Dondena Working Papers

Carlo F. Dondena Centre for Research on
Social Dynamics and Public Policy

Population Dynamics and Health Unit

Optimal income taxation without single-crossing

Spencer Bastani
Sören Blomquist
Luca Micheletto

Working Paper No. 129
June 2019

Università Bocconi • The Dondena Centre
Via Guglielmo Röntgen 1, 20136 Milan, Italy
http://www.dondena.unibocconi.it

ISSN-2035-2034

Optimal income taxation without single-crossing

Spencer Bastani* Sören Blomquist† Luca Micheletto‡

June 25, 2019

Abstract

In this paper we set up a simple two-type optimal nonlinear income tax model where the single-crossing condition is violated, and we characterize the properties of a second-best optimum by considering the entire second-best Pareto frontier. The violation of single-crossing is generated by the assumption that agents differ both in terms of market abilities and in terms of their needs for a work-related good. Our analysis highlights several non-standard features of a second-best optimum. In particular, we show that a nonlinear income tax may allow the government to convert a pooling laissez-faire equilibrium into a separating equilibrium, that the second-best Pareto frontier may be discontinuous, and that a second-best optimum may not preserve the income ranking prevailing under laissez-faire. Finally, we also show that at a second-best optimum the labor supply of some agents might be distorted even though no self-selection constraint is (locally) binding in equilibrium.

JEL classification: H21, D82, D86.

Keywords: Optimal nonlinear income taxation; single-crossing condition; multidimensional heterogeneity; redistribution.

*Department of Economics and Statistics, Linnaeus University; Linnaeus University Center for Integration and Discrimination Studies; Uppsala Center for Fiscal Studies; Uppsala Center for Labor Studies; CESifo, Germany. E-mail: spencer.bastani@lnu.se.

†Department of Economics, Uppsala University; Uppsala Center for Fiscal Studies; CESifo, Germany. E-mail: soren.blomquist@nek.uu.se.

‡Department of Law "C. Beccaria", University of Milan; Dondena Center for Research on Social Dynamics, Bocconi University; Uppsala Center for Fiscal Studies; CESifo, Germany. E-mail: luca.micheletto@unibocconi.it.

1 Introduction

The important and influential literature growing out of Mirrlees' (1971) seminal paper on optimal income taxation has stressed the trade-offs between incentive and distributional considerations in the design of income tax schedules. These trade-offs arise from an information friction that endogenizes the feasible tax instruments: the government knows the distribution of types in the population and it can also observe the actual earned income of each individual, but is not able to observe the specific type of any given individual. Personalized lump-sum taxes and transfers are therefore not available but public observability of earned income at the individual level allows the government to tax earned income on a nonlinear scale.

Given that any nonlinear income tax schedule defines a link between earned income and after-tax income, the government's problem can be equivalently formulated as the problem of assigning, to each type of agent, a bundle in the (pre-tax income, after-tax income)-space, subject to a public budget constraint and to a set of self-selection (incentive-compatibility) constraints, which require that each individual be better off with the bundle intended for him/her than with any other available bundle.

The vast majority of papers in the optimal tax literature assume that agents differ along a single dimension (market ability). This is due to tractability considerations. Given certain assumptions on the utility function, it enables a monotonic relationship between an agent's unobserved type and the slope of his/her indifference curve in the earnings-consumption space. This property, referred to as 'single-crossing' (hereafter, SC), allows the researcher to provide a full characterization of the set of implementable contracts while restricting attention to local incentive constraints linking adjacent types. In the case of a continuum of types, it also implies that the incentive constraints can conveniently be expressed in terms of differential equations. When agents differ along multiple dimensions, however, the SC property will generally be violated, as there is no natural way to order agents in a multidimensional space.[1]

A comparatively small literature analyzes optimal income taxation with multidimensional unobserved heterogeneity, and these contributions can roughly be divided into four strands. A first strand assumes that the additional dimensions of heterogeneity enters additively separable in the utility function, thereby not affecting individuals' trade-offs between pre-tax and after-tax income (see e.g., Kleven, Kreiner and Saez, 2009; Jacquet et al., 2013; Scheuer, 2014; Bastani, Blomquist and Micheletto, 2017). A second strand

[1] Multidimensional heterogeneity is however not a necessary condition to generate violation of SC. See, for instance, Gahvari (2007) and Ho and Pavoni (2018).

imposes restrictions such that the various dimensions of heterogeneity can be collapsed into one dimension and parameterized by a single index (see, e.g., Boadway et al.; 2002; Choné and Laroque, 2010; Golosov et al., 2013; Rothschild and Scheuer, 2014; Lockwood and Weinzierl, 2015). A third strand analyzes more general forms of heterogeneity, but focuses attention to a quantitative analysis of models with a small discrete number of types (see, e.g., Bastani, Blomquist and Micheletto, 2013; Judd et al., 2018). Finally, a fourth strand comprises papers that provide a characterization of optimal marginal tax rates while remaining agnostic about which incentive-compatibility constraints are binding in equilibrium (see, e.g., Cremer, Gahvari and Ladoux, 1998; Cremer and Gahvari, 2002; Micheletto, 2008).

Compared to the existing literature referred above, the purpose of this paper is to provide a more thorough investigation of the consequences descending from abandoning the SC condition. For this purpose, we set up a simple two-type model where the SC condition is naturally violated, and we characterize the properties of a second-best optimum by considering the entire second-best Pareto frontier (hereafter, PF).[2] The model that we consider is a standard intensive-margin optimal income tax model where agents have identical preferences and heterogeneous market abilities, but where we also allow for heterogeneity in "needs" for a work-related good/service, i.e. a good/service that some agents need to purchase in order to work.[3] It is this bi-dimensional heterogeneity that implies a violation of the SC condition.

Our analysis highlights several results, each of them representing an anomaly with respect to what is obtained in an optimal income tax model under SC. First of all, a second-best optimum might not preserve the earned-income ranking that prevails under laissez-faire. Second, redistribution via income taxation might be feasible even when the laissez-faire equilibrium is a pooling equilibrium. Third, a second-best optimum might not be unique, in the sense that there might be more than one set of allocations in the (pre-tax income, after-tax income)-space that solve the government's maximization problem. Fourth, the support of the function describing the PF might be a non-connected set. Fifth, supplementing an optimal nonlinear income tax with an optimal subsidy on work-related expenses may imply that at a second-best optimum redistribution is achieved through a separating- or pooling equilibrium where both self-selection constraints are binding. A

[2]A similar exercise has been undertaken by Bierbrauer and Boyer (2014) for a two-type optimal nonlinear income tax model where individuals have linear effort costs and the SC-condition holds.

[3]Several interpretations are possible. One example is day care services which are needed by parents of young kids in order to work. Other groups who might face needs constraints include workers with relatives who require elderly care, or workers who incur commuting costs or work-related health costs.

final result that we show is that at a second-best optimum it might be the case that the labor supply of some agents is distorted even though no self-selection constraint is (locally) binding in equilibrium.

The paper is organized as follows. In section 2 we present our setting and highlight how it implies that the SC condition does not hold. In Section 3 we characterize the optimal distortions under various assumptions regarding the redistributive goals of the government. To simplify the exposition in this section we make the assumption that, for agents who have to incur a cost for the purchase of a work-related good, the cost is proportional to their labor supply. This assumption is relaxed in Section 4 where we allow for nonlinear, convex or concave, cost functions. In section 5 we discuss how our results would be affected by subsidizing work-related expenses. Finally, section 6 offers concluding remarks.

2 The model

Consider an economy populated by two groups of individuals who have identical preferences represented by the quasi-linear utility function

$$U = c - \frac{1}{1+1/\beta} h^{1+1/\beta}, \tag{1}$$

where c denotes consumption, h denotes labor supply, and where β is a positive constant representing the elasticity of labor supply.[4]

The two groups of agents are assumed to differ with respect to their market ability, reflected in their hourly wage rate, and their needs for a work-related good. One group has no need for any work-related good, whereas agents belonging to the other group incur a monetary cost $\varphi(h)$ which is a (weakly) increasing function of labor supply h. Throughout the paper we will refer to these groups of agents as "non-users" and "users" and denote their hourly wage rates by, respectively, w^n and w^u (superscript "n" referring to non-users, and superscript "u" referring to users). Furthermore, we will assume that $w^u > w^n$, i.e. that the high-skilled agents are disadvantaged along our second dimension of heterogeneity.

Assume that the government levies a nonlinear income tax $T(wh)$ and let earned income be denoted by Y (i.e., $Y \equiv wh$) and after-tax income be denoted by B (i.e., $B \equiv Y - T(Y)$). It is straightforward to notice that the SC property is not satisfied

[4] The specific iso-elastic form of the utility function is here mainly adopted for analytical convenience but has been used extensively in the optimal tax literature as well as in the empirical literature estimating behavioral responses to tax changes.

in the two-type economy that we are considering. This property requires that, at any bundle in the (Y, B)-space, the indifference curves are flatter the higher the wage rate of an agent. In our model, and for a given (Y, B)-bundle, users and non-users have utilities that are respectively given by:

$$\begin{aligned} U^u &= B - \varphi(\frac{Y}{w^u}) - \frac{1}{1+1/\beta}\left(\frac{Y}{w^u}\right)^{1+1/\beta}, \\ U^n &= B - \frac{1}{1+1/\beta}\left(\frac{Y}{w^n}\right)^{1+1/\beta}. \end{aligned}$$

Therefore, for a given (Y, B)-bundle, users have an indifference curve with slope equal to

$$MRS^u_{YB}(Y, B) \equiv -\frac{\partial U^u/\partial Y}{\partial U^u/\partial B} = \frac{1}{w^u}\left[\varphi'(\frac{Y}{w^u}) + \left(\frac{Y}{w^u}\right)^{1/\beta}\right], \tag{2}$$

whereas non-users have an indifference curve with slope equal to

$$MRS^n_{YB}(Y, B) \equiv -\frac{\partial U^n/\partial Y}{\partial U^n/\partial B} = \frac{1}{w^n}\left(\frac{Y}{w^n}\right)^{1/\beta}. \tag{3}$$

Comparing (2) and (3), and taking into account that $\varphi' \geq 0$ and $w^u > w^n$, one can see that the sign of the difference $MRS^u_{YB} - MRS^n_{YB}$ will depend in general on the specific (Y, B)-bundle that is considered.

The fact that the SC property is not satisfied in our setting shows that our bi-dimensional heterogeneity (in skills and needs) cannot be reduced to one dimension. Albeit this complicates the analysis, it also allows us to highlight some interesting results. To keep the analysis as simple as possible and illustrate the anomalies that may arise when the SC does not hold, in the next section we will evaluate the properties of a second-best optimum under the assumptions that i) the work-related costs for users are proportional to labor supply, so that $\varphi(h) = qh$ (where q is a positive constant such that $q < w^u$), and ii) β in (1) is equal to 1. Before doing that, however, we will devote the remainder of this section to characterizing the laissez-faire equilibrium and the properties of the first-best PF.

Under laissez-faire with $\varphi(h) = qh$ and $\beta = 1$, users choose h to maximize $(w^u - q)h - h^2/2$, implying $h^u = w^u - q$, whereas non-users choose h to maximize $w^n h - h^2/2$, implying $h^n = w^n$.

Therefore, denoting by Y^i_{LF}, for $i = n, N$, the laissez-faire income of an individual i, we have that $Y^n_{LF} = (w^n)^2$, $Y^u_{LF} = (w^u - q)\, w^u$, and

$$Y^u_{LF} < (>)\, Y^n_{LF} \Longleftrightarrow (w^u - q)\, w^u < (>)\, (w^n)^2.$$

Regarding utilities, denoting by U^i_{LF}, for $i = n, N$, the utility of an individual i under laissez-faire, we have that $U^u_{LF} = (w^u - q)^2/2$, $U^n_{LF} = (w^n)^2/2$, and

$$U^u_{LF} < (>)\, U^n_{LF} \Longleftrightarrow w^u - q < (>)\, w^n.$$

One thing to notice is that the utility ranking and the income ranking may differ. In particular, while $Y^u_{LF} \leq Y^n_{LF}$ implies that $U^u_{LF} < U^n_{LF}$, knowing that $Y^u_{LF} > Y^n_{LF}$ is not enough to establish who is better off under laissez faire. When $Y^u_{LF} > Y^n_{LF}$, we can have that $U^u_{LF} < U^n_{LF}$ (when $(w^u - q)\, w^u > (w^n)^2 > (w^u - q)^2$), $U^u_{LF} = U^n_{LF}$ (when $(w^u - q)\, w^u > (w^n)^2 = (w^u - q)^2$), or $U^u_{LF} > U^n_{LF}$ (when $(w^u - q)\, w^u > (w^u - q)^2 > (w^n)^2$).

