Optimal Redistribution with Heterogeneous Preferences for Leisure

Robin BOADWAY[1], Maurice MARCHAND[2], Pierre PESTIEAU[3]
and María del Mar RACIONERO[4]

December 2000

Abstract

This paper examines the properties of the optimal nonlinear income tax when preferences are quasilinear in leisure and heterogeneous. Individuals differ in their ability and in their preferences for leisure. The government seeks to redistribute income. It can perfectly observe the level of endogenous income but cannot observe either ability or preferences. The heterogeneity of preferences leads to problems of comparability between individual utilities which challenge the design of redistributive schemes. In particular, we analyze the consequences of adopting a utilitarian social welfare function where the government is allowed to give different weights to individuals with different preferences. Under this particular social objective and given the quasilinearity of preferences, we are able to obtain closed-form solutions for the marginal tax rates and to examine the progressivity of the tax system according to the weights used.

Keywords: Optimal income taxation, quasi-linear preferences, asymmetric information

JEL Classification: H21, H41

[1]Queen's University, Canada.
[2]CORE, Université catholique de Louvain, Belgium.
[3]CREPP, Université de Liège, and CORE, Université catholique de Louvain, Belgium.
[4]School of Economics, Deakin University, Australia.

Research of the last author was supported by a "Marie Curie" fellowship of the European Community. We are grateful to Jean Hindriks, François Maniquet and John Roemer for their helpful comments. We have also benefited from comments of participants at the Public Economic Theory 2000 Conference in Warwick (July 2000) and the 56th Congress of the International Institute of Public Finance in Seville (August 2000).

1 Introduction

One of the most common assumptions in the standard nonlinear income taxation literature is that individuals have common preferences with respect to consumption and leisure. In this paper, we relax this assumption to allow for heterogeneous preferences. In particular, we consider individuals who differ both in productive ability and in preferences for leisure. The heterogeneity of preferences raises ethical questions which challenge the design of redistributive schemes. This is so even when individual characteristics are common knowledge. Two individuals that have the same ability and different disutility of labour will earn, at the laissez-faire, different gross incomes, with individuals having higher preference for leisure earning less. However, the government may be reluctant to redistribute towards this kind of individual, since they are also enjoying a higher quantity of leisure.

A strand of the social choice literature has devoted attention to the study of ethical principles for redistribution when individuals differ in several characteristics (see Fleurbaey and Maniquet (1999) for a survey on this subject). Among characteristics that make individuals heterogeneous, some may be deemed "relevant" and call for compensation, while others may fall under the responsibility of individuals and be considered "irrelevant" from the point of view of redistribution. The principles for redistribution analyzed in this literature are, accordingly, of two basic types: compensation for "relevant" characteristics and responsibility for "irrelevant" characteristics. The first aims at neutralizing the influence on individuals' outcomes of the characteristics that deserve compensation while the second calls for abstaining from acting about inequalities due to the other characteristics. These two objectives cannot, in most cases, be satisfied simultaneously. This incompatibility leads to an interesting ethical dilemma since the redistributive schemes will necessarily be a compromise between the two principles.

The incompatibility between ethical principles for redistribution arises even when the government has at its disposal all the information about individual characteristics. In models similar to ours, but in which individual utility is quasi-linear in consumption and quadratic in effort, Roemer (1998) investigates how redistributive policy should be designed according to the equal-opportunity principle he proposes to adopt. To satisfy this principle, individual welfare should be equalized across skills for those persons who supply the same *degree of effort*. In Roemer's definition, two individuals with different skills supply the same *degree of effort* if they exert *levels of effort* that are at the same percentile of the two distributions that characterize the *effort levels* of the two subgroups of individuals having these respective skills. However, the redistributive policy that satisfies this principle generally depends upon the percentile chosen, which calls for some compromise. With the compromise he proposes, Roemer shows that his equal-opportunity objective generates an amount of redistribution between what the utilitarian and Rawlsian principles recommend.

Some recent contributions have also dealt with optimal redistribution when there is asymmetric information about individuals' ability and preferences for leisure. Fleurbaey and Maniquet

(1998) use an ordinal approach. They build social objective functions which do not require any utility comparison nor any cardinal measurement of utility. These social welfare functions are required to satisfy compensation for inequalities in skills and responsibility for preferences. They obtain ethical foundations for the non-welfarist social objective of maximizing the minimum disposable income when agents with zero productivity are present in the population. However, this result does not directly apply when all agents have positive productivities.

In the present paper we depart somewhat from the above strand in the social choice literature. We adopt a utilitarian social welfare function (SWF) where different weights are allowed to be assigned to individuals with different preferences for leisure (this amounts to using different cardinalizations of individual utility functions). Our main purpose is to investigate how maximizing such a SWF affects the optimal redistributive policy, in particular the progressivity of the income tax schedule. Except for the fact that we do not consider SWFs that assign higher weight to epicurian than to hard-working persons (see later), we do not impose any a priori restriction on those weights. Societies can indeed hold different views about the way individuals with different preferences for leisure should be treated, and we do not want here to enter this debate.

As is standard in the literature on optimal income taxation, we assume that the government cannot observe individual characteristics. To simplify the presentation, we consider the case where skill can only take two values, and likewise for preference for leisure. Given the informational asymmetry, the redistributive policy must satisfy the so-called incentive-compatibility (or self-selection) constraints: the fiscal treatments must be designed so that no individual is enticed to apply for the fiscal treatment intended for another type of individuals. In the standard model of optimal income taxation with a utilitarian SWF, the binding incentive-compatibility constraints are generally easy to identify: high-skill individuals must be prevented from claiming that they are low-skill (which corresponds to the so-called downwards incentive-compatibility constraint). This is no more the case in our model with two individual characteristics. As minimal restrictions are imposed on the relative weights assigned to epicurian and hard-working persons in our SWF, either one of the downwards and upwards incentive-compatibility constraints that relate two different types of individuals may be binding at the optimal solution. This significantly complicates the analysis and calls for some simplification in the specification of the model.

Problems of optimal income taxation with two unknown characteristics are indeed notoriously difficult to solve. In order to make the model tractable we assume that individual preferences can be represented by a utility function that is quasi-linear in leisure. The two-characteristics model then reduces to a uni-dimensional one. The realism of this preference representation is certainly questionable, but it makes the problem tractable and enables us to obtain interesting insights into the progressivity of the optimal tax schedule. This assumption of quasi-linearity was also adopted by Lollivier and Rocher (1983) to solve the optimal income tax problem with uniform preference in the continuous-skill case, and by Weymark (1986a, 1986b, 1987) in the discrete case. A synthesis of this literature and further results are provided in Boadway, Cuff and Marchand (2000).

The main result of the paper can best be understood by reference to the case where the epicurian and hard-working individuals would be living on two separate islands. Then, incentive-compatibility constraints would only involve mimickers claiming to be identical to individuals living on the same island, i.e. sharing the same preference for leisure. Accordingly, with a utilitarian SWF, progressive taxation would prevail on each island (i.e., within each preference group): income would be redistributed from high-skill to low-skill individuals. And this qualitative result would not be affected by any transfer between the two islands. However, in our economy, epicurian and hard-warking individuals interact with each other, and some incentive-compatibility constrains are then related to individuals in a preference group who claim to be individuals belonging to the other group. As a consequence, we show that the standard result of progressive taxation no longer holds for plausible weights assigned to epicurian and hard-working people in the utilitarian SWF. In particular, if the relative weights are such that there is no transfer between the two preference groups, income redistribution is progressive within one preference group and regressive within the other.

For the sake of realism, we analyze the situation where individuals with high skill and high preferences for leisure can never be distinguished from individuals with low skill and low preferences for leisure. In other words, these two types of individuals share the same indifference curves in the space of pre-tax and post-tax incomes. In the public debate on redistributive policy it is indeed frequently argued that the incentive to make transfers to hard-working low-skill persons should be mitigated by the fact that these transfers are also benefiting less deserving epicurian high-skill persons, since their similar pre-tax incomes make them difficult to separate.[1] This imposes a strong constraint on the design of optimal taxes that it is most relevant to take into account.

