Optimal minimum wage, income taxation and redistribution

Martin Van der Linden

April 18, 2013

1 Introduction

1.1 Conventional wisdom : the inefficiency of minimum wage policies

The standard treatment of minimum wage policies in terms of Marshalian surupluses, common in introductory economics textbooks, helped to popularize the idea that minimum wage policies generate so-called "*deadweight*" welfare loss and are, as a consequence, inefficient.

Proposition 1. *Given a perfectly competitive (untaxed) labor market, if workers' utility functions is as defined in* (2) *(see Frame I), worker are homogeneous, profits are fully redistributed to workers and enter their utility functions in a linear manner, and the social planner has a classical utilitarian objective, then a minimum wage policy is suboptimal.*

The classical illustration supporting the last proposition is reproduced in figure 1.[1]

1.2 Beyond the standard argument

The argument in proposition 1 is "standard" in the sense that it does not account for the specific characteristics of the labor market (in fact, introductory textbooks tend to apply the same analysis to virtually every problem of quantity or price rationing). However, as it is clear from proposition 1, the validity of the standard argument against minimum wage policies rely on very specific assumptions which may not be appropriate to analyze the labor market. In fact, it is not hard to think of situations in which the argument breaks down as illustrated in the following example.

Example 1 (Imperfect profit redistribution)**.** *Assume that profits are not fully redistributed to workers. For instance, it could be that the production sector is composed of internationally owned companies, and the profits are repatriated by the owners of the firms. Suppose, as it is the case in many industries, that despite relocation possibilities on behalf of the firms, the demand for domestic labor is not perfectly wage-elastic. Suppose also that the supply elasticity is*

[1] We will see later, when we consider Lee and Saez (2012) that the inefficiency can actually be more sever than depicted, depending on the effectiveness of the rationing process.

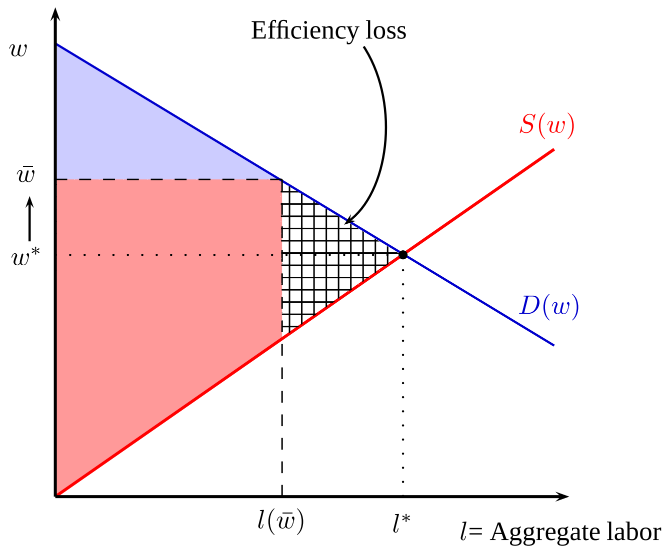

Figure 1: The usual argument for the inefficiency of minimum wage policies

positive. In this case, a minimum wage can be welfare improving, depending on the efficiency of the rationing process (more on this in section 3.2.2). This is illustrated in figure 2 where for simplicity we depict the case of a complete absence of profit redistribution to domestic workers.

A substantive amount of research efforts have been invested in identifying, from a theoretical point of view, when a minimum wage policy is efficient. Part of the literature, coming from labor economics, focuses on market imperfections (such as monopsonistic competition, efficiency wages and search models, see the introduction of Lee and Saez (2012) for a list of references) and identifies relevant settings in which a minimum wage policy allows for a pure efficiency gain.

In this notes, we are interested in situations in which a minimum wage is useful policy tool because of its redistributive implications. If the social planner values redistribution (or alternatively if it is assumed that people have decreasing marginal utility), a minimum wage policy may, in some settings, help redistribution by correcting the market distortions caused by a pre-existing tax system (Guesnerie and Roberts (1987)), or by overcoming issues of asymmetric information (Marceau and Boadway (1994), Boadway and Cuff (2001), Lee and Saez (2012)). This note discusses these issues based focusing on Lee and Saez (2012), Boadway and Cuff (2001) and Guesnerie and Roberts (1987).

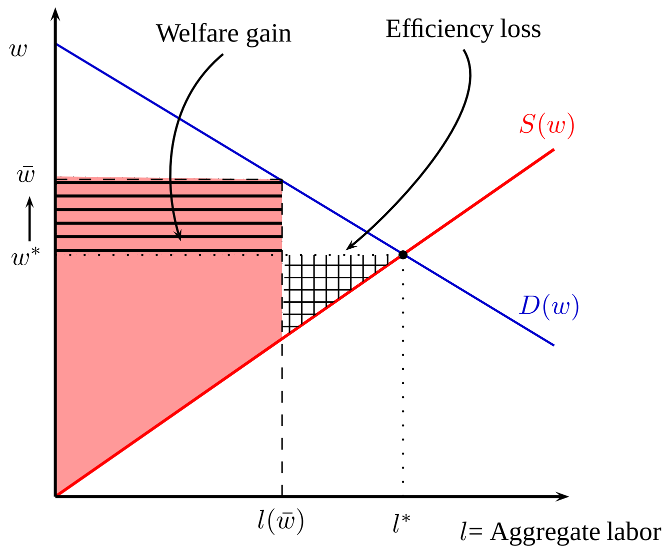

Figure 2: Minimum wage with imperfect profit redistribution

Frame I : The standard Marshalian surplus framework

The labor demand. The firms, indexed by $j = 1, \ldots, J$ produce some consumption good using only labor as an input. Firms are endowed with production functions $F_j(l_j)$ where l_j is the quantity of labor used by the firm j. The price of the consumption good p is usually normalized to $p = 1$. By the perfect competition assumption, firm take prices (p, w) as given and maximize profits by solving

$$\max_{l_j} F_j(l_j) - wl_j$$

$F_j(l_j)$ is assumed to be increasing and concave $(\partial^2 F_j / \partial^2 l_j \leq 0)$ for every firm so that the optimality condition is

$$\left.\frac{\partial F_j}{\partial l_j}\right|_{l_j = l_j^*} = w,$$

which, implicitly defines the optimal quantity of labor l_j^* as a negative function of the wage

$$\frac{\partial l_j^*(w)}{\partial w} < 0. \tag{1}$$

1 holding for every firm j yields the usual downward sloping aggregate demand function $D(w) = \sum_{j=1}^{J} l_j^*$ depicted in figure 1 (where linearity of $D(w)$ is assumed for illustration convenience).

The standard Marshalian surplus framework (continued)

The labor supply. Every worker is assumed to have the same skill. Despite labor being an homogeneous good, individuals are not homogeneous. Following the treatment of Lee and Saez (2012), individuals are assumed to differ in their disutility of labor. As noted by Lee and Saez (2012), these discrepancies in labor disutility may be the interpreted as differences in the level of effort, taste for work/leisure, or in the cost of skill acquisition through education. Formally, for every workers $h = 1, ..., H$, utility is assumed to take the linear form

$$u_h = \begin{cases} w - \theta_h, & \text{if } l_h = 1 \\ 0, & \text{if } l_h = 0 \end{cases}, \tag{2}$$

where $\theta_h > 0$ is the disutility of labor parameter and is distributed across the population according tho some cumulative distribution $F(\theta)$.
From 2 an individual h supplies her labor on the market if $w \geq \theta_h$. Given $F(\theta)$, the number of individuals supplying labor on the market is an increasing function of the wage $S(w)$, which give the usual upward sloping labor supply depicted in 1 (where linearity and continuity of $S(w)$ are again assumed for illustration convenience).
The Marshalian surplus welfare analysis. The standard case against minimum wage policy relies on a very specific concept of social welfare, namely Marshalian surpluses.
Given the specification of the utility function, we are allowed a money metric interpretation of utility. Now assume that profits are eventually redistributed to workers. Due to the linearity in income of the utility functions, *if* one assumes that the social planner seeks to maximize a classical utilitarian utility function (sum of the utilities), one can meaningfully identify social welfare with the sum of profits and utilities. In this case, a minimum wage policy is clearly suboptimal as depicted in figure 1.[a]

[a]The form of the utility functions allows for the convenient representation of the sum of utilities as the area between the supply curve and a horizontal line placed at the level of the prevailing wage (the red region on 1). The profit of the firms is represented similarly as the area between the demand curve and the same horizontal line (the blue region on 1).