In a first-best setting where asymmetric information is not an issue, the shape of the PF can be straightforwardly characterized. For this purpose, normalize to one the size of the total population, and let π denote the proportion of users. The first-best PF goes through the point with coordinates (U^n_{LF}, U^u_{LF}) and has slope $dU^u/dU^n = -(1-\pi)/\pi$ for values of U^n such that $-(w^n)^2/2 \leq U^n \leq \left[(w^u - q)^2 \pi/(1-\pi)\right] + (w^n)^2/2$; for $U^n > \left[(w^u - q)^2 \pi/(1-\pi)\right] + (w^n)^2/2$ the slope of the PF is such that $dU^u/dU^n < -(1-\pi)/\pi$; for $U^n < -(w^n)^2/2$ the slope is such that $-(1-\pi)/\pi < dU^u/dU^n < 0$.

The intuition is easy to grasp. Starting from the laissez-faire equilibrium, a 1\$ lump-sum tax levied on non-users, which reduces by 1 the utility of each of them, allows the government to collect \$$1-\pi$, which implies that each user can receive a lump-sum transfer of \$$(1-\pi)/\pi$, raising by $(1-\pi)/\pi$ their per capita utility. This kind of income- and utility-redistribution, from non-users to users, can go on until all the income earned by non-users under laissez-faire, i.e. $(w^n)^2$, is confiscated by the government. At that point we have that $U^n = -(w^n)^2/2$ (consumption for non-users is equal to zero and, with no income effects on labor supply, their labor supply remains undistorted at the laissez-faire level) and $U^u = \left[(w^n)^2 (1-\pi)/\pi\right] + (w^u - q)^2/2$. Once this point on the first-best PF is reached, and assuming that zero represents the lower bound for individual consumption,[5] a further increase in U^u can only be obtained by pushing the labor supply of non-users above its undistorted level $h^n = w^n$ (while keeping at zero their consumption), so that additional resources can be transferred to users.[6] However, due to the distortion on the labor supply of non-users, redistribution becomes costlier and the slope of the PF

[5] One can think that individual consumption cannot fall below a subsistence level $\overline{c}$. From this perspective, assuming that $\overline{c} = 0$ is simply a matter of normalization.

[6] The fact that the non-negativity constraint on consumption may become binding along the first-best PF is an artifact of our assumption that utility is linear in consumption. The constraint could be safely disregarded if the marginal utility of consumption goes to infinity as consumption approaches zero.

becomes equal to $dU^u/dU^n = -(1-\pi)\, w^n/\pi h^n$, which is greater than $-(1-\pi)/\pi$ when h^n exceeds w^n, i.e. its laissez-faire value.[7]

3 Pareto-efficient taxation when the cost of the work-related good is proportional to labor supply

Let's now consider the government's problem under the assumption that an agent's type is not directly observable. As customary in the optimal income tax literature, we will adopt a mechanism design approach assuming that the government chooses optimally two bundles in the (Y, B)-space subject to the requirement that the chosen set of bundles satisfies public-budget balance, incentive-compatibility, and non-negativity constraints on both consumption and labor supply. Denoting by (Y^u, B^u) the bundle intended for users and by (Y^n, B^n) the one intended for non-users, a Pareto-efficient tax problem can then be formalized as follows:

$$\max_{Y^u, B^u, Y^n, B^n} \quad B^u - \frac{q}{w^u} Y^u - \frac{1}{2}\left(\frac{Y^u}{w^u}\right)^2$$

subject to:

$$B^n - \frac{1}{2}\left(\frac{Y^n}{w^n}\right)^2 \geq \overline{V}^n, \qquad (\nu)$$

$$(Y^u - B^u)\,\pi + (Y^n - B^n)(1-\pi) \geq 0, \qquad (\mu)$$

$$B^n - \frac{1}{2}\left(\frac{Y^n}{w^n}\right)^2 \geq B^u - \frac{1}{2}\left(\frac{Y^u}{w^n}\right)^2, \qquad (\lambda)$$

$$B^u - \frac{q}{w^u} Y^u - \frac{1}{2}\left(\frac{Y^u}{w^u}\right)^2 \geq B^n - \frac{q}{w^u} Y^n - \frac{1}{2}\left(\frac{Y^n}{w^u}\right)^2, \qquad (\phi)$$

and subject to a set of non-negativity constraints on consumption and labor supply for both agents:

$$Y^u \geq 0,\; Y^n \geq 0,\; B^n \geq 0,\; B^u - \frac{q}{w^u} Y^u \geq 0.$$

In the problem above, the ν-constraint prescribes a lower bound for the utility of non-users, the μ-constraint represents the government's budget constraint (the resource

[7] A similar reasoning can be adopted to show that the slope of the first-best PF is equal to $-(1-\pi)/\pi$ for values of $U^n > U^n_{LF}$ and such that $(w^n)^2/2 < U^n \leq \left[(w^u - q)^2 \pi/(1-\pi)\right] + (w^n)^2/2$. When $U^n = \left[(w^u - q)^2 \pi/(1-\pi)\right] + (w^n)^2/2$, all the resources available for consumption by users under laissez-faire have been transferred to non-users. Since consumption for users has then reached its lower bound, a further increase in the utility of non-users can only be obtained by requiring users to increase their labor supply, while keeping at zero their consumption, so that additional resources can be transferred to non-users. However, since the required increase in h^u entails a distortion on the labor supply of users, redistribution becomes costlier and the slope of the PF becomes equal to $dU^u/dU^n = -(1-\pi)\, h^u/\pi (w^u - q)$, which is lower than $-(1-\pi)/\pi$ when h^u exceeds $w^u - q$, i.e. its laissez-faire value.

constraint of the economy), the λ-constraint the self-selection constraint requiring non-users not to be tempted to choose the bundle intended for users, and finally the ϕ-constraint the self-selection constraint requiring users not to be tempted to choose the bundle intended for non-users. For a given value of $\overline{V}^n$, the corresponding set of admissible bundles is the set of bundles satisfying the remaining constraints (including the non-negativity constraints on consumption and labor supply for each agent). Notice that, by varying the value selected for $\overline{V}^n$, one can characterize the entire second-best PF.[8] Denoting by $U^u\left(\overline{V}^n\right)$ the function describing the second-best Pareto frontier, its domain will be given by the values for $\overline{V}^n$ such that the set of admissible bundles is non-empty and the ν-constraint is binding.

At any given bundle in the (Y, B)-space, the marginal rate of substitution for a user is given by: $MRS^u_{YB} = \left(q + \frac{Y}{w^u}\right)/w^u$ whereas for non-users it is given by $MRS^n_{YB} = Y/\left(w^n\right)^2$. Thus, users and non-users will have equally sloped indifference curves at bundles where

$$Y = \frac{q}{w^u}\left[\frac{1}{\left(w^n\right)^2} - \frac{1}{\left(w^u\right)^2}\right]^{-1} = \frac{qw^u}{\left(w^u\right)^2 - \left(w^n\right)^2}\left(w^n\right)^2 \equiv \Omega > 0, \tag{4}$$

whereas at bundles where $Y > (<)\,\Omega$, users will have flatter (steeper) indifference curves than non-users.

Given that for both types of agents the slope is increasing in Y, and since users have steeper indifference curves for $Y < \Omega$, it follows that there are three possible configurations of a laissez-faire equilibrium: $Y^u_{LF} < Y^n_{LF} < \Omega$; $\Omega < Y^n_{LF} < Y^u_{LF}$; $Y^u_{LF} = Y^n_{LF} = \Omega$.[9] A graphical illustration of the violation of SC that occurs in our model is provided in Figure 1 below for the case when $Y^u_{LF} < Y^n_{LF}$.

We will denote by respectively $T'\left(Y^u_{SB}\right)$ and $T'\left(Y^n_{SB}\right)$ the marginal income tax rate faced by users and non-users at a second-best optimum. As customary in the optimal tax literature, the marginal income tax rate faced by an individual at a given point in the (Y, B)-space is defined as $1 - MRS_{YB}$.

[8] Alternatively and equivalently, we could have skipped the ν-constraint and considered an objective function of the form $\alpha^N f\left(B^u - \frac{q}{w^u}Y^u - \frac{1}{2}\left(\frac{Y^u}{w^u}\right)^2\right) + \alpha^n f\left(B^n - \frac{1}{2}\left(\frac{Y^n}{w^n}\right)^2\right)$ with α^N and α^n being social welfare weights and f being an increasing and concave function. Following this approach one could characterize the entire second-best PF by properly varying α^N and α^n.

[9] Formally, assume that $Y^u_{LF} < Y^n_{LF}$, i.e. $\left(w^u - q\right)w^u < \left(w^n\right)^2$; given that $w^u > w^n$, we have that $\left(w^u - q\right)w^u < \left(w^n\right)^2 < \left(w^u\right)^2$ and therefore $\left[\left(w^u\right)^2 - \left(w^n\right)^2\right]/qw^u < 1$. It then follows that Ω in (4) is greater than $\left(w^n\right)^2$, and therefore $Y^u_{LF} < Y^n_{LF} < \Omega$. Similarly, assume that $Y^u_{LF} > Y^n_{LF}$, i.e. $\left(w^u - q\right)w^u > \left(w^n\right)^2$; given that $w^u > w^n$, we have that $\left(w^n\right)^2 < \left(w^u - q\right)w^u < \left(w^u\right)^2$ and therefore $\left[\left(w^u\right)^2 - \left(w^n\right)^2\right]/qw^u > 1$. It then follows that Ω in (4) is smaller than $\left(w^n\right)^2$, and therefore $Y^u_{LF} > Y^n_{LF} > \Omega$.

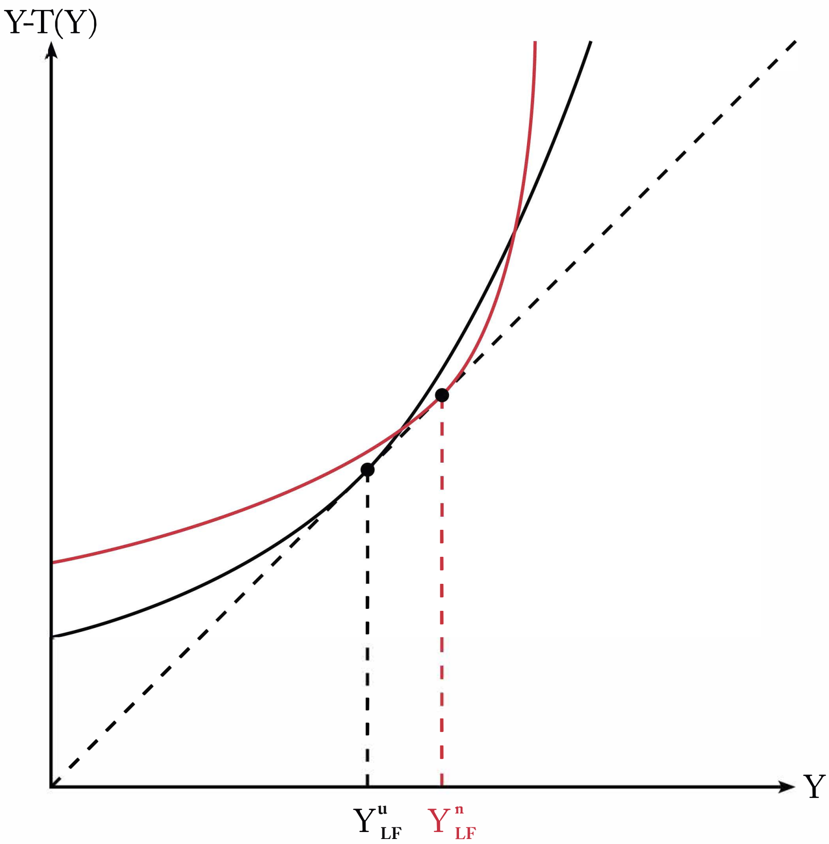

Figure 1

In the following Proposition we will characterize the properties of a second-best optimum for values of $\overline{V}^n$ such that $\overline{V}^n < U^n_{LF}$, so that the intended direction redistribution goes from non-users to users. In a separate Proposition we will characterize the properties of a second-best optimum for values of $\overline{V}^n$ such that $\overline{V}^n > U^n_{LF}$, so that redistribution goes in the opposite direction. The reason for treating separately the two cases ($\overline{V}^n < U^n_{LF}$ and $\overline{V}^n > U^n_{LF}$) is that, as we will be clear from our discussion, most of the anomalies due to the violation of SC arise when redistribution goes from non-users to users. For each Proposition we will distinguish between the three possible income rankings under laissez-faire.