The paper is organized as follows. In section 2, we describe the model and provide the characterization of the individuals' behaviour. We assume that low-ability individuals with low disutility of labour cannot be distinguished from high-ability individuals with high disutility of labour. We analyze the constraints faced by the government under this informational setting and characterize the second best problem in general. In section 3.1, we consider a weighted utilitarian objective where the government gives different weights given to individuals with different preferences. We investigate the optimal nonlinear income tax schedule. In particular, we explicitly analyze the pattern of binding self-selection constraints and the progressivity of the tax system for different weights given to individuals with different preferences in the SWF. In section 3.2 we then extend the analysis to the case where the social welfare function is of a weighted maximin type. Section 4 concludes the paper. In the main text we abstract from bunching situations: these are studied in the appendix to the paper.

[1] In the case of continuous skills, such impossibility of separating some persons would necessarily apply over a range of skills and preferences.

2 The model

2.1 The basic model

We consider individuals that differ in two characteristics: the level of ability w and the preference for leisure α. The consumption and labour supply of an individual are denoted by c and ℓ, respectively. We assume that the preferences of an individual with taste for leisure α can be described by a quasilinear utility function:[2]

$$u(c) - \alpha\ell, \tag{1}$$

where $u(.)$ is a continuous, differentiable, strictly increasing and strictly concave function with $u(0) = 0$, $u'(0) = \infty$ and $\lim_{c\to\infty} u'(c) = 0$. Since the pretax income of an individual with ability w is $y = wl$, the quasilinear utility can be rewritten in terms of consumption and pretax income:

$$u(c) - \frac{\alpha}{w}y. \tag{2}$$

We assume that the ability can take two positive values w_ℓ and w_h, with $w_h > w_\ell$. Ability h individuals have thus a higher productivity in the private labour market than ability ℓ individuals. The proportion of individuals with ability w_j is p_j $(j = \ell, h)$, with $p_\ell + p_h = 1$. The preference-for-leisure parameter α can also take two positive values α_h and α_ℓ with $\alpha_h > \alpha_\ell$. So individuals with preferences α_h have a higher disutility of labour. The proportion of individuals with taste parameter α_k is π_k $(k = h, \ell)$, with $\pi_h + \pi_\ell = 1$. We have then four types of individuals differing in two characteristics: ability and preference for leisure. Assuming that these two characteristics are distributed independently in the population, the proportions of each type of individual is given in Table 1. Without loss of generality, we normalize the total population to 1.

		Disutility for labour α_i	
		α_h	α_ℓ
Ability w_j	w_ℓ	$p_\ell\pi_h$	$p_\ell\pi_\ell$
	w_h	$p_h\pi_h$	$p_h\pi_\ell$

Table 1 Proportions of the various types of individuals in the population

The government knows the distribution of both ability and preferences in the population, given by Table 1, and observes the level of pretax income. However, it cannot observe either the ability and preferences of a particular individual or his labour supply. As noted in the introduction, we will focus on the case in which the intermediate individuals are, in addition, indistinguishable. Given the quasilinearity of the utility function, this is the case when $w_\ell/\alpha_\ell = w_h/\alpha_h$. That is, the high productivity individuals with high preference for leisure cannot be distinguished from the low productivity individuals with low preference for leisure. We will denote by $\widehat{w}$ the ratio

[2] This comes from the utility function: $u(c) + \alpha(L - \ell)$, where L denotes the exogenous time endowment.

of an individual's ability w to his preference parameter α. We are then left with three groups of distinguishable individuals characterized by:

$$\widehat{w}_1 = \frac{w_\ell}{\alpha_h} < \widehat{w}_2 = \frac{w_\ell}{\alpha_\ell} = \frac{w_h}{\alpha_h} < \widehat{w}_3 = \frac{w_h}{\alpha_\ell},$$

with proportions in the population

$$\widehat{\pi}_1 = p_\ell \pi_h, \quad \widehat{\pi}_2 = p_h \pi_h + p_\ell \pi_\ell \quad \text{and} \quad \widehat{\pi}_3 = p_h \pi_\ell,$$

respectively. The quasilinear utility expressed in terms of pretax income and consumption (or post tax income) of group-i individual is then written as $u(c) - y/\widehat{w}_i$. However it will be more convenient to work with a monotone transformation of these functions, namely:

$$U_i(c, y) = \widehat{w}_i u(c) - y. \tag{3}$$

The marginal rate of substitution of a group-i individual at bundle (y, c) is:

$$MRS_i(y, c) = \frac{1}{\widehat{w}_i u'(c)}.$$

This represents the slope of the indifference curve of a group-i individual at income-consumption bundle (y, c) in the (y, c)-space. As the slope is independent of income due to the quasilinearity assumption, the indifference curves are horizontally parallel. Individuals with low ability and high preference for leisure (i.e., group-1 individuals) have the steepest curves. The individuals with high ability and low preference for leisure (i.e., group-3) have the flattest curves. At any point in the (y, c)-space, the indifference curve of group-2 individuals (either high-ability high-preference for leisure individuals or low-ability low-preference individuals) has a slope that lies between those of individuals of groups 1 and 3. This implies that the single crossing (or monotonicity) property holds (i.e., the indifference curves cross only once) and that at any point (y, c) the slopes of indifference curves decline with index i. The utility of individuals increases as they move to indifference curves left- or upwards.

2.2 The second-best problem

The second-best problems in standard models of nonlinear income taxation include self-selection or incentive-compatibility constraints. Their objective is to ensure that no individual of one group has incentives to mimic an individual of another group. Weymark (1986ab, 1987) analyzes the optimal nonlinear income tax problem for a finite population where individuals differ in productive ability and have homogeneous quasilinear preferences. He considers a weighted utilitarian social welfare function where the weights given to individual utilities are restricted to be such that only downwards self-selection relating constraint adjacent individuals need be taken into account. As mentioned in the introduction, the present model is however more complex. First, individuals differ in both ability and preference for leisure. Hence, we will also have to take into account the self-selection constraints relating pairs of individuals with different preferences. Second, we consider social objectives that differ from the standard utilitarian (i.e., the direct sum of individual utilities) and maxi-min objectives. Indeed we allow the government to give different cardinalization weights to the utilities of individuals with different preferences for leisure with no a priori restriction on these relative weights. In the following, we must therefore take into account both the upwards and downwards self-selection constraints relating pairs of individuals. The relative level of weights chosen will determine the set of self-selection constraints that will be binding at the optimum. In the following we identify the regime prevailing at the optimum with this set of constraints.

As a first step, let us therefore define the different types of constraints the government faces when choosing the tax imposed on each type of individual.[3] The first one is the government budget constraint:

$$\sum_i \widehat{\pi}_i \left(y_i - c_i\right) \geq R, \tag{4}$$

where R is an exogenous revenue requirement. The second is the set of incentive-compatibility or self-selection constraints due to the asymmetry of information. These require that the commodity bundle intended for each individual must be optimal for him or her when faced with the other bundles:

$$U_i\left(c_i, y_i\right) \geq U_i\left(c_j, y_j\right) \qquad \forall i, j. \tag{5}$$

Since the single crossing property holds, it suffices to take into account the self-selection constraints that relate pairs of adjacent individuals. Among these, there are the standard downwards self-selection constraints that are designed to ensure that an individual of group i has no incentive to mimic an individual of group $i-1$:

$$U_i\left(c_i, y_i\right) \geq U_i\left(c_{i-1}, y_{i-1}\right) \qquad i = 2, 3. \tag{6}$$

However, there are also here upwards self-selection constraints that are less common in the literature. They cannot be ruled out in the present context since we impose no restriction on the weights considered. They are designed to ensure that an individual of group i has no incentive to mimic an individual of group $i+1$:

$$U_i\left(c_i, y_i\right) \geq U_i\left(c_{i+1}, y_{i+1}\right) \qquad i = 1, 2. \tag{7}$$