2 A minimum wage to correct markets distortions (Guesnerie and Roberts, 1987)

Guesnerie and Roberts (1987) study the case in which a minimum wage has no implications on the incentive constraint. The article starts with consider a setting with intensive margin response to minimum wage policies : following the introduction of a minimum wage, every worker for which the minimum wage constraint is binding uniformly reduces the number of hours she works. Therefore, there is no involuntary *un*employment *per se*, but rather a situation of *under*employment : some would like to work more hours than they actually do given the prevailing wage. Guesnerie and Roberts (1987) find that a minimum wage policy can be optimal when the taxation is linear in income. A minimum wage policy, inducing rationing in the labor supply, may help correcting for market distortions induced by a linear labor tax scheme. However, this is not the case when a non-linear taxation scheme is available and extensive margin response are allowed (a similar result is obtained by and Allen (1987)).

Frame II : An introduction to the general theory of rationing

Forthcoming notes on Guesnerie and Roberts (1984) and its relations with Guesnerie and Roberts (1987).

The fact that workers react on the intensive margin (adapting the number of hours worked) rather than the extensive one (leaving the labor market) is essential to the ineffectiveness of minimum wage policies. In effect, if workers react on the extensive margin, then a minimum wage can help the government to identify who is a low skill. This relaxes the incentive compatibility constraint and allows for a higher level of redistribution triggering higher social welfare (when the government values redustribution or the marginal utility of consumption is decreasing). Guesnerie and Roberts (1987) assume away this possibility, while noticing that this assumption is crucial for a minimum wage to be useless under a non-linear taxation of income. They warn that their result should not be expected to hold if this assumption is relaxed, a possibility that we will discuss in the next sections.

2.1 Linear taxation[2]

Consider a 2-commodity world (consumption and labor) in which labor is taxed at a rate t to fund a lump-sum grant a paid to every individual. The consumption level of an individual is

$$c = a + (1 - t)y, \tag{3}$$

where y is the income.

Guesnerie and Roberts (1987) considers a simple model with two type of workers, a low skilled (type 1) and a high skilled (type 2), of which there are respectively n_1 and n_2 in the population. The consumption good is produced according to the production function

$$Y = f(n_1 l_1) + g(n_2 l_2),$$

where l_1 and l_2 are respectively the quantity of labor supplied by any low and any high skilled. The fact that type 1 has lower skills than type 2 is embodied in the assumption that, at an equilibrium,

$$w_1 = f'(n_1 l_1) < g'(n_2 l_2) = w_2, \tag{4}$$

[2]This part of the notes is based on lecture notes by Stéphane Gauthier

where l_i, $i = 1, 2$, is, for now on, the labor supply maximizing the worker's utility (see below) under constraint (3).

It is assumed that the government taxes the profits at 100%, so that the only source of income for the individuals is their labor income

$$y_i = w_i l_i. \tag{5}$$

All individuals have the same preferences over consumption and labor described by some common utility function

$$u = u(c, l).$$

The budget constraint of the government is

$$\underbrace{f(n_1 l_1) - w_1 n_1 l_1}_{\pi_1} + \underbrace{g(n_2 l_2) - w_2 n_2 l_2}_{\pi_2} + \underbrace{t(w_1 n_1 l_1 + w_2 l_2 n_2)}_{\text{labor tax revenue}} - \underbrace{a(n_1 + n_2)}_{\text{lump-sum grant}} \geq R, \tag{6}$$

We will assume that the government want to maximize a Social Welfare Function (SWF) which is increasing in a, so that at an optimum, (6) will be binding. We also assume that the government must fund no exogenous spending through labor taxation, hence $R = 0$. With these assumptions, (6) can be rewritten as

$$f(n_1 l_1) + g(n_2 l_2) + (t - 1)(w_1 n_1 l_1 + w_2 l_2 n_2) - a(n_1 + n_2) = 0. \tag{7}$$

2.1.1 Fixed tax rate

For the moment, let us assume that the labor tax rate has been predetermined, $t = \bar{t}$. The government seeks to increase social welfare through a minimum wage policy only, for a given $t = \bar{t}$ that is without re-optimizing with respect to t.

By a minimum wage policy, we mean a policy consisting in raising the wage which is smaller in equilibrium, i.e w_1 by assumption. We consider an infinitesimal increase dw_1, so that, as $w_1 < w_2$, the resulting $w_1 + dw_1$ is effectively the minimum wage in the economy.

Using (3) and (5), the government maximization problem can be written as

Problem 1. $max_{w_1} \quad W\Big(n_1 v(a, (1-\bar{t})w_1),\ n_2 v(a, (1-\bar{t})w_2)\Big), \quad s.t. \quad (7),$

where the $v(\cdot)$ are the indirect utility functions. The Lagrangian of the problem is

$$\begin{aligned}\mathcal{L} =& W\Big(n_1 v(a, (1-\bar{t})w_1),\ n_2 v(a, (1-\bar{t})w_2)\Big) \\ &+ \lambda\Big[f(n_1 l_1) + g(n_2 l_2) + (\bar{t} - 1)(w_1 n_1 l_1 + w_2 l_2 n_2) - a(n_1 + n_2)\Big].\end{aligned} \tag{8}$$

As suggested by Guesnerie and Roberts (1987), to determine whether a minimum wage is desirable, one must compare the effect of the policy on the welfare of the low skill (the group that is affected by the policy) with the impact it has on the budget constraint of the government.

Formally, this implies considering the first order conditions associated with problem 1. The question is whether or not it is optimal to increase w_1 by imposing a minimum wage. If w_1 is optimal, we must have

$$\tfrac{\partial \mathcal{L}}{\partial w_1} = 0,$$

and if w_1 is suboptimal and a minimum wage increases social welfare, it must be the case that

$$\tfrac{\partial \mathcal{L}}{\partial w_1} > 0.$$

Differentiating the Lagrandian with respect to w_1, we see that the optimality of w_1 depends on two elements:

The change in the objective function due to the change in the indirect utility of the households following the introduction of the minimum wage. For any function $g(x_1, \ldots, x_G)$, let us adopt the usual notation $\partial g / \partial x_j = g'_j$, $j = 1, \ldots, G$. As only the low-skilled are impacted by a minimum wage policy, we have

$$\begin{aligned} dW &= W'_1 n_1 \frac{\partial v(\cdot)}{\partial w_1} dw_1 \\ &= W'_1 n_1 (1 - \bar{t}) l_1 v'_1 dw_1, \end{aligned} \tag{9}$$

where the second line flows from the standard envelope argument associated with the differentiation of indirect utility functions (see Frame III), and we assumed differentiability of the the SWF.

Frame III : Differentiating the indirect utility function with respect to w_1

The worker maximization problem is

$$max_{\ell_1} \;\; u\Big(a + (1-\bar{t})w_1\ell_1, \; \ell_1)\Big),$$

where the constraint (3) has been replaced in the utility function. The FOC associated with this problem implies that, at an optimal choice of labor supply $\ell_1 = l_1$,

$$u_1'(l_1)(1-\bar{t})w_1 + u_2'(l_1) = 0. \tag{10}$$

The last equation (10) implicitly defines l_1 as a function of w_1, $l_1 = l_1(w_1)$. The indirect utility function is

$$v\Big(a, \; (1-t)w_1\Big) = u\Big(a + (1-\bar{t})w_1 l_1, \; l_1\Big).$$

Therefore,

$$\begin{aligned}
\frac{\partial v\Big(a, \; (1-t)w_1\Big)}{\partial w_1} &= \frac{\partial u\Big(a + (1-\bar{t})w_1 l_1, \; l_1)\Big)}{\partial w_1}, \\
&= u_1'(1-\bar{t})l_1 + u_1'(1-\bar{t})w_1\frac{\partial l_1}{\partial w_1} + u_2'\frac{\partial l_1}{\partial w_1}, \\
&= u_1'(1-\bar{t})l_1 - u_2'\frac{\partial l_1}{\partial w_1} + u_2'\frac{\partial l_1}{\partial w_1}, \\
&= u_1'(1-\bar{t})l_1, \\
&= v_1'(1-\bar{t})l_1,
\end{aligned}$$

where in the third line we have use the fact that, at an optimal choice of labor supply $l_1 = l_1$, the FOC (10) must be satisfied, and hence $u_1'(l_1)(1-\bar{t})w_1 = -u_2'(l_1)$.

