Should low skilled work be subsidized?
  an inquiry within the extensive model

Philippe Choné and Guy Laroque
INSEE-CREST

December 4, 2006
Very preliminary and incomplete; comments welcome

1We have benefited from the comments of Martin Hellwig, Bruno Jullien, Wojciech Kopczuk, Jean-Charles Rochet, Emmanuel Saez, Gilles Saint Paul, Bernard Salanié, and of the participants in the theory seminar in Toulouse.
Abstract

A number of countries have implemented variants of a negative income tax, to push the less skilled members of the economy into work, or to make work pay in comparison with welfare benefits. In most cases, these measures have resulted for the concerned groups in a decrease of the tax rates, that remain positive, rather than in a subsidy, in conformity with the standard recommendations of the theory of optimal taxation. Indeed in the Mirrlees setup (continuous labor supply or intensive margin, unobserved productivity, utilitarian planner) the marginal tax rate is non negative at the optimum. It is known however mainly through examples, see in particular Diamond (1980), that in the extensive model it may be optimal to subsidize low skilled labor.

The purpose of the paper is to study more systematically the second best optimal tax schedules in the extensive model when the planner has a utilitarian objective. We find a large class of economies and planner objectives where it is optimal to have the low skilled work more than in the laissez-faire: their net income at work is larger than the sum of the subsistence income they would get when not working and of their productivity (here the cost of their labor to their employer). This property holds when both income effects and the social weights attached to the unemployed are not too large. We hope that these results help towards providing some theoretical foundations for low skilled work subsidy, and extending the scope of welfare to work programs.

JEL classification numbers: H21, H31.
Keywords: optimal taxation, extensive model, welfare.

Please address correspondence to Guy Laroque, laroque@ensae.fr, INSEE-CREST J360, 15 boulevard Gabriel Péri, 92245 Malakoff Cedex, France.
1 Introduction

A number of countries have recently implemented variants of a negative income tax, to push the less skilled members of the economy into work, or to make work pay in comparison with welfare benefits. In most cases, these measures have resulted for the concerned groups in a decrease of the tax rates, that remain positive, rather than in a subsidy\textsuperscript{1}. This is in conformity with the current recommendations of the theory of optimal taxation. Indeed it is well established in the Mirrlees setup (continuous labor supply or intensive margin, unobserved productivity, constant opportunity cost of work, utilitarian planner) that the marginal tax rate is non negative at the optimum (Seade (1977), Seade (1982), Werning (2000), Hellwig (2005)).

In fact, early on Diamond (1980), more recently Saez (2002), Beaudry and Blackorby (2004), Boone and Bovenberg (2004), Boone and Bovenberg (2006), Choné and Laroque (2005) and Laroque (2005) have described setups where the positive tax rate result does not hold. A common feature of the (rather different) models used in these works is that labor supply decisions involve a zero-one component, an extensive margin. These studies exhibit cases where negative tax rates can occur at an optimum. But it is fair to say that their theoretical foundations remain unclear as well as their practical relevance. Also it is important to note that the implications of negative tax rates are quite different in an extensive model and in an intensive model. In the intensive model, they imply that labor supply is distorted upwards compared with the laissez-faire. In the extensive model it is the average tax rate which is relevant for the shape of labor supply, and to the best of our knowledge the previous literature has not studied the extent of labor supply distortions in this setup.

The purpose of this paper is to study the shape of the second best allocations that are consistent with a utilitarian criterion in the extensive model. The aim is to classify the shapes of the solutions according to economically meaningful assumptions, overcoming the difficulties associated with the intrinsic non convexity of the model.

An important feature of the model is the average social weight of the employees, i.e. the derivative of the social welfare function (per head of employee) with

\textsuperscript{1}In the United States, the negative marginal tax rate is at its maximum equal to 40% (see e.g. CBO (2000) for the United States). However, in the extensive model which concerns us here, it is not the marginal rate that is relevant, but the average rate, i.e. 1 minus the ratio (net income at work - subsistence income out of work)/(labor cost). Here 'net income at work' denotes disposable income when taking a full time job, 'subsistence income out of work' is disposable income when not working, including food stamps and temporary assistance for needy families, and 'labor cost' is the cost of labor to the employer, a proxy for productivity in a competitive environment. While I do not know of a microeconomic assessment of the average tax rates, subsistence income is certainly not zero, so that they are larger than -40%. Aggregate figures indicate that they are unlikely to be negative. In 2004, the EITC budget was $33 billions, with TANF and food stamps respectively $25 and $27 billions.
with respect to income, at a given productivity level. We give conditions under which this derivative decreases with income and is therefore highest for the low skilled.

The first order condition for optimality implies a negative tax rate for the employees whose average social weight is larger than the marginal cost of public funds. The argument is as follows. Let $\lambda$ be the marginal cost of public funds, the multiplier of the government budget constraint. Consider a small change in the tax schedule in favor of the working agents of (low) productivity $\omega$, keeping unchanged the situation of the other agents of productivity different from $\omega$. This reform has two effects: it gives more money to the agents that are already working, and it brings into the labor force some (pivotal) agents previously unemployed. By assumption, the social value of a marginal transfer to the working agents of low productivity is larger than the marginal cost of public funds: the first effect increases social welfare. The second effect comes from the pivotal agents that enter the labor force. They are essentially indifferent between working and not working, and their impact on social welfare comes from the difference between their productivity $\omega$ and the cost of putting them to work. For the first order condition to be satisfied, this difference must be negative: the financial incentive to work (income at work minus subsistence income) has to be larger than productivity.

Whether the optimal allocation involves distorting labor supply upwards with respect to laissez-faire for low skilled workers therefore depends on whether their social weight is larger than the marginal cost of public funds. In the simplest benchmark situation, no income effects on labor supply, social weight depends only on income net of work opportunity cost (not directly on characteristics, nor on employment status), non degenerate distribution of work opportunity costs independent of the distribution of productivities, it turns out to our surprise the social weight of the low skilled workers is always larger than the marginal cost of public funds: all utilitarian optima in the benchmark model involve upwards labor supply distortions for low productivity workers. The utilitarian optimal allocations have more ‘working poor’ than the competitive equilibrium.

