UNIVERSITEIT GENT

FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

A comparison of optimal tax policies when compensation or responsibility matter

Laurence Jacquet

Dirk Van de gaer

October 2009

2009/615

D/2009/7012/67

A comparison of optimal tax policies when compensation or responsibility matter*

Laurence Jacquet†
Norvegian School of Economics and Business Administration

Dirk Van de gaer‡
Ghent University and CORE

October 7, 2009

Abstract

This paper examines optimal redistribution in a model with high and low-skilled individuals with heterogeneous tastes for labor, that either work or not. With such double heterogeneity, traditional Welfarist criteria including Utilitarianism fail to take the compensation-responsibility trade-off into account. As a response, several other criteria have been proposed in the literature. This paper is the first to compare the extent to which optimal policies based on different normative criteria obey the principles of compensation (for differential skills) and responsibility (for preferences for labor), when labor supply is along the extensive margin. The criteria from the social choice literature perform better in this regard than the traditional criteria, both in first and second best. More importantly, these equality of opportunity criteria push the second best policy away from an Earned Income Tax Credit and in the direction of a Negative Income tax.

Key Words: optimal income taxation, equality of opportunity, heterogeneous preferences for labor.
JEL Classification: H21, D63

---

*We thank Marc Fleurbaey, Michael Hoy, Etienne Lehmann, Roland Iwan Luttens, Pierre Pestieau and Matohiro Sato for their valuable comments and suggestions. Laurence Jacquet would like to thank Sturla Amundsens fundings to NHH. This paper presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by its authors.

†Norwegian School of Economics and Business Administration, Economics Department, Helleveien 30, 5045 Bergen, Norway and CESifo. Email: laurence.jacquet@nhh.no

‡SHERPPA, Faculteit Economie en Bedrijfskunde, Ghent University, Ghent and Center of Operational Research and Econometrics (CORE), Louvain-la-Neuve, Belgium. Email: dirk.vandegaer@ugent.be.

1 Introduction

Assuming labor supply along the participation (also called extensive) margin implies that a larger transfer towards low-paid workers than inactive people, i.e. an Earned Income Tax Credit (EITC), may become part of an optimal tax system (Diamond, 1980; Saez, 2002; Brewer *et al.* (2008); Choné and Laroque, 2009). This well-known result is obtained under Utilitarian social preferences while agents differ in terms of skills as well as preferences. However, it is commonly admitted that preference heterogeneity poses ethical questions which challenge standard objective functions like Utilitarianism, see, e.g., Rawls (1971), Sen (1980) and Dworkin (1981). Other normative criteria based on fairness requirements have been proposed in the social choice literature. However, they are scarcely used to derive optimal tax policies. The optimal income tax literature itself considers alternative social preferences but always with labor supply along the intensive margin. For instance, Boadway *et al.* (2002) use a Utilitarian social welfare function where different weights can be assigned to individuals with different preferences for leisure. This amounts to using different cardinalizations of individual utility functions. Paternalistic criteria, in which the planner uses a reference value for the taste for work and maximizes the sum of these adjusted utilities have also been considered, by, e.g. Schokkaert *et al.* (2004). Assuming high and low-skilled agents with heterogeneous tastes for labor, labor supply along the participation margin, this paper compares the optimal tax policies under a large set of social preferences from the social choice and the optimal taxation literature. We show that the social choice inspired criteria provide an additional argument for an optimal tax system away from the EITC. A lower transfer towards low-paid workers than inactive people, i.e. a Negative Income Tax (NIT), is more likely to become optimal. Moreover, under the assumption that the low-skilled have at least as large a participation elasticity as the high-skilled agents, the labor supply distortion for the highly skilled is tempered.

The second contribution of this paper is to check the optimal tax policies against equality of opportunity requirements. The dominant branch of the equality of opportunity literature, liberal egalitarian theories of justice, argues that income or welfare inequalities arising from non-responsibility factors such as innate skills should be eliminated (the compensation principle) and inequalities arising from responsibility factors such as preferences should be respected (the responsibility principle).[1] This paper then checks the optimal schedules we obtain using the criteria from social choice and also the ones from the optimal income tax literature against the compensation and responsibility principles. Unsurprisingly, the criteria which originate from the social choice approach to equality of opportunity perform much better than the traditional criteria, both under full and asymmetric information. Under the latter assumption, we also consider an alternative strategy that restricts the search for an optimal tax policy satisfying one of the equality of opportunity principles.

The third contribution is to propose five new normative criteria which satisfy priority to the worst-off (and thus weak) versions of the compensation and responsibility principles. They rely on a cardinal or, alternatively, on an ordinal measures of welfare. We show that these criteria, just like the social choice inspired criteria push the optimal tax away from an earned income tax credit and temper the labor supply distortion of the highly skilled.

[1] For an overview of this literature, see Fleurbaey (2008) or Fleurbaey and Maniquet (2009). The (in-) compatibility of these two principles was first analyzed by Fleurbaey (1995a) and Bossert (1995).

The paper is organized as follows. In Section 2, we describe the model, provide the characterization of the individuals' behavior, and describe the decision variables of the government under full and asymmetric information. Section 3 states the axioms behind equality of opportunity and presents the distinct objective functions. Section 4 investigates the optimal tax policies under full information, which is in Section 5 extended to the asymmetric information economy. Sufficient conditions for a NIT or a EITC are given. Section 6 concludes the paper. All proofs are gathered in appendix.

2 The model

2.1 Individual behavior

Assume agents decide whether to work or not.[2] They differ along two dimensions: their skill and their disutility of work. Skills take two values, $w_H > w_L > 0$, which correspond with the wages given that the production function exhibits constant returns to scale. The disutility of work, α, is distributed according to the cumulative distribution function $F(\alpha) : \mathbb{R}^+ \to [0,1] : \alpha \to F(\alpha)$ and the corresponding density function $f(\alpha)$. The latter is continuous and positive over its domain.[3] These functions are common knowledge. The proportion of low-skilled agents (or w_L-type) in the population is given by γ, $1-\gamma$ is the proportion of high-skilled people (or w_H-type). We assume that productivity and labor disutility are independently distributed. Utility is quasilinear and represented by:

$$v(x) - \alpha \text{ if they work,}$$
$$v(x) \text{ if they do not work,}$$

where x is consumption, $v(x) : \mathbb{R}^+ \to \mathbb{R} : x \to v(x)$ with $v' > 0 \geq v''$ and $\lim_{x\to\infty} v'(x) = 0$.

2.2 The government's decisions

Under full information (so-called first best), the government implements a tax policy depending on α and $w_Y (Y := L, H)$ hence it also assigns individuals to low-skilled jobs (where the gross wage is w_L), to high-skilled jobs (where the gross wage is w_H) or to inactivity (activity u). Activity assignment is captured through the functions $\delta_L(\alpha) : \mathbb{R}^+ \to \{0,1\} : \delta_L(\alpha) = 1$ ($\delta_L(\alpha) = 0$) if w_L-agents with this value for α are employed (inactive) and $\delta_H(\alpha) : \mathbb{R}^+ \to \{0,1\} : \delta_H(\alpha) = 1$ ($\delta_H(\alpha) = 0$) if w_H-agents with this value for α are employed (inactive). As a consequence $n_L \stackrel{\text{def}}{\equiv} \int_0^\infty \delta_L(\alpha)\, dF(\alpha)$ ($n_H \stackrel{\text{def}}{\equiv} \int_0^\infty \delta_H(\alpha)\, dF(\alpha)$) is the fraction of w_L-agents (w_H-agents) that are employed. w_L-agents cannot get access to high-skilled jobs, and, since efficiency matters, it will never be optimal that w_H-agents work in low-skilled jobs. By putting these people in high-skilled jobs instead of low-skilled jobs, they produce more which can be used to increase someone's consumption. Hence, formally, the government determines four consumption functions: $x_L^w(\alpha)$ for

[2] There is growing evidence that the extensive margin matters a lot, e.g. Meghir and Philips (2008).

[3] We want to see whether an EITC or a NIT is optimal. This requires us to describe the participation tax rates only. Therefore, it is appropriate to assume a discrete support for skills, like in Saez (2002). For simplicity we assume two skill levels, but increasing the number of skills does not modify our main results. Continuity of α is assumed for simplicity.

the w_L-workers, $x_H^w(\alpha)$ for the w_H-workers, $x_L^u(\alpha)$ for the w_L-inactive agents and $x_H^u(\alpha)$ for the w_H-inactive. All these functions go from $\mathbb{R}^+$ to $\mathbb{R}^+$.

The Government budget constraint can be formulated as follows:

$$\begin{aligned}
&\gamma\left[\int_0^\infty \left[\delta_L(\alpha)\left(w_L - x_L^w(\alpha)\right) - \left(1-\delta_L(\alpha)\right)x_L^u(\alpha)\right]dF(\alpha)\right] \\
&+(1-\gamma)\left[\int_0^\infty \left[\delta_H(\alpha)\left(w_H - x_H^w(\alpha)\right) - \left(1-\delta_H(\alpha)\right)\right)x_H^u(\alpha)\right]dF(\alpha)\right] \geq R,
\end{aligned} \tag{1}$$

where R is an exogenous revenue requirement, which can be positive or negative. This budget constraint must be binding at the optimum as all government objectives considered in the paper are increasing in individuals' consumption.

The problem for the government in the first best is to determine the functions $x_L^w(\alpha)$, $x_H^w(\alpha)$, $x_H^u(\alpha)$, $x_L^u(\alpha)$ together with $\delta_L(\alpha)$ and $\delta_H(\alpha)$ that are normatively desirable and satisfy the government budget constraint (1).

In the second best, the tax schedule can depend only on income levels $(0, w_L$ or $w_H)$. The government then defines three consumption levels x^u, x_L and x_H, denoting consumption levels when not participating in the labor force, when working in low-skilled and in high-skilled jobs, respectively. These consumption levels have to meet the government budget constraint, the set of self-selection constraints (which will be stated in Section 5) and have to be normatively desirable. The next section discusses which normative principles or criteria the government can use.

3 Equality of opportunity

The next subsection formally defines equality of opportunity in order to study whether the normative criteria usually assumed in the optimal tax literature succeed in reaching it.

3.1 Two equality of opportunity principles

Define, for the case where $Y = L$ or H, the evaluation of the consumption bundle $(x_Y(\alpha), \delta_Y(\alpha))$ as

$$u\left(x_Y(\alpha), \delta_Y(\alpha), \alpha^G\right) = \begin{cases} v\left(x_Y^w(\alpha)\right) - \alpha^G & \text{if } \delta_Y(\alpha) = 1, \\ v\left(x_Y^u(\alpha,)\right) & \text{if } \delta_Y(\alpha) = 0, \end{cases}$$

where labor disutility is evaluated by parameter α^G. If $\alpha^G = \alpha$, $u\left(x_Y(\alpha), \delta_Y(\alpha), \alpha^G\right)$ coincides with the individual's own utility.

We assume throughout that people are responsible for their tastes for work α, but not for their skills[4]. We can then apply Fleurbaey (1994) 's approach and capture the intuitions of equality of opportunity in two axioms. The first equality of opportunity axiom expresses the idea of compensation:

[4]Two remarks can be made at this point. First, if people are not responsible for anything, from a perspective of equality of opportunity, the only possible objectives are full equality of utility levels or leximin. Second, it is possible to follow the suggestion by Pestieau and Racionero (2009) to disentangle the parameter α in two components: $\alpha = \alpha_P + \alpha_D$, where people are responsible for α_P (a preference parameter), but not for α_D (a disability parameter). The present framework can be adjusted to deal with this issue, without altering the main results of the paper.

EWEP (Equal Welfare for Equal Preferences):
$\forall\alpha \in \mathbb{R}^{+} : u\left(x_L\left(\alpha\right), \delta_L\left(\alpha\right), \alpha\right) = u\left(x_H\left(\alpha\right), \delta_H\left(\alpha\right), \alpha\right)$.

An allocation satisfying EWEP is such that differences in skills do not influence a person's welfare. The second axiom of equality of opportunity expresses the idea of responsibility:

ETES (Equal Transfers for Equal Skills):
$\forall\alpha, \alpha' : \delta_L\left(\alpha\right) = \delta_L\left(\alpha'\right) = 1$ and $\forall\alpha'' : \delta_L\left(\alpha''\right) = 0 :$
$x_L^w\left(\alpha\right) - w_L = x_L^w\left(\alpha'\right) - w_L = x_L^u\left(\alpha''\right) = x_L^u,$
$\forall\alpha, \alpha' : \delta_H\left(\alpha\right) = \delta_H\left(\alpha'\right) = 1$ and $\forall\alpha'' : \delta_H\left(\alpha''\right) = 0 :$
$x_H^w\left(\alpha\right) - w_H = x_H^w\left(\alpha'\right) - w_H = x_H^u\left(\alpha''\right) = x_H^u,$

with some abuse of notations for the last term in both expressions. The latter emphasizes that taxes only depend on skill level. People are hold responsible for their taste for leisure α. For each skill level all inactive get the same benefit, all workers pay the same tax, and the transfer received by the inactive is equal to minus the tax paid by the workers. Therefore, welfare differences that are caused by differential tastes are not compensated and fully respected.

We formally define full equality of opportunity as follows:

FEO (Full Equality of Opportunity):
An allocation satisfies full equality of opportunity if it satisfies both EWEP and ETES.

In the traditional framework, where the government only (re-)distributes consumption, even in the first best there does not exist a FEO allocation -see, e.g., Fleurbaey (1994) and Bossert (1995). For this reason, Fleurbaey (1995b) suggested weakening at least one of the axioms, while maintaining the other[5]. This allowed him to characterize two allocations. Keeping ETES but requiring EWEP only for situations where all agents have a reference value for the taste parameter $\widetilde{\alpha}$ characterize the conditional equality allocation defined below. Keeping EWEP but requiring ETES only for situations where all agents have a reference value for the resource bundle, here taken to be the consumption level $\widetilde{x}$ and $\delta_Y = 1$ ($Y = L$ or H), characterizes the egalitarian equivalent allocation.

CE (Conditional Equality):
An allocation is the conditional equality allocation if and only if for all α and all Y it equalizes $u\left(x_Y\left(\alpha\right), \delta_Y\left(\alpha\right), \widetilde{\alpha}\right)$ at the highest feasible level.

EE (Egalitarian Equivalence):
An allocation is egalitarian equivalent if and only if for all α and all $Y : u\left(x_Y\left(\alpha\right), \delta_Y\left(\alpha\right), \alpha\right) = u\left(\widetilde{x}, 1, \alpha\right)$ and $\widetilde{x}$ is at the highest feasible level.

The CE allocation ensures that all individuals are equally well off with their actual bundle of resources when this is evaluated using the reference preference $\widetilde{\alpha}$. The EE allocation makes all individuals indifferent between their actual resource bundle and the reference bundle which gives them $\widetilde{x}$ and where they have to work.[6] In our definition here, we incorporate that no resources are

[5] Of course, it is also possible to weaken both axioms simultaneously -see, e.g., Bossert and Fleurbaey (1996) or Fleurbaey and Maniquet (2009).

[6] This is similar to the "full-health equivalent income" proposed by Fleurbaey (2005). An alternative egalitarian equivalent allocation would determine for each individual the consumption level that he needs when he has to be inactive and that is such that he is indifferent to this bundle and his actual consumption bundle.

wasted by, in the CE allocation, equalizing at the highest possible level, and in the EE allocation pursuing indifference at the highest feasible level of $\widetilde{x}$. A CE or EE allocation need not exist. In particular, in the second best, it will not be possible to equalize the reference utilities as required by CE, and, even in the first best, indifference for all individuals with the reference bundle is not feasible in our model. We formulate maximin social orderings inspired by the CE and EE allocation at the end of the next subsection.

3.2 Different social objective functions

The paper will consider the following social objective functions extensively used in the optimal taxation literature.

The Utilitarian social objective function (used in a.o., Ebert (1992), Diamond and Sheshinski (1995), Boadway *et al.* (2000), Hellwig (2007)) is the average of all individual utilities, i.e.

$$S^U = \gamma \int_0^\infty \delta_L(\alpha) \left[v(x_L^w(\alpha)) - \alpha\right] dF(\alpha) + \gamma \int_0^\infty (1 - \delta_L(\alpha))\, v(x_L^u(\alpha)) dF(\alpha) +$$
$$(1-\gamma) \int_0^\infty \delta_H(\alpha) \left[v(x_H^w(\alpha)) - \alpha\right] dF(\alpha) + (1-\gamma) \int_0^\infty (1 - \delta_H(\alpha))\, v(x_H^u(\alpha)) dF(\alpha). \quad (2)$$

Our Welfarist social objective is the average of a concave transformation of individual utilities. The concave transformation allows the expression of inequality aversion with respect to the distribution of utilities. Let the function $\Psi : \mathbb{R} \to \mathbb{R} : a \to \Psi(a)$ be a strictly concave function. Our Welfarist objective function is

$$S^W = \gamma \int_0^\infty \delta_L(\alpha) \Psi\left(v(x_L^w(\alpha)) - \alpha\right) dF(\alpha) + \gamma \int_0^\infty (1 - \delta_L(\alpha)) \Psi\left(v(x_L^u(\alpha))\right) dF(\alpha) +$$
$$(1-\gamma) \int_0^\infty \delta_H(\alpha) \Psi\left(v(x_H^w(\alpha)) - \alpha\right) dF(\alpha) + (1-\gamma) \int_0^\infty (1 - \delta_H(\alpha)) \Psi\left(v(x_H^u(\alpha))\right) dF(\alpha). (3)$$

Assumed in the seminal article of Mirrlees (1971), this welfare function has been very popular since then (e.g., Atkinson and Stiglitz (1980), Diamond (1998), Choné and Laroque (2005), Kaplow (2008), Kleven *et al.* (2009)).

The Boadway *et al.* (2002)'s objective function allows to attach a weight to individuals' utilities that depends on their taste for leisure. Let $W(\alpha) : \mathbb{R}^+ \to \mathbb{R}^+ : \alpha \to W(\alpha)$ be the social welfare weight given to the utility of an individual with disutility of labor equal to α. The Boadway *et al.* objective function is given by

$$S^B = \gamma \int_0^\infty \delta_L(\alpha) W(\alpha) \left[v(x_L^w(\alpha)) - \alpha\right] dF(\alpha) + \gamma \int_0^\infty (1 - \delta_L(\alpha)) W(\alpha) v(x_L^u(\alpha)) dF(\alpha)$$
$$+ (1-\gamma) \int_0^\infty \delta_H(\alpha) W(\alpha) \left[v(x_H^w(\alpha)) - \alpha\right] dF(\alpha)$$
$$+ (1-\gamma) \int_0^\infty (1 - \delta_H(\alpha)) W(\alpha) v(x_H^u(\alpha)) dF(\alpha). \quad (4)$$

This objective function was explicitly introduced to deal with individuals that are heterogeneous in skills and preferences. Also used in Cremer *et al.* (2004 and 2007) for instance, this criterion adopts distinct cardinalizations of individual utilities depending on the individual's taste parameter α.

Our Non-Welfarist social objective function uses a paternalistic view for the valuation of labor disutility. We define the reference labor disutility as $\overline{\alpha} \geq 0$, which is the weight attached by the government to the α of every individual. The social objective becomes

$$S^{NW} = \gamma \left[\int_0^\infty \delta_L(\alpha) \left[v(x_L^w(\alpha)) - \overline{\alpha} \right] dF(\alpha) \right] + \gamma \int_0^\infty (1 - \delta_L(\alpha)) \, v(x_L^u(\alpha)) dF(\alpha)$$
$$+(1-\gamma) \int_0^\infty \delta_H(\alpha) \left[v(x_H^w(\alpha)) - \overline{\alpha} \right] dF(\alpha) + (1-\gamma) \int_0^\infty (1 - \delta_H(\alpha)) \, v(x_H^u(\alpha)) dF(\alpha). \quad (5)$$

With this objective function, the social planner has a different idea than the individuals themselves about the 'correct' or reasonable disutility of work. There is then a clear paternalistic motive for taxation which arises from differences between social and private preferences. Schokkaert *et al.* (2004) consider this social objective function. Marchand *et al.* (2003) and Pestieau and Racionero (2009) consider an alternative paternalistic approach in which the government attaches a larger weight to the labor disutility of disabled individuals. Maximization of non-welfarist social objective functions typically selects allocations that are not Pareto efficient.