Although standard, the envelope argument is particularly important in our case. We consider an infinitesimal increment in the wage of the low-skilled, starting from an equilibrium situation at which (10) holds. Thus, for a given wage, the low-skilled are indifferent to a small decrease of their labor, as it will be exactly off-set by the benefit of additional leisure. This means that, to a first-order approximation, the low-skilled are impacted by the policy *only* through the change in the wage. The impact of the fall in labor demand can be disregarded.

We want to rewrite (9) in order to get rid of the dw_1 and reformulate the condition in terms of elasticities, as it is done by Guesnerie and Roberts (1987). The equilibrium condition (4) defines a direct relation between a change in the (low-skilled) wages and a change in the quantity of labor offered (by the low-skilled),

$$\frac{dw_1}{dl_1} = f''(n_1 l_1) n_1. \tag{11}$$

Through equation 11, we see that a change in w_1 is proportional – to a factor defined in (11) – to a change in l_1. In terms of elasticities, if we define

$$\epsilon_1 = -\frac{n_1 l_1 f''}{f'}, \quad \text{the elasticity of the equilibrium wage with respect to the quantity of labor,}$$
$$\hat{w}_1 = \frac{dw_1}{w_1}, \quad \text{the percentage change in the wage chosen by the government,}$$
$$\hat{l}_1 = \frac{dl_1}{l_1}, \quad \text{the percentage change in } \hat{l}_1 \text{ induced by the choice of } \hat{w}_1,$$

(11) can be rewritten as

$$\hat{w}_1 = -\epsilon_1 \hat{l}_1. \tag{12}$$

With these notations, (9) can be written as

$$dW = -W_1' n_1 w_1 l_1 (1-\bar{t}) v_1' \epsilon_1 \hat{l}. \tag{13}$$

The change in the government revenue due to the change in wages, profits, and number of hours worked. Again, we have to consider only the part of (7) which is impacted by a change in w_1 (only the low-skilled are impacted). We have

$$\frac{\partial[f(n_1 l_1) - (1-\bar{t}) n_1 w_1 l_1]}{\partial w_1} = n_1 f' \frac{\partial l_1}{\partial w_1} - n_1 (1-\bar{t})(l_1 + \frac{\partial l_1}{\partial w_1}).$$

Eliminating the dw_1, using (11) and reformulating in terms of elasticities, this yields,

$$d[f(n_1 l_1) - (1-\bar{t}) n_1 w_1 l_1] = n_1 w_1 l_1 [1 - (1-\bar{t})(1-\epsilon_1)] \hat{l}_1. \tag{14}$$

All in all, combining (14) and (13), an increase in the wage of the low-skilled induces the following change in the value of the Lagrangian:

$$d\mathcal{L} = \underbrace{n_1 w_1 l_1}_{>0} \underbrace{\Big[-W_1'(1-\bar{t}) v_1' \epsilon_1 + \lambda \big(1 - (1-\bar{t})(1-\epsilon_1)\big) \Big]}_{\lesseqgtr 0} \underbrace{\hat{l}_1}_{<0, \text{ as we consider } dw_1 > 0}.$$

So the introduction of a minimum wage is desirable if the term under bracket is strictly negative, that is

$$\lambda\big(1-(1-\bar{t})(1-\epsilon_1)\big) < W_1'(1-\bar{t})v_1'\epsilon_1. \tag{15}$$

Let us assume that the social planner cares at least a little for the low-skilled (typically a social planner is assumed to care at least as much for the low-skilled that for the high-skilled), so that $W_1' > 0$. Also, let us take the utility of consumption to the low-skilled as *numéraire*, $v_1' = 1$, and define $\tilde{\lambda} = {}^{\lambda}\!/_{W_1'}$. Then (15) can be rewritten as

$$\frac{1-\tilde{\lambda}}{\tilde{\lambda}} > \frac{\bar{t}}{\epsilon_1(1-\bar{t})}, \tag{16}$$

which is the condition obtained by Guesnerie and Roberts (1987).

Equation (16) tells us that a minimum wage is more likely to increase social welfare:

The smaller $\tilde{\lambda}$: the larger the weight given by the social planner to the low-skilled in equilibrium, i.e. the larger W_1'.

The larger ϵ_1 : ϵ_1 is, *in negative terms*, the "increase" in equilibrium low-skilled labor following an increase in the wage. The larger this "increase" – that is actually the smaller the decrease – the less underemployment is generated by the minimum wage policy, and the more likely the desirability of the policy.

The smaller $\bar{t}$: the smaller the *direct* tax loss induced by the reduction in the labour supply. Beware, although the direct effect on the government revenue is negative, the *total* effect has to be positive for the minimum wage to be desirable (in order to compensate the welfare loss suffered by the low-skilled with an increase in the lump-sum grant).

2.1.2 Optimal tax rate

All the previous argument relied on the assumption that the tax rate was fixed, $t = \bar{t}$. To keep the analysis of an optimally chosen t as simple as can be, Guesnerie and Roberts (1987) considers a very extreme case build around two assumption:

The government cares only about the low-skilled (so-called Rawlsian SWF), so that the SWF $W\Big(n_1 v(a,(1-\bar{t})w_1),\ n_2 v(a,(1-\bar{t})w_2)\Big) = W\Big(n_1 v(a,(1-\bar{t})w_1))\Big)$.

The production of the low-skilled is "negligible". Guesnerie and Roberts (1987) description of the assumption in this sentence "minimum wage legislation has direct effects only on those whose production value is small so that the low-skilled input into production can be ignored".

As for the first argument in section 2.1.1, Guesnerie and Roberts (1987) is extremely terse on the description of the assumptions and arguments leading to their results. In this second case (optimal tax rate), I could not reconstruct the argument leading to the result presented in the article, mostly due to the fact that it is not clear how the negligibility assumption is to be interpreted.[3]

According to Guesnerie and Roberts (1987), the two additional assumption should lead to the following analog to (16)

$$\frac{n_2}{n_1}(1 - tG' l_a^2) > \frac{1}{\eta \epsilon_1}, \tag{17}$$

where $l^2(w,a)$ is the labor supply function of the high-skilled and $\eta = (1-t)G'\frac{l_w^2}{l^2}$ is a parameter measuring the disincentive to work induced by labor taxation. When (17) is satisfied, a minimum wage is desirable. Thus, a minimum wage is more likely to be deisrable if

$^{n_2}/_{n_1}$ **is large :** there are relatively more high-skilled than low-skilled.

η **is large :** the disensentive to work due to taxation is small.

ϵ **is large :** minimum wage legislation has little effect on the demand for labor.

3 A Minimum wage as a tool to relax the incentive constraints?

In a sense, a minimum wage policy can be viewed as introducing non-linearities in an otherwise linear taxation scheme. One might think that being able to impose a minimum wage does not allow the social planner to chose among a larger set of allocations (i.e. lists of consumption-income couples, see below) than if she were allowed to implement a non-linear taxation scheme. If this was to be true one could, so to speak "integrate" the desirable effect of a minimum wage policy into a suitably chosen non-linear taxation scheme.

3.1 The non-linear taxation model from Guesnerie and Roberts (1987)

Guesnerie and Roberts (1987) put this idea to the test by considering a similar framework as in the former section, with two types of skills. A difference with the former section is that we now allow for non-linear taxation in which the government choses the menu of consumption-income couples offered to the workers. The constraints defining a feasible menu are

Decentralized labor market : the wage is determined by the equilibrium of the profit maximizing firm so that wages equal marginal productivities given the total labor supply,

[3]The low-skilled input into production can be ignored but does it means that the low-skilled labor must be viewed as working zero hours and getting zero income? If not, how can we make sense of a positive labor supply and a positive wage on behalf of the low-skilled when their part of the production is arbitrarily small? Lack of details in the article prevented us from answering those questions and reconstructing the argument.

