Review of Economic Studies (2006) **73**, 55–83 0034-6527/06/00030055$02.00

Fair Income Tax

MARC FLEURBAEY
CNRS-CERSES, Paris and IDEP

and

FRANÇOIS MANIQUET
University of Louvain-La-Neuve and CORE

First version received February 2003; *final version accepted February* 2005 (*Eds.*)

In a model where agents have unequal skills and heterogeneous preferences over consumption and leisure, we look for the optimal tax on the basis of efficiency and fairness principles and under incentive-compatibility constraints. The fairness principles considered here are: (1) a weak version of the Pigou–Dalton transfer principle; (2) a condition precluding redistribution when all agents have the same skills. With such principles we construct and justify specific social preferences and derive a simple criterion for the evaluation of income tax schedules. Namely, the lower the greatest average tax rate over the range of low incomes, the better. We show that, as a consequence, the optimal tax should give the greatest subsidies to the working poor (the agents having the lowest skill and choosing the largest labour time).

1. INTRODUCTION

Fairness is a key concept in redistribution issues. In this paper, we study how particular requirements of fairness can shed light on the design of the optimal income tax schedule.

We consider a population of heterogeneous individuals (or households), who differ in two respects. First, they have unequal skills and, therefore, unequal earning abilities. Second, they differ in terms of their preferences about consumption and leisure and, as a consequence, typically make different labour time choices. Both kinds of differences generate income inequalities. We study how to justify and compute a redistribution income tax in this context.

Redistribution through an income tax usually entails distortions of incentives, but the resulting efficiency loss has to be weighed against potential improvements in the fairness of the distribution of resources. We address this efficiency–equity trade-off here by constructing social preferences which obey the standard Pareto principle in addition to fairness conditions.

Two fairness requirements are introduced below. Briefly, the first requirement, a qualification of the Pigou–Dalton principle, states that transfers reducing income inequalities are acceptable, *provided* they are performed between agents having identical preferences and identical labour time. Thanks to this proviso, this requirement (contrary to the usual Pigou–Dalton transfer principle which applies to all income inequalities) is still justified if we consider that incomes should not necessarily be equalized among agents having different labour time or, more generally, different willingness to work. The second fairness requirement is that the laisser-faire (that is, the absence of redistribution) should be the social optimum in the hypothetical case when all agents have equal earning abilities. The underlying idea is that income inequalities would then reflect free choices from different preferences on an identical budget set, and that such choices ought to be respected.

These two requirements, together with the Pareto principle and ancillary conditions of informational parsimony and separability (the idea that indifferent agents should not influence social preferences), lead us to single out a particular kind of social preferences. These social preferences measure individual well-being in terms of what we call "equivalent wage" (see Section 2). For any given individual, her equivalent wage, relative to a particular indifference curve, is the hypothetical wage rate which would enable her to reach this indifference curve if she could freely choose her labour time at this wage rate. This particular measure of well-being, which is induced by the fairness conditions, does not require any other information about individuals than their ordinal non-comparable preferences about their own consumption–leisure bundles.

It is then shown that, under some richness assumptions about the distribution of characteristics in the population, such social preferences yield a very simple criterion for the welfare comparison of tax schedules. This criterion is the maximal average tax rate over low incomes (*i.e.* incomes below the minimum wage). This criterion can be used for the comparison of any pair of tax schedules, no matter how far from the optimum, but it can also be used to seek the optimal tax schedule. As far as the optimal tax is concerned, the main result is that those individuals who have the lowest earning ability but work full time, namely, the *hardworking poor*, will be granted the greatest subsidy (*i.e.* the smallest tax) of the whole population.

The literature on optimal taxation has focused mostly on social objectives defined in terms of welfarist (typically, utilitarian) social welfare functions, based on interpersonal comparisons of utility. It has obtained valuable insights into the likely shape of the optimal tax, as can be grasped from the outstanding works of Mirrlees (1971), Atkinson (1973, 1995), Sadka (1976), Seade (1977) Tuomala (1990), Ebert (1992), and Diamond (1998), among many others. Many results depend on the particular choice of individual utility function and social welfare function. The social marginal utility of an individual's income may thus reflect various personal characteristics (individual utility) and ethical values embodied in the social welfare function, including, potentially, fairness requirements. But, apart from the important relationship between inequality aversion and (Schur-)concavity of the social welfare function, the link between fairness requirements and features of the social welfare function are not usually made explicit. In contrast, our approach starts from requirements of fairness, and derives social preferences on this basis.

This literature has traditionally assumed that agents differ only in one dimension (typically, their earning ability). Several authors (Choné and Laroque, 2001, Boadway, Marchand, Pestieau and Racionero, 2002) have recently examined optimal taxation under the assumption that agents may be heterogeneous in two dimensions, their consumption–leisure preferences and their earning ability, or skill. They immediately face a conceptual difficulty: there is no clear way to define the objective of a utilitarian planner, as summing utility levels of agents having different preferences requires a particular choice of utility functions. It seems therefore necessary to impose what Choné and Laroque (2001) appropriately call an ethical assumption. Boadway *et al*. (2002) consider a whole span of possible weights for various utility functions. In this paper we show that the relative weight of agents having different preferences does not need to be determined by assumption, but can be derived from fairness conditions. An additional notorious difficulty of multi-dimensional screening is the impossibility to derive simple solutions due to widespread bunching.[1] We are however able to describe some basic features of the optimal tax and to obtain a simple criterion for the comparison of taxes.

This recent literature suggests that, with double heterogeneity, negative marginal income tax rates are more likely to be obtained than if agents differ with respect to one parameter only. Our results go in the same direction. In Choné and Laroque (2001), however, the focus is on labour participation, so that agents work either zero or one unit, whereas we consider the whole interval. In addition, their social objective gives absolute priority to agents with the smallest income, so that negative tax rates may obtain for high incomes (and only for special distributions), whereas

1. See, *e.g.* Armstrong (1996) or Rochet and Choné (1998).

our social objective gives priority to the working poor, and non-positive tax rates are obtained on low incomes (for all distributions). In Boadway *et al*. (2002), negative marginal rates are obtained on low incomes and in a closer way to ours, since they arise in the case when the weights assigned to agents with a high aversion to work are lower than those assigned to agents with a low aversion to work. But their framework has only four types of agents, whereas our result is obtained for an unlimited domain.[2]

Our work also builds on previous studies of the same model (with unequal earning abilities and heterogeneous preferences) which dealt with first-best allocations (Fleurbaey and Maniquet, 1996*a*, 1999) or with linear tax (Bossert, Fleurbaey and Van de Gaer, 1999), or focused on different fairness concepts (Fleurbaey and Maniquet, 2005).

The paper is organized as follows. Section 2 introduces the model and the concept of social preferences. Section 3 contains the axiomatic analysis and derives social preferences. Section 4 develops the analysis of taxation. Concluding remarks are offered in the last section.

2. THE MODEL

There are two goods, labour and consumption.[3] A *bundle* for agent i is a pair $z_i = (\ell_i, c_i)$, where ℓ_i is labour and c_i consumption. The agents' *consumption set* X is defined by the conditions $0 \leq \ell_i \leq 1$ and $c_i \geq 0$.

The population contains $n \geq 2$ agents. Agents have two characteristics, their personal preferences over the consumption set and their personal skill. For any agent $i = 1, \ldots, n$, *personal preferences* are denoted R_i, and $z_i R_i z_i'$ (resp. $z_i P_i z_i'$, $z_i I_i z_i'$) means that bundle z_i is weakly preferred (resp. strictly preferred, indifferent) to bundle z_i'. We assume that individual preferences are continuous, convex and monotonic.[4]

The marginal productivity of labour is assumed to be fixed, as in a constant returns to scale technology. Agent i's earning ability is measured by her productivity or *wage rate*, denoted w_i, and is measured in consumption units, so that $w_i \geq 0$ is agent i's production when working $\ell_i = 1$ and, for any ℓ_i, $w_i\ell_i$ is the agent's pre-tax income (earnings). Figure 1 displays the consumption set, with typical indifference curves, and earnings as a function of labour time. As illustrated on the figure, an agent's consumption c_i may differ from her earnings $w_i\ell_i$. This is a typical consequence of redistribution.

An allocation is a collection $z = (z_1, \ldots, z_n)$. Social preferences will allow us to compare allocations in terms of fairness and efficiency. Social preferences will be formalized as a complete ordering over all allocations in X^n, and will be denoted R, with asymmetric and symmetric components P and I, respectively. In other words, $z\ R\ z'$ means that z is at least as good as z', $z\ P\ z'$ means that it is strictly better, and $z\ I\ z'$ that they are equivalent.

Social preferences may depend on the population profile of characteristics $(R_1, \ldots, R_n)$ and $(w_1, \ldots, w_n)$. Formally, they are a *mapping* from the set of population profiles to the set of complete orderings over allocations. For the sake of simplicity, we do not introduce additional no-

2. Another branch of the literature sometimes obtains similar results by studying social objectives disregarding individual leisure–consumption preferences and focusing on income maintenance. See Besley and Coate (1995) for a synthesis. Here we retain a concern for efficiency via the Pareto principle, so that the social preferences obtained respect individual preferences.

3. Introducing several consumption goods would not change the analysis much if prices were assumed to be fixed. The case of variable consumption prices would require a specific analysis. See Fleurbaey and Maniquet (1996*b*, 2001) for explorations of the problem of fair division of consumption goods.

4. Preferences are monotonic if $\ell_i \leq \ell_i'$ and $c_i > c_i'$ implies that $(\ell_i, c_i) P_i (\ell_i', c_i')$. Our analysis could be easily extended to the larger domain of preferences which are strictly monotonic in c, but not necessarily monotonic in ℓ. Assuming only local non-satiation, on the other hand, would require a more radical revision of the analysis (see footnote 5 below).

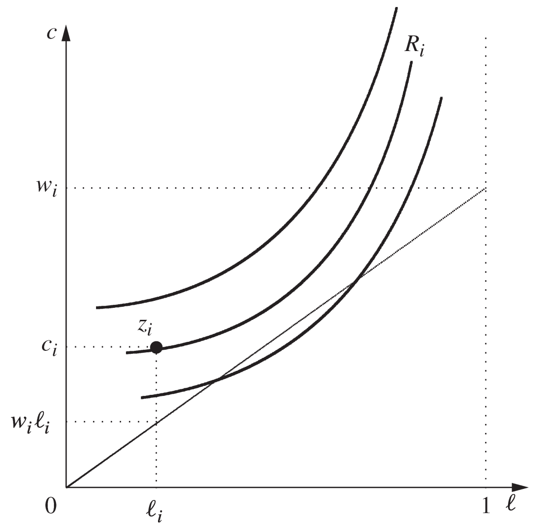

FIGURE 1

tations for these notions. The domain of economies for which we want social preferences to be defined contains all economies obeying the above conditions.

3. FAIR SOCIAL PREFERENCES

3.1. *Fairness requirements*

The main ethical requirement we will impose on social preferences, in this paper, is derived from the Pigou–Dalton transfer principle. Traditionally, however, this principle was applied to all income inequalities. This entails that no distinction is made between two agents with the same income but very different wage rates and different amounts of labour. We will be more cautious here, and apply it only to agents with identical labour. In addition, we will also restrict it to agents with identical preferences. There are two reasons for this additional restriction. First, applying the Pigou–Dalton principle to agents with different preferences would clash with the Pareto principle (to be defined more precisely below), as proved by Fleurbaey and Trannoy (2003). Second, when two agents have identical preferences one can more easily argue that they deserve to obtain similar incomes, whereas this is much less clear in the case of different preferences, as work disutility may differ. This gives us the following requirement:[5]

Transfer principle. If z and z' are two allocations, and i and j are two agents with identical preferences, such that $\ell_i = \ell_j = \ell'_i = \ell'_j$, and for some $\delta > 0$,

$$c'_i - \delta = c_i > c_j = c'_j + \delta,$$

whereas for all other agents k, $z_k = z'_k$, then $z \; R \; z'$.

Figure 2 illustrates the transfer. The axiom may sound too weak with the restriction $\ell_i = \ell_j$ if one thinks that an agent with higher skill and identical preferences is likely to work more in ordinary circumstances (like those of taxation described in the next section). But recall that, at

5. The transfer principle makes sense only when preferences are strictly monotonic in c. Otherwise, a transfer might fail to increase the receiver's satisfaction.

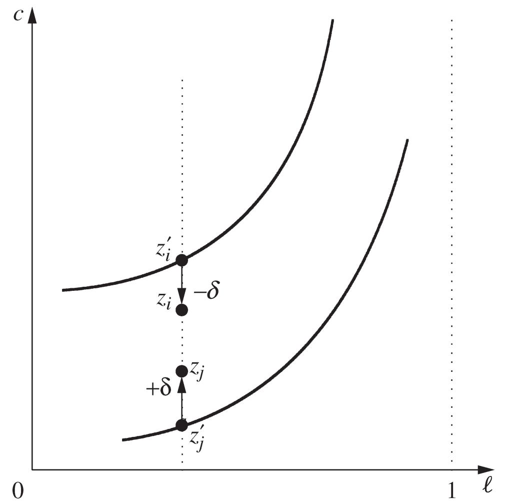

FIGURE 2

the stage of the construction of social preferences, we are only trying to find simple cases where our moral intuition is strong about how to improve the allocation. And we are not restricted to consider allocations that are likely to occur under specific institutions, since social preferences must rank all allocations. What this axiom says is simply that if, by whatever means, two agents with identical preferences and the same labour time happened to have different consumptions, then reducing this inequality would be socially acceptable. Independently of whether such a situation is likely or unlikely to occur (it is actually very common, in real life, for people who work full time), it is quite useful to consider it in order to put minimal constraints on social preferences.