With W denoting a general SWF (to be made more precise later on), the government's second-best problem is the following:

$$\max_{c_i, y_i} \quad W\left(U_1(c_1, y_1), U_2(c_2, y_2), U_3(c_3, y_3)\right)$$

[3] It may be necessary to introduce also either non-negativity or maximum bound constraints on labour. In a first-best environment, these contraints should be necessarily introduced. In a second-best framework, they can be ignored if the self-selection constraints suffice to prevent labour supplies at the optimal allocation from going below zero or above a certain maximum. We will assume throughout the paper that this is the case.

s.t.

$$
\begin{aligned}
(\lambda) \quad &: \quad \sum_i \widehat{\pi}_i \left(y_i - c_i\right) \geq R \\
(\lambda_i^d) \quad &: \quad U_i\left(c_i, y_i\right) \geq U_i\left(c_{i-1}, y_{i-1}\right), \quad i = 2, 3 \\
(\lambda_i^u) \quad &: \quad U_i\left(c_i, y_i\right) \geq U_i\left(c_{i+1}, y_{i+1}\right), \quad i = 1, 2
\end{aligned}
$$

where the non-negative dual variables of the budget and incentive-compatibility constraints are indicated on the left. The upperindices d and u of these variables stand for "downwards" and "upwards", respectively.

Therefore, in the above problem both downwards and upwards constraints relate individuals of groups 1 and 2 and likewise for individuals of groups 2 and 3. Given the quasilinearity of preferences, it will be shown that at least one of the two self-selection constraints relating two adjacent individuals is binding. Whether it is the downwards or the upwards constraint will depend on the parameters of the model and, especially, on the weights chosen. However, under certain circumstances, both constraints are binding. When the two constraints relating two adjacent individuals are binding, these individuals necessarily receive the same fiscal treatment (i.e., there is bunching). In the main text of this paper we shall abstract from bunching situations. A detailed analysis of these situations is provided in the appendix.

3 SWFs with different cardinalization of utilities

The objective function that the government is assumed to follow reflects an attempt to incorporate into the government's problem two separate, and perhaps conflicting, considerations. On the one hand, and in accordance with the principle of compensation, the government is assumed to want to redistribute from high-groups to low-skilled persons because of some non-negative aversion to inequality. To capture this, we follow the standard procedure of using a quasi-concave SWF within preference groups. Thus, let $W_h(U_1, U_2)$ and $W_\ell(U_2, U_3)$ be the SWF's for preference groups h and ℓ, respectively. At the same time, we allow the government to have varying attitudes toward persons of different preference by attaching a weight γ to the social welfare obtained by individuals with low preferences for leisure (those with α_ℓ), and $1-\gamma$ to those with high preferences for leisure. We posit that $0 \leq \gamma \leq 1$. Then, the government's objective function may be written $\gamma W_\ell + (1-\gamma) W_h$. To allow for the full range of possibilities, we consider two extreme cases for the SWFs within preference groups. In the next sub-section, we assume that the social welfare function is utilitarian (the sum of utilities) within each preference group. At the other extreme, within-group SWFs are assumed to be maximin, reflecting infinite aversion to inequality. That case is taken up in the following sub-section.

3.1 Weighted utilitarian welfare functions

As already said, the relative weights γ and $(1-\gamma)$ can be interpreted as relative cardinalizations of individual utility functions. When the SWF within preference groups is utilitarian, the overall weighted SWF is:

$$\sum_{i=1}^{3} \widehat{\mu}_i U_i\left(c_i, y_i\right),$$

where:

$$\widehat{\mu}_1 = \frac{(1-\gamma)\,p_\ell \pi_h}{\widehat{w}_1}, \quad \widehat{\mu}_2 = \frac{(1-\gamma)\,p_h \pi_h + \gamma p_\ell \pi_\ell}{\widehat{w}_2} \quad \text{and} \quad \widehat{\mu}_3 = \frac{\gamma p_h \pi_\ell}{\widehat{w}_3}$$

are the proportions $\widehat{\pi}_i$ adjusted for the above weights and divided by $\widehat{w}_i$ (recall the definition of U_i in (3)).

The government chooses the consumption-income bundle intended for each individual subject to the budget and the self-selection constraints studied previously. The Lagrangean for the weighted utilitarian social welfare function is then:

$$\begin{aligned} L &= \sum_{i=1}^{3} \widehat{\mu}_i U_i\left(c_i, y_i\right) + \lambda \left[\sum_i \widehat{\pi}_i\left(y_i - c_i\right) - R\right] + \\ &\quad + \sum_{i=2}^{3} \lambda_i^d \left(U_i\left(c_i, y_i\right) - U_i\left(c_{i-1}, y_{i-1}\right)\right) + \sum_{i=1}^{2} \lambda_i^u \left(U_i\left(c_i, y_i\right) - U_i\left(c_{i+1}, y_{i+1}\right)\right) \end{aligned} \tag{8}$$

The maximization yields the following first-order conditions:

$$FOC\left(y_1\right) : -\widehat{\mu}_1 + \lambda \widehat{\pi}_1 - \lambda_1^u + \lambda_2^d = 0 \tag{9}$$

$$FOC\left(c_1\right) : -\lambda \widehat{\pi}_1 + \left(\widehat{\mu}_1 + \lambda_1^u\right) \widehat{w}_1 u'\left(c_1\right) - \lambda_2^d \widehat{w}_2 u'\left(c_1\right) = 0 \tag{10}$$

$$FOC\left(y_2\right) : -\widehat{\mu}_2 + \lambda \widehat{\pi}_2 - \lambda_2^d - \lambda_2^u + \lambda_3^d + \lambda_1^u = 0 \tag{11}$$

$$FOC\left(c_2\right) : -\lambda \widehat{\pi}_2 + \left(\widehat{\mu}_2 + \lambda_2^d + \lambda_2^u\right) \widehat{w}_2 u'\left(c_2\right) - \lambda_3^d \widehat{w}_3 u'\left(c_2\right) - \lambda_1^u \widehat{w}_1 u'\left(c_2\right) = 0 \tag{12}$$

$$FOC\left(y_3\right) : -\widehat{\mu}_3 + \lambda \widehat{\pi}_3 - \lambda_3^d + \lambda_2^u = 0 \tag{13}$$

$$FOC\left(c_3\right) : -\lambda \widehat{\pi}_3 + \left(\widehat{\mu}_3 + \lambda_3^d\right) \widehat{w}_3 u'\left(c_3\right) - \lambda_2^u \widehat{w}_2 u'\left(c_3\right) = 0 \tag{14}$$

and the complementarity conditions:

$$\lambda_i^d \left(U_i\left(c_i, y_i\right) - U_i\left(c_{i-1}, y_{i-1}\right)\right) = 0 \qquad i = 2, 3 \tag{15}$$

$$\lambda_i^u \left(U_i\left(c_i, y_i\right) - U_i\left(c_{i+1}, y_{i+1}\right)\right) = 0 \qquad i = 1, 2 \tag{16}$$

From (9), (11) and (13), we obtain:

$$\lambda_2^d = \widehat{\mu}_1 - \lambda \widehat{\pi}_1 + \lambda_1^u \tag{17}$$

$$\lambda_3^d = \widehat{\mu}_1 + \widehat{\mu}_2 - \lambda\left(\widehat{\pi}_1 + \widehat{\pi}_2\right) + \lambda_2^u \tag{18}$$

$$\lambda = \sum_i \widehat{\mu}_i \tag{19}$$

where (18) can also be written as:

$$\lambda_3^d = \lambda \widehat{\pi}_3 - \widehat{\mu}_3 + \lambda_2^u. \tag{18'}$$

From (19), it can be easily seen that because of the quasilinearity in leisure of utility functions, λ only depends upon the distributions of skills and preferences and the weights γ and $(1-\gamma)$. Furthermore, λ can be given a very simple interpretation: it is the social cost of raising an additional unit of R, which is carried out by imposing on every individual a unit rise of his gross labour income, y_i $(i = 1, 2, 3)$. If the self-selection constraints are initially satisfied, such move keeps them satisfied since the indifference curves of any individual are horizontally parallel to each other.