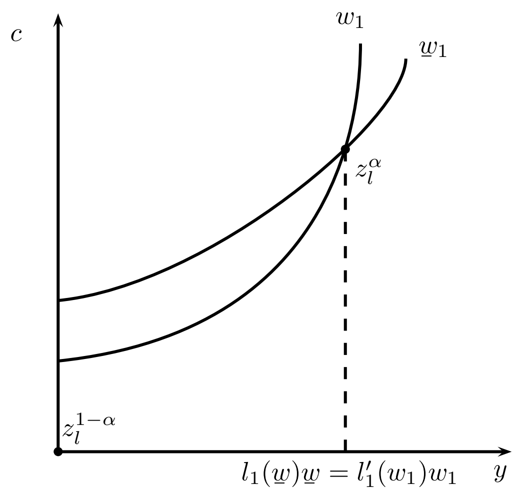

Figure 3: The set of incentive-compatible allocations is extended when a minimum wage can be enforced

Self selection constraint : given the wage, the agents chose teh consumption-income couple which is the best among those offered by the social planner,

Material balance constraint : total consumption can be produced from total labor.

Assume that the rationing is now organized on the on the extensive margin : if a minimum wage is enforced, given that the low-skilled labor supply exceeds the demand, a set of low-skilled workers is randomly selected to meet the demand, and the other workers are now *un*employed.

Consider the situation depicted in figure 3, where the two indifference curves refers to some wage w_1 and the minimum wage $\underline{w}_1$. Suppose that the government wants to implement an allocation in which a share α of the low-skilled get bundle z_l^{α} and the rest of low skilled is unemployed and get $z_l^{(1-\alpha)}$. Without a minimum wage, this allocation is not incentive compatible. The low-skilled which are supposed to receive $z_l^{(1-\alpha)}$ will prefer to chose z_l^{α}, although it lowers the wage to $w_1 < \underline{w}_1$ (as the low-skilled labor supply is increased) and forces them (as well asall the other low-skilled) to work $l_1(w_1) > l_1(\underline{w}_1)$.

When a minimum wage is enforceable, this allocation is feasible. As the wage cannot go lower than $\underline{w}_1$, there is a labor supply rationing for the low-skilled. Although the low-skilled bunched at $z_l^{(1-\alpha)}$ would want to move to z_l^{α}, they cannot as the wage is not allowed to go below $\underline{w}_1$. Therefore the random mechanism described above applies and the allocation is feasible (if α is precisely such that $f' = \underline{w}_1/\alpha n_1$).

The former argument shows that the set of feasible menus is larger when a minimum wage is available. Yet, the question remains as to whether any of these new possibilities are welfare improving. Guesnerie and Roberts (1987) shows that one can at least find a set of reasonable assumption under which they are not.

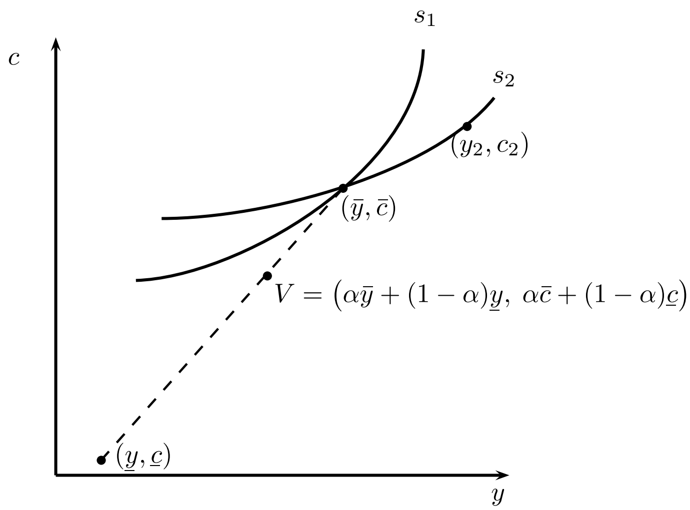

Figure 4: Case 1

Theorem 1 (Proposition p.496 (Guesnerie and Roberts, 1987)). *Assume that*

1. *Individuals have a concave utility functions over the income-consumption pairs,*
2. *The indifference curves in the* $(y, c) -$ *space satisfy single crossing,*
3. *The SWF is utilitarian,*

then the optimal income tax schedule does not involve a minimum wage.

Proof. The proof proceeds by showing that it is never optimal to treat equals unequally. Thus the argument also applies to other classes of rationing mechanisms than a minimum wage with random selection of those who work. Assume that through the rationing mechanism, some subset $(1-\alpha)n_1$ of the low-skilled are chosen to receive $(\underline{y}, \underline{c})$ and the other αn_1 low-skilled receive $(\bar{y}, \bar{c})$. Let $\bar{U} > \underline{U}$ be the corresponding utility levels. Consider replacing $(\underline{y}, \underline{c})$ and $(\bar{y}, \bar{c})$ in the tax schedule by $V = \big(\alpha\bar{y} + (1-\alpha)\underline{y},\ \alpha\bar{c} + (1-\alpha)\underline{c}\big)$. If the utility of the low-skilled at V is *lower* than $\bar{U}$, then by single crossing, the high skill will not chose V (see figure 4, bindingness of the incentive constraint is not necessary. The indifference curve are labeled s_1 for low-skilled and s_2 for high-skilled. Notice that when the low skill move to allocation V, their wage is impacted. In order to keep the illustation simple,[4] the figure does not account for this change. So the illustration should be viewed as a visualization tool only and does not represent the situation with full accuracy).

If the utility of the low-skilled at V is *higher* than $\bar{U}$, it might be the case that the high-skilled agent would want to change their bundle for V (see figure 5). In this case, lower the consumption level associated with V while maintaining the

[4] And as representing indifference curves for different w_1 would bring no help visualizing the preferences of the low skills

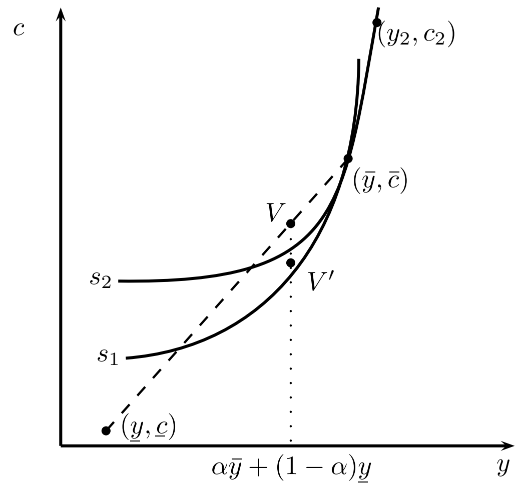

Figure 5: Case 2

revenue at $(\alpha\bar{c} + (1-\alpha)\underline{c})$, up to the point V' where utility is still higher than $\bar{U}$ to the low-skilled, but no high-skilled agent wants to trade for V' anymore (again see figure 5).

So offering V' to the low skills is incentive compatible. It is also production feasible as we assumed the low-skilled schedule $[(\underline{y}, \underline{c}); (\bar{y}, \bar{c})]$ was feasible. In effect, under $[(\underline{y}, \underline{c}); (\bar{y}, \bar{c})]$, the low-skilled aggregate production is $\alpha n_1 \bar{y} + (1-\alpha) n_1 \underline{y}$, and their consumption $\alpha n_1 \bar{c} + (1-\alpha) n_1 \underline{c}$ (in value terms). This is exactly equal to the aggregate production and consumption of the low-skilled at V. As the aggregate production is equal at V' and the aggregate consumption lower, V' is necessarily production feasible too.

It remains to show that one can achieve a higher level of social welfare by offering V or V' instead of the schedule in which some low-skilled receive $(\underline{y}, \underline{c})$ and some receive $(\bar{y}, \bar{c})$. When V' replaces $[(\underline{y}, \underline{c}); (\bar{y}, \bar{c})]$, the utility of every low-skilled is increases so that we actually have a Pareto improvement. When V replaces $[(\underline{y}, \underline{c}); (\bar{y}, \bar{c})]$, increase in social welfare follows from the concavity of the utility function,

$$\alpha n_1 u(\underline{y}, \underline{c}) + (1-\alpha) n_1 u(\bar{y}, \bar{c}) < n_1 u\big(\alpha\bar{y} + (1-\alpha)\underline{y},\ \alpha\bar{c} + (1-\alpha)\underline{c}\big).$$

□

The argument in the proof of theorem 1 lends itself to a nice interpretation : sharing labor time among the low-skilled is superior to rationing labor, e.g. through unemployment of some of the low-skilled.

