marginal rate of substitution between gross wages w_a and workers' ex-post surplus x_a that keep workers' expected surplus Σ_a unchanged differ from the one that keep Nash product N_a unchanged. Since $\beta < \gamma$, the Nash product criterion puts a higher relative weight on gross wages w. Hence, increasing the wage w_a for a given Nash product N_a increases workers' expected surplus Σ_a, while increasing the wage w_a for a given workers' expected surplus Σ_a decreases the Nash product N_a. This implies that the iso-Nash curves are steeper than the corresponding iso-expected-surplus curves. When the employment

[10] Formally $-Z_a$ is the welfare effect of a *relative* unit change in the Nash product N_a, keeping the function $t \mapsto w_t$ unchanged for $t \in [a, a_1]$.

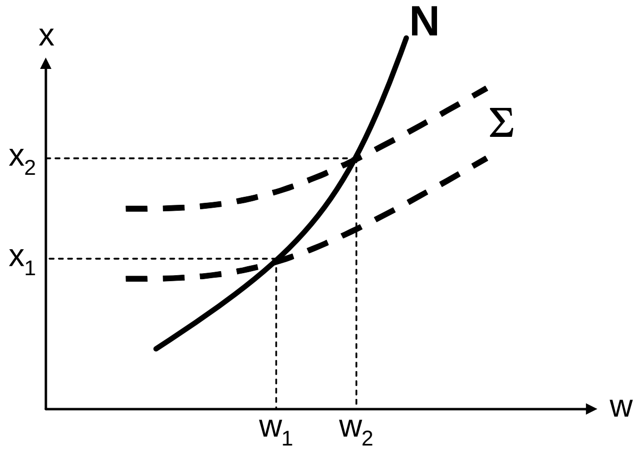

Figure 2: Expected Surplus and the Nash Product

level decreases by an increase in gross wages from w_1 to w_2, the net income has to rise from x_1 to x_2 to give the same Nash product as before to this firm-worker match. In terms of expected surplus, this increase in net income more than compensates for the employment loss due to the relatively small increase in gross wages. The expected surplus therefore increases.

The increase in Σ_a is valued at the marginal social utility of type a, namely Φ'_a. Since the government thus gives more resources to agents of type a, less resources can be affected to redistribution toward other individuals. This decrease in budgetary funds is valued at the marginal cost of public funds λ. An increase in the wage rate is therefore desirable if type a is a low-productivity type that gets a lower expected surplus than the high-productivity type because of the first-order incentive constraint. Giving resources to this low-productivity type is socially valuable because it increases equity. The contrary is true if type a is a high-productivity type from whom the government wants take resources in order to redistribute.

The sum of these three terms gives (22).

III.3 Optimal minimum wage $\underline{w}$

We are now in the position to derive the main result of this section. We interpret the minimum wage as a bunching of wages over an interval at the bottom of the skill distribution. Let a_m denote the upper bond of this interval (see Figure 3). We now consider the welfare effect of an increase in the minimum wage $\underline{w}$, that is of an increase of wages over $[a_d, a_m]$ up to w_{a_m}. In this policy change, we keep a_d unchanged. The tax function $T(.)$ is adjusted so that participation decisions remain unchanged.

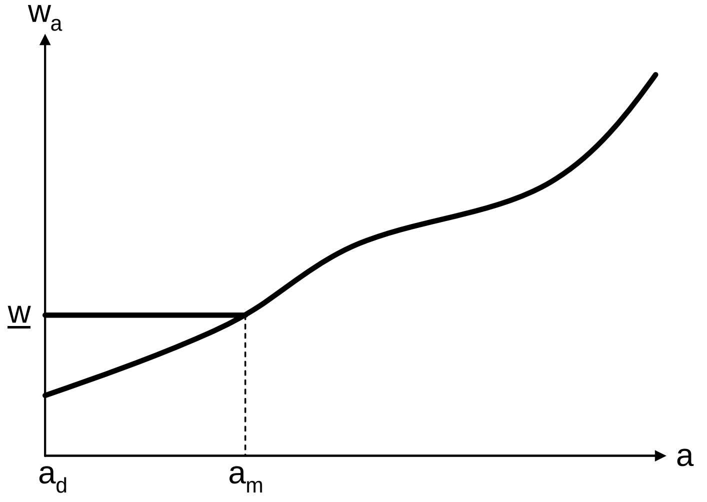

Figure 3: The Minimum Wage

We show in Appendix B.3 that:

$$\begin{aligned}\frac{\partial \Omega}{\partial \underline{w}} &= \lambda \int_{a_d}^{a_m} \frac{\partial Y_a}{\partial w_a}(\underline{w}) f(a)\, da \\ &+ \int_{a_d}^{a_m} \left(\Phi'(\Sigma_a + b) - \lambda\right) \Sigma_a \frac{1-\gamma}{\gamma}\left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}}\right) f(a)\, da \\ &+ \left(\frac{1-\beta}{\beta}\frac{1}{a_m - \underline{w}} - \frac{1-\gamma}{\gamma}\frac{1}{a_d - \underline{w}}\right) Z_{a_m}\end{aligned} \tag{24}$$

An increase in the minimum wage has three effects. The first term captures the efficiency effect. In accordance with common wisdom, a higher minimum wage decreases employment (and hence increases unemployment). Even though this leads to a lower gross output,the higher wage also implies that less resources are spent by firms investing in new workstations. Therefore, the effect on net output $\partial Y_a / \partial w_a(\underline{w})$ is ambiguous.

The second term captures the effect on the expected surplus of the workers of types a in $[a_d, a_m]$ whose wages are given by the minimum wage. The effect of a marginal increase $\Delta \underline{w}$ of the minimum wage on the expected surplus of these workers is twofold. First, the rise in the minimum wage reduces the labor demand. Equation (4) implies that a rise $\Delta \underline{w}$ of the minimum wage decreases the probability for a type-a searching worker to find a job by

$$\frac{\Delta L_a}{L_a} = \frac{1-\gamma}{\gamma}\frac{\Delta \underline{w}}{a - \underline{w}}$$

Second, the government has to increase $\underline{w} - T(\underline{w})$ to keep type-a_d workers in the labor force. Hence, one has

$$\frac{\Delta\left(\underline{w} - T(\underline{w}) - b\right)}{\underline{w} - T(\underline{w}) - b} = -\frac{\Delta L_{a_d}}{L_{a_d}} = \frac{1-\gamma}{\gamma}\frac{\Delta \underline{w}}{a_d - \underline{w}}$$

to keep type a_d workers indifferent between participating or not. The net effect on workers' expected surplus Σ_a is therefore

$$\frac{\Delta\Sigma_a}{\Sigma_a} = \frac{\Delta L_a}{L_a} + \frac{\Delta\left(\underline{w} - T\left(\underline{w}\right) - b\right)}{\underline{w} - T\left(\underline{w}\right) - b} = \frac{1-\gamma}{\gamma}\left(\frac{\Delta\underline{w}}{a_d - \underline{w}} - \frac{\Delta\underline{w}}{a - \underline{w}}\right)$$

At a given wage level, the probability of finding a job is relatively less sensitive to wage for workers of a higher productivity. Hence, for $a > a_d$, the net effect of a rise in minimum wage on workers' expected surplus is positive. The induced increase in Σ_a has a direct consequence on the government's objective that is valued at a rate $\Phi'\left(\Sigma_a\right)$, and a negative effect on public funds that is valued at rate λ. Summing this effect for all types between a_d and a_m gives the second term in the right-hand side of (24).