This property does not hold anymore when the government puts a lot of weight on the unemployed agents, as a Rawlsian planner would do: then all workers are taxed to contribute to the subsistence income of the non workers (see Choné and Laroque (2005)). Similarly, income effects on labor supply, assuming leisure to be a normal good, increase the (implicit) cost of transfers to the unemployed and seem to play against upwards labor supply distortions. All this should the subject of further research.
2 The model

2.1 Heterogeneity and the description of the agents

We consider an economy with a continuum of agents of measure 1. There is a single commodity, which serves as the numéraire. The agents’ main decision is whether to work or not. They differ along several dimensions. First, when they work full time, they produce a quantity of commodity at most equal to their productivity level $\omega$. The model is extensive: as long as one works, one’s utility is unaffected by the quantity $y$ effectively produced, $y \leq \omega$. Second, there is an heterogeneity parameter $\alpha$ that describes other idiosyncratic characteristics of the agents, such as their costs of going to work or pleasure in listening music. Let $c_E$ be their consumption of commodity if they decide to work, $c_U$ their consumption when unemployed. The agent’s choice is the one that delivers the greatest utility of

$$\begin{cases} u(c_E; \alpha, \omega) & \text{if she participates,} \\ v(c_U; \alpha, \omega) & \text{if she does not work.} \end{cases}$$

The utility functions $u$ and $v$ are assumed to be twice continuously differentiable. They are increasing and concave in their first argument, which takes its values in $\mathbb{R}_+$. The utility functions are normalized so that the objective of the utilitarian government is the sum of the utilities of the agents in the economy. In particular the social value of a marginal income transfer to the agent $(\alpha, \omega)$, which we shall also call the social weight of this agent, is $u'(c_E; \alpha, \omega)$ (resp. $v'(c_U; \alpha, \omega)$) when she works (resp. when she is unemployed).

Allowing for a lot of heterogeneity (i.e. a multidimensional parameter $\alpha$) accords with empirical studies where marital status, family composition, human capital and culture appear to be determinants of the labor supply decision. However to simplify the exposition, we shall stick here with a parameter of dimension 1. Furthermore we assume that

$$u(c_E; \alpha, \omega) - v(c_U; \alpha, \omega)$$

is a continuously differentiable decreasing function of $\alpha$. Let $a(c_E - c_U, c_U, \omega)$ be the unique root\footnote{When $\alpha$ belongs to some, possibly unbounded interval $[\underline{\alpha}, \overline{\alpha}]$, define $a(c_E - c_U, c_U, \omega) = \alpha$ (resp. $\overline{\alpha}$) when the left hand side of the equation is always negative (resp. positive).} of the equation

$$u(c_E; \alpha, \omega) - v(c_U; \alpha, \omega) = 0.$$

The function $a$ is continuously differentiable with respect to its two first arguments at each value in the interior of its range.
It is increasing in first argument, the financial incentive to work \(c_E - c_U\). Faced with an income profile \((c_E, c_U)\), where \(c_E\) (resp. \(c_U\)) denotes income when employed (resp. unemployed), an agent of type \((\alpha, \omega)\) works whenever
\[
\alpha < a(c_E - c_U, c_U, \omega),
\]
does not work when
\[
\alpha > a(c_E - c_U, c_U, \omega),
\]
and is indifferent between working or not when at the boundary. From an economic viewpoint, it is useful to define the work opportunity cost \(\delta(c_U; \alpha, \omega)\) of agent \((\alpha, \omega)\), when her income out of work is \(c_U\), as the solution of the equation
\[
a(\delta, c_U, \omega) = \alpha.
\]
The work opportunity cost is the (possibly negative) sum of money which, given to the agent if she works on top of the subsistence income she has while unemployed, makes her indifferent between working or not. Whenever possible, we shall use the distribution of work opportunity costs (marked with a tilde), which has a direct economic interpretation rather than positing assumptions on the abstract parameter \(\alpha\). Substituting \(\delta\) for \(\alpha\), this leads us to introduce
\[
\hat{u}(c_E; \delta, \omega, c_U) \equiv u(c_E; a(\delta, c_U, \omega), \omega),
\]
and
\[
\hat{v}(c_U; \delta, \omega) \equiv v(c_U; a(\delta, c_U, \omega), \omega).
\]
Similarly, the distribution of \(\delta\) will follow from that of \((\alpha, \omega)\) and will be denoted with a tilde. It depends on the subsistence income \(c_U\).

The dependence of \(a\) with respect to its second argument is linked to the income effects. If leisure is a normal good, keeping the financial incentives to work constant, an increase in income out of work \(c_U\) decreases labor supply: \(a\) is decreasing in its second argument, the derivative \(a'_2\) is negative\(^3\). If there are no income effects, \(a\) does not depend on its second argument.

An economy is defined by a distribution of agents characteristics \((\alpha, \omega)\), which implies a distribution on \((\delta, \omega)\). It will make things simpler to assume:

**Assumption 1.** The marginal distribution of productivities \(\omega\) has support \(\Omega = [\underline{\omega}, \overline{\omega}]\), an interval of the positive line. Its cumulative distribution function \(G\) has a continuous positive derivative \(g\) everywhere on the support.

\(^3\)Note that by construction the total derivative of \(a\) with respect to \(c_U\),
\[
\frac{da}{dc_U} = -a'_1 + a'_2 < 0
\]
is negative. Even when leisure is not a normal good, \(a'_2\) is bounded above by \(a'_1\).
The distribution of $\alpha$ (resp. $\delta$), conditional on $\omega$, is also continuous with support $[\alpha, \overline{\alpha}]$, $\overline{\alpha} > \alpha \geq 0$ (resp. $[\delta, \overline{\delta}]$, $\overline{\delta} > \delta > 0$), and cumulative distribution function $F(\cdot | \omega)$ (resp. $\tilde{F}(\cdot, \omega, c_U)$ where $c_U$ is subsistence income). Its probability distribution function $f(\cdot | \omega)$ (resp. $\tilde{f}(\cdot, \omega, c_U)$) is positive everywhere on its support.

**Benchmark case:** a simple specification of particular interest is the following. The utility functions are defined through a concave twice differentiable increasing function on $\mathbb{R}$, such that

$$u(c_E; \alpha, \omega) = U(c_E - \alpha),$$

$$v(c_U; \alpha, \omega) = U(c_U).$$

Then direct substitutions yield

$$a(d, c_U, \omega) = d, \quad \delta(c_U; \alpha, \omega) = \alpha$$

$$\tilde{u}(c_E; \delta, \omega, c_U) = U(c_E - \delta), \quad \tilde{v}(c_U; \delta, \omega) = U(c_U).$$

The work opportunity cost of agent $(\alpha, \omega)$ is simply equal to $\alpha$. It is independent of the value of the subsistence income $c_U$: there are no income effects. Also the value attached by the government to the welfare of the unemployed agents only depend on their income, and does not vary with their characteristics $(\alpha, \omega)$.

### 2.2 Tax schedules

The government announces an income schedule $R$, which associates to any before tax income $y$, $y > 0$, a non negative disposable income $R(y)$ and gives to the non workers a subsistence income $s$. To this income schedule, one can associates a wedge function $w(y)$ which describes the difference between labor cost to the firm and income of the worker

$$w(y) = y - R(y),$$

and the tax rate faced by the worker $\tau(y)$

$$\tau(y) = \frac{y - (R(y) - r)}{y} = 1 - \frac{R(y) - r}{y}.$$  

For some (low) income levels, $\tau(y)$ may be negative, in which case its absolute value is a subsidy and labor supply is distorted upwards compared to laissez-faire where $\tau(y) = 0$.