Let's then start our characterization of a second-best optimum by considering a setting where $\overline{V}^n < U^n_{LF}$, so that the ϕ-constraint in the government's problem is necessarily slack.[10]

[10]Notice also that, despite the fact that the single-crossing property does not hold, the λ- and ϕ-constraint cannot be both binding at an optimum unless the two groups are pooled under laissez faire (i.e. $Y^u_{LF} = Y^n_{LF}$) and separation of types cannot be achieved by means of a nonlinear income tax. The reason is the following. Assume that at a second-best optimum we obtain a separating equilibrium where both self-selection constraints are binding. With the μ-constraint being binding, one bundle will

Proposition 1 *Assume that* $\overline{V}^n < U^n_{LF}$*, so that the intended direction of redistribution is from non-users to users. We have that:*

i) When $Y^n_{LF} = Y^u_{LF}$*, the laissez-faire equilibrium will be second-best optimal provided that* $\pi \geq q/w^u$*. With* $\pi < q/w^u$*, we have that:*

a) If $\overline{V}^n \geq (1-\pi)U^n_{LF}$ *then there are two equivalent second-best optima where* $T'(Y^n_{SB}) = 0$*, one with* $T'(Y^u_{SB}) < 0$ *and another one with* $T'(Y^u_{SB}) > 0$*.*

b) If $-U^n_{LF} \leq \overline{V}^n < (1-\pi)U^n_{LF}$ *we have that* $T'(Y^n_{SB}) = 0$ *and* $T'(Y^u_{SB}) < 0$*.*

c) If $\overline{V}^n < -U^n_{LF}$*, we have that* $T'(Y^n_{SB}) < 0$ *and* $T'(Y^u_{SB}) < 0$*.*

ii) When $Y^n_{LF} < Y^u_{LF}$*, we have that:*

a) If $\overline{V}^n \geq U^n_{LF} - \frac{\pi}{2}\frac{(Y^u_{LF}-Y^n_{LF})^2}{Y^n_{LF}}$*, then* $T'(Y^n_{SB}) = 0$ *and* $T'(Y^u_{SB}) = 0$*.*

b) If $\overline{V}^n < U^n_{LF} - \frac{\pi}{2}\frac{(Y^u_{LF}-Y^n_{LF})^2}{Y^n_{LF}}$*, then* $T'(Y^n_{SB}) \leq 0$ *and* $T'(Y^u_{SB}) < 0$*.*

iii) When $Y^n_{LF} > Y^u_{LF}$*, we have that:*

a) If $\pi \geq 1 - \left(\frac{w^n}{w^u}\right)^2$*, then* $T'(Y^n_{SB}) = 0$ *and* $T'(Y^u_{SB}) \geq 0$*.*[11]

b) If $\pi < 1 - \left(\frac{w^n}{w^u}\right)^2$ *and* $\overline{V}^n \geq (1-\pi)U^n_{LF}$*, then* $T'(Y^n_{SB}) = 0$ *and* $T'(Y^u_{SB}) \geq 0$*.*[12]

c) If $\pi < 1 - \left(\frac{w^n}{w^u}\right)^2$ *and* $\overline{V}^n < (1-\pi)U^n_{LF}$*, it is then possible that* $T'(Y^n_{SB}) \leq 0$ *and* $T'(Y^u_{SB}) < 0$*.*

Proof. See Appendix A. ■

Part *i)* of Proposition 1 shows that, provided that the proportion of users is sufficiently small ($\pi < q/w^u$), it is feasible for the government to redistribute from non-users to users even in cases when both types earn the same amount of income under laissez-faire. This possibility hinges on the violation of the SC condition; under SC an anonymous nonlinear

be associated with a positive tax payment and another one with a negative tax payment. Then the government could improve upon the initial set of allocations by implementing a pooling allocation where all agents are offered the bundle to which is associated a positive tax payment (the utility of all agents would be unaffected and the government would run a positive surplus). But this cannot be an optimum either, since the government's budget constraint would be slack. Thus, either we have cases where no self-selection constraint is binding (and the second-best implements the first-best optimum) or cases where only one of the two self-selection constraints is binding. The argument relies on the assumption that a nonlinear income tax is the only instrument at disposal for the government. As we will see in Section 5, when other policy instruments affect the revenue collected by the government at a given (Y, B)-bundle, it might happen that at a second-best optimum both self-selection constraints are binding.

[11] $T'(Y^u_{SB}) > 0$ when $\overline{V}^n < U^n_{LF} - \frac{\pi}{2}\frac{(Y^u_{LF}-Y^n_{LF})^2}{Y^n_{LF}}$.

[12] $T'(Y^u_{SB}) > 0$ when $\overline{V}^n \in \left[(1-\pi)U^n_{LF}, U^n_{LF} - \frac{\pi}{2}\frac{(Y^u_{LF}-Y^n_{LF})^2}{Y^n_{LF}}\right)$.

income tax would not allow the government to convert a pooling laissez-faire equilibrium into a separating equilibrium.

Furthermore, part *i)* shows that, (for $\pi < q/w^u$ and) as long as $(1-\pi)U^n_{LF} \leq \overline{V}^n < U^n_{LF}$, the optimal bundle intended for users is not unique; in particular, for each $\overline{V}^n \in [(1-\pi)U^n_{LF}, U^n_{LF})$ there are two equivalent second-best optima (in the sense of entailing the same value for U^u_{SB}), one entailing a downward distortion on the labor supply of users ($T'(Y^u_{SB}) > 0$) and one entailing an upward distortion on their labor supply ($T'(Y^u_{SB}) < 0$). Intuitively, the reason why there are two equivalent second-best optima is that, for a given $\overline{V}^n \in [(1-\pi)U^n_{LF}, U^n_{LF})$, the magnitude of the distortion needed to achieve type separation is the same irrespective of its direction (downward or upward).

On the other hand, starting from $\overline{V}^n = (1-\pi)U^n_{LF}$, a reduction in $\overline{V}^n$ allows for the possibility to further raise the utility of users but in this case a second-best optimum necessarily requires an upward distortion on the labor supply of users ($T'(Y^u_{SB}) < 0$). The reason is that, for $\overline{V}^n = (1-\pi)U^n_{LF}$, the second-best optimum entailing a downward distortion on the labor supply of users requires to offer them the bundle $(Y, B) = (0, (1-\pi)U^n_{LF})$. At this bundle the labor supply of users is pushed to its lower bound, implying that a further reduction in $\overline{V}^n$ cannot be accommodated by magnifying the downward distortion on the labor supply of users. Thus, for $\overline{V}^n < (1-\pi)U^n_{LF}$ a second-best optimum will necessarily entail $T'(Y^u_{SB}) < 0$.

Regarding the labor supply of non-users, it will be left undistorted as long as $\overline{V}^n$ does not fall below $-U^n_{LF}$. For $\overline{V}^n < -U^n_{LF}$, instead, the labor supply of non-users will be upward distorted too. Intuitively, the reason is that, by leaving undistorted their labor supply (i.e. prescribing $Y^n = (w^n)^2$), it is not feasible to lower their utility below $-U^n_{LF} = -(w^n)^2/2$ (taking into account that $U^n = B^n - \frac{1}{2}\left(\frac{Y^n}{w^n}\right)^2$ and assuming that consumption can only take non-negative values, so that $B^n \geq 0$). Thus, in order to collect from each non-user a tax that is larger than $(w^n)^2$, their labor supply needs to be upward distorted.[13]

Part *ii)* considers the case when users earn more than non-users under laissez-faire. It shows that when the extent of redistribution from non-users to users is, loosely speaking, small, the second-best optimum will coincide with the first-best optimum and no distortion is needed to maintain incentive-compatibility. For intermediate degrees of redistribution only the labor supply of users will be distorted (by letting them face a negative marginal tax rate). Finally, if the redistributive goals pursued by the government

[13]Notice that, for $\overline{V}^n < -U^n_{LF}$, besides having $T'(Y^u_{SB}) < 0$ and $T'(Y^n_{SB}) < 0$, we also have that the average income tax rate for non-users is 100%.

are sufficiently strong ($\overline{V}^n < -U^n_{LF}$), it might be the case that both types face an upward distortion on their labor supply. However, as shown in the appendix (see the proof of Proposition 1), for this possibility to occur it is necessary that the proportion of users in the population is not too large.[14]

An aspect that is worth emphasizing about part *ii)* of Proposition 1 is that $Y^n_{LF} < Y^u_{LF}$ does not imply that $U^n_{LF} < U^u_{LF}$. In particular, when $(w^u - q)^2 < (w^n)^2 < (w^u - q)\, w^u$, we have that $Y^n_{LF} < Y^u_{LF}$ and $U^n_{LF} > U^u_{LF}$, so that redistribution from non-users to users represents the "normal" direction of redistribution.[15] Thus, the fact that agents are heterogeneous in both skills and needs implies that one does not require unconventional redistributive tastes to rationalize negative marginal tax rates.[16] Nonetheless, notice that according to part *ii)* of Proposition 1, it is still true that, if incentive-compatibility considerations require to distort the bundle offered to the group that benefits from redistribution, the sign of the distortion is "coherent" with the income ranking prevailing under laissez-faire: when users earn more than non-users under laissez-faire and redistribution is in their favor, $T'(Y^u_{SB}) \leq 0$, and the income ranking under a second-best optimum is consistent with the income ranking under laissez-faire. As we will see below, however, income re-ranking can sometimes occur when, as in our setting, the SC condition is violated.

Part *iii)* considers the case when users earn less than non-users under laissez-faire. It shows that when the extent of redistribution from non-users to users is small, the second-best optimum will coincide with the first-best optimum and no distortion is needed to maintain incentive-compatibility. For intermediate degrees of redistribution only the labor supply of users will be distorted (downwards, by letting them face a positive marginal tax rate). Finally, if the redistributive goals pursued by the government are sufficiently strong ($\overline{V}^n < (1-\pi)\, U^n_{LF}$) and π sufficiently small ($\pi < 1 - \left(\frac{w^n}{w^u}\right)^2$), it might be the case that either $T'(Y^u_{SB}) < 0$ and $T'(Y^n_{SB}) = 0$ or that both types face an upward distortion on their labor supply. Thus, according to part *iii)* of Proposition 1, when $\pi < 1 - \left(\frac{w^n}{w^u}\right)^2$, the

[14]In particular, it should be either that $\pi \leq 1 - \left(\frac{w^n}{w^u}\right)^2$ or that π is sufficiently close to $1 - \left(\frac{w^n}{w^u}\right)^2$ so that the following condition is satisfied: $\sqrt{\frac{2}{\pi}}\left[(w^n)^2 - (1-\pi)(w^u)^2\right] < w^u(w^u - q) - (w^n)^2$.

[15]We are here referring to the terminology introduced by Stiglitz (1982), where the "normal" direction of redistribution is from those who are better off under laissez-faire towards those who are worse off.

[16]When $(w^u - q)^2 < (w^n)^2 < (w^u - q)\, w^u$ (so that $Y^u_{LF} > Y^n_{LF}$ but $U^u_{LF} < U^n_{LF}$), even a max-min planner would let users face a negative marginal tax rate. Moreover, as shown in Appendix A, when $(w^u - q)^2 < (w^n)^2 < (w^u - q)\, w^u$, a max-min planner may succeed, if the proportion of users is sufficiently small, in equalizing the utility for the two groups despite the fact that users, who benefit from the redistribution enacted by the government, face a negative marginal tax rate. This stands in contrast to standard models where a max-min planner can achieve utility-equalization only by discouraging the labor supply of the transfer-recipients to the point where they choose not to work.

support of the function $U^u\left(\overline{V}^n\right)$, describing the PF, may be a non-connected set: once $\overline{V}^n$ reaches $(1-\pi)\,U^n_{LF}$ a further increase in the utility of users may not be feasible or, if feasible, it necessarily requires a discrete downward jump in $\overline{V}^n$. If the support of the function $U^u\left(\overline{V}^n\right)$ is a non-connected set, the distortion imposed on the labor supply of users changes direction as the extent of redistribution from non-users to users increases: from a downward distortion $(T'\left(Y^u_{SB}\right)>0)$ one enters a region of the second-best frontier where the labor supply of users is distorted upwards $(T'\left(Y^u_{SB}\right)<0)$. Finally, when $T'\left(Y^u_{SB}\right)$ turns from positive to negative, the income ranking under a second-best optimum is no longer consistent with the income ranking under laissez-faire: whereas $Y^n_{LF}>Y^u_{LF}$, $Y^n_{SB}<Y^u_{SB}$. Both the possibility that the support of $U^u\left(\overline{V}^n\right)$ is a non-connected set and the possibility of income re-ranking (when comparing the laissez-faire equilibrium with the second-best optimum) follow from the circumstance that in our setting the SC condition is violated.[17]

Having discussed the properties of a second-best optimum when the socially desirable direction of redistribution is from non-users to users, in the next Proposition we provide a characterization of a second-best optimum for the opposite case.