Finally, a rise in the minimum wage changes the level of the Nash product for the workers of the limiting type a_m. Firms' ex-post surplus $a_m - w_{a_m}$ decreases by $-\Delta\underline{w}$ while workers' ex-post surplus increases by $\Delta\left(\underline{w} - T\left(\underline{w}\right) - b\right)$, so the net effect equals to

$$\begin{aligned} \frac{\Delta N_{a_m}}{N_{a_m}} &= \frac{1-\beta}{\beta}\frac{\Delta\left(a_m - w_{a_m}\right)}{a_m - w_{a_m}} + \frac{\Delta\left(\underline{w} - T\left(\underline{w}\right) - b\right)}{\underline{w} - T\left(\underline{w}\right) - b} \\ &= \left(\frac{1-\gamma}{\gamma}\frac{1}{a_d - \underline{w}} - \frac{1-\beta}{\beta}\frac{1}{a_m - \underline{w}}\right)\Delta\underline{w} \end{aligned} \tag{25}$$

However, this relative change in the Nash product of these workers spills over the whole distribution of productivity to prevent the mimicking off worker-firm pairs of all type t above a_m (see 15). Since Z_a is the shadow cost of a relative increase in the Nash product, the last term of equation (24) represents the effect of the minimum wage on the individuals whose wage is not constraint by the minimum wage.

From this optimality condition on the minimum wage, we can derive the following proposition:

Proposition 1 *If $\beta < \gamma$, a binding minimum wage is optimal.*

Proof. Assume by contradiction that bunching at the bottom is not optimal. For this, we consider a threshold a_d and a function $a \mapsto w_a$, such that this function is increasing in the neighborhood of $a = a_d$. Hence we get $a_m = a_d$. So, Equation (24) simplifies to

$$\frac{\partial\Omega}{\partial\underline{w}} = \left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right)\frac{Z_{a_d}}{a_d - w_{a_d}} \tag{26}$$

Since $\beta < \gamma$, the terms in the bracket is positive. Furthermore, Z_{a_d} is positive too from Lemma 2. Therefore, one finds that, starting from any situation with no minimum wage, there exists a binding level of minimum wage that is welfare-improving. ■

To recommend a minimum wage even though it increases unemployment might seem slightly counterintuitive at first sight. To explain the mechanisms that lie behind this result, start from the case where there is no binding minimum wage (hence $a_m = a_d$). Now, consider an increase in the wage of type a_d above the wage w_{a_d} that solves equation (22). Given that the participation constraint for a_d has still to hold with equality, equation (12) implies that this increase in the wage decreases the maximized Nash product N_{a_d}. This initial level of the maximized Nash product is however important: Through the first-order incentive constraint (15), it determines the evolution of the Nash product for all types. These Nash products then again determine the surplus given to individuals by (12). Hence, decreasing the initial maximized Nash product N_{a_d} by an increase in w_{a_d} implies that the rents for all types $a > a_d$ decrease. This mechanism thus allows the government to extract more resources from the high-productivity individuals ($a > a_d$). These resources are then available for redistributive purposes of the government. However, since all wages for higher-productivity types solve equation (22), the increase in w_{a_d} implies that there is a downward jump in wages right above a_d. This clearly violates the second-order incentive constraint (18). As a consequence, the optimal solution exhibits bunching at the bottom of the wage distribution. Such a bunching can be interpreted as a binding minimum wage.

Technically speaking, the new feature of our model is that while the worker and the firm maximize the Nash product, the participation constraint depends on another variable, the expected surplus. In the traditional adverse selection model in contract theory, the variables concerning the agent's maximization problem and the participation constraint are the same. By the participation constraint, the maximized utility of the lowest participating type has to equal the (exogenous) outside option of the individual. Therefore, the principal cannot affect the maximized utility of the lowest participating type. From this utility of the lowest participating type, the incentive constraints then determine the evolution of the maximized utility for all other types, and hence the informational rent given to these types. In our model, things are different. Even though the expected surplus of the lowest participating type must equal the outside option (i.e. the utility of leisure in our case), the government can decrease the level of the maximized Nash product by imposing a very high wage on the lowest participating individual (see equation (12)). This implies through the incentive constraints that the Nash products of all types decrease, and, again by equation (12), expected surplus of all individuals above the lowest type decreases (See the effect of w_{a_d} in Equation (16)). Hence, this allows the government to decrease informational rents of the agents and use these resources for redistributive purposes, which in turn increases social welfare. However, choosing a very high wage for the lowest participating individual inevitably violates the second-order incentive constraint. As a consequence, there is bunching at the bottom of the wage distribution in the second-order approach, and since this constraint is on the wage, we can interpret this bunching as a minimum wage.

III.4 Optimal participation a_d

For completeness, we finally consider a marginal change in the threshold a_d, keeping the function $a \mapsto w_a$ unchanged. The optimal participation decision sets the optimal value of the threshold type a_d and writes (see Appendix B.4)

$$\lambda \left(Y_{a_d}\left(w_{a_d}\right) - d\right) f\left(a_d\right) = \frac{1-\gamma}{\gamma}\left(\frac{1}{a_d - \overline{w}} - \frac{\dot{\kappa}_{a_d}}{\kappa_{a_d}}\right) Z_{a_d} \tag{27}$$

The left-hand side of (27) corresponds to the efficiency part of the trade-off. When a_d decreases, participation increases by an amount that is proportional to the density $f\left(a_d\right)$. The participation of these workers increases total net output by Y_{a_d} but requires that they receive an expected surplus at least equal to d to participate. This net budgetary gain is socially valued at the shadow cost of public funds, λ. The right-hand side of (27) is the equity part of the trade-off. When a_d decreases, worker-firm pairs with productivity a_d have the possibility to mimic the additional participants. To avoid this mimicking, the government has to give an additional informational rent to type-a_d matches. The term in front of Z_{a_d} is equal to the relative increase in the Nash product that should be given to type-a_d matches to prevent them from mimicking the new entrants. The equity cost multiplies this increase by the shadow cost Z_{a_d} of a relative increase in the Nash product N_{a_d}.