To state the next two objective functions, which are less standard, we define an operator that takes the first element of a set with two elements if $\delta(\alpha)$ equals one, and the second element otherwise. Formally, we define the operator as

$$\underset{\delta(\alpha)}{\text{oper}} \{a, b\} = a \text{ if } \delta(\alpha) = 1 \text{ and } \underset{\delta(\alpha)}{\text{oper}} \{a, b\} = b \text{ if } \delta(\alpha) = 0.$$

Roemer (1993 and 1998) proposes that equality of opportunity for welfare holds when the utilities of all those who exercised a comparable degree of responsibility are equal, irrespective of their skills. Assuming that those that have the same preferences have exercised a comparable degree of responsibility, the ideal is to give the same utility to those with the same preferences, irrespective of their skills. Since utilities have to be equal for each preference, it will usually (except, as we will see in the first best) not be possible to achieve this. Roemer therefore suggests to maximize a weighted average of the minimal utilities across individuals having the same tastes. As a result, Fleurbaey (2008) calls this the mean of mins criterion. Roemer' s (1998) objective function can be written as

$$S^R = \int_0^\infty \min \left\{ \underset{\delta_L(\alpha)}{\text{oper}} \{ v(x_L^w(\alpha)) - \alpha, v(x_L^u(\alpha)) \}, \underset{\delta_H(\alpha)}{\text{oper}} \{ v(x_H^w(\alpha)) - \alpha, v(x_H^u(\alpha)) \} \right\} dF(\alpha). \quad (6)$$

For each α, the government assigns low and high skilled individuals to employment or inactivity. The min function in the integral term takes, for each α level, the smallest utility across skill types. The Roemer rule maximizes the sum (over α) of these minimal utility levels. It has been used by Roemer *et al.* (2003) to empirically compare the extent to which fiscal policies manage to equalize opportunities for income acquisition in a set of countries.

While Roemer's proposal is well known, an obvious alternative was proposed by Van de gaer (1993). The starting point is that for each level of skill, utility as a function of the taste parameter can be interpreted as the utilities to which someone with that skill level has access. The proposal is then to maximize the value of the smallest opportunity set, where the opportunity set is the surface under utilities to which he has access, weighted by the frequency with which the corresponding preference parameter occurs. Hence the proposed social objective function, labeled the min of means criterion by Fleurbaey (2008), is

$$S^V = \min\left\{\int_0^\infty \underset{\delta_L(\alpha)}{\text{oper}}\{v\left(x_L^w\left(\alpha\right)\right) - \alpha, v\left(x_L^u\left(\alpha\right)\right)\}dF\left(\alpha\right), \right.$$
$$\left.\int_0^\infty \underset{\delta_H(\alpha)}{\text{oper}}\{v\left(x_H^w\left(\alpha\right)\right) - \alpha, v\left(x_H^u\left(\alpha\right)\right)\}dF\left(\alpha\right)\right\}. \tag{7}$$

This criterion and Roemer's criterion were used to compute optimal linear income taxes in Bossert *et al.* (1999) and Schokkaert *et al.* (2004).[7]

We formulate the maximin objective function inspired by the Conditional Equality (CE) allocation:

$$S^{CE} = \min_{\alpha, w_Y} u\left(x_Y\left(\alpha\right), \delta_Y\left(\alpha\right), \widetilde{\alpha}\right), \tag{8}$$

meaning that the optimal policy is determined such that the lowest level of utility that someone in the population gets with his actual allocation, evaluated at the reference preferences $\widetilde{\alpha}$, is as high as possible. The resulting optimal allocation is not necessarily Pareto efficient. The criterion was explicitly considered by Bossert *et al.* (1999).

Finally, we formulate a maximin objective function inspired by the Egalitarian Equivalent (EE) allocation. For each individual, we determine the consumption level that he needs when he has to work and is such that he is indifferent to this bundle and his actual consumption bundle. Evidently, for workers, this is simply their actual consumption level. Inactive people require a consumption level equal to $v^{-1}\left(v\left(x_Y^u\left(\alpha\right)\right) + \alpha\right)$, where $x_Y^u\left(\alpha\right)$ is their actual consumption level. Hence, we can define an EE ordering as maximizing

$$S^{EE} = \min_{\alpha, w_Y}\left\{x_L^w\left(\alpha\right), x_H^w\left(\alpha\right), v^{-1}\left(v\left(x_L^u\left(\alpha\right)\right) + \alpha\right), v^{-1}\left(v\left(x_H^u\left(\alpha\right)\right) + \alpha\right)\right\}. \tag{9}$$

In our framework, this social ordering is the natural counterpart of the ordering proposed by Fleurbaey and Maniquet (2005 and 2006). In their papers, the equivalent wage for an individual is defined as the wage rate such that he is indifferent between his actual bundle and the bundle that he could reach if he had his equivalent wage. Their proposed social ordering is then to maximize the minimal equivalent wage. Fleurbaey and Maniquet work in an intensive labor supply choice model; the computation of the equivalent wage involves a counterfactual labor supply choice lying between inactivity and full time employment. In our extensive labor supply model, such a choice is not available. However, we can adjust the concept by comparing the actual consumption bundle with the wage making the individual indifferent with full time employment. Formally, in our extensive margin model, the equivalent wage is defined for the employed as $x_Y^{wE}\left(\alpha\right) = x_Y^w\left(\alpha\right)$ and for the inactive as $x^{uE}\left(\alpha\right) : v\left(x^{uE}\left(\alpha\right)\right) - \alpha = v\left(x_Y^u\left(\alpha\right)\right)$, which implies that $x^{uE}\left(\alpha\right) = v^{-1}\left(v\left(x_Y^u\left(\alpha\right)\right) + \alpha\right)$. Maximinning this equivalent wage leads to the social ordering defined in (9).

4 First best

This section studies the optimal policies under full information with criteria that are only loosely based on equality of opportunity principles as well as criteria directly inspired by equality of oppor-

[7] Axiomatic characterizations of these criteria can be found in Ooghe *et al.* (2007) and Fleurbaey (2008).

tunity axioms. We state the analytical properties,[8] interpret them and check whether the EWEP and ETES axioms are satisfied. The Lagrangian multiplier associated to the budget constraint is denoted by λ. The superscript $X = U$, W, B, NW, R, V, CE or EE denotes the variables at the optimum under the Utilitarian, Welfarist, Boadway *et al.*, Non-Welfarist, Roemer, Van de gaer, Conditional Equality and Egalitarian Equivalent objectives, respectively. The following assignment rule, denoted by AAR, often (but not always) defines the optimal activity assignment:

AAR^X (Activity Assignment Rule under social objective X): there exist α_L^{X*} and α_H^{X*} such that $\delta_L^X(\alpha) = 1$ for all $\alpha \leq \alpha_L^{X*}$, $\delta_L^X(\alpha) = 0$ otherwise, $\delta_H^X(\alpha) = 1$ for all $\alpha \leq \alpha_H^{X*}$, $\delta_H^X(\alpha) = 0$ otherwise and $\alpha_L^{X*} \leq \alpha_H^{X*}$.

Under this activity assignment rule, those low- (high-) skilled with disutility from work smaller than α_L^{X*} (α_H^{X*}) are employed, while those with a higher disutility from work are inactive and more highly than lowly skilled are employed.

Theorem 1a. *With full information, the following configuration of policies is optimal:*
(a) Utilitarian planner:
Consumption bundles: $\overline{x}^U \overset{\text{def}}{\equiv} x_L^{wU}(\alpha) = x_L^{uU}(\alpha) = x_H^{wU}(\alpha) = x_H^{uU}(\alpha)$.
Activity assignment: AAR^U.
(b) Welfarist planner:

$$\textit{Consumption bundles:}\begin{cases} \overline{x}^{uW} \overset{\text{def}}{\equiv} x_L^{uW}(\alpha) = x_H^{uW}(\alpha) = x_L^{wW}(0) = x_H^{wW}(0), \\ x_L^{wW}(\alpha) = x_H^{wW}(\alpha), \end{cases}$$

Activity assignment: AAR^W.
(c) Boadway et al. planner:
Consumption bundles: $x^B(\alpha) \overset{\text{def}}{\equiv} x_L^{wB}(\alpha) = x_L^{uB}(\alpha) = x_H^{wB}(\alpha) = x_H^{uB}(\alpha)$,

$$\textit{Activity assignment:}\begin{cases} \text{Case 1: } \frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)} > -1 : \text{AAR}^B. \\ \text{Case 2: } \frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)} = -1 \text{ (i.e. } W(\alpha)\alpha \text{ is constant):} \\ \quad \lambda^B = \int_0^\infty W(\alpha)\, v'\left(x^B(\alpha)\right) dF(\alpha), \\ \quad w_H\lambda^B > w_L\lambda^B > W(\alpha)\alpha \Rightarrow n_H^B = n_L^B = 1. \\ \quad w_H\lambda^B > w_L\lambda^B = W(\alpha)\alpha \Rightarrow n_H^B = 1, 0 < n_L^B < 1 \\ \quad w_H\lambda^B > W(\alpha)\alpha > w_L\lambda^B \Rightarrow n_H^B = 1, n_L^B = 0. \\ \quad w_H\lambda^B = W(\alpha)\alpha > w_L\lambda^B \Rightarrow 0 < n_H^B < 1, n_L^B = 0. \\ \quad W(\alpha)\alpha > w_H\lambda^B > w_L\lambda^B \Rightarrow n_H^B = n_L^B = 0, \\ \text{Case 3: } \frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)} < -1 : \\ \delta_L(\alpha) = 1 \text{ for all } \alpha \geq \alpha_L^{B**},\ \delta_H(\alpha) = 1 \text{ for all } \alpha \geq \alpha_H^{B**} \text{ and } \alpha_L^{B**} > \alpha_H^{B**}. \end{cases}$$

(d) Non-Welfarist planner:
Consumption bundles: $\overline{x}^N \overset{\text{def}}{\equiv} x_L^{wN}(\alpha) = x_L^{uN}(\alpha) = x_H^{wN}(\alpha) = x_H^{uN}(\alpha)$

$$\textit{Activity assignment:}\begin{cases} \lambda^N = v'\left(\overline{x}^N\right) \\ w_H\lambda^N > w_L\lambda^N > \overline{\alpha} \Rightarrow n_H^N = n_L^N = 1. \\ w_H\lambda^N > w_L\lambda^N = \overline{\alpha} \Rightarrow n_H^N = 1, 0 < n_L^N < 1. \\ w_H\lambda^N > \overline{\alpha} > w_L\lambda^N \Rightarrow n_H^N = 1, n_L^N = 0. \\ w_H\lambda^N = \overline{\alpha} > w_L\lambda^N \Rightarrow 0 < n_H^N < 1, n_L^N = 0. \\ \overline{\alpha} > w_H\lambda^N > w_L\lambda^N \Rightarrow n_H^N = n_L^N = 0. \end{cases}$$

(e) Roemer planner:

$$\textit{Consumption bundles:}\begin{cases} \forall\alpha \in [0, \alpha_L^{*R}) \cup [\alpha_H^{*R}, \infty) : \overline{x}^R \overset{\text{def}}{\equiv} x_L^{wR}(\alpha) = x_H^{wR}(\alpha) = x_L^{uR}(\alpha) = x_H^{uR}(\alpha), \\ \forall\alpha \in [\alpha_L^{*R}, \alpha_H^{*R}) : x_L^{uR}(\alpha) = v^{-1}\left(v\left(x_H^{wR}(\alpha)\right) - \alpha\right) < \overline{x}^R. \end{cases}$$

[8] All proofs are given in Appendix A.

Activity assignment: AAR^R.
(f) Van de gaer planner:
Consumption bundles: $\overline{x}^V \stackrel{\text{def}}{\equiv} x_L^{wV}(\alpha) = x_L^{uV}(\alpha) < \overline{\overline{x}}^V \stackrel{\text{def}}{\equiv} x_H^{wV}(\alpha) = x_H^{uV}(\alpha)$.
Activity assignment: AAR^V.

(a) Utilitarian planner
A Utilitarian planner gives the same consumption $\overline{x}^U$ to everyone, irrespective of his skill level and his taste parameter. Workers are clearly worse off than inactive people; the worst off will be the high-skilled workers with taste parameter α_H^*.
(b) Welfarist planner
The main difference between a Welfarist and a Utilitarian planner is that a Welfarist planner will give different consumption bundles to workers, depending on their disutility of labor. More precisely, it can be shown that $\partial x_Y^{wW}(\alpha)/\partial\alpha > 0$ with $Y = L, H$[9]. A Welfarist planner tries to compensate workers with a higher disutility of labor by giving them additional consumption, but the compensation is insufficient to make utility independent of labor disutility: $\partial\left(v\left(x_y^{wW}(\alpha)\right) - \alpha\right)/\partial\alpha < 0$. As a result, the high-skilled worker with taste α_H^* remains the worst off, as under the Utilitarian criterion. Moreover, consumption of workers is equalized at each α level, i.e. $x_L^{wW}(\alpha) = x_H^{wW}(\alpha)$.
(c) Boadway *et al.* planner
The Boadway *et al.* planner's consumption function depends on tastes only. If the weight given to individuals with a higher disutility of labor α increases (decreases), those with a higher (lower) α get more consumption, i.e. $\partial x^B(\alpha)/\partial\alpha \geq (\leq) 0$ if $W'(.) \geq (\leq) 0$. Activity assignment can take many forms, depending on the elasticity of the social welfare function with respect to the taste parameter, $(\partial W(\alpha)/\partial\alpha)(\alpha/W(\alpha))$. If this elasticity is larger than -1 (as in the Utilitarian case where $W(\alpha)$ is a constant and so the elasticity is zero), the usual assignment to activities, AAR, occurs as in the Utilitarian and Welfarist cases. However, if this elasticity is smaller than -1 (which requires that $W(\alpha)$ is sufficiently declining in α), the Boadway *et al.* planner wants to keep those with a high disutility of labor in work. If the elasticity is exactly -1, corner solutions prevail in which at least everyone in one skill group works or is inactive. If there exists a group for which no corner solution occurs, the planner is indifferent to who (i.e. which value for the taste parameter) is assigned to work. Which case occurs if $W(\alpha)\alpha$ is constant depends crucially on the level of this constant.
(d) Non-Welfarist planner
The Non-Welfarist consumption function has the same features as the Utilitarian one: everyone receives the same consumption, $\overline{x}^N$, irrespective of his skill and his taste parameter. The activity assignment crucially depends on the level of $\overline{\alpha}$. Moreover, the Non-Welfarist and Boadway *et al.* criterion, with elasticity of $W(\alpha)$ equal to -1, both lead to similar activity assignment, with the reference $\overline{\alpha}$ playing the role of the constant $W(\alpha)\alpha$.
(e) Roemer planner
Roemer planner's consumption function depends on tastes only. The Roemer planner satisfies AAR. However, w_H-workers and w_L-inactive having the same $\alpha \in [\alpha_L^{*R}, \alpha_H^{*R})$ receive the same utility level since $x_L^{uR}(\alpha) = v^{-1}\left(v\left(x_H^{wR}(\alpha)\right) - \alpha\right)\ \forall\alpha \in [\alpha_L^{*R}, \alpha_H^{*R})$.

[9] For this and the other formal properties stated in this discussion, see appendix A.

(f) Van de gaer planner

The main difference between Roemer and Van de gaer planners is that the latter gives different consumption bundles to people with identical α and the same activity choice, when their skills differ. In particular, high-skilled people receive a larger consumption level (denoted by $\overline{\overline{x}}^V$) than low-skilled people (denoted by $\overline{x}^V$), $\overline{\overline{x}}^V \geq \overline{x}^V$.

Among all allocations listed in theorem 1a, only the one derived under Roemer's criterion satisfies EWEP. With all criteria, there exist values for α for which high-skilled, contrary to low-skilled, have to work. By definition, EWEP then requires $v\left(x_L^w\left(\alpha\right)\right) - \alpha = v\left(x_L^u\left(\alpha\right)\right) \Leftrightarrow x_L^u\left(\alpha\right) = v^{-1}\left(v\left(x_H^w(\alpha)\right) - \alpha\right) \forall \alpha \in [\alpha_L^{*X}, \alpha_H^{*X})$. This equality never occurs except under Roemer's criterion. Since, in addition, Roemer's allocation is also such that $\forall \alpha \in [0, \alpha_L^{*R})$ ($\forall \alpha \in [\alpha_H^{*R}, \infty)$), everyone works (is inactive) and receives the same consumption, Roemer's allocation satisfies EWEP.

For all above optimal allocations $\exists \alpha, \alpha' \in \mathbb{R}^+ : x_Y^w\left(\alpha\right) = x_Y^u\left(\alpha'\right)$ for $Y = L$ or H. Since $w_Y > 0$, this violates ETES.

We next turn to the FEO, CE and EE criteria described in the previous section. These criteria are directly inspired by the equality of opportunity axioms.

Theorem 1b. *With full information, the following configuration of policies is optimal:*
(a) FEO:
(i) $n_H = n_L = 1$ *and* $x_L^w = x_H^w = \gamma w_L + (1-\gamma) w_H - R$.
(ii) $n_H = n_L = 0$ *and* $x^u = -R$.
(b) CE: There are five types of optimal allocations possible:
(i) $n_H = n_L = 1$ *and* $x_L^w = x_H^w = \gamma w_L + (1-\gamma) w_H - R$.
(ii) $n_H = 1, 0 < n_L < 1, -x^u = [w_L - x_L^w]$ *and* $x_L^w = x_H^w = v^{-1}\left(v\left(x^u\right) - \widetilde{\alpha}\right)$.
(iii) $n_H = 1, n_L = 0$ *and* $x^u = \left[(1-\gamma)\left(w_H - x_H^w\right) - R\right]/\gamma$ *and* $x_H^w = v^{-1}\left(v\left(x^u\right) - \widetilde{\alpha}\right)$.
(iv) $0 < n_H < 1, n_L = 0, -x^u = [w_H - x_H^w]$ *and* $x_H^w = v^{-1}\left(v\left(x^u\right) - \widetilde{\alpha}\right)$.
(v) $n_H = n_L = 0$ *and* $x^u = -R$.
(c) EE:
$x_L^w = x_H^w = \gamma w_L + (1-\gamma) w_H - R$ *and* $x^u = 0$,
$\alpha_L^* = \alpha_H^* = v\left(\gamma w_L + (1-\gamma) w_H\right) - v(0)$.

(a) FEO allocation

By construction, the FEO allocations satisfy both EWEP and ETES, however they are quite trivial. FEO (i) assigns everyone to work while FEO (ii) implies that everyone is inactive. FEO (i) and (ii) give everyone the same consumption. Note that, contrary to FEO (i), FEO (ii) gives everyone the same utility. It is easy to verify that neither of the FEO allocations is Pareto efficient.

(b) CE allocation

With the CE criterion, the two FEO allocations can be optimal as well as three others. The latter are denoted by CE (ii), (iii) and (iv). The CE allocation equalizes $u\left(x_Y\left(\alpha\right), \delta_Y\left(\alpha\right), \widetilde{\alpha}\right)$ for all α and $Y = L, H$. Therefore, welfares are equalized when bundles are evaluated with references preferences, but not with actual preferences (see, (ii), (iii), (iv)). EWEP is thus not satisfied. We will now check the validity of ETES. In the CE allocation (ii), all high-skilled work and a fraction of the low-skilled agents work. All high-skilled people receive the same consumption bundle x_H^w and all low-skilled people receive the same transfer $-x^u = w_L - x_L^w$. This CE allocation thus satisfies

ETES. CE allocation (iii) has all high-skilled and no low-skilled working. ETES is then satisfied. The CE allocation (iv) has only a fraction of the high-skilled working. Again the planner does not care which high-skilled. Since no low-skilled work, and for the high-skilled $-x^u = [w_H - x_H^w]$, this allocation satisfies ETES too.

Which of these CE allocations is the optimal one depends on the parameters of the model. For $\widetilde{\alpha}$ sufficiently low, the optimum will be of type (i). As $\widetilde{\alpha}$ increases, we move over cases (ii), (iii) and (iv) to (v). The properties of the CE allocation clearly show that it is possible to find allocations that have attractive properties from the perspective of equality of opportunity in the first best.

Moreover, note the qualitative similarity of activity assignment with the CE objective in theorem 1b, the Boadway *et al.* objective in case 2 of theorem 1a and the Non-Welfarist objective in theorem 1a. The crucial difference between these allocations in theorem 1a and the CE allocation is the determination of the consumption bundles: the Welfarist planner gives the same consumption to everyone, the Boadway *et al.* planner in case 2 gives lower consumption to the individuals with less deserving tastes (i.e. with a higher α), while the CE planner determines the consumption bundles such that they satisfy the ETES axiom.

(c) EE allocation

Under the EE allocation, all workers receive the same consumption bundle, irrespective of their skill level. The inactive get zero benefits. This looks harsh at first sight, but in terms of equivalent wages, the metric used by the planner in this case, these individuals are best off, and, in the present framework people are responsible for their preference. Observe that this policy satisfies EWEP. All high-skilled pay the same tax, all low-skilled pay the same tax, and all inactive get the same zero transfer. The tax paid is not equal to minus the transfer received, however. Hence ETES is not satisfied.