Another possible objection is that if two agents have the same preferences, same labour but different productivity, it may seem normal that the more productive consumes more, whereas Pigou–Dalton transfers tend to eliminate inequality. In effect, the above axiom is justified only when agents cannot be held responsible for their differential productivity. This raises in particular the issue of whether the low-skilled may be considered to have responsibly chosen their lower productivity, or instead have suffered from various handicaps which have prevented them from acquiring higher skills. The Transfer Principle axiom is consistent with the latter view. We leave for future research the study of a richer model in which agents could be held partially responsible for their wage rate, via their educational or occupational choices.

The second fairness requirement we introduce has to do with providing opportunities and respecting individual preferences. Although reducing income inequalities is a generous goal, it is not obvious how to deal with agents who "choose" poverty out of a budget set which contains better income opportunities. In particular, when all agents have the same wage rate, it can be argued that there is no need for redistribution, as they all have access to the same labour–consumption bundles (Dworkin, 1981). Any income difference is then a matter of personal preferences. A laisser-faire allocation z^* is such that for every agent i, z_i^* is the best for R_i over the budget set defined by $c_i \leq w_i \ell_i$. The following requirement says that a laisser-faire allocation,[6] in this particular case of uniform earning ability, is (one of) the best among all feasible allocations.

6. There may be several laisser-faire allocations if preferences are not strictly convex. But all laisser-faire allocations, in a given economy, give agents the same satisfaction.

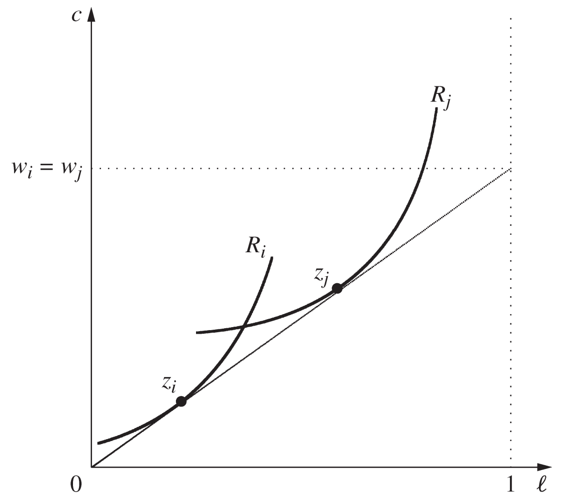

FIGURE 3

Laisser-faire. If all agents have the same wage rate w, then for any laisser-faire allocation z^* and any allocation z' such that $\sum_i c_i' \leq w \sum_i \ell_i'$, one has $z^*\ R\ z'$.

A laisser-faire allocation in a two-agent equal-skill economy is illustrated in Figure 3. Both agents have the same budget. Agent i, on the figure, may choose to have more leisure and less consumption, and the axiom of Laisser-Faire declares this to be unproblematic. One sees that this principle is acceptable if individual preferences are fully respectable, but should be treated with caution if some individual preferences are influenced by questionable social factors (*e.g.* apparent laziness may be due to discouragement and social stigma; workaholism may be due to social pressure).

The other requirements are basic conditions derived from the theory of social choice. First, we want social preferences to obey the standard Pareto condition. This condition is essential in order to take account of efficiency considerations. Social preferences satisfying the Pareto condition will never lead to the selection of inefficient allocations. In this way we are preserved against excessive consequences of fairness requirements, such as equality obtained through levelling-down devices.

Weak pareto. If z and z' are such that for all i, $z_i\ P_i\ z_i'$, then $z\ P\ z'$.

Second, we want our social preferences to use minimal information about individual preferences, in the spirit of Arrow's (1951) condition of independence of irrelevant alternatives. Arrow's condition is, however, much too restrictive, and leads to the unpalatable results of his impossibility theorem. Arrow's independence of irrelevant alternatives requires social preferences over two allocations to depend only on individual preferences over these two allocations. This condition makes it impossible, for instance, to check that two agents have the same preferences, or that an allocation is a laisser-faire allocation, etc. For extensive discussions of how excessive Arrow's independence is, see Fleurbaey and Maniquet (1996*b*, 2001) and Fleurbaey, Suzumura and Tadenuma (2003). We will instead follow Hansson (1973) and Pazner (1979) who have proposed a weaker condition still consistent with the idea that information needed to make social choices should be as parsimonious as possible. That condition requires social preferences over two allocations to depend only on individual *indifference curves* at these two allocations. More formally, it requires social preferences over two allocations to be the same in two different profiles of preferences when agents' indifference curves through the bundles they are assigned in these allocations are the same.

Hansson independence. Let z and z' be two allocations, and R, R' be the social orderings for two profiles $(R_1, \ldots, R_n)$ and $(R'_1, \ldots, R'_n)$, respectively. If for all i, and all $q \in X$,

$$z_i \ I_i \ q \Leftrightarrow z_i \ I'_i \ q$$
$$z'_i \ I_i \ q \Leftrightarrow z'_i \ I'_i \ q,$$

then

$$z \ R \ z' \Leftrightarrow z \ R' \ z'.$$

Finally, we want our social preferences to have a separable structure, as is usual in the literature on social index numbers. The intuition for separability requirements is that agents who are not concerned by a social decision need not be given any say in it. This is not only appealing because it simplifies the structure of social preferences, but also because it can be related to a standard conception of democracy, implying that unconcerned populations need not intervene in social decisions. This is often called the subsidiarity principle. We retain the following condition, requiring social preferences over two allocations to be unchanged if an agent receiving the same bundle in both allocations is removed from the economy.

Separability. Let z and z' be two allocations, and i an agent such that $z_i = z'_i$. Then

$$z \ R \ z' \Rightarrow z_{-i} \ R_{-i} \ z'_{-i},$$

where $z_{-i} = (z_1, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n)$, and R_{-i} is the social preference ordering for the economy with reduced population $\{1, \ldots, i-1, i+1, \ldots, n\}$.

3.2. *Social preferences*

The fairness conditions introduced above do not convey a strong aversion to inequality. Actually, the only redistributive condition here is the Transfer Principle, which, in the above weak formulation, is compatible with any degree of inequality aversion, including zero. Nonetheless, the combination of all the properties entails an infinite aversion to inequality, and forces social preferences to rely on the maximin criterion. Moreover, the maximin criterion needs to be applied to a precise evaluation of individual situations, as stated in the following theorem.

Theorem 1. *Let social preferences satisfy Transfer Principle, Laisser-Faire, Weak Pareto, Hansson Independence and Separability. For any allocations z, z', one has $z \ P \ z'$ if one of the following conditions holds:*

(i) $z_i \ P_i \ (0,0)$ *and* $z'_i \ R_i \ (0,0)$ *for all* i, *and*

$$\min_i W_i(z_i) > \min_i W_i(z'_i),$$

where $W_i(z_i) = \max\{w \in \mathbb{R}_+ \mid \forall \ell, \ z_i \ R_i \ (\ell, w\ell)\}$*;*

(ii) $z_i \ P_i \ (0,0)$ *for all* i *and* $(0,0) \ P_i \ z'_i$ *for some* i.

When $z_i \ R_i \ (0,0)$, the set $\{w \in \mathbb{R}_+ \mid \forall \ell, \ z_i \ R_i \ (\ell, w\ell)\}$ is not empty (it contains at least 0), and by monotonicity and continuity of preferences, it is compact, so that its maximum is well

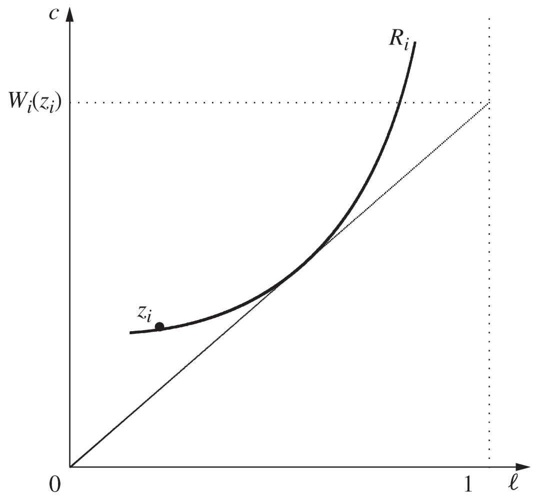

FIGURE 4

defined. The computation of $W_i(z_i)$ is illustrated in Figure 4. Concretely, $W_i(z_i)$ is the wage rate which would enable agent i to reach the same satisfaction as in z_i, if she were allowed to choose her labour time freely, at this wage rate: "What wage rate would give you the same satisfaction as your current situation?" Of course, we cannot think of using this question as a practical device for assessing individuals' situations. First, they may have a hard time working out what the true answer is. Second, they would have incentives to misrepresent their situation. The next section will examine how this kind of measure can be practically implemented.

Another interpretation of $W_i(z_i)$ relates it more directly to the axiom of Laisser-Faire. Consider an agent i who is indifferent between z_i and the bundle z_i^* she would choose in a laisser-faire allocation that would be socially optimal if all agents had an equal wage rate w^*. Then $W_i(z_i) = w^*$. In other words, $W_i(z_i)$ is the hypothetical common wage rate which would render this agent indifferent between z_i and an optimal allocation.[7]

The function $W_i(z_i)$ is a particular utility representation of agent i's preferences (for a part of the consumption set). It makes it possible to compare the situations of individuals who have identical or different preferences, on the basis of their current indifference curves. In addition, the social preferences described in Theorem 1 give absolute priority to agents with the lowest $W_i(z_i)$. In this way, this result suggests a solution to the problem of weighting different utility functions, mentioned in the introduction. By giving priority to the worst-off, such social preferences also escape Mirrlees' criticism of utilitarian social welfare functions. Mirrlees (1974), indeed, proved that utilitarian first-best allocations had to display the property that high-skilled agents envy low-skilled agents, that is, the former are assigned bundles on lower indifference curves than the latter.[8] In contrast, a first-best allocation maximizing $\min_i W_i(z_i)$ would have the property that all agents have the same $W_i(z_i)$. Consequently, two agents having the same preferences would be assigned bundles on the same indifference curve, independently of their skills, and no one would envy the other.

The proof of the theorem is in the Appendix. We provide the intuition for it here (the rest of this section may be skipped without any problem for understanding the rest of the paper). Let us

7. This concept is closely related to the Equal Wage Equivalent first-best allocation rule characterized on different grounds in Fleurbaey and Maniquet (1999).

8. Choné and Laroque (2001) generalize the criticism to the case where agents also differ in terms of their preferences, and use it as a justification for adopting social preferences of the maximin kind.

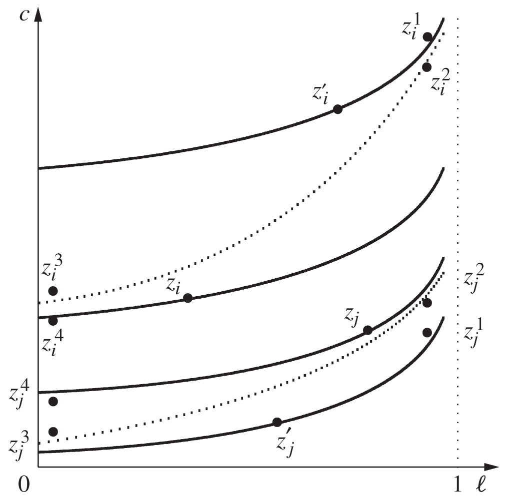

FIGURE 5

first show how the combination of Weak Pareto, Transfer Principle and Hansson Independence entails an infinite aversion to inequality. Consider two agents i and j with identical preferences R_0, and two allocations z and z' such that

$$z_i' \; P_0 \; z_i \; P_0 \; z_j \; P_0 \; z_j'.$$

The related indifference curves are shown in Figure 5, and one sees in this particular example that the axiom of Transfer Principle cannot directly entail that z is preferable to z', because agent i's loss of consumption between z' and z is much greater than agent j's gain, and also because their labour times differ. By Hansson Independence, social preferences over z and z' can only depend on the indifference curves through those allocations, so that they must coincide with what they would be if the dotted indifference curves represented in Figure 5 were also part of agents i and j's preferences.

In this particular case, one can construct intermediate allocations such as z^1, z^2, z^3, z^4 in the figure. By Weak Pareto, $z^1 \; P \; z'$. By Transfer Principle, $z^2 \; R \; z^1$. By Weak Pareto again, $z^3 \; P \; z^2$. By Transfer Principle again, $z^4 \; R \; z^3$. Finally, Weak Pareto implies $z \; P \; z^4$, so, by transitivity, one can conclude that $z \; P \; z'$. Since this kind of construction can be done even when the gain is very small for j while i's loss is huge, one then obtains an infinite inequality aversion regarding indifference curves of agents with identical preferences.

The second central part of the argument consists in proving that the maximin has to be applied to $W_i(z_i)$. The crucial axioms are now Laisser-Faire and Separability. Let us illustrate the proof in the case of two agents i and j and two allocations z and z' such that $z_k = z_k'$ for all $k \neq i, j$, and

$$W_i(z_i') > W_i(z_i) > W_j(z_j) > W_j(z_j').$$

We need to conclude that z is better than z'. Introduce two new agents, a and b, whose identical wage rate w is such that $W_i(z_i) > w > W_j(z_j)$, and whose preferences are $R_a = R_i$ and $R_b = R_j$. Let z^* denote a laisser-faire allocation for the two-agent economy formed by a and b, and (z_a, z_b) be another allocation which is feasible but inefficient in this two-agent economy, and such that

$$W_i(z_i) > W_a(z_a) > w > W_b(z_b) > W_j(z_j).$$

Figure 6 illustrates these allocations.