As noted by Guesnerie and Roberts (1987), the argument relies heavily on its assumptions. The conjunction of a concave utility function and a classical utilitarian SWF is particularly important here.[5]

[5]Although one might want to assume that social planners are at least as worried for the unemployed as they are for the employed, Guesnerie and Roberts (1987) notes that this is not necessarily the case. For instance, when both wages and unemployement benefits are

More important for what follows, it is implicitly assumed that the introduction of a rationing mechanism on low-skilled labor has no impact on the willingness of the high-skilled to mimic a low-skilled outcome. This is far from being obvious. Assume that the high-skilled have to chose between an high-skill occupation and the low-skilled occupation. The work decision is on the extensive margin exclusively so that once one chooses an occupation, one must supply full-time labor to this occupation. As no adjustment in labor time is allowed, the high-skilled cannot mimic low-skilled outcomes by working less but remaining in the high-skilled occupation. So their only mimicking option consists in actually working in the low-skilled occupation. Then they face the possibility of unemployment. If the rationing mechanism is random, the fact that mimicking implies a *stochastic* outcome might help deterring risk-averse high-skilled from mimicking low-skilled outcomes, hence loosening the incentive compatibility constraint. These kinds of mechanism through which the introduction of a minimum wage helps relaxing the incentive constraint are studied in the two next sections.

3.2 Efficient rationing (Lee and Saez, 2012)

(When referring to a new article, we will change the notation to match with the article.)

3.2.1 The model

The demand side. Consider again a two input economy where the production function $F(h_1, h_2)$ depends on the number of low-skilled h_1 and high-skilled h_2 employed. It is assumed that the production function has *constant returns to scale*, and that the labor market is perfectly competitive, with the firm taking the equilibrium wages (w_1, w_2) as given. The production sector chooses labor demand (h_1, h_2) optimally in order to maximize profit which lead to the usual FOC

$$w_j = \frac{\partial F}{\partial h_j}, \qquad j = 1, 2. \tag{18}$$

As in the former model, we formalize the difference in skills by assuming that $w_1 < w_2$ at an equilibrium.

The supply side Again, all individuals have the same utility function and differ in the cost they face when working in one of the two occupations. It is an important distinctive feature of this model that *individuals are allowed to chose between working in occupation 1* (low-skilled), *occupation 2* (high-skilled), *or not working at all.* For now, individuals are not allowed to vary their labor time : when they commit to an occupation, they have to supply all their labor time to this occupation. The author refer to the choice between working and not

collectively bargained over by trade unions and firms representatives, a classical assumption in the labor economics litterature is to consider the trade union to more concerned by the wage of those who work than by the level of the unemployment benefits. The case for a concave utility function is also hard to vindicate in general.

working as the extensive margin, and to the choice between either occupations as the intensive margin.

The cost of working in a certain occupation is determined by the idiosyncratic vector

$$\theta = (\theta_1, \theta_2), \quad \theta \in \Theta, \quad \theta \sim H(\theta),$$

where $H(\theta)$ is a smooth cumulative distribution function and the population size is normalized to 1.

As we allow people to switch from an occupation to another, this model is more suited to long term horizon. The distribution of θ may for instance represent the distribution of education and training costs.

In this framework, the SWF can be understood as encapsulating the normative view of the social planner regarding people's responsibility in the determination of their disutility of work. Roughly speaking, a social planner who consider that the work disutility mostly result from a costly investment in human capital would advocate *laissez-faire* – as, in this case, costs of work are the mere consequence of inter-temporal preferences. When the cost of work parameter is viewed as endowments more than a prior investments, the social planner may want to compensate the low-skilled by adopting a SWF supporting redistribution.

The government *can* observe earnings $w = 0, w_1, w_2$ (as labor time is fixed) but does not know the individual cost of work θ. Thus the government can condition taxes and transfers based on the occupation, and the net effect of taxes and transfers can be written T_i, where $i = 0, 1, 2$. Each household can chose among the consumption levels

$$c_i = w_i - T_i. \tag{19}$$

A linear money-metric utility function is assumed

$$u_i = c_i - \theta_i, \quad \text{where } \theta_0 = 0. \tag{20}$$

The results are preserved with concave utility functions of the form $u(c_i) - \theta_i$, but the linear case allows for analogies with the classical Marhsalian surplus analysis.

Individuals choose an occupation in order to maximize (20) subject to (19). Let $\Theta_i = \{\theta \in \Theta \mid u_i = max_j u_j,\ j = 0, 1, 2\}$ be the set of θ's for which choosing occupation i is optimal. Let $h_i(c) = |\Theta_i|$ be the aggregate supply function, that is the fraction of the population working in occupation i as a function of the consumption level (19) offered in this occupation. The distribution $H(\theta)$ is assumed to be smooth enough for $h_i(c)$ to be smooth.

The equilibrium Given constant return to scale, the optimality condition (18) imply that

$$\frac{w_2}{w_1} = \frac{F_2'(1, {}^{h_2}/_{h_1})}{F_1'(1, {}^{h_2}/_{h_1})}. \tag{21}$$

Assuming *decreasing marginal productivities along each skill*, the R.H.S of (21) is a decreasing function of ${}^{h_2}/_{h_1}$. So we can write

$$ {}^{h_2}/_{h_1} = \rho({}^{w_2}/_{w_1}), \tag{22} $$

where $\rho(\cdot)$ is a decreasing function. Also, with constant returns to scale we have that there are not profits

$$ \Pi = F(h_1, h_2) - w_1 h_1 - w_2 h_2 = 0, $$

which we can rewrite as

$$ w_1 + w_2 \rho({}^{w_2}/_{w_1}) = F(1, \rho({}^{w_2}/_{w_1})). $$

For a fixed ${}^{h_2}/_{h_1}$, the last expression implicitly defines, w_2 as a decreasing function of w_1 (and vice versa), which we write $w_2 = w_2(w_1)$. The intuition is clear : if the proportion of low-skilled and high-skilled is fixed, any change in the wage of some type is compensated by an inverse change in the wage of the other type in order to maintain the zero profit condition.

We want to establish a more general relation between dw_1, dw_2, dh_1 and dh_2. To do so we differentiate the last expression to get

$$ \begin{aligned} & dw_1 + \rho dw_2 + w_2 d\rho && = F_2' d\rho, \\ \Leftrightarrow\ & dw_1 + \rho dw_2 && = 0, \qquad \text{(using } w_2 = F_2'\text{)} \\ \Leftrightarrow\ & \frac{dw_2}{dw_1} && = -\frac{h_1}{h_2}. \end{aligned} \tag{23} $$

(23) give the general equilibrium relation between the two markets. If we know about partial equilibrium variations on the market for low-skilled labor, we can infer the general equilibrium implications on the high-skilled market through this equation.

The social planner objective Assuming no exogenous spending requirement, the social planner seeks to maximize

$$ \begin{aligned} SW = \int G(u) dh(\theta), s.t. \\ h_0 c_0 + h_1 c_1 + h_2 c_2 \leq h_1 w_1 + h_2 w_2. \end{aligned} $$

Where $G(\cdot)$ is an increasing concave transformation of the money metric utilities, representing social preference for redistribution or decreasing utility of money. Given the particular structure of our problem, the social welfare can be written as

$$ SW = (1 - h_1 - h_2) G(c_0) + \int_{\Theta_1} G(c_1 - \theta_1) dH(\theta) + \int_{\Theta_2} G(c_2 - \theta_2) dH(\theta). $$

3.2.2 Desirability of a minimum wage with no taxes

Assume taxes and transfers are zero so that $T_0 = T_1 = T_2 = 0$ and $c_0 = 0$, $c_1 = w_1$, $c_2 = w_2$. Suppose we start from an equilibrium (w_1, w_2, h_1, h_2) and we introduce a minimum wage infinitesimally larger than the competitive wage, $\bar{w} = w_1 + d\bar{w}$. As a consequence, some low-skilled lose their job, and either become unemployed, or move to the high-skill occupation, enduring a higher cost of work.