The EE allocation assigns the same consumption bundle to workers as allocations FEO (i) and CE (i), but contrary to these allocations, those with high disutility of labor are not working. They are inactive, and are, actually better off (both in terms of utility and equivalent wages) than under allocations FEO (i) and CE (i).

We summarize the performance of the criteria in theorem 1a-b from the equality of opportunity principles in the following corollary.

Corollary 1: equality of opportunity axioms and social objectives in the first best.

<table>
<tr><th>Social Objective</th><th>Satisfies EWEP?</th><th>Satisfies ETES?</th></tr>
<tr><td>Utilitarian</td><td rowspan="5">No</td><td rowspan="5">No</td></tr>
<tr><td>Welfarist</td></tr>
<tr><td>Boadway et al.</td></tr>
<tr><td>Non-Welfarist</td></tr>
<tr><td>Van de gaer</td></tr>
<tr><td>Roemer</td><td>Yes</td><td>No</td></tr>
<tr><td>FEO</td><td>Yes</td><td>Yes</td></tr>
<tr><td>Egalitarian Equivalent</td><td>Yes</td><td>No</td></tr>
<tr><td>Conditional Egalitarian</td><td>No</td><td>Yes</td></tr>
</table>

Given the origin of these social orderings, it is unsurprising to see that those criteria which originate from the social choice approach to equality of opportunity perform much better than the traditional criteria. They were designed to do so. It is of course possible to search for the first best optimal allocation over the set of allocations satisfying one of the equality of opportunity principles, using the objective functions that do not satisfy the equality of opportunity principle under consideration. However, to keep the size of the paper within reasonable limits and since the second best context is more relevant, we only illustrate this procedure in the second best context [10].

Finally, all the first best solutions listed in Theorem 1a and 1b depend on both α and w_i. Therefore, they are not implementable when the government only observes income (second best). The next section deals with this issue.

5 Second best optima

5.1 Second best constraints and their implications

In second best, the Government needs to take into account the set of self-selection or incentive compatibility constraints (hereafter ICC) in order to prevent individuals from a given type from mimicking (i.e. taking the tax-treatment designed for) individuals of other types. We first state these IC constraints and then discuss their implications for the social objective functions.

Agents of w_L-type choose between $v(x^u)$ and $v(x_L) - \alpha$. Introducing the threshold value a_L^*, and dropping the superscripts U, W, B, NW, R, V, CE and EE for notational simplicity, the ICC[11] on w_L-agents can be written as:

$$v(x_L) - \alpha_L^* = v(x^u), \tag{10}$$

such that a low skilled with taste parameter α chooses low skilled employment instead of inactivity if and only if $\alpha < \alpha_L^*$.

Agents of w_H-type choose between $v(x^u)$, $v(x_L) - \alpha$ and $v(x_H) - \alpha$. Since all our objective functions are increasing in individuals' consumption, it will, just like in the first best, never be optimal that high-skilled people work in low-skilled jobs. By putting these people in high-skilled jobs instead of low-skilled jobs, they produce more which can be used to increase everyone's consumption in a way that respects the ICC and hence increases the social objective's value. Consequently, to induce high-skilled people to work in high-skilled jobs,

$$x_H \geq x_L, \tag{11}$$

and, introducing the threshold value α_H^*, the ICC on agents of w_H-type states

$$v(x_H) - \alpha_H^* = v(x^u), \tag{12}$$

such that a high skilled agent with taste parameter α prefers high-skilled employment to inactivity if and only if $\alpha < \alpha_H^*$. Moreover, from (10), (11) and (12), we have that

$$\alpha_H^* \geq \alpha_L^*. \tag{13}$$

[10] We only impose one principle at a time, as imposing both principles simultaneously leads to the FEO allocation.

[11] The set of IC constraints for each agent of type (w_Y, α) (with $Y := L, H$ and $\alpha \in \mathbb{R}^+$) can be rewritten as constraints (10)-(12). Moreover, since the labor supply decision is restricted to be binary, the (direct truthful) mechanism that implements the optimal allocations is not fully revealing. Each agent fully reveals his w_Y information but not his α value; he announces only whether α is larger or lower than α_Y^*.

As a result of the second best constraints (10), (12) and (13), irrespective of the social objective function, activity assignment has to be of type AAR^X. Moreover, because of (11), utility of w_H-workers is at least as high as of w_L-workers. Hence, the utilities as a function of α, for w_L- and w_H-skilled agents, look as in the following Figure.

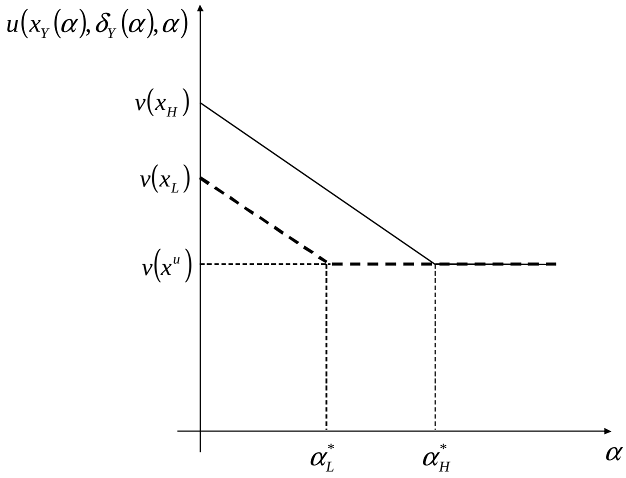

Figure 1: utilities in the second best.

The full line is the utility of a w_H-individual. He works if his disutility of work $\alpha \leq \alpha_H^*$, and he is inactive otherwise. Similarly the bold dotted line is the utility of a w_L-individual. The latter works for $\alpha \leq \alpha_L^*$ and is inactive otherwise. Different planners choose different values for $(x^u, x_L, x_H, \alpha_L^*, \alpha_H^*)$, but the qualitative shape of the utilities as a function of α, for high- and low-skilled individuals, is always as indicated in the graph.

The second best framework has important implications for the equality of opportunity principles, as stated in the following lemma.

Lemma 1. Equality of opportunity principles in the second best.
(a) A necessary and sufficient condition to fully satisfy EWEP is that $\alpha_L^* = \alpha_H^*$, which requires that $x_L = x_H$.
(b) A necessary and sufficient condition to fully satisfy ETES is that $x_L - w_L = x^u = x_H - w_H$.

Part (a) says that the threshold values α_L^* and α_H^* have to be the same. To accomplish this, the government has to offer the same consumption level to high and low-skilled workers. It implies that the same number of high and low-skilled individuals will work. Part (b) of the corollary follows immediately from application of the ETES axiom and has two noteworthy implications. First, since $x_L - w_L = x^u$ and $x_H - w_H = x^u$, the government cannot subsidize or tax the participation decision. Since it cannot do this at the bottom end of the skill distribution, there is neither a negative income tax nor an earned income tax credit. Second, since $x_L - w_L = x_H - w_H$, the government cannot redistribute between low and high-skilled workers. This is a very severe restriction, which makes the ETES axiom difficult to defend in the second best context.

As a result of the second best constraints, the second best optimal tax problem in its general form reduces to the following maximization problem.

GSBP (General Second Best Problem):

$$\max_{x_L, x_H, x^u, \alpha_L^*, \alpha_H^*} \widetilde{S}^X \left(x_L, x_H, x^u, \alpha_L^*, \alpha_H^*\right),$$

subject to the government budget constraint,

$$\gamma \left[\left(w_L - x_L\right) F\left(\alpha_L^*\right) - x^u \left(1 - F\left(\alpha_L^*\right)\right)\right]$$
$$+ \left(1 - \gamma\right) \left[\left(w_H - x_H\right) F\left(\alpha_H^*\right) - x^u \left(1 - F\left(\alpha_H^*\right)\right)\right] - R = 0,$$

and constraints (10), (11) and (12).

The second best framework has important consequences for the specification of the social objective functions. Combining the expressions for the social objective functions (2), (3), (4), (5), (6), (7), (8), (9) with expression (10), (11), (12) and (13) results in the following writing of the objective functions, as shown in Appendix B. Again, we skip the superscripts U, W, B, NW, R, V, CE and EE for notation simplicity.

(a) Utilitarian

$$\widetilde{S}^U = \gamma \int_0^{\alpha_L^*} \left[v(x_L) - \alpha\right] dF(\alpha) + \gamma \int_{\alpha_L^*}^{\infty} v(x^u) dF(\alpha)$$
$$+ (1-\gamma) \int_0^{\alpha_H^*} \left[v(x_H) - \alpha\right] dF(\alpha) + (1-\gamma) \int_{\alpha_H^*}^{\infty} v(x^u) dF(\alpha).$$

(b) Welfarist

$$\widetilde{S}^W = \gamma \int_0^{\alpha_L^*} \Psi\left(v(x_L) - \alpha\right) dF(\alpha) + \gamma \int_{\alpha_L^*}^{\infty} \Psi\left(v(x^u)\right) dF(\alpha)$$
$$+ (1-\gamma) \int_0^{\alpha_H^*} \Psi\left(v(x_H) - \alpha\right) dF(\alpha) + (1-\gamma) \int_{\alpha_H^*}^{\infty} \Psi\left(v(x^u)\right) dF(\alpha).$$

(c) Boadway *et al.*

$$\widetilde{S}^B = \gamma \int_0^{\alpha_L^*} W\left(\alpha\right) \left[v(x_L) - \alpha\right] dF(\alpha) + \gamma \int_{\alpha_L^*}^{\infty} W\left(\alpha\right) v(x^u) dF(\alpha)$$
$$+ (1-\gamma) \int_0^{\alpha_H^*} W\left(\alpha\right) \left[v(x_H) - \alpha\right] dF(\alpha) + (1-\gamma) \int_{\alpha_H^*}^{\infty} W\left(\alpha\right) v(x^u) dF(\alpha).$$

(d) Non-Welfarist

$$\widetilde{S}^{NW} = \gamma \left[\int_0^{\alpha_L^*} \left[v(x_L) - \overline{\alpha}\right] dF(\alpha)\right] + \gamma \int_{\alpha_L^*}^{\infty} v(x^u) dF(\alpha)$$
$$+ (1-\gamma) \int_0^{\alpha_H^*} \left[v(x_H) - \overline{\alpha}\right] dF(\alpha) + (1-\gamma) \int_{\alpha_H^*}^{\infty} v(x^u) dF(\alpha).$$

(e) Roemer and (f) Van de gaer

$$\widetilde{S}^R = \int_0^{\alpha_L^*} \left(v\left(x_L\right) - \alpha\right) dF\left(\alpha\right) + \int_{\alpha_L^*}^{\infty} v\left(x^u\right) dF\left(\alpha\right).$$

(g) Conditional Equality

$$\widetilde{S}^{CE} = v\left(x_L\right) - \widetilde{\alpha} \text{ subject to } \widetilde{\alpha} \geq \alpha_L^*.$$

(h) Egalitarian Equivalent

$$\widetilde{S}^{EE} = x_L.$$

Under asymmetric information, Roemer and Van de gaer's criterion are equal. Due to the second best constraint, utility as a function of the taste parameter of the low-skilled will never be below utility as a function of the taste parameter for the high-skilled. One implication of this is that the opportunity set for the lowly skilled is below the one for the highly skilled, hence, in the second best the mean of mins and min of means criterion will yield the same solution.

5.2 Optimal tax formula

Before we can characterize the optimal tax rates, we need to introduce more definitions. Let $T_L = w_L - x_L$, $T_H = w_H - x_H$, and $T_u = -x^u$, be the tax paid by the low-skilled workers, the high-skilled workers and the inactive, respectively. Define the elasticities of participation of the low-skilled with respect to x_L[12] and of the high-skilled with respect to x_H as

$$\eta\left(x_L, \alpha_L^*\right) \overset{\text{def}}{\equiv} \frac{x_L}{F\left(\alpha_L^*\right)} f\left(\alpha_L^*\right) v'\left(x_L\right), \tag{14}$$

$$\eta\left(x_H, \alpha_H^*\right) \overset{\text{def}}{\equiv} \frac{x_H}{F\left(\alpha_H^*\right)} f\left(\alpha_H^*\right) v'\left(x_H\right), \tag{15}$$

respectively. Next, observe that the average of the inverse of the private marginal utility of consumption, is given by

$$g_P^X \overset{\text{def}}{\equiv} \frac{\gamma F(\alpha_L^{X*})}{v'(x_L^X)} + \frac{\gamma(1 - F(\alpha_L^{X*})) + (1-\gamma)(1 - F(\alpha_H^{X*}))}{v'(x^{uX})} + \frac{(1-\gamma)F(\alpha_H^{X*})}{v'(x_H^X)}. \tag{16}$$

Let subscripts to the function S denote the partial derivative of S with respect to the argument in the subscript and note that the effect of a uniform increase in private utilities on the social objective function is given by

$$D^X = \frac{\widetilde{S}_{x_L}^X}{v'(x_L)} + \frac{\widetilde{S}_{x_H}^X}{v'(x_H)} + \frac{\widetilde{S}_{x^u}^X}{v'\left(x^u\right)}. \tag{17}$$

Finally, the average social marginal utility of consumption for workers of skill level Y ($Y = L$ or H) is

$$g_L^X = \frac{S_{x_L}^X}{\lambda \gamma F\left(\alpha_L^*\right)} \text{ and } g_H^X = \frac{S_{x_H}^X}{\lambda\left(1-\gamma\right) F\left(\alpha_H^*\right)}.$$

The following theorem states the solution for the general second best problem.

Theorem 2: *Under asymmetric information, the optimal consumption levels have to satisfy the budget constraint, constraints (10), (11) and (12) and the following equations:*

$$\begin{aligned}
\frac{(T_L - T_u)}{x_L} &= \frac{1}{\eta\left(x_L, \alpha_L^*\right)}\left[1 - g_L^X + \frac{\nu}{\lambda\gamma F\left(\alpha_L^*\right)}\right] - \frac{\widetilde{S}_{\alpha_L^*}^X}{\lambda\gamma f\left(\alpha_L^*\right) x_L}, \\
\frac{(T_H - T_u)}{x_H} &= \frac{1}{\eta\left(x_H, \alpha_H^*\right)}\left[1 - g_H^X - \frac{\nu}{\lambda\left(1-\gamma\right) F\left(\alpha_H^*\right)}\right] - \frac{\widetilde{S}_{\alpha_H^*}^X}{\lambda\left(1-\gamma\right) f\left(\alpha_H^*\right) x_H}, \\
\left(\lambda^X\right)^{-1} &= g_P^X / D^X,
\end{aligned}$$

[12] $\eta\left(x_L, \alpha_L^*\right) \overset{\text{def}}{\equiv} \left(x_L / \gamma F\left(\alpha_L^*\right)\right)\left(\partial\left(\gamma F\left(\alpha_L^*\right)\right) / \partial x_L\right)$. Since $\alpha_L^* = v\left(x_L\right) - v\left(x^u\right)$, we get $\partial \alpha_L^* / \partial x_L = v'\left(x_L\right)$ hence we obtain (14).

where ν is the Lagrangian multiplier associated with the constraint $x_H \geq x_L$.

The λ^{-1} equations are similar to Diamond and Sheshinsky (1995)'s equation (6), p.6. and are associated with an equal marginal change of the consumption of everyone in the economy. Consider a uniform increase in all private utilities of one unit. This does not change the activity decisions. To accomplish this uniform increase, we need per low-skilled worker $1/v'(x_L)$ extra units of consumption, per high-skilled worker we need $1/v'(x_H)$ extra units of consumption and per inactive person we need $1/v'(x^u)$ extra units of consumption. Weighting this by the frequencies of these groups in the population, we find that we need an additional $g_P(x^u, x_L, x_H, \alpha_L^*, \alpha_H^*)$ units of public means to finance this operation. In terms of social welfare, this is worth $\lambda g_P(x^u, x_L, x_H, \alpha_L^*, \alpha_H^*)$. This has to be equal to the increase in the social objective function caused by the uniform increase in utilities, which is equal to D. The equation for λ^{-1} thus equates the inverse of the marginal cost of public funds to the ratio between the average of the inverse of the private utilities and the marginal social utility of a uniform increase in all individual utilities.

Next, we give a simple heuristic interpretation of the optimal tax formulas in the spirit of Saez (2002). Consider a small increase of the consumption x_L (i.e. a small reduction of the income tax in low-skilled jobs), around the optimal tax schedule. This has a mechanical effect and a behavioral (or labor supply response) effect.

Mechanical effect

There is a mechanical decrease in tax revenue equal to $-\gamma F(\alpha_L^*)dx_L$ because low-skilled workers have dx_L additional consumption. Each unit of x_L is worth $\left(\widetilde{S}_{x_L}^X - v\right)/\lambda$ in terms of government revenue. Hence the total value of the decrease in tax revenue is worth $-\left(\gamma F(\alpha_L^*) - \left(\widetilde{S}_{x_L}^X - v\right)/\lambda\right) dx_L$ in terms of government revenue, which can be written as

$$-\left[1 - \frac{\widetilde{S}_{x_L}^X - v}{\lambda\gamma F(\alpha_L^*)}\right]\gamma F(\alpha_L^*)dx_L.$$

Behavioral effect

The change $dx_L > 0$ induces a change in α_L^* equal to $(\partial\alpha_L^*/\partial x_L)\, dx_L$. By (10), $\partial\alpha_L^*/\partial x_L = v'(x_L)$ and from the definition of the elasticity of participation (14), $v'(x_L) = [F(\alpha_L^*)\,\eta(x_L, \alpha_L^*)] / [x_L f(\alpha_L^*)]$, such that the induced change in α_L^* is $[F(\alpha_L^*)\,\eta(x_L, \alpha_L^*)] / [x_L f(\alpha_L^*)]\, dx_L$. A change in the critical value α_L^* has a welfare effect, worth $\widetilde{S}_{\alpha_L^*}^X/\lambda$ in terms of government revenue and increases government revenue by $\gamma[T_L - T_u] f(\alpha_L^*)$. Hence the total behavioral effect in terms of government revenue equals

$$\left[\frac{\widetilde{S}_{\alpha_L^*}^X}{\lambda} + \gamma(T_L - T_u) f(\alpha_L^*)\right]\frac{F(\alpha_L^*)\,\eta(x_L, \alpha_L^*)}{x_L f(\alpha_L^*)} dx_L$$

At the optimum, the sum of the mechanical and behavioral effects has to be nil. It is easy to verify that this yields the first equation in theorem 2. The second equation can be given a similar interpretation.

Observe that the optimal tax formula in the theorem contain three elements: the deviation of the average social marginal utility of consumption for workers of a particular skill level from unity,

$1-g_Y^X$, the Lagrangian multiplier ν and the term $\widetilde{S}^X_{\alpha_Y^*}$ ($Y = L$ or H). The last two terms have not been dealt with in the literature on optimal taxation with extensive labor supply, as they do not appear with the social objective functions U and W that have been considered so far. This is stated in the following lemmas.

Lemma 2: the value of $\widetilde{S}^X_{\alpha_Y^*}$ $(Y = L, H)$:
(a) $\widetilde{S}^X_{\alpha_Y^*} = 0$ for $X = U$, W, B, R, EE, CE.
(b) $\widetilde{S}^{NW}_{\alpha_L^*} = [\alpha_L^* - \overline{\alpha}]\, \gamma f(\alpha_L^*)$ and $S^{NW}_{\alpha_H^*} = [\alpha_H^* - \overline{\alpha}]\,(1-\gamma)\, f(\alpha_H^*)$.

Lemma 3: the value of the Lagrangian multiplier:
(a) $\nu = 0$ for $X = U$, W, B, and NW.
(b) $\nu \geq 0$ for $X = R$, EE and CE.

Lemma 2 follows from partially differentiating the expressions for the social objective functions with respect to α_Y^* $(Y = L, H)$. These terms represent direct effects of changes in the critical values on the social objective functions and occur only in the Non-Welfarist case. Lemma 3 is slightly less straightforward. Remember that ν is equal to the sum of the welfare and budget effect of decreasing x_H and, at the same time it is equal to the sum of the welfare and budget effect of increasing x_L. Hence these two effects have to be equal. Moreover, if $\nu > 0$, the sum of the welfare effects of decreasing x_H and increasing x_L and their government revenue effect have to be equal. If, evaluated at $x_H = x_L$ and $\alpha_H^* = \alpha_L^*$, the per capita welfare effects $\left(\widetilde{S}^X_{x_L} + \widetilde{S}^X_{\alpha_L^*}\right)/\gamma$ and $\left(\widetilde{S}^X_{x_H} + \widetilde{S}^X_{\alpha_H^*}\right)/(1-\gamma)$ are equal (as is the case for $X = U$, W, B and NW), then the per capita budget effect of decreasing x_H must be equal to the per capita budget effect of increasing x_L for $x_H = x_L$ and $\alpha_H^* = \alpha_L^*$. These per capita budget effects are $F(x_Y^*) - (w_Y - x_Y + x^u)\, f(\alpha_Y^*)\, v'(x_Y)$ for $Y = L$ or H. However, with $x_H = x_L$ and $\alpha_H^* = \alpha_L^*$ these cannot be equal, as $w_H > w_L$. The objective functions R, EE and CE have unequal per capita welfare effects, so an optimum with $\nu > 0$ does not require equal per capita budget effects; it is possible to obtain an optimum where $\nu > 0$.

