Fair optimal tax with endogenous productivities

Marc Fleurbaey*and Giacomo Valletta†

October 2012

Abstract

What is a good incentive-compatible policy when one wants to respect individual choices of labor and human capital but eliminate inequalities due to unequal access to human capital and different returns to human capital, and when earnings and human capital expenditures are the only verifiable variables? We propose a social ordering that incorporates this goal and we analyze the evaluation of tax reforms and the properties of optimal linear and non-linear taxes. For reform evaluation and for optimal non-linear taxation, the focus is on the situation of individuals with the most disadvantaged characteristics who work full time and spend a certain (high) amount in human capital.

Keywords: endogenous skills, human capital, tax reforms, optimal tax, Fairness.

JEL Classification: D63, D71.

1 Introduction

Since the seminal contribution of Mirrlees [29] most of the literature on income taxation has typically assumed that agents use a linear technology: their productive skills are fixed and independent of any factor that is subject to their choices. In spite of the numerous insights that this model has provided it suffers from two serious weaknesses. First, it is not very realistic. One can think for example of human capital. Agents can, at least to some extent, affect their productivity by making certain choices about their level of human capital. So, at least intuitively, a system of human capital subsidies can be used as a mean for redistributing income across agents and help to better tackle the trade off between efficiency and equity. Second, the fact that productivity

---

*Princeton University, USA. Email: mfleurba@princeton.edu.

†Dep. of Economics, University of Maastricht. Email: g.valletta@maastrichtuniversity.nl.

is exogenous partially hinders the possibility of examining the consequences of certain normative considerations. Assume one wants to compensate agents for their lack of productivity as long as they are not responsible for it. A model where productivity is exogenous would not allow to distinguish between a low-skilled agent who has *responsibly* chosen her lower productivity and another low-skilled agent who has instead suffered from various impediments which have prevented her from acquiring a higher skill.

We propose a model where agents differ in three characteristics. First, agents differ in their human capital-dependent earning ability. Typically human capital positively affects productivity, and a higher (marginal) productivity is typically reflected by an increased wage.[1] However the impact of human capital on earning ability can vary across agents. In particular, an individual might be more productive than another one for any possible level of human capital and such heterogeneity might be due to innate factors, or social connections, which are beyond their responsibility.

Second, in order to reach a certain level of human capital, different agents might have to face different intrinsic costs. That is, agents have a different human capital disposition. For instance, one could think of health. Two agents might both be willing to be in good health but one of them suffers from a congenital disease and only an expansive treatment allows her to be in good health: she has a worst human capital disposition than the other one. Once again these differences are typically due to factors that are beyond agents' responsibility: agents may have a different human capital disposition because of genetic factors[2], their upbringing or the social context they live in.[3] We will refer to these two characteristics, the earning ability and the human capital disposition, as to the *circumstances* of a certain agent.

Third, agents have heterogeneous preferences over consumption, labour and human capital: they typically make different choices about their consumption, about their labor time, and about their human capital level.

In such a framework we study criteria to evaluate policies aiming at combining income tax with a system of subsidies on human capital expenditure. More precisely, a tax policy is a function

---

[1] This is certainly the case for health care as shown, among others, by Mushkin [30], Grossman and Benham [20], Luft [26].

[2] Christensen et al. [9] shows that approximately a quarter of the variation in the liability to self-reported health and the number of hospitalizations could be attributed to genetic factors.

[3] A high socio-economic status is typically associated with better health and longer life, see for example Reid et al. [32], or Marmot el al. [28].

defining a transfer of income depending *both* on the level of earnings and on the human capital expenditure. So two agents with the same earnings might be subject to a different taxation (for example, one is subsidized and the other is taxed) solely because of a different human capital expenditure.

To perform our analysis we use social preferences that incorporate efficiency and fairness concerns. As also recently stressed by Piketty and Saez [31], the classical utilitarian social welfare function would not allow us to incorporate the value judgment that inequalities due to circumstances are more offensive than inequalities due to differences in preferences. Boadway et al. [4] introduced weights in the social welfare function in order to accommodate this idea, but did not provide a precise methodology to determine suitable weights. Such a methodology has been developed by Fleurbaey and Maniquet,[4] and we borrow it here. Specifically, we derive a social welfare ordering which is based on a particular welfare representation of agents' preferences. An individual's well-being is measured by the amount of money that would leave her indifferent between her current situation and being free to choose her labor time and her human capital expenditure from a hypothetical budget set where both her earning ability and her human capital disposition are equal to the average ones. Such a measure of individual well-being does not require any other information about individuals' utilities than their ordinal non-comparable preferences, a convenient property that follows from the attribution of responsibility for preferences and subjective utility to the individuals themselves. The well-being levels of agents, at any allocation, are then aggregated using the leximin criterion.

In spite of the complexity of the model we are able to provide a criterion for the comparison of non-linear tax policies in a setting where only earned income and human capital expenditure are observable. This task is made possible by the fact that we are using social preferences of the leximin type. Indeed, in order to understand which part of the tax function should be changed in priority, in case of reform, one needs to spot the worst off agent at the allocation generated by the tax function under examination. It turns out that the focus should always be on a specific region of the budget set shaped by the tax function. In order to evaluate a certain tax scheme the policy maker should look primarily at the part of the budget set that is attainable by an agent with the worst earning ability and the worst human capital disposition. However the worst off agent at the allocation generated by a certain tax policy does not actually need to be an agent

---

[4]Fleurbaey (2008), Fleurbaey and Maniquet (2011a,b).

whose circumstances are the worst in society. Nonetheless, giving priority to the region of the budget set that is attainable by poor agents means that the main concern of the policy maker should not necessarily be to improve the condition of agents with a low level of human capital but rather look, more in general, at agents for whom acquiring human capital is particularly difficult (i.e., agents with a bad human capital disposition). We also study the shape of the optimal income schedule. We consider first the case of linear taxes. It turns out that in certain instances taxing human capital expenditure might be optimal. Interestingly the occurrence of a negative subsidy rate on human capital expenditure is closely related to the distribution of preferences across the population. For example an high sensitivity of human capital expenditures to subsidies together with a low sensitivity of earnings to tax will bring high tax rates and low subsidy rates. The same happens if the worst off agent spends considerably less than average in human capital or if she earns considerably more than average. On the other hand, when we turn our attention to the nonlinear case we find that the agents who are at the focus of social preferences are typically subsidized. As a matter of fact the agent who receives the highest subsidy is some unskilled agent who works full time and who, among all the agents with the poorest health disposition, has the highest human capital expenditure.

The paper is organized as follows. Section 2 reviews the related literature. Section 3 introduces the model and the notation. Section 4 introduces the notion of social preferences used in the paper and the ethical requirements they bear. Section 5 proposes a way to compare different arbitrary tax policies. Section 6 describes some features of the optimal linear tax schedule. Section 7 describes some features of the optimal non-linear tax schedule. Section 8 deals with the case of observable human capital. Section 9 concludes. The appendix briefly resumes the axiomatic analysis.

2   Related literature

The generality of our model allows to build several links with different strands of the literature on optimal income taxation depending on how one interprets human capital.

One can think of human capital as the level of education of agents. There are few papers about the subsidization of education that allow for endogenous productivities (see among others, Jacobsyand and Lans [25] Bovenbergz, Guo and Krause [22], Maldonado [27]). The underlying idea is that education should be either taxed or subsidized depending on whether the elasticity of

earnings with respect to education is positively influenced by the labor supply or by the earning ability.[5] Our results also suggest that human capital expenditure might be taxed however both the scope and the rationale of our findings differ from those just mentioned. First, we find that taxing human capital might be optimal in the particular case of linear taxes. Second, whether or not taxing human capital is optimal depends not only on the technology available to agents (marginal benefit vs. marginal cost of human capital) but, as mentioned earlier, it also quite relevantly depends on the distribution of preferences across the population. This feature is specific to our setting, indeed in the papers mentioned above agents have (homogeneous) preferences over consumption and labor and are indifferent about education (which is merely instrumental in increasing productivity).

One can also interpret human capital as the level of health of the individuals. Interestingly most of the literature focusing on the taxation/subsidization of health care does not allow for endogenous earning abilities. Health is rather considered as a factor that can randomly affect the amount of resources available to an individual (see, among others, Blomqvist and Horn [2], Cremer and Pestieau [8], Rochet [33] and Henriet and Rochet [23]). The main objective of these papers is to understand whether covering people against such a risk, by means of a public health insurance, is welfare improving or not from an ex ante perspective.

All the papers quoted so far have focused on social objectives defined in terms of utilitarian-type social welfare functions. Such social welfare functions are typically not precisely specified, and the objective of redistributing resources only depends on their degree of concavity. Moreover they rely on specific assumptions about preferences such as separability, and generally assume that all individuals have the same utility function. A common result of this approach is that marginal tax rates are everywhere positive. Interestingly, this result still holds in situations where the maximin criterion is used. Boadway and Jacquet [3] provide minimal conditions for the marginal tax rate to be not only positive but also decreasing throughout the whole skill distribution.

Things become considerably more complex if individual preferences are assumed to be heterogeneous. As a matter of fact papers that deal with this assumption exclusively focus on income taxation, and human capital is not part of the analysis. Agents' productivities are heterogeneous but exogenous (see, among others, Boadway, Marchand, Pestieau and Racionero [4], Choné and Laroque [6] [7], Jacquet and Van de gaer [24], Saez [35]). As noted in the introduction, a key

---

[5] If for example the elasticity of earnings with respect to education depends more strongly on labor supply than on ability then it is optimal to subsidize education for the sake of efficiency.

difficulty that comes with preference heterogeneity is that in order to sum the utility levels of agents endowed with different preferences one needs a cardinalization of utilities. Moreover, as pointed by Jacquet and Van de gaer [24], with double heterogeneity traditional welfarist criteria (including utilitarianism) might lead to policy recommendations that are unappealing in at least two respects. First, they fail to compensate agents for inequalities deriving from characteristics they cannot be held responsible for. Second, depending on the weights assigned to different kinds of preferences the optimal policy might require to redistribute income even if all agents have the same earning ability and the same disposition to acquire human capital.

In order to tackle these difficulties we use a precise definition of social welfare for a population that is heterogeneous in three dimensions. Such a social objective is derived from fairness principles that capture the idea that inequalities due to circumstances are unfair whereas inequalities due to differences in preferences and utilities are acceptable.[6]

Such a methodology was introduced by Fleurbaey and Maniquet [18] and has been used for the evaluation of public policies in several frameworks already.[7] In particular this approach has given interesting insights about the evaluation of tax policies. Fleurbaey and Maniquet [15], [16] propose an array of social preferences (over the allocation of consumption and labor) and an array of criteria for the welfare evaluation of tax policies but their analysis is limited by the assumption that labor productivity is exogenous so that agents cannot be considered, to any extent, responsible for it.[8] A common feature of the (optimal) tax policies they propose is that poor hardworking agents should be granted the greatest subsidy in the whole population. From the standpoint of our more complex model, imposing the same tax burden (or subsidy) to individuals with the same human capital but a different human capital expenditure might lead, *de facto*, to income inequalities that are particularly undesirable among low income earners.

Valletta [36] introduced a simplified version of our model, in which health influences productivity, the choice of the health status is dichotomous, and there are only two types of earning abilities

---

[6]A frequent criticism is that differences in utilities may reflect different capacities for enjoyment that should be compensated as well. But if individuals differ in such capacities this should be explicitly introduced in the model, as additional objects of preferences. We assume that our model fully describes the object of individual preferences, so that differences in utilities cannot be due to inequalities in additional internal resources that the individuals care about. For a study of compensation for inequalities in internal resources, see Fleurbaey [13].

[7]A broad and detailed description of this methodology and its possible applications is provided by Fleurbaey and Maniquet .[17]

[8]See, however, Fleurbaey [13] (p. 149–150) for a brief analysis of endogenous skills.

and health dispositions in the population. His paper provides an axiomatic characterization of a social ordering function that can be easily extended to our model, and we will retain it here.

Finally, our paper also relates to the literature on commodity taxation. In their seminal contribution Atkinson and Stiglitz [1] showed that within a population of individuals who differ only in their labor productivity, if preferences are separable between labor and consumption of other goods, then commodity taxation cannot increase welfare above the level obtained with an optimal income tax alone. Many studies have examined the robustness of this result (see Boadway and Pestieau [5] for an overview). For the case of heterogeneous preferences and non-linear commodity taxation, Fleurbaey [12], using an approach similar to ours, shows that poor hardworking agnets should be submitted to a uniform or null commodity tax. This result does not hold in our framework if applied to human capital expenditures. This is explained by the fact that human capital is here a special commodity which affects the agents' productivity, and agents have to face unequal costs in order to acquire it.

3   The model

We consider a set of economies, each with a finite set of agents $N \subset \mathbb{N}$. There are three goods: consumption, labor and human capital. A bundle, for agent $i \in N$, is a triple $z_i = (c_i, l_i, h_i)$, where c_i is consumption, l_i is labor, and h_i is human capital. In particular, $c_i \in \mathbb{R}_+$ will be interpreted here as the expenditure on ordinary consumption goods, excluding human capital expenditure. As usual for this kind of analysis, $l_i \in [0, 1]$. Human capital is also a continuous variable, and for simplicity we assume $h_i \in [0, 1]$.[9] To sum up, the consumption set is $X = \mathbb{R}_+ \times [0, 1] \times [0, 1]$. An allocation describes each agent's bundle, and will be denoted by $z = (z_i)_{i \in N}$.