An essential issue is to determine the rationing mechanism out of equilibrium. We need to know who losses her job. Given the distribution of θ, there are all kinds of agents who choose a low-skilled occupation. Chiefly relevant to the social planner problem is the fact that the surplus they get from this low-skilled occupation (u_i) differ. This surplus ranges from 0, when $\theta \in \Theta_1$ and $\theta_1 = w_1 - T_1$, up to possibly $w_1 - T_1$ if some agents are endowed with a $\theta_1 = 0$. Call the lowest surplus agents the "marginal" agents.

Of course, from the social planner point of view, it is better that the marginal agent be unemployed or switch to a high-skilled occupation. However, as a minimum wage is set, the agent with the higher surplus cannot undercut the marginal agent by offering to work at a wage which would deter the marginal agent from working as a low-skilled. The allocation of low-skilled job is not solved by wage adaptation anymore and there is rationing. Lee and Saez (2012) set this issue by assuming that some rationing mechanism operates, and that this mechanism is as favorable as can be for a minimum wage to be desirable.

Assumption 1 (Efficient rationing)**.** *Workers who involuntarily lose their low-skilled jobs (either becoming unemployed or by moving to an high-skilled occupation) due to the minimum wage are those with the least surplus from working as a low-skilled.*

See Lee and Saez (2012) for an empirical justification of this assumption. Under Efficient rationing, one obtain that a minimum wage is desirable with no taxes and transfers.

Proposition 2 (Desirability of a minimum wage with no transfers Lee and Saez (2012))**.** *With no taxes/transfers, if*

1. *Efficient rationing is satisfied,*
2. *The government values redistribution from the high-skilled to the low-skilled,*
3. *The demand elasticity for low-skilled is finite,*
4. *and the supply elasticity of low-skilled workers is positive,*

then introducing a minimum wage increases social welfare.

The intuition for this result is very similar to figure 2, yet with in infinitesimal changes and is illustrated in figure 6. Because of the Efficient rationing assumption, those who loose their low-skilled job have infinitesimal surpluses (as long as the elasticity conditions are met) so that the welfare loss from those losing jobs is of second order. The welfare gain to the low-skilled from the wage increase is, on the other hand a first order gain. So, up to a first order approximation, the total welfare effect on the low-skilled is the effect of higher wage alone and is positive.

Now consider the general equilibrium implications of the minimum wage policy. Remember from (23) that, due to the zero profit condition, a earning gain in the low-skilled sector is compensated by an identical earning loss in the high-skilled one ($dw_1 h_1 = -dw_2 h_2$). So a minimum wage effectively works as a tool for redistribution from the high-skilled to the low-skilled and a government favoring this kind of redistribution – see frame IV – would want to implement it.

Frame IV : When is redistribution from the high-skilled to the low-skilled desirable?

In the framework studied by Lee and Saez (2012), it is not at all obvious that redistribution from the high-skilled to the low-skilled is in general desirable. In particular, this is not a necessary consequence of the concavity of $G(\cdot)$. The desirability of such redistribution will notably depend on the distribution of skills $H(\theta)$. The intuition for this is clear : if the θ's are distributed in such a way that a vast majority of the low utility individuals are those choosing a high-skill occupation (for instance because many agents dislike high-skilled work, but hate low-skilled work, i.e. they have a high θ_2 but a yet higher θ_1), then redistribution might not be desirable.

More precisely, the relevant parameter is the so-called "social marginal welfare weight" for each occupation (Lee and Saez (2012)) defined as follows

$$\begin{aligned} g_0 &= \frac{G'(C_0)}{\lambda}, \\ g_j &= \frac{\int_{\Theta_j} G'(c_j - \theta_j) dH(\theta)}{\lambda h_j}, \end{aligned} \tag{24}$$

where λ is the multiplier of the Lagrangian associated with the social planner maximization problem. The g_i measure the average welfare weight of an individual in occupation i. They are a measure of the social marginal value of redistributing one dollar uniformly across all individuals in occupation i, which is what we do when we raise the minimum wage. Given that $G(\cdot)$ is concave the value of this redistribution is higher for occupations in which individuals have low surpluses. If some $g_i > 1$, this means that the benefits of redistribution to occupation i are higher than the cost λ (in terms of social welfare). A redistribution from the high-skilled to the low-skilled is desirable if $g_1 > g_2$.

3.2.3 With taxes an transfers

A stronger result obtains when a minimum wage is combined with a system of non-zero taxes an transfers.

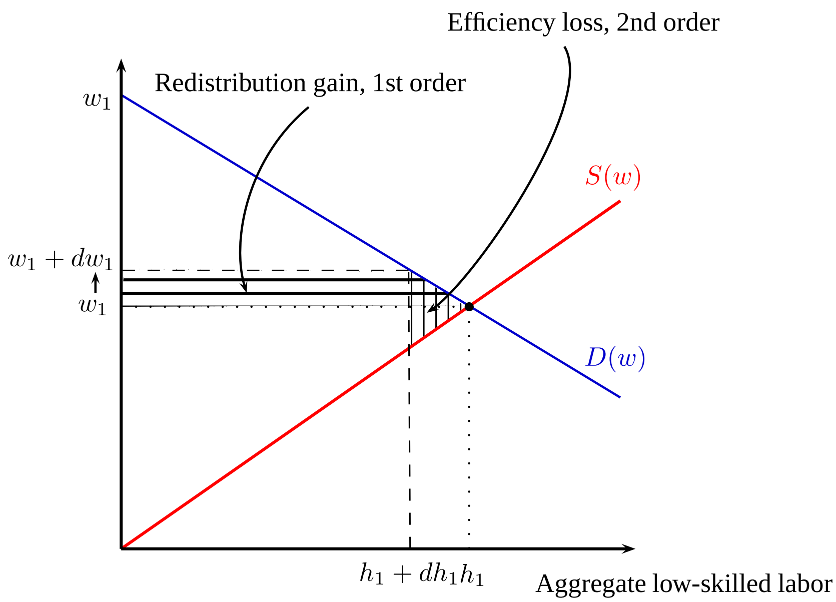

Figure 6: Desirability of a minimum wage with no taxes

Proposition 3 (Desirability of minimum wage with taxes)**.** *At an optimal tax allocation (without a minimum wage), if*

1. *Efficient rationing is satisfied,*
2. *The government would like to redistribute to the low-skilled if it could assume away labor shifting from the high-skilled and the unemployed,*
3. *The demand elasticity for low-skilled job is finite,*

then introducing a minimum wage increases social welfare and at an optimum, the redistribution to the low-skilled is complete (it reaches the level it would be without labor shifting).

We give a sketch of the argument, which again relies heavily on Efficient rationing. Suppose that at an optimum without minimum wage, the government would like to redistribute more to the low skilled but is prevented to do so by the fact that it will trigger labor shifting. The fear is that, if more consumption is awarded to the low-skilled, this will attract the high-skilled and the unemployed, which will lower the wage of the low-skilled and eventually decrease social welfare. However, this is avoided if Efficient rationing holds and a minimum wage is set at the current minimum wage. As the low-skilled wage w_1 is fixed, we know from (23) that neither w_2 nor h_1 or h_2 will change. As consumption is increased for the low-skilled, all those who previously chose to work as low-skilled still do. Therefore, as h_1 is fixed, if some want to shift to a low-skilled job, this will create rationing on the low-skilled job market.

For an change dc_1 in the consumption of the low-skilled, those who would like to shift are those who would get a negligible surplus if they were to receive a low-skilled job. Therefore, under Efficient rationing, the priority in the rationing process will be given to those who worked as low-skilled before the change dc_1 as they have a larger surplus. So no shifter will actually get a job if shifting and every labor shifting is thus precluded. By Efficient rationing again, those who will have to leave their low-skilled job will also be those who surplus is negligible. Given this argument, the government may increase the consumption of the low-skilled up to the point were it would be set was labor shifting simply banned ($g_1 = 1$, see frame IV) : redistribution to the low-skilled can effectively be made lump-sum.