As mentioned above, in this section we abstract from bunching situations. It is worth noticing from (17) that, in the absence of bunching, which of the two self-selection constraints relating groups 1 and 2 is binding depends upon the sign of $\widehat{\mu}_1 - \lambda\widehat{\pi}_1$. If it is positive, $\lambda_2^d > 0$ and $\lambda_1^u = 0$: the downwards self-selection constraint designed to ensure that a group-2 individual has no incentive to mimic a group-1 individual is binding. On the contrary, if it is negative, $\lambda_1^u > 0$ and $\lambda_2^d = 0$: it is now the upwards self-selection constraint designed to ensure that a group-1 individual has no incentive to mimic a group-2 individual that is binding. It can be shown[4] that $\widehat{\mu}_1 - \lambda\widehat{\pi}_1$ is positive for $\gamma = 0$ and negative for $\gamma = 1$. Therefore, as $\widehat{\mu}_1 - \lambda\widehat{\pi}_1$ is linear in γ, there is a switch in the binding incentive constraint when γ is such that $\widehat{\mu}_1 - \lambda\widehat{\pi}_1 = 0$. Let us denote this threshold value by γ_1. It is the value of γ given by:

$$\gamma_1 = \frac{\dfrac{1-p_\ell\pi_h}{\widehat{w}_1} - \dfrac{p_h\pi_h}{\widehat{w}_2}}{\dfrac{1-p_\ell\pi_h}{\widehat{w}_1} - \dfrac{p_h\pi_h}{\widehat{w}_2} + \dfrac{p_\ell\pi_\ell}{\widehat{w}_2} + \dfrac{p_h\pi_\ell}{\widehat{w}_3}} \tag{20}$$

Furthermore, since $\widehat{\mu}_1 - \lambda\widehat{\pi}_1$ decreases monotonically with γ, the downwards self-selection constraint of group-2 individuals is binding for values of $\gamma < \gamma_1$ while the upwards incentive constraint of group-1 individuals is binding for values of $\gamma > \gamma_1$. It is also easy to show[5] that

$$\frac{1}{2} < \gamma_1 < 1. \tag{21}$$

Similarly, from (18) the sign of $\widehat{\mu}_1 + \widehat{\mu}_2 - \lambda(\widehat{\pi}_1 + \widehat{\pi}_2)$ determines which of the two self-selection constraints relating individuals 2 and 3 is binding. If it is positive, $\lambda_3^d > 0$ (and $\lambda_2^u = 0$), and if it is negative, $\lambda_2^u > 0$ (and $\lambda_3^d = 0$), with the associated self-selection constraints binding. Two cases must be distinguished here according to whether $\pi_\ell p_\ell/\widehat{w}_2 - (1-\pi_\ell p_h)/\widehat{w}_3$ is positive or negative.[6]

[4] Substituting for $\widehat{\mu}_1$ and $\widehat{\pi}_1$ and also for λ from (19), we obtain: $\widehat{\mu}_1 - \lambda\widehat{\pi}_1 = (1-\gamma)$ $p_\ell\pi_h[(1-p_\ell\pi_h)/\widehat{w}_1 - p_h\pi_h/\widehat{w}_2] - \gamma p_\ell\pi_h\pi_\ell(p_\ell/\widehat{w}_2 + p_h/\widehat{w}_3)$, in which the expression in brackets is positive (since $\widehat{w}_2 > \widehat{w}_1$ and $1 - p_\ell\pi_h > p_h\pi_h$). Therefore, $\widehat{\mu}_1 - \lambda\widehat{\pi}_1 > 0$ for $\gamma = 0$ and < 0 for $\gamma = 1$.

[5] Equating to 0 the expression of $\widehat{\mu}_1 - \lambda\widehat{\pi}_1$ in the previous footnote and using $\widehat{w}_3 > \widehat{w}_2 > \widehat{w}_1$ yields $\gamma_1/(1-\gamma_1) > 1$ and so $\gamma_1 > 1/2$.

[6] The results that follow can be inferred from the following relation: $\widehat{\mu}_1 + \widehat{\mu}_2 - \lambda(\widehat{\pi}_1 + \widehat{\pi}_2) = \gamma p_h\pi_\ell[\pi_\ell p_\ell/\widehat{w}_2 - (1-\pi_\ell p_h)/\widehat{w}_3] + (1-\gamma)p_h\pi_\ell[\pi_h(p_\ell/\widehat{w}_1 + p_h/\widehat{w}_2)]$, that is positive for any $0 \leq \gamma \leq 1$ if $\pi_\ell p_\ell/\widehat{w}_2 - (1-\pi_\ell p_h)/\widehat{w}_3 > 0$. In the opposite case where $\pi_\ell p_\ell/\widehat{w}_2 - (1-\pi_\ell p_h)/\widehat{w}_3 < 0$, the above expression is positive for $\gamma = 0$ and negative for $\gamma = 1$. Therefore its sign switches from positive to negative for some critical value $0 < \gamma_2 < 1$. From (22) we infer that in this later case, $\gamma_2/(1-\gamma_2) > 1$ and so $1/2 < \gamma_2 < 1$.

If it is negative, there is a critical value of γ, labeled γ_2, at which $(\mu_1 + \widehat{\mu}_2) - \lambda(\widehat{\pi}_1 + \widehat{\pi}_2) = 0$. It is given by:

$$\gamma_2 = \frac{\dfrac{p_\ell \pi_h}{\widehat{w}_1} + \dfrac{p_h \pi_h}{\widehat{w}_2}}{\dfrac{p_\ell \pi_h}{\widehat{w}_1} + \dfrac{p_h \pi_h}{\widehat{w}_2} + \dfrac{1 - p_h \pi_\ell}{\widehat{w}_3} - \dfrac{p_\ell \pi_\ell}{\widehat{w}_2}}. \tag{22}$$

from which we infer that $1/2 < \gamma_2 < 1$. For $\gamma < \gamma_2$, the downwards self-selection constraint relating individuals of groups 2 and 3 is binding, while for $\gamma > \gamma_2$ it is the upwards one. In the opposite case where $\pi_\ell p_\ell / \widehat{w}_2 - (1 - \pi_\ell p_\ell)/\widehat{w}_3$ is positive, $(\widehat{\mu}_1 + \widehat{\mu}_2) - \lambda(\widehat{\pi}_1 + \widehat{\pi}_2) > 0$ for any $0 \leq \gamma \leq 1$, and so the downwards self-selection constraint relating individuals of groups 2 and 3 is always binding. To harmonize the presentation in the following, we adopt the convention that $\gamma_2 > 1$ in this case.

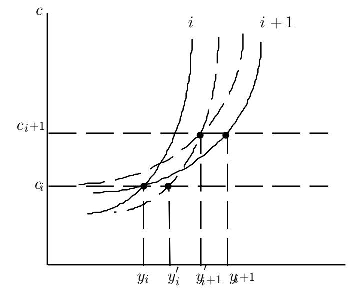

Figure 1: Switch of regime

It is worth noticing that the critical values of γ_1 and γ_2 depend only on the distribution of skills and preferences. To sum up, at the threshold weight levels γ_i $(i = 1, 2)$ there is a switch from a regime where the downwards self-selection constraint relating individuals of group $i + 1$ and group i is binding to a regime where the upwards self-selection constraint is binding. Recall that in the absence of bunching assumed here, at most one of these incentive constraints is binding. The switch of regime is described in Figure 1 where the pairs of plain and dotted indifference curves relate to the regimes before and after the switch, respectively. At this switch, the values of consumption stay the same.[7] However, the income values change since they are crucially determined, in this quasilinear setting, by the self-selection constraints that are binding.