The after tax income schedule $R(\cdot)$ can be taken to be non decreasing without loss of generality. Then, when she participates, the agent always produces at her

---

4Take any, possibly sometimes decreasing, function $\tilde{R}(y)$. Let $R(y) = \max_{y \geq z} \tilde{R}(z)$. The agents have the same behavior under $R$ and $\tilde{R}$. Such a property would not hold anymore if there were variable opportunity costs to work, which would smooth the labor supply of the agents.
full productivity, so that her before tax income $y$ is equal to $\omega$. As a consequence, before tax income takes its values in the support $\Omega$ of productivities. The function $R$ is to be defined on $\Omega$.

An agent of productivity $\omega$ works whenever $\delta$ is less than or equal to $R(\omega) - s$. We assume that the utilitarian government maximizes

$$\int_{\Omega} \left[ \int_{a(R(\omega) - s, \omega)}^{a(\omega)} u(R; \alpha, \omega) \, dF(\alpha|\omega) + \int_{a(R(\omega) - s, \omega)}^{a(\omega)} v(s; \alpha, \omega) \, dF(\alpha|\omega) \right] \, dG(\omega),$$

subject to the feasibility constraint

$$\int_{\Omega} \left[ \omega - R(\omega) + s \right] F(a(R(\omega) - s, s, \omega)|\omega) \, dG(\omega) = s.$$  

or equivalently, using work opportunity costs

$$\int_{\Omega} \left[ \int_{\delta}^{R(\omega) - s} \tilde{u}(R; \delta, \omega, s) \, d\tilde{F}(\delta|\omega, s) + \int_{R(\omega) - s}^{\delta} \tilde{v}(s; \delta, \omega) \, d\tilde{F}(\delta|\omega, s) \right] \, dG(\omega),$$

subject to the feasibility constraint

$$\int_{\Omega} \left[ \omega - R(\omega) + s \right] \tilde{F}(R(\omega) - s|\omega, s) \, dG(\omega) = s.$$  

The optimization takes place over the couple $(R, s)$, where $R$ is a non negative non decreasing function and $s$ is non negative.

Note that the objective does not depend on the precise value taken by $R(\omega)$ in the region where nobody works, i.e. whenever $R(\omega) - s$ is smaller than the minimal work opportunity cost $\delta(s; \alpha, \omega)$, or $a(R(\omega) - s, s, \omega)$ is equal to $\alpha$.

### 3 Marginal utility of income or social weights

The marginal utilities of income play an important role in the analysis. We denote the average marginal utility of income, or social weight, of the employed agents of productivity $\omega$ when their income is $R$ as $p_E(R, s|\omega)$

$$p_E(R, s|\omega) = E \left[ \tilde{u}'_1(R; \delta, \omega, s) \mid \delta \leq R - s, \omega \right]$$

$$= \frac{1}{F(R - s|\omega, s)} \int_{\delta}^{R - s} \tilde{u}'_1(R; \delta, \omega, s) \, d\tilde{F}(\delta|\omega, s).$$  

\footnote{For efficiency, the agents that are indifferent between working and not working should be put to work when their productivity is larger than the induced cost to the government, $\omega > R(\omega) - s$ and left on the dole when the inequality is in the other direction, $\omega < R(\omega) - s$. To avoid an overburden of notations, in the following equations, we suppose that all those agents, typically a set of measure zero, are working.}
Similarly, the average social weight of the *unemployed* agents of productivity $\omega$ when the net wage is $R$ is denoted $p_U(R, s|\omega)$

$$p_U(R, s|\omega) = E\left[\tilde{v}_1'(s; \delta, \omega) \mid \delta > R - s, \omega\right]$$

$$= \frac{1}{1 - \tilde{F}(R - s|\omega, s)} \int_{R-s}^{\infty} \tilde{v}_1'(s; \delta, \omega) d\tilde{F}(\delta|\omega, s). \quad (8)$$

How the social weight of the employed agents $p_E(R, s|\omega)$ depends on income at work $R$ turns out to be an important determinant of the shape of the optimal tax schedule. A natural assumption under utilitarianism, with which we shall work in most of the paper, is that $p_E$ decreases with $R$: the larger income, the smaller the social weight put on the agents who receive this income. However a brute differentiation yields

$$\frac{\partial p_E}{\partial R} = \frac{1}{\tilde{F}(R - s|\omega, s)} \int_{\delta}^{R-s} \tilde{u}''_{11}(R; \delta, \omega, s) d\tilde{F}(\delta|\omega, s)$$

$$- \frac{\tilde{f}(R - s|\omega, s) a_1'}{\tilde{F}(R - s|\omega, s)} p_E(R, s|\omega)$$

$$+ \frac{a_1'}{\tilde{F}(R - s|\omega, s)} \tilde{u}'_1(R; R - s, \omega, s) \tilde{f}(R - s|\omega, s).$$

The two first terms are negative, following the intuition, but the last term is positive: an increase in $R$ brings newcomers into employment, whose weights may be high.

**Proposition 1.** Assume that the marginal utility of income of the new entrants into the labour force is non increasing with $R$, that is

$$\frac{d\tilde{u}}{dR}[R; R - s, \omega, s]$$

is non increasing in $R$ for all $s$ and $\omega$. Assume furthermore that the cumulative distribution of work opportunity costs $\tilde{F}(\cdot|s, \omega)$ has a concave logarithm.

Then $p_E(R, s|\omega)$ is a non increasing function of $R$.

**Proof of Proposition 1**

We have to show that

$$p_E(R, s|\omega) = \frac{1}{\tilde{F}(R - s|\omega, s)} \int_{\delta(\omega)}^{R-s} \frac{\partial \tilde{u}}{\partial R}(R; \delta) d\tilde{F}(\delta|\omega, s) = E_{\delta}[\frac{\partial \tilde{u}}{\partial R} \mid \delta \leq R - s, \omega, s]$$

$$= \int_{\delta(\omega)}^{R-s} \frac{\partial \tilde{u}}{\partial R}(R; \delta) d\tilde{F}(\delta|\omega, s). \quad (9)$$

decreases with $R$. To simplify notations, we drop the variables $s$ and $\omega$ which are kept constant in the proof. Now an integration by parts yields

$$p_E(R) = \frac{\partial \tilde{u}}{\partial R}(R; R - s) - \frac{1}{\tilde{F}(R - s)} \int_{\delta(\omega)}^{R-s} \frac{\partial^2 u}{\partial R \partial \delta}(R; \delta) \tilde{F}(\delta) d\delta,$$
or
\[ p_E(R) = \frac{\partial \tilde{u}}{\partial R}(R; R - s) - E_{\delta} \left[ \frac{\partial^2 \tilde{u}}{\partial R^2}(R; \delta) \tilde{F}(\delta|\omega) \frac{\tilde{f}(\delta|\omega)}{f(\delta)} \mid \delta \leq R - s \right] . \]