Lemma 3 combined with lemma 1 has implications for the performance of the different social objective functions from the perspective of the equality of opportunity principles. Since the U, W, B, and NW criteria have a zero value for ν, their solution will have $x_H > x_L$ (as shown in lemma C of Appendix C), and so $\alpha_H^* > \alpha_L^*$, such that their solution violates EWEP. However, with the R, EE and CE criteria, ν may be strictly positive, in which case $x_H = x_L$ and $\alpha_H^* = \alpha_L^*$ such that EWEP is satisfied.

In order to obtain optimal tax rates with the different social objective functions, we use the relevant properties of these social objective functions and plug them in the equations of the theorem. Lemma 2 gives us the values for $\widetilde{S}^X_{\alpha_Y^*}$ $(Y = L, H)$, and lemma 3 the values for the Lagrangian multipliers. The average social marginal utility of consumption g_Y^X under objective functions X $(= U, W, B, NW, R, V, CE$ or $EE)$ for agents of skill level Y (=L or H) are given in the following table.

X	g_L^X	g_H^X
U, NW	$\frac{v'\left(x_L^X\right)}{\lambda^X}$	$\frac{v'\left(x_H^X\right)}{\lambda^X}$
W	$\frac{v'\left(x_L^W\right)}{\lambda^W} \frac{\int_0^{\alpha_L^{W*}} \Psi'\left(v\left(x_L^W\right)-\alpha\right)dF(\alpha)}{F(\alpha_L^{W*})}$	$\frac{v'\left(x_H^W\right)}{\lambda^W} \frac{\int_0^{\alpha_H^{W*}} \Psi'\left(v\left(x_H^W\right)-\alpha\right)dF(\alpha)}{F(\alpha_H^{W*})}$
B	$\frac{v'\left(x_L^B\right)}{\lambda^B} \frac{\int_0^{\alpha_L^{B*}} W(\alpha)dF(\alpha)}{F(\alpha_L^{B*})}$	$\frac{v'\left(x_H^B\right)}{\lambda^B} \frac{\int_0^{\alpha_H^{B*}} W(\alpha)dF(\alpha)}{F(\alpha_H^{B*})}$
$R(=V)$	$\frac{v'\left(x_L^R\right)}{\lambda^R \gamma}$	0
EE	$\frac{1}{\lambda^{EE}\gamma F(\alpha_L^{EE*})}$	0
CE	$\frac{v'\left(x_L^{CE}\right)}{\lambda^{CE}\gamma F(\alpha_L^{CE*})}$	0

Using these expression in the equations of theorem 2, together with $\nu = 0$ for $X = U$, W, B and NW results in the following corollary.[13]

Corollary 2. Under asymmetric information, the optimal consumption levels have to satisfy the budget constraint, constraints (10), (11) and (12) and the following equations:

X	$\left(\lambda^X\right)^{-1}$	$\frac{T_H^X - T_u^X}{x_H^X}$	$\frac{T_L^X - T_u^X}{x_L^X}$
U	g_P^X	$\frac{1}{\eta\left(x_H^X, \alpha_H^{X*}\right)}\left(1 - g_H^X\right)$	$\frac{1}{\eta\left(x_L^X, \alpha_L^{X*}\right)}\left(1 - g_L^X\right)$
W	g_P^X / D^X		
B	g_P^X / D^X		
$R = V$	g_P^X	$\frac{1}{\eta\left(x_H^X, \alpha_H^{X*}\right)}\left(1 - \frac{\nu^X}{\lambda^X(1-\gamma)F\left(\alpha_H^{X*}\right)}\right)$	$\frac{1}{\eta\left(x_L^X, \alpha_L^{X*}\right)}\left(1 - g_L^X + \frac{\nu^X}{\lambda^X \gamma F\left(\alpha_L^{X*}\right)}\right)$
EE	g_P^X / D^X		
CE	g_P^X		$\frac{1}{\eta\left(x_L^X, \alpha_L^{X*}\right)}\left(1 - (1-\xi)\, g_L^X + \frac{\nu^X}{\lambda^X \gamma F\left(\alpha_L^{X*}\right)}\right)$
NW	g_P^X	$\frac{1}{\eta\left(x_H^X, \alpha_H^{X*}\right)}\left(1 - g_H^X\right) - \frac{\alpha_H^{X*} - \overline{\alpha}}{\lambda^X x_L^X}$	$\frac{1}{\eta\left(x_L^X, \alpha_L^{X*}\right)}\left(1 - g_L^X\right) - \frac{\alpha_L^{X*} - \overline{\alpha}}{\lambda^X x_L^X}$

with

$$
\begin{aligned}
D^W = {} & \gamma \left[\int_0^{\alpha_L^{W*}} \Psi'\left(v\left(x_L^W\right) - \alpha\right) dF(\alpha) + \int_{\alpha_L^{W*}}^{\infty} \Psi'\left(v\left(x^{uW}\right)\right) dF(\alpha) \right] \\
& + (1-\gamma) \left[\int_0^{\alpha_H^{W*}} \Psi'\left(v\left(x_H^W\right) - \alpha\right) dF(\alpha) + \int_{\alpha_H^{W*}}^{\infty} \Psi'\left(v\left(x^{uW}\right)\right) dF(\alpha) \right], \\
D^B = {} & \int_0^{\infty} W(\alpha)\, dF(\alpha) \qquad \text{and} \qquad D^{EE} = 1/v'\left(x_L^{EE}\right).
\end{aligned}
$$

Taking as a benchmark case the formula for the popular objective functions U and W, we see three sources of adjustment. Let us focus on the formula for the lowly skilled. First, with the Non-Welfarist criterion, the extra term $\left[\alpha_L - \overline{\alpha}\right] / \left[\lambda x_L\right]$ appears. It captures the social value of the divergence between private and social preferences. Second in the Conditional Equality planner's optimal policy, the multiplier associated with the constraint $\widetilde{\alpha} \geq \alpha_L^*$ enters in $\left(T_L^{CE} - T_u^{CE}\right)/x_L^{CE}$. If the constraint is binding, the planner needs to bring α_L^* down, for which it has to decrease x_L^{CE} or increase x^{uCE}. The former increases T_L^{CE}, the latter decreases T_u^{CE}, and so $\left(T_L^{CE} - T_u^{CE}\right)/x_L^{CE}$ must increase. This explains why an increase in ξ increases the right hand side of the equation in the table for $\left(T_L^{CE} - T_u^{CE}\right)/x_L^{CE}$. Third, with the social policies inspired by equality of opportunity

[13] The optimal activity assignments are characterized by $\alpha_H^* > \alpha_L^* > 0$ under the U, W, B and NW criteria, while $\alpha_H^* \geq \alpha_L^*$ under the R, EE and CE criteria. Moreover, $\alpha_H^* < \infty$ for all criteria. Appendix C states the proofs.

principles ($X = R$, CE or EE), it is possible that the constraint $x_H \geq x_L$ is binding and $\nu > 0$ hence this term enters the optimal tax formula. In this case, the planner would like to decrease x_L, which requires an increase in T_L, such that the multiplier enters positively in the right hand side of the equation in the table for $\left(T_L^X - T_u^X\right)/x_L^X$.

Since Diamond (1980), it is well known that subsidizing the low-skilled workers more than inactive people (i.e. $T_L < T_u$) can be optimal when the labor supply is modeled along the extensive margin. Using the definition of Saez (2002), an Earned Income Tax Credit (EITC) is then optimal. On the contrary, when $T_L > T_u$ a Negative Income Tax (NIT) is optimal. Alternatively, since $T_L < (>) T_u$ can be rewritten as $w_L < (>) x_L - x^u$, i.e. the income gain when a low-skilled agent enters the labor force ($x_L - x^u$) is larger (lower) than the gross labor income (w_L). In other words, the labor supply of the low-skilled is distorted upwards (downwards), compared to laissez faire. Theorem 2 can be used to study the necessary conditions for an EITC or a NIT under other criteria than the standard Utilitarian and Welfarist ones. Corollary 3 emphasizes that the Roemer, EE, CE and Non-Welfarist criteria challenge the standard necessary conditions.

Corollary 3. Optimality of EITC or NIT in second best.

Social Objective	ETES/ NIT/ EITC?
Utilitarian, Welfarist, Boadway *et al.*	NIT (EITC) if $g_L^X < (>)\, 1$
Roemer (=V), Egalitarian Equivalent	NIT (EITC) if $g_L^X - \frac{\nu^X}{\lambda \gamma F\left(\alpha_L^{X*}\right)} < (>)\, 1$
Conditional Egalitarian	NIT (EITC) if $(1-\xi)\, g_L^{CE} - \frac{\nu^{CE}}{\lambda^{CE} \gamma F\left(\alpha_L^{CE*}\right)} < (>)\, 1$
Non-Welfarist	NIT (EITC) if $\frac{1}{\eta\left(x_L^{NW}, \alpha_L^{NW*}\right)} \left(1 - g_L^{NW}\right) > (<) \frac{\alpha_L^{NW*} - \overline{\alpha}}{\lambda^{NW} x_L^{NW}}$
Maximin	NIT

Under the Utilitarian, the Welfarist and the Boadway *et al.* objectives, we retrieve the result that the average social weight of the low-skilled workers larger than one is a necessary condition for the EITC to be optimal. For the Roemer, Egalitarian Equivalent and Conditional Egalitarian objective functions, this condition has to be adjusted since the constraint that $x_H \geq x_L$ may be binding. If this constraint is binding, a NIT can be optimal even when g_L^X is larger than one. In that sense, these social objective functions that find their inspiration in equality of opportunity theories are more in favor of a NIT.

The necessary condition to obtain unambiguous results under the Non-Welfarist criterion is clearly more complicated: there is no simple relationship between the average social weight of the low-skilled workers being larger than one and the optimality of an EITC. The EITC (NIT) encourages (discourages) participation of the marginal worker, which results in an increased (decreased) utility of consumption equal to α_L^{N*}, which is desirable if this is larger (smaller) than $\overline{\alpha}$, the utility cost of work in the eyes of the Non-Welfarist planner. The extra term $\left(\alpha_L^{N*} - \overline{\alpha}\right) / \left(\lambda^N x_L^N\right)$ which appears at the right hand side in corollary 3 is used as a device to correct undesirable social outcomes. It corrects individual labor supply to correspond to social preferences[14]. Hence, if social

[14]Put differently, the other planners (including the Conditional Egalitarian when the constraint $\widetilde{\alpha} > \alpha_L^{C*}$ is not

preferences are characterized by $\alpha_L^{N*} > (<) \overline{\alpha}$, the government encourages (discourages) participation, the right hand side of the inequality in the corollary is positive, such that the EITC (NIT) then becomes more attractive for the Non-Welfarist planner. This term is sometimes called the paternalistic or first best motive for taxation since it arises from differences between social and private preferences (Kanbur *et al.,* 2006). Assuming $\alpha_L^{N*} > \overline{\alpha}$, when the Non-Welfarist government's views on working becomes more "Calvinistic", i.e. when $\overline{\alpha}$ decreases, the term at the right hand side becomes larger and hence plays in favor of an EITC to promote participation of more people.

As a final point of reference, we compare our policy prescriptions with the policy prescription of the Maximin social objective function. Maximin, which is a subcase of the Welfarist criterion plays in favor of a NIT, as shown in Choné and Laroque (2005). Under Maximin, only the least-well off receive a positive average social marginal utility of consumption. Due to the IC constraints, the least-well off are the inactive hence $g_L^W = g_H^W = 0$, and it can be shown that $\nu = 0$.[15] Substituting the latter into the optimal Welfarist tax formulas of theorem 2 yields

$$\left(T_L^M - T_u^M\right)/x_L^M = \left(1/\eta\left(x_L^M, \alpha_L^{U*}\right)\right) \quad \text{and} \quad \left(T_H^M - T_u^M\right)/x_H^M = \left(1/\eta\left(x_H^M, \alpha_H^{U*}\right)\right),$$

where M has been used as superscript for Maximin. Therefore $T_L^M - T_u^M > 0$ hence a NIT is always optimal under Maximin.

Empirical studies suggest that participation decisions are more elastic at the bottom of the skill distribution (see the empirical evidence surveyed by Immervoll *et al.*, 2007, and Meghir and Phillips, 2008) which motivates the following assumption:

Assumption 1: $\eta\left(x_L, \alpha_L^*\right) \geq \eta\left(x_H, \alpha_H^*\right)$.

Corollary 4. Under assumption 1, for the Utilitarian, Welfarist and Boadway *et al.* when $W(\alpha)$ is a decreasing function and for the Roemer, EE, CE (when $\xi < 1$):[16]

$$\left(T_L - T_u\right)/x_L < \left(T_H - T_u\right)/x_H.$$

Our model is an extensive model of labor supply. We have that the degree to which labor supply is distorted downwards depends on the difference between taxes paid when working and taxes paid when inactive (the latter is $-x^u$). The larger is this difference, the more labor supply is distorted downwards; if the difference is negative, labor supply is distorted upwards. We now have the following corollary.

Corollary 5. Under assumption 1, the Utilitarian, Welfarist, and Boadway *et al.* criteria when $W(\alpha)$ is a decreasing function and the Roemer, EE and CE (when $\xi < 1$) criteria, the labor supply of the high-skilled is more distorted downwards than the labor supply of the low-skilled.

The statement that labor supply of the high-skilled is more downwardly distorted, also allows for the possibility that it is less upwardly distorted than the labor supply of the low-skilled. Which one of these possibilities happens, depends crucially on the amount of external resources the economy has at its disposal.

binding) all respect individual's preferences, and so they evaluate the marginal individual's disutility at α_L^*, such that this term drops out.

[15] This follows from lemma B given in the proof of lemma 3 in the appendix, as for Maximin $\widetilde{S}^M_{\alpha_L^*} = \widetilde{S}^M_{\alpha_H^*} = \widetilde{S}^M_{x_L} = \widetilde{S}^M_{x_H} = 0$.

[16] For the Roemer, EE and CE criteria when $\nu > 0$ we have $\eta\left(x_L, \alpha_L^*\right) = \eta\left(x_H, \alpha_H^*\right) = \eta\left(x, \alpha^*\right)$.

5.3 Restricted second best

Lemma 1 lists the requirement of the equality of opportunity principles in the second best. In this section, we search for the optimal policies when these policies are restricted such that they satisfy at least one of the principles fully. Appendix D gives all proofs of this section.

In the discussion following lemma 1, we already noted the restrictive nature of ETES. The severity of the ETES axiom in the context of our model also appears clearly in the following theorem, which shows that, in the second best, there is only one possible allocation that satisfies ETES.

Theorem 3. *Second best optima satisfying ETES.*

There exists only one second best allocation satisfying ETES. The corresponding values for $(x^u, \alpha_L^*, \alpha_H^*)$ *are determined by* $x_L = w_L + x^u$, $x_H = w_H + x^u$ *and*

$$\begin{aligned} x^u \left[1 - 2\left[\gamma F\left(\alpha_L^*\right) + (1-\gamma) F\left(\alpha_H^*\right)\right]\right] &= R, \\ \alpha_L^* &= v\left(w_L + x^u\right) - v\left(x^u\right), \\ \alpha_H^* &= v\left(w_H + x^u\right) - v\left(x^u\right). \end{aligned}$$

Due to the severe implications of the ETES axiom in our model, we think that, in the second best model, priority should be given to the EWEP principle. We now show which allocations are second best optimal under the different criteria, when the optimum is sought under the allocations satisfying EWEP. Of course, when the optimal policies under the equality of opportunity inspired social objective functions automatically satisfy EWEP (i.e. when $\nu > 0$), the optima derived in this section for $X = R$, EE and CE will be identical to the optima in the previous subsection.

From lemma 1, (a) we know that the critical values and the consumption levels for both types of workers have to be the same. We denote this critical value by α^* and the workers' consumption by x^w:

$$v\left(x^w\right) - \alpha^* = v\left(x^u\right). \tag{18}$$

The only policy instruments of the planner are now x^w and x^u which prevents any redistribution between w_L and w_H-workers. Hence, the following programming problem describes the EWEP-restricted general second best problem:

ERGSBP (EWEP Restricted General Second Best Problem):

$$\max_{x^w, x^u, \alpha^*} \widehat{S}^X\left(x^w, x^u, \alpha^*\right),$$

subject to the government budget constraint,

$$\left[\gamma w_L + (1-\gamma) w_H - x^w\right] F\left(\alpha^*\right) - x^u\left(1 - F\left(\alpha^*\right)\right) - R = 0,$$

and constraint (18).

We define the following elasticity of participation (which is any of the previous elasticities where $x^L = x^H = x^w$ is substituted):

$$\eta\left(x^w, \alpha^*\right) \stackrel{\text{def}}{\equiv} \frac{x^w}{F\left(\alpha^*\right)} f\left(\alpha^*\right) v'\left(x^w\right). \tag{19}$$

The average of the inverse of the private marginal utility of consumption is now given by

$$g_P^X \stackrel{\text{def}}{\equiv} \frac{F(\alpha^{X*})}{v'(x^{wX})} + \frac{(1 - F(\alpha^{X*}))}{v'(x^{uX})} \ , \tag{20}$$

the effect of a uniform increase in private utilities on the social objective function equals

$$D^X = \frac{\widehat{S}^X_{x^u}}{v'\left(x^u\right)} + \frac{\widehat{S}^X_{x^w}}{v'\left(x^w\right)}, \tag{21}$$

and the average social marginal utility of workers' consumption is

$$g^X = \frac{\widehat{S}^X_{x^w}}{\lambda F\left(\alpha^*\right)}.$$

The following theorem states the solution for the EWEP restricted General Second Best Problem.

Theorem 4: *Under asymmetric Information, the optimal consumption levels have to satisfy the budget constraint, constraint (18), and the following equations:*

$$\begin{aligned} \frac{\gamma w_L + (1-\gamma)\, w_H - x^w + x^u)}{x^w} &= \frac{1}{\eta\left(x^w, \alpha^*\right)} \left[1 - g^X\right] - \frac{\widehat{S}^X_{\alpha^*}}{\lambda f\left(\alpha^*\right) x^w}. \\ \lambda^{-1} &= g_P^X / D^X. \end{aligned}$$

The interpretation of the equation for λ^{-1} is similar to the interpretation in the previous section. To obtain more specific expressions for the different social objective functions, observe that $\widehat{S}^X_{\alpha^*} = 0$ for all objective functions, except for the Non-Welfarist, for which $\widehat{S}^X_{\alpha^*} = \left(\alpha^* - \overline{\alpha}\right) f\left(\alpha^*\right)$. The average social marginal utility of consumption g^X under objective function X ($= U$, W, B, N, R, CE or EE) is given in the following table:

X	g^X
U, NW	$\frac{v'\left(x_L^X\right)}{\lambda^X}$
W	$\frac{v'\left(x^{wW}\right)}{\lambda^W} \frac{\int_0^{\alpha^{W*}} \Psi'\left(v\left(x^{wW}\right) - \alpha\right) dF(\alpha)}{F(\alpha^{W*})}$
B	$\frac{v'\left(x^{wB}\right)}{\lambda^B} \frac{\int_0^{\alpha^{B*}} W(\alpha) dF(\alpha)}{F(\alpha^{B*})}$
$R(= V)$	$\frac{v'\left(x^{wR}\right)}{\lambda^R}$
EE	$\frac{1}{\lambda^{EE} F(\alpha^{EE*})}$
CE	$\frac{v'\left(x^{wCE}\right)}{\lambda^{CE} F(\alpha^{CE*})}$

All these ingredients are used in the expressions in theorem 4 to obtain the following corollary which gives the optimal consumption levels in the restricted second best.