What this reasoning also shows is that, absent possibilities of labor shift, there is little room for a minimum wage. This is roughly the content of Proposition 3 in Lee and Saez (2012). When the frontiers of the sectors are watertight, any reduction in an existing minimum wage is a Pareto improvement.

Finally, notice that these results do not carry over in a setting with variable hours of work, unless the government can enforce occupation specific taxes, which seems unlikely. This was the case in the model we just studied. However, it was justified by our assumption of a fixed labor time. Because labor time was fixed, the usual assumption that the government observes income allowed it to infer occupations and wages. In a model with variable labor time, this is not the case anymore, and occupation specific taxes are a very strong assumption. Most people would therefore consider that the results we presented do not carry over when variable labor time is allowed.

3.3 Selectivity in access to unemployement benefits (Boadway and Cuff, 2001)

3.3.1 The model

The agents

- $U(c, l) = u(c) - l$, quasi-linear in labor.
- $u(c)$ is strictly concave.
- Abilities to convert l in y are
  - Observed by the firms but not by the government,
  - Distributed in the population according to $F(w)$ with $w \in [\underline{w}, \bar{w}]$, where $\underline{w} > 0$, $\bar{w} < \infty$, $F(\bar{w}) = 1$ and $F'(w) \equiv f(w) > 0$, for all w.
- People that want to work can (at the offered wage w).

3.3.2 Linear income taxation

- Government observes income *and* employment status.
- The government taxes proportionally the income of those who work, so that $c = (1 - \tau)y$, and redistribute T to every *non-working* household (non-working household consume T).

Optimal consumer behavior

- Working individuals $\max_y u\Big((1-\tau)y\Big) - {}^{y}\!/\!{}_{w}$.
- FOC:

$$u'\Big((1-\tau)y\Big)(1-\tau) = \frac{1}{w}, \tag{25}$$

which implicitly defines y as a function of w and τ, $y(w,\tau)$.

- By total differentiation of the FOC (setting $d\tau = 0$), one finds that ${}^{\partial y(w,\tau)}\!/\!{}_{\partial w} > 0$. However ${}^{\partial y(w,\tau)}\!/\!{}_{\partial \tau}$ cannot be signed.[6]
- The participation constraint is determined by the government choices of τ and T : an individual will work if $u(c) - {}^{y}\!/\!{}_{w} \geq u(T)$.
- Maximized utility is increasing in w, so there is only one type for which the participation condition is binding. This type is denoted $w_\circ$ and assumed to be in the interior of $[\underline{w}, \bar{w}]$.
- So individuals with $w < w_\circ$ do not work (they choose not to work) and individuals with $w > w_\circ$ do work.
- $w_\circ$ is implicitly defined by $u\big((1-\tau)y(w_\circ,\tau)\big) - {}^{y(w_\circ,\tau)}\!/\!{}_{w_\circ} = u(T)$. The individual with $w_\circ$ is called the marginal worker.
- Finally, note that utility is decreasing in the tax rate τ (for fixed T as a larger τ might ultimately raise the utility of the unemployed).

The planner optimization problem

- For expository purposes, suppose a classical utilitarian SWF. The analysis is qualitatively identical if the SWF displays finite inequality aversion, but not in the case of a maximin or leximin SWF.
- The Lagrangian of the government problem is given in figure 7, where the first constraint is the budget constraint of the government, and the second the participation condition. The government implicitly choses $w_\circ$ but cannot force people to work if they are better-off with their transfer only. So there are only two real degrees of freedom for the government.
- Using the fact that the participation constraint is binding for the marginal worker, the FOC of this problem are as given in figure 8.

[6] By total differentiation again, we have $\frac{\partial y(w,\tau)}{\partial \tau} = \frac{u'(\cdot)}{u''(\cdot)(1-\tau)^2} + \frac{y}{(1-\tau)}$, so ${}^{\partial y(w,\tau)}\!/\!{}_{\partial \tau}$ depends on the curvature of $u(\cdot)$ and the level of y.

$$\max_{\tau, T, w_o} \mathscr{L} = \int_{w_o}^{\bar{w}} \left[u\left(y(w,\tau)(1-\tau)\right) - \frac{y(w,\tau)}{w} \right] f(w)\,\mathrm{d}w + u(T)F(w_o)$$
$$+ \lambda \left[\int_{w_o}^{\bar{w}} \tau y(w,\tau) f(w)\,\mathrm{d}w - TF(w_o) \right]$$
$$+ \phi \left[u\left(y(w_o,\tau)(1-\tau)\right) - \frac{y(w_o,\tau)}{w_o} - u(T) \right],$$

Figure 7: Lagrangian of the government maximization program

$$- \int_{w_o}^{\bar{w}} u'(c(w,\tau)) y(w,\tau) f(w)\,\mathrm{d}w + \lambda \int_{w_o}^{\bar{w}} \left[y(w,\tau) + \tau \frac{\partial y(w,\tau)}{\partial \tau} \right] f(w)\,\mathrm{d}w$$
$$- \phi u'(c(w_o,\tau)) y(w_o,\tau) = 0, \qquad (\tau)$$
$$u'(T)F(w_o) + \lambda F(w_o) - \phi u'(T) = 0, \qquad (T)$$
$$- \lambda[\tau y(w_o,\tau) + T] f(w_o) + \phi \frac{y(w_o,\tau)}{w_o^2} = 0. \qquad (w_o)$$

Figure 8: FOC

- Denote the solutions of the problem without minimum wage $\tau^\circ, T^\circ, w_\circ$.
- At an equilibrium, those with $w = w^\circ$ are exactly indifferent between working and not working, and as maximized utility is increasing in w, everyone with $w > w^\circ$ is better-off.
- We also have that the marginal cost of public funds (in terms of social utility) $\lambda = u'(T^\circ)[F(w_\circ) - \phi]$. Using $(w_\circ)$ we can sign $\phi > 0$ ($\lambda > 0$ as it is the multiplier of a binding inequality constraint), so that $\lambda < u'(T^\circ)F(w_\circ)$, that is we ideally would like to increase the level of redistribution if we did not have to respect the participation constraint.

Effect of minimal wage legislation

- The legislation takes the following form
  - Firms cannot offer a wage below w_m.
  - But as firms can observe abilities, they offer a job only to those who's abilities are at least w_m.
  - It is assumed that the minimum wage policy is perfectly enforceable *and* that people who are offered a job have no choice but to take it. As suggested in the title of the section, this could be the case if as an individual refuses a job offer, she cannot access unemployment benefits.
- So the government solves its information asymmetry problem by using the information of the firms : through this legislation, the government implicitly recover information on abilities as it knows that those who work have skill levels above w_m, and those who do not work below skill levels below w_m. This allows to ignore the participation constraint.
- So the problem is as described in figure 7, but with $\phi = 0$.
- Consider what would happen if the government fixed $w_m = w_\circ$.
- By looking at the second FOC, (T), before the minimum wage is introduced, we have $u'(T^\circ)F(w_\circ) = \phi u'(T^\circ) - \lambda F(w_\circ)$. So to restore the equality when $\phi = 0$ one must decrease $u'(T^\circ)$, that is increase T° up to say T^m.
- Therefore, to keep the budget balanced, one must also increase taxes to say τ^m. This is illustrated in figure 9. In the end, as depicted in the figure, all working individuals are worst off and those who do not work are better-off.
- Social welfare is increased by the minimum wage, even when set at $w_\circ = w_m$, as introducing the minimum wage triggers a change in the optimal solution.
- It is not clear whether or not it is desirable to raise the minimum wage above $w_\circ$. From the Envelope Theorem, deriving the maximized social welfare gives figure 10 (where ζ is likely a typo and should be f).