IV Varying bargaining power

The previous section considered the optimal policy if the bargaining power is exogenous. This section has a look at what happens when the government can also influence the worker's bargaining power β. There might be different ways how the government can affect the bargaining power. The law can change the bargaining procedures, or the way unions are financed, etc. Explaining in more details how bargaining power is changed is beyond the scope of the paper. We simply assume that the government has some degree of latitude about the bargaining power through institutional settings.

The envelope theorem implies that $\frac{d\Omega}{d\beta} = \frac{\partial\Omega}{\partial\beta}$. Analyzing this last expression, we can derive the following proposition, which is proved in Appendix C:

Proposition 2 *A worker's bargaining power β that is below the elasticity of the matching function, γ, is never optimal.*

To understand this result, one might identify the distortions that are present in the (second-best) optimum. In the absence of redistribution and under the Hosios condition (i.e. $\beta = \gamma$), negotiated wages maximize net output. Since the government wants to redistribute from high- to low-income individuals, it wants to install a high marginal tax rate. According to equation (10), this distorts the wage downwards. However, the bargaining power also distorts the wage levels. A higher bargaining power increases the wage, again according to equation (10). Therefore, a rise in the bargaining power induces a distortion on the wage that (partly) offsets the one induced by the redistributive taxation. The equity-efficiency trade-off becomes less severe the higher the worker's bargaining power β. An increase in β is thus always desirable, at least up to the point where $\gamma = \beta$.

We are not able to find analytical results for the case where $\beta > \gamma$. In fact, Appendix C shows that increasing the workers's bargaining power is welfare-improving as long as Σ_a is increasing in a, that is, as long as the desired redistribution goes from high- to low-productive workers. Lemma 1 only proves that this is the case when $\beta \leq \gamma$. If $\beta > \gamma$, then the worker's expected surplus at the optimal solution might not be monotonically increasing in a anymore.

To understand why we need workers' expected surplus Σ_a to be increasing in types to obtain our result that a rise in workers' bargaining power is welfare-improving, we have to analyze whether a rise in the bargaining power β relaxes or strengthens the relevant incentive constraints. When workers' expected utility Σ_a increases with their productivity a, the relevant incentive constraint is that a type-a match does not want to mimic slightly less productive matches. In other words, when they negotiate the wage, the relevant incentive constraint induces the worker and the firm of a type-a job to choose the wage w_a designed for them, and not the wage w_{a-da} designed for slightly less productive jobs of type $a-da$. Obviously, the higher the worker's bargaining power β, the harder it becomes for the firm to obtain w_{a-da} as the bargaining outcome instead of w_a. Therefore, a rise in the bargaining power β relaxes the incentive constraints that prevent worker-firm pairs from mimicking less productive worker-firm pairs, which explains why in such a context, the government can achieve a better outcome.

V Conclusion

We have given a sufficient condition for the minimum wage to be a part of the optimal redistributive policy: If the bargaining power is lower than the elasticity of the matching function, then the introduction of a binding minimum wage is welfare-improving. However, if the government can also control the workers' bargaining power, it should increase it, at least up to a point where our argument in favor of minimum wage does no longer apply. Hence, our argument in favor of the introduction of a binding minimum wage only holds if the government cannot control the bargaining power, and if this parameter is relatively low. In other words, the minimum wage is an imperfect substitute for a rise in workers' bargaining power. Hence, the government should prefer the latter, in combination with redistributive taxation.

Whether and how the government can affect the bargaining power is still an open question. It would also be interesting to determine the optimal level of the bargaining power when income taxation is simultaneously optimized. We have shown that as long as the workers' expected surplus remains increasing in types, increasing the bargaining power is welfare-improving because it relaxes the relevant incentive constraint. Hence, since the workers' expected surplus is increasing whenever the bargaining power is not higher than the elasticity of the matching function, this suggest that the optimal bargaining power in our redistributive context is higher than the one prescribed by Hosios (1990) in a pure efficiency context. We left the further characterization of the optimal bargaining power for future research. Furthermore, it might be interesting to see whether it is possible to generalize the framework developed in the present paper to other models of adverse selection.

References

[1] Allen, F., 1982, Optimal linear income taxation with general equilibrium effects on wages, *Journal of Public Economics*, 17, 135-143.

[2] Allen, S., 1987, Taxation, Redistribution, and the minimum wage: a theoretical analysis, *Quarterly Journal of Economics*, 101, 477-489.

[3] Boadway, R., and K. Cuff, 2001, A minimum wage can be welfare-improving and employment-enhancing, *European Economic Review*, 45, 553-576.

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

[5] Cahuc, P. and G. Laroque, 2007, "Optimal Taxation and Monopsonistic Labor Market: Does Monopsony justify the Minimum Wage?", Working Paper CREST.

[6] Drèze, J. and Gollier, C., 1993, Risk Sharing on the Labour Market and Second-Best Wage Rigidity, *European Economic Review*, 37, 1457-1482.

[7] Guesnerie, R. and Roberts, K., 1987, Minimum Wage Legislation as a second best policy, *European Economic Review*, 31, 490-498.

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

[9] Hungerbühler, M., Lehmann, E., Parmentier, A., Van der Linden, B., 2006, Optimal redistributive taxation in a search equilibrium model, *Review of Economic Studies*, 73(3), 715-739.

[10] Lollivier, S. and J-C. Rochet, 1983, Bunching and Second-Order Conditions: A note on Optimal Tax Theory, *Journal of Economic Theory*, 31, 392-400.

[11] Marceau, N. and Boadway R., 1994, Minimum Wage Legislation and Unemployment Insurance as Instruments for Redistribution, *Scandinavian Journal of Economics*, 96, 67-81.

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

[13] Mortensen, D. and Pissarides, C., 1999, "New developments in models of search in the Labor Market", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol 3 (North-Holland, Amsterdam).

[14] Petrongolo, B. and Pissarides, C. A. , 2001, "Looking into the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, 38, 390-431.

[15] Pissarides, 2000, *Equilibrium Unemployment Theory*, MIT Press, Cambridge, MA, US.

[16] Robinson, J., 1933, *The Economics of Imperfect Competition*, MacMillan, New York.