Proposition 2 *Assume that the intended direction of redistribution is from users to non-users. Then,*

i) When $Y^n_{LF}=Y^u_{LF}$ the laissez-faire equilibrium will be second-best optimal;

ii) When $Y^n_{LF}<Y^u_{LF}$, then $T'\left(Y^n_{SB}\right)\geq 0$ and $T'\left(Y^u_{SB}\right)=0$.[18]

iii) When $Y^n_{LF}>Y^u_{LF}$, then $T'\left(Y^n_{SB}\right)\leq 0$ and $T'\left(Y^u_{SB}\right)\leq 0$.[19]

Proof. See Appendix B ■

Part *i)* of Proposition 2 shows that when the two types are pooled at the laissez-faire equilibrium, it is never possible to use a nonlinear income tax to redistribute from users

[17] Notice that in a model without income effects on labor supply, as the one that we have been considering, the income-ranking under a first-best optimum is always consistent with the income ranking under laissez-faire (since whenever their labor supply is left undistorted, agents will always work the same amount as under laissez-faire, no matter how large is the tax that they pay or the transfer that they receive). Thus, the fact that the income ranking under a second-best optimum may differ with respect to the one prevailing under laissez-faire also implies that the income ranking under a second-best optimum may differ with respect to the one prevailing under a first-best optimum.

[18] $T'\left(Y^n_{SB}\right)>0$ if $U^n_{LF}+\frac{\pi}{2}\frac{\left(Y^u_{LF}-Y^n_{LF}\right)^2}{(w^u)^2}<\overline{V}^n\leq U^n_{LF}+\frac{\pi}{2}\frac{\left(Y^u_{LF}-Y^n_{LF}\right)^2}{\left[(w^u)^2-\pi(w^n)^2\right]}$, where the right hand side of the inequality defines the upper bound of the utility that can be enjoyed by non-users.

[19] If $U^n_{LF}<\overline{V}^n\leq U^n_{LF}+\frac{\pi}{2}\frac{\left(Y^u_{LF}-Y^n_{LF}\right)^2}{(w^u)^2}$, then $T'\left(Y^n_{SB}\right)=0$ and $T'\left(Y^u_{SB}\right)=0$; if $\overline{V}^n>U^n_{LF}+\frac{\pi}{2}\frac{\left(Y^u_{LF}-Y^n_{LF}\right)^2}{(w^u)^2}$, then $T'\left(Y^n_{SB}\right)<0$ and $T'\left(Y^u_{SB}\right)\leq 0$.

to non-users. This is in contrast with what we have obtained in part *i)* of Proposition 1, where we have shown that redistribution from non-users to users was feasible provided that the proportion of users was below a given threshold. What explains this difference is the fact that, when $Y^n_{LF} = Y^u_{LF}$, the indifference curve on which non-users locate under laissez-faire lies everywhere above the indifference curve on which users locate under laissez-faire (except at the point $Y^n_{LF} = Y^u_{LF}$ where the two indifference curves are tangent). This makes it impossible to move users on a lower indifference curve without violating incentive-compatibility.

Part *i)* of Proposition 2, coupled with part *i)* of Proposition 1, allows concluding that when the laissez-faire equilibrium features pooling, the first-best- and the second-best PF share only one point, the laissez-faire equilibrium. When $\pi \geq q/w^u$, the second-best frontier consists of one single point, the laissez-faire equilibrium; when $\pi < q/w^u$, the only feasible direction of redistribution is from non-users to users and the labor supply of the latter will necessarily be distorted to implement a separating equilibrium.

Parts *ii)* and *iii)* provide instead results that mirror those that would be obtained in a setting where the SC condition holds. For this reason we will not discuss these results at length.

According to part *ii),* if redistribution goes from agents earning a relatively high income under laissez-faire towards agents earning a relatively low income under laissez-faire, incentive-compatibility considerations call for distorting downwards the labor supply of those benefiting from redistribution (provided that redistribution is sufficiently large).[20]

Similarly, according to part *iii),* if redistribution goes from agents earning a relatively low income under laissez-faire towards agents earning a relatively high income under laissez-faire, incentive-compatibility considerations call for distorting upwards the labor supply of those benefiting from redistribution (again, provided that redistribution is sufficiently large). Moreover, albeit part *iii)* of Proposition 2 indicates that negative marginal tax rates may be a feature of a second-best optimum, this can only occur by postulating unconventional redistributive tastes (given that $Y^n_{LF} > Y^u_{LF} \Longrightarrow U^n_{LF} > U^u_{LF}$).[21]

[20]Notice that $Y^n_{LF} < Y^u_{LF}$ is a necessary (but not sufficient) condition for a max-min planner to pursue redistribution from users to non-users. When $w^u - q > w^n$, so that $U^u_{LF} > U^n_{LF}$, a max-min planner will redistribute from users to non-users and incentive-compatibility considerations will require to impose a downward distortion on the labor supply of non-users. However, as shown in Appendix B, the magnitude of the downward distortion will never be so large to induce non-users to stop working (and in particular $Y^n_{SB} > \Omega$, where Ω is defined in (4)). Nevertheless, it is possible that a max-min planner succeeds in equalizing the utility of the two groups even though non-users, i.e. the transfer-recipients, work a positive amount of hours. This stands in contrast to standard models where a max-min planner can achieve utility-equalization only by discouraging the labor supply of the transfer-recipients to the point where they choose not to work.

[21]Notice that this is in contrast with what we obtained in Proposition 1 where one did not require

Finally, notice that, with redistribution going from users to non-users, income re-ranking will never occur.

4 Pareto-efficient taxation when the cost of the work-related good is nonlinear in labor supply

In the previous Section we have assumed that the cost of the work-related good for users was proportional to their labor supply and we have emphasized four main results: i) an anonymous nonlinear income tax may allow the government to convert a pooling laissez-faire equilibrium into a separating equilibrium; ii) negative marginal income tax rates can be rationalized even without resorting to unconventional redistributive tastes; iii) the support of the function $U^u\left(\overline{V}^n\right)$, describing the second-best PF may be a non-connected set; iv) a second-best optimum may not preserve the income ranking prevailing under laissez-faire. Similar qualitative results generalize, with some nuances, to the case when the cost of the work-related good is nonlinear in hours of work. Below we discuss this possibility by focusing on the two following functional forms for $\varphi(h)$:

$$\text{Case 1:} \quad \varphi(h) = q_1 h + q_3 \frac{h^3}{3}, \tag{5}$$

$$\text{Case 2:} \quad \varphi(h) = q_1 h + q_2 2h^{1/2}, \tag{6}$$

where we assume $q_1 \geq 0$, $q_2 > 0$, $q_3 > 0$, and $w^u > q_1$ (where the last assumption represents a necessary condition for $Y_{LF}^u > 0$). In Case 1, φ is convex; in Case 2 it is concave.

For Case 1 we have that, at any given bundle in the (Y, B)-space, the marginal rate of substitution for a user is given by:

$$MRS_{YB}^u = \frac{q_1}{w^u} + \left(\frac{Y}{w^u}\right)^2 \frac{q_3}{w^u} + \frac{1}{(w^u)^2} Y, \tag{7}$$

whereas for Case 2 we have:

$$MRS_{YB}^u = \frac{q_1}{w^u} + \left(\frac{Y}{w^u}\right)^{-1/2} \frac{q_2}{w^u} + \frac{1}{(w^u)^2} Y. \tag{8}$$

In both cases the marginal rate of substitution for a non-user is given $MRS_{YB}^n = Y/(w^n)^2$. Thus, when $w^n \geq w^u$ users will always have steeper indifference curves than non-users at all bundles in the (Y, B)-space. When $w^n < w^u$, instead, the SC property is no longer satisfied (similar to the case with proportional costs of work).

unconventional redistributive tastes to rationalize negative marginal tax rates.

In Case 1 users will have flatter indifference curves at bundles where

$$\left[\frac{1}{(w^n)^2} - \frac{1}{(w^u)^2}\right] Y - \frac{q_1}{w^u} > \left(\frac{Y}{w^u}\right)^2 \frac{q_3}{w^u}. \tag{9}$$

When $q_1 > 0$, (9) implies that users will have flatter indifference curves at bundles where

$$Y_{low} < Y < Y_{high},$$

and Y_{low} and Y_{high} are the values associated with the following expression:[22]

$$w^u \frac{(w^u)^2 - (w^n)^2 \pm \sqrt{\left[(w^u)^2 - (w^n)^2\right]^2 - 4q_1 q_3 (w^n)^4}}{2q_3 (w^n)^2}.$$

When instead $q_1 = 0$, users will have flatter indifference curves at bundles where

$$Y < \frac{(w^u)^2 - (w^n)^2}{(w^n)^2} \frac{w^u}{q_3}. \tag{10}$$

In Case 2, users will have flatter indifference curves at bundles where

$$\left[\frac{1}{(w^n)^2} - \frac{1}{(w^u)^2}\right] Y - \frac{q_1}{w^u} > \left(\frac{w^u}{Y}\right)^{1/2} \frac{q_2}{w^u}, \tag{11}$$

which requires Y to be sufficiently large.

Consider first Case 1 when $q_1 > 0$. Apart from the fact that, on the contrary to what happened in Section 3, two indifference curves, one pertaining to a user and one pertaining to a non-user, may cross more than twice, if one were to characterize the properties of a second-best optimum there would be three main differences with the qualitative results stated in Propositions 1 and 2.[23]

The first difference refers to part *i)* of Proposition 1. In particular, even though it is still true that, provided π is sufficiently small, it is feasible for the government to convert a pooling laissez-faire equilibrium into a separating equilibrium where redistribution favors users, it is no longer true that there exists a range of values for $\overline{V}^n$ such that the second-best optimum is not unique (in the sense that there exist two different allocations that can be offered to users and that maximize their utility). Intuitively, the reason is that, whereas in the model considered in Section 3 the convexity of the indifference curves was

[22]Hereafter we will assume that $\left[(w^u)^2 - (w^n)^2\right]^2 - 4q_1 q_3 (w^n)^4 > 0$.

[23]Notice that there are five possible configurations for a laissez-faire equilibrium: $Y^u_{LF} < Y^n_{LF} < Y_{low}$; $Y_{low} < Y^n_{LF} < Y^u_{LF} < Y_{high}$; $Y_{high} < Y^u_{LF} < Y^n_{LF}$; $Y^u_{LF} = Y^n_{LF} = Y_{low}$; $Y^u_{LF} = Y^n_{LF} = Y_{high}$.

constant in Y for both users and non-users,[24] this is no longer true for users when $\varphi(h)$ is nonlinear in labor supply.[25]

The second difference refers to part *i)* of Proposition 2. In particular, it may now be feasible to use a nonlinear income tax to redistribute from users to non-users even when the laissez-faire equilibrium is a pooling one. As we observed in Section 3, the reason why this was not possible with $\varphi(h) = qh$ was that, when $Y^n_{LF} = Y^u_{LF}$, the indifference curve on which non-users locate under laissez-faire lies everywhere above the indifference curve on which users locate under laissez-faire (except at the point $Y^n_{LF} = Y^u_{LF}$ where the two indifference curves are tangent). Under condition (5), however, this is no longer necessarily true. In fact, assume that $Y^n_{LF} = Y^u_{LF} = Y_{low}$; even though the indifference curves associated with utility levels U^u_{LF} and U^n_{LF}, which are tangent at Y_{low}, do not cross at $Y < Y_{low}$ or $Y_{low} < Y < Y_{high}$, they will cross at some value $Y > Y_{high}$.[26] Similarly, assume that $Y^n_{LF} = Y^u_{LF} = Y_{high}$; even though the indifference curves associated with utility levels U^u_{LF} and U^n_{LF}, which are tangent at Y_{high}, do not cross at $Y > Y_{high}$ or $Y_{low} < Y < Y_{high}$, they might cross at some value $Y < Y_{low}$.[27]

The last difference refers to the kind of income re-ranking that may arise at a second-best optimum. In particular, whereas Propositions 1 and 2 never contemplated the possibility that $Y^u_{LF} > Y^n_{LF}$ while $Y^u_{SB} < Y^n_{SB}$, this may occur in Case 1. The proof of this result is presented in Appendix C.