Agents have three characteristics: their personal preferences, their earning ability and their human capital disposition.

For each agent $i \in N$, 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. Let $R = (R_i)_{i \in N}$ denote the population profile of preferences. We restrict our attention to preferences which are continuous, strictly monotonic (increasing in c_i and h_i, decreasing in l_i) and convex. Let $R =$

---

[9] A more general model, describing human capital as a multidimensional variable, would be certainly more realistic and most of the results, in principle, would still hold. However this would render the analysis quite cumbersome.

$(R_i)_{i \in N}$ denote the profile of preferences of the whole population.

The marginal productivity of labour is assumed to be an increasing function of human capital, $w_i(h_i)$ with $w(0) \geq 0$. It is measured in consumption units per full time labor, so that for any l_i, $w_i(h_i)l_i$ is the agent's pre-tax income (earnings). Agents are endowed with different such functions. For some $i, j \in N$ we say that agent i is more productive than agent j if i's productivity function dominates j's, that is, if $w_i(h) \geq w_j(h)$ for all h. Let $w(.) = (w_i(.))_{i \in N}$ denote the profile of individual productivity functions for the whole population.

Finally, every individual i has a mapping $m_i(h_i)$ describing how much of human capital expenses must be made in order to bring her to a human capital level h_i. We assume that this function satisfies $m_i(h) = 0$ for $h \leq \underline{h}_i$, and belongs to two possible classes. In the first class, m_i is increasing over $\left[\underline{h}_i, \overline{h}_i\right]$, and is equal to $+\infty$ for $h > \overline{h}_i$. In the second class, it is increasing over $\left[\underline{h}_i, \overline{h}_i\right)$, tends to $+\infty$ for $h \to \overline{h}_i$, and is equal to $+\infty$ for $h \geq \overline{h}_i$. It is possible to have $\underline{h}_i = 0$ and/or $\overline{h}_i = 1$. We define the inverse function $m_i^{-1} : \mathbb{R}_+ \to \left[\underline{h}_i, \overline{h}_i\right]$ (first class) or $m_i^{-1} : \mathbb{R}_+ \to \left[\underline{h}_i, \overline{h}_i\right)$ (second class) by $m_i^{-1}(m) = \underline{h}_i$ for $m = 0$, $m_i^{-1}(m) = h$ such that $m_i(h) = m$ for $0 < m < m_i\left(\overline{h}_i\right)$, and in the first class, $m_i^{-1}(m) = \overline{h}_i$ for $m \geq m_i\left(\overline{h}_i\right)$. The function $m_i(.)$ captures all factors that determine the human capital costs, including not only purely human capital features but also social and economic characteristics which influence i's human capital. Again, agents are eventually endowed with different such functions. For some $i, j \in N$ we say that agent i has a (weakly) worse human capital disposition than agent j if $m_i(h) \geq m_j(h)$ for all h. Let $m(.) = (m_i(.))_{i \in N}$ denote the profile of human capital dispositions for the whole population.

An economy is denoted $e = (R, w(.), m(.))$. We let the population N remain implicit in this description of an economy. Let $\mathcal{D}$ denote the set of economies complying with our assumptions.

The set of allocations to be ranked by a social ordering is $Z(e) = X^{|N|}$. This set includes feasible and non-feasible allocations. An allocation is feasible if

$$\sum_{i=1}^{n} c_i + \sum_{i=1}^{n} m_i(h_i) \leq \sum_{i=1}^{n} w_i(h_i)l_i.$$

In absence of redistribution, the budget set of each agent $i \in N$ is equal to the possible combinations of consumption, labor and human capital that are attainable for her, given her earning ability and her human capital disposition. In the first best context, one can use lump-sum transfers in order to redistribute income across agents. Then, agent i's first-best budget set is, letting t_i denote the

transfer:

$$B(t_i, w_i(.), m_i(.)) = \{(c_i, l_i, h_i) \in X \mid c_i + m_i(h_i) \le t_i + w_i(h_i)l_i\}.$$

It is important to notice that agents have preferences for human capital so that they may choose a certain level of human capital just because they care about it and not necessarily because this choice is instrumental to the attainment of a higher level of consumption (the higher the level of human capital, the higher the earning ability).

In order to compare allocations in terms of fairness and efficiency we will use a specific measure of social welfare. We will define a complete ordering over all the (feasible and not feasible) allocations and this will be denoted by $\overline{R}$ for the weak preferences, with related strict preferences $\overline{P}$ and indifference $\overline{I}$. In other words, $z'\overline{R}z$ means that the allocation z' is (socially) at least as good as z, $z'\overline{P}z$ means that it is strictly better, and $z'\overline{I}z$ that they are equivalent. As the social ordering will depend on the profile of the population, we in fact need a social ordering function (SOF), i.e., a mapping from the set of economies to the set of complete orderings over allocations. So, for each $e \in \mathcal{D}$, we write $\overline{R}(e)$, $\overline{P}(e)$, and $\overline{I}(e)$ in order to express the fact that particular social preferences are specific to the economy $e \in \mathcal{D}$.

4 From fairness requirements to social welfare

This section introduces the specific notion of social welfare we use in this paper. Before that let us introduce the two main fairness requirements that single out this specific way of ranking social alternatives.

The need for redistribution comes first of all from the idea that one would like to compensate agents for differences in their circumstances that are beyond their responsibility. In our framework this amounts to saying that inequalities deriving solely from someone's human capital disposition or someone's earning ability are not acceptable. In other words:

It is a strict social improvement to change an allocation by modifying the consumption levels of two agents i and j who have identical preferences $R_i = R_j$, the same amount of labor time and the same level of human capital, from c_i,c_j to c'_i,c'_j such that

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

where Δ is some strictly positive real number.

9

If compared to previous contributions (Fleurbaey and Maniquet [15] [16]), our richer model allows us to refine the normative analysis about the way personal responsibility relates to unequal achievements. Indeed, one should notice that the main implication of this fairness requirement is that agents here are only partially responsible for (the level of) their marginal productivity because in part they can affect it by changing their human capital level. For example, an agent with a low level of human capital is compensated for her lack of productivity as long as she is compared with another agent with the same level of human capital but with a better earning ability function. On the other hand this fairness requirement is silent in those situations when a different marginal productivity is solely due to different choices made by different individuals (for example, an agent with a low level of human capital is not compensated for her lack of productivity if she is compared with an agent who has chosen to increase her marginal productivity by acquiring a higher level of human capital).

Redistribution should anyway have a limit: inequalities solely due to different choices might be acceptable since individuals should, at least to some extent, be held responsible for their goals. If all agents had the same human capital disposition mapping and the same earning ability mapping, then they should be let free to choose a different amount of labor, a different level of human capital and hence, indirectly, a different productivity.

It is a strict social improvement, in a society where all agents have the same circumstances, to change an allocation obtained via lump-sum transfers by modifying the lump-sum transfers of any two agents i and j, from t_i,t_j to t'_i,t'_j such that

$$t_i - \Delta = t'_i > t'_j = t_j + \Delta,$$

where Δ is some strictly positive real number.

Notice that this requirement implies that the laissez-faire allocation (i.e., no redistribution) should be the social optimum in this particular case of uniform earning ability and uniform human capital disposition.

The rest of the paper analyzes the consequences of such fairness requirements on the evaluation of policies. In order to do so we rely on a certain notion of social welfare that stems directly from the requirements we have just presented, together with efficiency, informational and robustness requirements. The appendix provides the complete list of the axioms. These axioms single out

both a specific measure of individual well-being (i.e., a way to perform interpersonal comparisons) and a way to aggregate these individual measures.

The index of well-being that is obtained is the lump sum transfer that would leave the agent indifferent between her current bundle and being free to choose her labor time and her human capital expenditure from a (hypothetical) budget set where both her earning ability and her human capital disposition are equal to the average ones. More formally, let $\overline{w}(.) = \frac{1}{|N|} \sum_{j \in N} w_j(.)$ and $\overline{m}(.) = \frac{1}{|N|} \sum_{j \in N} m_j(.)$ denote respectively the average earning ability function and the average human capital disposition function. Then, the implicit transfer associated with an agent's indifference curve, a hypothetical human capital disposition $\overline{m}(.)$, and an hypothetical earning ability $\overline{w}(.)$ is defined by[10]

$$IT_i(z_i, R_i, \overline{w}(.), \overline{m}(.)) = t \Leftrightarrow z_i I_i \max |_{R_i} B(t, \overline{w}(.), \overline{m}(.)).$$

This expression, if considered as a function of z_i, corresponds (for a given R_i) to a particular money-metric utility function. This measure of individual well-being does not require any information about individuals' subjective utility, it depends only on ordinal non-comparable preferences.

For any given allocation we can compute the vector of implicit transfers associated with the bundles received by each agent. The level of social welfare is then measured by the lowest implicit transfer in society at such an allocation. Two different allocations will be ranked applying the leximin criterion to the vector of the corresponding implicit transfers.

Average Circumstances Egalitarian Equivalent Leximin SOF(ACEE). For all $e \in \mathcal{D}$, $z, z' \in Z$,

$$z'\overline{R}(e)z \Leftrightarrow (IT(z_i', R_i, \overline{w}(.), \overline{m}(.)))_{i \in N} \geq_{lex} (IT(z_i, \overline{w}(.), \overline{m}(.), R_i))_{i \in N}.$$

Let us stress that that the particular reference budget set (the average one) used for the computation of the implicit transfer is, at least to some extent, dictated by the axioms we have used for the characterization. The axioms actually considerably constrain the set of potential options. The natural appeal of using average circumstances as a reference derives from the fact that, ideally, all agents are entitled to an equal split of the overall production possibility set.

---

[10]The expression $\max |_{R_i} B(t, \overline{w}(.), \overline{m}(.))$ denotes the subset of $B(t, \overline{w}(.), \overline{m}(.))$ that contains the best allocations for R_i. Under our assumptions, this subset is always non-empty.

5 Tax evaluation

We will use the notion of social welfare just described, first of all, as a tool for the evaluation of different arbitrary tax policies. As is well known in the taxation literature since Feldstein [10], the reform problem is often more relevant to policy makers than knowing the features of the optimal tax policy. In such a case the policy maker is primarily interested in determining which part of the tax policy should be changed first in order to obtain a social improvement.

5.1 Incentive-compatible allocations

Consider a given economy $e = (R, w(.), m(.))$. As $\overline{w}(.), \overline{m}(.)$ are now fixed, for ease of notation we write $IT(z_i, R_i)$ instead of $IT(z_i, R_i, \overline{w}(.), \overline{m}(.))$ to denote the implicit transfer of agent i at the bundle z_i. The policy maker is assumed to know the distribution of types in the population but ignores the characteristics of any particular agent. We assume that, in the second best context, only earned income, $y_i = w_i(h_i)l_i$, and human capital expenditure, $m_i = m_i(h_i)$, are observable.

A tax policy is a function $T(y, m)$ defining a transfer of income depending on the level of earnings and on the human capital expenditure. The tax turns into a subsidy when $T(y, m) < 0$. Individuals are free to choose their labor time and their human capital status in the budget set modified by the tax function namely, the set of bundles $(c, l, h) \in \mathbb{R}_+ \times [0, 1]^2$ such that

$$c_i \leq w_i(h_i)l_i - m_i(h_i) - T(w_i(h_i)l_i, m_i(h_i)).$$

Let $B_i(T)$ denote this set. In what follows we will focus on the space of consumption, earnings, human capital expenditure where agent's i budget set becomes

$$c_i \leq y_i - m_i - T(y_i, m_i),$$

so that the laissez-faire tax induces the budget set $c = y - m$. In addition to the budget constraint, every agent is submitted to the constraints $c, m \geq 0$, $y \leq w_i^*(m)$, where the function $w_i^*(m) = w_i \circ m_i^{-1}(m)$ determines the earning ability that i obtains with any amount of human capital expenditure m. A laissez-faire budget is represented in Figure 1. The part of the plane $c = y - m$ that lies above $c = 0$ is the uplifted triangle on the right-hand side of the figure (numbers in parenthesis are the slopes of the lines). The budget upper boundary is the subset of this triangle delineated by the points OABO. Note that the tax cannot affect the function $w_i^*(m)$, therefore it cannot change the projection of the curve AB on the (y, m) subspace (this projection is the dotted