[17] Salanié, 2002, *Contract Theory, a primer*, 2nd edition, MIT Press, Cambridge, MA, US.

[18] Stigler, G., 1946, The economics of minimum wage legislation, *American Economic Review*, 36(3), 358-365.

[19] Stiglitz, J. E., 1982, "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics*, 17, 213-240.

Appendix

A Proof of Lemma 1

Let I_a be a function defined for $a \in [a_d, a_1]$ by

$$I_a \equiv \sup_w (a - w) \cdot (w - T(w) - b)^{\frac{\beta}{1-\beta}}$$

From (11), one has

$$N_a = (I_a)^{\frac{1-\beta}{\beta}} \tag{28}$$

Without any mathematical restrictions on the tax function, for any w, the function $a \mapsto (a - w)(w - T(w) - b)^{\frac{\beta}{1-\beta}}$ is linear in a. Therefore, the function $a \mapsto I_a$ is the convex envelope of linear and increasing functions of a. This ensures that function $a \mapsto I_a$ is convex and increasing in a. Therefore, $a \mapsto I_a$ is continuous on $[a_d, a_1)$ and admits a derivative everywhere except on a countable set (we use henceforth the acronym *a.e.* for "almost everywhere"). Furthermore, by the envelope theorem, we get that whenever $a \mapsto I_a$ admits a derivative $\dot{I}_a$, this derivative verifies

$$\dot{I}_a \overset{a.e}{=} (x_a)^{\frac{\beta}{1-\beta}} \qquad \Leftrightarrow \qquad \frac{\dot{I}_a}{I_a} \overset{a.e}{=} \frac{1}{a - w_a} \tag{29}$$

Since $a \mapsto I_a$ is convex, one has that $a \mapsto x_a$ is non-decreasing. To show that $a \mapsto w_a$ is non-decreasing too, we assume by contradiction it is not and show that the mechanism $a \mapsto (w_a, x_a)$ then violates (13). Let then assume there exists $a' > a$ such that $w_{a'} < w_a$. Recall that according to the definition of the Nash product in (8), $\mathcal{N}(w, x, a)$ is decreasing in the wage w and increasing in workers' ex-post surplus x. Hence we get $\mathcal{N}(w_a, x_a, a) < \mathcal{N}(w_{a'}, x_a, a)$. Since $a' > a$, we have that $x_{a'} \geq x_a$, which induces in turn $\mathcal{N}(w_a, x_a, a) < \mathcal{N}(w_{a'}, x_{a'}, a)$. This last inequality contradicts (13). Therefore $a \mapsto w_a$ is non-decreasing.

We now prove that Σ_a is increasing in a. Consider two skill levels $a' > a$. Then, from (8) and (13), one gets

$$\log x_{a'} - \log x_a \geq \frac{1-\beta}{\beta} \left[\log(a' - w_a) - \log(a' - w_{a'})\right]$$

From (6), one has

$$\begin{aligned} \log \Sigma_{a'} - \log \Sigma_a &= \frac{1-\gamma}{\gamma}\left[\log\left(\frac{a' - w_{a'}}{\kappa_{a'}}\right) - \log\left(\frac{a - w_a}{\kappa_a}\right)\right] \\ &\quad + \log x_{a'} - \log x_a \end{aligned}$$

So,

$$\begin{aligned} \log \Sigma_{a'} - \log \Sigma_a &\geq \frac{1-\gamma}{\gamma}\left[\log\left(\frac{a' - w_{a'}}{\kappa_{a'}}\right) - \log\left(\frac{a - w_a}{\kappa_a}\right)\right] \\ &\quad + \frac{1-\beta}{\beta}\left[\log(a' - w_a) - \log(a' - w_{a'})\right] \end{aligned}$$

which gives

$$\begin{aligned} \log \Sigma_{a'} - \log \Sigma_a &\geq \frac{1-\gamma}{\gamma}\left[\log\left(\frac{a' - w_a}{\kappa_{a'}}\right) - \log\left(\frac{a - w_a}{\kappa_a}\right)\right] \\ &\quad + \left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right)\left[\log(a' - w_a) - \log(a' - w_{a'})\right] \end{aligned}$$

and finally

$$\begin{aligned} \log \Sigma_{a'} - \log \Sigma_a &\geq \frac{1-\gamma}{\gamma}\int_a^{a'}\left(\frac{1}{t - w_{a'}} - \frac{\dot{\kappa}_t}{\kappa_t}\right)dt \\ &\quad + \left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right)\left[\log(a' - w_a) - \log(a' - w_{a'})\right] \end{aligned}$$

Since $t - w_{a'} < t$, Equation (1) ensures that the first term in the right-hand side of the last expression is positive. Since $a' > a$, one has $a' - w_a \geq a' - w_{a'}$ from (18). Hence, whenever $\beta \leq \gamma$, the second term in the last expression is the product of two non-negative terms. So $\log \Sigma_{a'} - \log \Sigma_a$ is larger than the sum of a positive term and a non-negative term, which ends the proof that Σ_a is increasing in a. As a consequence, there exists a unique[11] productivity level $a_d \in [a_0, a_1]$ such that $\Sigma_{a_d} = d$. Moreover, for any $a \geq a_d$, one has $\Sigma_a \geq d$, so these workers of type a choose to search for a job. Conversely, for any $a < a_d$, one has $\Sigma_a < d$ so these workers choose not to search for a job.

Finally, since I_a is a continuous function of a, the integration of (29) between a_d and a gives $\log I_a = \log I_{a_d} + \exp\left[\int_{a_d}^{a} \frac{dt}{t - w_t}\right]$. Together with (28), this last equality gives (15). Using (12) on a and on a_d with $d = \Sigma_{a_d}$ finally leads to (16). Finally (17) is directly obtained with either (6) or (8).

We now show the converse. Take a threshold value $a_d \in [a_0, a_1]$ and a non-decreasing function $a \mapsto w_a$ defined on $[a_d, a_1]$. We verify whether the allocation $a \mapsto (w_a, N_a, \Sigma_a, x_a)$ defined by Equations (15) (16) and (17) satisfies the incentive constraints (13) and the participation constraint (7) for all $a \in [a_d, a_1]$.