Now differentiating (9) gives
\[
\frac{\partial p_E}{\partial R}(R) = E_{\delta} \left[ \frac{\partial^2 \tilde{u}}{\partial R^2}(R; \delta) \frac{\tilde{f}(\delta|\omega)}{f(\delta)} \tilde{F}(\delta|\omega) \tilde{f}(\delta) \mid \delta \leq R - s \right] - \tilde{f}(R - s) \tilde{F}(R - s) \frac{\partial u}{\partial R}(R; R - s) p_E(R) + \tilde{f}(R - s) \frac{\partial \tilde{u}}{\partial R}(R; R - s). \]

Substituting the expression obtained for \( p_E(R) \):
\[
\frac{\partial p_E}{\partial R}(R) = E_{\delta} \left[ \frac{\partial^2 \tilde{u}}{\partial R^2}(R; \delta) + \frac{\partial^2 \tilde{u}}{\partial R \partial \delta}(R; \delta) \frac{\tilde{f}(R - s)}{F(R - s)} \frac{\tilde{F}(\delta)}{f(\delta)} \mid \delta \leq R - s \right] . \]

By the log-concavity of \( \tilde{F} \), the factor
\[
\frac{\tilde{f}(R - s) \tilde{F}(\delta)}{\tilde{F}(R - s) f(\delta)}
\]

is smaller than 1. By concavity of the utility function, the first term is negative. By the assumption on the marginal utility of income of the newcomer into the labor force
\[
\frac{\partial^2 \tilde{u}}{\partial R^2}(R; \delta) + \frac{\partial^2 \tilde{u}}{\partial R \partial \delta}(R; \delta)
\]
is non positive. The result follows.

In the benchmark case described at the end of Section 2.1, the derivative \( \frac{d\tilde{u}}{dR} \) is identically equal to zero. Then a sufficient condition for \( p_E \) to decrease with \( R \) is the log-concavity of the cumulative distribution of work opportunity costs.

In the remainder of the paper, we shall mainly consider situations where the following assumption holds:

**Assumption 2.** The average social weight of the employees \( p_E(R, s|\omega) \) is a decreasing function of \( R \).

### 4 The second best programme

Let \( \lambda \) be the multiplier associated with the feasibility constraint (11). The Lagrangian of the problem is
\[
\int_\omega L(R(\omega), s; \omega) \, dG(\omega)
\]
where

\[ L(R, s; \omega) = \int_\alpha^{a(R-s, s, \omega)} u(R; \alpha, \omega) \, dF(\alpha|\omega) + \int_{a(R-s, s, \omega)}^\omega v(s; \alpha, \omega) \, dF(\alpha|\omega) + \lambda[\omega - R - s] F(a(R - s, s, \omega)|\omega) - \lambda s. \]

Of course, \( L \) can be written equivalently with work opportunity costs, and both forms are useful, depending on circumstances:

\[
L(R, s; \omega) = \int_\delta^{R-s} \tilde{u}(R; \delta, \omega, s) \, d\tilde{F}(\delta|\omega, s) + \int_{R-s}^\omega \tilde{v}(s; \delta, \omega) \, d\tilde{F}(\delta|\omega, s) + \lambda[\omega - R + s] \tilde{F}(R - s|\omega, s) - \lambda s.
\]

Unfortunately, the function \( L \) is not concave in its arguments. Any utilitarian optimum however must satisfy first order necessary conditions. At any point \( \omega \) where \( R \) is strictly increasing and \( R - s \) is in the interval \((\delta, \overline{\delta})\), i.e. some, but not all, agents of productivity \( \omega \) want to work, it satisfies the first order condition for a pointwise maximum:\footnote{The second order condition is

\[
\frac{\partial^2 L}{\partial R^2}(R, s; \omega) = \lambda[\omega - R + s] \tilde{f}'(R - s|\omega, s) - 2\lambda \tilde{f}(R - s|\omega, s) + \int_\delta^{R-s} \tilde{u}''(R; \delta, \omega, s) \, d\tilde{F}(\delta|\omega, s) + \int_{R-s}^\omega \tilde{u}'(R; \delta, \omega, s) \, d\tilde{F}(\delta|\omega, s).
\]

} \( \frac{\partial L}{\partial R}(R, s; \omega) = \lambda[\omega - R + s] \tilde{f}(R - s|\omega, s) - \tilde{F}(R - s|\omega, s)[\lambda - p_E(R, s|\omega)] = 0, \) \( (11) \)

In general, there may exist several solutions to the first order condition, corresponding to local maxima or minima. \footnote{The optimum may involve pooling, with regions where \( R \) stays constant because of the monotonicity condition. In a pooling interval \([\omega_1, \omega_2]\), whenever \( R - s \) does not hit the lower bound \( \max_{\omega \in [\omega_1, \omega_2]} \delta(\omega) \), the first order conditions become

\[
\int_{\omega_1}^{\omega_2} \frac{\partial L}{\partial R}(R, s; \omega) \, dG(\omega) = 0
\]

and

\[
\int_{\omega_1}^{\omega_2} \frac{\partial L}{\partial R}(R, s; \omega) \, dG(\omega) \leq 0
\]

for all \( \omega_1 \leq \omega \leq \omega_2 \).}
where \( p_E(R, s|\omega) \) is the previously defined average social weight of the working agents of productivity \( \omega \).

The expression of \( \partial L/\partial R \) has a direct economic interpretation. The first term \( \lambda[\omega - R + s]f(R - s) \) is the gain in government income obtained from the new \( f(R - s) \) workers who participate following an increase in \( R \): they produce \( \omega \), they are paid \( R(\omega) \) but do not receive the subsistence income \( s \) anymore. The second term \( F(R - s)[\lambda - p_E(R, s)] \) is the loss on the existing workers: the marginal cost is \( \lambda \) per worker, while the social value of this distribution of income is equal to the average social weight of the employees of productivity \( \omega \).