curve on the figure), but it can change the level of consumption and enable the agent to obtain positive consumption on the left part of the figure. In particular, with subsidies it may become possible to have $m_i > w_i\left(\bar{h}_i\right)$.
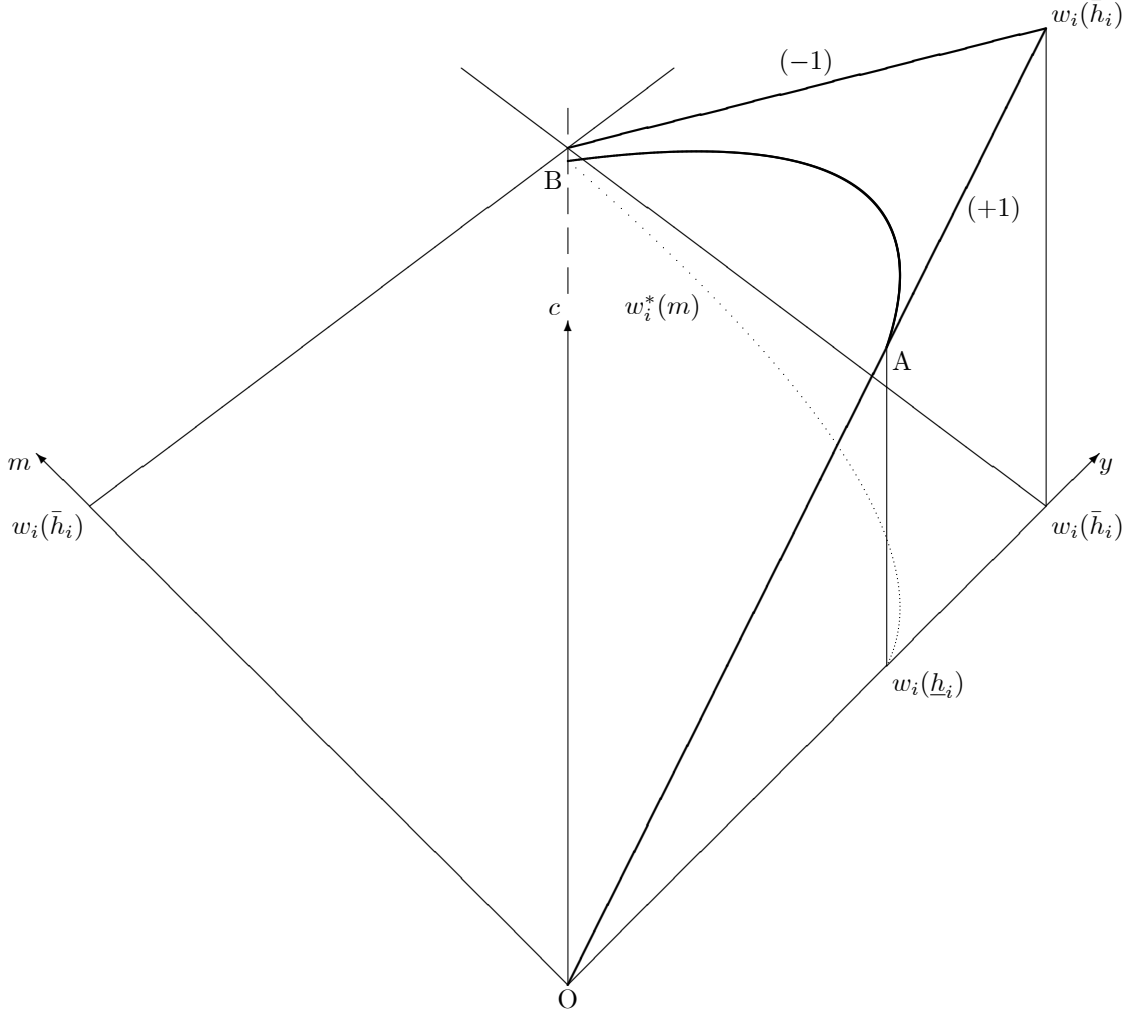

Figure 1: Laissez-faire budget

Let R_i^* define agent i's preferences over consumption, earnings and human capital expenditure. These are derived from the ordinary preferences R_i defined in the (c, l, h)-space as follows:

$$(c, y, m)\, R_i^*\left(c', y', m'\right) \Leftrightarrow \left(c, \frac{y}{w_i^*(m)}, m_i^{-1}(m)\right) R_i \left(c', \frac{y'}{w_i^*(m')}, m_i^{-1}(m')\right).$$

These preferences are continuous, convex, increasing in c, non-decreasing in m, and decreasing in y. In addition, they satisfy the following restriction:

$$(c, y, m)\, R_i^*\left(c, y', m'\right) \text{ if } \frac{y}{w_i^*(m)} \leq \frac{y'}{w_i^*(m')} \text{ and } m \geq m'. \tag{1}$$

13

This restriction comes from the fact that in the (c, l, h) space, (1) amounts to

$$(c, l, h) \, R_i \, (c, l', h') \; \text{ if } l \leq l' \text{ and } h \geq h', \tag{2}$$

which is a direct consequence of monotonicity of preferences in l and h[11].

The restriction described by (1) has important consequences on the agents' behavior. They will never choose a bundle (c, y, m) if they are given the possibility to choose another bundle which entails the same labor supply $y/w^*(m)$, a greater m (therefore greater h), and no lower c. It is important to stress that an agent might be confronted with this kind of choice in many ordinary situations. For instance, at the laissez-faire allocation, where $T(y, m) \equiv 0$, the bundle $(w^*(m) - m, w^*(m), m)$, corresponding to working full time and spending m in human capital, is dominated by another bundle $(w^*(m') - m', w^*(m'), m')$ if $m' > m$ and $w^*(m') - m' \geq w^*(m) - m$. This is a situation such that the extra human capital expenditure is more than repaid by the extra earnings it makes possible: $w^*(m') - w^*(m) \geq m' - m$. This restriction hence is a clear consequence of the fact that we are assuming endogenous productivity and imposes quite important changes in the analysis compared to the simpler model in which productivity is exogenous.

In order to better understand how this restriction affects the way people make choices on their budget set let us focus on the example depicted in Figure 2. Consider again the laissez-faire budget already introduced. Let us consider all the pairs (c, m) that are attainable by agent i if her labor time is fixed to a certain amount ($l = 0.5$ in the figure). This locus of points is represented by the (dotted) curve CD. In order to better show how this curve evolves we also draw the projection of it on (c, m) space (the curve $C'D'$ in the figure). Clearly there is a part of the $C'D'$ curve along which c is increasing in m. This means that there is a corresponding part of CD where consumption is increasing in health expenditure. Along this path agent i can increase her consumption just by increasing her human capital (and working the same amount of time) because the extra amount of money devoted to human capital is more than repaid by the fact the her productivity increases. Hence, the increasing part of CD will never be chosen by the agent, whatever her preferences. Finally, one should also notice that agent i might still choose a point on the decreasing part of CD: she could be willing, given her preferences, to give away consumption in order to acquire a

---

[11] An additional restriction is that

$$(c, y, m) \, I_i^* \, (c, y, m') \; \text{ if } m, m' \geq m_i \left(\overline{h}_i \right),$$

because this corresponds to a situation in which the corresponding (c, l, h) bundles are the same.
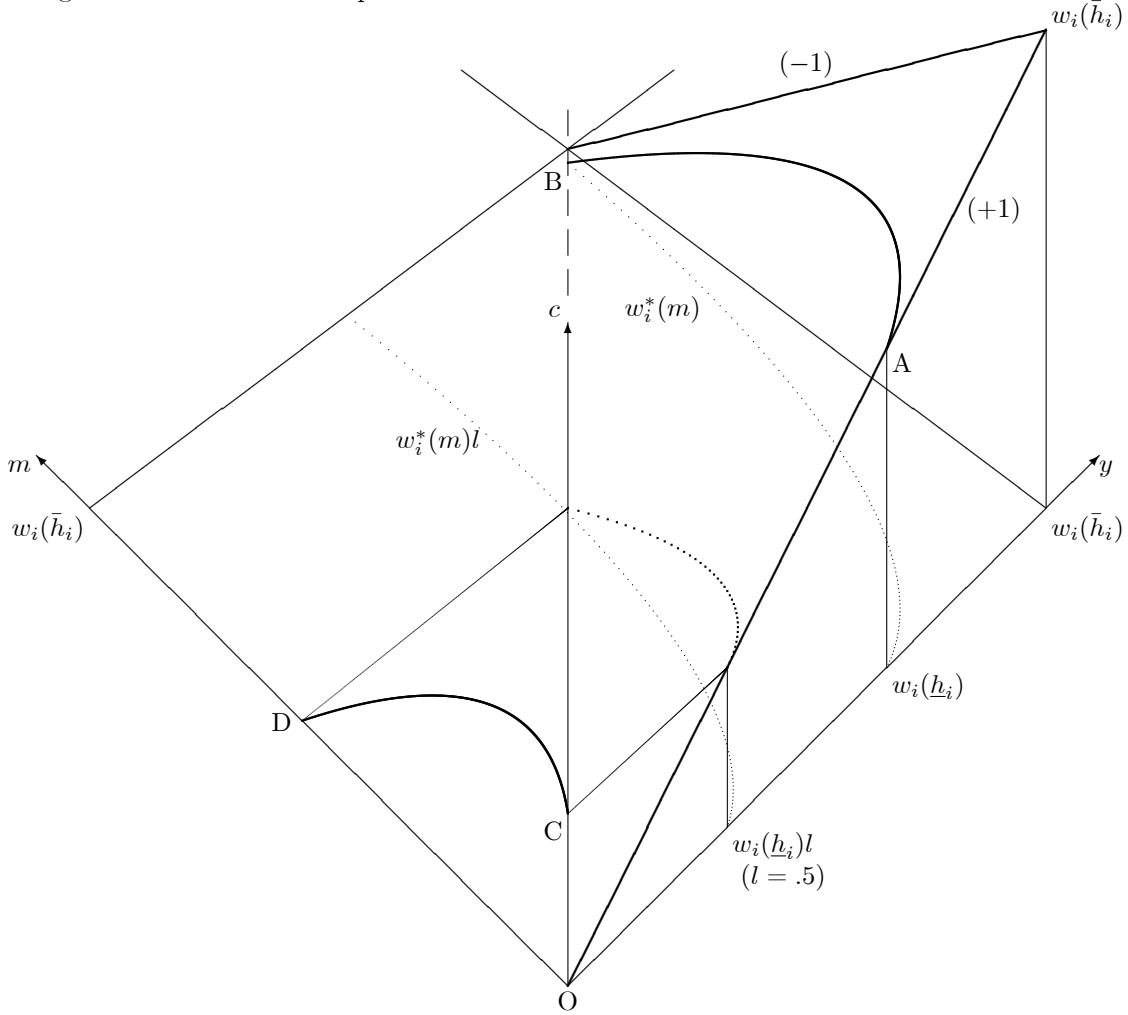
higher amount of human capital.

Figure 2: Budget curve in (c, m) space for fixed l

An allocation $z \in Z(e)$ is incentive compatible if and only if no agent envies the bundle of any other agent provided that such a bundle is feasible for her: for all $i, j \in N$,

$$(c_i, y_i, m_i)\, R_i^*\, (c_j, y_j, m_j) \ \text{ or } y_j > w_i^*\, (m_i)\,.$$

In other words agent i has to receive an allocation that she prefers to the allocation received by agent j unless it is not possible for her to mimic agent j. This implies that any incentive-compatible allocation can be obtained by letting every agent $i \in N$ choose her best bundle, under the constraint $y \leq w_i^*\, (m)$, in a budget set modified by a well chosen tax function $T(y, m)$ such that the locus of points

$$S(T) = \left\{ (c, y, m) \in \mathbb{R}_+^3 \mid c \leq y - m - T(y, m) \right\},$$

15

lies nowhere above the envelope curve of the indifference curves in the (c, y, m)-space, and intersects this envelope curve at all points (c_i, y_i, m_i) for each $i \in N$. Conversely, any allocation obtained by letting all agents choose from a budget set $S(T)$, under the constraint, $y \leq w_i^*(m)$, is incentive compatible. In other words, the taxation principle (Guesnerie [21], Rochet [34]) holds in this model. An incentive compatible allocation so obtained is feasible if and only if $\sum_{i=1}^{N} T(y_i, m_i) \geq 0$.

For every incentive-compatible allocation, there is a *minimal tax* that implements it, namely, the tax T such that $y - m - T(y, m)$ follows the lower envelope of agents' upper contour sets in (c, y, m) space at the allocation. For such a tax, $S(T)$ coincides with the intersection of the closed lower contour sets of the agents.

It is worth mentioning that minimal taxes, in this model, form a narrow class of tax functions, because of (1). In the classical Mirrlees model (with exogenous productivities), any non-decreasing function $y - T(y)$ can be arbitrarily close to the budget curve for a minimal tax for a sufficiently large population with sufficiently diverse preferences. In contrast, in this model, a function $y - m - T(y, m)$ that is non-decreasing in y and non-increasing in m may never be close to a minimal tax configuration. Let us develop this point. It holds true, by monotonicity of preferences R_i^*, that we can restrict attention to tax functions T such that $y - m - T(y, m)$ is non-decreasing in y and non-increasing in m. But the restriction (1) implies that some parts of such a budget set may never be chosen by the agents. Take a bundle (c, y, m) from the budget frontier such that for all i, the expression

$$\frac{y}{w_i^*(m)} w_i^*(x) - x - T\left(\frac{y}{w_i^*(m)} w_i^*(x), x\right)$$

is increasing in x. No agent will choose such a point, whatever her preferences, because for a fixed labor $l_i = y/w_i^*(m)$, consumption is increasing with human capital expenditure (one could think of the curve CD in the previous example).

5.2 Estimating social welfare

The evaluation of a policy clearly hinges on its social consequences. It turns out that evaluating the consequences of a certain policy is made tractable by the fact that we are using a social ordering function of the leximin type. Indeed, given the allocation generated by a given tax policy we just need to spot the worst off agent at such an allocation. Once we have this piece of information we know which part of the budget set modified by the tax function has to be changed (and how) in

order to obtain a social improvement. Let T be an arbitrary tax function such that $y-m-T(y,m)$ is non-decreasing in y and non-increasing in m. We want to compute $\min_i IT(z_i, R_i)$ at the allocation z generated by T.