Corollary 6: Under asymmetric information, the second best optimal consumption levels satisfying EWEP have to satisfy the budget constraint, constraint (18) and the following equations:

X	$\left(\lambda^X\right)^{-1}$	$\frac{\gamma w_L+(1-\gamma)x_H-x^w+x^u)}{x^w}$
U	g_P^U	
W	g_P^W/D^W	
B	g_P^B/D^B	$\frac{1}{\eta(x^{wX},\alpha^{X*})}\left(1-g^X\right)$
$R=V$	g_P^R	
EE	g_P^{EE}/D^{EE}	
CE	g_P^{CE}	$\frac{1}{\eta(x^{wCE},\alpha^{CE*})}\left[1-(1-\xi)\,g^{CE}\right]$
NW	g_P^{NW}	$\frac{1}{\eta(x^{wNW},\alpha^{NW*})}\left(1-g^{NW}\right)-\frac{\alpha^{NW*}-\overline{\alpha}}{\lambda^{NW}x^{wNW}}$

with

$$D^W=\left[\int_0^{\alpha^{W*}}\Psi'\left(v\left(x^{wW}\right)-\alpha\right)dF(\alpha)+\int_{\alpha^{W*}}^{\infty}\Psi'\left(v\left(x^{uW}\right)\right)dF\left(\alpha\right)\right],$$

$$D^B=\int_0^{\infty}W\left(\alpha\right)dF\left(\alpha\right)\qquad\text{and}\qquad D^{EE}=1/v'\left(x^{wEE}\right).$$

Unsurprisingly, the optimal tax formulas have the same shape as in the previous subsection, but now the constraint $x_H=x_L$ is imposed. The major difference is due to the fact that EWEP impedes the government to distinguish between low and high-skilled workers, such that the formula now have to hold for an imaginary worker who has average productivity and thus average wage $\gamma w_L+(1-\gamma)\,w_H$.

5.4 Priority principles

The social choice literature to equality of opportunity argues that, since compensation and responsibility cannot be fully satisfied in general, only a maximin variant makes sense (Fleurbaey, 2008). Therefore, rather than strictly imposing one of the equality of opportunity principles and searching for the optimal allocation satisfying it, this section examines the optimal tax policies when priority to the worst off is given. The strict equality demanded by each of the principles is weakened and replaced with maximin and we search for social orderings that embody this weak version of the principle.[17]

EWEP requires that for each value of α, welfares are equalized. Rather than insisting on full equality, the priority principle requires that social states are judged, for each α, by the welfare level obtained by the skill level L or H, that has the lowest welfare. It expresses the idea that the allocation of consumption levels and jobs between two individuals with identical tastes should be such it is impossible to redistribute among them and increase the level of well-being of the least well off.

The question then becomes how to measure individuals' welfare. A first possibility is to measure welfare by individual utilities. Roemer's criterion applies a Utilitarian aggregation to these minimal levels of welfare, but other aggregation procedures are possible, such as a Welfaristic and a Boadway

[17] For an interesting alternative, social choice approach, starting from such priority principles see Fleurbaey and Maniquet (2005 and 2007).

et al. variant, leading to the Priority Welfare weighted Utility ordering

$$\widetilde{S}^{PWU} = \int_0^\infty \Omega^R \left(\min\{\underset{\delta_L(\alpha)}{\text{oper}}\{v\left(x_L^w\left(\alpha\right)\right) - \alpha, v\left(x_L^u\left(\alpha\right)\right)\}, \right.$$

$$\left. \underset{\delta_H(\alpha)}{\text{oper}}\{v\left(x_H^w\left(\alpha\right)\right) - \alpha, v\left(x_H^u\left(\alpha\right)\right)\}\} \right) dF\left(\alpha\right), \tag{22}$$

where $\Omega^R(\cdot)$ is a welfare function with $\Omega^{R\prime}(\cdot) > 0$ and the Priority Taste weighted Utility ordering

$$\widetilde{S}^{PTU} = \int_0^\infty \Phi^R\left(\alpha\right) \left[\min\{\underset{\delta_L(\alpha)}{\text{oper}}\{v\left(x_L^w\left(\alpha\right)\right) - \alpha, v\left(x_L^u\left(\alpha\right)\right)\}, \right.$$

$$\left. \underset{\delta_H(\alpha)}{\text{oper}}\{v\left(x_H^w\left(\alpha\right)\right) - \alpha, v\left(x_H^u\left(\alpha\right)\right)\}\} \right] dF\left(\alpha\right), \tag{23}$$

where $\Phi^R(\alpha) > 0$ weights different tastes. These two objective functions are clearly distinct: S^{PTU} allows the planner to express inequality aversion (preference) with respect to utility differences that arise due to differences in tastes if $\Omega^{R\prime\prime}(\cdot) < (>)\, 0$, while in S^{PWU} the planner gives different weights to different tastes as such, irrespective of their welfare levels. Both are generalizations of Roemer's criterion but they do not respect the utilitarian reward principle (see Fleurbaey (2008)), which requires zero aversion to inequalities due to different preferences. However, if the planner wants to express an opinion about welfare inequality that arizes due to differences in tastes, these specifications allow the planner to do so.

A second approach consists in taking an ordinal measure of welfare. We can find here inspiration with the reasoning that leads to the Egalitarian Equivalent ordering, and take that consumption level a person requires when he works that makes him indifferent to his actual consumption bundle. The aggregation of these welfare levels can occur again in a Welfarist or a Boadway *et al.* way, leading to the Priority Welfare weighted Equivalent ordering

$$S^{PWE} = \int_0^\infty \Omega^O \left(\min\{\underset{\delta_L(\alpha)}{\text{oper}}\{x_L^w\left(\alpha\right), v^{-1}\left(v\left(x_L^u\left(\alpha\right)\right) + \alpha\right)\}, \right.$$

$$\left. \underset{\delta_H(\alpha)}{\text{oper}}\{x_H^w\left(\alpha\right), v^{-1}\left(v\left(x_H^u\left(\alpha\right)\right) + \alpha\right)\} \right) dF\left(\alpha\right), \tag{24}$$

where $\Omega^O(\cdot)$ is a welfare function with $\Omega^{O\prime}(\cdot) > 0$ and the Priority Taste weighted Equivalent ordering

$$S^{PTE} = \int_0^\infty \Phi^O\left(\alpha\right) \left[\min\{\underset{\delta_L(\alpha)}{\text{oper}}\{x_L^w\left(\alpha\right), v^{-1}\left(v\left(x_L^u\left(\alpha\right)\right) + \alpha\right)\}, \right.$$

$$\left. \underset{\delta_H(\alpha)}{\text{oper}}\{x_H^w\left(\alpha\right), v^{-1}\left(v\left(x_H^u\left(\alpha\right)\right) + \alpha\right)\}\} \right] dF\left(\alpha\right), \tag{25}$$

where $\Phi^O(\alpha) > 0$ weights different tastes. If the welfare function $\Omega^O(\cdot)$ becomes infinitely inequality averse, the social welfare function (24) reduces to the egalitarian equivalent ordering (9).[18]

[18] In a recent contribution Hodler (2009) proposes to measure inequality in societies with unequal earning abilities and tastes for work by computing traditional inequality indices (Gini, Atkinson-Kolm ,Theil, ...) for equivalent wages in the *entire* population. When interested in inequality, one can do something similar here, but the priority principle forces us to take, for each value of tastes, only the lowest equivalent wage into account.

ETES requires that transfers are the same for all those that have equal skills. To apply the priority principle here, for each level of skill we have to consider the lowest transfer received by an individual with that skill level. Since we have only two levels of skill, a social ordering embodying the priority principle would be the following Priority Transfer ordering

$$\begin{aligned} S^{PT} &= \rho \min_{\alpha \in \mathbb{R}^+} \left\{ \underset{\delta_L(\alpha)}{\text{oper}} \left\{ x_L^w(\alpha) - w_L, x_L^u(\alpha) \right\} \right\} \\ &\quad + (1-\rho) \min_{\alpha \in \mathbb{R}^+} \left\{ \underset{\delta_H(\alpha)}{\text{oper}} \left\{ x_H^w(\alpha) - w_H, x_H^u(\alpha) \right\} \right\}, \end{aligned} \tag{26}$$

where $\rho \in [0, 1]$ gives the relative importance attached to the low-skilled agents.

The following lemma gives expressions for these new objective functions in the second best framework.

Lemma 4: priority social objective functions in the second best.

$$\widetilde{S}^{PWU} = \int_0^{\alpha_L^*} \Omega^R \left(v(x_L) - \alpha\right) dF(\alpha) + \int_{\alpha_L^*}^{\infty} \Omega^R \left(v(x^u)\right) dF(\alpha).$$

$$\widetilde{S}^{PTU} = \int_0^{\alpha_L^*} \Phi^R(\alpha) \left[v(x_L) - \alpha\right] dF(\alpha) + \int_{\alpha_L^*}^{\infty} \Phi^R(\alpha)\, v(x^u)\, dF(\alpha).$$

$$\widetilde{S}^{PWE} = \int_0^{\alpha_L^*} \Omega^O(x_L)\, dF(\alpha) + \int_{\alpha_L^*}^{\infty} \Omega^O\left(v^{-1}\left(v(x^u) + \alpha\right)\right) dF(\alpha).$$

$$\widetilde{S}^{PTE} = x_L \int_0^{\alpha_L^*} \Phi^O(\alpha)\, dF(\alpha) + \int_{\alpha_L^*}^{\infty} \Phi^O(\alpha) \left(v^{-1}\left(v(x^u) + \alpha\right)\right) dF(\alpha).$$

$$\widetilde{S}^{PT} = \rho\left(x_L - w_L\right) + (1-\rho)\left(x_H - w_H\right).$$

The problem of finding the optimal tax rates with these objective functions has exactly the same structure as the General Second Best Problem formulated in section 5.1, and whose solution is given by theorem 2.

Lemma 5: the value of $\widetilde{S}^X_{\alpha_Y^*}$ $(Y = L, H)$:
$\widetilde{S}^X_{\alpha_Y^*} = 0$ for $X = PWU$, PTU, PWE, PTE and PT.

Lemma 6: the value of the Lagrangian multiplier:
$\nu \geq 0$ for $X = PWU$, PTU, PWE, PTE and PT.

Combining lemma 6 with lemma 1 (a) we see how the different criteria perform from the EWEP-perspective: for PWU, PTU, PWE, PTE and PT the constraint $x_H \geq x_L$ can be binding, in which case $x_H = x_L$, $\alpha_H^* = \alpha_L^*$ and their solution satisfies EWEP.

The average social marginal utility of consumption g_Y^X under objective function X $(= PWU$, PTU, PWE, PTE or $PT)$ for agents of skill level Y $(= L$ or $H)$ are given in the following table.

Using these expressions in theorem 2 results in the following corollary.

Corollary 7. Under asymmetric information, the optimal consumption levels have to satisfy the budget constraint, constraints (10), (11) and (12) and the following equations:

X	g_L^X	g_H^X
PWU	$\frac{v'\left(x_L^{PWU}\right)}{\lambda^{PWU}} \frac{\int_0^{\alpha_L^{PWU*}} \Omega^{R\prime}\left(v\left(x_L^{PWU}\right)-\alpha\right)dF(\alpha)}{\gamma F(\alpha_L^{PWU*})}$	0
PTU	$\frac{v'\left(x_L^{PTU}\right)}{\lambda^{PTU}} \frac{\int_0^{\alpha_L^{PTU*}} \Phi^{R}(\alpha)dF(\alpha)}{\gamma F(\alpha_L^{PTU*})}$	0
PWE	$\frac{\Omega^{O\prime}\left(x_L^{PWE}\right)}{\lambda^{PWE}\gamma}$	0
PTE	$\frac{1}{\lambda^{PTE}\gamma} \frac{\int_0^{\alpha_L^{PTE*}} \Phi^{O}(\alpha)dF(\alpha)}{F(\alpha_L^{PTE*})}$	0
PT	$\frac{\rho}{\lambda^{PT}\gamma F\left(\alpha_L^{PT*}\right)}$	$\frac{1-\rho}{\lambda^{PT}(1-\gamma)F(\alpha_H^{PT*})}$

<table>
<tr><th>X</th><th>$\left(\lambda^X\right)^{-1}$</th><th>$\frac{T_H^X-T_u^X}{x_H^X}$</th><th>$\frac{T_L^X-T_u^X}{x_L^X}$</th></tr>
<tr><td>PWU</td><td>g_P^{PWU}/D^{PWU}</td><td rowspan="4">$\frac{1}{\eta\left(x_H^X,\alpha_H^{X*}\right)}\left[1-\frac{\nu^X}{\lambda^X(1-\gamma)F\left(\alpha_H^{X*}\right)}\right]$</td><td rowspan="5">$\frac{1}{\eta\left(x_L^X,\alpha_L^{X*}\right)}$ $\times$ $\left(1-g_L^X+\frac{\nu^X}{\lambda^X\gamma F\left(\alpha_L^{X*}\right)}\right)$</td></tr>
<tr><td>PTU</td><td>g_P^{PTU}/D^{PTU}</td></tr>
<tr><td>PWE</td><td>g_P^{PWE}/D^{PWE}</td></tr>
<tr><td>PTE</td><td>g_P^{PTE}/D^{PTE}</td></tr>
<tr><td>PT</td><td>g_P^{PT}/D^{PT}</td><td>$\frac{1}{\eta\left(x_H^{PT},\alpha_H^{PT*}\right)}\left(1-g_H^{PT}-\frac{\nu}{\lambda^{PT}(1-\gamma)F\left(\alpha_H^{PT*}\right)}\right)$</td></tr>
</table>

with

$$D^{PWU}=\int_0^{\alpha_L^{PWU*}}\Omega^{R\prime}\left(v(x_L^{PWU})-\alpha\right)dF(\alpha)+\int_{\alpha_L^{PWU*}}^{\infty}\Omega^{R\prime}\left(v\left(x^{uPWU}\right)\right)dF(\alpha),$$

$$D^{PTU}=\int_0^{\alpha_L^{PTU*}}\Phi^{R}(\alpha)\,dF(\alpha)+\int_{\alpha_L^{PTU*}}^{\infty}\Phi^{R}(\alpha)\,dF(\alpha),$$

$$D^{PWE}=\frac{\Omega^{O\prime}\left(x_L^{PWE}\right)F\left(\alpha_L^{PWE*}\right)}{v'\left(x_L^{PWE}\right)}+\frac{\int_{\alpha_L^{PWE}}^{\infty}\Omega^{O\prime}\left(v^{-1}\left(v\left(x^{uPWE}\right)+\alpha\right)\right)\frac{\partial\left(v^{-1}\left(v\left(x^{uPWE}\right)+\alpha\right)\right)}{\partial x^{uPWE}}dF(\alpha)}{v'\left(x^{uPWE}\right)}$$

$$D^{PTE}=\frac{\int_0^{\alpha_L^{PTE*}}\Phi^{O}(\alpha)\,dF(\alpha)}{v'(x_L^{PTE})}+\frac{\int_{\alpha_L^{PTE*}}^{\infty}\Phi^{O}(\alpha)\frac{\partial\left(v^{-1}\left(v\left(x^{uPTE}\right)+\alpha\right)\right)}{\partial x^{uPTE}}dF(\alpha)}{v'\left(x^{uPTE}\right)}$$

$$D^{PT}=\frac{\rho}{v'\left(x_L^{PT}\right)}+\frac{1-\rho}{v'(x_H^{PT})}$$

The optimal tax rates have the same structure under the PWU, PTU, PWE, PTE and PT social objective functions in the sense that the multiplier ν pushes the tax system away from the EITC. This is shown in the next lemma.

Corollary 8. Optimality of EITC or NIT in second best with a priority requirement.

Social Objective	ETES/ NIT/ EITC?
PWU, PTU, PWE, PTE and PT	NIT (EITC) if $g_L^X<(>)\,1+\frac{\nu^X}{\lambda^X\gamma F\left(\alpha_L^{X*}\right)}$

Under assumption 1, we can derive the following counterparts to corollaries 4 and 5.[19]

[19] Again, $\nu>0$ we have $\eta\left(x_L,\alpha_L^*\right)=\eta\left(x_H,\alpha_H^*\right)=\eta\left(x,\alpha^*\right)$. Moreover, the proof of the corollaries is similar to the proofs of corollaries 4 and 5 for the Roemer, EE and CE social objective functions, and is suppressed.

Corollary 9. Under assumption 1, for the PWU, PTU, PWE, PTE and PT social objective functions:

$$(T_L - T_u)/x_L < (T_H - T_u)/x_H$$

Corollary 10. Under assumption 1, for the PWU, PTU, PWE, PTE and PT social objective functions , the labor supply of the high-skilled is more distorted downwards than the labor supply of the low-skilled.

6 Conclusion

This paper has studied optimal tax policies when agents differ in terms of skills and tastes for labor. We assumed quasilinear utility and that labor supply decision is at the extensive margin. The optimal tax policies under distinct objective functions have been derived, in full and asymmetric information.

The determination of appealing social criteria is important if one looks for social preferences applicable in public economics, in particular when dealing with redistribution. When agents differ in terms of skills and tastes for labor, the equality of opportunity approach is inspiring (Fleurbaey, 1995a) and broadly accepted (Alesina and Angeletos, 2005).

This paper has shown that many criteria in the optimal tax literature (Utilitarianism, Welfarism, Boadway *et al.*, Van de gaer and Non-Welfarist criteria) fail the requirements of equality of opportunity, i.e. the compensation (EWEP) and responsibility (ETES) principles. It has been shown that, in the first best, criteria respecting one of these principles are Roemer's, the Conditional Equality and the Egalitarian Equivalent criterion, the latter two advocated by Fleurbaey (1995b). Given that these criteria were designed so as to meet one of the principles in the first best, this should not come as a surprise. We also showed that in the second best, these criteria might satisfy EWEP, while the standard criteria from the optimal tax literature never satisfy it. The difference between the standard approaches and the equality of opportunity approach is not just a difference between the way social marginal utilities of incomes of individuals are aggregated, but goes much deeper.

In this paper, we explore two ways to deal with the equality of opportunity principles in the second best model. One is to search for optimal policies over the allocations that satisfy one of the principles. The other is to weaken the full equality demanded in the equality of opportunity principles, and replace them by priority principles, as advocated in social choice (Fleurbaey, 2008). We therefore build up new criteria; one satisfying an ETES-priority principle and several others satisfying EWEP-priority principles leading to generalizations of Roemer's criterion and the egalitarian equivalent allocation. They have similar properties to the other equality of opportunity principles, but allow the researcher to express different kinds and extents of inequality aversion. Throughout we find that the equality of opportunity approach tends to work against an Earned Income Tax Credit and in favor of a Negative Income Tax.

References

Alesina A. and G-M. Angeletos (2005), Fairness and Redistribution, *American Economic Review*, 95(4): 960-980.

Atkinson, A.B. and J.E. Stiglitz (1980), Lectures on Public Economics, New York: Mc Graw-Hill.

Boadway, R., Cuff, K. and M. Marchand (2000), Optimal Income Taxation with Quasi-linear Preferences Revisited, *Journal of Public Economic theory*, 2(4): 435-460.

Boadway, R., Marchand, M., Pestieau, P. and M. Del Mar Racionero (2002), Optimal redistribution with heterogeneous preferences for leisure, *Journal of Public Economic Theory*, 4: 475-498.

Bossert, W. (1995), Redistribution mechanisms based on individual characteristics, *Mathematical Social Sciences,* 29: 1-17.

Bossert, W. and M. Fleurbaey (1996), Redistribution and compensation, *Social Choice and Welfare* 13: 343-355.

Bossert, W., Fleurbaey, M. and D. Van de gaer (1999), Responsibility, talent, and compensation: a second best analysis, *Review of Economic Design,* 4: 35-55.

Choné, P. and G. Laroque (2005), Optimal Incentives for Labor Force Participation, *Journal of Public Economics*, 89: 395-425.

Choné, P. and G. Laroque (2009), Should labor force participation be subsidized?, mimeo INSEE-CREST.

Cremer, H., Lozachmeur, J-M. and P. Pestieau, (2004), Optimal retirement and disability benefits with audit, *Finanzarchiv*, 60(3): 278—295.

Cremer, H. Lozachmeur, J-M. and P. Pestieau (2007), Disability Testing and Retirement, *The B. E. Journal of Theoretical Economics* (Contributions), 7(1), Article 10.

Diamond P. (1980), Income Taxation with Fixed Hours of Work, *Journal of Public Economics*, 13: 101-110.

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

Diamond P. and E. Sheshinski (1995), Economic Aspects of Optimal Disability Benefits, *Journal of Public Economics*, 57: 1-23.

Dworkin, R. (1981), What is Equality? Part 1. Equality of Welfare. *Philosophy and Public Affairs*, 10: 185-246.

Ebert, U. (1992), A reexamination of the optimal nonlinear income tax, *Journal of Public Economics*, 49: 47-73.

Fleurbaey, M. (1994), On fair compensation, *Theory and Decision*, 36: 277-307.

Fleurbaey, M. (1995a), The Requisites of Equal Opportunity. In: Barnett, W.A., Moulin, H., Salles, M. and Schofield, N.J. (Eds.), *Social Choice, Welfare and Ethics*, Cambridge: Cambridge University Press: 37-53.

Fleurbaey, M. (1995b), Three solutions for the compensation problem, *Journal of Economic Theory*, 65: 505-521.