The average tax rate supported by the workers of productivity \( \omega \) is \( \tau(\omega) = [\omega - R(\omega) + s]/\omega \), so that the first order condition can be rewritten as

\[
\omega - R + s = \omega \tau(\omega) = \frac{\tilde{F}[R - s|\omega, s]}{\tilde{f}[R - s|\omega, s]}[1 - \frac{p_E(R, s|\omega)}{\lambda}].
\]  

(12)

An immediate consequence is:

**Proposition 2.** Suppose that Assumptions 1 and 2 hold, and consider an optimal allocation at a point \( \omega \) where the tax schedule \( R(\omega) \) is strictly increasing. Then

either \( p_E(\omega, s; \omega) < \lambda \), the financial incentive to work \( R(\omega) - s \) is smaller than before tax income \( \omega \): labor supply is distorted downwards compared to laissez-faire;

or \( p_E(\omega, s; \omega) = \lambda \), the financial incentive to work \( R(\omega) - s \) equals before tax income \( \omega \): labor supply is the same as in laissez-faire;

or \( p_E(\omega, s; \omega) > \lambda \), the financial incentive to work \( R(\omega) - s \) is larger than before tax income \( \omega \): labor supply is distorted upwards compared to laissez-faire.

Social weights larger than \( \lambda \), corresponding to a group of employees whose average social weight is larger than the marginal cost of public funds, receive a financial incentive to work \( R(\omega) - s \) larger than their productivity \( \omega \). Their labor supplies are distorted upwards, compared with laissez-faire.

To complete the study of the optimal tax problem, the value of the marginal cost of public funds \( \lambda \) is linked to another necessary condition, associated with a translation of the overall income, i.e. an equal marginal change in both \( s \) and \( R(\omega) \) for all \( \omega \) (for this differentiation, it is more transparent to use the initial expression in the (structural) parameter \( \alpha \)):

\[
\int_\Omega \left\{ F[a(R(\omega) - s, s, \omega)|\omega]p_E(R(\omega), s|\omega) + (1 - F[a(R(\omega) - s, s, \omega)|\omega])p_U(R(\omega), s|\omega) + \lambda \{[\omega - R + s]f(a(R(\omega) - s, s, \omega)|\omega)d_2'(R(\omega) - s, s, \omega) - 1\} \right\} \, dG(\omega) = 0.
\]  

(13)

*Since the optimal schedule is locally (strictly) increasing, the first order condition (12) for a pointwise maximization holds.*
If one supposes that the first order condition (11) holds everywhere\(^9\), one can substitute the element coming out of the budget constraint which involves the income effect\(^10\) to get

\[
\int_{\Omega} \left\{ F \left(1 - \frac{a_2'}{a_1'} \right) p_E(R(\omega), s|\omega) + (1 - F) p_U(R(\omega), s|\omega) \right. \\
- \left. \lambda \left\{ 1 - \frac{a_2'}{a_1'} F \right\} \right\} dG(\omega) = 0. 
\]

(14)

In the above equation, for each \(\omega\), the coefficients of \(p_E\) and of \(p_U\) are non negative (recall (2)) and sum up to the coefficient of \(\lambda\), \(F(1 - a_2'/a_1') + (1 - F) = (1 - F a_2'/a_1')\). The marginal cost of public funds is the overall average of the social weights, the weights being a bit tilted to account for the effects of subsistence income on labor supply. When leisure is a normal good, the employees are valued more than in the absence of income effect.

**Remark 4.1.** A limit case of some technical interest is the situation where everyone has the same work opportunity cost, say \(\delta_0\) (note that then Assumption \(\text{II}\) does not hold). This situation is studied by Homburg (2002). We limit our attention to the benchmark case of the end of Section 2.1. Then (10) is to be maximized over \(R\), \(R\) non decreasing, with

\[
L(R, s; \omega) = \begin{cases} 
U(s) - \lambda s & \text{if } R - s < \delta_0, \\
U(R - \delta_0) + \lambda [\omega - R] & \text{if } R - s > \delta_0,
\end{cases}
\]

and any of the two quantities when \(R - s = \delta_0\): the agent is indifferent between working or not, and the planner can choose the preferred outcome. Then the first order condition (11) with respect to \(R(\omega)\) for an unconstrained worker of productivity \(\omega\) is:

\[
U'(R(\omega) - \delta_0) = \lambda.
\]

This implies that there is pooling: the incomes of the workers \(R(\omega)\) do not depend on their productivities and satisfy the above equality. Furthermore the feasibility constraint (11) here becomes

\[
\int_{\Omega} [\omega - R + s] \mathbf{1}_{\text{workers}} dG(\omega) = s.
\]

An optimum is characterized by two quantities \((R, s)\) linked by the feasibility constraint, and by the set of workers which has to be chosen when \(R - s\) is equal to \(\delta_0\). Consequently, there are three possibilities:

\(^9\)One must consider separately the situations where \(R - s\) is less than or equal to \(\delta\), belongs to \((\delta, \bar{\delta})\), or is larger than or equal to \(\bar{\delta}\). The formula (14) is valid provided that \(\partial L/\partial R\) is equal to zero for almost all \(\omega\).

\(^10\)By definition, \(\tilde{F}(\delta|\omega, s)\) is equal to \(F(a(\delta, s, \omega)|\omega)\). Therefore \(\tilde{f}\) is equal to \(fa_1'\).
1. If $R - \delta_0 < s$, nobody works. By feasibility $s$ is equal to zero. This is probably a global minimum.

2. If $R - \delta_0 > s$, everybody works. The constant income $R$ is equal to the average productivity in the economy $\int \omega \, dG(\omega)$. The social welfare is equal to

$$U \left( \int \omega \, dG(\omega) - \delta_0 \right).$$

3. Finally, when there is indifference between working or not, $R - \delta_0 = s$, and looking at the expression of $L(R, s; \omega)$, the planner decides to put to work the agents with productivity $\omega$ at least as large as $R - s$. The feasibility condition gives the value of $s$

$$\int_{\Omega} \max[\omega - \delta_0, 0] \, dG(\omega) = s.$$ 

In all circumstances, the utilitarian optimum does not leave any surplus to the workers and everyone is treated equally. Heterogeneity in the form of some dispersion of work opportunity costs gives more scope for redistribution, based on the unknown value of $\delta$.

5 When is labor supply distorted upwards?

The extensive model has imbedded at its heart two dimensions of heterogeneity, which cannot be reduced to one. This gives a lot of leeway to get results of the type ‘any kind of tax function can occur’ manipulating (12): one can play with the distribution $F(\delta|\omega)$, as in Choné and Laroque (2005) for a Rawlsian planner. Our aim here is to (partially) classify the various shapes of tax schemes that may arise, based on natural assumptions on the distribution of characteristics and on the social welfare function.

We put restrictions on the distribution of the agents characteristics. First, we do not want the correlation between productivity and the opportunity cost of work to play a major role, and in the main analysis we assume independence of the two distributions: $F(\delta|\omega)$ does not depend on $\omega$. It simplifies the exposition to suppose that the lower bound of productivity is not larger than the lower bound of the work opportunity cost. Also we suppose that the distribution of heterogeneity is well behaved. Formally, on top of Assumption 1, we assume

**Assumption 3.** The pdf $F(\delta|\omega, s)$ of the work opportunity cost $\delta$ and the average social weight of the employees $p_E(R, s|\omega)$ are independent of productivity. Furthermore $\ln(F(\delta|s))$ is concave on its support $[\delta, \bar{\delta}]$, and $\bar{\delta} > \omega$.