Let us first consider the budget set, modified by the tax function T, of some agent $i \in N$ in (c,l,h) space. As explained above, the upper frontier of $B_i(T)$ may contain dominated parts (in terms of consumption). Indeed, increasing h may entail an increase in productivity that pays more than its cost. It is therefore better to focus on the undominated parts of the budget set since this gives a more accurate picture of the well-being opportunities of the worst type. More precisely, let us define a new budget set which flattens the dominated parts of the budget surface. For an arbitrary function $f(h)$, let $f^+(h)$ be the lowest non-increasing cover of f, i.e., the lowest function that is non-increasing and never below f. For a given l and T, let

$$b_{ilT}(h) = w_i(h)\,l - m_i(h) - T(w_i(h)\,l, m_i(h)).$$

The new budget is defined as the set of (c,l,h) such that

$$c \le b_{ilT}^+(h).$$

Let us call this new budget set $B_i^+(T)$. This step is exemplified in figure 3 where the thin line depicts, for a given l, the increasing part of $b_{ilT}(h)$ and the thick line depicts $b_{ilT}^+(h)$.
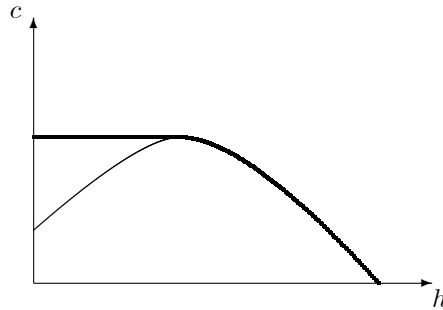

Figure 3: Budget set for given l

Let $B_\cap^+(T) = \bigcap_{i \in N} B_i^+(T)$. Finally, consider the budget $B(t_0, \overline{w}(.), \overline{m}(.))$ of the average type that would be obtained under laissez-faire but with a lump-sum transfer $t_0 \in \mathbb{R}$:

$$c \le \overline{w}(h)\,l - \overline{m}(h) + t_0.$$

17

For t_0 small enough (possibly negative), this budget $B(t_0, \overline{w}(.), \overline{m}(.))$ is contained in $B_\cap^+(T)$. Let t_0^* be the maximum level at which this property is satisfied (as exemplified in figure 4). This maximum level is well defined because both $B_\cap^+(T)$ and $B(t_0^*, \overline{w}(.), \overline{m}(.))$ are compact, and the latter varies continuously with t_0.
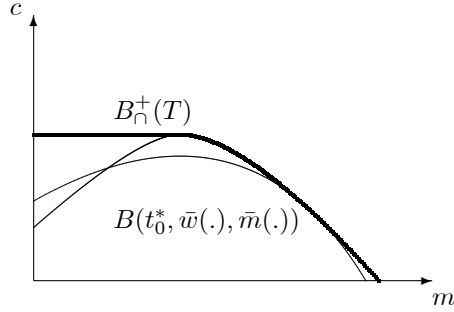

Figure 4: Budget tangency for some l

Agents, given their preferences, choose their bundle on the budget set modified by the tax function (see the indifference curves depicted in figure 5). Clearly the indifference surface passing through each bundle lies always above such a budget set. Moreover, no agent will choose a bundle that is in the interior of $B_\cap^+(T)$ so that, by construction, the allocation generated by the tax function T will grant to any agent a level of well being no lower than t_0^*. This observation enables us to bracket the value of $\min_i IT(z_i, R_i)$, as stated below. Observe on the figures that the boundary of $B(t_0^*, \overline{w}(.), \overline{m}(.))$ crosses that of (non-flattened) $B_\cap^+(T)$. This explains why it is important to work with $B_i^+(T)$ rather than $B_i(T)$ in order to obtain a tighter lower bound.

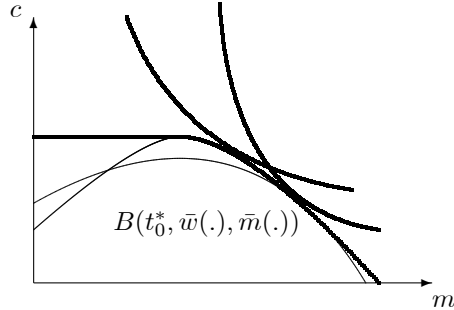Figure 5: Individuals choosing a particular l

Proposition 1 *Let z be an incentive-compatible allocation generated by the tax function T. Then*

$$t_0^* \leq \min_i IT(z_i, R_i) \leq \min_{i \in N} \left[c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{w_i^*(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i) \right].$$

Proof. Lower bound: By construction, for all $i \in N$,

$$B(t_0^*, \overline{w}(.), \overline{m}(.)) \subseteq B_{\cap}^+(T) \subseteq B_i^+(T).$$

Moreover, for all $i \in N$, the indifference surface passing through z_i lies always (weakly) above $B_i^+(T)$, therefore (weakly) above $B(t_0^*, \overline{w}(.), \overline{m}(.))$. Then, necessarily,

$$t_0^* \leq \min_i IT(z_i, R_i).$$

Upper bound: From the definition of $IT(z_i, R_i)$ it follows that, for each $i \in N$, at the allocation z,

$$IT(z_i, R_i) \leq c_i - \overline{w}(h_i) l_i + \overline{m}(h_i).$$

The right-hand side of this inequality is, for $i \in N$, equal to

$$c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{w_i^*(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i),$$

so that

$$\min_i IT(z_i, R_i) \leq \min_{i \in N} \left[c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{w_i^*(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i) \right].$$

∎

Proposition 1 provides an imperfect way to compare different policies. First, it may be silent in some applications because it only brackets the value of $\min_i IT(z_i, R_i)$. Second, the computation of the upper bound needs identifying the distribution of bundles of the individuals of all types. As the planner knows the distribution of characteristics in the population, this is a quantity that can be computed, i.e., it does not require identifying the characteristics of any particular individual. Nonetheless it would be more convenient to obtain formulae that require even less information. We study refinements below.

Let us briefly mention a favorable case that immediately follows from Proposition 1.

Corollary 2 *If there is $i \in N$ such that*

$$t_0^* = c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{w_i^*(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i),$$

then $t_0^* = \min_i IT(z_i, R_i)$. For such agent i, $IT(z_i, R_i) = t_0^*$.

The literature on fair income tax in the Mirrlees model, as summarized in Fleurbaey and Maniquet [17], suggests focusing on the most disadvantaged individuals. It turns out that this does not work so well in our extended framework. Let us explain, by exploring how the results from this literature could be translated into our setting.

First, one can introduce the assumption that there is a worst type, in terms of circumstances, in the population profile, i.e., a *poor* agent that is uniformly disadvantaged with respect to earning ability and human capital disposition.

Assumption 1 (Worst Type): There is a nonempty subset $P \subset N$ and two functions $\underline{w}(.), \underline{m}(.)$, such that for all $i \in P$, all $j \in N$, $w_i(h) = \underline{w}(h) \leq w_j(h)$ and $m_i(h) = \underline{m}(h) \geq m_j(h)$ for all h.

This assumption implies in particular that $\underline{w}(h) \leq \overline{w}(h)$ and $\underline{m}(h) \geq \overline{m}(h)$ and it simplifies the analysis because it enables us to identify the worst off agents more easily. Notice that, for every $i \in N$,

$$w_i(h) l - m_i(h) - T(w_i(h) l, m_i(h)) \geq \underline{w}(h) l - \underline{m}(h) - T(\underline{w}(h) l, \underline{m}(h)),$$

because the expression $wl - m - T(wl, m)$ is non-decreasing in w and non-increasing in m. This means that the worst type has a budget set that is always included in the budget sets of all other types of agents. For $i \in P$, the budget $B_i^+(T)$ will be denoted $\underline{B}^+(T)$. For every $i \in N$, $B_i^+(T)$ includes $\underline{B}^+(T)$ because whenever for two arbitrary functions f and g one has $f \geq g$, then necessarily $f^+ \geq g^+$. Therefore $\underline{B}^+(T) = B_\cap^+(T)$.

Another assumption is needed to obtain that the lower bound expressed in Proposition 1 coincides with the well being level of the worst off agent. It requires that whenever an individual is willing to choose a bundle that is accessible to the worst type, there is an agent of the worst type who is willing to choose the same bundle.

Assumption 2 (Preference Diversity): For all $i \in N$, there exists $j \in P$ such that $R_j^* = R_i^*|_{\{(c,y,m)|y \leq \underline{w}^*(m)\}}$.

This assumption is easier to satisfy with a large population (there must be as many sorts of preferences R_i^* in P as there are in the rest of the population), but it may be satisfied with a finite number of agents. Note that it may impose restrictions on the agents' preferences, because their indifference surfaces must not contain dominated areas for P agents, i.e., there must not

exist x, x' such that $x R_i^* x'$ for some $i \in N$ and x' dominates x for any $j \in P$ (in the sense that x' involves greater consumption, leisure, human capital). When the w^* functions for all types are proportional to one another, however, this possibility vanishes. Indeed, if $x' = (c', y', m')$ dominates $x = (c, y, m)$ for some $j \in P$, this means that

$$\left(c', -\frac{y'}{\underline{w}^*(m')}, m'\right) > \left(c, -\frac{y}{\underline{w}^*(m)}, m\right).$$

For $i \in N \setminus P$ such that $w_i^* = \lambda \underline{w}^*$, with $\lambda > 1$, this implies

$$\left(c', -\frac{y'}{w_i^*(m')}, m'\right) > \left(c, -\frac{y}{w_i^*(m)}, m\right),$$

i.e., x' dominates x for i as well. Therefore, in this case, Preference Diversity imposes no restriction on the preferences of agents from $N \setminus P$.

The following result is similar to results obtained by Fleurbaey and Maniquet [15],[16].

Corollary 3 *Assume Worst Type and Preference Diversity hold. Then for every minimal tax T one has $\min_i IT(z_i, R_i) = t_0^*$.*

Proof. By Preference Diversity, the lower envelope of upper contour sets of all $i \in N$, for bundles satisfying $y \leq \underline{w}^*(m)$, coincides with the lower envelope of upper contour sets of all $i \in P$. By construction, the lower envelope of upper contour sets for all $i \in P$ contains no dominated part. Moreover, the fact that T is minimal implies that the budget frontier coincides with the lower envelope of upper contour sets for all $i \in N$. For bundles such that $y \leq \underline{w}^*(m)$, the budget frontier therefore coincides with the lower envelope of upper contour sets for all $i \in P$. This implies that $\underline{B}^+(T) = B_i(T)$ for any $i \in P$ —otherwise there would be a gap between the budget frontier and the lower envelope of upper contour sets. Therefore the intersection of the upper frontier of $B(t_0^*, \overline{w}(.), \overline{m}(.))$ with the upper frontier of $\underline{B}^+(T)$ belongs to the upper contour set of some $i \in P$, which implies that for this particular i, $IT(z_i, R_i) = t_0^*$. By Proposition 1, necessarily $\min_{i \in N} IT(z_i, R_i) = t_0^*$. ∎

The fact that, due to (1), both the Preference Diversity assumption and the minimality assumption are rather special in our framework restricts the scope of Corollary 3, in contrast with the Mirrlees model in which they are unexceptional. In the framework of this paper, it is harder to spot the worst off agents at arbitrary tax policies.

5.3 More on the bounds

Let us examine in more details the bounds obtained in Proposition 1, starting with the lower bound t_0^*. Its computation in (c, l, h) space is not transparently connected with the tax function T. It is instructive to examine how t_0^* can be computed when one looks at the budget set in the (c, y, m) space. We retain the Worst Type assumption in the sequel, as it allows us to focus on the low incomes $y \leq \underline{w}^*(m)$. Moreover, when it comes to the computation of t_0^* the Worst Type assumption switches the focus on $\underline{B}^+(T)$ since, as mentioned earlier, $\underline{B}^+(T) = B_\cap^+(T)$. Hence, the first thing to do is to locate $\underline{B}^+(T)$ in (c, y, m) space. In (c, l, h) space, it is defined by

$$c \leq \underline{b}_{lT}^+(h),$$

for all $l \in [0, 1]$, where $\underline{b}_{lT}^+(h)$ denotes $\underline{b}_{ilT}^+(h)$ for any $i \in P$. In (c, y, m) one has then to look at the curve

$$\underline{b}_{lT}^*(m) = \underline{w}^*(m)\, l - m - T(\underline{w}^*(m)\, l, m)$$

(this could be exemplified by the curve CD in Figure 2). In order to complete the transposition of the budget $c \leq \underline{b}_{lT}^+(h)$ into (c, y, m) space one has simply to set $l = y/\underline{w}^*(m)$ and require $y \leq \underline{w}^*(m)$. Note that

$$\underline{b}_{lT}^+\left(\underline{m}^{-1}(m)\right) = \underline{b}_{lT}^{*+}(m).$$

The relevant budget is then

$$c \leq \underline{b}_{\frac{y}{\underline{w}^*(m)}T}^{*+}(m), \quad y \leq \underline{w}^*(m).$$

The set $B(t_0, \overline{w}(.), \overline{m}(.))$, that represents the budget set of a hypothetical agent with average circumstances who receives a lump sum transfer t_0, is defined in (c, l, h) space by

$$c \leq \overline{w}(h)\, l - \overline{m}(h) + t_0,$$

which corresponds in (c, y, m) space to

$$c \leq y - m + t_0 \text{ and } y \leq \overline{w}^*(m),$$

where $\overline{w}^*(m) = \overline{w}\left(\overline{m}^{-1}(m)\right)$.[12]