Fleurbaey, M. (2005), Health, Wealth and Fairness, *Journal of Public Economic Theory*, 7: 253-284.

Fleurbaey, M. (2008), Fairness, Responsibility and Welfare, Oxford: Oxford University Press.

Fleurbaey, M. and F. Maniquet (2005), Fair social orderings when agents have unequal production skills, *Social Choice and Welfare*, 24: 93-127.

Fleurbaey, M. and F. Maniquet (2006), Fair Income Tax, *Review of Economic Studies*, 73: 55-83.

Fleurbaey, M. and F. Maniquet (2007), Help the Low Skilled or Let the Hardworking Thrive? A Study of Fairness in Optimal Income Taxation, *Journal of Public Economic Theory*, 9(3): 467-500.
Fleurbaey, M. and F.Maniquet (2009), Compensation and responsibility, in Handbook of Social Choice and Welfare Vol.2, EDS: Arrow, K.J., Sen, A.K. and Suzumura, K., Elsevier North Holland.
Hellwig, M.F. (2007), A contribution to the theory of optimal utilitarian income taxation, *Journal of Public Economics*, 91: 1449–1477.
Hodler, R. (2009), Redistribution and Inequality in a Heterogeneous Society, forthcoming in *Economica.*
Immervoll, H., Kleven, H., Kreiner, C. T. and E. Saez (2007), Welfare reforms in European countries: a microsimulation analysis, *Economic Journal*, 117: 1-44.
Kanbur, R., Pirttilä, J. and M. Tuomala (2006), Non-Welfarist Optimal Taxation and Behavioural Public Economics, *Journal of Economic Surveys*, 20(5): 849-868.
Kaplow, L. (2008), Optimal Policy with Heterogeneous Preferences, *The B.E. Journal of Economic Analysis & Policy*, 8(1) (Advances), Art. 40.
Kleven H.J. , Kreiner, C.T. and E. Saez (2009), The Optimal Income Taxation of Couples, *Econometrica*, 77: 537-560.
Marchand, M. Pestieau, P. and M. Racionero (2003), Optimal Redistribution When Different Workers Are Indistinguishable, *Canadian Journal of Economics*, 36(4): 911-922.
Meghir C. and D. Phillips (2008), Labour Supply and Taxes, IZA discussion paper 3405.
Ooghe, E., Schokkaert, E. and D. Van de gaer (2007), Equality of opportunity versus equality of opportunity sets, *Social Choice and Welfare*, 28: 209-230.
Pestieau, P. and M. Racionero (2009), Optimal redistribution with unobservable disability: welfarist versus non-welfarist social objectives, *European Economic Review* forthcoming and CORE DP 2007/35.
Rawls, J. (1971), A Theory of Justice, Oxford: Oxford University Press.
Roemer, J. (1993), A pragmatic theory of responsibility for the egalitarian planner, *Philosophy and Public Affairs*, 22: 146-166.
Roemer, J. (1998), Equality of Opportunity. Harvard Univ. Press, Cambridge.
Roemer, J., Aaberge, R., Colombino, U., Fritzell, J., Jenkins, S.P., Lefranc, A., Marx, I., Page, M., Pommer, E., Ruiz-Castillo, J., San Segundo, M.J., Tranaes, T., Trannoy, A., Wagner, G.G., Zubiri, I. (2003), To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?, *Journal of Public Economics*, 87: 539–565.
Saez, E. (2002), Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply responses, *Quarterly Journal of Economics*, 117: 1039-1073.
Schokkaert, E., Van de gaer, D., Vandenbroucke, F. and R.I. Luttens (2004), Responsibility sensitive egalitarianism and optimal linear income taxation, *Mathematical Social Sciences*, 48: 151-182.
Sen, A. (1980), Equality of What? In: Sen, A. (Ed), Choice, Welfare and Measurement. Oxford: Blackwell: 123-162.
Van de gaer, D. (1993), Equality of opportunity and investment in human capital. *PhD Thesis, KULeuven*, Leuven, Belgium.

Appendix A: Proofs of Section 4.

PROOF OF THEOREM 1a.

The Lagrangian functions for each of the social objective functions are formed by combining the expressions for the social objective functions given in Section 3.2, the government budget constraint (1) and its associated Lagrangian multiplier λ.

(a) Welfarist and (b) Utilitarian planners

We discuss the Welfarist case first, and show how the properties of the Utilitarian case follow. The first-order conditions of the constrained optimization problem with respect to the four consumption functions are:

$$\begin{aligned}
\delta_L(\alpha)\left[\Psi'(v(x_L^w(\alpha))-\alpha)v'(x_L^w(\alpha))-\lambda\right] &= 0, \\
(1-\delta_L(\alpha))\left[\Psi'(v(x_L^u(\alpha))v'(x_L^u(\alpha))-\lambda\right] &= 0, \\
\delta_H(\alpha)\left[\Psi'(v(x_H^w(\alpha))-\alpha)v'(x_H^w(\alpha))-\lambda\right] &= 0, \\
(1-\delta_H(\alpha))\left[\Psi'(v(x_H^u(\alpha))v'(x_H^u(\alpha))-\lambda\right] &= 0.
\end{aligned}$$

Since $\delta_L(\alpha)$ and $\delta_H(\alpha)$ are equal to 1 or 0, for each value of α, only two of these first-order conditions matter; for those that matter the corresponding social marginal utilities of consumption have to be equal, for the other two the consumption function does not matter (as nobody with this value for α is receiving it). So we get for all those that do not work:

$$\Psi'(v(x_L^u(\alpha)))v'(x_L^u(\alpha))=\lambda=\Psi'(v(x_H^u(\alpha)))v'(x_H^u(\alpha)). \tag{27}$$

Due to the strict concavity of $\Psi'(\cdot)$ and $v'(\cdot)$, this can only hold true if

$$\overline{x}^u=x_L^u(\alpha)=x_H^u(\alpha).$$

For those that work, we get

$$\Psi'(v(x_L^w(\alpha))-\alpha)v'(x_L^w(\alpha))=\lambda=\Psi'(v(x_H^w(\alpha))-\alpha)v'(x_H^w(\alpha)). \tag{28}$$

For a given value for α, the requirement is exactly the same for w_L- and w_H-workers. Hence, for a given value of α, both get the same consumption bundle and so, for all α :

$$x_L^w(\alpha)=x_H^w(\alpha). \tag{29}$$

Hence worker's consumption bundles depend on α. Moreover, from the implicit function theorem:

$$\frac{\partial x_L^w(\alpha)}{\partial\alpha}=\frac{\Psi''(v(x_L^w(\alpha))-\alpha)v'(x_L^w(\alpha))}{\Psi''(v(x_L^w(\alpha))-\alpha)\left[v'(x_L^w(\alpha))\right]^2+\Psi'(v(x_L^w(\alpha))-\alpha)v''(x_L^w(\alpha))}>0, \tag{30}$$

Therefore, for $\alpha_1<\alpha_2$, due to the concavity of $v(.)$ we have:

$$v'(x_L^w(\alpha_1))>v'(x_L^w(\alpha_2)).$$

Combining the last inequality with (28) requires that $\Psi'(v(x_L^w(\alpha_1)) - \alpha_1) < \Psi'(v(x_L^w(\alpha_2)) - \alpha_2)$. Since Ψ is strictly concave, this requires that

$$v\left(x_L^w(\alpha_1)\right) - \alpha_1 > v(x_L^w(\alpha_2)) - \alpha_2,$$

and so low-skilled workers with a higher disutility of labor are not fully compensated for this higher disutility. Due to (29), the same holds for high-skilled workers. Note that from (28) with $\alpha = 0$ and (27) we get that

$$x_L^w(0) = x_H^w(0) = \overline{x}_u.$$

The government budget constraint only depends on the number of high and low-skilled that work, not on which high and low-skilled. From (30) workers' consumption is increasing in their disutility of work, and so it is cheapest and hence optimal for the government to make those work with the lowest α. In view of (29), putting high-skilled and low-skilled at work is equally expensive for the government, but since high-skilled contribute more to the budget than low-skilled, more high-skilled than low-skilled will have to work. Hence, there exist critical values for α_L^* and α_H^* such that

$$\delta_L(\alpha) = 1 \text{ for all } \alpha \leq \alpha_L^*, \delta_H(\alpha) = 1 \text{ for all } \alpha \leq \alpha_H^* \text{ and } \alpha_L^* < \alpha_H^*. \tag{31}$$

The Welfarist criterion reduces to the Utilitarian one when $\Psi(.) \overset{def}{\equiv} id(.)$ hence $\Psi'(.) = 1$. Therefore, under the Utilitarian criterion, (27)-(28) yield that the first-order conditions with respect to consumption reduce to $(\forall\alpha)$ (since λ is a constant):

$$\begin{aligned} &v'(x_L^{wU}(\alpha)) = v'(x_L^{uU}(\alpha)) = v'(x_H^{wU}(\alpha)) = v'(x_H^{uU}(\alpha)) = \lambda \\ &\Longleftrightarrow \overline{x} = x_L^{wU}(\alpha) = x_L^{uU}(\alpha) = x_H^{wU}(\alpha) = x_H^{uU}(\alpha)\ . \end{aligned} \tag{32}$$

Since all individuals get the same consumption bundle, it follows from the reasoning leading to (31) that $\alpha_L^{U*} < \alpha_H^{U*}$.

(c) Boadway *et al.* planner

The first-order conditions with respect to consumption functions (assuming an interior solution) are:

$$\begin{aligned} &\int_0^\infty \delta_L(\alpha)\left[W(\alpha)\, v'(x_L^w(\alpha)) - \lambda\right] dF(\alpha) = 0, \\ &\int_0^\infty (1 - \delta_L(\alpha))\left[W(\alpha)\, v'(x_L^u(\alpha)) - \lambda\right] dF(\alpha) = 0, \\ &\int_0^\infty \delta_H(\alpha)\left[W(\alpha)\, v'(x_H^w(\alpha)) - \lambda\right] dF(\alpha) = 0, \\ &\int_0^\infty (1 - \delta_H(\alpha))\left[W(\alpha)\, v'(x_H^u(\alpha)) - \lambda\right] dF(\alpha) = 0. \end{aligned}$$

Consequently, we get

$$\begin{aligned} &v'(x_L^w(\alpha)) = v'(x_L^u(\alpha)) = v'(x_H^w(\alpha)) = v'(x_H^u(\alpha)) = \frac{\lambda}{W(\alpha)} \\ &\Longleftrightarrow x(\alpha) = x_L^w(\alpha) = x_L^u(\alpha) = x_H^w(\alpha) = x_H^u(\alpha)\ . \end{aligned} \tag{33}$$

Given α, it is equally costly to have high and low-skilled at work, but since high skilled workers contribute more to the government budget, the government always prefers to have more high than low-skilled at work. From (33), consumption depends on taste for leisure. Application of the implicit function theorem to the equation $v'(x(\alpha)) = \frac{\lambda}{W(\alpha)}$ yields:

$$\frac{\partial x(\alpha)}{\partial \alpha} = -\frac{\lambda}{[W(\alpha)]^2}\frac{W'(\alpha)}{v''(x(\alpha))} \geq (\leq)\, 0 \text{ if } W'(.) \geq (\leq)\, 0.$$

Using (33) in the government budget constraint (1) yields that the function $x(\alpha)$ must be such that

$$\int_0^\infty x(\alpha)\, dF(\alpha) = \gamma w_L n_L + (1-\gamma) w_H n_H - R.$$

For the government budget constraint it only matters how many high and low-skilled people work, it does not matter which high and low-skilled people work. Hence, differential treatment in job assignment between equally skilled people must be based on the objective function. Using (4), the value of the objective function is given by:

$$S^B = \int_0^\infty W(\alpha)\, v(x(\alpha))\, dF(\alpha) - \gamma \int_0^\infty W(\alpha)\, \delta_L(\alpha)\, \alpha dF(\alpha) - (1-\gamma)\int_0^\infty W(\alpha)\, \delta_H(\alpha)\, \alpha dF(\alpha).$$

Whether people with high or low disutility of effort should be working depends on the last two terms of this expression. If $W(\alpha)\, a$ is increasing, having people with a high disutility working is not a good idea. From this it follows that, if the elasticity of the weight function ($\frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)}$) is larger than -1, then it is optimal for the government not to employ people that have a high disutility of work. If this elasticity is smaller than -1, it will be optimal to employ people with a high disutility of work. Consequently, the functions $\delta_L(\alpha)$ and $\delta_H(\alpha)$ can have different shapes:

Case 1: $\frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)} > -1 : \delta_L(\alpha) = 1$ for all $\alpha \leq \alpha_L^*$, $\delta_H(\alpha) = 1$ for all $\alpha \leq \alpha_H^*$ and $\alpha_L^* < \alpha_H^*$,

Case 2: $\frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)} = -1$ (i.e. $W(\alpha)\,\alpha$ is a constant): see discussion below,

Case 3: $\frac{\partial W(\alpha)}{\partial \alpha}\frac{\alpha}{W(\alpha)} < -1 : \delta_L(\alpha) = 1$ for all $\alpha \geq \alpha_L^{**}$, $\delta_H(\alpha) = 1$ for all $\alpha \geq \alpha_H^{**}$ and $\alpha_L^* > \alpha_H^*$.

Analyzing case 2 in more detail, the problem facing the planner with $W(\alpha)\,\alpha$ constant has the following Lagrangian:

$$\begin{aligned} &\pounds(x(\alpha), n_L, n_H, \lambda) = \int_0^\infty W(\alpha)\, v(x(\alpha))\, dF(\alpha) - \gamma W(\alpha)\, \alpha n_L - (1-\gamma) W(\alpha)\, \alpha n_H \\ &+\lambda\left[\gamma w_L n_L + (1-\gamma) w_H n_H - \int_0^\infty x(\alpha)\, dF(\alpha) - R\right], \end{aligned}$$

which leads to the following:

$$\begin{aligned} \frac{\partial \pounds}{\partial x(\alpha)} &= 0 \Leftrightarrow \lambda = \int_0^\infty W(\alpha)\, v'(x(\alpha))\, dF(\alpha), \\ \frac{\partial \pounds}{\partial n_L} &= -\gamma W(\alpha)\, \alpha + \lambda \gamma w_L, \\ \frac{\partial \pounds}{\partial n_H} &= (1-\gamma) W(\alpha)\, \alpha + \lambda (1-\gamma) w_H. \end{aligned}$$

Note that the second and third condition cannot hold simultaneously with equality:

$$\begin{aligned} \frac{\partial \pounds}{\partial n_L} &\geq (\leq)\, 0 \Leftrightarrow [\lambda w_L - W(\alpha)\, \alpha] \geq (\leq)\, 0, \\ \frac{\partial \pounds}{\partial n_H} &\geq (\leq)\, 0 \Leftrightarrow [\lambda w_H - W(\alpha)\, \alpha] \geq (\leq)\, 0. \end{aligned}$$

Hence, since $w_H > w_L$, we always have that $\frac{\partial \pounds}{\partial n_L} \geq 0 \Rightarrow \frac{\partial \pounds}{\partial n_H} > 0$ and $\frac{\partial \pounds}{\partial n_H} \leq 0 \Rightarrow \frac{\partial \pounds}{\partial n_L} < 0$. We then get the possibilities listed in case 2 of the theorem.

(d) Non-Welfarist social planner

It is easy to see that we obtain the same first-order conditions as with the Utilitarian objective, and so the consumption functions are similar to (32): everybody receives the same level of consumption $\overline{x}$, which, because of the government budget constraint equals $\gamma w_L n_L + (1-\gamma) w_H n_H - R$. Consequently, using (5), the value of our Non-Welfaristic objective function becomes

$$v\left(\gamma w_L n_L + (1-\gamma) w_H n_H\right) - \gamma \overline{\alpha} n_L - (1-\gamma) \overline{\alpha} n_H.$$

This expression only depends on the number of low and high-skilled that are employed; the planner determines n_L and n_H so as to maximize this expression. The derivatives of this expression with respect to n_H and n_L are, respectively

$$(1-\gamma)\left[w_H v'(\overline{x}) - \overline{\alpha}\right] \quad \text{and} \quad \gamma\left[w_L v'(\overline{x}) - \overline{\alpha}\right].$$

Since $w_H > w_L$, we can distinguish the cases listed in the theorem.

(e) Roemer planner

There is no point in allowing the two elements in the min operator of Roemer's objective function to be different in the first best. Hence there are in principle four possibilities:

(i) $\delta_L(\alpha) = \delta_H(\alpha) = 1 \Rightarrow x_L^w(\alpha) = x_H^w(\alpha)$,

(ii) $\delta_L(\alpha) = 0, \delta_H(\alpha) = 1 \Rightarrow v(x_L^u(\alpha)) = v(x_H^w(\alpha)) - \alpha \Rightarrow x_L^u(\alpha) < x_H^w(\alpha)$,

(iii) $\delta_L(\alpha) = \delta_H(\alpha) = 0 \Rightarrow x_L^u(\alpha) = x_H^u(\alpha)$,

(iv) $\delta_L(\alpha) = 1, \delta_H(\alpha) = 0 \Rightarrow v(x_L^w(\alpha)) - \alpha = v(x_H^u(\alpha)) \Rightarrow x_L^w(\alpha) > x_H^u(\alpha)$.

There is equivalence between the maximin approach and the revenue-maximizing approach. Maximizing tax revenue subject to a minimal utility level is equivalent to maximizing the minimum of utility subject to the revenue constraint. Here, the objective function maximizes the sum of the minimal utility levels but the logic is similar. The government maximizes the tax revenue subject to minimal utility levels. The tax revenue will be maximized the more people are working, in particular productive people. The minimal utility levels avoid that people with large α work. Therefore, if anyone, we would like the ones with low values for α to work, and since highly skilled have a higher productivity, we want more highly skilled to work ($\alpha_H^* \geq \alpha_L^*$); for α increasing, we move from (i) over (ii) to (iii). If we plug this in, we get the following objective function:

$$\int_0^{\alpha_L^*} \min\left\{v\left(x_L^w(\alpha)\right) - \alpha, v\left(x_H^w(\alpha)\right) - \alpha\right\} dF(\alpha)$$
$$+ \int_{\alpha_L^*}^{\alpha_H^*} \min\left\{v\left(x_L^u(\alpha)\right), v\left(x_H^w(\alpha)\right) - \alpha\right\} dF(\alpha) + \int_{\alpha_H^*}^{\infty} \min\left\{v\left(x_L^u(\alpha)\right), v(x_H^u(\alpha))\right\} dF(\alpha).$$

Maximizing this objective function implies

$$x_L^w(\alpha) = x_H^w(\alpha) \quad \forall \alpha \in [0, \alpha_L^*), \tag{34}$$
$$x_L^u(\alpha) = v^{-1}\left(v\left(x_H^w(\alpha)\right) - \alpha\right) \quad \forall \alpha \in [\alpha_L^*, \alpha_H^*), \tag{35}$$
$$x_L^u(\alpha) = x_H^u(\alpha) \quad \forall \alpha \in [\alpha_H^*, \infty). \tag{36}$$

Therefore, the objective function can be rewritten as

$$\int_0^{\alpha_L^*} \left(v\left(x_L^w\left(\alpha\right)\right) - \alpha\right) dF\left(\alpha\right) + \int_{\alpha_L^*}^{\infty} v\left(x_L^u\left(\alpha\right)\right) dF\left(\alpha\right). \tag{37}$$

The government budget constraint (1) can be formulated as follows:

$$\begin{aligned}
&\gamma\left[\int_0^{\alpha_L^*} \left(w_L - x_L^w\left(\alpha\right)\right) dF(\alpha) - \int_{\alpha_L^*}^{\alpha_H^*} x_L^u\left(\alpha\right) dF(\alpha)\right] \\
&+(1-\gamma)\left[\int_0^{\alpha_L^*} \left(w_H - x_L^w(\alpha)\right) dF(\alpha) + \int_{\alpha_L^*}^{\alpha_H^*} \left(w_H - v^{-1}\left(v\left(x_L^u\left(\alpha\right) + \alpha\right)\right)\right) dF(\alpha)\right] \\
&-\int_{\alpha_H^*}^{\infty} x_L^u\left(\alpha\right) dF(\alpha) \geq R.
\end{aligned}$$

Forming the Lagrangian with objective function (37), the previous government budget constraint and the Lagrangian multiplier λ, the first-order conditions with respect to $x_L^w\left(\alpha\right)$ and $x_L^u\left(\alpha\right)$ are:

$$\begin{aligned}
&\alpha \leq \alpha_L^* : v'\left(x_L^w\left(\alpha\right)\right) = \lambda, \\
&\alpha_H^* < \alpha : v'(x_L^u\left(\alpha\right)) = \lambda, \\
&\alpha_L^* < \alpha \leq \alpha_H^* : v'(x_L^u\left(\alpha\right)) = \lambda\left[\gamma + (1-\gamma)\frac{v'\left(x_L^u\left(\alpha\right)\right)}{v'\left(x_H^w\left(\alpha\right)\right)}\right].
\end{aligned}$$