Consider now Case 1 when $q_1 = 0$. With some nuances, this case delivers results that are opposite to those obtained in Propositions 1 and 2. Intuitively, this is due to the fact that, whereas in Section 3 users had flatter indifference curves for values of Y exceeding a given threshold, exactly the opposite pattern holds for (5) when $q_1 = 0$.[28] This implies that it is never feasible to convert a pooling laissez-faire equilibrium into a

[24]In the (Y, B)-space, the indifference curves of users have equation $B = U^u + \frac{q}{w^u}Y + \frac{1}{2}\left(\frac{Y}{w^u}\right)^2$, so that $\frac{\partial^2 B}{\partial Y \partial Y} \mid_{U^u} = \frac{1}{(w^u)^2}$, and the indifference curves of non-users have equation $B = U^n + \frac{1}{2}\left(\frac{Y}{w^n}\right)^2$, which implies $\frac{\partial^2 B}{\partial Y \partial Y} \mid_{U^n} = \frac{1}{(w^n)^2}$.

[25]When $\varphi(h) = \frac{Y}{w^u}q_1 + \left(\frac{Y}{w^u}\right)^3 \frac{q_3}{3}$, we have that $\frac{\partial^2 B}{\partial Y \partial Y} \mid_{U^u} = \left(1 + 2q_3\frac{Y}{w^u}\right) / (w^u)^2$, implying that $\frac{\partial^2 B}{\partial Y \partial Y} \mid_{U^u}$ is increasing in Y.

[26]Notice that when $Y^n_{LF} = Y^u_{LF} = Y_{low}$ the indifference curve U^u_{LF} lies everywhere below (except at Y_{low}) the indifference curve U^n_{LF} for $Y \leq Y_{high}$. Thus, when it is feasible to convert a pooling laissez-faire equilibrium where $Y^n_{LF} = Y^u_{LF} = Y_{low}$ into a separating equilibrium where redistribution favors non-users, the labor supply of non-users will be upward distorted (they will face a negative marginal tax rate).

[27]Notice that, as remarked in footnote 23, a pooling equilibrium under laissez-faire can only occur at either Y_{low} or Y_{high}. Notice also that when $Y^n_{LF} = Y^u_{LF} = Y_{high}$ the indifference curve U^u_{LF} lies everywhere above (except at Y_{high}) the indifference curve U^n_{LF} for $Y \geq Y_{low}$.

[28]For $\varphi(h) = \left(\frac{Y}{w^u}\right)^3 \frac{q_3}{3}$, there are three possible configurations of a laissez-faire equilibrium: $Y^n_{LF} < Y^u_{LF} < \frac{(w^u)^2-(w^n)^2}{(w^n)^2}\frac{w^u}{q_3}$; $\frac{(w^u)^2-(w^n)^2}{(w^n)^2}\frac{w^u}{q_3} < Y^u_{LF} < Y^n_{LF}$; $Y^u_{LF} = Y^n_{LF} = \frac{(w^u)^2-(w^n)^2}{(w^n)^2}\frac{w^u}{q_3}$.

separating equilibrium where redistribution favors users, while it is feasible, provided π is sufficiently large, to break a pooling laissez-faire equilibrium and redistribute towards non-users.[29] Moreover, in contrast to what happened in Section 3, where a second-best optimum might have been non-unique when $Y_{LF}^n = Y_{LF}^u$, here the bundle offered to each of the two groups at a second-best optimum is uniquely determined. Once again, the reason is that the convexity of the indifference curves is not constant in Y for users when $\varphi(h)$ is nonlinear in labor supply. In particular, given that the aforementioned convexity is for users increasing in Y, when $Y_{LF}^n = Y_{LF}^u$ and π is sufficiently large so that it is feasible to redistribute towards non-users, it is more efficient to achieve type separation by distorting upwards the labor supply of non-users (letting them face a negative marginal tax rate). Finally, and again in contrast to what happened in Section 3, the only type of income re-ranking that may occur is $Y_{LF}^u > Y_{LF}^n \wedge Y_{SB}^u < Y_{SB}^n$.

Consider now Case 2 where $\varphi(h)$ is given by (6). In this case we have that $Y_{LF}^u > 0$ provided that q_2 is not too large. In any case, $Y_{LF}^u < (w^u - q_1) w^u$, where $(w^u - q_1) w^u$ represents the laissez-faire level of income earned by users when $q_2 = 0$ (so that $\varphi(h) = q_1 h$, which would bring us back to the case analyzed in the Section 3). Assuming that q_2 is such that $Y_{LF}^u > 0$, if one were to characterize the properties of a second-best optimum, there would be three main differences with the qualitative results stated in Propositions 1 and 2.

The first difference refers to part *i)* of Proposition 1. In particular, even though it is still true that, provided π is sufficiently small, it is feasible for the government to convert a pooling laissez-faire equilibrium into a separating equilibrium where redistribution favors users, it is no longer true that there exists a range of values for $\overline{V}^n$ such that the second-best optimum is not unique. Once again, this is due to the fact that the convexity of the indifference curves is not constant in Y for users when $\varphi(h)$ is nonlinear in labor supply. With $\varphi(h)$ given by (6), the convexity is increasing in Y.[30] Thus, when $Y_{LF}^n = Y_{LF}^u$ and π is sufficiently low so that it is feasible to redistribute towards users, it is more efficient to achieve type separation by distorting downwards the labor supply of users (letting them face a positive marginal tax rate).

The second difference refers to the kind of income re-ranking that may arise at a second-best optimum: whereas Propositions 1 and 2 never contemplated the possibility

[29] This is because, when $\varphi(h) = \left(\frac{Y}{w^u}\right)^3 \frac{q_3}{3}$ and $Y_{LF}^n = Y_{LF}^u$, the indifference curve on which users locate under laissez-faire lies everywhere above the indifference curve on which non-users locate under laissez-faire (except at the point $Y_{LF}^n = Y_{LF}^u$ where the two indifference curves are tangent).

[30] When $\varphi(h) = \frac{Y}{w^u} q_1 + \left(\frac{Y}{w^u}\right)^{1/2} 2q_2$, we have that $\frac{\partial^2 B}{\partial Y \partial Y} |_{U^u} = \frac{1}{(w^u)^2} - \frac{q_2}{2(w^u)^{1/2}} (Y^u)^{-3/2}$, implying that $\frac{\partial^2 B}{\partial Y \partial Y} |_{U^u}$ is increasing in Y.

that $Y^u_{LF} > Y^n_{LF}$ while $Y^u_{SB} < Y^n_{SB}$, this may occur under condition (6). The proof of this result is presented in Appendix D.

Finally, the last difference refers to the fact that, when redistribution goes from non-users to users, it is possible that a second-best optimum entails a distortion on the labor supply of users even when no self-selection constraint is (locally) binding in equilibrium. The reason for this last result is related to the fact that, with $\varphi(h)$ given by (6), it is no longer the case that MRS^u_{YB} is monotonically increasing in Y.[31] More precisely, since $MRS^u_{YB} = \frac{q_1}{w^u} + \left(\frac{Y}{w^u}\right)^{-1/2}\frac{q_2}{w^u} + \frac{1}{(w^u)^2}Y$, we have that $\partial MRS^u_{YB}/\partial Y = \frac{1}{(w^u)^2} - \frac{q_2}{2(w^u)^{1/2}}(Y^u)^{-3/2}$, which implies that $\partial MRS^u_{YB}/\partial Y > 0$ for $Y^u > (q_2/2)^{2/3} w^u$. Thus, while the value of MRS^u_{YB} is always positive for $Y \geq 0$, it starts at $+\infty$ when $Y = 0$, then it gradually decreases until it reaches a minimum, and only after that it monotonically increases. In particular, the fact that $MRS^u_{YB} > 1$ for very low values of Y implies that, when $\overline{V}^n$ is sufficiently low so that incentive-compatibility considerations require that Y^u must be very small,[32] it may be optimal for the government to offer users a bundle where $Y^u = 0$ even though it would be incentive-compatible to let users increase to some extent their labor supply (and enjoy a slightly larger value of consumption).[33]

This possibility is illustrated in Figure 2 below and a numerical example is provided in Appendix E.

In the figure above pre-tax income Y is represented on the horizontal axis and after-tax income $Y - T(Y) = B$ is represented on the vertical axis. The dashed 45 degree line represents the laissez-faire budget line (no taxes nor transfers); the point labelled I on

[31] In other words, the indifference curves for users are not everywhere convex.

[32] This happens when $\overline{V}^n$ is set larger than but very close to $(1-\pi)(w^n)^2/2$. As explained in Appendix A, with $U^n_{SB} < U^n_{LF}$ the government collects from each non-user a maximum amount of revenue equal to $Y^n - B^n = (1/2)(w^n)^2 - U^n_{SB}$. This implies that the revenue that can be transferred to each user is equal to $(1-\pi)\left[(1/2)(w^n)^2 - U^n_{SB}\right]/\pi$, which in turn implies that users can be offered a bundle on the virtual budget line $B = \frac{1-\pi}{\pi}\left[\frac{1}{2}(w^n)^2 - U^n_{SB}\right] + Y$. On this virtual budget line some bundles cannot be offered since they would induce mimicking by non-users. To identify the set of incentive-compatible bundles on the virtual budget line, one has to find the two values for Y at which the relevant indifference curve for non-users (i.e. the one associated with utility U^n_{SB}) intersects the virtual budget line. The only bundles on the virtual budget line that do not violate incentive-compatibility are those to the left of the first intersection point and to the right of the second intersection point. When $\overline{V}^n$ is close to $(1-\pi)(w^n)^2/2$ the first intersection point occurs at a value for Y which is close to zero.

[33] More precisely, it may be optimal for the government to offer users the bundle $(Y^u, B^u) = \left(0, \frac{1-\pi}{\pi}\left[\frac{1}{2}(w^n)^2 - \overline{V}^n\right]\right)$ even though it would be incentive-compatible to let users increase their labor supply to reach the bundle $(Y^u, B^u) = \left((w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}, \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}\overline{V}^n - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}\right)$, where the latter represents the bundle at which we have the first intersection between the relevant indifference curve for non-users (i.e. the one associated with utility $\overline{V}^n$) and the virtual budget line for users implied by the tax revenue collected from the group of non-users.

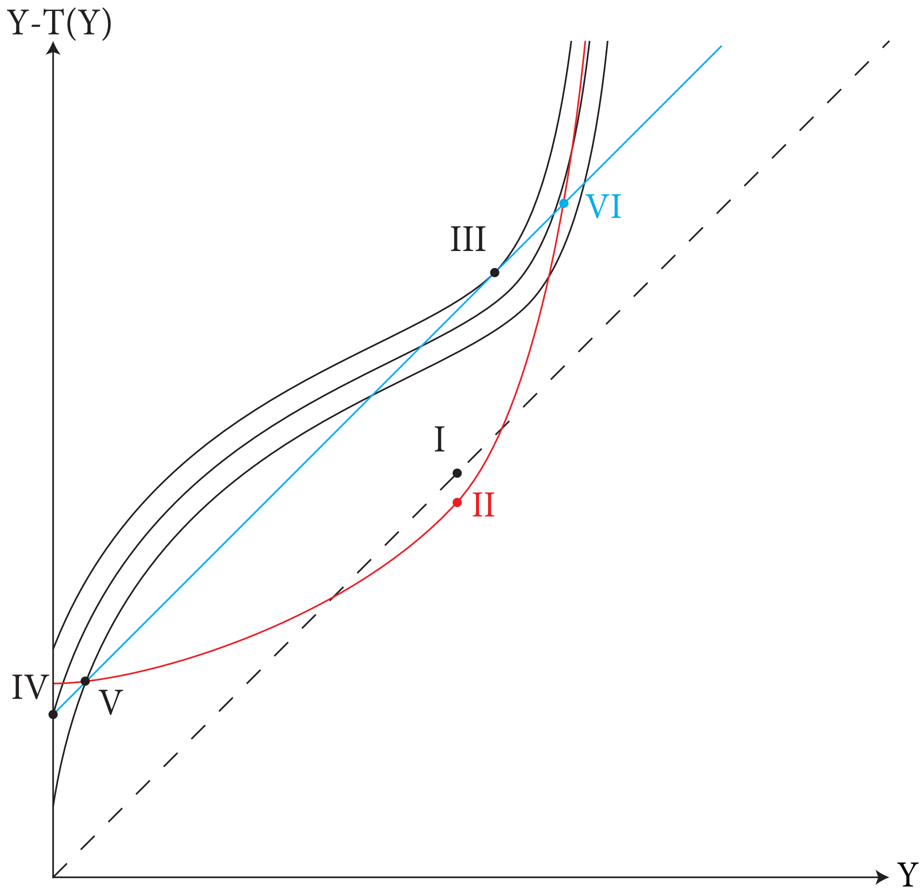

Figure 2

this line represents the bundle selected by non-users under laissez-faire. Bundle labelled II represents the undistorted bundle offered to non-users and lying on the indifference curve where $U^n = \overline{V}^n$. The blue 45 degree line represents the virtual budget line on which a bundle for users can be offered given the revenue extracted from non-users. Incentive compatibility requires that, on the blue virtual budget line, users can only be offered bundles to the left of bundle V and to the right of bundle VI, with both V and VI belonging to the set of admissible bundles. The three black curves passing through bundles V, IV and III are three different indifference curves pertaining to users.