12
It should be clear that the shape of the optimal tax schedule is related to the value $R_m$ of the financial incentives to work which makes the average social weight of the employed agents equal to the marginal cost of public funds. Formally

**Definition 1.** Let $R_m$ be the smallest root of the equation $p_E(R, s) = \lambda$ if any, with $R_m - s = \delta$ if $p_E(R, s) < \lambda$ for all $R$, and $R_m = +\infty$ when $p_E(R, s) > \lambda$ for all $R$.

The location of $R_m - s$ with respect to the support of the distribution of work opportunity costs $[\delta, \bar{\delta}]$ is an important determinant of the shape of the optimal tax schedule. The following proposition indicates what happens when $R_m$ is very large (proof in the Appendix):

**Proposition 3.** Assume that the average social weight of the unemployed is at least as large as that of the employed for all $R > \delta$, and that the average social weight of the employees $p_E(R, s)$ decreases with $R$.

Under Assumption 3, when $R_m - s \geq \delta$, all the agents in the economy work and receive the same after-tax income $R_m$.

Typically, a redistributive government puts as much weight on the unemployed as on the workers. Then an allocation such that $R_m - s \geq \delta$ probably can only occur in rich economies with high productivities: then everyone receives the very high wage $R_m$, possibly larger than one’s productivity. These are not the allocations of interest to us in the remainder of the paper, where we shall limit ourselves to the (more realistic) situations where $R_m$ is smaller than $\bar{\delta}$.

What remains to be done is to separate the cases where $R_m - s$ is larger than the minimum work opportunity cost, implying an upward distortion of the labor supply of the low skilled workers, from those where $R_m - s$ is stuck at the lower end of the range of work opportunity costs. Both cases may occur:

**Proposition 4.** Consider an economy satisfying Assumptions 1 to 3. Suppose furthermore that

1. there are no income effects;
2. the social weight put on the unemployed is equal to the weight given to a marginal employee:

$$p_U(R, s|\omega) = \lim_{R \to s+\delta} p_E(R, s).$$

Then, if $R_m - s = \delta < \bar{\delta}$, nobody works at the optimum allocation: $R(\omega) = s+\delta$ for all $\omega$. 
Proof: Since there are no income effects, (13) simplifies into

$$\int_{\Omega} \left[ F(a(R(\omega) - s, s, \omega))p_E(R(\omega), s) + (1 - F(a(R(\omega) - s, s, \omega)))p_U(R, s|\omega) \right] dG(\omega) = \lambda. \quad (15)$$

All the social weights on the left hand side of (15) are at most equal to \( \lambda \). They must therefore be all equal to \( \lambda \), i.e. \( p_E(R(\omega)) = \lambda \) for all \( \omega \). The strict monotonicity of \( p_E(R) \) yields the desired result.

Theorem 5. Consider an economy satisfying Assumptions [1 to 3]. Suppose furthermore that

1. there are no income effects;
2. the social weight put on the unemployed is equal to the weight given to a marginal employee:

$$p_U(R, s|\omega) = \lim_{R \to s + \delta} p_E(R, s).$$

Any optimal allocation where there is work is such that the labor supply of the less skilled workers is distorted upwards compared with laissez-faire.

Under utilitarianism, in the benchmark model, apart from the special situation where no one works (!), it is optimal to subsidize low skilled work.

When 2. is not satisfied, there are optimal allocations with no upwards distortions of labour supply. This is likely to be the case when the social weight on the unemployed agents is (much) larger than the weight put on the marginal employee.

6 Typical shapes of optimal tax schemes

Under the assumptions posited earlier, one can describe more precisely the qualitative properties of the optimal tax schedule. A possible shape of the optimal incentive schedule is drawn on Figure 1 which obtains in the cases described in the following proposition.

Proposition 6. Consider an economy that satisfies Assumptions [1 to 3] with a finite \( R_m \), \( R_m - s \leq \delta \).

Assume that

$$M(R, s) = R - s + \frac{\hat{F}(R - s)}{f(R - s)}[\lambda - p_E(R, s)]$$
is strictly increasing in $R$ for $s + \delta \leq R \leq R_m$. Then there is no pooling at the optimum for $\delta \leq \omega \leq M(\delta + s, s)$. The optimal tax schedule $R(\omega)$ is uniquely defined for all $\omega$ such that $M(\delta + s, s) \geq \omega \geq \delta$, and is solution to the equation

$$M(R(\omega), s) = \omega,$$

on this interval. Furthermore $R(\omega)$ is an increasing function of $\omega$ on $s + [\delta, M(\delta, s)]$ which satisfies

$$R(\delta) - s = \delta,$$

$$R(\omega) - s \geq \omega \text{ whenever } R \leq R_m.$$

Finally, when $\delta$ is finite, $R(\omega)$ is constant, equal to $s + \delta$, for $\omega$ larger than $s + M(\delta, s)$.

**Proof:** Note that $M(R, s)$ is increasing for $R > R_m$, since the last term in its expression is the product of two nonnegative positive non-decreasing functions.

We look for the function $R(\omega)$ which maximizes $\int_\omega L(R(\omega), s; \omega) \, dG(\omega)$. We have

$$\frac{\partial L}{\partial R}(R, s; \omega) = \begin{cases} f(R - s)[\omega - M(R, s)] & \text{for } \delta \geq R - s \geq \delta \\ p_E(R, s) - \lambda & \text{for } R - s > \delta. \end{cases}$$

At the lower end of the domain, when $\delta > \delta$:

$$\frac{\partial L}{\partial R}(s + \delta, s; \omega) = (\omega - \delta)f(\delta) \geq 0.$$
The preceding computations imply that, under the monotonicity assumption of \( M(R, s) \), the function \( L(., s; \omega) \) has a single maximum in the interval \( s + [\delta, \overline{\delta}] \), which is the unique root \( \rho(\omega) \) of the equation \( M(., s) = \omega \) when \( \omega \leq M(s + \delta, s) \), and is equal to \( s + \overline{\delta} \) for \( \omega \geq M(s + \overline{\delta}, s) \). On the half line \( s + [\delta, +\infty) \), \( L(., s; \omega) \) also has a single maximum, which is equal to \( \max(s + \delta, R_m) \), and is decreasing whenever \( R_m \) is smaller than \( s + \delta \).

Following Proposition 3, we focus on the case \( R_m \leq s + \delta \). Then the function \( L(., s; \omega) \) has a unique root \( \rho(\omega) \) of the equation \( M(., s) = \omega \) when \( \omega \leq M(s + \delta, s) \), and is equal to \( s + \delta \) for \( \omega \geq M(s + \delta, s) \).