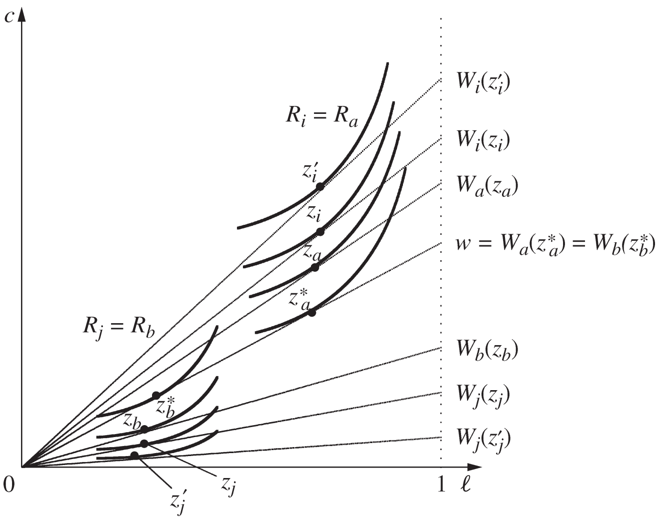

FIGURE 6

Let $R_{\{a,b\}}$, $R_{\{a,b,i,j\}}$ and $R_{\{i,j\}}$ denote the social preferences for the economies with population $\{a,b\}$, $\{a,b,i,j\}$ and $\{i,j\}$, respectively. By Laisser-Faire and Weak Pareto, a laisser-faire allocation is strictly better than any inefficient feasible allocation, so $z^* \; P_{\{a,b\}} \; (z_a, z_b)$. Therefore, by Separability, it must necessarily be the case that

$$(z_a^*, z_b^*, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a, z_b, z_i, z_j).$$

By the above argument producing an infinite inequality aversion among agents with identical preferences (from Transfer Principle and Hansson Independence), one also sees that, by reducing the inequality between agents a and i,

$$(z_a, z_b^*, z_i, z_j') \; P_{\{a,b,i,j\}} \; (z_a^*, z_b^*, z_i', z_j')$$

and between agents b and j,

$$(z_a, z_b, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a, z_b^*, z_i, z_j').$$

As a consequence, by transitivity one has

$$(z_a^*, z_b^*, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a^*, z_b^*, z_i', z_j'),$$

from which Separability entails that

$$(z_i, z_j) \; R_{\{i,j\}} \; (z_i', z_j').$$

We would have obtained the desired strict preference $(z_i, z_j) \; P_{\{i,j\}} \; (z_i', z_j')$ by referring, in the previous stages of this argument, to another allocation (z_i'', z_j'') Pareto-dominating z', instead of z' itself. Then, from Separability again, one can finally derive the conclusion that $z \; P \; z'$ in the initial economy.

From this intuitive proof, one sees that it is the combination of Transfer Principle and Hansson Independence which leads to focusing on the worst-off, and that it is the combination of Laisser-Faire and Separability which singles out $W_i(z_i)$ as the proper measure of individual situations.

This theorem does not give a full characterization of social preferences, because it does not say how to compare allocations for which $\min_i W_i(z_i) = \min_i W_i(z_i')$. But for the purpose of evaluating taxes and finding the optimal tax, the description given in the theorem is sufficient to yield precise results, as we will show in the next section. Moreover, the theorem does not say how to define the social ranking within the subset of allocations such that $(0,0)\ P_i\ z_i$ for some i, but it says that such allocations are low in the social ranking and again that is sufficient for the purpose of tax applications.

As an additional illustration of this result, let us briefly examine how other kinds of social preferences fare with respect to the axioms. In order to simplify the discussion, we restrict our attention to how social preferences rank allocations z such that $z_i\ R_i\ (0,0)$ for all i. First, consider social preferences based on $\sum_i W_i(z_i)$ instead of $\min_i W_i(z_i)$:

$$z\ R\ z' \Leftrightarrow \sum_i W_i(z_i) \geq \sum_i W_i(z_i').$$

Such social preferences violate Transfer Principle and Laisser-Faire. Social preferences based on the median $W_i(z_i)$ would, in addition, violate Separability. Now, consider social preferences similar to those retained in Choné and Laroque (2001), and based on $\text{leximin}_i\ C_i(z_i)$,[9] where

$$C_i(z_i) = \max\{c \in \mathbb{R}_+ \mid z_i\ R_i\ (0,c)\}.$$

Such social preferences satisfy all our axioms except Laisser-Faire. Consider social preferences based on $\text{leximin}_i\ V_i(z_i)$, where

$$V_i(z_i) = \max\{t \in \mathbb{R} \mid \forall \ell,\ z_i\ R_i\ (\ell, t + w_i \ell)\}.$$

These social preferences satisfy all our axioms except Transfer Principle. As a final example, consider utilitarian social preferences based on $\sum_i U_i(z_i)$, where U_i is an exogenously given utility function representing R_i. Such social preferences require more information (the U_i functions) than the social preferences studied in this paper, and therefore do not fit exactly in our framework. One can nonetheless examine whether they satisfy some of our axioms. They fully satisfy Weak Pareto and Separability. They also satisfy Transfer Principle when the utility functions are concave in c (and when two agents with identical preferences also have identical utility functions). They do not satisfy Laisser-Faire, except on the subdomain of utility functions which are quasi-linear in c, and do not satisfy Hansson Independence on any reasonable domain.

4. TAX REDISTRIBUTION

4.1. *Setting*

In this section, we examine the issue of devising the redistribution system under incentive-compatibility constraints and with the objective of achieving the best possible consequences according to the above social preferences. As is standard in the second-best context, whose formalism dates back to Mirrlees (1971), we assume that only earned income $y_i = w_i \ell_i$ is observed, so that redistribution is made via a tax function $\tau(y_i)$. This tax is a subsidy when $\tau(y_i) < 0$. Individuals are free to choose their labour time in the budget set modified by the tax schedule. The government is assumed to know the distribution of types (preferences, earning abilities) in the population but ignores the characteristics of any particular agent. Since it is easy to forecast the behaviour of any given type of agent under a tax schedule, knowing the distribution of

9. Leximin is the lexicographic extension of maximin (when the smallest value is equal, one looks at the second smallest value, and so on).

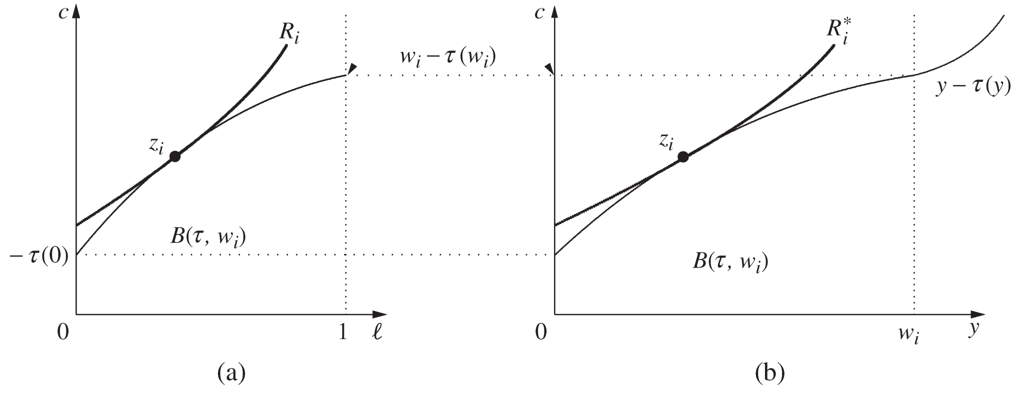

FIGURE 7

types enables the government to forecast the social consequences of any tax function. It may then evaluate or choose a tax function in view of the foreseen social consequences.

Under this kind of redistribution, agent i's budget set is defined by (see Figure 7(a)):

$$B(\tau, w_i) = \{(\ell, c) \in X \mid c \leq w_i \ell - \tau(w_i \ell)\}.$$

Notice that $-\tau(0)$ is the minimum income granted to agents with no earnings. It is convenient to focus on the earnings–consumption space, in which the budget is defined by (see Figure 7(b)):

$$B(\tau, w_i) = \{(y, c) \in [0, w_i] \times \mathbb{R}_+ \mid c \leq y - \tau(y)\}.$$

We retain the same notation for the two sets since no confusion is possible. Similarly, in our figures z_i will simultaneously denote the bundle (ℓ_i, c_i) in one space and the bundle $(y_i, c_i) = (w_i \ell_i, c_i)$ in the other space.

In the earnings–consumption space, one can define individual preferences R_i^* over earnings–consumption bundles, and they are derived from ordinary preferences over labour–consumption bundles via

$$(y, c)\ R_i^*\ (y', c') \Leftrightarrow \left(\frac{y}{w_i}, c\right)\ R_i\ \left(\frac{y'}{w_i}, c'\right).$$

The fact that all agents are submitted to the same constraint $c \leq y - \tau(y)$ implies that for any pair of agents i, j, when i chooses (y_i, c_i) in $B(\tau, w_i)$ and j chooses (y_j, c_j) in $B(\tau, w_j)$, one must have $(y_i, c_i) R_i^* (y_j, c_j)$ or $y_j > w_i$.

Conversely,[10] any allocation z satisfying

$$\text{for all } i, j,\ (y_i, c_i)\ R_i^*\ (y_j, c_j) \text{ or } y_j > w_i \quad \text{(self-selection constraints)}$$

is incentive-compatible and can be obtained by letting every agent i choose her best bundle in a budget set $B(\tau, w_i)$ for some well-chosen tax function τ. This tax function must be such that $y - \tau(y)$ lies nowhere above the envelope curve of the indifference curves of the population in the (y, c)-space, and intersects this envelope curve at all points (y_i, c_i) for $i = 1, \ldots, n$. By monotonicity of individual preferences, we may restrict attention to tax functions τ such that $y - \tau(y)$ is non-decreasing.

10. See, *e.g.* Stiglitz (1987, pp. 1002–1004) or Boadway and Keen (2000, pp. 737–738).

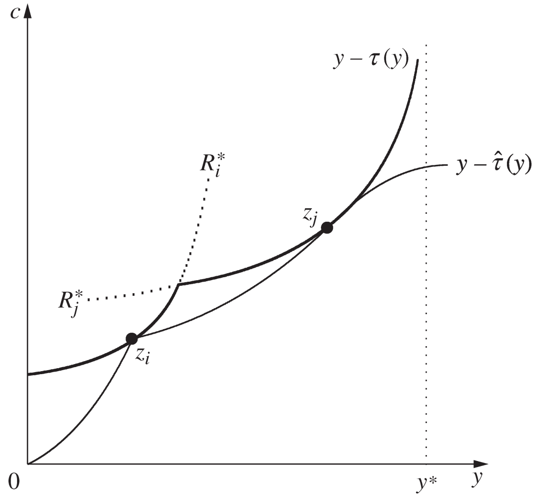

FIGURE 8

An allocation is *feasible* if it satisfies

$$\sum_{i=1}^{n} c_i \leq \sum_{i=1}^{n} y_i.$$

A tax function τ is *feasible* if it satisfies

$$\sum_{i=1}^{n} \tau(w_i \ell_i) \geq 0$$

when all agents choose their labour time by maximizing their satisfaction over their budget set.

Consider an incentive-compatible allocation z. By the assumptions made on individual preferences, the envelope curve of the agents' indifference curves in (y, c)-space, at z, is then the graph of a non-decreasing, non-negative function f defined on an interval $S(z) \subset [0, \max_i w_i]$. Let τ be a tax function yielding the allocation z. It is called *minimal* when $y - \tau(y) = f(y)$ for all $y \in S(z)$, or equivalently when any tax function τ' which yields the same incentive-compatible allocation z is such that $\tau'(y) \geq \tau(y)$ for all $y \in S(z)$. Concretely, when a tax τ is not minimal, one can devise tax cuts which have no consequence on the agents' behaviour and on tax receipts (because no agent has earnings in the range of the tax cuts). Figure 8 illustrates this, with a minimal tax τ and a non-minimal tax $\hat{\tau}$. When $\max_i w_i \notin S(z)$, then there is y^* such that $\lim_{y \to y^*, y < y^*} f(y) = +\infty$ (see Figure 8). In this case, by convention we let any corresponding minimal tax have $\tau(y) = -\infty$ (or equivalently $y - \tau(y) = +\infty$) for all $y \geq y^*$. When $z_i \; R_i \; (0, 0)$ for all i, then $0 \in S(z)$, so on the interval $[0, \max_i w_i]$ there is only one minimal tax τ corresponding to z.[11]

In the following, we explore the evaluation of taxes for the class of social preferences highlighted in Theorem 1. This means that an incentive-compatible allocation z is socially preferred to another incentive-compatible allocation z' whenever

$$\min_i W_i(z_i) > \min_i W_i(z_i').$$

The way $W_i(z_i)$ is computed in the earnings–consumption space is illustrated in Figure 9.

11. The definition of $\tau(y)$ for $y > \max_i w_i$ does not matter. By convention, for all tax functions considered in this paper, we let $\tau(y) = \tau(\max_i w_i)$ for all $y > \max_i w_i$.