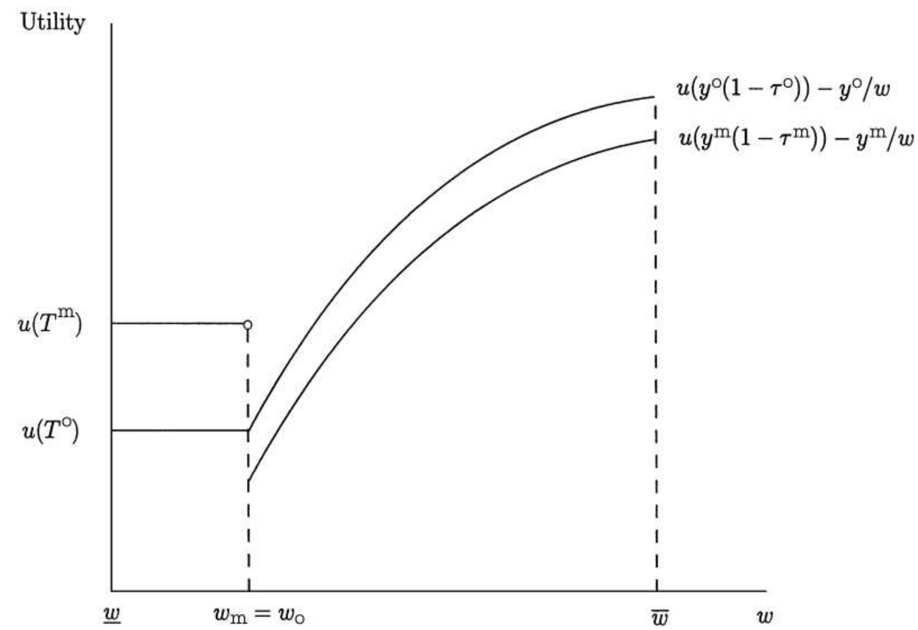

Fig. 1. Proportional tax with and without a minimum wage ($w_{\mathrm{m}} = w_{\mathrm{o}}$).

Figure 9: With and without minimum wage

$$\frac{\mathrm{d}SW}{\mathrm{d}w_{\mathrm{m}}} = \left.\frac{\partial \mathscr{L}}{\partial w_{\mathrm{m}}}\right|_{w_{\mathrm{m}} = w_{\mathrm{o}}} = \left[u\left(T^{\mathrm{m}}\right) - u(c(w_{\mathrm{o}}, \tau)) + \frac{y(w_{\mathrm{o}}, \tau)}{w_{\mathrm{o}}}\right]\zeta\left(w_{\mathrm{o}}\right)$$

$$- u'(T^{\mathrm{m}})[\tau y(w_{\mathrm{o}}, \tau) + T^{\mathrm{m}}]f(w_{\mathrm{o}}).$$

Figure 10:

- The first term shows the difference between what the marginal worker get when she lives out of transfers and the utility she got – disutility of labor included – when she was working at the minimum wage (this is positive), weighted by the density at $w_{\circ}$. The second term is negative and represent the additional revenue that government must raise as a worker is lost and an additional transfer is to be paid (weighted by the social value of this revenue loss). This can take different values depending on the form of the utility function and the distribution of wages. As the minimum wage can be either higher or lower than $w_{\circ}$, unemployment might either rise of fall following the introduction of a minimum wage.

3.3.3 Non-linear income tax

Fixed labour supply

- If they work, people supply a fixed amount of labour d and produce $y = wd$, depending on their productivity.
- Firms observe ability perfectly (they do not make mistake paying wages which would be too high) so y is equivalently the income one receives.
- So incomes are distributed across $[\underline{y}, \bar{y}]$ according to some distribution function $H(y)$, $H'(y) > 0$ and $\underline{y} < 0, \bar{y} < \infty$.
- Utility is $u(c) - d$, the same for every agent.
- The tax schedule can now be conditioned on the observed income $c = y - t(y)$ (non-linear taxation).
- As the government observes y and knows that everyone's working time is d, it can infer w.
- Yet *potential income of those who do not work remains non-observable*, and everyone who does not work must the same transfer T. So there remains some asymmetric information due to the fact that to observe someone's wage, the person has to be working (one can still mimic unemployment).
- Therefore a participation constraint still applies $u(y - t(y)) - d \geq u(T)$ for all those who work.

$$\max_{t(y),T,y_p} \mathscr{L} = \int_{\underline{y}}^{y_p} u(T)h(y)\,\mathrm{d}y + \int_{y_p}^{\bar{y}} [u(y-t(y)) - d\,]h(y)\,\mathrm{d}y$$
$$+ \lambda\left(\int_{y_p}^{\bar{y}} t(y)h(y)\,\mathrm{d}y - \int_{\underline{y}}^{y_p} Th(y)\,\mathrm{d}y - R\right)$$
$$+ \int_{y_p}^{\bar{y}} \phi(y)\,[u(y-t(y)) - d - u(T)]\,\mathrm{d}y,$$

Figure 11:

$$-u'(y-t(y))h(y) + \lambda h(y) - \phi(y)u'(y-t(y)) = 0, \quad y \geqslant y_{\mathrm{p}}, \qquad (t)$$
$$u'(T)H(y_{\mathrm{p}}) - \lambda H(y_{\mathrm{p}}) - \int_{y_p}^{\bar{y}} \phi(y)u'(T)\,\mathrm{d}y = 0, \qquad (T)$$
$$-\lambda[T + t(y_{\mathrm{p}})] = 0, \qquad (y_{\mathrm{p}})$$

Figure 12:

- The difference with respect to the linear tax setting is that $t(y)$ begin conditional on y allows the government to make the constraint just binding for all workers.
- y_p is the cut-off seperating those who work from those who don't, determined endogenously by the choice of T and $t(y)$.
- The new maximization problem is represented in figure 11, for which the FOC are in figure 12.
- The solutions of this problem are denoted $t^p(y), T^p, y_p$.
- In equilibrium, all the participation constraint bind $(u(y - t(y)) - d = u(T))$, due to the concavity of $u(\cdot)$. Consumption and utility are equalized among the *working* population.
- Because of the inability to observe *potential* income among those who do not work implies that again, at an optimum, one would want to redistribute more if the participation constraint was not binding. So one might want to consider the desirability of a minimum wage for the same reason as in the linear tax model.
- Again, with a minimum wage, those who work are those with wage $w > w_m$. As in the linear taxation case, the introduction of a minimum wage allows the government to ignore the participation constraints for all the individuals.
- In this case, the first best allocation can be implemented, in which the utility is equalized among the *whole* population. Those who work would

love to be unemployed and live out of the unemployment benefits. But they cannot do so as to access unemployment benefits they must conduct a job search in which they will be offered a job as their wage is above the minimum wage (and this is perfectly observable by the government).

3.3.4 The variable labor time

- Boadway and Cuff (2001) uses calculus of variation to solve their continuum model. Here we present only the main feature of their results.
- We are in the standard model in which individuals chose c, y to maximize $u(c) - {}^{y}/_{w}$ given their wage level and $y = wl$.
- Single-crossing is assumed.
- Incentive compatibility constraint is required.
- The classical result of income increasing with wage holds.
- The government faces a non-negativity constraint on income and a revenue feasibility constraint.

Without minimum wage

- At a solution, the highest ability individuals face a zero marginal tax rate and the marginal individuals (indifferent between working and not working) have a positive marginal tax rate.

With a minimum wage

- The author proceed similarly as before. First set $w_m = w_\circ$.
- In this case, the highest ability and the marginal individuals face a zero marginal tax rate.
- One can show that the introduction of a minimum wage reduces the marginal cost of public funds. As a consequence, the transfer to the unemployed will be higher.
- Once again, one finds that the marginal individual are worse-off under a minimum wage policy.
- In a second step, the authors investigate whether choosing $w_m \neq w_\circ$ can be optimal. They find that a higher minimum wage unambiguously increases social welfare. This will trigger a higher unemployment than under $w_m = w_\circ$.

References

Allen, S. (1987). Taxes, redistribution, and the minimum wage: a theoretical analysis. *The Quarterly Journal of Economics*, 102(3):477–490.

Boadway, R. and Cuff, K. (2001). A minimum wage can be welfare-improving and employment-enhancing. *European Economic Review*, 45:553–576.

Guesnerie, R. and Roberts, K. (1984). Effective policy tools and quantity controls. *Econometrica*, 52(1):59–86.

Guesnerie, R. and Roberts, K. (1987). Minimum wage legislation as a second best policy. *European Economic Review*, 31:490–498.

Lee, D. and Saez, E. (2012). Optimal minimum wage policy in competitive labor markets. *Journal of Public Economics*, 96(9-10):739–749.

Marceau, N. and Boadway, R. (1994). Minimum wage legislation and unemployment insurance as instruments for redistribution. *The Scandinavian Journal of Economics*, 96(1):67–81.