In order to seek the intersection between $B(t_0, \overline{w}(.), \overline{m}(.))$ and $\underline{B}^+(T)$ in (c, y, m) space, the former budget has to be rescaled so that the configuration of the two budgets correctly represents

---

[12]Notice that this is not generally the same as $\frac{1}{n}\sum_i w_i^*(m)$.

in this space what happens in (c, l, h) space. This is done as follows. Take any point (c, y, m) that is attainable by a worst-type agent, i.e., such that $y \leq \underline{w}^*(m)$. This point corresponds to some point (c, l, h) such that $c = c$, $l = y/\underline{w}^*(m)$, $h = \underline{m}^{-1}(m)$. This point (c, l, h) would correspond, for a an agent with average circumstances, to some point (c', y', m') such that $c' = c$, $y' = \overline{w}(h) l$, $m' = \overline{m}(h)$. We therefore obtain the transformation

$$c' = c$$

$$y' = \overline{w}\left(\underline{m}^{-1}(m)\right) y/\underline{w}^*(m)$$

$$m' = \overline{m}\left(\underline{m}^{-1}(m)\right).$$

Therefore, in (c, y, m) space, the budget $B(t_0, \overline{w}(.), \overline{m}(.))$, as it appears to a worst-type agent, is rescaled to

$$c \leq \frac{\overline{w} \circ \underline{m}^{-1}(m)}{\underline{w}^*(m)} y - \overline{m} \circ \underline{m}^{-1}(m) + t_0.$$

Finally consider any intersection point (c^*, y^*, m^*) between $B(t_0, \overline{w}(.), \overline{m}(.))$ and $\underline{B}^+(T)$. For the sake of simplicity let us start with the assumption that $\underline{b}^*_{\frac{y^*}{\underline{w}^*(m)}T}(m^*) = \underline{b}^{*+}_{\frac{y^*}{\underline{w}^*(m)}T}(m^*)$. That is, the intersection does not occur in a dominated part of the budget $\underline{B}(T)$. Then we have

$$y^* \frac{\overline{w} \circ \underline{m}^{-1}(m^*)}{\underline{w}^*(m^*)} - \overline{m} \circ \underline{m}^{-1}(m^*) + t_0 = y^* - m^* - T(y^*, m^*)$$

implying

$$t_0 = y^* \left[1 - \frac{\overline{w} \circ \underline{m}^{-1}(m^*)}{\underline{w}^*(m^*)}\right] + \overline{m} \circ \underline{m}^{-1}(m^*) - m^* - T(y^*, m^*).$$

The value of t_0^* corresponds to the minimum of this expression, i.e., the lowest value of t_0 such that the upper boundaries of the two budget sets have a non-empty intersection.

Proposition 4 *If* $\underline{b}^*_{\frac{y^*}{\underline{w}^*(m)}T}(m^*) = \underline{b}^{*+}_{\frac{y^*}{\underline{w}^*(m)}T}(m^*)$ *at the intersection point between* $B(t_0, \overline{w}(.), \overline{m}(.))$ *and* $\underline{B}^+(T)$, *then* t_0^* *is the minimum of*

$$y^* \left[1 - \frac{\overline{w}_N \circ \underline{m}^{-1}(m^*)}{\underline{w}^*(m^*)}\right] + \overline{m}_N \circ \underline{m}^{-1}(m^*) - m^* - T(y^*, m^*). \tag{3}$$

Otherwise, the minimum of (3) is less or equal to t_0^*, *and therefore still provides a lower bound for* $\min_i IT(z_i, R_i)$.

Proof. The first part has been proved in the text.

Suppose that $\underline{b}^*_{\frac{y^*}{\underline{w}^*(m)}T}(m^*) \neq \underline{b}^{*+}_{\frac{y^*}{\underline{w}^*(m)}T}(m^*)$ at the intersection point. This means that the intersection between $B(t_0, \overline{w}(.), \overline{m}(.))$ and $\underline{B}^+(T)$ occurs for a greater t_0 than the intersection

between $B(t_0, \overline{w}(.), \overline{m}(.))$ and $\underline{B}(T)$. The minimum of (3) gives t_0 for the latter intersection. This proves the second part. ∎

It is worth examining further how (3) can be used to compute t_0^*. Let us focus on the particular case in which T is non-decreasing in y. Then the expression (3) is decreasing in y^* so that $y^* = \underline{w}^*(m^*)$. Let us substitute $\underline{w}^*(m^*)$ for y^* in (3); this yields:

$$\underline{w}^*(m^*) - \overline{w} \circ \underline{m}^{-1}(m^*) + \overline{m} \circ \underline{m}^{-1}(m^*) - m^* - T(\underline{w}^*(m^*), m^*). \tag{4}$$

If the assumption $\underline{b}^*_{\frac{y^*}{\underline{w}^*(m)}T}(m^*) = \underline{b}^{*+}_{\frac{y^*}{\underline{w}^*(m)}T}(m^*)$ holds, we are in a part of the space for which $\underline{w}^*(m^*) - m^* - T(\underline{w}^*(m^*), m^*)$ is non-increasing in m^*. The behavior of $\overline{m} \circ \underline{m}^{-1}(m^*) - \overline{w} \circ \underline{m}^{-1}(m^*)$ is less obvious, as it may decrease for low m^* and increase for high m^*. Therefore, under the assumption that $\underline{b}^*_{\frac{y^*}{\underline{w}^*(m)}T}(m^*) = \underline{b}^{*+}_{\frac{y^*}{\underline{w}^*(m)}T}(m^*)$, one has to consider two possible alternatives. Either the minimum is obtained for some m^* satisfying $\underline{w}^*(m^*) - m^* - T(\underline{w}^*(m^*), m^*) > 0$ and

$$(1 - T_y(\underline{w}^*(m^*), m^*))\underline{w}^{*\prime}(m^*) - T_m(\underline{w}^*(m^*), m^*) - 1 = \frac{\overline{w}' \circ \underline{m}^{-1}(m^*) - \overline{m}' \circ \underline{m}^{-1}(m^*)}{\underline{m}' \circ \underline{m}^{-1}(m^*)} \tag{5}$$

where T_x denotes the partial derivative of T with respect to x (assuming that these functions are differentiable), or

$$t_0^* = -\overline{w} \circ \underline{m}^{-1}(m^*) + \overline{m} \circ \underline{m}^{-1}(m^*), \tag{6}$$

for m^* such that $\underline{w}^*(m^*) - m^* - T(\underline{w}^*(m^*), m^*) = 0$.

Let us interpret these results, in light of the fact that the intersection of the two budget sets identifies the worst off agent if there is $i \in P$ whose upper contour set contains this point. The left hand side of equation (5) features the impact of an increase in m^* on consumption for an agent from P. On the right-hand side, it displays the impact on consumption, of the same increase in human capital, for an agent with average characteristics at laissez-faire. Hence, the worst off agent, if his bundle lies at the intersection, belongs to P, works full time and enjoys a post-tax productivity of human capital expenditures equal to the pre-tax productivity of an average type agent with the same amount of labor and human capital (note that the same level of h would be obtained by the average type agent for a different level of m). The second case, equation (6), is obtained when the intersection of the two budget sets occurs at a bundle where consumption is null (a corner solution). This corresponds to a case in which the worst off situation corresponds to the greatest affordable human capital expenditure (and full-time work) for a worst-type agent.

Let us now turn our attention to the upper bound that appears in Proposition 1. This term is informationally demanding because it requires the computation of the situation of all types of agents induced by T. In order to reduce the computational requirements of the upper bound, one can invoke Preference Diversity and then look only at all i such that $y_i \leq \underline{w}^* (m_i)$. Thus, knowing the distribution of earnings and human capital expenditures in this low-income bracket is enough, and no information about which type consumes which bundle is needed.

However, as we have already explained, this assumption is restrictive because it imposes restrictions on preferences. An alternative is to look at all i such that $y_i \leq \underline{w}^* (m_i)$ and $\underline{b}^*_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i) = \underline{b}^{*+}_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i)$. For this to yield an upper bound, one only needs to assume that all points (c_i, y_i, m_i) such that (y_i, m_i) satisfy these conditions could be chosen by some agents from P (they can also be chosen by some other agent). This is formulated in the following assumption.

Assumption 3 (Weak Preference Diversity): For all $i \in N$, if $y_i \leq \underline{w}^* (m_i)$ and $\underline{b}^*_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i) = \underline{b}^{*+}_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i)$, then there exists $j \in P$ such that $(c_i, y_i, m_i) R_j^* (c_j, y_j, m_j)$.

Note that by incentive compatibility, necessarily $(c_j, y_j, m_j) R_j^* (c_i, y_i, m_i)$, so that we could as well write $(c_i, y_i, m_i) I_j^* (c_j, y_j, m_j)$ in the assumption . Assumption 3 is a logically weaker variant of the Preference Diversity assumption. Indeed it only assumes that, for the particular tax function T, and the particular allocation it generates, the undominated part of the budget set of agents in P contains no bundle that is chosen by some agent with better circumstances and would not be acceptable to any of the agents in P. This is weaker than Preference Diversity in two ways. First, it depends on the specific tax function T, and it may not hold for some other tax function, whereas Preference Diversity is independent of T. Second, unlike Preference Diversity, it imposes no restriction on the $N \setminus P$ agents' preferences. In a nutshell, our new assumption is much weaker, and obtains results for more economies, but it does not apply to *all* tax functions.

Proposition 5 *Let z be an incentive-compatible allocation generated by the tax function T. Under Preference Diversity,*

$$\min_i IT(z_i, R_i) \leq \min_{i: y_i \leq \underline{w}^*(m_i)} \left[c_i - \frac{\overline{w} \circ \underline{m}^{-1}(m_i)}{\underline{w}^*(m_i)} y_i + \overline{m} \circ \underline{m}^{-1}(m_i) \right].$$

Under Weak Preference Diversity,

$$\min_i IT(z_i, R_i) \leq \min_{\substack{i: y_i \leq \underline{w}^*(m_i), \\ \underline{b}^*_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i) = \underline{b}^{*+}_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i)}} \left[c_i - \frac{\overline{w} \circ \underline{m}^{-1}(m_i)}{\underline{w}^*(m_i)} y_i + \overline{m} \circ \underline{m}^{-1}(m_i) \right].$$

Proof. First, observe that, for the P subpopulation defined in the Worst Type assumption,

$$\min_{i \in P} \left[c_i - \frac{\overline{w} \circ \underline{m}^{-1}(m_i)}{\underline{w}^*(m_i)} y_i + \overline{m} \circ \underline{m}^{-1}(m_i) \right] = \min_{i \in P} \left[c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{w_i^*(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i) \right]$$

$$\geq \min_{i \in N} \left[c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{w_i^*(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i) \right]$$

is an upper bound for $\min_i IT(z_i, R_i)$. We focus on the specification of this upper bound.

Under Preference Diversity, the intersection of the closed lower contour sets of all $i \in N$, on the subset of (c, y, m) such that $y \leq \underline{w}^*(m)$, coincides with the intersection of the closed lower contour sets of all $i \in P$. Therefore, for all chosen bundles (c_i, y_i, m_i) such that $y_i \leq \underline{w}^*(m_i)$, there is $j \in P$ such that $(c_i, y_i, m_i) I_j^*(c_j, y_j, m_j)$. For such j, one has

$$IT(z_j, R_j) \leq c_i - \frac{\overline{w} \circ \underline{m}^{-1}(m_i)}{\underline{w}^*(m_i)} y_i + \overline{m} \circ \underline{m}^{-1}(m_i).$$

Therefore

$$\min_{j \in P} IT(z_j, R_j) \leq c_i - \frac{\overline{w} \circ \underline{m}^{-1}(m_i)}{\underline{w}^*(m_i)} y_i + \overline{m} \circ \underline{m}^{-1}(m_i).$$

The conclusion follows.

Under Weak Preference Diversity, the same reasoning as in the previous paragraph holds on the subset of chosen (c_i, y_i, m_i) such that $y_i \leq \underline{w}^*(m_i)$ and $\underline{b}^*_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i) = \underline{b}^{*+}_{\frac{y_i}{\underline{w}^*(m_i)}T}(m_i)$. ∎

This corollary is important because it reduces the amount of information needed for policy evaluation to data that are more easily available to the policy-maker than the full distribution of population characteristics. It suffices to look at bundles in a well-defined area —the function $\underline{w}^*$ is known— and, regarding the second part of the result, it is not difficult to locate and exclude the dominated part of the worst-type budget.

Apart from Corollary 3, in our results there is no guarantee that the worst off individuals actually belong to the worst type. But this is plausible and it would not be difficult to make assumptions to this effect (e.g., assuming that all types of preference orderings R_i found in the population are represented in P). However, this would not provide any different bounds than those obtained here, it would only probably make them closer to $\min_i IT(z_i, R_i)$. Nonetheless, the introduction of the Worst Type assumption considerably sharpens the scope of Proposition 1 as reflected by Propositions 4 and 5 whose practical implications are quite simple to grasp: even if the well being of the worst off agent cannot necessarily be measured and she is not necessarily a member of P, still she is located in the region of the budget set that is attainable by a poor agent. This part of the budget set should be the focus of any reform.

6 Optimal tax: the linear case

We consider the previous section to be the most relevant for practical policy-making. However, the classical literature on optimal taxation has by and large focused on the design of the optimal tax scheme (that is, a tax scheme that maximizes the social welfare function under incentive compatibility constraints). In this section and the next one, we study some properties of the optimal tax scheme under the particular notion of social welfare we are using.