We first verify that $a \mapsto \Sigma_a$ is increasing in a. Let $a < a'$. We get from (16)

$$\begin{aligned} \log \Sigma_{a'} - \log \Sigma_a &= \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right) [\log (a' - w_{a'}) - \log (a - w_a)] \\ &\quad + \frac{1-\beta}{\beta} \int_a^{a'} \frac{dt}{t - w_t} - \frac{1-\gamma}{\gamma} [\log \kappa_{a'} - \log \kappa_a] \\ &= \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right) \left[\log\left(\frac{a' - w_{a'}}{a' - w_a}\right) + \log\left(\frac{a' - w_a}{a - w_a}\right)\right] \\ &\quad + \frac{1-\beta}{\beta} \int_a^{a'} \frac{dt}{t - w_t} - \frac{1-\gamma}{\gamma} [\log \kappa_{a'} - \log \kappa_a] \end{aligned}$$

However,

$$\log\left(\frac{a' - w_a}{a - w_a}\right) = \int_a^{a'} \frac{dt}{t - w_a} \leq \int_a^{a'} \frac{dt}{t - w_t}$$

To obtain the last inequality, one should note that, since $t > a > w_a$ and $a \mapsto w_a$ is non-decreasing, the inequality $1/(t - w_a) \leq 1/(t - w_t)$ holds. For $\beta \leq \gamma$, one has $\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta} \leq 0$, thereby:

$$\begin{aligned} &\left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right) \left[\log\left(\frac{a' - w_{a'}}{a' - w_a}\right) + \log\left(\frac{a' - w_a}{a - w_a}\right)\right] \\ \geq\ &\left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right) \log\left(\frac{a' - w_a}{a' - w_{a'}}\right) + \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right) \int_a^{a'} \frac{dt}{t - w_t} \end{aligned}$$

Since, $\log \kappa_{a'} - \log \kappa_a = \int_a^{a'} (\dot{\kappa}_t/\kappa_t)\, dt$, we get

$$\log \Sigma_{a'} - \log \Sigma_a \geq \left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right) \log\left(\frac{a' - w_a}{a' - w_{a'}}\right) + \frac{1-\gamma}{\gamma} \int_a^{a'} \left(\frac{1}{t - w_t} - \frac{\dot{\kappa}_t}{\kappa_t}\right) dt$$

[11] With the convention that if $\Sigma_{a_0} > d$ this solution is $a_d = a_0$ and if $\Sigma_{a_1} < d$, this solution is $\alpha_d = a_1$.

Since $a \mapsto w_a$ is non decreasing (thereby $a' - w_a \geq a' - w_{a'}$), the first term in the right hand side of the inequality is non-negative. From assumption (1), the second term is positive. Hence $\Sigma_{a'} > \Sigma_a$ and the participation constraint (7) holds for all $a \geq a_d$.

We now show that this mechanism also verifies (13). Take $a' \neq a$. If $w_{a'} \geq a$, then $\mathcal{N}(w_{a'}, x_{a'}, a) \leq 0 < N_a$. If conversely $w_{a'} < a$, it is equivalent to prove (13), or to prove that $\log N_a - \log \mathcal{N}(w_{a'}, x_{a'}, a)$ is non-negative. Given (8), this last expression equals

$$\log N_a - \log \mathcal{N}(w_{a'}, x_{a'}, a) = \log N_a - \log N_{a'} + \frac{1-\beta}{\beta}\left[\log(a' - w_{a'}) - \log(a - w_{a'})\right]$$

Given (15), we get

$$\begin{aligned} \log N_a - \log \mathcal{N}(w_{a'}, x_{a'}, a) &= \frac{1-\beta}{\beta}\left[\log(a' - w_{a'}) - \log(a - w_{a'}) - \int_a^{a'} \frac{dt}{t - w_t}\right] \\ &= \frac{1-\beta}{\beta}\left[\int_a^{a'} \left(\frac{1}{t - w_{a'}} - \frac{1}{t - w_t}\right) dt\right] \end{aligned}$$

So,

- if $a' < a$, then for all $t \in [a', a]$, one has $0 < w_{a'} \leq w_t < t$, so the integrand is non-positive. Since $a' < a$, the integral is therefore non-negative which ensures that (13) holds.
- if $a' > a$ and $w_{a'} < a$, then for all $t \in [a, a']$, one has $0 < w_t \leq w_{a'} < t$, and the integrand is non-negative. Since $a < a'$, the integral is therefore non-negative, too.

B The government's problem

It is convenient to rewrite (16) as

$$\begin{aligned} \log \Sigma_a &= \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right)\left[\log(a - w_a) - \log(a_d - w_{a_d})\right] \\ &\quad + \frac{1-\beta}{\beta}\int_{a_d}^{a} \frac{dt}{t - w_t} - \frac{1-\gamma}{\gamma}\log\left(\frac{\kappa_a}{\kappa_{a_d}}\right) + \log d \end{aligned} \tag{30}$$

B.1 Optimality conditions with respect to negotiated wages w_a for all $a \in (a_m, a_1]$

We consider the effect of a variation δw in the wage w_a for the agents of type $[a, a + \delta a[$ with δw and δa being infinitesimally small and $a > a_m$. From Equation (30), the implied variation $\delta \Sigma_t$ of workers' expected surplus equals

$$\delta \Sigma_t = \frac{1-\beta}{\beta\ (a - w_a)^2}\Sigma_t\ \delta w\ \delta a$$

for $a + da < t$,

$$\delta\Sigma_t = -\left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right)\frac{1}{a-w_a}\ \Sigma_t\ \delta w + \frac{1-\beta}{\beta\left(a-w_a\right)^2}\Sigma_t\ \delta w\ \delta a$$

for[12] $a \leq t < a + \delta a$ and $\delta\Sigma_t = 0$, for $t < a$.

From the budget constraint (20), the variation of the welfare benefit is

$$\delta b = \int_{a_d}^{a_1} \left(\delta Y_t\left(w_t\right) - \delta\Sigma_t\right) f\left(t\right) dt$$

Hence, for any $a \in (a_m, a_1]$

$$\begin{aligned}\delta b &= \frac{\partial Y_a}{\partial w_a}\left(w_a\right)\ \delta w\ \delta a + \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta}\right)\frac{1}{a-w_a}\ \Sigma_a\ f\left(a\right)\ \delta w\ \delta a \\ &\quad -\frac{1-\beta}{\beta\left(a-w_a\right)^2}\left(\int_a^{a_1}\Sigma_t f\left(t\right) dt\right)\ \delta w\ \delta a\end{aligned}$$

From the government's objective (19), one has

$$\begin{aligned}\delta\Omega &= \lambda\ \delta b + \int_a^{a_1}\Phi'\left(\Sigma_t + b\right)\ \delta\Sigma_t\ f\left(t\right)\ dt \\ &= \lambda\frac{\partial Y_a}{\partial w_a}\left(w_a\right) f\left(a\right)\int_a^{a_1}\left(\Phi'\left(\Sigma_t\right) - \lambda\right)\ \delta\Sigma_t\ f\left(t\right)\ dt\end{aligned}$$

where we define the shadow cost of public funds λ through Equation (21). Hence

$$\begin{aligned}\frac{\delta\Omega}{\delta w\delta a} &= \lambda\frac{\partial Y_a}{\partial w_a}\left(w_a\right) f\left(a\right) + \left(\Phi'\left(\Sigma_a + b\right) - \lambda\right)\left(\frac{1-\beta}{\beta} - \frac{1-\gamma}{\gamma}\right)\frac{1}{a-w_a}\Sigma_a f\left(a\right) \\ &\quad +\frac{1-\beta}{\beta\left(a-w_a\right)^2}\int_a^{a_1}\left(\Phi'\left(\Sigma_t + b\right) - \lambda\right)\ \Sigma_t\ f\left(t\right) dt\end{aligned}$$

which gives (22), together with (23).