The location of \( R(\omega) \) with respect to the 45 degree line is a straightforward consequence of the shape of \( M(R, s) \).

\[ \rho(\omega) = \max(s + \delta, \lambda) \] for all \( \omega \)’s. This pointwise maximization yields a non-decreasing function \( R(\omega) \), and therefore corresponds to the global optimum. The location of \( R(\omega) \) with respect to the 45 degree line is a straightforward consequence of the shape of \( M(R, s) \).

Figure 1 illustrates the two foregoing propositions in the ‘well-behaved’ situation. The financial incentives to work are a continuous increasing function of productivity. Under Assumptions 1 and 2, there is a low skilled region, \( \delta \leq \omega \leq \lambda - s \), where labor supply is distorted upwards, while for higher productivities labor is taxed and the marginal tax rate is positive. In the more restricted case of Proposition 6, the marginal tax rate is negative for low enough productivities (indeed, since \( R(\delta) = s + \delta \) and \( R(\omega) - s > \omega \) in a neighborhood, \( R' \) has to be larger than one in the region). Of course, if one is interested in subsidizing low skilled work, the significance of these results hinges on the extent of the region \( [\delta, \lambda - s] \).

One can examine situations more complicated than the ones described in Proposition 6 or Figure 1. Specifically, the function \( M(R, s) \) may very well be non-increasing for \( R < R_m \), in which case the first order condition \( \omega = M(R, s) \) typically has several solutions. The proof of Proposition 6 goes through by selecting the solution \( R(\omega) \) associated with the global maximum of \( L(R, s; \omega) \), provided this selection is increasing in \( \omega \). The shape of the tax schedule in the region below \( R_m \) then could look quite different, for instance having \( R(\delta) > s + \delta \) and possibly exhibiting upward discontinuities at solution switches. This is illustrated in the following example (proof in the Appendix):

**Proposition 7.** Consider a benchmark economy satisfying Assumptions 1 and 3. Suppose that the opportunity cost \( \delta \) is uniformly distributed on \( [\delta, \overline{\delta}] \) and that \( s + \delta < R_m < s + \overline{\delta} \).

Then \( R(\omega) \) is increasing and concave whenever some agents of productivity \( \omega \) work, i.e. on the set \( \{ \omega | R(\omega) - s > \delta \} \). Moreover:

1. If \( p_U(s) \leq 2\lambda \), the conditions of Proposition 6 are verified, \( R(\delta) - s = \delta \) and \( R'(\delta) = 1/(2\lambda - p_U(s)) > 1 \). At the optimum, none of the agents of productivity smaller than \( \delta \) work.

11It is similar to Figure IIa in Saez (2002), who discusses from a more applied perspective the occurrence of negative marginal tax rates.
2. If \( p_U(s) > 2\lambda \), there exists \( \omega_0 \), \( \omega \leq \omega_0 < \delta \), such that \( R(\omega) - s > \delta \) for all \( \omega \geq \omega_0 \) and \( R(\omega) - s \leq \delta \) for productivities smaller than \( \omega_0 \). There is an upwards discontinuity in the incentives to work at \( \omega_0 \).

The situation where the social weights of the unemployed agents are high \((p_U(s) > 2\lambda)\) is shown on Figure 2. None of the agents with very low productivities, \( \omega < \omega_0 \), work. But for all \( \omega \) larger than or equal to \( \omega_0 \), a fraction of the agents do some work. In fact the upwards distortion to labor supply here is particularly strong: some agents with productivity smaller than the minimal cost of going to work participate in the labor force. The curve AB on the Figure has equation \( M(R, s) = \omega \): it describes the roots of the first order condition. There is a single root, corresponding to a global maximum of \( L \) for \( \omega \) larger than \( \delta \), but there are two roots in a part of the low productivity region. The bold line describes the solution. The curve is concave, implying a progressive tax system. It is not always the case that there are negative marginal tax rates at the beginning of the curve, close to \( \omega_0 \), contrary to the situation when \( p_U(s) < 2\lambda \) of Figure 1. But there is an upwards discontinuity in the tax schedule at \( \omega_0 \), indeed an infinite negative marginal tax rate.

Remark 6.1. We have focused on the shape of incentives in the low productivity region. We do not attempt a full classification of the optimal tax schedules, which satisfy other properties. For instance, Theorem 6 of Choné and Laroque (2005) applies here: all the utilitarian optimal allocations correspond to incentive schemes located above the Rawlsian (Laffer) curve. Theorem 3 of Laroque (2005)
also applies: any incentive scheme above the Laffer curve which does not overtax and such that \( R(\omega) - s \leq \omega \) corresponds to a second best optimal allocation. Note that in a benchmark model, from the above results, none of these allocations satisfy a utilitarian criterion. All the utilitarian optimal allocations are such that \( R(\omega) - s > \omega \) for some \( \omega \)'s, a property discussed in Remark 2.3 of Laroque (2005).

7 Discussion and extensions

In practice, situations where the social weight attached to the unemployed agents is larger than that attached to the employees abound: for instance this would be the case in the presence of ‘involuntary’ unemployment, or, in the spirit of the discussion of the intensive model, when a large opportunity cost to work is associated with a handicap (the marginal social value \( \hat{v}'(s; \delta, \omega) \) is increasing with \( \delta \)). It is then easy to think of economies where the average social weight of the lowest paid workers is smaller than the marginal value of public funds \( (R_m - s = \hat{\delta}) \). In these economies, after tax income is everywhere smaller than productivity.

Similarly, the analysis has proceeded under the assumption that the social welfare function is smooth, so that the distribution of the agents’ weights has no mass point. When it does have a mass point on the unemployed agents, Theorem 5 does not apply and there is no warranty that \( R_m - s > \hat{\delta} \). In particular, the case of a Rawlsian planner who puts a Dirac mass on the least favored agent in the economy corresponds here to a situation where \( p_E(D) \) is equal to zero for all \( D \), and \( R_m - s = \hat{\delta} \). Then the optimal incentive scheme satisfies 1. of Proposition 2: it is everywhere smaller than productivity and the marginal tax rate is always positive. This is in line with the results of Choné and Laroque (2005).

It would be of interest to know whether and when the subsidy result still holds in the mixed situation where both the extensive and intensive margins operate. Boone and Bovenberg (2004) analyze such a model where the utility is quasi linear (but linear in hours of work, rather than in consumption as here). There is a fixed cost of searching for a job work which is constant across the population, and the random outcome of search creates heterogeneity. They find cases where work is subsidized (Section 4.3), but do not characterize them in terms of economic fundamentals. More work is needed in this area.
References


A Appendix

Lemma 1. Under Assumption 1, suppose that \( R(\omega) = \bar{R} \) is constant on some non degenerate interval of \([\omega, \overline{\omega}]\) (pooling), with \( \bar{R} + s \in (\delta, \delta) \).