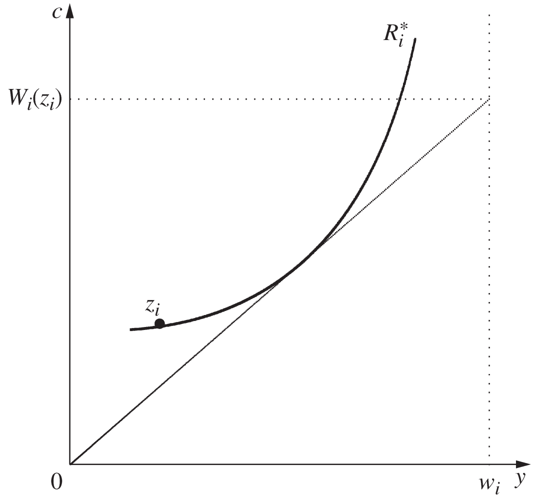

FIGURE 9

4.2. *Two agents*

As an introductory analysis, consider the case of a two-agent population $\{1, 2\}$. Assume that $w_1 < w_2$. As a consequence, agent 2's budget set always contains agent 1's one. And if agent 1's labour time is positive at the laisser-faire allocation z^*, necessarily $W_1(z_1^*) < W_2(z_2^*)$ since $W_i(z_i^*) \geq w_i$ for $i = 1, 2$, with equality $W_i(z_i^*) = w_i$ when the agent has a positive labour time. (If an agent is so averse to labour that $\ell_i^* = 0$, then $W_i(z_i^*)$ equals the marginal rate of substitution at $(0, 0)$, which is greater than or equal to w_i.)

If the agents have the same preferences $R_1 = R_2$, then the optimal tax is the one which maximizes the satisfaction of agent 1 (since agent 2's budget set contains agent 1's one, in the case of identical preferences one has $W_2(z_2) \geq W_1(z_1)$ in any incentive-compatible allocation). This result extends immediately to a larger population: When all agents have the same preferences, an optimal tax is one which, among the feasible tax functions, maximizes the satisfaction of the agents with the lowest wage rate.

In the general case when the agents may have the same or different preferences (assuming that agent 1 has a positive labour time at the laisser-faire allocation), then either the optimal tax achieves an allocation such that $W_1(z_1) = W_2(z_2)$, or it maximizes the satisfaction of agent 1 over the set of feasible taxes. The argument for this fact is the following. Starting from the laisser-faire z^* where $W_1(z_1^*) < W_2(z_2^*)$, one redistributes from agent 2 to agent 1, and this increases $W_1(z_1)$ and decreases $W_2(z_2)$, following the second-best Pareto frontier. When one reaches the equality $W_1(z_1) = W_2(z_2)$, redistribution has to stop, since, by Pareto-efficiency, there is no other allocation with a greater $\min_i W_i(z_i)$. But an alternative possibility is that the incentive-compatibility constraint $(y_2, c_2) R_2^* (y_1, c_1)$ puts a limit on redistribution, which occurs when the point maximizing agent 1's satisfaction is reached. Then, the inequality $W_1(z_1) < W_2(z_2)$ remains at the optimal tax.

Figure 10 illustrates these two possibilities. In (a), the optimal allocation has $W_1(z_1) = W_2(z_2)$. The fact that it does not maximize the satisfaction of agent 1 is transparent in this example because agent 2's self-selection constraint is not binding—note that the allocation is then first-best efficient. In (b), the optimal allocation maximizes the satisfaction of agent 1 and $W_1(z_1) < W_2(z_2)$.

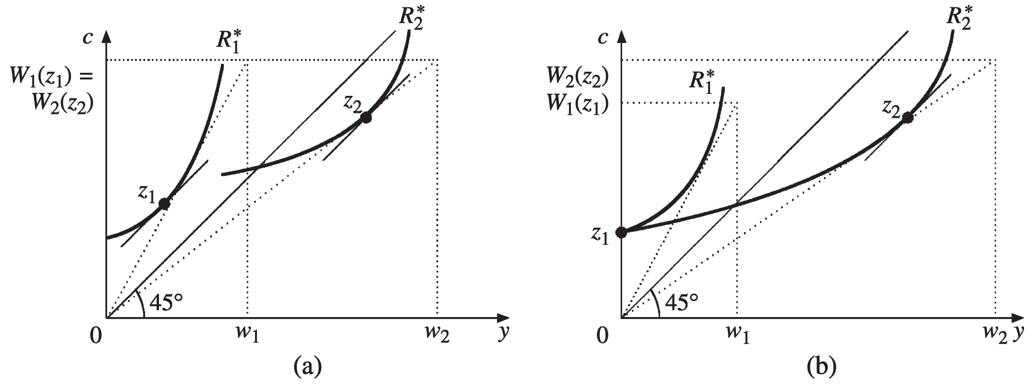

FIGURE 10

4.3. *General population*

Let us now turn to the case of a larger population. The computation of the optimal tax is quite complex in general, in particular because the population is heterogeneous in two dimensions, preferences and earning ability.[12] We will, however, be able to derive some conclusions about, first, the part of the tax schedule which should be the focus of the social planner and, second, some features of the optimal tax.

The main difficulty in such an analysis comes from the theoretical possibility of observing ranking reversals, with high-skilled agents earning lower incomes than low-skilled agents. In the standard setting with agents differing only in the skill dimension, this is usually excluded by the Spence–Mirrlees single-crossing assumption. In the current multi-dimensional setting, it would be exceedingly artificial to exclude such reversals, since agents with slightly different wages may obviously have quite different preferences, and it would be questionable to assume that high-skilled agents are always more hardworking than low-skilled agents. Fortunately, it appears that the real difficulty does *not* lie with individual reversals, that is, with the fact that *some* high-skilled agent may earn less than *some* low-skilled agent. For our purposes, we only need to exclude the possibility of observing gaps in the distribution of earnings of low-skilled agents, with such gaps filled only with high-skilled agents. That is, we need to exclude the possibility of having, say, a succession of intervals $[0, y_1]$, (y_1, y_2), $[y_2, w]$, such that agents with wage rate w earn only incomes in the intervals $[0, y_1]$ and $[y_2, w]$, whereas in the earnings interval (y_1, y_2) one only finds agents with skill $w' > w$. Excluding this possibility is quite natural. This can be done by assuming that whenever some high-skilled agents are ready to have earnings in some intermediate interval (y_1, y_2), there are also low-skilled agents with locally similar preferences in the (y, c)-space who are willing to earn similar levels of income.

Formally, let $uc((y_i, c_i), w_i, R_i^*)$ denote the closed upper contour set for R_i^* at (y_i, c_i):

$$uc((y_i, c_i), w_i, R_i^*) = \left\{(y, c) \in [0, w_i] \times \mathbb{R}_+ \mid (y, c) R_i^* (y_i, c_i)\right\}.$$

The assumption that we introduce says that a high-skilled agent, when contemplating low earnings, always finds low-skilled agents who have *locally* similar preferences in the (y, c)-space. Let $w_m = \min_i w_i$. We assume throughout this section that $w_m > 0$.

12. Actually, since the set of individual preferences is itself infinitely multi-dimensional, this is a problem of screening with, *a priori*, infinitely many dimensions of heterogeneity (but the population is finite in our model). The fact that the complexity of the multi-dimensional screening problem increases with the number of dimensions is shown in Matthews and Moore (1987).

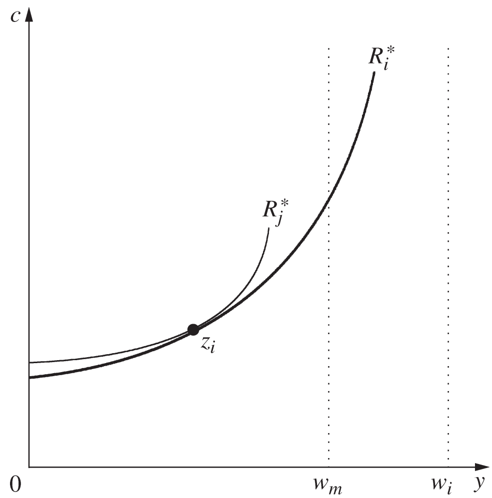

FIGURE 11

Assumption (Low-Skill Diversity). *For every agent i, and every (y, c) such that $y \leq w_m$, there is an agent j such that $w_j = w_m$ and $uc((y, c), w_j, R_j^*) \subseteq uc((y, c), w_i, R_i^*)$.*

Figure 11 illustrates this configuration. The inclusion of upper contour sets means that whenever agent i chooses (y_i, c_i) in a budget set, there is a low-skilled agent j who is willing to choose the same bundle (y_i, c_i) from the same budget set (for another bundle, it may be another low-skilled agent).

This assumption is of course rather strong for small populations. As explained above, however, what is needed for the results below is only that there be no gap in the distribution of earnings for low-skilled agents. More precisely, the consequence of Low-Skill Diversity that is used below is that for all incentive-compatible and feasible allocations, the envelope curve in (y, c)-space of the indifference curves of low-skilled agents coincides over the interval $[0, w_m]$ of earnings with the envelope curve of the whole population. This weaker assumption is quite natural for large populations, and Low-Skill Diversity is probably the simplest assumption on the primitives of the model which guarantees that it will be satisfied.

The first result in this section has to do with translating the abstract objective of maximizing $\min_i W_i(z_i)$ into a more concrete objective about the part of the agents' budget set which should be maximized.

Theorem 2. *Consider two incentive-compatible allocations z and z' obtainable with two minimal tax functions τ and τ', respectively, such that $\tau(0) < 0$ and $\tau'(0) \leq 0$. If social preferences satisfy Transfer Principle, Laisser-Faire, Weak Pareto, Hansson Independence and Separability, then z is socially preferred to z' whenever the maximal average tax rate over low incomes $y \in [0, w_m]$ is smaller in z:*

$$\max_{0 \leq y \leq w_m} \frac{\tau(y)}{y} < \max_{0 \leq y \leq w_m} \frac{\tau'(y)}{y}.$$

The proof of this result (see the Appendix) goes by showing that this inequality on tax rates entails that

$$\min_i W_i(z_i) > \min_i W_i(z_i'),$$

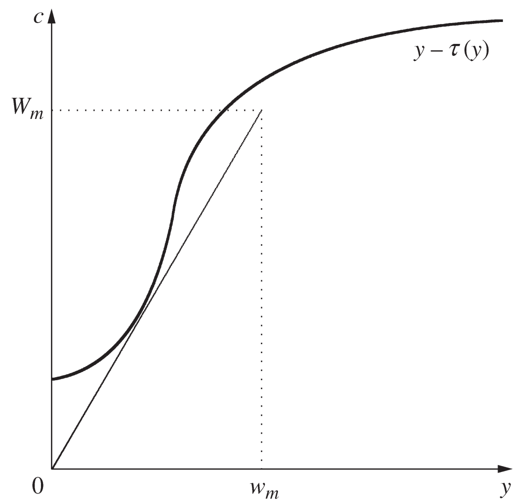

FIGURE 12

so one may apply Theorem 1 to conclude that z is socially preferred. Note that $\tau(0) < 0$ implies that $z_i \; P_i \; (0, 0)$ for all i. The priority of the worst-off in social preferences, combined with the assumption of Low-Skill Diversity, is the key factor that leads to focusing on earnings in the range $[0, w_m]$. The measure of individual situations by $W_i(z_i)$, on the other hand, is the key ingredient for taking the average tax rate $\tau(y)/y$ as the relevant token. Indeed, consider on Figure 12 that the graph of $y - \tau(y)$ over the range $[0, w_m]$ coincides (by Low-Skill Diversity and the assumption that the tax function is minimal) with the envelope curve of low-skilled agents' indifference curves. As shown in the figure, the smallest value of $W_i(z_i)$ for the low-skilled agents is then found by looking for the ray that is tangent to this portion of the graph, and therefore equals

$$W_m = w_m \times \min_{0 \leq y \leq w_m} \frac{y - \tau(y)}{y}.$$

It turns out that this is actually the smallest value of $W_i(z_i)$ over the whole population. The conclusion of Theorem 2 immediately follows.

This result has three features which deserve some comments. First, this result does not only provide information about the optimal tax, saying that it must minimize

$$\max_{0 \leq y \leq w_m} \frac{\tau(y)}{y},$$

but also gives a criterion for the assessment of suboptimal taxes. Given the fact that political constraints and disagreements often make the computation of the optimal tax look like an ethereal exercise, it is quite useful to be able to say something about realistic taxes and piecemeal reforms in an imperfect world.

Second, it provides a very simple criterion for the observer who wants to compare taxes. The application of the criterion requires no information about the population characteristics, except the value of w_m, which, in practice, may be thought to coincide with the legal minimum wage.[13] Therefore, there is no need to measure $W_i(z_i)$ for every individual, nor even to estimate the

13. Except, perhaps, when there is more than frictional unemployment. See below.

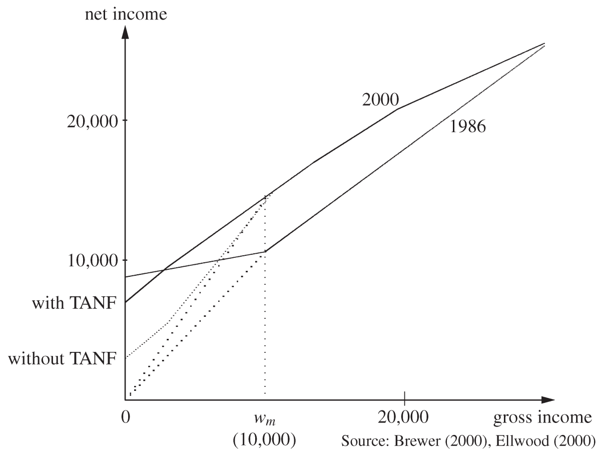

FIGURE 13

The U.S. reform

distribution of $W_i(z_i)$ over the population. A simple examination of the tax schedule yields a definite answer.