The equations that determine the threshold weight values γ_1 and γ_2 can be interpreted in terms of the social disutility of earnings across groups. Let us focus on γ_1 . As mentioned, for $\gamma < \gamma_1$, $\widehat{\mu}_1 > \lambda \widehat{\pi}_1$. This inequality means that the social cost of a unit increase in y_1 (measured by $\widehat{\mu}_1 = p_\ell \pi_h (1 - \gamma) \alpha_h / w_\ell$) is higher than the social benefit of redistributing uniformly the revenue so raised ($\widehat{\pi}_1 = p_\ell \pi_h$) to all individuals (measured by $\lambda \widehat{\pi}_1$). Therefore, if $\widehat{\mu}_1 > \lambda \widehat{\pi}_1$, it

[7]This can be inferred from first-order conditions (10), (12) and (14).

is socially optimal to reduce y_1 as far as the self-selection constraints allow. This correspond to the plain indifference curves in Figure 1 and the associated incomes y_i and y_{i+1}. On the contrary, for $\gamma > \gamma_1$, the opposite holds and it is socially optimal to raise y_1 as far as possible. This corresponds to the dotted indifference curves in Figure 1 and the associated incomes y'_i and y'_{i+1}.

For given distributions of skills and preferences, the set of self-selection constraints that are binding at the optimum will depend on the weights chosen in the social welfare function. We have four possible regimes:

Regime I:	$\lambda_2^d > 0, \lambda_3^d > 0$	if $\gamma < \min[\gamma_1, \gamma_2]$
Regime IIa:	$\lambda_1^u > 0, \lambda_3^d > 0$	if $\gamma_1 < \gamma < \gamma_2$
Regime IIb:	$\lambda_2^d > 0, \lambda_2^u > 0$	if $\gamma_2 < \gamma < \gamma_1$
Regime III:	$\lambda_1^u > 0, \lambda_2^u > 0$	if $\gamma > \max[\gamma_1, \gamma_2]$

which are displayed in Figure 2.

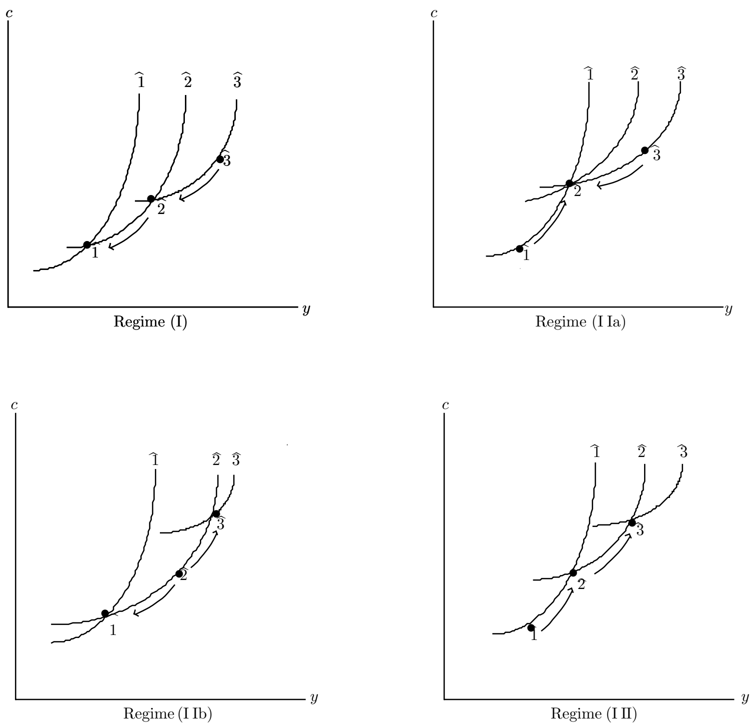

Figure 2: Sets of binding self-selection constraints in the four regimes

In each of the four diagrams, the arrows indicate the directions (either downwards or upwards) of the binding self-selection constraints. It is worth noticing that the government wants to redistribute income between two adjacent groups of individuals in the direction of the arrows.

It is now easy to see how the optimal regime evolves when γ rises, i.e. when more resources are made available to hard-working persons (with low preference for leisure) at the expense of epicurian (with high preference for leisure). Starting from $\gamma = 0$, Regime I stays optimal until γ reaches some critical value (either γ_1 or γ_2) that is larger than 1/2. It means that this regime prevails with a fully utilitarian objective, that gives equal weight to all individuals ($\gamma = 1/2$). In Regime I, the tax system is progressive within each preference-for-leisure group i.e. it redistributes from high-ability persons to low-ability ones. As γ reaches $\min(\gamma_1, \gamma_2)$, a switch to either Regime IIa or Regime IIb occurs, depending on whether $\gamma_1 < \gamma_2$ or $\gamma_1 > \gamma_2$ respectively (see Figure 3). After this switch, the tax system becomes regressive for one of the preference-for-leisure group; this is the epicurian group in Regime IIa, and the hard-working one in Regime IIb. It is caused by the impossibility of separating low-ability hard-working individuals from high-ability epicurian ones in group 2: raising the welfare of one of these types also benefits the other type, and progressivity has to be given up for either one of the preference-for-leisure groups.

When γ further rises, a new change, now to Regime III, may occur. In this regime, the tax system is regressive for both preference-for-leisure groups, and for the same reason as above. Notice that this further change does not occur if $\gamma_2 > 1$. In this case, Regime IIa stays optimal even for values of γ that are close or equal to 1.

These results stand in strong contrast to the ones that would prevail if the two preference-for-leisure groups were living on separate islands. In this situation, whatever the transfer between the two islands, the tax system would be progressive within the two preference-for-leisure groups

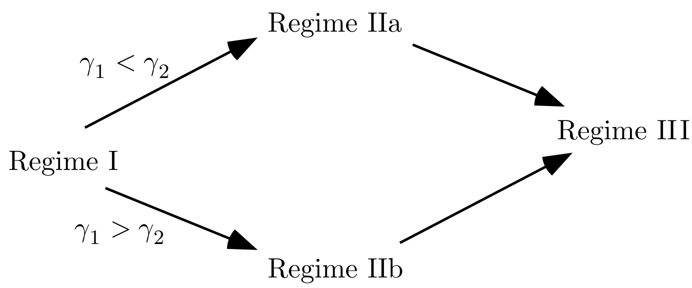

Figure 3: Changes of regime as γ rises

As shown above in Figure 3, whether it is Regime IIa or IIb that prevails as an intermediate regime depends upon the relative values of γ_1 and γ_2. From (20) and (22) we obtain the following:

$$\gamma_1 < \gamma_2 \Leftrightarrow \frac{w_h}{w_\ell} > \frac{p_h \pi_h}{p_\ell \pi_\ell}$$

where the relative proportions of the two types of individuals that are present in group 2 play an important role. In particular, if the individuals with low preference for leisure outweigh those with high preference in this group, Regime IIa is the intermediate optimal regime: the tax system is progressive in the low-preference-for-leisure group and regressive in the other preference group. For Regime IIb to prevail instead, it must be the case that there is a higher proportion of high ability individuals with high preferences for leisure in group 2, and that the gap between productivities is small enough.
To get further insights in the intermediate regime, it is worth noticing that Regime IIa is characterized by:

$$\lambda > \frac{1-\gamma}{\widehat{w}_1} \text{ and } \lambda > \frac{\gamma}{\widehat{w}_3},$$

which imply that it is optimal to increase both y_1 and y_3 as far as the incentive compatibility constraints permit. In this regime, our setting implies that redistribution within the high-preference-for-leisure group is regressive. We tend to redistribute towards individuals of group 2. On the contrary, Regime IIb is characterized by:

$$\lambda < \frac{1-\gamma}{\widehat{w}_1} \text{ and } \lambda < \frac{\gamma}{\widehat{w}_3},$$

and it is optimal to decrease both y_1 and y_3 as far as possible. In this regime, redistribution within the low-preference-for-leisure is regressive. When the number of high ability individuals with high preferences for leisure in the intermediate class is high enough, we tend to redistribute from individuals of group 2.