B.2 Proof of Lemma 2

Since $t \mapsto \Sigma_t$ admits everywhere a left and right derivative, so does the function $t \mapsto \left(\lambda - \Phi'\left(\Sigma_t + b\right)\right)\Sigma_t f\left(t\right)$. The integral in (23) is therefore well defined, is a continuous function of t, and admits everywhere a left and a right derivative.

From Lemma 1, we know that Σ_a is increasing in a. Hence the marginal social welfare $\Phi'\left(b + \Sigma_a\right)$ is decreasing in a. The shadow cost of public funds λ equals the average of all marginal social welfare (see equation 21). Hence there exists a unique a_c for which $\Phi'\left(b + \Sigma_{a_c}\right) = \lambda$. For all $t < a_c$, we get $\Phi'\left(b + \Sigma_t\right) > \lambda$ and $\Sigma_t < \Sigma_{a_c}$, while for $t > a_c$, we get $\Phi'\left(b + \Sigma_t\right) < \lambda$ and $\Sigma_t > \Sigma_{a_c}$. Therefore, for any $t \neq a_c$ $\left[\Phi'\left(b + \Sigma_t\right) - \lambda\right]\Sigma_t < \left[\Phi'\left(b + \Sigma_t\right) - \lambda\right]\Sigma_{a_c}$. Hence we get from (23):

$$Z_a = \int_a^{a_1}\left[\lambda - \Phi'\left(b + \Sigma_t\right)\right]\Sigma_t f\left(t\right) dt < \int_a^{a_1}\left[\lambda - \Phi'\left(b + \Sigma_t\right)\right]\Sigma_{a_c}\cdot f(a)da$$

[12] The difference between $(t - a)$and da tends to 0 as da tends to 0 is therefore neglected.

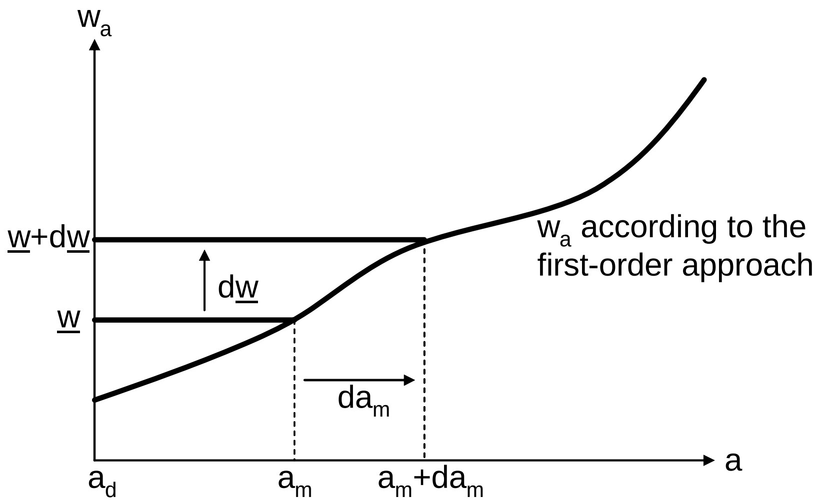

Figure 4: Covariations of a_m and of minimum wage $\underline{w}$.

However, given equation (21)

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

where $\mathbb{E}_f$ is the expectation operator under distribution f for t, $\Phi'_t = \Phi'(b + \Sigma_t)$ for $t \geq a_d$ and $\Phi'_t = \Phi'(b + d)$ otherwise. Hence Φ'_t is non-increasing and decreasing in t over $[a_d, a_1]$. From Lemma 1 this implies that $\mathbb{E}_f [\Phi'_t | t \geq a]$ decreases in a over $[a_d, a_1]$ so $\mathbb{E}_f [\Phi'_t | t \geq a] < \mathbb{E}_f [\Phi'_t]$. Hence $Z_a > 0$.

B.3 Optimality condition with respect to the minimum wage $\underline{w}$

Consider a variation of the minimum wage of $d\underline{w}$. This implies an increase in the wage, but also an increase in the amount of types for whom the minimum wage is relevant, as illustrated in Figure 4.

Hence, we calculate the direct effects of a variation in a_m at a given minimum wage $\underline{w}$, and then the direct effects of a variation in the minimum wage $\underline{w}$ for a given a_m.

Effect of a_m at given $\underline{w}$ According to equation (30), $\partial \Sigma_t / \partial a_m = 0$ for all $a \in [a_d, a_1]$. From (20)

$$\frac{\partial b}{\partial a_m} = (Y_{a_m} - \Sigma_{a_m}) f(a_m) - (Y_{a_m} - \Sigma_{a_m}) f(a_m) - \int_{a_d}^{a_m} \frac{\partial \Sigma_a}{\partial a_m} f(a)\, da - \int_{a_m}^{a_1} \frac{\partial \Sigma_a}{\partial a_m} f(a)\, da = 0$$

Finally, from (19) $\partial \Omega / \partial a_m = 0$. Hence we can concentrate on the direct effect of $\underline{w}$ for a given a_m for all variables of interest Σ_t, b, Ω.