Then \( R(\omega) \) is equal to \( \bar{R} \) for all \( \omega \)'s (full pooling) and \( p_E(\bar{R}) \leq \lambda \).

Proof of Lemma 1

In the pooling interval \([\omega_1, \omega_2]\), the partial \( \partial L(R(\omega), s; \omega) / \partial R \) is an increasing and affine function of \( \omega \). Since we must have

\[
\int_{\omega_1}^{\omega_2} \frac{\partial L}{\partial R}(R(\omega); \omega) \, dG(\omega) = 0 \tag{16}
\]

on every pooling interval, it must be the case that \( \partial L / \partial R \) is negative at \( \omega_1 \) and positive at \( \omega_2 \). Suppose \( \omega_1 > \omega_2 \). Then \( R \) must be continuous at \( \omega_1 \) (otherwise, decreasing \( R \) around \( \omega_1 \) would increase the Lagrangian and respect the monotonicity constraint). It follows that \( \partial L / \partial R \) is zero at \( \omega_1 \), implying that \( \partial L / \partial R \geq 0 \) on \([\omega_1, \omega_2]\) and contradicting (16).

Then we must have \( \omega_1 = \omega_2 \). The same argument applies at the top of the distribution, showing that \( \omega_2 = \overline{\omega} \). So if there is pooling, it must be full pooling, that is \( R = \bar{R} \) on the whole range of \( \omega \).

Now integrating (11) over \( \omega \), it follows from the feasibility constraint (4) and the non negativity of \( s \) that \( p_E(\bar{R}) \leq \lambda \).

Proof of Proposition 3

1) We first show that if \( R(\omega) - s \geq \delta \), then \( R(\omega) = R_m \). At any \( R \) such that \( R - s \geq \delta \), the derivative of the Lagrangian (11) reduces to

\[
\frac{\partial L}{\partial R} = \tilde{F}(R - s|s)[p_E(R, s) - \lambda]. \tag{17}
\]

Consider the optimum \( R \) and argue by contradiction. If \( R(\omega) \) were not equal to \( R_m \) whenever \( R(\omega) - s \geq \delta \), then putting \( R(\omega) = R_m \) for all \( \omega \)'s in the region would satisfy the monotonicity requirement on \( R \) and increase the Lagrangian, the desired contradiction.

2) We now show that there are no \( \omega \)'s such that \( R(\omega) - s < \delta \). From Lemma 1, there are two possibilities: either there is full pooling, or the first order condition holds at every point in the considered domain.

In the former case, using (13) and the first order condition associated with pooling yields (all quantities being evaluated at \( \bar{R} \))

\[
F(1 - \frac{a_2'}{a_1'})p_E(\bar{R}, s) + (1 - F)p_U(\bar{R}, s) = \lambda(1 - \frac{a_2'}{a_1'}F).
\]

Recalling (2), the coefficient of \( p_E \) on the left hand side is positive. It follows that \( \lambda \) is a weighted average of \( p_E \) and \( p_U \), a contradiction with the fact that \( p_U \geq p_E > \lambda \).
In the latter case, $\[\text{11}\]$ holds everywhere, the equality $\[\text{14}\]$ also holds, which implies that all the agents in the economy have a social weight, or marginal utility of income, equal to $\lambda$. Consequently there are no $\omega$’s with $R(\omega) - s < \delta$. ■

**Proof of Proposition 7** Let

$$\lambda = \int \int U'[s + \max(0, R(\omega) - s - \delta)] d\tilde{F}(\delta) dG(\omega).$$

Then

$$p_E(R) = \frac{1}{F(R - s)} \int_{\delta}^{R - s} U'[R - \delta] \frac{d\delta}{\delta - \delta}.$$ Integrating and substituting yields

$$M(R, s) = (1 + \lambda)(R - s) - \lambda\delta - [U(R - \delta) - U(s)].$$

The function $M(R, s)$ is strictly convex in $R$ and $M_R(\delta, s) = 1 + \lambda - p_U(s)$.

1) Case $p_U(s) \leq 1 + \lambda$. $M(R, s)$ is strictly increasing in $R$ and Proposition 6 applies. The convexity of $M(\cdot, s)$ implies the concavity of $R(\omega)$.

2) Case $p_U(s) > 1 + \lambda$. As in the proof of Proposition 6 we consider the pointwise maximum of $L(R, s; \omega)$ for $R - s \geq \delta$. Since it is increasing in $\omega$, it satisfies the monotonicity condition and is the optimum.

Recall that $L(s + \hat{\delta}, s; \omega) = 0$. Now,

$$\frac{\partial L}{\partial R}(R, s; \omega) = (\omega - M(R, s)) f(R - s) = \frac{1}{\delta - \delta} (\omega - M(R, s))$$

for $\delta \leq R - s \leq \bar{\delta}$ is a concave function of $R$ which becomes negative for large enough $R$. We consider three cases:

a. For $\omega > \delta$, $\partial L/\partial R(s + \hat{\delta}; \omega)$ is positive. There is a single zero $R(\omega)$ of the derivative, solution to $\omega = M(R, s)$, which maximizes $L(R, s; \omega)$.

b. For $\omega = \delta$, $\partial L/\partial R(s + \hat{\delta}; s; \omega)$ is equal to zero. $\partial^2 L/\partial R^2(s + \hat{\delta}; s; \omega) = (p_U(s) - 1) \delta(\delta - \bar{\delta})$ is positive, so that there is another root $R(\delta)$, larger than $s + \delta$ ($R = s + \delta$ is a local minimum of $L$). Recall that $L(s + \hat{\delta}, s; \omega)$ is equal to zero for all $\omega$: the maximum is positive.

c. Finally consider $\omega < \delta$. The function $\partial L/\partial R(\cdot, s; \omega)$ is linear increasing in $\omega$: when $\omega$ decreases from $\delta$, its smallest root increases, its largest root (a local maximum of $L$), say $\Delta(\omega)$, decreases, until eventually they both disappear, say at $\omega_1, \omega_1 < \delta$. Note that $L(\Delta(\omega), s; \omega)$ is an increasing function of $\omega$. Since $L(\delta, s; \delta) = 0, L(\Delta(\omega_1), s; \omega_1)$ is negative. Let $\omega_2, \omega_2 > \omega_1$, be such that $L(\Delta(\omega_2), s; \omega_2)$ is equal to zero. Define $\omega_0 = \max(\omega, \omega_2)$, $R(\omega) - s = \Delta(\omega)$ for $\omega_0 \leq \omega \leq \delta$, and $R(\omega) - s = \delta$ for $\omega$ smaller than $\omega_0$.

It is easy to check that the $R(\omega)$ function thus defined indeed is the solution of the problem. ■