From the figure one can see that bundle labelled IV is strictly preferred by users to both the bundle labelled V and the bundle labelled VI. But if users are offered the bundle IV, the self-selection constraint requiring non-users not to mimic users is slack. Notice also that users would be better off if they could get bundle III on the blue virtual budget line, i.e. the bundle at which their labor supply is undistorted. However, offering them this bundle would induce mimicking by non-users. Therefore, at a second-best optimum users are offered bundle IV and non-users are offered bundle II. The labor supply of users is downward distorted even though no self-selection constraint is binding at the second-best optimum. Nonetheless, the reason why users are offered a distorted bundle is ultimately due to the need to prevent mimicking from non-users and ensure proper self-selection by agents.

5 Subsidizing work-related expenses

In our analysis we have so far maintained the assumption that the only policy instrument for the government is a nonlinear income tax. In this setting we have highlighted the consequences descending from the violation of the SC condition. Most governments, however, allow special tax treatments for work-related expenses.[34] As we will show below, in general this does not imply that the SC condition is restored. To illustrate this point, assume that job-related expenses are subsidized at a flat rate s.[35] The first thing to notice

[34]Recent contributions that have analyzed the optimal tax treatment of work-related expenses include Koehne and Sachs (2017), Bastani et al. (2017) and Ho and Pavoni (2018), where the last two papers explicitly focus on the case of child care expenditures. A common feature of these papers is that they consider a setting where all agents are, according to our terminology, "users". Both Koehne and Sachs (2017) and Ho and Pavoni (2018) assume that agents only differ in market ability and allow for the possibility that an individual tax liability is a nonlinear and non-separable function of earned income and work-related expenses. However, whereas Ho and Pavoni (2018) focus on deriving properties of particular welfare optima, Koehne and Sachs (2017) focus on Pareto-improving reforms and provide an *incentive-adjusted no-arbitrage principle* that is a necessary requirement for Pareto-efficiency. They also show that the optimal work-related consumption wedge generally differs across agents, which implies that the required wedge cannot be created by standard forms of linear commodity taxation or by allowing all agents to deduct from their income tax base the same proportion of work-related expenses. Bastani et al. (2017) evaluates the desirability of uniform tax deductions and/or uniform tax credits for child care expenditures in a setting where agents differ in market ability and nurturing ability, and where the two dimensions of heterogeneity are perfectly and positively correlated. They show that, when child care expenditures are also a function of the quality of the service and this can be freely chosen by agents, subsidizing child care expenditures may be undesirable.

[35]We are implicitly assuming that job-related expenses are not observable by the government at the individual level so that a nonlinear subsidy scheme is not an option. This is an assumption that is often made in the literature (see, e.g., Anderberg and Balestrino, 2000, Cremer, Pestieau and Rochet, 2001, Micheletto, 2008, Blomquist, Christiansen and Micheletto, 2010, Jacobs and Boadway, 2014). Lack of public observability of personal purchases appears a realistic case to consider since individuals have often the possibility to misreport their true work-related expenses to the tax authority. For purchases of work-related goods, as opposed to work-related services, the possibility of reselling by agents exacerbates the problem of observing consumption at the individual level. If job-related expenses were costlessly observable by the government at the individual level, the government could instead devise a nonlinear tax schedule that is a joint function of earned income and job-related expenses. In such a case, if as in our model job-related expenses were not adjustable by the individuals, a first-best optimum could be implemented. However, a first-best optimum would not be implementable if the work-related good had a value for individuals that goes beyond its role as a necessary requirement for working. For instance, assume that X is a work-related good that users need to purchase in an amount that is given by the increasing function $\kappa(h)$, but it is also a good that positively contributes to the agents' utility for the amount exceeding $\kappa(h)$. Denoting by $\widehat{x}$ the total amount of X that is purchased by an agent and denoting by c the amount consumed of a separate composite consumption good, utility would be given by $U = c + v(\widehat{x} - \kappa(h)) - h^2/2$, where $v(\cdot)$ represents an increasing and concave function and where $\kappa(h) = 0$ for non-users. In such a setting, even if the government were able to observe total expenditures on good X at the individual level (i.e. $\widehat{x}$), a tax function that jointly depends on earned income and expenditures on X would not be enough to implement a first-best optimum. The reason is that non-users

is that, by setting $s = 1$, the SC condition would be restored.[36] The issue, however, is that $s = 1$ is not necessarily the optimal strategy for the government. To understand why this is the case, we will here refer to the set-up considered in Section 3 where $U = c - h^2/2$ and $\varphi(h) = qh$. Since a subsidy on job-related expenses works in favor of users, assume that the socially desirable direction of redistribution is from non-users to users ($\overline{V}^n = U^n_{SB} < (w^n)^2/2 = U^n_{LF}$). Moreover, assume that $(w^u - q)\, w^u \neq (w^n)^2$ and[37]

$$-\frac{(w^n)^2}{2} \leq \overline{V}^n < \frac{(w^n)^2}{2} - \frac{\pi\left[(w^u - q)\, w^u - (w^n)^2\right]^2}{2\,(w^n)^2}, \tag{12}$$

so that when $s = 0$ the government needs to offer users a distorted bundle to prevent mimicking by non-users,[38] whereas the latter are offered the undistorted bundle[39]

$$(Y, B) = \left((w^n)^2, \overline{V}^n + (1/2)\,(w^n)^2\right). \tag{13}$$

Notice that, if incentive-compatibility considerations were not an issue, the government would offer users the undistorted bundle (Y^u, B^u) where

$$(Y, B) = \left((w^u - q)\, w^u, (w^u - q)\, w^u + \frac{1-\pi}{\pi}\left[\frac{1}{2}(w^n)^2 - \overline{V}^n\right]\right), \tag{14}$$

allowing them to achieve a higher utility.

What we want to ascertain is whether, by properly choosing the subsidy rate s, the government may indeed offer users an undistorted bundle while preventing mimicking from non-users. For this purpose, assume that the government introduces a subsidy at rate $s > 0$ and that it offers users the bundle

$$(Y, B) = \left((w^u - q)\, w^u, (w^u - q)\, w^u + \frac{1-\pi}{\pi}\left[\frac{1}{2}(w^n)^2 - \overline{V}^n\right] - (w^u - q)\, qs\right), \tag{15}$$

would be able to adjust their purchases of good X to mimic the purchases of users.

[36] With $s = 1$ users would have flatter (steeper) indifference curves at any point in the (Y, B)-space whenever $w^u > (<)\, w^n$. From the perspective of agents, $s = 1$ is equivalent to granting them a refundable tax credit for all their work-related expenses (since offering agents a refundable tax credit for a fraction s of their work-related expenses is equivalent to subsidize work-related expenses at the rate s). Tax deductions represent an alternative possibility to offer agents a tax break. However, full deductibility, besides being not necessarily optimal, does not always imply that the single-crossing condition is restored. Denote taxable income by M, the nonlinear income tax by $T(M)$, and define $B \triangleq M - T(M)$. What one can show is that full deductibility of work-related expenses allows restoring the single-crossing condition in the (M, B)-space when $\varphi(h) = qh$. In this case users would have flatter (steeper) indifference curves at any point in the (M, B)-space whenever $w^u - q > (<)\, w^n$. If instead work-related expenses scale nonlinearly in labor supply, full deductibility of work-related expenses does not allow restoring the single-crossing condition.

[37] The assumption $(w^u - q)\, w^u \neq (w^n)^2$ implies that $Y^u_{LF} \neq Y^n_{LF}$.

[38] The fact that the government needs to distort the bundle offered to users when $\overline{V}^n < \frac{(w^n)^4 - \pi\left[(w^u - q)w^u - (w^n)^2\right]^2}{2(w^n)^2}$ is shown in Appendix A.

[39] As shown in Appendix A, non-users would be offered a distorted bundle if $\overline{V}^n < -(w^n)^2/2$.

while keeping unchanged the bundle intended for non-users.

Comparing the two bundles given by (14) and (15), we can see that, while Y is the same, the value of B in (15) has been lowered by an amount $(w^u - q)\, qs = (Y/w^u)\, qs = h^u qs$, which exactly offsets the saving that users enjoy due to the subsidy on job-related expenses. Thus, under a subsidy at rate s, the bundle (15) would represent an undistorted bundle that allows users to achieve the same utility as under the bundle (14) when $s = 0$. The difference is that, while offering (14) with $s = 0$ is not incentive-compatible, offering (15) with $s > 0$ prevents mimicking by non-users when the following condition holds:

$$\overline{V}^n \geq (w^u - q)\, w^u + \frac{1-\pi}{\pi} \left[\frac{1}{2} (w^n)^2 - \overline{V}^n\right] - \frac{1}{2} \left(\frac{(w^u - q)\, w^u}{w^n}\right)^2 - (w^u - q)\, qs. \quad (16)$$

Solving (16) to find the minimum value for s, denoted by s^*, that satisfies the inequality above, one gets:

$$s^* = \frac{(w^u - q)\, w^u + \frac{1}{\pi} \left[\frac{1}{2} (w^n)^2 - \overline{V}^n\right] - \frac{1}{2} \frac{(w^n)^4 + [(w^u - q) w^u]^2}{(w^n)^2}}{(w^u - q)\, q}. \quad (17)$$

Summarizing what we have obtained so far, when s is chosen according to (17) the government could offer users an undistorted bundle, without inducing mimicking by non-users, even when $\overline{V}^n < \frac{(w^n)^2}{2} - \frac{\pi\left[(w^u - q) w^u - (w^n)^2\right]^2}{2 (w^n)^2}$, a result that cannot be achieved if the government only relies on a nonlinear income tax.