From the first and second first-order conditions, we have (since λ is constant):

$$\forall \alpha \in [0, \alpha_L^*) \cup [\alpha_H^*, \infty) : x_L^w\left(\alpha\right) = x_L^u\left(\alpha\right) = \overline{x}.$$

For $\alpha_L^* < \alpha \leq \alpha_H^*$, from (35), it follows that $x_L^u\left(\alpha\right) < x_H^w\left(\alpha\right)$ and so $v'\left(x_L^u\left(\alpha\right)\right) > v'\left(x_H^w\left(\alpha\right)\right)$, such that $v'(x_L^u\left(\alpha\right)) > \lambda$ and

$$\forall \alpha \in [\alpha_L^*, \alpha_H^*) : x_L^u\left(\alpha\right) < \overline{x}.$$

(f) Van de gaer planner:

In the first best, there is no reason for having different values for opportunity sets of different skill-types. For the same reasons as usual, if anybody works, it will be those with a low disutility of work. Hence the objective function reduces to:

$$\int_0^{\alpha_L^*} \left[v(x_L^w\left(\alpha\right)) - \alpha\right] dF\left(\alpha\right) + \int_{\alpha_L^*}^{\infty} v\left(x_L^u\left(\alpha\right)\right) dF\left(\alpha\right). \tag{38}$$

This objective function must be maximized subject to two constraints. The first is that both opportunity sets must have the same value:

$$\begin{aligned}
&\int_0^{\alpha_L^*} \left[v(x_L^w\left(\alpha\right)) - \alpha\right] dF\left(\alpha\right) + \int_{\alpha_L^*}^{\infty} v\left(x_L^u\left(\alpha\right)\right) dF\left(\alpha\right) \\
&= \int_0^{\alpha_H^*} \left[v\left(x_H^w\left(\alpha\right)\right) - \alpha\right] dF\left(\alpha\right) + \int_{\alpha_H^*}^{\infty} v\left(x_H^u\left(\alpha\right)\right) dF\left(\alpha\right).
\end{aligned} \tag{39}$$

The second is the budget constraint:

$$\gamma\left[\int_0^{\alpha_L^*}\left(w_L-x_L^w(\alpha)\right)dF(\alpha)-\int_{\alpha_L^*}^{\infty}x_L^u(\alpha)\,dF(\alpha)\right]$$
$$+(1-\gamma)\left[\int_0^{\alpha_H^*}\left(w_H-x_H^w(\alpha)\right)dF(\alpha)-\int_{\alpha_H^*}^{\infty}x_H^u(\alpha)\,dF(\alpha)\right]=R. \tag{40}$$

Forming the Lagrangian with objective function (38), the equality of opportunity set constraint (39) with the associated Lagrangian multiplier μ and the government budget constraint (40) with its Lagrangian multiplier λ, the first-order conditions with respect to $x_L^w(\alpha)$, $x_L^u(\alpha)$, $x_H^w(\alpha)$ and $x_H^u(\alpha)$ are:

$$v'\left(x_L^w(\alpha)\right)(1+\mu)=\lambda\gamma, \tag{41}$$
$$v'\left(x_L^u(\alpha)\right)(1+\mu)=\lambda\gamma, \tag{42}$$
$$-\mu v'\left(x_H^w(\alpha)\right)=\lambda(1-\gamma), \tag{43}$$
$$-\mu v'\left(x_H^u(\alpha)\right)=\lambda(1-\gamma). \tag{44}$$

From (41)-(42) and (43)-(44) respectively, we have:

$$x_L^w(\alpha)=x_L^u(\alpha)=\overline{x}\ \text{ and }\ x_H^w(\alpha)=x_H^u(\alpha)=\overline{\overline{x}}.$$

Substituting these two equations into the equality of opportunity sets constraint (39) gives:

$$v(\overline{x})-\int_0^{\alpha_L^*}\alpha dF(\alpha)=v(\overline{\overline{x}})-\int_0^{\alpha_H^*}\alpha dF(\alpha).$$

If $\alpha_L^*=\alpha_H^*$, then $\overline{x}=\overline{\overline{x}}$. However, such a situation cannot be optimal, as high-skilled workers contribute more to the government budget than low-skilled workers. Therefore, $\alpha_L^*<\alpha_H^*$ hence $\overline{x}<\overline{\overline{x}}$.

LEMMA A.

Lemma A: for an allocation that satisfies EWEP and ETES, there cannot exist an $\alpha\in\mathbb{R}^+$: $\delta_L(\alpha)\neq\delta_H(\alpha)$.

Proof. If such an α existed, we would have by EWEP that for this value either $v\left(x_L^u(\alpha)\right)=v\left(x_H^w(\alpha)\right)-\alpha$ or $v\left(x_L^w(\alpha)\right)-\alpha=v\left(x_H^u(\alpha)\right)$, both of which are impossible since by ETES the consumption bundles cannot depend on α.

PROOF OF THEOREM 1b.

(a) FEO allocation: in view of Lemma A, we have that for all $\alpha:\delta_L(\alpha)=\delta_H(\alpha)$. Suppose there exists an allocation satisfying EWEP and ETES in which some people work and others do not work. From ETES we know that all low-skilled in work have to get the same consumption bundle, which with some abuse of notation we denote as x_L^w. Similarly, all high-skilled in work get the same consumption bundle, denoted as x_H^w. In addition, by ETES, we need (i) $x_L^w-w_L=x_L^u$ and (ii) $x_H^w-w_H=x_H^u$. EWEP requires that $x_L^u=x_H^u$. Combining this with (i) and (ii) we get that $x_L^w = w_L - w_H + x_H^w$, which because EWEP requires $x_L^w = x_H^w$, reduces to $w_L = w_H$, which was excluded by assumption. Hence an allocation that satisfies EWEP and ETES cannot have some people working and others not working.

It is easy to verify that the allocations (i) and (ii) satisfy both axioms. Their consumption bundles follow from the government budget constraint (1).

(b) CE allocation: a first thing to note is that for the allocation to equalize $u\left(x_Y(\alpha), \delta_Y(\alpha), \widetilde{\alpha}\right)$ for all α and $Y = L, H$ requires that utility is independent of w_Y. This has the following implications:

i) for all α such that $\delta_L(\alpha) = \delta_H(\alpha) = 1 \Rightarrow x_L^w(\alpha) = x_H^w(\alpha)$. In addition, all those assigned in a job have to get the same level of utility, which implies that their consumption bundle cannot depend on α, and thus $x_L^w = x_L^w(\alpha) = x_H^w(\alpha) = x_H^w$;

ii) for all α such that $\delta_L(\alpha) = \delta_H(\alpha) = 0 \Rightarrow x_L^u(\alpha) = x_H^u(\alpha)$. In addition, all those that are inactive have to get the same level of utility, implying that their consumption bundle cannot depend on α, such that $x_L^u(\alpha) = x_H^u(\alpha) = x^u$;

iii) for all α such that $\delta_L(\alpha) = 1$ and $\delta_H(\alpha) = 0 \Rightarrow x_L^w(\alpha) = v^{-1}\left(v\left(x_H^u(\alpha)\right) + \widetilde{\alpha}\right)$, which combined with case (i) and (ii) gives $x_L^w = v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)$;

iv) for all α such that $\delta_L(\alpha) = 0$ and $\delta_H(\alpha) = 1 \Rightarrow x_H^w(\alpha) = v^{-1}\left(v\left(x_L^u(\alpha)\right) + \widetilde{\alpha}\right)$, which combined with case 1 and 2 gives $x_H^w = v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)$.

Combining these results, we get

$$x_L^w = x_H^w = v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right).$$

Everybody gets the same level of utility $v\left(x^u\right)$ in the optimum, and so the problem of the first best allocation then amounts to maximize the equal utility level $v\left(x^u\right)$ with respect to x^u, n_L and n_H subject to the budget constraint

$$\begin{aligned} R \leq{}& \gamma\left(w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right) n_L - \gamma x^u\left[1 - n_L\right] \\ &+ (1-\gamma)\left(w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right) n_H - (1-\gamma) x^u\left[1 - n_H\right]. \end{aligned}$$

The Lagrangian for this problem is

$L = v\left(x^u\right) + \lambda[\gamma\left(w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right) n_L - \gamma x^u\left[1 - n_L\right]$
$+ (1-\gamma)\left(w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right) n_H - (1-\gamma) x^u\left[1 - n_H\right] - R]$.

Taking derivatives, we get :

$\frac{\partial L}{\partial x^u} = v'\left(x^u\right) - \lambda\gamma\frac{\partial v^{-1}\left(v\left(x^u\right)+\widetilde{\alpha}\right)}{\partial x^u} n_L - \lambda(1-\gamma)\frac{\partial v^{-1}\left(v\left(x^u\right)+\widetilde{\alpha}\right)}{\partial x^u} n_H$
$-\lambda\gamma\left[1 - n_L\right] - \lambda(1-\gamma)\left[1 - n_H\right] = 0,$

$\frac{\partial L}{\partial n_L} = \lambda\gamma\left[w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right] + \lambda\gamma x^u = \lambda\gamma\left[x^u + \left[w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right]\right],$

$\frac{\partial L}{\partial n_H} = \lambda(1-\gamma)\left[w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right] + \lambda(1-\gamma) x^u$
$= \lambda(1-\gamma)\left[x^u + \left[w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right]\right].$

The two last first-order derivatives cannot possibly both be equal to zero at the same time:

$w_H > w_L \Rightarrow w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right) > w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)$
$\Rightarrow x^u + \left[w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right] > x^u + \left[w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right].$

Hence we either have that

(i) $\frac{\partial L}{\partial n_L} > 0 \Rightarrow \frac{\partial L}{\partial n_H} > 0$, implying that $n_H = 1 = n_L$,

(ii) $-x^u = \left[w_L - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right]$ and $\frac{\partial L}{\partial n_H} > 0$, implying $n_H = 1$ and n_L follows from the budget constraint,

(iii) $\frac{\partial L}{\partial n_H} > 0$ and $\frac{\partial L}{\partial n_L} < 0$, implying that $n_H = 1$ and $n_L = 0$,

(iv)$-x^u = \left[w_H - v^{-1}\left(v\left(x^u\right) + \widetilde{\alpha}\right)\right]$ and $\frac{\partial L}{\partial n_L} < 0$, implying $n_L = 0$ and n_H follows from the budget constraint or

(v) $\frac{\partial L}{\partial n_H} < 0 \Rightarrow \frac{\partial L}{\partial n_L} < 0$, implying that $n_H = 0 = n_L$.

Which of these allocations yields the highest value for $v\left(x^u\right)$ depends on the parameters of the model. If $\widetilde{\alpha}$ is sufficiently low, the optimum will be case (i), as $\widetilde{\alpha}$ rises, we move from (i) to (ii), as it increases further we move to (iii) and (iv) and for values of $\widetilde{\alpha}$ sufficiently high, the optimum will be case (v).

(c) EE allocation: we want everybody to be indifferent between his actual resources (consumption and activity) and a reference resource bundle where he works and gets consumption $\widetilde{x}$. The best thing to do is to give all employed exactly this reference consumption bundle: $x_L^w = x_H^w = \widetilde{x}$. Clearly, to bring the equivalent wage of the inactive with a very high α down can lead to negative consumption levels. To prevent this, we impose that $x_Y^u\left(\alpha\right) \geq 0$. If this constraint is binding, these individuals get an equivalent wage larger than $\widetilde{x}$; we have to give up the ideal of equalizing equivalent incomes. The logical alternative then becomes Fleurbaey and Maniquet's maximin solution.

To get an equivalent wage of exactly $\widetilde{x}$, a person with taste parameter α needs an inactivity transfer equal to $v^{-1}\left(v\left(\widetilde{x}\right) - \alpha\right)$, which is independent of his skill level. Since we maximin the equivalent wages, the transfer for the inactive is $x^u\left(\alpha\right) = \min\left\{v^{-1}\left(v\left(\widetilde{x}\right) - \alpha\right), 0\right\}$. There exists a value for α, say $\widehat{\alpha}$, such that, if $\alpha \leq \widehat{\alpha}$ we have $x^u\left(\alpha\right) = v^{-1}\left(v\left(\widetilde{x}\right) - \alpha\right) \geq 0$, and if $\alpha > \widehat{\alpha}$, $x^u\left(\alpha\right) = 0$. In both cases, $x^u\left(\alpha\right) \leq \widetilde{x}$ such that it is cheaper to have people inactive than to have them working.

However, working people produce w_L or w_H, while inactive people produce nothing. As a consequence, it can never be optimal to have people inactive for which $\alpha \leq \widehat{\alpha}$: they cost $v^{-1}\left(v\left(\widetilde{x}\right) - \alpha\right) \geq 0$, but produce nothing. The best policy that maximizes S^{EE} under budget constraint is therefore $x_L^w = x_H^w = \gamma w_L + \left(1-\gamma\right) w_H - R$, $x^u = 0$ and $\alpha_L^* = \alpha_H^* = v\left(\gamma w_L + \left(1-\gamma\right) w_H\right) - v\left(0\right)$.

Appendix B: Proofs of section 5.1.

PROOF OF LEMMA 1.

(a) Suppose the proposition does not hold true. By (13), we then have that $\alpha_H^* > \alpha_L^*$. In that case, there exist α, $\alpha_L^* < \alpha < \alpha_H^*$ for which highly skilled workers get utility $v\left(x_H\right) - \alpha$ and lowly skilled workers get $v\left(x^u\right)$. Since the former depends on α, but the latter doesn't these two can never be equal for all α, $\alpha_L^* < \alpha < \alpha_H^*$, and so EWEP must be violated.

(b) Follows immediately from the second best context and the definition of ETES.

PROOF OF SOCIAL OBJECTIVE FUNCTIONS IN SECOND BEST.

Parts (a), (b), (c) and (d) are straightforward to prove.

To see part (e), observe that (11) (due to incentive constraints) implies that for all α, $v(x_L)-\alpha \leq v(x_H)-\alpha$. Therefore, Roemer's objective function

$$\int_0^\infty \min\{\underset{\delta_L(\alpha)}{\operatorname{oper}}\{v(x_L)-\alpha, v(x^u)\}, \underset{\delta_H(\alpha)}{\operatorname{oper}}\{v(x_H)-\alpha, v(x^u)\}\}dF(\alpha)$$

becomes

$$\int_0^{\alpha_L^*} (v(x_L)-\alpha)\, dF(\alpha) + \int_{\alpha_L^*}^\infty v(x^u)\, dF(\alpha). \tag{45}$$

To see part (f), note that, in second best, Van de gaer's objective function is

$$\min\{\int_0^\infty \underset{\delta_L(\alpha)}{\operatorname{oper}}\{v(x_L)-\alpha, v(x^u)\}dF(\alpha), \int_0^\infty \underset{\delta_H(\alpha)}{\operatorname{oper}}\{v(x_H)-\alpha, v(x^u)\}dF(\alpha)\}.$$

Due to the incentive constraints, this reduces to (45).

To see part (g), observe that, since the policy can no longer depend on α, (8) reduces to

$$\widetilde{S}^C = \min\{u(x_L(\delta_L), \delta_L, \widetilde{\alpha}), u(x_H(\delta_H), \delta_H, \widetilde{\alpha})\},$$

where, for $Y = L$ or H, $\delta_Y = 1$ or 0 and $x_Y(\delta_Y) = x_Y$ if $\delta_Y = 1$ and $x_Y(\delta_Y) = x^u$ if $\delta_Y = 0$. However, since (11) holds true, the first element in the set behind the min sign is always the smallest; the low skilled will always be the worst off and

$$\widetilde{S}^C = \min\{v(x_L)-\widetilde{\alpha}, v(x^u)\}. \tag{46}$$

If maximization of $v(x_L)-\widetilde{\alpha}$ yields a value $\alpha_L^* > \widetilde{\alpha}$, then $v(x_L)-\widetilde{\alpha} > v(x_L)-\alpha_L^* = v(x^u)$, and so objective function (46) was not maximized. To prevent this from occurring, we maximize $v(x_L)-\widetilde{\alpha}$ subject to the constraint that $\widetilde{\alpha} \geq \alpha_L^*$. The multiplier associated to this constraint is denoted by ξ.

To see part (h), note that the equivalent wages for the employed are equal to x_Y ($Y = H$ or L) and for the inactive $v^{-1}(v(x^u)+\alpha)$. The objective is to maximize the lowest equivalent wage. Consider the inactive. Since $v^{-1}(.)$ is an increasing function, the equivalent wage is lowest for those inactive having the lowest value for α; which are those with $\alpha = \alpha_L^*$. Hence the lowest value for the equivalent wage is $v^{-1}(v(x^u)+\alpha_L^*) = v^{-1}(v(x_L)) = x_L$.

Appendix C: Proofs of Section 5.2.

PROOF OF THEOREM 2.

The Lagrangian function for the general second best problem is

$$\begin{aligned}
&\pounds = \widetilde{S}^X(x_L, x_H, x^u, \alpha_L^*, \alpha_H^*, \lambda, \mu_L, \mu_H, \nu, c) \\
&+\lambda\{\gamma(w_L - x_L)F(\alpha_L^*) - \gamma x^u(1-F(\alpha_L^*)) \\
&+(1-\gamma)(w_H - x_H)F(\alpha_H^*) - (1-\gamma)x^u(1-F(\alpha_H^*)) - R\} \\
&+\mu_L[v(x_L) - \alpha_L^* - v(x^u)] \\
&+\mu_H[v(x_H) - \alpha_H^* - v(x^u)] \\
&+\nu(x_H - x_L - c),
\end{aligned}$$

which has to be maximized with respect to x_L, x_H, x^u, α_L^*, α_H^* and c, taking into account that $c \geq 0$. This leads to the following first-order conditions:

$$\widetilde{S}_{x_L}^X - \lambda\gamma F(\alpha_L^*) - \nu = -\mu_L v'(x_L), \tag{47}$$

$$\widetilde{S}_{x^u}^X - \lambda[\gamma(1-F(\alpha_L^*)) - (1-\gamma)(1-F(\alpha_H^*))] = (\mu_L + \mu_H)v'(x^u), \tag{48}$$

$$\widetilde{S}_{x_H}^X - \lambda(1-\gamma)F(\alpha_H^*) + \nu = -\mu_H v'(x_H), \tag{49}$$

$$\widetilde{S}_{\alpha_L^*}^X + \lambda\gamma f(\alpha_L^*)(w_L - x_L + x^u) = \mu_L, \tag{50}$$

$$\widetilde{S}_{\alpha_H^*}^X + \lambda(1-\gamma) f(\alpha_H^*)(w_H - x_H + x^u) = \mu_H, \tag{51}$$

$$-\nu \leq 0 \text{ and } \nu c = 0. \tag{52}$$

Solving (47) for μ_L and equating the resulting expression to the LHS of (50), we obtain

$$\begin{aligned} & \widetilde{S}_{\alpha_L^*}^X + \lambda\gamma f(\alpha_L^*)(w_L - x_L + x^u) \\ = & \frac{\lambda\gamma F(\alpha_L^*)}{v'(x_L)} - \frac{\widetilde{S}_{x_L}^X}{v'(x_L)} + \frac{\nu}{v'(x_L)}, \end{aligned}$$

from which

$$w_L - x_L + x^u = \frac{F(\alpha_L^*)}{f(\alpha_L^*)v'(x_L)}\left[1 - \frac{\widetilde{S}_{x_L}^X - \nu}{\lambda\gamma F(\alpha_L^*)}\right] - \frac{\widetilde{S}_{\alpha_L^*}^X}{\lambda\gamma f(\alpha_L^*)}.$$

Using the definition (14), we get

$$\frac{w_L - x_L + x^u}{x_L} = \frac{1}{\eta(x_L, \alpha_L^*)}\left[1 - \frac{\widetilde{S}_{x_L}^X - \nu}{\lambda\gamma F(\alpha_L^*)}\right] - \frac{\widetilde{S}_{\alpha_L^*}^X}{\lambda\gamma f(\alpha_L^*)\, x_L}.$$

Similarly, solving (49) for μ_H, equating the resulting expression to the LHS of (51) and using definition (15), we get

$$\frac{w_H - x_H + x^u}{x_H} = \frac{1}{\eta(x_H, \alpha_H^*)}\left[1 - \frac{\widetilde{S}_{x_H}^X + \nu}{\lambda(1-\gamma)F(\alpha_H^*)}\right] - \frac{\widetilde{S}_{\alpha_H^*}^X}{\lambda(1-\gamma) f(\alpha_H^*)\, x_H}.$$

Divide the equations (47)-(49) by the marginal utility on their RHS, adding the resulting equation for (47) and (49) and equating the result to (49) yields

$$\begin{aligned} & \frac{\lambda\gamma F(\alpha_L^*)}{v'(x_L)} - \frac{\widetilde{S}_{x_L}^X}{v'(x_L)} + \frac{\nu}{v'(x_L)} + \frac{\lambda(1-\gamma)F(\alpha_H^*)}{v'(x_H)} - \frac{\widetilde{S}_{x_H}^X}{v'(x_H)} - \frac{\nu}{v'(x_H)} \\ = & \frac{\widetilde{S}_{x^u}^X}{v'(x^u)} - \frac{\lambda[\gamma(1-F(\alpha_L^*)) + (1-\gamma)(1-F(\alpha_H^*))]}{v'(x^u)}. \end{aligned}$$

Collecting the terms in λ gives

$$\begin{aligned} & \lambda\left[\frac{\gamma F(\alpha_L^*)}{v'(x_L)} + \frac{(1-\gamma)F(\alpha_H^*)}{v'(x_H)} + \frac{[\gamma(1-F(\alpha_L^*)) + (1-\gamma)(1-F(\alpha_H^*))]}{v'(x^u)}\right] \\ = & \frac{\widetilde{S}_{x_L}^X}{v'(x_L)} + \frac{\widetilde{S}_{x_H}^X}{v'(x_H)} + \frac{\widetilde{S}_{x^u}^X}{v'(x^u)} + \nu\left[\frac{1}{v'(x_H)} - \frac{1}{v'(x_L)}\right]. \end{aligned}$$

Now, note that from (52), if $\nu > 0$, then $c = 0$, such that $x_H = x_L$ and the last term in the above equation always drops out. Using definitions (16) and (17) gives $\lambda g_P^X = D^X$, and thus $\lambda^{-1} = g_P^X / D^X$.