Third, the content of the criterion itself is quite intuitive. It says that the focus should be on the maximum average tax rate $\tau(y)/y$ over low earnings. Near the optimal tax, low earnings will actually be subsidized, that is, $\tau(y)$ will be negative over this range. Then, the criterion means that the *smallest average rate of subsidy should be as high as possible*, over this range. Interestingly, when w_m tends to zero, the criterion boils down to comparing the value of the minimum income (or demogrant) $-\tau(0)$, and advocates that it should be as high as possible.

It may be useful here to illustrate how the simple comparison criterion provided in the above result can be applied. The next figure presents the 2000 budget set for a lone parent with two children in the U.S.[14] Net income is computed including income tax, social security contributions, food stamps and Temporary Assistance to Needy Families (TANF), a scheme which replaced the Aid to Families with Dependent Children (AFDC) programme in 1996.[15] Since the TANF is temporary (it has a five-year limit), it is also relevant to look at the budget set after withdrawal of TANF. This is drawn on the figure with a dotted line. An approximate representation of a 1986 (pre-reform) budget is also provided, in order to assess the impact of the reform. The reform has had a positive impact according to the criterion provided in Theorem 2, as shown by the dotted rays from the origins. The conclusion remains even when withdrawal of TANF is considered.

In the following theorem, we provide more information about the optimal tax.

Theorem 3. *Assume that there exists a feasible (not necessarily incentive-compatible) allocation z such that z_i P_i $(0,0)$ for all i. If z^* is an optimal (incentive-compatible) allocation for social preferences satisfying Transfer Principle, Laisser-Faire, Weak Pareto, Hansson Independence and Separability, then it can be obtained with a tax function τ^* which, among all feasible tax functions, maximizes the net income of the hardworking poor, $w_m - \tau(w_m)$, under the constraints that*

14. In this paper, we do not deal with the issue of unequal household sizes. Theorem 2 does however apply to any subpopulation of households of a certain kind. The case of lone parents with children is probably the most relevant if one wants to focus on the subgroup of the population which is the worst-off in all respects.

15. The TANF programme is managed at the State level. Figures corresponding to Florida are retained in Figure 8.

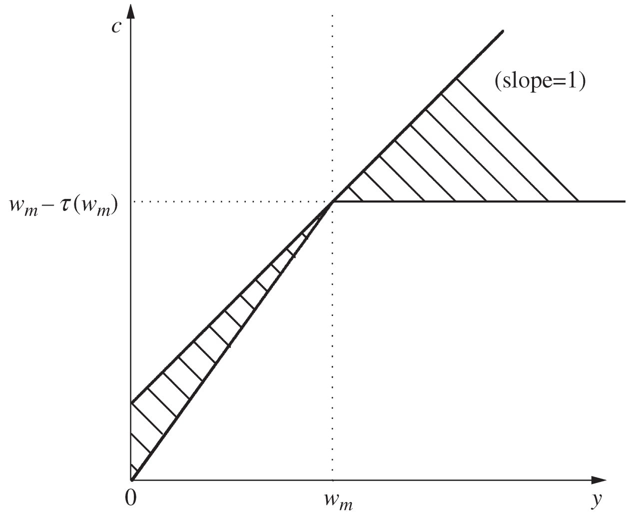

FIGURE 14

$$\frac{\tau(y)}{y} \leq \frac{\tau(w_m)}{w_m} \quad \textit{for all } y \in (0, w_m],$$
$$\tau(y) \geq \tau(w_m) \quad \textit{for all } y,$$
$$\tau(0) \leq 0.$$

The initial assumption made in the theorem simply excludes the case when the zero allocation is efficient and there is therefore no interesting possibility of redistribution. The three constraints listed at the end of the statement mean, respectively, that the average tax rate on low incomes is always lower than at w_m, that the tax (subsidy) is the smallest (largest) at w_m, and that the tax (subsidy) is non-positive (non-negative) at 0.

This result does not say that every optimal tax must satisfy these constraints, but it says, quite relevantly for the social planner, that there is no problem, *i.e.* no welfare loss, in restricting attention to taxes satisfying those constraints, when looking for the optimal allocation. This result shows how the social preferences defined in this paper lead to focusing on the hardworking poor, who should get, in the optimal allocation, the greatest *absolute amount* of subsidy, among the whole population. However, the taxes computed for those with a lower income than w_m also matter, as those agents must obtain at least as great a *rate of subsidy* as the hardworking poor.

Theorem 3 is illustrated in Figure 14. From the point $(w_m, w_m - \tau(w_m))$ one can construct the hatched area delimited by an upper line of slope 1 and a lower boundary made of the ray to the origin (on the left) and a flat line (on the right). Now, Theorem 3 says that computing the optimal tax may, without welfare loss, be done by maximizing the second coordinate of the point $(w_m, w_m - \tau(w_m))$ under the constraint that the income function $y - \tau(y)$ is located in the corresponding hatched area. It is useless to consider income functions which lie outside this area.

Interestingly, the shape of this area implies that the marginal tax rate over incomes below w_m is, on average, non-positive.

As explained above, when there are agents with almost zero earning ability, our results boil down to a simple maximization of the minimum income. The case of a zero w_m can be related to productive disabilities but also to involuntary unemployment. Since unemployment may be viewed as nullifying the agents' earning ability, this result should best be interpreted as suggesting that the focus of redistributive policies should shift from the hardworking poor to the low-income households when the extent of unemployment is large, and especially when long-term unemployment is a significant phenomenon. Then, for instance, the assessment of the welfare reform in the U.S., as illustrated in Figure 13, would be much less positive since the minimum income has been reduced (and the temporariness of TANF would appear quite questionable in this context). On the other hand, physical disabilities and unemployment are more or less observable characteristics, which may elicit special policies toward those affected by such conditions, as can be witnessed in many countries.[16] If this is the case, then the above result should apply to the rest of the population, and the relevant value of w_m is then likely to be the minimum legal hourly wage. Nonetheless when unemployment takes the form of constrained part time jobs (a less easily observable form than ordinary unemployment), this should also be tackled by considering it as a reduction of the agents' earning ability.

5. CONCLUSION

In this paper, we have examined how two fairness principles, a weak version of the Pigou–Dalton transfer principle and a laisser-faire principle for equal-skill economies, single out particular social preferences and a particular measure of individual situations. Such social preferences grant absolute priority to the worst-off, in the maximin fashion. This result[17] might contribute to lending more respectability to the maximin criterion, which is sometimes criticized for its extreme aversion to inequality.[18]

The measure of individual situations obtained here is the tax-free wage rate which would enable an agent to maintain her current satisfaction. This may be viewed as a special money-metric utility representation of individual preferences. The choice of this measure, however, did not rely on introspection or a philosophical examination of human well-being. It derived from the fairness principles (especially the laisser-faire principle), and the analysis did not require any other informational input about individual welfare than ordinal non-comparable preferences. The famous impossibility of social choice (Arrow's theorem) was avoided by weakening Arrow's axiom of Independence of Irrelevant Alternatives in order to take account of the shape of individual indifference curves at the allocations under consideration. It must be stressed that we do not consider W_i as the only reasonable measure of individual situations. In Fleurbaey and Maniquet (2005), alternative social preferences, using different measures, are defended on the basis of other ethical principles. Our purpose is not to defend a single view of social welfare, but to clarify the link between fairness principles and concrete policy evaluations. "It is a legitimate exercise of economic analysis to examine the consequences of various value judgments, whether or not they are shared by the theorist" (Samuelson, 1947, p. 220).

The second part of the paper studied the implications of such social preferences for the evaluation of income tax schedules, under incentive constraints due to the unobservability of skills and the possibility for agents to freely choose their labour time in their budget set. The main result was the discovery of a simple criterion for the comparison of tax schedules, based on the smallest average subsidy (or greatest average tax rate) for low incomes. Another result was that the average marginal tax rate for low incomes should optimally be non-positive, and that the hardworking poor should receive maximal subsidies, under the constraint that lower incomes

16. Observation of disabilities and involuntary unemployment is, however, imperfect. For an analysis of optimal taxation under imperfect tagging, see Salanié (2002).

17. Similar derivations of the maximin criterion have also been obtained in different contexts by Fleurbaey (2001) and Maniquet and Sprumont (2004).

18. It has always been, however, one of the prominent criteria in the literature of optimal taxation. See, *e.g.* Atkinson (1973, 1995) and, more recently, Choné and Laroque (2001).

should not have a lower rate of subsidy than the hardworking poor. This constraint is important. It forbids policies which harshly punish the agents working part time and give exclusive subsidies to full-time jobs. In addition, various forms of unemployment can be taken into account by revising the distribution of earning abilities in the population, leading to a reduction of w_m and therefore to a more generous policy toward low incomes.

There are many directions in which this line of research can be pursued. In particular, the model can be enriched so as to study such issues as savings and the taxation of unearned income, or different consumption goods and the interaction between income taxation and commodity taxation.

APPENDIX: PROOFS

Lemma 1. *If social preferences satisfy Transfer Principle, Weak Pareto and Hansson Independence, then for any pair of allocations z, z' and any pair of agents i, j with identical preferences R_0, such that*

$$z'_i \; P_0 \; z_i \; P_0 \; z_j \; P_0 \; z'_j \; R_0 \; (0,0)$$

and $z_k \; P_k \; z'_k$ for all $k \neq i, j$, one has $z \; P \; z'$.

Proof. Let z, z' satisfy the above conditions. By Hansson Independence, we can arbitrarily modify the preferences R_0 at bundles which are not indifferent to one of the four bundles z_i, z'_i, z_j, z'_j. Let f_i, g_i, f_j, g_j be the functions whose graphs are the indifference curves for R_0 at these four bundles, respectively. Let f_i^* be the function whose graph is the lower boundary of the convex hull of

$$(0, f_i(0)) \cup \{(\ell, c) \mid c \geq g_i(\ell)\},$$

and f_j^* be the function whose graph is the lower boundary of the convex hull of

$$(0, g_j(0)) \cup \{(\ell, c) \mid c \geq f_j(\ell)\}.$$

These functions are convex, and their graphs can be arbitrarily close (w.r.t. the sup norm) to two indifference curves for R_0. We will indeed assume that there is an indifference curve for R_0, between f_i and g_i, arbitrarily close to the graph of f_i^*, and another one, between f_j and g_j, arbitrarily close to f_j^*.

By construction there exists ℓ_1 such that

$$g_i(\ell_1) - f_i^*(\ell_1) < f_j^*(\ell_1) - g_j(\ell_1),$$

and similarly

$$f_i^*(0) - f_i(0) = 0 < f_j(0) - f_j^*(0) = f_j(0) - g_j(0).$$

Therefore, one can find $c_i^1, c_i^2, c_j^1, c_j^2$ such that

$$g_i(\ell_1) - f_i^*(\ell_1) < c_i^1 - c_i^2 = c_j^2 - c_j^1 < f_j^*(\ell_1) - g_j(\ell_1),$$

$$c_i^2 < f_i^*(\ell_1) \leq g_i(\ell_1) < c_i^1,$$

$$g_j(\ell_1) < c_j^1 < c_j^2 < f_j^*(\ell_1),$$

and $c_i^3, c_i^4, c_j^3, c_j^4$ such that

$$0 < c_i^3 - c_i^4 = c_j^4 - c_j^3 < f_j(0) - f_j^*(0),$$

$$c_i^4 < f_i(0) = f_i^*(0) < c_i^3,$$

$$f_j^*(0) < c_j^3 < c_j^4 < f_j(0).$$

Define z^1, z^2, z^3, z^4 by

$$z_k \; P_k \; z_k^4 = z_k^3 \; P_k \; z_k^2 = z_k^1 \; P_k \; z'_k$$

for all $k \neq i, j$, and

$$z_k^1 = (\ell_1, c_k^1),\; z_k^2 = (\ell_1, c_k^2),\; z_k^3 = (0, c_k^3),\; z_k^4 = (0, c_k^4)$$

for all $k = i, j$.

By Transfer Principle, one has

$$z^2 \; R \; z^1 \text{ and } z^4 \; R \; z^3.$$

By Weak Pareto (and the assumption about indifference curves close to f_i^* and f_j^*),

$$z \; P \; z^4, z^3 \; P \; z^2 \text{ and } z^1 \; P \; z'.$$

By transitivity, one concludes that $z \; P \; z'$. ‖

Lemma 2. *If social preferences satisfy Transfer Principle, Laisser-Faire, Weak Pareto, Hansson Independence and Separability, then for any pair of allocations z, z' and any pair of agents i, j, such that*

$$W_i(z_i') > W_i(z_i) > W_j(z_j) > W_j(z_j'),$$

$z_i \; P_i \; (0,0)$ *and* $z_k = z_k'$ *for all* $k \neq i, j$, *one has* $z \; P \; z'$.