The important thing to notice, however, is that this does not allow concluding that the second-best optimum will coincide with the first-best optimum. In fact, once s is chosen according to (17), the other self-selection constraint, i.e. the one requiring users not to be tempted to mimic non-users, may become binding.[40] The reason is that, since non-users are still offered the undistorted bundle (13), a subsidy on job-related expenses implies that the consumption available for a user behaving as a mimicker, i.e. choosing the bundle intended for non-users, increases by the amount $(w^n)^2 sq/w^u$, where $(w^n)^2 / w^u$ represents the labor supply of a user behaving as a mimicker. In particular, users will not have an incentive to mimic non-users if the following condition holds:

$$\begin{aligned} & (w^u - q)\, w^u + \frac{1-\pi}{\pi} \left[\frac{1}{2} (w^n)^2 - \overline{V}^n\right] - (w^u - q)\, q - \frac{1}{2} (w^u - q)^2 \\ \geq\ & \overline{V}^n + \frac{1}{2} (w^n)^2 - (1 - s)\, q \frac{(w^n)^2}{w^u} - \frac{1}{2} \left(\frac{(w^n)^2}{w^u}\right)^2, \end{aligned} \quad (18)$$

where the left hand side of the inequality above represents the utility achieved by users at the undistorted bundle offered to them by the government, and the right hand side

[40] This self-selection constraint was trivially non-binding at a second-best optimum when $\overline{V}^n < U^n_{LF}$ and the only policy instrument used by the government was a nonlinear income tax.

represents the utility that they would achieve if they were to choose the bundle (13) intended for non-users.[41]

As shown in Appendix F, substituting in (18) the value for s provided by (17), we can rewrite the no-mimicking condition as follows:

$$\frac{2w^u}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\left[w^u(w^u-q)-(w^n)^2\right]\geq(q-2w^u)\left[w^u(w^u-q)-(w^n)^2\right]^2. \quad (19)$$

It follows that, when $w^u(w^u-q)-(w^n)^2\geq 0$, users have no incentive to mimic non-users. When instead $w^u(w^u-q)-(w^n)^2<0$, users have no incentive to mimic non-users when the following condition holds:

$$\frac{2}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]w^u\leq-(2w^u-q)\left[w^u(w^u-q)-(w^n)^2\right],$$

namely when

$$\overline{V}^n\geq\frac{1}{2}\left\{(w^n)^2-\pi\frac{2w^u-q}{w^u}\left[(w^n)^2-w^u(w^u-q)\right]\right\}. \quad (20)$$

Therefore, when $w^u(w^u-q)-(w^n)^2<0$ and (20) is violated, an optimal nonlinear income tax coupled with an optimal subsidy on job-related expenses will not allow implementing the first-best allocation. The reason is that in this case the optimal value for s will be the result of a trade-off between the desirable effects in terms of deterring mimicking by non-users and the undesirable effects of making more tempting for users to mimic non-users. This implies that at the resulting second-best optimum both self-selection constraint will be binding. For $\overline{V}^n$ lower than but sufficiently close to $U^n_{LF}-\frac{\pi}{2}\frac{2w^u-q}{w^u}(Y^n_{LF}-Y^u_{LF})$, the second-best optimum will be a separating equilibrium where the labor supply of both types is downward distorted ($Y^u_{SB}<Y^u_{LF}$, $Y^n_{SB}<Y^n_{LF}$ and $Y^u_{SB}<Y^n_{SB}$). As one keeps lowering $\overline{V}^n$, one eventually reaches a lower bound below which the government is no longer able to make users better off by implementing a separating equilibrium.[42] Once this lower bound for $\overline{V}^n$ is reached, a further increase in U^u

[41] Notice that the right hand side of (18) can be rewritten as $B^n-(1-s)\,q\frac{Y^n}{w^u}-\frac{1}{2}\left(\frac{Y^n}{w^u}\right)^2$, where the term $(1-s)\,q\frac{Y^n}{w^u}$ represents the effective outlay for job-related costs when users mimic non-users and job-related expenses are subsidized at rate s.

[42] The intuition is the following. Suppose to start from a separating equilibrium where both self-selection constraints are binding and both h^u and h^n are downward distorted. If s and h^n are kept fixed, a reduction in $\overline{V}^n$ requires to exacerbate the downward distortion on h^u that is needed to prevent mimicking from non-users. But the required increase in the downward distortion on h^u might be so large that U^u fails to increase. Alternatively, suppose that s is increased to mitigate the required increase in the downward distortion on h^u that is needed to prevent mimicking from non-users. The minimum increase in s needed to make users better off might induce them to behave as mimickers. To prevent this from happening, the government should further downward distort the labor supply of non-users along their indifference curve corresponding to the new, lower value for $\overline{V}^n$. But given that h^n was already downward distorted at the initial equilibrium, it follows that, by reducing $\overline{V}^n$, the government may not succeed in collecting extra-revenue from non-users.

can be achieved but it may require a discrete downward jump in $\overline{V}^n$, and it necessarily requires a switch from a separating- to a pooling regime where both types earn the same pre-tax income Y^{pool}, where $Y^u_{LF} < Y^{pool} < Y^n_{LF}$.

Notice also that, since the right hand side of (20) defines a value for $\overline{V}^n$ that is lower than the one defined by the right hand side of (12) when $(w^n)^2 > w^u(w^u - q)$,[43] it follows that, provided that the desired degree of redistribution from non-users to users is not too large, an optimal subsidy allows implementing the first-best allocation even when $(w^n)^2 > w^u(w^u - q)$. In particular, the range of values for $\overline{V}^n$ for which this occurs is given by:

$$\frac{1}{2}\left\{(w^n)^2 - \pi\frac{2w^u - q}{w^u}\left[(w^n)^2 - w^u(w^u - q)\right]\right\}$$
$$\leq \overline{V}^n < \frac{1}{2}\left\{(w^n)^2 - \frac{\pi\left[(w^u - q)w^u - (w^n)^2\right]^2}{(w^n)^2}\right\}. \quad (21)$$

So far, our analysis has relied on the assumption that $(w^u - q)w^u \neq (w^n)^2$ so that $Y^u_{LF} \neq Y^n_{LF}$. If instead $Y^u_{LF} = Y^n_{LF}$, it is easy to see that supplementing a nonlinear income tax with a subsidy on job-related expenses allows implementing a first-best optimum. In fact, assume that $-(w^n)^2/2 \leq \overline{V}^n < (w^n)^2/2 = U^n_{LF}$.[44] By offering to all agents, users and non-users, the bundle $(Y, B) = \left((w^n)^2, \frac{(w^n)^2}{2} + \overline{V}^n\right)$ and setting $s = \frac{\frac{(w^n)^2}{2} - \overline{V}^n}{(w^u - q)q\pi}$, one achieves redistribution ($U^n_{SB} = \overline{V}^n < U^n_{LF}$; $U^u_{SB} = U^u_{LF} + \frac{1-\pi}{\pi}\left(U^n_{LF} - \overline{V}^n\right) > U^u_{LF}$), while at the same time leaving undistorted the labor supply of all agents ($Y^n_{LF} = Y^n_{SB} = Y^u_{LF} = Y^u_{SB}$), maintaining incentive-compatibility (given that all agents are offered the same bundle in the (Y, B)-space), and satisfying the public budget constraint (since the cost of the subsidy benefiting users, i.e. $(w^u - q)sq\pi$, is exactly matched by the total revenue collected through the income tax, i.e. $\frac{(w^n)^2}{2} - \overline{V}^n$).

The following Proposition summarizes our results.

Proposition 3 *Assume that* $-U^n_{LF} \leq \overline{V}^n < U^n_{LF} - \frac{\pi}{2}\frac{(Y^u_{LF} - Y^n_{LF})^2}{Y^n_{LF}}$.

i) If $Y^u_{LF} \geq Y^n_{LF}$*, a nonlinear income tax coupled with a properly chosen subsidy on job-related expenses allows implementing a first-best optimum; the same will be true when* $Y^u_{LF} < Y^n_{LF}$ *and* $\overline{V}^n \geq U^n_{LF} - \frac{\pi}{2}\frac{2w^u - q}{w^u}\left(Y^n_{LF} - Y^u_{LF}\right)$.

[43] Assume that $(w^n)^2 - w^u(w^u - q) > 0$. The condition $(w^n)^2 - \pi\frac{2w^u - q}{w^u}\left[(w^n)^2 - w^u(w^u - q)\right] < (w^n)^2 - \pi\frac{\left[(w^n)^2 - (w^u - q)w^u\right]^2}{(w^n)^2}$ can be restated as $\left[(w^n)^2 - (w^u - q)w^u\right]^2 w^u - (2w^u - q)\left[(w^n)^2 - w^u(w^u - q)\right](w^n)^2 < 0$ and therefore $\left[(w^n)^2 - (w^u - q)w^u\right]w^u < (2w^u - q)(w^n)^2$. Simplifying terms one gets $-(w^u - q)(w^u)^2 < (w^u - q)(w^n)^2$.

[44] The reason to assume $\overline{V}^n \geq -(w^n)^2/2$ is that for $\overline{V}^n < -(w^n)^2/2$ it is not possible to leave the labor supply of non-users undistorted without pushing B^n below zero.

ii) If instead $Y^u_{LF} < Y^n_{LF}$ and $\overline{V}^n < U^n_{LF} - \frac{\pi}{2}\frac{2w^u - q}{w^u}\left(Y^n_{LF} - Y^u_{LF}\right)$, an optimal nonlinear income tax coupled with an optimal subsidy on job-related expenses will implement a second-best optimum where both self-selection constraints are binding and both types of agents face a distortion on their labor supply. If the second-best optimum is a separating equilibrium ($Y^u_{SB} < Y^n_{SB}$), both types of agents will face a downward distortion on their labor supply. If the second-best optimum is a pooling equilibrium ($Y^u_{SB} = Y^n_{SB}$), the labor supply of users will be upward distorted and the labor supply of non-users will be downward distorted.

As we have remarked above, one interesting finding is the possibility of a second-best optimum which is a pooling equilibrium. This possibility is illustrated by the following numerical example. Assume that $w^u = 12$, $w^n = 10$, $q = 5$ and $\pi = 1/2$. Under laissez-faire we have that $Y^u_{LF} = (w^u - q)\, w^u = 84$ and $Y^n_{LF} = (w^n)^2 = 100$, with $U^u_{LF} = (w^u - q)^2/2 = 24.5$ and $U^n_{LF} = (w^n)^2/2 = 50$. Assume that the social welfare function maximized by the government is of a maximin type. When only a nonlinear income tax is used, the second-best optimum is a separating equilibrium where the only binding self-selection constraint is the one requiring non-users not to be tempted to mimic users, and where $Y^u_{SB} = 42.86$, $Y^n_{SB} = 100$, $U^u_{SB} = 26.79$, $U^n_{SB} = 41.84$, $T'\left(Y^u_{SB}\right) = 28.57\%$, $T'\left(Y^n_{SB}\right) = 0$, and the average income tax rate, defined as $(Y - B)/Y$, is equal to 8.16% for non-users and to −8.16% for users.

When a nonlinear income tax is used jointly with an optimal subsidy on job-related expenses, the second-best optimum is a pooling equilibrium where both self-selection constraints are binding, $s^* = 0.65$, $Y^u_{SB} = Y^n_{SB} = 95.45$, $U^u_{SB} = U^n_{SB} = 36.97$, and the average income tax rate, defined as $(Y - B)/Y$, is equal to 13.54% for both groups (but where users get a benefit from the subsidy on job-related expenses so that their effective average tax rate, defined by $[(Y - B)/Y] - s^* q / w^u$, is equal to −13.54%). At the second-best optimum the labor supply of users is distorted upwards (it was distorted downwards under a second-best optimum when only a nonlinear income tax was used) and the labor supply of non-users is distorted downwards (it was left undistorted under a second-best optimum when only a nonlinear income tax was used).

6 Concluding remarks

In this paper, we have considered a two-type optimal nonlinear income tax model where agents differ both in terms of market ability and in terms of "needs" for a work-related good/service, i.e. a good/service that some agents need to purchase in order to work.

Because of this bi-dimensional heterogeneity, the single-crossing conditions fails to hold. Ruling out public observability of individual types, we have characterized the properties of a second-best optimum by looking at the entire second-best Pareto frontier.

We have highlighted that, due to the violation of single-crossing, some non-standard results arise. First of all, a second-best optimum might not preserve the earned-income ranking that prevails under laissez-faire (with the corollary that the distortion imposed on the labor supply of transfer-recipients is not necessarily "coherent" with the income ranking prevailing under laissez-faire). Second, redistribution via income taxation might be feasible even when the laissez-faire equilibrium is a pooling equilibrium. Third, a second-best optimum might not be unique, in the sense that there might be more than one set of allocations in the (pre-tax income, after-tax income)-space that solve the government's maximization problem. Fourth, the support of the function describing the second-best Pareto frontier may be a non-connected set. Fifth, supplementing an optimal nonlinear income tax with an optimal subsidy on work-related expenses may imply that at a second-best optimum redistribution is achieved through a separating- or pooling equilibrium where both self-selection constraints are binding. Sixth, we have shown that at a second-best optimum the labor supply of some agents might be distorted even though no self-selection constraint is (locally) binding in equilibrium.

Before concluding, a final remark is in order. For tractability reasons, we have focused our analysis to a simplified two-type model where skills and needs are perfectly correlated. However, insofar as our non-standard results hinge on the violation of the single-crossing condition, they generalize, with some nuances, to settings with a larger number of types and imperfect correlation between skills and needs.

Appendix A

Proof of Proposition 1: Assuming that the intended direction of redistribution is from non-users to users, it follows that the government aims at offering to users a (Y, B)-bundle such that $Y^u - B^u < 0$ and to non-users a (Y, B)-bundle such that $Y^n - B^n > 0$. If the government succeeds in achieving its redistributive goals, non-users will then obtain a utility that is lower than the utility they would obtain under laissez-faire. Denoting their utility at a second-best optimum by U^n_{SB}, we have $U^n_{SB} < U^n_{LF} = (w^n)^2/2$.