Effect of $\underline{w}$ for a given a From (30) and $\underline{w} \equiv w_{a_d}$, we get for all $a \in [a_d, a_m]$:

$$\begin{aligned}
\frac{\partial \Sigma_a}{\partial \underline{w}} &= \Sigma_a \left\{ \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta} \right) \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}} \right) + \frac{1-\beta}{\beta} \int_{a_d}^{a} \frac{dt}{(t-\underline{w})^2} \right\} \\
&= \Sigma_a \left\{ \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta} \right) \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}} \right) + \frac{1-\beta}{\beta} \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}} \right) \right\} \\
&= \frac{1-\gamma}{\gamma} \Sigma_a \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}} \right) > 0
\end{aligned}$$

From (30), we have for all $a \in [a_m, a_1]$

$$\begin{aligned}
\frac{\partial \Sigma_a}{\partial \underline{w}} &= \Sigma_a \left\{ \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta} \right) \left(\frac{1}{a_d - \underline{w}} \right) + \frac{1-\beta}{\beta} \int_{a_d}^{a_m} \frac{dt}{(t-\underline{w})^2} \right\} \\
&= \Sigma_a \left\{ \left(\frac{1-\gamma}{\gamma} - \frac{1-\beta}{\beta} \right) \left(\frac{1}{a_d - \underline{w}} \right) + \frac{1-\beta}{\beta} \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a_m - \underline{w}} \right) \right\} \\
&= \Sigma_a \left\{ \frac{1-\gamma}{\gamma} \frac{1}{a_d - \underline{w}} - \frac{1-\beta}{\beta} \frac{1}{a_m - \underline{w}} \right\}
\end{aligned}$$

From (20) we find

$$\frac{\partial b}{\partial \underline{w}} = \int_{a_d}^{a_m} \left(\frac{\partial Y_t}{\partial w_t} (\underline{w}) - \frac{\partial \Sigma_t}{\partial \underline{w}} \right) f(t)\, dt - \int_{a_m}^{a_1} \frac{\partial \Sigma_t}{\partial \underline{w}} f(t)\, dt$$

Hence

$$\begin{aligned}
\frac{\partial b}{\partial \underline{w}} &= \int_{a_d}^{a_m} \left\{ \frac{\partial Y_a}{\partial w_a} (\underline{w}) - \frac{1-\gamma}{\gamma} \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}} \right) \Sigma_a \right\} f(a)\, da \\
&\quad + \left(\frac{1-\beta}{\beta} \frac{1}{a_m - \underline{w}} - \frac{1-\gamma}{\gamma} \frac{1}{a_d - \underline{w}} \right) \int_{a_m}^{a_1} \Sigma_a \cdot f(a) \cdot da
\end{aligned}$$

Finally, from (19) and (21):

$$\frac{\partial \Omega}{\partial \underline{w}} = \lambda \frac{\partial b}{\partial \underline{w}} + \int_{a_d}^{a_m} \Phi'(\Sigma_a + b) \frac{\partial \Sigma_a}{\partial \underline{w}} f(a)\, da + \int_{a_m}^{a_1} \Phi'(\Sigma_a + b) \frac{\partial \Sigma_a}{\partial \underline{w}} f(a)\, da$$

After some manipulations, one then gets

$$\begin{aligned}
\frac{\partial \Omega}{\partial \underline{w}} &= \int_{a_d}^{a_m} \left\{ \lambda \frac{\partial Y_a}{\partial w_a} (\underline{w}) + (\Phi'(\Sigma_a + b) - \lambda) \Sigma_a \frac{1-\gamma}{\gamma} \left(\frac{1}{a_d - \underline{w}} - \frac{1}{a - \underline{w}} \right) \right\} f(a)\, da + \\
&\quad + \left(\frac{1-\gamma}{\gamma} \frac{1}{a_d - \underline{w}} - \frac{1-\beta}{\beta} \frac{1}{a_m - \underline{w}} \right) \left(\int_{a_m}^{a_1} (\Phi'(\Sigma_a + b) - \lambda) \Sigma_a f(a)\, da \right)
\end{aligned}$$

Together with (23), we obtain (24).

B.4 Optimality condition with respect to the threshold a_d

We consider now a variation δa_d for the threshold a_d, keeping the function $a \mapsto w_a$ unchanged. Using Proposition 1, we know that there is bunching of wages at the bottom of the skill distribution. Therefore, a marginal change of the threshold a_d keeps unchanged the level of the lowest wage w_{a_d}. Hence, from (30), we get for $a \in [a_d, a_1]$

$$\frac{\delta\Sigma_a}{\delta a_d} = -\frac{1-\gamma}{\gamma}\left(\frac{1}{a_d - w_{a_d}} - \frac{\dot{\kappa}_{a_d}}{\kappa_{a_d}}\right)\Sigma_a$$

One then gets from the budget constraint (20)

$$\frac{\delta b}{\delta a_d} = -\left(Y_{a_d} - \Sigma_{a_d}\right) - \int_{a_d}^{a_1} \frac{\delta\Sigma_a}{\delta_{a_d}} f(a)\, da$$

Since $\Sigma_{a_d} = d$, one gets

$$\frac{\delta b}{\delta a_d} = -\left(Y_{a_d} - d\right) + \frac{1-\gamma}{\gamma}\left(\frac{1}{a_d - w_{a_d}} - \frac{\dot{\kappa}_{a_d}}{\kappa_{a_d}}\right)\int_{a_d}^{a_1} \Sigma_a f(a)\, da$$

Finally, from the government's objective one gets $\delta\Omega = \lambda\ \delta b + \int_{a_d}^{a_1} \delta\Sigma_a\ f(a)\ \ da$. Taking equation (23) into account, we then obtain (27).

C Proof of Proposition 2

To prove Proposition 2, we apply the envelope theorem and prove that the partial derivative of Ω with respect to $(1-\beta)/\beta$ for a given function $a \mapsto w_a$ is negative.

From equation (30), one gets for all $a \in [a_d, a_m]$

$$\frac{\partial\Sigma_a}{\partial\frac{1-\beta}{\beta}} = 0$$

From equation (30), one has, for a given function $a \mapsto w_a$ that for all $a \in [a_m, a_1]$

$$\frac{\partial\Sigma_a}{\partial\frac{1-\beta}{\beta}} = \left\{-\log(a - w_a) + \log(a_d - w_{a_d}) + \int_{a_d}^{a}\left(\frac{1}{x - w_x} - \frac{\dot{\kappa}_x}{\kappa_x}\right)dx\right\}\Sigma_a$$

So:

$$\begin{aligned}\frac{\partial\Sigma_a}{\partial\frac{1-\beta}{\beta}} &= \left\{\log\left(\frac{a - w_{a_d}}{a - w_a}\cdot\frac{a_d - w_{a_d}}{a - w_{a_d}}\right) + \int_{a_d}^{a}\frac{dx}{x - w_x}\right\}\Sigma_a \\ &= \left\{\log\left(\frac{a - w_{a_d}}{a - w_a}\right) + \int_{a_d}^{a}\left\{\frac{1}{x - w_x} - \frac{1}{x - w_{a_d}}\right\}dx\right\}\Sigma_a\end{aligned}$$