Proof. Let z and z' be two allocations satisfying the above conditions. Necessarily $W_j(z_j) > 0$, and since $z_i \; P_i \; (0,0)$ there exists $\hat{z}_i$ such that $W_i(z_i) > W_i(\hat{z}_i) > W_j(z_j)$. Let a and b be two new agents with $w_a = w_b = w = W_i(\hat{z}_i)$, and with preferences $R_a = R_i$ and $R_b = R_j$. Let z^* be a laisser-faire allocation for the two-agent economy $\{a, b\}$, and (z_a, z_b) be another allocation such that

$$W_i(z_i) > W_a(z_a) > w > W_b(z_b) > W_j(z_j)$$

and

$$c_a + c_b < w(\ell_a + \ell_b),$$

which means that (z_a, z_b) is inefficient. Therefore, there exists $(\widetilde{z}_a, \widetilde{z}_b)$ such that $\widetilde{c}_a + \widetilde{c}_b \leq w(\widetilde{\ell}_a + \widetilde{\ell}_b)$ and $\widetilde{z}_a \; P_a \; z_a$, $\widetilde{z}_b \; P_b \; z_b$.

For any population A, let R_A denote the social preferences for the economy with population A. By Laisser-Faire,

$$z^* \; R_{\{a,b\}} \; (\widetilde{z}_a, \widetilde{z}_b)$$

and by Weak Pareto,

$$(\widetilde{z}_a, \widetilde{z}_b) \; P_{\{a,b\}} \; (z_a, z_b)$$

so by transitivity,

$$z^* \; P_{\{a,b\}} \; (z_a, z_b).$$

Therefore, by Separability,

$$(z_a^*, z_b^*, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a, z_b, z_i, z_j).$$

We will now use the fact that

$$z_i' \; P_i \; z_i \; P_i \; z_a \; P_i \; z_a^*$$

and

$$z_b^* \; P_j \; z_b \; P_j \; z_j \; P_j \; z_j'.$$

Let z_i^-, $z_i'^+$ and z_a^- be such that

$$z_i'^+ \; P_i \; z_i' \; P_i \; z_i \; P_i \; z_i^- \; P_i \; z_a \; P_i \; z_a^- \; P_i \; z_a^*,$$

and similarly, let $z_j'^+$, $z_j'^{++}$ and z_b^{*+} be such that

$$z_b^{*+} \; P_j \; z_b^* \; P_j \; z_b \; P_j \; z_j \; P_j \; z_j'^{++} \; P_j \; z_j'^+ \; P_j \; z_j'.$$

Since

$$z_i'^+ \; P_i \; z_i^- \; P_i \; z_a^- \; P_i \; z_a^*$$

and $z_b^{*+} \; P_j \; z_b^*$, $z_j'^{++} \; P_j \; z_j'^+$, one can refer to Lemma 1, and conclude that

$$(z_a^-, z_b^{*+}, z_i^-, z_j'^{++}) \; P_{\{a,b,i,j\}} \; (z_a^*, z_b^*, z_i'^+, z_j'^+).$$

Similarly, since

$$z_b^{*+} \; P_j \; z_b \; P_j \; z_j \; P_j \; z_j'^{++}$$

and $z_i \; P_i \; z_i^-$, $z_a \; P_i \; z_a^-$, one obtains

$$(z_a, z_b, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a^-, z_b^{*+}, z_i^-, z_j^{\prime++}).$$

By transitivity, one then has

$$(z_a, z_b, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a^*, z_b^*, z_i^{\prime+}, z_j^{\prime+}),$$

and therefore

$$(z_a^*, z_b^*, z_i, z_j) \; P_{\{a,b,i,j\}} \; (z_a^*, z_b^*, z_i^{\prime+}, z_j^{\prime+}).$$

Separability then entails that

$$(z_i, z_j) \; R_{\{i,j\}} \; (z_i^{\prime+}, z_j^{\prime+}),$$

and by Weak Pareto one actually gets

$$(z_i, z_j) \; P_{\{i,j\}} \; (z_i', z_j').$$

From Separability again, one can finally derive the conclusion that $z \; P \; z'$ in the initial economy. ‖

Proof of Theorem 1.

(i) Let z and z' be two allocations such that $z_i \; P_i \; (0,0)$ and $z_i' \; R_i \; (0,0)$ for all i, and

$$\min_i W_i(z_i) > \min_i W_i(z_i').$$

Then, by monotonicity of preferences, one can find two allocations x, x' such that for all i, $z_i \; P_i \; x_i \; P_i \; (0,0)$, $x_i' \; P_i \; z_i'$, and there exists i_0 such that for all $i \neq i_0$

$$W_i(x_i') > W_i(x_i) > W_{i_0}(x_{i_0}) > W_{i_0}(x_{i_0}').$$

Let $(x^k)_{1 \leq k \leq n+1}$ be a sequence of allocations such that for all $i \neq i_0$,

$$x_i^i = \ldots = x_i^1 = x_i',$$
$$z_i \; P_i \; x_i^{n+1} = \ldots = x_i^{i+1} = x_i,$$

while

$$x_{i_0} = x_{i_0}^{n+1} \; P_{i_0} \; x_{i_0}^{n-1} \; P_{i_0} \ldots P_{i_0} x_{i_0}^{i_0+1} = x_{i_0}^{i_0} \; P_{i_0} \ldots P_{i_0} \; x_{i_0}^1 = x_{i_0}'.$$

One sees that for all $k \neq i_0$, $x_k^{k+1} \; P_k \; (0,0)$ and

$$W_k(x_k^k) > W_k(x_k^{k+1}) > W_{i_0}(x_{i_0}^{k+1}) > W_{i_0}(x_{i_0}^k),$$

while for all k, and all $i \neq i_0, k$, $x_i^{k+1} = x_i^k$. By Lemma 2, this implies that $x^{k+1} \; P \; x^k$ for all $k \neq i_0$, while $x^{i_0+1} = x^{i_0}$.

By Weak Pareto, $x^1 \; P \; z'$, and $z \; P \; x^{n+1}$. By transitivity, $z \; P \; z'$.

(ii) Consider allocations z and z' such that $z_i \; P_i \; (0,0)$ for all i and $(0,0) \; P_{i_0} \; z_{i_0}'$ for some i_0. By Hansson Independence, social preferences over $\{z, z'\}$ are not altered if the indifference curve for i_0 at $(0,0)$ is assumed to be such that $W_{i_0}(0,0) < \min_i W_i(z_i)$. Let z'' be such that $z_{i_0}'' = (0,0)$ and for all $i \neq i_0$, $z_i'' \; P_i \; (0,0)$ and $z_i'' \; P_i \; z'$. One has $\min_i W_i(z_i'') \leq W_{i_0}(0,0) < \min_i W_i(z_i)$ and by Theorem 1(i), $z \; P \; z''$. By Weak Pareto, $z'' \; P \; z'$. Therefore, $z \; P \; z'$. ‖

Proof of Theorem 2. Consider an allocation z such that $z_i \; R_i \; (0,0)$ for all i, and the (unique) related minimal tax function τ. Since τ is minimal, the income function $y - \tau(y)$ coincides with the envelope curve of the population's indifference curves in (y, c)-space, at z.

We first prove the following fact: Over $[0, w_m]$, the income function $y - \tau(y)$ coincides with the envelope curve of the indifference curves of the agents from the w_m subpopulation. Consider the set delimited by the envelope curve of all agents' indifference curves over this range:

$$([0, w_m] \times \mathbb{R}_+) \cap \left(\bigcup_i uc((w_i \ell_i, c_i), w_i, R_i^*) \right) = \bigcup_i \left(uc((w_i \ell_i, c_i), w_i, R_i^*) \cap ([0, w_m] \times \mathbb{R}_+) \right).$$

If the stated fact did not hold, then one would find some (y_0, c_0) such that

$$(y_0, c_0) \in \bigcup_i \left(uc((w_i\ell_i, c_i), w_i, R_i^*) \cap ([0, w_m] \times \mathbb{R}_+)\right),$$

$$(y_0, c_0) \notin \bigcup_{i: w_i = w_m} \left(uc((w_i\ell_i, c_i), w_i, R_i^*) \cap ([0, w_m] \times \mathbb{R}_+)\right).$$

The first statement means that there is some i such that

$$(y_0, c_0) \in uc((w_i\ell_i, c_i), w_i, R_i^*) \cap ([0, w_m] \times \mathbb{R}_+),$$

implying

$$uc((y_0, c_0), w_i, R_i^*) \subseteq uc((w_i\ell_i, c_i), w_i, R_i^*).$$

By the Low-Skill Diversity assumption, there is j with $w_j = w_m$ such that

$$uc((y_0, c_0), w_j, R_j^*) \subseteq uc((y_0, c_0), w_i, R_i^*),$$

and therefore

$$uc((y_0, c_0), w_j, R_j^*) \subseteq uc((w_i\ell_i, c_i), w_i, R_i^*).$$

A consequence of this inclusion is that for any (y, c) such that $(y, c) P_j^* (y_0, c_0)$, one has $(y, c) P_i^* (w_i\ell_i, c_i)$. Since

$$(y_0, c_0) \notin \bigcup_{i: w_i = w_m} \left(uc((w_i\ell_i, c_i), w_i, R_i^*) \cap ([0, w_m] \times \mathbb{R}_+)\right),$$

one must have $(w_j\ell_j, c_j) P_j^* (y_0, c_0)$, and therefore $(w_j\ell_j, c_j) P_i^* (w_i\ell_i, c_i)$. Now, this violates the incentive-compatibility condition. We obtain a contradiction, which proves the stated fact.

Let

$$W_m = w_m \min_{0 \le y \le w_m} \frac{y - \tau(y)}{y}.$$

By the above fact,

$$W_m = w_m \min \left\{ \frac{c}{y} \;\middle|\; (y, c) \in \bigcup_{i: w_i = w_m} uc((w_i\ell_i, c_i), w_i, R_i^*) \right\},$$

which equivalently reads

$$W_m = w_m \min_{i: w_i = w_m} \min \left\{ \frac{c}{y} \;\middle|\; (y, c) \in uc((w_i\ell_i, c_i), w_i, R_i^*) \right\}.$$

Now, one has, by definition:

$$W_i(z_i) = w_i \min \left\{ \frac{c}{y} \;\middle|\; (y, c) \in uc((w_i\ell_i, c_i), w_i, R_i^*) \right\}.$$

Therefore, W_m is the minimum value of $W_i(z_i)$ over the w_m subpopulation.

Similarly, for agents with a higher w, the minimum value of $W_i(z_i)$ is greater or equal to

$$W_{(w)} = w \min_{0 \le y \le w} \frac{y - \tau(y)}{y}.$$

It may be strictly greater than $W_{(w)}$ because, contrary to the case of $w = w_m$ where the Low-Skill Diversity assumption applied, the envelope curve of indifference curves for agents with wage rate $w > w_m$ may be above the envelope curve of all agents' indifference curves over the range $[0, w]$. Notice that, for any w, either

$$W_{(w)} = w - \tau(w)$$

or

$$W_{(w)} = w \frac{y_0 - \tau(y_0)}{y_0} \quad \text{for } y_0 < w.$$

Since $y - \tau(y)$ is non-decreasing, the first expression is non-decreasing in w, and this is also trivially true for the second expression. As a consequence, $W_{(w)}$ is non-decreasing in w, so that $W_{(w)} \ge W_m$, and *a fortiori* W_m is indeed the minimum value of $W_i(z_i)$ over the whole population.

We want to compare z and z', as given in the statement of the theorem. Let

$$W'_m = w_m \min_{0 \le y \le w_m} \frac{y - \tau'(y)}{y}.$$

As $\tau(0) < 0$ and $\tau'(0) \le 0$, one has $z_i \; P_i \; (0,0)$ and $z'_i \; R_i \; (0,0)$ for all i, so that Theorem 1(i) applies: Allocation z is socially preferred to z' whenever $W_m > W'_m$. This inequality is equivalent to

$$\min_{0 \le y \le w_m} \frac{y - \tau(y)}{y} > \min_{0 \le y \le w_m} \frac{y - \tau'(y)}{y},$$

or equivalently,

$$\max_{0 \le y \le w_m} \frac{\tau(y)}{y} < \max_{0 \le y \le w_m} \frac{\tau'(y)}{y}.$$

This concludes the proof. ‖

We need three lemmas for the proof of Theorem 3. These lemmas deal with the possibility of finding incentive-compatible allocations in a neighbourhood of allocations satisfying some properties.

Lemma 3. *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be an arbitrary non-decreasing function, and z an incentive-compatible (not necessarily feasible) allocation.*

(i) *Assume that $c_i < f(y_i)$ for all i. Then, for any $\varepsilon > 0$ there exists an incentive-compatible allocation z' such that $(\ell_i, c_i + \varepsilon) \; P_i \; z'_i \; P_i \; z_i$ and $c'_i \le f(y'_i)$ for all i, and $\sum_i (c'_i - y'_i) \le \sum_i (c_i - y_i) + \varepsilon$.*
(ii) *Assume that $0 < c_i \le f(y_i)$ for all i. Then for any ε such that $0 < \varepsilon < \min_i c_i$ there exists an incentive-compatible allocation z' such that $z_i \; R_i \; z'_i \; P_i \; (\ell_i, c_i - \varepsilon)$ and $c'_i \le f(y'_i)$ for all i, and $\sum_i (c'_i - y'_i) < \sum_i (c_i - y_i)$.*

Proof.

(i) Let $z^+_i = (\ell_i, c_i + \varepsilon/n)$ for all i. Let g be a function whose graph in (y,c)-space coincides with the envelope curve of the agents' indifference curves at z^+. Since $z^+_i \; P_i \; z_i$ for all i and $(y_i, c_i) \; R^*_i \; (y_j, c_j)$ for all i, j such that $y_j \le w_i$, implying $(y_i, c^+_i) \; P^*_i \; (y_j, c_j)$ for all i, j such that $y_j \le w_i$, one has $c_i < g(y_i)$ for all i. Let $\eta = \min_i (\min\{f(y_i), g(y_i)\} - c_i)_{i=1,\ldots,n}$. One has $(y_i, c^+_i) \; R^*_i \; (y_j, c_j + \eta)$ for all i, j such that $y_j \le w_i$, and therefore $(\ell_i, c_i + \varepsilon) \; P_i \; (y_j/w_i, c_j + \eta)$ for all i, j such that $y_j \le w_i$.