We first restrict our attention to linear taxes. Namely,

$$T(y, m) = \tau y - \rho m - \theta,$$

where $\theta \in \mathbb{R}$ is a universal lump-sum grant while $\tau \in \mathbb{R}$ and $\rho \in \mathbb{R}$ are the parameters for marginal income tax rate and human capital subsidy rate. This implies that, for each agent $i \in N$, the budget set modified by the tax function is

$$c_i \leq \theta + (1 - \tau) y_i - (1 - \rho) m_i,$$

or, in (c, l, h) space,

$$c_i \leq \theta + (1 - \tau) w_i (h_i) l_i - (1 - \rho) m_i (h_i).$$

Each agent i chooses a bundle (c_i, l_i, h_i) that is the best for her preferences in this budget. This choice depends (apart from her preferences) on her circumstances and on the parameters θ, τ and ρ. To sum up we can define the functions

$$(c_i(\theta, \tau, \rho, w_i, m_i), l_i(\theta, \tau, \rho, w_i, m_i), h_i(\theta, \tau, \rho, w_i, m_i)),$$

that describe the choice made by each agents as a function of the parameters of the problem. This representation enables us to deduce, from the preference ordering R_i, the preference ordering over $(\theta, \tau, \rho, w, m)$. For each $i \in N$ let us denote such an ordering by $R_i^\diamond$:

$$(\theta, \tau, \rho, w, m) \, R_i^\diamond \, (\theta', \tau', \rho', w', m') \Leftrightarrow$$

$$(c_i(\theta, \tau, \rho, w, m), l_i(\theta, \tau, \rho, w, m), h_i(\theta, \tau, \rho, w, m)) \, R_i$$

$$(c_i(\theta', \tau', \rho', w', m'), l_i(\theta', \tau', \rho', w', m'), h_i(\theta', \tau', \rho', w', m'))$$

The general budget constraint requires

$$n\theta + \rho \sum_i m_i(h_i(\theta, \tau, \rho, w_i, m_i)) \leq \tau \sum_i w_i(h_i(\theta, \tau, \rho, w_i, m_i)) l_i(\theta, \tau, \rho, w_i, m_i). \tag{7}$$

Let $\theta(\tau, \rho)$ denote the maximum θ compatible with a given pair (τ, ρ). Plugging $\theta(\tau, \rho)$ into equation (7) one obtains a budget constraint that is function of two parameters only (τ, ρ). That is,

$$n\theta(\tau, \rho) + \rho \sum_i m_i (h_i (\theta(\tau, \rho), \tau, \rho, w_i, m_i)) = \tau \sum_i w_i (h_i (\theta(\tau, \rho), \tau, \rho, w_i, m_i)) l_i (\theta(\tau, \rho), \tau, \rho, w_i, m_i).$$

Or, equivalently

$$n\theta(\tau, \rho) + \rho M(\theta(\tau, \rho), \tau, \rho) = \tau Y(\theta(\tau, \rho), \tau, \rho), \tag{8}$$

where

$$M(\theta(\tau, \rho), \tau, \rho) = \sum_i m_i (h_i (\theta(\tau, \rho), \tau, \rho, w_i, m_i)),$$

and

$$Y(\theta(\tau, \rho), \tau, \rho) = \sum_i w_i (h_i (\theta(\tau, \rho), \tau, \rho, w_i, m_i)) l_i (\theta(\tau, \rho), \tau, \rho, w_i, m_i).$$

Differentiating both sides of (8) by τ one obtains

$$n\theta_\tau + \rho(M_\theta \theta_\tau + M_\tau) = Y + \tau(Y_\theta \theta_\tau + Y_\tau). \tag{9}$$

Similarly, differentiating both sides of (8) by ρ one obtains

$$n\theta_\rho + M + \rho(M_\theta \theta_\rho + M_\rho) = \tau(Y_\theta \theta_\rho + Y_\rho). \tag{10}$$

From these two equations, after some manipulations, one obtains

$$\tau = \frac{(M_\theta \theta_\tau + M_\tau)(M + n\theta_\rho) + (M_\theta \theta_\rho + M_\rho)(Y - n\theta_\tau)}{(M_\theta \theta_\tau + M_\tau)(Y_\theta \theta_\rho + Y_\rho) - (M_\theta \theta_\rho + M_\rho)(Y_\theta \theta_\tau + Y_\tau)}, \tag{11}$$

$$\rho = \frac{(Y - n\theta_\tau)(Y_\theta \theta_\rho + Y_\rho) + (M + n\theta_\rho)(Y_\theta \theta_\tau + Y_\tau)}{(M_\theta \theta_\tau + M_\tau)(Y_\theta \theta_\rho + Y_\rho) - (M_\theta \theta_\rho + M_\rho)(Y_\theta \theta_\tau + Y_\tau)}. \tag{12}$$

We will focus here on the ratio τ/ρ, which is a good summary of the instrument mix:

$$\frac{\tau}{\rho} = \frac{(M_\theta \theta_\tau + M_\tau)(M/n + \theta_\rho) + (M_\theta \theta_\rho + M_\rho)(Y/n - \theta_\tau)}{(Y_\theta \theta_\tau + Y_\tau)(M/n + \theta_\rho) + (Y_\theta \theta_\rho + Y_\rho)(Y/n - \theta_\tau)}. \tag{13}$$

We will assume that $\theta_\tau > 0 > \theta_\rho$. In order to better understand the determinants of the optimal ratio between redistribution (of income) and subsidization (of human capital), we will also have a close look to the case of quasi-linear preferences where $Y_\theta = M_\theta = 0$. This simplifies (13) into

$$\frac{\tau}{\rho} = \frac{M_\tau(M/n + \theta_\rho) + M_\rho(Y/n - \theta_\tau)}{Y_\tau(M/n + \theta_\rho) + Y_\rho(Y/n - \theta_\tau)}. \tag{14}$$

If there is a sufficient diversity of preferences, the optimal tax will be close to being the best tax for one of the worst off individuals. If agent i's preferences over (c, y, m) are represented by

$u(c, y, m)$, given the budget constraints for the individual and society the agent chooses his bundle by maximizing

$$u\left(\theta\left(\tau, \rho\right) + (1 - \tau) y - (1 - \rho) m, y, m\right)$$

w.r.t. (y, m). Assuming an interior solution, one then obtains, by the envelope theorem, that the evolution of u when tax parameters change is described by:

$$\frac{\partial u}{\partial \tau} = u_c \left[\theta_\tau - y\right] \quad \text{and} \quad \frac{\partial u}{\partial \rho} = u_c \left[\theta_\rho + m\right].$$

The optimal tax for i then satisfies:

$$\theta_\tau = y_i \quad \text{and} \quad \theta_\rho = -m_i.$$

Plugging this into equation (13), one gets:

$$\frac{\tau}{\rho} = \frac{\left(M_\theta y_i + M_\tau\right)\left(M/n - m_i\right) + \left(-M_\theta m_i + M_\rho\right)\left(Y/n - y_i\right)}{\left(Y_\theta y_i + Y_\tau\right)\left(M/n - m_i\right) + \left(-Y_\theta m_i + Y_\rho\right)\left(Y/n - y_i\right)}, \tag{15}$$

and, in the case of quasi-linear preferences (equation (14)):

$$\frac{\tau}{\rho} = \frac{M_\tau\left(M/n - m_i\right) + M_\rho\left(Y/n - y_i\right)}{Y_\tau\left(M/n - m_i\right) + Y_\rho\left(Y/n - y_i\right)}. \tag{16}$$

We will focus on the standard situation in which

$$Y_\tau \leq 0; \ M_\rho \geq 0$$

and, moreover,

$$Y_\rho \geq 0; \ M_\tau \leq 0,$$

because if ρ increases, M increases, makes agents more productive, so that they work more and earn even more; if τ increases, agents work less, which reduces the payoff of human capital, therefore leading to less expenditure in m.

Similarly, if income effects are not too strong, we can then assume that

$$M_\theta y_i + M_\tau \ \leq \ 0 \leq -M_\theta m_i + M_\rho,$$

$$Y_\theta y_i + Y_\tau \ \leq \ 0 \leq -Y_\theta m_i + Y_\rho.$$

If $m_i > M/n$ and $y_i < Y/n$, then τ and ρ have the same sign, which will be positive if[13]

$$\left(M_\theta y_i + M_\tau\right)\left(-Y_\theta m_i + Y_\rho\right) > \left(-M_\theta m_i + M_\rho\right)\left(Y_\theta y_i + Y_\tau\right)$$

---

[13] This determines the sign of the denominator in (11) and (12).

a plausible condition. In particular, in the quasi-linear case, this condition boils down to $M_\tau Y_\rho >$ $M_\rho Y_\tau$, which is very likely to occur because Y should be more sensitive to τ than to ρ, whereas the opposite holds for M.

So, under the assumptions we have listed so far what does ultimately determine the mix of income redistribution and human capital subsidies? If M_ρ is much greater than $|M_\tau|$ and $|Y_\tau|$ is much greater than Y_ρ, and if $M_\theta, |Y_\theta|$ are small, the prominent terms in (15) and (16) form the ratio

$$\frac{M_\rho \left(Y/n - y_i\right)}{Y_\tau \left(M/n - m_i\right)},$$

and provide a simple message. More redistribution than subsidizing will take place if human capital expenditures react strongly to subsidies whereas earnings react less to tax (this reflects the incentive concern), and if the gap between the relevant worst-off agent and the average situation is greater in earnings and smaller in human capital expenditures (this reflects the inequality concern).

This simple message is refined by adding the other components of the ratios in (15) and (16). In particular, more income redistribution in the mix will be pushed by a greater sensitivity of M to income tax and a lower sensitivity of Y to human capital subsidies. These results also suggest that it may be optimal to tax human capital expenditures if the worst-off spend less than average in human capital or, alternatively, earn more than average. The former case does not appear unrealistic in the context of education. The following simulations illustrate this possibility.

Consider an economy with eight equally sized subgroups, varying in three dimensions: preferences, earning function, human capital expenditure function. Preferences are either $c + \sqrt{(1 - l)\, h}$ or $c + 1.5\sqrt{(1 - l)\, h}$; wages are either $\sqrt{h}$ or $2\sqrt{h}$; expenditures are either h^2 or $2h^2$. The average functions are therefore $\overline{w}(h) = 1.5\sqrt{h}$ and $\overline{m}(h) = 1.5h^2$. In this economy the optimal policy is approximately $\tau = .35$ and $\rho = -.13$, with $\theta = .21$. The worst-off type's (first preferences, low wage, high cost of human capital) human capital expenditure is .09 units below the average. The particular feature of this example is that the worst-off agents have preferences with less concern for leisure but also for human capital than the other type of preferences.

It must be emphasized, however, that even in this kind of economy, lower-than-average human capital expenditures on behalf of the worst-off is not sufficient to induce an optimal tax (i.e., a negative subsidy) on such expenditures, because the other terms in (16) can dominate. If preferences are either $c + \sqrt{(1 - l)\, h}$ or $c + 2\sqrt{(1 - l)\, h}$ and wages are either $2\sqrt{h}$ or $4\sqrt{h}$, then the optimal

policy is approximately $\tau = .39$ and $\rho = .03$, with $\theta = .69$. The worst-off type's human capital expenditure is .22 units below the average but ρ is positive. Nevertheless, it is strikingly low, and this is because the term $Y_\tau (M/n - m_i)$ is strongly negative in (16), even though it ends up being counterbalanced by the positive term $Y_\rho (Y/n - y_i)$ because of the great gap $Y/n - y_i$.

Note that, except in the simulations, little use has been made in this analysis of the specific social ordering introduced earlier in the paper. The only feature of this ordering that has been invoked is its focus on the worst-off (independently of how the worst-off is identified). Thus, in light of the previous section, it is worth exploring how the results are further specified when the relevant worst-off agent i is from P and when it is a good approximation to consider $IT_i (\theta, \tau, \rho) = t_0^*$ and i is working (approximately)[14] full time, so that $y_i = \underline{w}^* (m_i)$. Then, the comparison between y_i and average earnings may depend on the profile of the population, but it should remain quite standard to have $y_i < Y/n$.

In addition, if $c_i > 0$, then, from equation (5), $m_i = \underline{m} (h_i)$, for h_i such that:

$$(1 - \tau) \underline{w}' (h_i) - (1 - \rho) \underline{m}' (h_i) = \overline{w}' (h_i) - \overline{m}' (h_i) .$$

This also reads

$$\overline{w}' (h_i) - (1 - \tau) \underline{w}' (h_i) = \overline{m}' (h_i) - (1 - \rho) \underline{m}' (h_i) .$$

If $\underline{w}' (h_i)$ is much lower than $\overline{w}' (h_i)$, and if $\underline{m}' (h_i)$ is significantly greater than $\overline{m}' (h_i)$, this equation requires a high ρ. This shows that the optimal policy then does not only depend on the earnings and expenditure gap between the worst-off and the average, but also on the gap in terms of marginal benefit and marginal cost of human capital.

7   Optimal tax: the non-linear case

Describing the optimal non-linear tax policy is extremely hard when the individuals differ in many dimensions and their behavior unfolds in a three-dimensional space. We will focus on a very specific aspect of the optimal tax, which is limited but nevertheless quite central in understanding the shape of the optimal policy. Our goal is to determine what sort of agent (type, behavior) will receive the greatest subsidy at the second-best optimum. Let z^* be an optimal incentive-compatible allocation. In the following proposition, t_0^* is defined as in Section 5.2. This proposition identifies

---

[14] An interior choice of bundle was assumed earlier.

a way to cut subsidies above a certain level without harming the role of t_0^* as a lower bound for well-being as measured by IT.