As indicated in the introductory section, a strand of the social choice literature investigates the consequences for redistribution of the responsibility-for-preferences principle. According to this, individuals should not be treated differently on the ground that they have different preferences. In our setting one might interpret this principle as requiring that there is no transfer between the two preference groups taken as a whole (each preference group must only rely on its own resources). This imposes a further constraint on the income tax scheme to be chosen since redistribution is now limited to taxes and transfers within each preference group. It implies that within each of these two preference groups, individuals of one ability type have their bundle located in the (y, c)-space above the 45^o line through the orgin while the bundle of the other type is located below this line. It is then straightforward to see that this implies that tax progressivity holds for one of the preference groups, and tax regressivity for the other one. This is only consistent with a value of γ that is larger than $1/2$, i.e. with a SWF that gives more weight to hard-working than to epicurian persons.[8]

[8]As numerical simulations have shown, with a quasi-linear utility function, it might be impossible to find a value of γ such that no redistribution occurs across preference groups. Depending upon the distribution of skills and preferences, when increasing γ from $1/2$ and approaching the critical value of γ for which there is no redistribution between preference groups, a change of regime may occur at which redistribution from the low-preference-for-leisure group to the high-preference-for-leisure one abruptly shifts in the opposite direction. In this case, no value of γ exists at which there is no redistribution between preference groups.

The optimal consumptions of the three groups can be inferred from the above first-order conditions (9) to (14). As these conditions indicate, they do not depend upon revenue requirement R. They are given by:

$$u'(c_1) = \frac{1}{\widehat{w}_1 + \frac{\widehat{w}_1 - \widehat{w}_2}{\lambda \widehat{\pi}_1}\lambda_2^d} \tag{23}$$

$$u'(c_2) = \frac{1}{\widehat{w}_2 + \frac{1}{\lambda \widehat{\pi}_2}\left[(\widehat{w}_2 - \widehat{w}_3)\lambda_3^d + (\widehat{w}_2 - \widehat{w}_1)\lambda_1^u\right]} \tag{24}$$

$$u'(c_3) = \frac{1}{\widehat{w}_3 + \frac{\widehat{w}_3 - \widehat{w}_2}{\lambda \widehat{\pi}_3}\lambda_2^u} \tag{25}$$

In fact, because of the quasi-linearity of utility function a rise in the required revenue makes gross incomes y_1, y_2 and y_3 increase uniformly while keeping c_1, c_2 and c_3 constant. The self-selection constraints are kept satisfied since the indifference curves of an individual are horizontally parallel.

As is standard, the optimal allocation resulting from solving the above problem can be decentralized by means of a tax schedule: $T(y) = y - c(y)$. The marginal tax rate on individual i is:

$$T'(y_i) = 1 - \frac{1}{\widehat{w}_i u'(c_i)},$$

which yields, using (9) to (12):

$$T'(y_1) = \frac{\left(\frac{\widehat{w}_2}{\widehat{w}_1} - 1\right)\lambda_2^d}{\lambda \widehat{\pi}_1} \tag{26}$$

$$T'(y_2) = \frac{\left(\frac{\widehat{w}_3}{\widehat{w}_2} - 1\right)\lambda_3^d - \left(1 - \frac{\widehat{w}_1}{\widehat{w}_2}\right)\lambda_1^u}{\lambda \widehat{\pi}_2} \tag{27}$$

$$T'(y_3) = -\frac{\left(1 - \frac{\widehat{w}_2}{\widehat{w}_3}\right)\lambda_2^u}{\lambda \widehat{\pi}_3} \tag{28}$$

the signs of which are indicated in Table 2. Thus, $T'(y_1) > 0$ in Regimes I and IIb, when $\lambda_2^d > 0$,

	Regime I	Regime IIa	Regime IIb	Regime III
$T'(y_1)$	> 0	$= 0$	> 0	$= 0$
$T'(y_2)$	> 0	?	$= 0$	< 0
$T'(y_3)$	$= 0$	$= 0$	< 0	< 0

Table 2 Marginal tax rates by regimes

and $T'(y_1) = 0$, otherwise. The marginal tax rate on group-1 individuals is different from zero and we distort these individuals' choice when the self-selection constraint that precludes group-2 individuals from mimicking group-1 individuals is binding. Thus, we impose a positive marginal tax rate on individuals of group 1 to relax the self-selection constraint. In Regimes IIa and III, $T'(y_3) < 0$ when $\lambda_2^u > 0$, and $T'(y_3) = 0$ otherwise. $T'(y_2)$ could be of either sign: $\lambda_3^d > 0$ makes the marginal tax rate positive whereas $\lambda_1^u > 0$ makes the marginal tax rate negative. In particular, $T'(y_2)$ is positive in Regime I, nil in Regime IIb and negative in Regime III, whereas it is ambiguous in regime (IIa) when individuals of both groups 1 and 3 try to mimic individuals of group 2. These results are standard: with a binding downwards incentive compatibility constraint, taxing marginal labour income enables one to relax the constraint, and with a binding upwards incentive compatibility constraint, the same outcome obtains by subsidizing marginal income.

3.2 Weighted maximin

In this section, we provide some results for the maximin social objective[9]. The heterogeneity of preferences makes it controversial to determine which of the two low-ability individuals is the Rawlsian individual whose utility should be maximized. Low-ability individuals with different disutility of labour will earn different levels of private income at the laissez-faire. Individuals with higher disutility of labour will earn less. However, the government may be reluctant to redistribute towards this kind of individuals, since they are also enjoying a higher quantity of leisure.

We consider a SWF of the maximin type within each preference group, but where the government is allowed to give different weights to the utilities of low-ability individuals with different preferences for leisure. As above, we denote by γ the weight given to individuals with low preference for leisure (i.e., individuals with α_ℓ) and by $(1-\gamma)$ the weight given to individuals with high preference for leisure, and we posit that $0 \leq \gamma \leq 1$. The social welfare function is then:

$$\sum_{i=1}^{2} \widehat{\mu}_i U_i \left(c_i, y_i\right),$$

where:

$$\widehat{\mu}_1 = \frac{(1-\gamma)}{\widehat{w}_1} \quad \text{and} \quad \widehat{\mu}_2 = \frac{\gamma}{\widehat{w}_2}.$$

The government chooses the consumption-income bundle intended for each of the three groups of individuals subject to the budget and the self-selection constraints introduced previously. The

[9]Our analysis is similar to Cuff (2000) where a maximin criterion is also considered with the most deserving individual being different according to whether a high disutility of labour comes from disability or laziness. In her analysis, preferences differ only for the low-skill individuals.

Lagrangean for the weighted social welfare function is then:

$$
\begin{aligned}
L \;=\; & \sum_{i=1}^{2} \widehat{\mu}_i U_i\left(c_i, y_i\right) + \lambda\left[\sum_i \widehat{\pi}_i\left(y_i - c_i\right) - R\right] + \\
& + \sum_{i=2}^{3} \lambda_i^d\left(U_i\left(c_i, y_i\right) - U_i\left(c_{i-1}, y_{i-1}\right)\right) + \sum_{i=1}^{2} \lambda_i^u\left(U_i\left(c_i, y_i\right) - U_i\left(c_{i+1}, y_{i+1}\right)\right)
\end{aligned} \tag{29}
$$