PROOF OF LEMMA 2.

Follows immediately from partially differentiating the expressions for S^X with respect to α_L^* and α_H^*.

PROOF OF LEMMA 3.

Step 1: we proof the following lemma:

Lemma B: If, evaluated at $x_H = x_L$ and $\alpha_H^* = \alpha_L^*$, $\frac{\widetilde{S}_{x_H}^X + \widetilde{S}_{\alpha_H^*}^X v'(x)}{1-\gamma} = \frac{\widetilde{S}_{x_L}^X + \widetilde{S}_{\alpha_L^*}^X v'(x)}{\gamma}$, then $\nu = 0$.

Proof:

Using (51) in (49) and solving for ν, we obtain

$$\begin{aligned}\nu &= -\widetilde{S}_{x_H}^X - \widetilde{S}_{\alpha_H^*}^X v'(x_H) \\ &\quad +\lambda (1-\gamma) \left[F(\alpha_H^*) - f(\alpha_H^*)(w_H - x_H + x^u) v'(x_H) \right].\end{aligned}$$

Hence, $\nu > 0$ (such that $x_H = x_L = x$ and $\alpha_H^* = \alpha_L^* = \alpha^*$) if and only if

$$F(\alpha^*) - (w_H - x + x^u) f(\alpha^*) v'(x) > \frac{\widetilde{S}_{x_H}^X + \widetilde{S}_{\alpha_H^*}^X v'(x)}{\lambda (1-\gamma)}. \tag{53}$$

Similarly, using (47) in (50) and solving for ν,

$$\begin{aligned}\nu &= \widetilde{S}_{x_L}^X + \widetilde{S}_{\alpha_L^*}^X v'(x_L) \\ &\quad -\lambda\gamma \left[F(\alpha_L^*) - f(\alpha_L^*)(w_L - x_L + x^u) v'(x_L) \right],\end{aligned}$$

we find that $\nu > 0$ if and only if

$$F(\alpha^*) - [w_L - x + x^u] f(\alpha^*) v'(x) < \frac{\widetilde{S}_{x_L}^X + \widetilde{S}_{\alpha_L^*}^X v'(x)}{\lambda\gamma}. \tag{54}$$

If the antecedent of lemma B holds true, the right hand sides of (53) and (54) are equal, such that $\nu > 0$ requires

$$F(\alpha^*) - (w_H - x + x^u) f(\alpha^*) v'(x) > F(\alpha^*) - [w_L - x + x^u] f(\alpha^*) v'(x),$$

but this can only hold true if $w_H < w_L$, which goes against the model's assumptions.

Step 2: we compute the expressions that occur in lemma B. They are given in the following table.

X	$\frac{\widetilde{S}_{x_L}^X}{\gamma}$	$\frac{\widetilde{S}_{x_H}^X}{1-\gamma}$	$\frac{\widetilde{S}_{\alpha_L^*}^X}{\gamma}$	$\frac{\widetilde{S}_{\alpha_H^*}^X}{1-\gamma}$
U	$v'(x) F(\alpha^*)$		0	
W	$v'(x) \int_0^{\alpha^*} \Psi'(v(x) - \alpha) dF(\alpha)$		0	
B	$v'(x) \int_0^{\alpha^*} W(\alpha) dF(\alpha)$		0	
$R = V$	$v'(x) F(\alpha^*)$	0	0	
EE	$1/\gamma$	0	0	
CE[20]	$v'(x)$	0	0	
NW	$v'(x) F(\alpha^*)$		$[\alpha^* - \overline{\alpha}] f(\alpha^*)$	

[20] Under the assumption that the constraint $\widetilde{\alpha} \geq \alpha^*$ is not binding.

Clearly, for $X = U, W, B$ and NW, by lemma B, $\nu = 0$.

PROOF OF FOOTNOTE 13.

Step 1: we proof the following Lemma:

Lemma C: $x_H^X > x_L^X$ for $X = U, W, B$ and NW.

Proof:

Under $X = U$, W, B and NW, $\nu = 0$ from lemma 3. Assume $x_H = x_L = x$ hence $\alpha_H^* = \alpha_L^* = \alpha^*$. Combining equations (47) and (50) gives:

$$\frac{\widetilde{S}_{x_L}^X}{\gamma} = \lambda F(\alpha^*) - \frac{v'(x)}{\gamma}\left[\widetilde{S}_{\alpha_L^*}^X + \lambda\gamma f(\alpha^*)(w_L - x + x^u)\right].$$

Combining equations (49) and (51) we can write:

$$\frac{\widetilde{S}_{x_H}^X}{1-\gamma} = \lambda F(\alpha^*) - \frac{v'(x)}{1-\gamma}\left[\widetilde{S}_{\alpha_H^*}^X + \lambda(1-\gamma) f(\alpha^*)(w_H - x + x^u)\right].$$

Using $\widetilde{S}_{x_L}^X/\gamma = \widetilde{S}_{x_H}^X/(1-\gamma)$, $\widetilde{S}_{\alpha_L^*}^X/\gamma = \widetilde{S}_{\alpha_H^*}^X/(1-\gamma)$ for $X = U, W, B$ and NW from the previous table, the two previous equations yield $\lambda f(\alpha^*)(w_L - x + x^u) = \lambda f(\alpha^*)(w_H - x + x^u)$ but this can only hold true if $w_H = w_L$, which leads to a contradiction. We can conclude that $x_H > x_L$.

From lemma C, (10) and (12) we have $\alpha_H^* > \alpha_L^*$ under the U, W, B and NW criteria.

Step 2: in second best, α_H^*, $\alpha_L^* < \infty$.

Proof. As $\forall\alpha : f(\alpha) > 0$, all low-ability (high-ability) people work means $\alpha_L^* \to \infty$ ($\alpha_H^* \to \infty$) at the optimum. Since consumption levels are finite, from (10) and (resp. (12)), α_L^* and α_H^* cannot tend to ∞.

Step 3: $\alpha_L^* > 0$ when $\nu = 0$.

Proof. Suppose $\alpha_L^* = 0$. From (10), evaluated at $\alpha_L^* = 0$, we have $x_L = x^u$. Since $\nu = 0$ and $F(0) = 0$, from first-order condition (47), $\mu_L = -\widetilde{S}_{x_L}^X/v'(x^u)$. The value $\alpha_L^* = 0$ can only be optimal if $\partial\pounds/\partial\alpha_L^*|_{\alpha_L^*=0} \leq 0$, which requires, using the previous results

$$\lambda\gamma f(0) w_L \leq -\widetilde{S}_{\alpha_L^*}^X - \widetilde{S}_{x_L}^X/v'(x^u),$$

Going back to the table in this appendix, it is clear that for all the criteria the right hand side is negative, such that $\alpha_L^* = 0$ can only be optimal if $w_L < 0$, which, however, was excluded by assumption.

Step 4: to complete the proof, note that we have shown that, for the U, W, B, and NW criterion, $\nu = 0$, such that $x_H > x_L$ (using lemma C) and thus $\alpha_H^* > \alpha_L^*$. For $X = R$, EE and CE, we have shown that $\nu \geq 0$, such that $\alpha_H^* \geq \alpha_L^*$.

PROOF OF COROLLARY 3.

The proof is obvious from the table in corollary 2.

PROOF OF COROLLARY 4.

Welfarist optimum.

Since $x_H > x_L$, $v(x_L)-\alpha < v(x_H)-\alpha$ and since $\Psi'' < 0$, $\Psi'(v(x_L)-\alpha_1) > \Psi'(v(x_H)-\alpha_1) > \Psi'(v(x_H)-\alpha_2)$ when $\alpha_2 > \alpha_1$, such that $g_H^W < g_L^W$. Combined with $\eta(x_L,\alpha_L^*) \geq \eta(x_H,\alpha_H^*)$, it follows from the expressions in theorem 2, that $(T_L - T_u)/x_L < (T_H - T_u)/x_H$.

Boadway et al. optimum.

Note that

$$\frac{\int_0^{\alpha_L^*} W(\alpha)\,dF(\alpha)}{F(\alpha_L^*)} \geq (\leq) \frac{\int_0^{\alpha_H^*} W(\alpha)\,dF(\alpha)}{F(\alpha_H^*)} \Leftrightarrow$$

$$\frac{\int_0^{\alpha_L^*} W(\alpha)\,dF(\alpha)}{F(\alpha_L^*)} \geq (\leq) \frac{\int_0^{\alpha_L^*} W(\alpha)\,dF(\alpha)}{F(\alpha_L^*)}\frac{F(\alpha_L^*)}{F(\alpha_H^*)} + \frac{\int_{\alpha_L^*}^{\alpha_H^*} W(\alpha)\,dF(\alpha)}{F(\alpha_H^*)} \Leftrightarrow$$

$$\frac{\int_0^{\alpha_L^*} W(\alpha)\,dF(\alpha)}{F(\alpha_L^*)} \geq (\leq) \frac{\int_{\alpha_L^*}^{\alpha_H^*} W(\alpha)\,dF(\alpha)}{F(\alpha_H^*) - F(\alpha_L^*)},$$

which holds as $\geq$ automatically if $W(\alpha)$ is a decreasing function, and as $\leq$ if $W(\alpha)$ is an increasing function.

Therefore, assume that $W(\alpha)$ is a decreasing function hence $g_L^B > g_H^B$. Since $x_H > x_L$, such that $v'(x_H) < v'(x_L)$, and the assumption that $\eta(x_L,\alpha_L^*) \geq \eta(x_H,\alpha_H^*)$, it follows from the expressions in theorem 2, that $(T_L - T_u)/x_L < (T_H - T_u)/x_H$.

Roemer, EE and CE (when $\xi < 1$).

There are two cases to consider:

(i) When $\nu = 0$, the proof is straightforward from the table in corollary 2, $\eta(x_L,\alpha_L^*) \geq \eta(x_H,\alpha_H^*)$ and $g_L^X > 0$.

(ii) When $\nu > 0$, $x_H = x_L = x$ hence, $T_Y = w_Y - x$ $(Y = L, H)$ which, combined with $w_H > w_L$, yields the inequality $(T_L - T_u)/x_L < (T_H - T_u)/x_H$ (where $x_H = x_L$).[21]

PROOF OF COROLLARY 5.

By definition, $\frac{T_L - T_u}{x_L} < \frac{T_H - T_u}{x_H} \Leftrightarrow \frac{w_L - x_L + x^u}{x_L} < \frac{w_H - x_H + x^u}{x_H}$. Therefore under assumption 1, from corollary 4, we have that for the planners considered in the corollary, $x_H(w_L - x_L + x^u) < x_L(w_H - x_H + x^u)$. Since $x_H \geq x_L$ (from (11)), we have: $w_L - x_L + x^u < w_H - x_H + x^u$.

Appendix D: Proofs of Section 5.3

PROOF OF THEOREM 3.

From the government budget constraint, we have that

$$\gamma \int_0^{\alpha_L^*} (w_L - x_L)\,dF(\alpha) + \gamma \int_{\alpha_L^*}^{\infty} x^u dF(\alpha)$$
$$+ (1-\gamma) \int_0^{\alpha_H^*} (w_H - x_H)\,dF(\alpha) + (1-\gamma) \int_{\alpha_H^*}^{\infty} x^u dF(\alpha) = R.$$

Substituting the ETES constraints $w_L - x_L = -x^u$ and $w_H - x_H = -x^u$ and rearranging gives the first expression in the lemma. The second and third expression follow from (10) and (12), the definitions of the critical values a_L^* and α_H^*.

[21] Note that using $x_L = x_H$ and $\alpha_L^* = \alpha_H^*$ into (14)-(15) yields $\eta(x_L,\alpha_L^*) = \eta(x_L,\alpha_L^*)$.

PROOF OF THEOREM 4.

The Lagrangian is

$$
\begin{aligned}
&\pounds = \widehat{S}^X \left(x^w, x^u, \alpha^*, \lambda, \mu\right) \\
&+\lambda \left\{\left[\gamma w_L + (1-\gamma) w_H - x^w\right] F\left(\alpha^*\right) - x^u \left(1 - F\left(\alpha^*\right)\right) - R\right\} \\
&+\mu \left[v\left(x^w\right) - \alpha^* - v\left(x^u\right)\right].
\end{aligned}
$$

The first-order conditions are

$$
\widehat{S}^X_{x^w} - \lambda F\left(\alpha^*\right) = -\mu v'\left(x^w\right), \tag{55}
$$

$$
\widehat{S}^X_{x^u} - \lambda \left(1 - F\left(\alpha^*\right)\right) = \mu v'\left(x^u\right), \tag{56}
$$

$$
\widehat{S}^X_{\alpha^*} + \lambda \left[\gamma w_L + (1-\gamma) w_H - x^w + x^u\right] f\left(\alpha^*\right) = \mu. \tag{57}
$$

Combining (55) and (57),

$$
\gamma w_L + (1-\gamma) w_H - x^w + x^u = \frac{F\left(\alpha^*\right)}{v'\left(x^w\right) f\left(\alpha^*\right)} \left[1 - \frac{\widehat{S}^X_{x^w}}{\lambda F\left(\alpha^*\right)}\right] - \frac{\widehat{S}^X_{\alpha^*}}{\lambda f\left(\alpha^*\right)},
$$

which, after using (19) yields

$$
\frac{\gamma w_L + (1-\gamma) w_H - x^w + x^u}{x^w} = \frac{1}{\eta\left(x^w, \alpha^*\right)} \left[1 - \frac{\widehat{S}^X_{x^w}}{\lambda F\left(\alpha^*\right)}\right] - \frac{\widehat{S}^X_{\alpha^*}}{\lambda f\left(\alpha^*\right) x^w}.
$$

Dividing equations (55)-(56) by the marginal utilities on the right hand side and adding, we obtain

$$
\lambda \left[\frac{F\left(\alpha^*\right)}{v'\left(x^w\right)} + \frac{1 - F\left(\alpha^*\right)}{v'\left(x^u\right)}\right] = \frac{\widehat{S}^X_{x^u}}{v'\left(x^u\right)} + \frac{\widehat{S}^X_{x^w}}{v'\left(x^w\right)},
$$

from which, using definitions (20) and (21), we get $\lambda g^X = D^X$, and so $\lambda^{-1} = g^X / D^X$.

Appendix E: Proofs of Section 5.4

PROOF OF LEMMA 4.

(a) Proof for $\widetilde{S}^{PWU}$, $\widetilde{S}^{PTU}$, $\widetilde{S}^{PWE}$ and $\widetilde{S}^{PTU}$.

Observe that in the second best, for $\alpha < \alpha^*_L \le \alpha^*_H$, $\delta_L(\alpha) = \delta_H(\alpha) = 1$, $x^w_L(\alpha) = x_L$, $x^w_H(\alpha) = x_H$ and that $x_L \le x_H$. For $\alpha^*_L \le \alpha \le \alpha^*_H$, $\delta_L(\alpha) = 0$, and $\delta_H(\alpha) = 1$ and by (12), $v(x^u) = v(x_H) - \alpha^*_H$, which for $\alpha^*_L \le \alpha \le \alpha^*_H$ gives $v(x^u) \le v(x_H) - \alpha$. For $\alpha > \alpha^*_H$, $\delta_L(\alpha) = \delta_H(\alpha) = 0$, and $x^u_L(\alpha) = x^u_H(\alpha) = x^u$.

Substituting these properties into S^{PWU} and S^{PTU} yields $\widetilde{S}^{PWU}$ and $\widetilde{S}^{PTU}$, respectively. Substituting these properties into S^{PWE} and S^{PTE} leads to $\widetilde{S}^{PWE}$ and $\widetilde{S}^{PTE}$. In the procedure, for $\alpha^*_L \le \alpha \le \alpha^*_H$ we use $v^{-1}(v(x^u) + \alpha) \le x_H$ from $v(x^u) \le v(x_H) - \alpha$.

(b) Proof for $\widetilde{S}^{PT}$.

Since consumption levels do not depend on α in the second best, S^{PT} reduces to

$$
\rho \min\left\{x_L - w_L, x^u\right\} + (1-\rho) \min\left\{x_H - w_H, x^u\right\}.
$$

Hence, with the ETES priority principle, the Lagrangian is

$$\begin{aligned}
&\pounds\left(x_L, x_H, x^u, \alpha_L^*, \alpha_H^*, \lambda, \mu_L, \mu_H, \nu\right) = \rho \min\left\{x_L - w_L, x^u\right\} + (1-\rho)\min\left\{x_H - w_H, x^u\right\} \\
&+\lambda\left\{\gamma F(\alpha_L^*)\left(w_L - x_L\right) + (1-\gamma)\, F(\alpha_H^*)\left(w_H - x_H\right)\right. \\
&\left.-\left[\gamma\left(1 - F(\alpha_L^*)\right) + (1-\gamma)\left(1 - F(\alpha_H^*)\right)\right] x^u - R\right\} \\
&+\mu_H\left[v(x_H) - \alpha_H^* - v\left(x^u\right)\right] + \mu_L\left[v(x_L) - \alpha_L^* - v\left(x^u\right)\right] + \nu\left(x_H - x_L - c\right)
\end{aligned}$$

with $0 \leq \rho \leq 1$.

(i) Suppose $x_L - w_L \geq x^u$. The first-order condition of the Lagrangian with respect to x_L then becomes $-\lambda\gamma F\left(\alpha_L^*\right) - \nu = -\mu_L v'\left(x_L\right)$, from which $\mu_L > 0$. However, the first-order condition with respect to α_L^* gives $\mu_L = \lambda\gamma f\left(\alpha_L^*\right)\left(w_L - x_L + x^u\right) \leq 0$ under the assumption made. Hence we obtain a contradiction, such that we know that $x_L - w_L < x^u$.

(ii) Suppose $x_H - w_H \geq x^u$. Then we get $-T_H \geq x^u$; the high-skilled workers receive a larger subsidy than the inactive people which cannot not be optimal. Consequently, $x_H - w_H < x^u$.

As a result of (i) and (ii), the ETES priority principle reduces to $\rho\left(x_L - w_L\right) + (1-\rho)\left(x_H - w_H\right)$.

PROOF OF LEMMA 5.

That for all objective functions $\widetilde{S}^X_{\alpha_H^*} = 0$ and that $\widetilde{S}^{PT}_{\alpha_H^*} = \widetilde{S}^{PT}_{\alpha_L^*} = 0$ is evident. Simple differentiation yields $\widetilde{S}^{PWU}_{\alpha_L^*} = \left[\Omega^R\left(v\left(x_L\right) - \alpha_L^*\right) - \Omega^R\left(v\left(x^u\right)\right)\right]$. Due to (10), $v\left(x_L\right) - \alpha_L^* = v\left(x^u\right)$, and so $\widetilde{S}^{PWU}_{\alpha_L^*} = 0$. Similarly it can be shown that $\widetilde{S}^X_{\alpha_L^*} = 0$ for $X = PTU$, PWE and PTE.

PROOF OF LEMMA 6.

The proof follows the reasoning for lemma 3 (using lemma 5) so is skipped here.