Proposition 6 *Assume Worst Type holds. Let $T^*(y, m)$ implement z^*, let t_0^* be the greatest value of t_0 such that $B(t_0, \overline{w}(.), \overline{m}(.)) \subset B_\cap^+(T^*)$, and let r^* be equal to the maximum of $\overline{w}(h) - \underline{w}(h) - \overline{m}(h) + \underline{m}(h)$ for $h \in [0, 1]$ such that $t_0^* + \overline{w}(h) - \overline{m}(h) \geq 0$. Then the tax function*

$$T^{**}(y, m) = \max \{T^*(y, m), -t_0^* - r^*\}$$

is feasible and satisfies

$$t_0^* \leq \min_i IT_i(z_i^{**}, R_i),$$

for any allocation z^{**} that T^{**} induces.

Proof. The new tax reduces the budget set by cutting all subsidies above $t_0^* + r^*$. Consider any $i \in N$. If $T^*(y_i, m_i) \geq -(t_0^* + r^*)$, then z_i^* is still in the smaller budget, therefore still a best bundle for i. If $T^*(y_i^*, m_i^*) < -(t_0^* + r^*)$, then z_i^* is no longer accessible to i, and at the new best bundle z_i^{**} in the smaller budget, $T^{**}(y_i^{**}, m_i^{**}) \geq -(t_0^* + r^*)$, therefore $T^{**}(y_i^{**}, m_i^{**}) > T^*(y_i^*, m_i^*)$. Therefore if there are agents i for whom z_i^* is no longer accessible in their budget, the new allocation z^{**} generates a surplus. In any case, the new allocation is feasible.

Suppose that $B(t_0^*, \overline{w}(.), \overline{m}(.)) \subset B_\cap^+(T^{**})$. Then, by Prop. 1, $t_0^* \leq \min_i IT_i(z_i^{**}, R_i)$. It is therefore sufficient to prove that $B(t_0^*, \overline{w}(.), \overline{m}(.)) \subset B_\cap^+(T^{**})$.

By Worst Type, $B_\cap^+(T^{**}) = \underline{B}^+(T^{**})$. By construction,

$$\underline{B}^+(T^{**}) = \underline{B}^+(T^*) \cap \{(c, l, h) \in X \mid c \leq \underline{w}(h) l - \underline{m}(h) + t_0^* + r^*\}.$$

As $B(t_0^*, \overline{w}(.), \overline{m}(.)) \subset \underline{B}^+(T^*)$, it is sufficient to prove that

$$B(t_0^*, \overline{w}(.), \overline{m}(.)) \subset \{(c, l, h) \in X \mid c \leq \underline{w}(h) l - \underline{m}(h) + t_0^* + r^*\}.$$

Suppose this does not hold. Then there is $(c, l, h) \in B(t_0^*, \overline{w}(.), \overline{m}(.))$ such that $c > \underline{w}(h) l - \underline{m}(h) + t_0^* + r^*$, implying

$$t_0^* + \overline{w}(h) l - \overline{m}(h) \geq c > \underline{w}(h) l - \underline{m}(h) + t_0^* + r^*,$$

therefore

$$r^* < \overline{w}(h) l - \underline{w}(h) l - \overline{m}(h) + \underline{m}(h).$$

The right-hand side is increasing in l, so one must have

$$r^* < \overline{w}(h) - \underline{w}(h) - \overline{m}(h) + \underline{m}(h).$$

In addition, as $(c, l, h) \in B(t_0^*, \overline{w}(.), \overline{m}(.))$,

$$0 \leq t_0^* + \overline{w}(h)l - \overline{m}(h) \leq t_0^* + \overline{w}(h) - \overline{m}(h).$$

One therefore obtains a contradiction with the definition of r^*. ■

The previous proposition proves that constructing T^{**} from T^* does not necessarily entail a large welfare loss in the sense that t_0^* remains a lower bound for the worst off at the allocation generated by both tax functions. The following corollary identifies the conditions under which the two tax functions are welfare equivalent, namely, the conditions under which T^{**} is actually optimal.

Corollary 7 *Under the conditions of Proposition 6, if $t_0^* = \min_i IT_i(z_i^*, R_i)$, then $t_0^* = \min_i IT_i(z_i^{**}, R_i)$ and z^* is implemented by T^{**}.*

Proof. This derives from the fact that by construction, $IT_i(z_i^{**}, R_i) \leq IT_i(z_i^*, R_i)$ for all i, and by Proposition 6, $t_0^* \leq \min_i IT_i(z_i^{**}, R_i)$. In the proof of Proposition 6 it was shown that if T^{**} cannot implement z^* (because z_i^* is no longer affordable for some i), then z^{**} generates a surplus. But if this is the case, it is possible to distribute the surplus so as to raise IT_i for every i.[15] This would contradict the fact that z^* is optimal and therefore maximizes $\min_i IT_i(z_i^*, R_i)$. ■

These results suggest that it is interesting to study T^{**}. Note that even if $t_0^* < \min_i IT_i(z_i^*, R_i)$, one has

$$t_0^* \leq \min_i IT_i(z_i^{**}, R_i) \leq \min_i IT_i(z_i^*, R_i),$$

so that if t_0^* is close to $\min_i IT_i(z_i^*, R_i)$, the allocation z^{**} is close to being optimal. Therefore, when looking at the optimal tax scheme, there is no loss, or a limited loss, of social welfare if one restricts his attention to taxes that share the salient features of T^{**}. In what follows we describe some of these features.

---

[15] Doing such a distribution while preserving incentive compatibility is not trivial. See [15] for a rigorous proof in the Mirrlees model. The argument can be extended to the present model, as the dimension of the other goods than c does not matter.

What is interesting about T^{**} is that it generates a budget $c = y - m - T^{**}(y, m)$ which lies between the hyperplane $c = y - m + t_0^* + r^*$ and the manifold defined by

$$y \le \underline{w}^*(m) \quad \text{and} \quad c = y \frac{\overline{w} \circ \underline{m}^{-1}(m)}{\underline{w}^*(m)} - \overline{m} \circ \underline{m}^{-1}(m) + t_0^*. \tag{17}$$

The former fact is a direct consequence of $T^{**}(y, m) \ge -(t_0^* + r^*)$; the latter is nothing but the translation, in (c, y, m) space, of the fact that $B(t_0^*, \overline{w}(.), \overline{m}(.)) \subset \underline{B}^+(T^{**})$. Indeed, in (c, l, h) space the equation defining $B(t_0^*, \overline{w}(.), \overline{m}(.))$ is $c = \overline{w}(h) l - \overline{m}(l) + t_0^*$. Substituting $l = y/\underline{w}(h) \le 1$ and $h = \underline{m}^{-1}(m)$ yields the manifold described by (17).

The intersection between the hyperplane and the manifold determines the sort of individual who receives the greatest subsidy. The intersection is determined by the equation

$$y - m + t_0^* + r^* = y \frac{\overline{w} \circ \underline{m}^{-1}(m)}{\underline{w}^*(m)} - \overline{m} \circ \underline{m}^{-1}(m) + t_0^*,$$

but, more simply, corresponds to the point which defines r^*, i.e., the maximum of $\overline{w}(h) - \underline{w}(h) - \overline{m}(h) + \underline{m}(h)$ for $h \in [0, 1]$ such that $t_0^* + \overline{w}(h) - \overline{m}(h) \ge 0$. This means in particular that $l = 1$ at this point, i.e., the greatest subsidy goes to full-time work. It remains to determine the value of h or m at the maximum.

The maximum can be obtained either at a point h^* satisfying

$$\overline{w}'(h^*) - \underline{w}'(h^*) - \overline{m}'(h^*) + \underline{m}'(h^*) = 0,$$

or at a point such that $t_0^* + \overline{w}(h) - \overline{m}(h) = 0$. The latter case will be obtained in particular if $\overline{w}'(h) > \underline{w}'(h)$ and $\underline{m}'(h) \ge \overline{m}'(h)$ for all h. In such a case the greatest subsidy is obtained by a P agent who works full time and has a null consumption because of great human capital expenditures.

The P agents who work full time and have lower human capital than h^*, i.e., lower expenditures than $\underline{m}(h^*)$, face a non-negative marginal rate of subsidy for human capital expenditures (on average over this part of their budget), whereas those who have greater expenditures face a non-positive rate of subsidy on average. This is due to the fact that their budget set under T^{**} has to lie below the hyperplane at which the rate of subsidy is null.

Similarly, the agents (from P or not from P) who spend $\underline{m}(h^*)$ and earn less than $\underline{w}(h^*)$ face on average over this range of earnings a non-positive marginal tax rate. But note that when $t_0^* + \overline{w}(h^*) - \overline{m}(h^*) = 0$ there are no such agents because consumption is below zero in this area.

Let us briefly compare our results to those of Fleurbaey and Maniquet [16] for the Mirrlees model (with exogenous human capital). They obtained the general conclusion that at an optimal

allocation for a similar social ordering (egalitarian-equivalent with reference wage equal to the average) it was possible to have a tax function with a marginal rate that is non-positive on average over income below the lowest wage, with a greatest subsidy granted to the least skilled individuals working full time. They relied on a preference diversity assumption.

Here we have avoided this assumption because it is restrictive in our setting, and nevertheless obtained a similar focus on the hardworking poor. It is technically interesting to understand what can be said in absence of this assumption. But there are two more important differences. First, the focus is no longer on the least skilled agents but on the agents with the least favorable dispositions. Agents with the lowest skills but better dispositions than the worst type are not considered among the worst-off here. Second, for the same reason, the advantage given to hardworking agents is now restricted to those who have important human capital expenditures. The healthy who have low skills may face substantial tax rates on earnings if this helps funding the human capital subsidy.

Of course, this more complex configuration comes in part from the fact that we studied the most general tax function $T(y,m)$ with any possible interdependence between the tax on y and the subsidy on m. The study of the special but interesting case of separate non-linear instruments $T(y), S(m)$ is not undertaken in this paper.

8 Observable human capital

We will turn now our attention to a different informational context. Let us assume that h is observed, together with c, y, m. This amounts to saying that, for instance, when it comes to education, the policy maker can observe the diplomas an agent has. Alternatively one could think of health. In this case our assumption implies that the social planner can rely on the physicians's evaluation in order to assess agents' health status. In such an informational framework the incentive-compatibility constraint becomes: for all i, j,

$$(c_i, y_i, m_i) \, R_i^* \, (c_j, y_j, m_j) \ \text{ or } \ y_j > w_i^* (m_i) \ \text{ or } \ m_i(h_j) \neq m_j(h_j).$$

As in the previous setting, agent i still has to receive an allocation that she prefers to the allocation received by agent j unless it is not possible for her to mimic agent j. This occurs either if $y_j > w_i^* (m_i)$ (exactly as in the previous framework) or if $m_i(h_j) \neq m_j(h_j)$. That is, agent i can pretend to have agent j's human capital disposition only if her human capital disposition function

crosses j's function at $h = h_j$.[16] To simplify the analysis and to better analyze the consequences of using an egalitarian social objective we introduce the following assumption:

Assumption 4 (Nested Types) The m_i functions do not cross, i.e., there is no i, j such that for some h, h', $m_i(h) > m_j(h)$ and $m_i(h') < m_j(h')$.

This assumption allows us to partition the population into different subgroups of agents having the same human capital disposition. Let K denote the set of subgroups resulting from such a partition. The fact that the human capital level is observable entails that one can conceive a different tax policy $T_k(y, m)$ for each $k \in K$.

We also introduce a further assumption which is meant to rule out a strict relation between having a good earning ability and a good human capital disposition. Whatever the human capital disposition, there is always some agent with the worst earning ability belonging to such group. Correlation is however permitted.

Assumption 5 (Uniformity): For every $k = 1, ..., K$, there is i in subgroup k such that $w_i(h) = \underline{w}(h)$.

This assumption just rules out the possibility for the policy maker of conceiving a tax scheme that is particularly harsh to some specific subgroup k just because she happens to know that no unskilled agents belong to that subgroup. Let P_k denote the subset of i from subgroup k such that $w_i = \underline{w}$. Let also $\underline{B}_k^+(T)$ denote the budget set of some agent belonging to P_k, for $k \in K$.

Consider the budget $B(t_0^*, \overline{w}(.), \overline{m}(.))$ of a hypothetical agent with average circumstances,

---

[16] An alternative specification would allow agents to "inflate" their expenditures and pretend they have a worse m function than they really have. In this case the incentive-compatibility constraint would become: for all i, j,

$$(c_i, y_i, m_i) R_i^* (c_j, y_j, m_j) \text{ or } y_j > w_i^*(m_i) \text{ or } m_i(h_j) > m_j(h_j).$$

This alternative setting would give some protection to agents with a better disposition. However the practical implications would not be very different since we rely on an egalitarian social welfare function anyway. Hence we stick to the setting presented in the main text which is simpler.

under laissez-faire except for a lump-sum transfer $t_k \in \mathbb{R}$:

$$c \leq \overline{w}(h) l - \overline{m}(h) + t_k.$$

For any $k \in K$ and for t_k small enough (possibly negative), this budget $B(t_k^*, \overline{w}(.), \overline{m}(.))$ is contained in $\underline{B}_k^+(T)$. Let t_k^* be the maximum level at which this property is satisfied. This maximum level is well defined because both $\underline{B}_k^+(T)$ and $B(t_k^*, \overline{w}(.), \overline{m}(.))$ are compact, and the latter varies continuously with t_k.