The maximization yields the following first-order conditions:

$$FOC\left(y_1\right) \;:\; -\widehat{\mu}_1 + \lambda\widehat{\pi}_1 - \lambda_1^u + \lambda_2^d = 0 \tag{30}$$

$$FOC\left(c_1\right) \;:\; -\lambda\widehat{\pi}_1 + \left(\widehat{\mu}_1 + \lambda_1^u\right)\widehat{w}_1 u'\left(c_1\right) - \lambda_2^d \widehat{w}_2 u'\left(c_1\right) = 0 \tag{31}$$

$$FOC\left(y_2\right) \;:\; -\widehat{\mu}_2 + \lambda\widehat{\pi}_2 - \lambda_2^d - \lambda_2^u + \lambda_3^d + \lambda_1^u = 0 \tag{32}$$

$$FOC\left(c_2\right) \;:\; -\lambda\widehat{\pi}_2 + \left(\widehat{\mu}_2 + \lambda_2^d + \lambda_2^u\right)\widehat{w}_2 u'\left(c_2\right) - \lambda_3^d \widehat{w}_3 u'\left(c_2\right) - \lambda_1^u \widehat{w}_1 u'\left(c_2\right) = 0 \tag{33}$$

$$FOC\left(y_3\right) \;:\; \lambda\widehat{\pi}_3 - \lambda_3^d + \lambda_2^u = 0 \tag{34}$$

$$FOC\left(c_3\right) \;:\; -\lambda\widehat{\pi}_3 + \lambda_3^d \widehat{w}_3 u'\left(c_3\right) - \lambda_2^u \widehat{w}_2 u'\left(c_3\right) = 0 \tag{35}$$

and the complementarity conditions:

$$\lambda_i^d\left(U_i\left(c_i, y_i\right) - U_i\left(c_{i-1}, y_{i-1}\right)\right) = 0 \qquad i = 2, 3 \tag{36}$$

$$\lambda_i^u\left(U_i\left(c_i, y_i\right) - U_i\left(c_{i+1}, y_{i+1}\right)\right) = 0 \qquad i = 1, 2 \tag{37}$$

From (30), (32) and (34), we obtain:

$$\lambda_2^d = \widehat{\mu}_1 - \lambda\widehat{\pi}_1 + \lambda_1^u \tag{38}$$

$$\lambda_3^d = \widehat{\mu}_1 + \widehat{\mu}_2 - \lambda\left(\widehat{\pi}_1 + \widehat{\pi}_2\right) + \lambda_2^u \tag{39}$$

$$\lambda = \sum_{i=1}^{2} \widehat{\mu}_i = \frac{(1-\gamma)}{\widehat{w}_1} + \frac{\gamma}{\widehat{w}_2} \tag{40}$$

From (40), it can be easily seen that the shadow price of public funds λ is once again completely determined by the distribution of skills and preferences and the weights γ and $(1-\gamma)$.

From (38) we obtain that, in the absence of bunching, which of the two self-selection constraints relating groups 1 and 2 is binding depends upon the sign of $\widehat{\mu}_1 - \lambda\widehat{\pi}_1$. There is a switch in the binding incentive constraint when γ is such that $\widehat{\mu}_1 = \lambda\widehat{\pi}_1$. Let us denote this threshold value by γ_1. Similarly, from (39) the sign of expression $\widehat{\mu}_1 + \widehat{\mu}_2 - \lambda\left(\widehat{\pi}_1 + \widehat{\pi}_2\right)$ determines which of the two self-selection constraints relating individuals 2 and 3 is binding. However, note that now we have that $\lambda = \widehat{\mu}_1 + \widehat{\mu}_2$, so the above expression reduces to $\lambda\left(1 - \widehat{\pi}_1 - \widehat{\pi}_2\right) = \lambda\widehat{\pi}_3$. Since this is always positive, $\lambda_3^d > 0$, with the associated downwards self-selection constraint always binding.

Now we have only two possible regimes:

Regime I	$\lambda_2^d > 0$, $\lambda_3^d > 0$	if $\gamma < \gamma_1$
Regime IIa	$\lambda_1^u > 0$, $\lambda_3^d > 0$	if $\gamma_1 < \gamma$

The critical values of γ_1 that determines the switch from one regime to the other is:

$$\gamma_1 = \frac{\widehat{w}_2 \left(1 - p_\ell \pi_h\right)}{\widehat{w}_1 p_\ell \pi_h + \widehat{w}_2 \left(1 - p_\ell \pi_h\right)} \tag{41}$$

Note that Regime I holds for the polar case in which the Rawlsian individual is a low-ability low-working individual (i.e. $\gamma = 0$) and Regime IIa holds for the polar case in which the Rawlsian individual is a low-ability hard-working individual (i.e. $\gamma = 1$). Regime I reflects the "standard" case in which the government wants to redistribute from higher to lower income persons, and tax liabilities rise with income. Of course, in so doing the government implicitly redistributes from low-taste-for-leisure high-skill persons to high-taste-for-leisure low-skill persons. As the weight attached to low-taste-for-leisure persons increases, the economy eventually moves into Regime IIa where the pattern of redistribution changes. In this regime, redistribution goes from the highest and lowest income persons towards the middle income persons. In this case, the government cannot avoid redistributing from high-taste-for-leisure low-skill persons to low-taste-for-leisure high-skill persons.

4 Concluding remarks

Our purpose in this paper has been to analyze the optimal nonlinear income tax schedule when individuals have heterogeneous and quasilinear preferences for leisure. The heterogeneity of preferences raises ethical questions regarding the choice of the social objective. We have studied the consequences of adopting variable cardinalizations of individual utilities. In particular, we have analyzed maximin and utilitarian social welfare functions where the government was allowed to give different weights to individuals with different preferences.

In this paper, individuals differ both in ability and in preferences for leisure, the government perfectly observes the private labour income but does not observe neither ability nor preferences. When preferences are quasilinear, both utilitarian and maxi-min objective functions yield closed-form solutions. We are able to derive the progressivity of the tax schedule as function of the distributions of skills and preferences, and of the weights chosen. For all the social objectives considered, we have explicitly analyzed the pattern of binding self-selection constraints for different weights (and the conditions for bunching in the appendix).

When the SWF is the weighted sum of individual utilities and no a priori restrictions are placed on the weights, both downwards and upwards self-selection constraints relating pairs of adjacent individuals need be taken into account. When the SWF consists in the unweighted sum of the individual utilities, the optimal nonlinear tax schedule implies that only downwards self-selection constraints are binding and a net transfer of resources from the group of individuals with low preferences for leisure towards the group of individuals with high preferences for leisure takes place. As the weight given to individuals with low preferences for leisure increases, the pattern of binding self-selection constraints may change and this net transfer diminishes. We have identified the threshold weight values for which changes in the binding constraints take place.

We have assumed throughout the paper that high-ability individuals with high preferences for leisure cannot be distinguished from low-ability individuals with low preferences for leisure. As a consequence, they receive the same fiscal treatment. The composition of this intermediate group plays a crucial role in determining the direction of redistribution inside each group of preferences. When the intermediate group is composed by a high proportion of low-ability individuals with low preferences for leisure, a kind of regressive redistribution (from low- to high-ability) takes place inside the group of individuals with high preferences for leisure. Alternatively, if the proportion of high-ability individuals with high preferences for leisure is high enough, the low-ability individuals with low preferences for leisure get penalized since this kind of regressive redistribution takes place inside the group of individuals with low preferences for leisure.

The analysis can be extended to the situation where skill w is a continuous variable on the interval $(\underline{w}, \overline{w})$. If $\overline{w}/\underline{w} > \alpha_h/\alpha_\ell$, there is now a range $(\underline{w}/\alpha_\ell, \overline{w}/\alpha_h)$ of values of $\widehat{w}$ such that two types of individuals are not distinguishable: at any point in this range, the same value of $\widehat{w}$ and so the same indifference curve map are shared by two types of individuals, namely a low-skill hard-working and a high-skill epicurian. The same qualitative results as in the discrete case obtain in this continuous case: when the weight given to hard-working in the weighted utilitarian welfare function is large enough, there are ranges of income for which the marginal tax rates is negative, implying that one wishes to redistribute from people having lower income to people having higher income (regressive redistribution).

Given the impossibility of satisfying responsibility for preferences while compensating for ability in all groups of individuals having the same preferences, Fleurbaey and Maniquet (1998) propose to use some common reference preferences in an unweighted utilitarian social welfare function for which compensation will be carried out as much as possible. It would be interesting to consider the possibility of using this approach in our setting and analyze how the extent of redistribution depends on the particular reference preferences chosen in the social welfare objective to be maximized.