For any i, k in $\{1, \ldots, n\}$, let

$$v_i(y_k, c) = \begin{cases} \min\{x \ge 0 \mid (0,x) \; R^*_i \; (y_k, c)\} \text{ if } y_k \le w_i \text{ and } (y_k, c) \; R^*_i \; (0,0), \\ -\max\{y \ge 0 \mid (y,0) \; R^*_i \; (y_k, c)\} \text{ if } y_k \le w_i \text{ and } (0,0) \; P^*_i \; (y_k, c), \\ -w_i - y_k/(1+c) \text{ if } y_k > w_i. \end{cases}$$

For all y_k, this "value function" is continuous and strictly increasing in $c \ge 0$, and it represents i's preferences R^*_i over the subset of (y_k, c) such that $y_k \le w_i$.

We now focus on allocations $(y'_i, c'_i)_{i=1,\ldots,n}$ such that for some permutation π on $\{1, \ldots, n\}$ and for some vector $(d_1, \ldots, d_n) \ge 0$, one has $(y'_i, c'_i) = (y_{\pi(i)}, c_{\pi(i)} + d_{\pi(i)})$ for all i. The initial allocation $(y_i, c_i)_{i=1,\ldots,n}$ is obtained by π being the identity mapping and $(d_1, \ldots, d_n) = 0$. It is "envy-free" in the sense that for all i, k, $v_i(y_i, c_i) \ge v_i(y_k, c_k)$. This is an immediate consequence of the fact that for any i, k, $(y_i, c_i) \; R^*_i \; (y_k, c_k)$ if $y_k \le w_i$, and $v_i(y_i, c_i) \ge -w_i > v_i(y_k, c_k)$ if $y_k > w_i$.

We can then apply the "Perturbation Lemma" in Alkan, Demange and Gale (1991, p. 1029) to conclude that there is another envy-free allocation $(y'_i, c'_i)_{i=1,\ldots,n}$, for some π and some d such that $0 < d_i < \eta$ for all i.

The allocation z' defined by $z'_i = (y'_i/w_i, c'_i)$ for all i satisfies the desired properties. By envy-freeness one has $v_i(y'_i, c'_i) \ge v_i(y'_k, c'_k)$ for all i, k, and in particular for k such that $\pi(k) = i$, $v_i(y'_i, c'_i) \ge v_i(y_i, c_i + d_i) > v_i(y_i, c_i)$. Since $v_i(y_i, c_i) \ge -w_i$ for all i, this implies $v_i(y'_i, c'_i) > -w_i$ for all i. Therefore, $y'_i \le w_i$ for all i and $(y'_i, c'_i) \; R^*_i \; (y'_k, c'_k)$ for all i, k such that $y_k \le w_i$, which means that z' is incentive-compatible.

By construction, $(y'_i, c'_i) < (y_{\pi(i)}, c_{\pi(i)} + \eta)$. Since $(\ell_i, c_i + \varepsilon) \; P_i \; (y_j/w_i, c_j + \eta)$ for all i, j such that $y_j \le w_i$, it follows that $(\ell_i, c_i + \varepsilon) \; P_i \; z'_i$ for all i. In addition $v_i(y'_i, c'_i) > v_i(y_i, c_i)$ implies $z'_i \; P_i \; z_i$.

Finally,

$$\sum_i (c'_i - y'_i) \le \sum_i (c_i - y_i) + n\eta < \sum_i (c_i - y_i) + \varepsilon.$$

(ii) Let $m = \max_i (c_i - y_i)$. Let $M = \{i \mid c_i - y_i > m - \varepsilon/2\}$. Notice that for all i,

$$\frac{\varepsilon}{2} < \varepsilon < c_i \leq y_i + m.$$

For all $i \in M$, let $\hat{c}_i = y_i + m - \varepsilon/2 > 0$, and let (y'_i, c'_i) be a best bundle for i in the subset $\{(y_k, \hat{c}_k)_{k \in M}, (y_k, c_k)_{k \notin M}\}$. For $i \notin M$, let $z'_i = z_i$.

The allocation z' is incentive-compatible. Indeed, for every $i \in M$, (y'_i, c'_i) is her best bundle in $\{(y_k, \hat{c}_k)_{k \in M}, (y_k, c_k)_{k \notin M}\}$ and therefore also in $\{(y_k, c'_k)_{k \in M}, (y_k, c_k)_{k \notin M}\} \subseteq \{(y_k, \hat{c}_k)_{k \in M}, (y_k, c_k)_{k \notin M}\}$. And since $\hat{c}_k < c_k$ for $k \in M$, the fact that for any $i \notin M$, (y'_i, c'_i) is a best bundle in the subset $\{(y_k, c_k)_{k \in M}, (y_k, c_k)_{k \notin M}\}$ entails that it is *a fortiori* a best bundle in $\{(y_k, c'_k)_{k \in M}, (y_k, c_k)_{k \notin M}\} \subseteq \{(y_k, \hat{c}_k)_{k \in M}, (y_k, c_k)_{k \notin M}\}$.

For every $i \in M$, $z_i \; R_i \; z'_i \; P_i \; (\ell_i, c_i - \varepsilon)$, because $(y'_i, c'_i) \; R^*_i \; (y_i, \hat{c}_i)$ and $\hat{c}_i = y_i + m - \varepsilon/2 \geq y_i + (c_i - y_i) - \varepsilon/2 > c_i - \varepsilon$. For every $i \notin M$, $z_i = z'_i \; P_i \; (\ell_i, c_i - \varepsilon)$.

The fact that $\hat{c}_i \leq c_i$ for all $i = 1, \ldots, n$ implies that $\hat{c}_i \leq f(y_i)$ for all i and thereby guarantees that $c'_i \leq f(y'_i)$ for all i.

Finally,

$$\begin{aligned}\sum_i (c'_i - y'_i) &= \sum_{i \in M} (c'_i - y'_i) + \sum_{i \notin M} (c'_i - y'_i) \\ &\leq \sum_{i \in M} (m - \varepsilon/2) + \sum_{i \notin M} (c_i - y_i) < \sum_i (c_i - y_i). \quad \|\end{aligned}$$

Lemma 4. *Let A be the set of allocations z which are feasible, incentive-compatible, and such that $z_i \; P_i \; (0,0)$ for all i. Let B be the set of allocations z which are feasible, incentive-compatible, and such that $z_i \; R_i \; (0,0)$ for all i. Let U_i be a continuous representation of R_i, and let $U(z)$ denote $(U_1(z_1), \ldots, U_n(z_n))$. If A is not empty, then for any $z \in B \setminus A$, there is $z' \in A$ such that $U(z')$ is arbitrarily close to $U(z)$.*

Proof. Let $z \in B \setminus A$ and assume $A \neq \emptyset$.

(1) If $\sum_i (c_i - y_i) < 0$, then by Lemma 3(i), for any $\varepsilon > 0$ there exists an incentive-compatible allocation z' such that $(c_i + \varepsilon, \ell_i) \; P_i \; z'_i \; P_i \; z_i$ for all i, and $\sum_i (c'_i - y'_i) \leq \sum_i (c_i - y_i) + \varepsilon$. If $\varepsilon < \left|\sum_i (c_i - y_i)\right|$, then z' is feasible and belongs to A. Since ε is arbitrarily small, $U(z')$ is arbitrarily close to $U(z)$.

(2) If $\sum_i (c_i - y_i) = 0$:
(2-i) If $\max_i (c_i - y_i) = 0$, then $c_i = y_i$ for all i. Let τ be the minimal tax implementing z (*i.e.* $y - \tau(y)$ coincides on $[0, \max_i w_i]$ with the envelope curve of the agents' indifference curves in (y, c) space).
(2-i-a) If there exists y_0 such that $\tau(y_0) > 0$, then consider agent i such that $z_i \; I_i \; (\frac{y_0}{w_i}, y_0 - \tau(y_0))$, and let $z_i^- = (\frac{y_0}{w_i}, y_0 - \tau(y_0))$. One has $c_i^- - y_i^- = -\tau(y_0) < 0$, so that the allocation (z_i^-, z_{-i}) may be dealt with as in case (1) in order to find $z' \in A$ with arbitrarily close utilities.
(2-i-b) If $\tau(y) \leq 0$ for all y, let $J = \{i \mid z_i \; P_i \; (0,0)\}$ and $K = \{i \mid z_i \; I_i \; (0,0)\}$. The set J is non-empty, otherwise A would be empty (because, then, for all i, $(0,0)$ is a best allocation for R^*_i in $\{(y,c) \mid c \leq y\}$). And $c_i > 0$ for all $i \in J$. Define $z_i^- = (0,0)$ for all $i \in K$ and $z_i^- = z_i$ for all $i \in J$. The allocation z^- is feasible (recall that $c_i = y_i$ for all i), incentive-compatible and Pareto-indifferent to z. Let f be a function whose graph, in (y,c)-space, coincides with the envelope curve of the indifference curves at z^- of agents from K. Since for any $i \in K$, $j \in J$, either $(0,0) \; R^*_i \; (y_j, c_j)$ or $y_j > w_i$, one has $c_j \leq f(y_j)$ or $y_j > \max_{i \in K} w_i$ for all $j \in J$. By extending $f(.)$ over the interval $(\max_{i \in K} w_i, +\infty)$, one can easily have $c_j \leq f(y_j)$ for all $j \in J$. By Lemma 3(ii), for any η such that $0 < \eta < \min_{j \in J} c_j$ there exists $\left(z'_j\right)_{j \in J}$ which is incentive-compatible for the subpopulation J and such that $z_j \; R_j \; z'_j \; P_j \; (\ell_j, c_j - \eta)$ and $c'_j \leq f(y'_j)$ for all $j \in J$, and $\sum_{j \in J} \left(c'_j - y'_j\right) < \sum_{j \in J} \left(c_j - y_j\right)$. By taking η sufficiently small one can have $z'_j \; P_j \; (0,0)$ for all $j \in J$. Let g be a function whose graph, in (y,c)-space, coincides with the envelope curve of the indifference curves at z' of agents from J. One has $c_i^- = 0 < g(y_i^-) = g(0)$ for all $i \in K$. Let $0 < \varepsilon < \sum_{j \in J} \left(c_j - y_j\right) - \sum_{j \in J} \left(c'_j - y'_j\right)$. By Lemma 3(i), there exists $\left(z'_i\right)_{i \in K}$ which is incentive-compatible and such that $(c_i^- + \varepsilon, \ell_i^-) \; P_i \; z'_i \; P_i \; z_i^-$ and $c'_i \leq g(y'_i)$ for all $i \in K$, and $\sum_{i \in K} (c'_i - y'_i) \leq \sum_{i \in K} (c_i - y_i) + \varepsilon$. The allocation z' is in A and is as desired.
(2-ii) If $\max_i (c_i - y_i) > 0$: Let

$$M = \left\{ i \in \{1, \ldots, n\} \mid c_i - y_i = \max_j \left(c_j - y_j\right) \right\}.$$

Since $\sum_j \left(c_j - y_j\right) \leq 0$, $M \subsetneq \{1, \ldots, n\}$.

(2-ii-a) If there is $i \in M$ such that $(y_i, c_i)\ I_i^*\ (y_{j_0}, c_{j_0})$ for some $j_0 \notin M$, then let $(y_i^-, c_i^-) = (y_{j_0}, c_{j_0})$. Let $0 < \varepsilon < c_i - y_i - \left(c_i^- - y_i^-\right)$. By Lemma 3(i) there exists an incentive-compatible allocation z' such that $(c_i^- + \varepsilon, \ell_i^-)\ P_i\ z_i'\ P_i\ z_i^-$ and $(c_j + \varepsilon, \ell_j)\ P_j\ z_j'\ P_j\ z_j$ for all $j \neq i$, and

$$\sum_i \left(c_i' - y_i'\right) \leq c_i^- - y_i^- + \sum_{j \neq i} \left(c_j - y_j\right) + \varepsilon < \sum_j \left(c_j - y_j\right) \leq 0.$$

(2-ii-b) If there is $i \in M$ such that $z_i\ I_i\ (0,0)$, then let $z_i^- = (0,0)$. The rest of the argument is as in case a.