We are now able to bracket the value of $\min_i IT(z_i, R_i)$, as stated below.

Proposition 8 *Let z be an incentive-compatible allocation generated by the tax function T. Then*

$$\min_k t_k^* \leq \min_i IT(z_i, R_i) \leq \min_k \min_{i \in P_k} \left[c_i - \frac{\overline{w} \circ m_i^{-1}(m_i)}{\underline{w} \circ m_i^{-1}(m_i)} y_i + \overline{m} \circ m_i^{-1}(m_i) \right].$$

Proof. Proposition 1 implies that , for every k,

$$t_k^* \leq \min_{i \in k} IT(z_i, R_i) \leq \min_{i \in P_k} \left[c_i - \frac{\overline{w}_N \circ m_i^{-1}(m_i)}{\underline{w} \circ m_i^{-1}(m_i)} y_i + \overline{m}_N \circ m_i^{-1}(m_i) \right].$$

The conclusion then follows from the fact that

$$\min_{i \in N} IT(z_i, R_i) = \min_k \min_{i \in k} IT(z_i, R_i),$$

and from the fact that when for all $k \in K$, $a_k \leq x_k \leq b_k$, then

$$\min_k a_k \leq \min_k x_k \leq \min_k b_k.$$

∎

As far as optimal tax is concerned, the result of the previous section applies to every subgroup k separately. What is new is that an optimal tax will equalize $\min_{i \in k} IT(z_i, R_i)$ across k. This is not the same as equalizing t_k^* across k, because in absence of preference diversity, one may have $t_k^* < \min_{i \in k} IT(z_i, R_i)$ for some k.

9   Conclusion

This paper proposes a very general model in which earnings and human capital expenditures provide the basis for a redistribution policy that respects individual choices on labor and human capital, but seeks to eliminate inequalities due to inter-individual differences in the intrinsic cost to acquire human capital and in earning ability conditional on human capital.

Our main intent is to contribute to the theory of tax reform (Feldstein [10]). The idea is to see how certain ideas of fairness lead to the evaluation of arbitrary tax policies in such a general model. We find that the policy maker should primarily be interested in the part of the budget set that is attainable by agents endowed with the worst personal circumstances. Interestingly, the worst off agent, at any arbitrary incentive compatible allocation, need not be one of such agents. In typical circumstances (in particular, when the marginal tax rate on income is less than 100 percent), the part of the budget set that should be the focus of attention corresponds to the full time earnings of an agent from the worst type, at a level of human capital expenditures defined in terms of post-tax productivity of human capital expenditures.

The paper also contributes to the theory of optimal taxation. We look both at linear and non linear tax schemes. The main difference between the two cases is that in the former one human capital expenditure might actually be taxed while in the latter case human capital expenditures are subsidized on the margin, up to a level of expenditures defined in reference to the agents who receive the greatest absolute amount of subsidy.

Several extensions of this analysis can be considered. First, our analysis has ignored risk in the production of human capital and in the returns to human capital on the labor market. However, we believe that our analysis covers the most relevant case of pure idiosyncratic risk, i.e., when the policy-maker is able to predict the distribution of individual situations. It is then more respectful of the individuals' preferences to take account of this distribution rather than just the individual ex ante prospects, because what the individuals care about is their final situation.

Another extension would consider more than one dimension of human capital. While our model can be applied to education or health, it cannot be applied to both dimensions simultaneously, unless they are lumped together into a single human capital variable. The extension of the social ordering function to multidimensional human capital is straightforward, but the application to tax evaluation is less obvious because two kinds of expenditures can then be distinguished by the tax function.

A key feature of our approach, which helps a lot in obtaining results in such a general model, is the absolute priority granted to the worst-off. One may find that indexing well-being by money-metric utilities $IT_i\left(z_i^*, R_i\right)$ is sensible but resist the absolute priority. It would be interesting to see what happens to the results when a strong but finite degree of priority replaces the maximin criterion in the evaluation of taxes. This would imply paying attention to levels of income above

the levels accessible to the worst type.

Finally, actual policies are segmented and specific tax-subsidy functions operate separately on income and human capital expenditures. Our analysis of reform evaluation, fortunately, carries over to this case which is a subclass of the arbitrary tax functions studied here. The analysis of optimal linear tax, by construction, happens to satisfy this separation property. But such is not the case for optimal non-linear taxation. The methodology of Proposition 6 cannot be applied because for an optimal tax function T^* that is additively separable in earnings and human capital expenditures, the new tax function T^{**} that cuts all subsidies above a fixed level loses this property.

References

[1] A. B. Atkinson, J.E. Stiglitz, The design of tax structure: direct versus indirect taxation, Journal of Public Economics, 6(1976), 55-75.

[2] A. Blomqvist, H. Horn, Public Health Insurance and Optimal Income Taxation, Journal of Public Economics, 24(1984), 352-371.

[3] R. Boadway, L. Jacquet, Optimal Marginal and Average Income Taxation under Maximin, Journal of Economic Theory, 143(2008), 425-441.

[4] R. Boadway, M. Marchand, P. Pestieau, M. Racionero, Optimal Redistribution with Heterogeneous Preferences for Leisure, Journal of Public Economic Theory, 4(2002), 475-498.

[5] R. Boadway, P. Pestieau, Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem, in Richard Arnott, Bruce Greenwald, Ravi Kanbur and Barry Nalebuff (eds.), Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz, MIT Press (2003).

[6] P. Choné, G. Laroque, Optimal incentives for labor force participation, Journal of Public Economics, 89(2005), 395-425.

[7] P. Choné, G. Laroque, Negative Marginal Tax Rates and Heterogeneity, American Economic Review, 100(2010), 2532 47.

[8] H. Cremer, P. Pestieau, Redistributive Taxation and Social Insurance, International Tax and Public Finance, 3(1996), 281-295.

[9] K. Christensen, N. V. Holm, M. Mcgue, L. Corder, J. W. Vaupel, A Danish Population-Based Twin Study on General Health in the Elderly, Journal of Aging and Health, 11(1999), 49-64.

[10] M.S. Feldstein, On the theory of tax reform, Journal of Public Economics 6(1976), 77-104.

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

[12] M. Fleurbaey, Is commodity taxation unfair?, Journal of Public Economics 90(2006), 1765-1787.

[13] M. Fleurbaey, Fairness, responsibility and welfare. Oxford University Press (2008).

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

[15] M. Fleurbaey, F. Maniquet, Fair Income Tax, Review of Economic Studies, 73(2006), 55-83.

[16] M. Fleurbaey, F. Maniquet, Help the low-skilled or let the hardworking thrive? A study of fairness in optimal income taxation, Journal of Public Economic Theory 9(2007) 467-500.

[17] M. Fleurbaey, F. Maniquet, A Theory of Fairness and Social Welfare, Cambridge University Press (2011).

[18] M. Fleurbaey, F. Maniquet, 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 U. Press, forthcoming.

[19] M. Fleurbaey, F. Maniquet, Compensation and responsibility, in K.J. Arrow, A.K. Sen and K. Suzumura (Eds), Handbook of Social Choice and Welfare V. II, North-Holland, forthcoming.

[20] M. Grossman, L. Benham, Health, Hours and Wages, M. Perlman (ed.), The Economics of Health and Medical Care. Macmillan and Co., London (1974).

[21] R. Guesnerie, A contribution to the pure theory of taxation, Cambridge University Press (1995).

[22] J. Guo, A. Krause, Optimal Nonlinear Taxation of Income and Education Expenditures, Oxford Economic Papers, forthcoming.

[23] D. Henriett, J.C. Rochet, Is Public Health Insurance an Appropriate Instrument for Redistribution?, Annales d'economie et de statistique, 83-84(2006), 61-88.

[24] L. Jacquet, D. Van de gaer, A Comparison of Optimal Tax Policies when Compensation or Responsibility Matter, Journal of Public Economics (forthcoming).

[25] B. Jacobsyand, A. L. Bovenbergz, Optimal Taxation of Human Capital and the Earnings Function, Journal of Public Economic Theory (forthcoming).

[26] H. Luft, The Impact of Poor Health on Earnings. Review of Economics and Statistics, 57 (1975), 43-57.

[27] D. Maldonado, Education Policies and Optimal Taxation, International Tax and Public Finance, 15(2008), 131-143.

[28] M.G. Marmot, G. Davey Smith, S. Stansfeld, C. Patel, F. North, J. Head, I. White, E. Brunner, A. Feeney, Health inequalities among British civil servants: the Whitehall II study, Lancet 337(1991), 1387-1393.

[29] J. Mirrlees, An exploration in the theory of optimum income taxation, Review of Economic Studies 38(1971), 175-208.

[30] S. Mushkin, Health as an Investment, Journal of Political Economy 70(1962), 129-157.

[31] T. Piketty, E. Saez, Optimal Labor Income Taxation, NBER Working Paper 18521.

[32] D.D. Reid, G.Z. Brett, P.J.S. Hamilton, R.J. Jarrett, H. Keen, G. Rose, Cardio-respiratory disease and diabetes among middle-aged male civil servants: a study of screening and intervention. Lancet 303(1974), 469-473.

[33] J.C. Rochet, Incentives, Redistribution and Social Insurance, Geneva Papers on Risk and Insurance Theory, 16(1991), 143-165.

[34] J.C. Rochet,The taxation principle and multi-time Hamilton-Jacobi equations, Journal of Mathematical Economics, 14(1985), 113-128.

[35] E. Saez, Optimal income transfer programs: intensive versus extensive labor supply responses, Quarterly Journal of Economics, 117(2002) , 1039-1073.

[36] G. Valletta, Health, fairness and taxation, CORE DP 16 (2012)

Appendix

We list here the axioms that characterize the social welfare function used in the paper. For a description of the normative implications of the axioms and for a formal proof of the characterization one can see Valletta [36]).

Strong Pareto: For all $e \in \mathcal{D}$, $z, z' \in Z$ if, for all $i \in N$, $z'_i R_i z_i$ then $z' \widetilde{R}(e) z$. If moreover, for some $j \in N$, $z'_j P_j z_j$ then $z' \widetilde{P}(e) z$.

Equal Well-being for Equal Preferences For all $e \in \mathcal{D}$, $z, z' \in Z$, if there exist $i, j \in N$ and some $\Delta > 0$, such that $R_i = R_j$, $l_i = l_j = l'_i = l'_j$, $h_i = h_j = h'_i = h'_j$ with $z_k = z'_k$ for all $k \neq i, j$,

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

then $z' \widetilde{P}(e) z$; if otherwise

$$c'_i = c_j \text{ and } c_i = c'_j$$

then $z' \widetilde{I}(e) z$.

Uniform Circumstances Neutrality: For all $e \in \mathcal{D}$, $z, z' \in Z$, if for all $i, j \in N$, $w_i(.) = w_j(.)$, $m_i(.) = m_j(.)$, and if there exist $m, n \in N$ and some $\Delta > 0$ such that, $z_m \in \max|_{R_m} B(t_m, w_m(.), m_m(.))$, $z'_m \in \max|_{R_m} B(t'_m, w_m(.), m_m(.))$, $z_n \in \max|_{R_n} B(t_n, w_n(.), m_n(.))$, $z'_n \in \max|_{R_n} B(t'_n, w_n(.), m_n(.))$, with $z_k = z'_k$ for all $k \neq m, n$,

$$t_m - \Delta = t'_m > t'_n = t_n + \Delta$$

then $z' \widetilde{P}(e) z$; if otherwise

$$t'_m = t_n \text{ and } t_m = t'_n$$

then $z' \widetilde{I}(e) z$.

Independence: For all $z, z' \in Z$, $e, e' \in \mathcal{D}$, with $e = (R, w, m)$ and $e' = (R', w, m)$, if for all $i \in N$ and $q \in Z$,

$$z_i I_i q \iff z_i I_i' q$$

$$z_i' I_i q \iff z_i' I_i' q$$

then

$$z' \widetilde{R}(e) z \iff z' \widetilde{R}(e') z.$$

For any $S \subseteq N$ let $\overline{m}_S(.) = \frac{1}{|S|} \sum_{j \in S} m_j(.)$ and $\overline{w}_S(.) = \frac{1}{|S|} \sum_{j \in S} w_j(.)$ denote, respectively, the average health disposition mapping and the average earning ability mapping within the group S. For the ease of notation let $\overline{m}_N(.) = \overline{m}(.)$ and $\overline{w}_N(.) = \overline{w}(.)$.

Separation: For all $e \in \mathcal{D}$ and for all $z, z' \in Z$, if there is $S \subseteq N$ such that $\overline{m}_S(.) = \overline{m}(.)$, $\overline{w}_S(.) = \overline{w}(.)$ and, for all $i \in S$, $z_i = z_i'$ then

$$z' \widetilde{R}(e) z \iff z'_{-S} \widetilde{R}(R_{-S}, w_{-S}(.), m_{-S}(.)) z_{-S}.$$

Theorem: *On the domain $\mathcal{D}$ a social ordering function satisfies* Strong Pareto, Equal Well-being for Equal Preferences, Uniform Circumstances Neutrality, Independence *and* Separation, *if and only if it is an Average Circumstances Egalitarian Equivalent Leximin function.*