Since there is no more bunching in the right-neighborhood of a_m, one has together with (18) that $w_a > w_{a_m} = w_{a_d}$. This induces that $a - w_a < a - w_{a_d}$, so the first term in the bracket is positive. Furthermore, for all $x \in (a_m, a]$, one has $w_{a_d} = w_{a_m} < w_x < x$, so

$1/(x-w_x) > (1/(x-w_{a_d})) > 0$. Therefore the second term in the bracket is positive too and

$$\frac{\partial \Sigma_a}{\partial \frac{1-\beta}{\beta}} > 0$$

We now prove that the function $a \mapsto \partial \Sigma_a / \partial((1-\beta)/\beta)$ is increasing. This function is the product of two positive functions, and we already know from Lemma 1 that $a \mapsto \Sigma_a$ is increasing. We now proove that $a \mapsto \log\left(\frac{a-w_{a_d}}{a-w_a}\right) + \int_{a_d}^{a}\left\{\frac{1}{x-w_x} - \frac{1}{x-w_{a_d}}\right\} dx$ is increasing too. Its partial derivative with respect to w_a holding a constant is

$$\left.\frac{\partial\left\{\log\left(\frac{a-w_{a_d}}{a-w_a}\right) + \int_{a_d}^{a}\left\{\frac{1}{x-w_x} - \frac{1}{x-w_{a_d}}\right\} dx\right\}}{\partial w_a}\right|_a = \frac{1}{a-w_a} > 0$$

Moreover, its partial derivative with respect to a holding w_a constant is

$$\left.\frac{\partial\left\{\log\left(\frac{a-w_{a_d}}{a-w_a}\right) + \int_{a_d}^{a}\left\{\frac{1}{x-w_x} - \frac{1}{x-w_{a_d}}\right\} dx\right\}}{\partial a}\right|_{w_a} = \frac{1}{a-w_{a_d}} - \frac{1}{a-w_a} + \frac{1}{a-w_a} - \frac{1}{a-w_{a_d}} = 0$$

Since wage w_a is nondecreasing in a (See 18), the function $a \mapsto \partial \Sigma_a / \partial((1-\beta)/\beta)$ is the product of two positive function, one being increasing in a, and the other one being nondecreasing. Hence, it is increasing in a.

From the budget constraint (20), one gets

$$\frac{db}{d\frac{1-\beta}{\beta}} = \frac{\partial b}{\partial \frac{1-\beta}{\beta}} = -\int_{a_m}^{a_1} \frac{\partial \Sigma_a}{\partial \frac{1-\beta}{\beta}} f(a)\, da < 0$$

Finally, from the social objective (19), we find

$$\begin{aligned}
\frac{d\Omega}{d\frac{1-\beta}{\beta}} &= \lambda \frac{\partial b}{\partial \frac{1-\beta}{\beta}} + \int_{a_m}^{a_1} \Phi'(b+\Sigma_a) \cdot \frac{\partial \Sigma_a}{\partial \frac{1-\beta}{\beta}} \cdot f(a)\, da \\
&= \int_{a_m}^{a_1} \left[\Phi'(b+\Sigma_a) - \lambda\right] \cdot \frac{\partial \Sigma_a}{\partial \frac{1-\beta}{\beta}} \cdot f(a)\, da
\end{aligned}$$

Since the function $a \mapsto \Phi'(\Sigma_a + b)$ is decreasing from the concavity of $\Phi(.)$ and Lemma 1, two cases are possible:

1. $\lambda \geq \Phi'(b+\Sigma_{a_m})$. Then for all $a > a_m$, $\lambda \geq \Phi'(b+\Sigma_a)$ where this inequality is strict for at least some of these types. Hence, one has

$$\left[\Phi'(b+\Sigma_a) - \lambda\right] \cdot \frac{\partial \Sigma_a}{\partial \frac{1-\beta}{\beta}} \cdot f(a) \leq 0$$

and this inequality holds strictly for at least some of these types. This implies

$$\frac{d\Omega}{d\frac{1-\beta}{\beta}} < 0$$

2. Otherwise, $\lambda < \Phi'(b+\Sigma_{a_m})$. From equation (21) and Lemma 1, we know that there exists a $a_c \in (a_m, a_1)$ such that $\lambda = \Phi'(b+\Sigma_{a_c})$. Furthermore, the function $\partial\Sigma_a/\partial((1-\beta)/\beta)$ is positive and increasing.

   - For $a \in (a_m, a_c)$, we have $\Phi'(b+\Sigma_a) \geq \lambda$ and thus
   $$[\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_a}{\partial\frac{1-\beta}{\beta}} \leq [\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_{a_c}}{\partial\frac{1-\beta}{\beta}}$$
   - for $a \in (a_c, a_1)$, we have $\Phi'(b+\Sigma_a) \leq \lambda$ and thus
   $$[\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_a}{\partial\frac{1-\beta}{\beta}} \leq [\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_{a_c}}{\partial\frac{1-\beta}{\beta}}$$

   Hence for all a
   $$[\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_a}{\partial\frac{1-\beta}{\beta}} \leq [\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_{a_c}}{\partial\frac{1-\beta}{\beta}}$$
   and therefore
   $$\begin{aligned}\frac{d\Omega}{d\frac{1-\beta}{\beta}} &\leq \int_{a_m}^{a_1}[\Phi'(b+\Sigma_a)-\lambda]\cdot\frac{\partial\Sigma_{a_c}}{\partial\frac{1-\beta}{\beta}}\cdot f(a)\,da\\ \frac{d\Omega}{d\frac{1-\beta}{\beta}} &\leq \frac{\partial\Sigma_{a_c}}{\partial\frac{1-\beta}{\beta}}\left\{\int_{a_m}^{a_1}[\Phi'(b+\Sigma_a)-\lambda]\,f(a)\,da\right\}\\ \frac{d\Omega}{d\frac{1-\beta}{\beta}} &\leq \frac{\partial\Sigma_{a_c}}{\partial\frac{1-\beta}{\beta}}(1-F(a_m))\left\{\mathbb{E}[\Phi'(b+\Sigma_a)\,|a \geq a_m]-\lambda\right\}\end{aligned}$$
   From Lemma 1, we now that function $a \mapsto \Phi'(\Sigma_a+b)$ is decreasing in a. Hence, given (21), one has $\lambda > \mathbb{E}[\Phi'(b+\Sigma_a)\,|a \geq a_m]$, therefore,
   $$\frac{d\Omega}{d\frac{1-\beta}{\beta}} < 0$$

Hence, in both cases, one finds
$$\frac{d\Omega}{d\beta} > 0$$