(2-ii-c) If none of the cases a–b holds, then for all $i \in M$, $z_i\ P_i\ (0,0)$ and $(y_i, c_i)\ P_i^*\ (y_j, c_j)$ (or $y_j > w_i$) for all $j \notin M$. This case is dealt with similarly as the case i-b, by taxing agents from M at the benefit of the others. ‖

Lemma 5. *If there is a feasible allocation z such that $z_i\ P_i\ (0,0)$ for all i, then there is a feasible and incentive-compatible allocation z such that $z_i\ P_i\ (0,0)$ for all i.*

Proof. Let $z^* = ((0,0), \ldots, (0,0))$. It is feasible and incentive-compatible. Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a function whose graph coincides on $[0, \max_i w_i]$ with the envelope curve in (y, c)-space of individual indifference curves at z^*. If $f(y) \geq y$ for all $y \in [0, \max_i w_i]$, then z^* is Pareto-efficient and there is no feasible allocation z such that $z_i\ P_i\ (0,0)$ for all i. Therefore, $f(y_0) < y_0$ for some $y_0 \leq \max_i w_i$. Let i_0 be an agent such that $(0,0)\ I_{i_0}^*\ (y_0, f(y_0))$. The allocation z' such that $(y_{i_0}', c_{i_0}') = (y_0, f(y_0))$ and $(y_j', c_j') = (0,0)$ for all $j \neq i_0$ is incentive-compatible and such that $\sum_i c_i' < \sum_i y_i'$. By Lemma 3, there exists another feasible and incentive-compatible allocation z such that for all i, $z_i\ P_i\ z_i'$. ‖

Proof of Theorem 3. Consider an optimal allocation z^*. Suppose there is i such that $(0,0)\ P_i\ z_i^*$. Since by Lemma 5, there is a feasible and incentive-compatible allocation z such that $z_i\ P_i\ (0,0)$ for all i, then by Theorem 1(ii) $z\ P\ z^*$ and this contradicts the assumption that z^* is optimal. Therefore, one must have $z_i^*\ R_i\ (0,0)$ for all i. There is a (unique) minimal tax function τ such that the income function $y - \tau(y)$ coincides with the envelope curve of the population's indifference curves in (y, c)-space, at z^*. In particular, $\tau(0) \leq 0$.

Let the sets A and B be defined as in Lemma 4. We have just proved that $z^* \in B$. We now show that

$$\min_i W_i(z_i^*) = \max_{z \in B} \min_i W_i(z_i).$$

Suppose not. This may be either because $\max_{z \in B} \min_i W_i(z_i)$ does not exist, or because $\min_i W_i(z_i^*) < \max_{z \in B} \min_i W_i(z_i)$. In both cases, there exists $z \in B$ such that $\min_i W_i(z_i^*) < \min_i W_i(z_i)$. By Lemma 4, there exists $z' \in A$ such that $W(z')$ is arbitrarily close to $W(z)$, so $\min_i W_i(z_i^*) < \min_i W_i(z_i')$. This implies $z'\ P\ z^*$, which contradicts the assumption that z^* is optimal.

The fact that $\min_i W_i(z_i^*) = \max_{z \in B} \min_i W_i(z_i)$ means that z^* is obtained by a tax which, among all feasible taxes such that $\tau(0) \leq 0$ (and $y - \tau(y)$ is non-decreasing), maximizes $\min_i W_i(z_i)$ at the resulting allocation z. It remains to show that there is no restriction in adding the other conditions stated in the theorem, and that under these conditions maximizing $\min_i W_i(z_i)$ is equivalent to maximizing $w_m - \tau(w_m)$.

By the proof of Theorem 2,

$$\min_i W_i(z_i^*) = W_m = w_m \min_{0 \leq y \leq w_m} \frac{y - \tau(y)}{y}.$$

At a laisser-faire allocation z^{LF}, one has $W_i(z_i^{LF}) \geq w_i$ for all i, so

$$\min_i W_i(z_i^{LF}) \geq w_m.$$

A fortiori, at the optimum,

$$W_m = \min_i W_i(z_i^*) \geq w_m.$$

Let a new tax be defined by

$$\tau^*(y) = \max\{\tau(y), w_m - W_m\}.$$

This tax function is feasible, because it cuts all subsidies greater than a constant, $W_m - w_m \geq 0$, so that no agent's tax may decrease (and no subsidy increase), even after she adjusts her choice. Moreover,

$$\begin{aligned}\min_{0\leq y\leq w_m} \frac{y-\tau^*(y)}{y} &= \min_{0\leq y\leq w_m} \left\{ \frac{y-\tau(y)}{y}, \frac{y-(w_m-W_m)}{y} \right\} \\ &= \min\left\{ \min_{0\leq y\leq w_m} \frac{y-\tau(y)}{y}, \min_{0\leq y\leq w_m} 1+\frac{W_m-w_m}{y} \right\} \\ &= \min\left\{ \frac{W_m}{w_m}, 1+\frac{W_m-w_m}{w_m} \right\} \\ &= \frac{W_m}{w_m} = \min_{0\leq y\leq w_m} \frac{y-\tau(y)}{y}.\end{aligned}$$

The tax function τ^* need not be minimal. Let z^{**} be an allocation obtained with τ^*, and chosen so that $z_i^{**} = z_i^*$ for all i such that $\tau(y_i^*) \geq w_m - W_m$. Let τ^{**} be the corresponding minimal tax function. One has $\tau^{**} \leq \tau^*$, and therefore

$$\min_i W_i(z_i^{**}) = w_m \min_{0\leq y\leq w_m} \frac{y-\tau^{**}(y)}{y} \geq w_m \min_{0\leq y\leq w_m} \frac{y-\tau^*(y)}{y} = W_m.$$

In addition, since $\tau^* \geq \tau$, necessarily $z_i^* \; R_i \; z_i^{**}$ for all i, implying $W_i(z_i^{**}) \leq W_i(z_i^*)$ for all i. Therefore

$$\min_i W_i(z_i^{**}) \leq \min_i W_i(z_i^*) = W_m,$$

and then

$$\min_i W_i(z_i^{**}) = W_m.$$

The allocation z^{**} has been constructed so that for every i, either $z_i^{**} = z_i^*$ and $\tau^*(y_i^{**}) = \tau(y_i^*)$, or $\tau^*(y_i^{**}) > \tau(y_i^*)$. Suppose there is i such that $\tau^*(y_i^{**}) > \tau(y_i^*)$. Then one has $\sum_i \tau^*(y_i^{**}) > 0$, meaning that $\sum_i c_i^{**} < \sum_i y_i^{**}$. By Lemma 3(i), this inequality contradicts the fact that z^{**} maximizes $\min_i W_i$ over B. Therefore, there is no i such that $\tau^*(y_i^{**}) > \tau(y_i^*)$, and for all i, $z_i^{**} = z_i^*$. This means that τ^* implements z^*.

By construction, for all $y \geq 0$, $\tau^*(y) \geq w_m - W_m$, and as shown above, for all $y \leq w_m$,

$$W_m \leq w_m \frac{y-\tau^*(y)}{y},$$

so

$$w_m - W_m \leq \tau^*(y) \leq y\left(1-\frac{W_m}{w_m}\right).$$

For $y = w_m$, this entails: $\tau^*(w_m) = w_m - W_m$. Therefore, $\tau^*(y) \geq \tau^*(w_m)$ for all $y \geq 0$.

Moreover, for all $y \in [0, w_m]$,

$$\tau^*(y) \leq y\left(1-\frac{W_m}{w_m}\right) = y\left(1-\frac{w_m-\tau^*(w_m)}{w_m}\right) = y\left(\frac{\tau^*(w_m)}{w_m}\right),$$

entailing $\tau(0) \leq 0$ and, for $y \in (0, w_m]$,

$$\frac{\tau^*(y)}{y} \leq \frac{\tau^*(w_m)}{w_m}.$$

Since $W_m = w_m - \tau^*(w_m)$, maximizing W_m is equivalent to maximizing $w_m - \tau^*(w_m)$. ‖

Acknowledgements. We thank Dilip Abreu, Faruk Gul, Eric Maskin, Pierre Pestieau, John Roemer and H. Peyton Young for very useful discussions, two anonymous referees, Juuso Välimäki, and seminar participants at CORE, U. de Lyon, U. de Montpellier, U. de Pau, Yale U. and the Fourgeaud-Roy seminar (Paris) for their comments, and Giunia Gatta for her help with the presentation of the paper. The paper was written while the second author was a member of the Institute for Advanced Study (Princeton, NJ); the exceptional research atmosphere at the Institute is gratefully acknowledged. This paper presents research results of the Belgian Programme on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming.

REFERENCES

ALKAN, A., DEMANGE, G. and GALE, D. (1991), "Fair Allocation of Indivisible Goods and Criteria of Justice", *Econometrica*, **59**, 1023–1040.

ARMSTRONG, M. (1996), "Multiproduct Nonlinear Pricing", *Econometrica*, **64**, 51–75.

ARROW, K. J. (1951) *Social Choice and Individual Values* (New York: Wiley).

ATKINSON, A. B. (1973), "How Progressive Should Income Tax Be?", in M. Parkin and A. R. Nobay (eds.) *Essays in Modern Economics* (London: Longmans).

ATKINSON, A. B. (1995) *Public Economics in Action* (Oxford: Clarendon Press).

BESLEY, T. and COATE, S. (1995), "The Design of Income Maintenance Programs", *Review of Economic Studies*, **62**, 187–221.

BOADWAY, R. and KEEN, M. (2000), "Redistribution", in A. B. Bourguignon and F. Bourguignon (eds.) *Handbook of Income Distribution*, Vol. 1 (Amsterdam: North-Holland).

BOADWAY, R., MARCHAND, M., PESTIEAU, P. and RACIONERO, M. (2002), "Optimal Redistribution with Heterogeneous Preferences for Leisure", *Journal of Public Economic Theory*, **4** (4), 475–498.

BOSSERT, W., FLEURBAEY, M. and VAN DE GAER, D. (1999), "Responsibility, Talent, and Compensation: A Second-Best Analysis", *Review of Economic Design*, **4**, 35–55.

BREWER, M. (2000), "Comparing In-Work Benefits and Financial Work Incentives for Low-Income Families in the U.S. and the U.K." (WP #00/16 Institute for Fiscal Studies).

CHONE, P. and LAROQUE, G. (2001), "Optimal Incentives for Labor Force Participation", *Journal of Public Economics* (forthcoming).

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

DWORKIN, R. (1981), "What is Equality? Part 2: Equality of Resources", *Philosophy and Public Affairs*, **10**, 283–345.

EBERT, U. (1992), "A Reexamination of the Optimal Nonlinear Income Tax", *Journal of Public Economics*, **49**, 47–73.

ELLWOOD, D. T. (2000), "The Impact of the Earned Income Tax Credit and Social Policy Reforms on Work, Marriage, and Living Arrangements", *National Tax Journal*, **53**, 1063–1075.

FLEURBAEY, M. (2001), "The Pazner–Schmeidler Social Ordering: A Defense", *Review of Economic Design* (forthcoming).

FLEURBAEY, M. and MANIQUET, F. (1996*a*), "Fair Allocation with Unequal Production Skills: The No-Envy Approach to Compensation", *Mathematical Social Sciences*, **32**, 71–93.

FLEURBAEY, M. and MANIQUET, F. (1996*b*), "Utilitarianism Versus Fairness in Welfare Economics", in M. Salles and J. A. Weymark (eds.) *Justice, Political Liberalism and Utilitarianism: Themes from Harsanyi and Rawls* (Cambridge: Cambridge University Press) (forthcoming).

FLEURBAEY, M. and MANIQUET, F. (1999), "Fair Allocation with Unequal Production Skills: The Solidarity Approach to Compensation", *Social Choice and Welfare*, **16**, 569–583.

FLEURBAEY, M. and MANIQUET, F. (2001), "Fair Social Orderings", *Economic Theory* (forthcoming).

FLEURBAEY, M. and MANIQUET, F. (2005), "Fair Orderings with Unequal Production Skills", *Social Choice and Welfare*, **24**, 93–127.

FLEURBAEY, M., SUZUMURA, K. and TADENUMA, K. (2003), "Arrovian Aggregation in Economic Environments: How Much Should We Know About Indifference Surfaces?", *Journal of Economic Theory* (forthcoming).

FLEURBAEY, M. and TRANNOY, A. (2003), "The Impossibility of a Paretian Egalitarian", *Social Choice and Welfare*, **21**, 243–264.

HANSSON, B. (1973), "The Independence Condition in the Theory of Social Choice", *Theory and Decision*, **4**, 25–49.

MANIQUET, F. and SPRUMONT, Y. (2004), "Fair Production and Allocation of an Excludable Nonrival Good", *Econometrica*, **72**, 627–640.

MATTHEWS, S. and MOORE, J. H. (1987), "Optimal Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening", *Econometrica*, **55**, 441–467.

MIRRLEES, J. (1971), "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, **38**, 175–208.

MIRRLEES, J. (1974), "Notes on Welfare Economics, Information and Uncertainty", in M. Balch, D. McFadden and S. Wu (eds.) *Essays in Equilibrium Behavior under Uncertainty* (Amsterdam: North-Holland) 243–258.

PAZNER, E. (1979), "Equity, Nonfeasible Alternatives and Social Choice: A Reconsideration of the Concept of Social Welfare", in J. J. Laffont (ed.) *Aggregation and Revelation of Preferences* (Amsterdam: North-Holland).

ROCHET, J.-C. and CHONE, P. (1998), "Ironing, Sweeping and Multidimensional Screening", *Econometrica*, **66**, 783–826.

SADKA, E. (1976), "On Income Distribution Incentive Effects and Optimal Income Taxation", *Review of Economic Studies*, **43**, 261–268.

SALANIE, B. (2002), "Optimal Demogrants with Imperfect Tagging", *Economics Letters*, **75**, 319–324.

SAMUELSON, P. A. (1947) *Foundations of Economic Analysis* (Cambridge, MA: Harvard University Press).

SEADE, J. K. (1977), "On the Shape of Optimal Tax Schedules", *Journal of Public Economics*, **7**, 203–236.

STIGLITZ, J. E. (1987), "Pareto Efficient and Optimal Taxation and the New New Welfare Economics", in A. J. Auerbach and M. Feldstein (eds.) *Handbook of Public Economics*, Vol. 2 (Amsterdam: North-Holland).

TUOMALA, M. (1990) *Optimal Income Tax and Redistribution* (Oxford: Oxford University Press).