Even though the gist of this paper is more conceptual than practical it has some policy implications. If one believes that some non observable and yet important characteristics are the result of individual choices, then one should move away from income tax schedules based on a "naive" utilitarian criterion, where the two preference groups are given the same weight in the SWF ($\gamma = 1 - \gamma = 1/2$). This can imply adopting schedules which appear regressive at least in terms of observed earnings. In particular, when the characteristic is as here the individuals' preference for leisure, whether this means that more leisure brings more utility or that more labor generates more pain is very crucial. In the first instance, there is no reason to penalize someone because he works more; this results from a decision taken at an earlier stage about the life style the individual will adopt. In the second instance, pain at work does not result from a deliberate choice and clearly calls for some compensation from a utilitarian social planner.

References

[1] Boadway, R., Cuff, K. and M. Marchand, 2000. Optimal income taxation with quasi-linear preferences revisited. *Journal of Public Economic Theory* 2, 435–460.

[2] Cuff, K., 2000. Optimality of workfare with heterogeneous preferences, Institute for Economic Research. *Canadial Journal of Economics*, 33 149–174.

[3] Fleurbaey, M. and F. Maniquet, 1998. Optimal income taxation: an ordinal approach. CORE DP 9865.

[4] Fleurbaey, M. and F. Maniquet, 1999. Compensation and responsibility. Mimeo.

[5] Lollivier, S. and J-C. Rochet, 1983, Bunching and second-order conditions: A note on optimal tax theory, Journal of Economic Theory 31, 392-400.

[6] Roemer, J., 1998. *Equality of Opportunity.* Harvard University Press, Cambridge.

[7] Weymark, J., 1986a. A reduced-form optimal nonlinear income tax problem. *Journal of Public Economics* 30, 199–217.

[8] Weymark, J., 1986b. Bunching properties of optimal nonlinear income taxes. *Social Choice and Welfare* 3, 213–232.

[9] Weymark, J., 1987. Comparative static properties of optimal nonlinear income taxes. *Econometrica* 55(5), 1165–1185.

Appendix: Characterization of bunching situations

In the main text we have abstracted from bunching situations. The purpose of this appendix is to analyze these situations. Bunching occurs when individuals belonging to different groups are given the same bundle (y, c).

It is worth noticing that the first-order conditions (9) to (14), the complementary conditions (15) and (16) and the requirement that the Lagrange multipliers be nonnegative are not sufficient to insure that the non-binding incentive constraints are satisfied with the required inequality. The self-selection constraints that relate individuals of groups i and $i+1$ are:

$$U_i\left(c_i, y_i\right) \geq U_i\left(c_{i+1}, y_{i+1}\right)$$
$$U_{i+1}\left(c_{i+1}, y_{i+1}\right) \geq U_{i+1}\left(c_i, y_i\right)$$

In figure 4 are displayed the indifference curves of groups i and $i+1$ that cross each other at bundle (y_i, c_i), with the indifference curve of group $i+1$ being flatter than that of group i: the above self-selection constraints are satisfied if and only if bundle (y_{i+1}, c_{i+1}) lies to the North-East of bundle (y_i, c_i) between the two indifference curves.

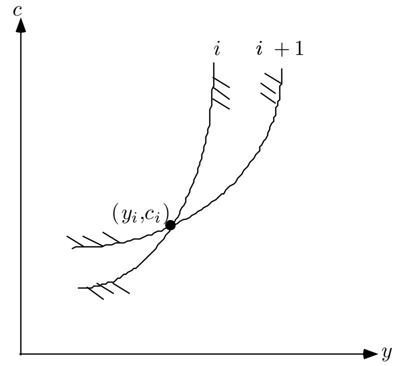

Figure 4: Necessary condition for absence of bunching

Hence, a necessary condition for an allocation to satisfy the self-selection constraints is:

$$(y_1, c_1) \leq (y_2, c_2) \leq (y_3, c_3)\,, \text{ with } (y_{i-1}, c_{i-1}) < (y_i, c_i) \quad \text{if} \quad (y_{i-1}, c_{i-1}) \neq (y_i, c_i)\,.$$

Given the quasilinearity of preferences, it suffices to check whether the previous condition holds for the consumption levels. Consumption should be nondecreasing (i.e., $c_i \geq c_{i-1},\ i = 2, 3$). These requirements define what are called the *second-order incentive constraints* (SOIC). If some are violated, bunching occurs.

We can infer the conditions for nondecreasing consumption[10] from (26) to (28):

$$c_3 \quad \geq \quad c_2 \Leftrightarrow \frac{w_2}{w_1}\left[1 + \frac{\lambda_2^u}{\lambda \widehat{\pi}_3} + \frac{\lambda_3^d}{\lambda \widehat{\pi}_2}\right] \geq \frac{\lambda_1^u}{\lambda \widehat{\pi}_2} \tag{42}$$

$$c_2 \quad \geq \quad c_1 \Leftrightarrow \frac{w_2}{w_1}\frac{\lambda_3^d}{\lambda \widehat{\pi}_2} \leq 1 + \frac{\lambda_2^d}{\lambda \widehat{\pi}_1} + \frac{\lambda_1^u}{\lambda \widehat{\pi}_2} \tag{43}$$

As a first result, we obtain that total bunching is ruled out. The denominator of (27) is indeed always larger than that of (25), so $c_3 > c_1$. We will analyze the conditions for partial bunching first at the bottom (i.e., bunching between 1 and 2), second at the top (between 2 and 3). We will employ equations (42) and (43) to derive the conditions for bunching.

(i) **Bunching at the bottom**

It can be easily shown that $\gamma < \gamma_2$ is necessary for bunching at the bottom. This is done by contradiction. If $\gamma > \gamma_2$, we have $\lambda_2^u > 0$, and we thus consider $\lambda_3^d = 0$ to focus on bunching at the bottom, since total bunching is impossible. However, if we substitute $\lambda_3^d = 0$ into the SOIC condition (43), we obtain that $c_2 > c_1$. Therefore, to analyze bunching at the bottom, we can restrict attention to $\gamma < \gamma_2$.

[10]To obtain these conditions, we also use the following equality: $(\widehat{w}_3 - \widehat{w}_2)/(\widehat{w}_2 - \widehat{w}_1) = w_2/w_1$.

For $\gamma \leq \gamma_1 < \gamma_2$, from (43) the condition for bunching is:

$$\frac{w_2}{w_1} \geq \frac{\widehat{\pi}_2}{\widehat{\pi}_1} \frac{\widehat{\mu}_1}{\lambda \widehat{\pi}_3 - \widehat{\mu}_3}. \tag{44}$$

To obtain this inequality, we have used relations (8) and (13) to substitute the λ_d^2 and λ_d^3 in (43). Likewise, for $\gamma_1 \leq \gamma < \gamma_2$ bunching prevails if

$$\frac{w_2}{w_1} \geq \frac{\lambda \left(\widehat{\pi}_1 + \widehat{\pi}_2\right) - \widehat{\mu}_1}{\lambda \widehat{\pi}_3 - \widehat{\mu}_3}. \tag{45}$$

Those conditions only depend upon the distribution of skill and preference.

(ii) **Bunching at the top**

It can be easily shown that $\gamma > \gamma_1$ is necessary for bunching at the top, once again by contradiction. If $\gamma < \gamma_1$, $\lambda_2^d > 0$, and we assume then $\lambda_1^u = 0$ to focus on bunching at the top, since total bunching is impossible. However, if we substitute $\lambda_1^u = 0$ into the SOIC condition (42), we obtain that $c_3 > c_2$. To analyze bunching at the top, we should thus restrict attention to situations with $\gamma > \gamma_1$.

Following the same procedure as before, we can determine the condition for bunching at the top. For $\gamma_1 < \gamma \leq \gamma_2$, this condition is

$$\frac{w_2}{w_1} \leq \frac{\lambda \widehat{\pi}_1 - \widehat{\mu}_1}{\lambda(\widehat{\pi}_2 + \widehat{\pi}_3) - \widehat{\mu}_3} \tag{46}$$

while for $\gamma_2 \leq \gamma < 1$, it is

$$\frac{w_2}{w_1} \leq \frac{\widehat{\pi}_3}{\widehat{\pi}_2} \frac{\lambda \widehat{\pi}_1 - \widehat{\mu}_1}{\widehat{\mu}_3}. \tag{47}$$