With income tax revenue collected from each non-user being equal to $Y^n - B^n$, the revenue that can be transferred to each user is equal to $(1-\pi)(Y^n - B^n)/\pi$. With non-users being offered a bundle on their indifference curve with associated value $\overline{V}^n$, the maximum revenue that the government can collect from each of them is obtained at the bundle where their labor supply is undistorted, implying a zero implicit marginal income tax rate for non-users, at least as long as the undistorted bundle on the indifference curve U_{SB}^n does not violate the constraint $B^n \geq 0$. Assume for the moment that this is indeed the case.[45] Then, independently on the value of $\overline{V}^n$, we will have that $Y^n = (w^n)^2$.

With $\overline{V}^n < U_{LF}^n$ and $Y^n = (w^n)^2$, the government collects from each non-user an amount of revenue equal to $Y^n - B^n = (w^n)^2 - \left[\overline{V}^n + (1/2)(w^n)^2\right] = (1/2)(w^n)^2 - \overline{V}^n$. This implies that the revenue that can be transferred to each user is equal to $(1-\pi)\left[(1/2)(w^n)^2 - \overline{V}^n\right]/\pi$, which in turn implies that users will be offered a bundle on the virtual budget line

$$B = \frac{1-\pi}{\pi}\left[\frac{1}{2}(w^n)^2 - \overline{V}^n\right] + Y \tag{A1}$$

On this virtual budget line, however, some bundles cannot be offered since they would induce mimicking by non-users. To find the set of incentive-compatible bundles on the virtual budget line (A1), one has to identify the values for Y at which the relevant indifference curve for non-users (i.e. the one associated with utility $\overline{V}^n$) intersects the virtual budget line.

Taking into account that the relevant indifference curve for non-users has equation

$$B = \overline{V}^n + \frac{1}{2}\left(\frac{Y}{w^n}\right)^2, \tag{A2}$$

by equating (A1) and (A2) one can find two values for Y. These are given by:

$$\begin{aligned} Y &= (w^n)^2\left\{1 \pm \sqrt{1 - \frac{2}{(w^n)^2}\frac{1}{\pi}\left[\overline{V}^n - \frac{1}{2}(1-\pi)(w^n)^2\right]}\right\} \\ &= (w^n)^2\left\{1 \pm \sqrt{\frac{1}{\pi} - \frac{2}{(w^n)^2}\frac{1}{\pi}\overline{V}^n}\right\} \\ &= (w^n)^2\left\{1 \pm \sqrt{\frac{1}{\pi}\left(1 - \frac{2}{(w^n)^2}\overline{V}^n\right)}\right\} \\ &= (w^n)^2 \pm w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}, \end{aligned} \tag{A3}$$

where the term within square root is positive due to the initial assumption that $\overline{V}^n < U_{LF}^n = (w^n)^2/2$.

On the virtual budget line (A1) only the bundles with $Y \leq (w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}$ and $Y \geq (w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}$ are incentive-compatible (do not induce the non-

[45]This is equivalent to assume that $\overline{V}^n \geq -(w^n)^2/2$, since this would be the utility for a non-user at the undistorted bundle $(Y, B) = ((w^n)^2, 0)$. We will later relax the assumption that $\overline{V}^n \geq -(w^n)^2/2$.

users to behave as mimickers).[46] If incentive-compatibility considerations were not an issue, users could be offered on the virtual budget line (A1) the undistorted bundle

$$(Y, B) = \left((w^u - q) w^u, (w^u - q) w^u + \frac{1-\pi}{\pi} \left[\frac{1}{2} (w^n)^2 - \overline{V}^n \right] \right).$$

Thus, if it is either the case that

$$(w^u - q) w^u \geq (w^n)^2 + w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]}, \tag{A4}$$

or that

$$(w^u - q) w^u \leq (w^n)^2 - w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]}, \tag{A5}$$

the second-best optimum would entail no distortion on the labor supply of users ($T'(Y_{SB}^u) = 0$). Solving (A4) and (A5) for $\overline{V}^n$, one finds that $T'(Y_{SB}^u) = 0$ when

$$\overline{V}^n \geq \frac{(w^n)^4 - \pi \left[(w^u - q) w^u - (w^n)^2 \right]^2}{2 (w^n)^2}, \tag{A6}$$

where the right hand side of (A6) is strictly lower than $(w^n)^2 / 2 = U_{LF}^n$ as long as $(w^u - q) w^u \neq (w^n)^2$. (The case when $(w^u - q) w^u = (w^n)^2$ will be considered later.)

Suppose instead that the socially optimal degree of redistribution from non-users to users is sufficiently large, so that inequality (A6) does not hold. Offering users an undistorted bundle along the virtual budget line (A1) would then violate the incentive-compatibility constraint for non-users. This implies that users will either be offered the bundle (Y_A, B_A) where

$$\begin{aligned} Y_A &= (w^n)^2 - w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]}, \qquad (A7) \\ B_A &= \frac{1-\pi}{\pi} \left[\frac{1}{2} (w^n)^2 - \overline{V}^n \right] + (w^n)^2 - w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]} \\ &= \frac{1+\pi}{2\pi} (w^n)^2 - \frac{1-\pi}{\pi} \overline{V}^n - w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]}, \qquad (A8) \end{aligned}$$

and the labor supply of users is distorted downwards ($T'(Y_{SB}^u) > 0$), or the bundle (Y_B, B_B) where

$$\begin{aligned} Y_B &= (w^n)^2 + w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]}, \qquad (A9) \\ B_B &= \frac{1-\pi}{\pi} \left[\frac{1}{2} (w^n)^2 - \overline{V}^n \right] + (w^n)^2 + w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]} \\ &= \frac{1+\pi}{2\pi} (w^n)^2 - \frac{1-\pi}{\pi} \overline{V}^n + w^n \sqrt{\frac{1}{\pi} \left[(w^n)^2 - 2\overline{V}^n \right]}, \qquad (A10) \end{aligned}$$

[46] Notice that, for sufficiently low values of $\overline{V}^n$ (in particular, $\overline{V}^n < (1-\pi) (w^n)^2 / 2$), the lower root of (A3) might be negative; in that case the set of incentive-compatible bundles on the virtual budget line (A1) is given by those bundles where Y is greater or equal to the larger root.

and the labor supply of users is distorted upwards ($T'(Y^u_{SB}) < 0$).

For later purposes, notice that from (A7), since Y_A cannot take negative values, U^n can never fall below $\frac{(w^n)^2(1-\pi)}{2}$ when users are offered the bundle (Y_A, B_A). Notice also that (A9)-(A10) represents a valid characterization of an incentive-compatible bundle offered to users as long as $\overline{V}^n \geq -(w^n)^2/2$. The reason is that in deriving (A9)-(A10) we have assumed that the labor supply of non-users was left undistorted; this implies that, since their consumption must be non-negative, $\overline{V}^n \geq -(w^n)^2/2$. This does not mean that it is not possible to push the utility of non-users below $-(w^n)^2/2$ in an incentive-compatible way; it only means that to do that it is necessary to distort upwards the labor supply of non-users, which in turn would imply a different characterization than (A9)-(A10) for the incentive-compatible bundle offered to users.

Evaluating the utility of users at the bundle characterized by (A7)-(A8), we have:

$$\begin{aligned} U^u(Y_A, B_A) &= \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}\overline{V}^n - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]} \\ &\quad -\frac{q}{w^u}\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}\right\} \\ &\quad -\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}\right\}^2, \end{aligned} \tag{A11}$$

whereas the utility of users at the bundle characterized by (A9)-(A10) is

$$\begin{aligned} U^u(Y_B, B_B) &= \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}\overline{V}^n + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]} \\ &\quad -\frac{q}{w^u}\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}\right\} \\ &\quad -\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}\right\}^2. \end{aligned} \tag{A12}$$

Before comparing the utility of users at (Y_A, B_A) and (Y_B, B_B), notice that a necessary condition for (Y_A, B_A) to be part of the second-best PF is that $\frac{\partial U^u(Y_A,B_A)}{\partial \overline{V}^n} < 0$ (and similarly, a necessary condition for (Y_B, B_B) to be part of the second-best PF is that $\frac{\partial U^u(Y_B,B_B)}{\partial \overline{V}^n} < 0$).

Consider first $\partial U^u(Y_A, B_A)/\partial \overline{V}^n$. This is given by:

$$\begin{aligned} \frac{\partial U^u(Y_A, B_A)}{\partial \overline{V}^n} &= \left[-(1-\pi) + \left(\frac{w^n}{w^u}\right)^2\right]\frac{1}{\pi} \\ &\quad + \left[\frac{(w^u - q)w^u - (w^n)^2}{(w^u)^2}\right]\left\{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]\right\}^{-1/2}\frac{w^n}{\pi}. \end{aligned} \tag{A13}$$

To evaluate when (A13) takes a negative sign, two cases need to be distinguished: i) $(w^u - q)\, w^u - (w^n)^2 < 0$; ii) $(w^u - q)\, w^u - (w^n)^2 > 0$. Under case i) we have that $\frac{\partial U^u(Y_A,B_A)}{\partial \overline{V}^n} < 0$ when

$$\frac{(w^n)^2 - (1-\pi)(w^u)^2}{(w^n)^2 - (w^u - q)\, w^u} \frac{1}{w^n} < \frac{1}{\left\{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]\right\}^{1/2}}, \tag{A14}$$

whereas under case ii) we have that $\frac{\partial U^u(Y_A,B_A)}{\partial \overline{V}^n} < 0$ when

$$\frac{(w^n)^2 - (1-\pi)(w^u)^2}{(w^n)^2 - (w^u - q)\, w^u} \frac{1}{w^n} > \frac{1}{\left\{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]\right\}^{1/2}}. \tag{A15}$$

For case i), when $\pi \leq \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$ (implying $(w^n)^2 - (1-\pi)(w^u)^2 \leq 0$), (A14) is always satisfied, implying that one can keep raising the utility of users until $\overline{V}^n$ is pushed down to the value $\frac{(w^n)^2(1-\pi)}{2}$ (implying that Y_A, as defined by (A7), reaches its lower bound $Y_A = 0$, and $U^u\left(Y^A, B^A\right) = U^n_{SB} = \frac{(w^n)^2(1-\pi)}{2}$). When instead $\pi > \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, (A14) is satisfied as long as

$$\overline{V}^n > \frac{(w^n)^2}{2}\left\{1 - \pi\left[\frac{(w^n)^2 - (w^u - q)\, w^u}{(w^n)^2 - (1-\pi)(w^u)^2}\right]^2\right\}, \tag{A16}$$

where the right hand side of (A16) is larger than $\frac{(w^n)^2(1-\pi)}{2}$ when $\pi > q/w^u$.

Noticing that $(w^u - q)\, w^u - (w^n)^2 < 0 \Longrightarrow \frac{(w^u)^2-(w^n)^2}{(w^u)^2} < \frac{q}{w^u}$, we can conclude that, with $(w^u - q)\, w^u - (w^n)^2 < 0$, by offering users the bundle (Y_A, B_A) one can keep raising their utility up to the point where $\overline{V}^n$ is either lowered to the value $\frac{(w^n)^2(1-\pi)}{2}$, when $\pi < \frac{q}{w^u}$, or to the value (larger than $\frac{(w^n)^2(1-\pi)}{2}$) defined by (A16), when $\pi \geq \frac{q}{w^u}$.

For case ii), instead, we have that when $\pi \geq \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, (A15) is never satisfied, ruling out the possibility that the bundle (Y_A, B_A) is second-best optimal. When instead $\pi < \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, (A15) is satisfied as long as

$$\overline{V}^n < \frac{(w^n)^2}{2}\left\{1 - \left[\frac{(w^u - q)\, w^u - (w^n)^2}{(1-\pi)(w^u)^2 - (w^n)^2}\right]^2 \pi\right\}, \tag{A17}$$

where the right hand side of (A17) is larger than $\frac{(w^n)^2(1-\pi)}{2}$ for $\pi < q/w^u$.

Noticing that $(w^u - q)\, w^u - (w^n)^2 > 0 \Longrightarrow \frac{(w^u)^2-(w^n)^2}{(w^u)^2} > \frac{q}{w^u}$, we can conclude that, with $(w^u - q)\, w^u - (w^n)^2 > 0$, as long as $\pi \geq \frac{q}{w^u}$, offering users the bundle (Y_A, B_A) cannot be part of a second-best optimum.

Consider now $\partial U^u(Y_B, B_B)/\partial \overline{V}^n$. This is given by:

$$\frac{\partial U^u(Y_B, B_B)}{\partial \overline{V}^n} = \left[-(1-\pi) + \left(\frac{w^n}{w^u}\right)^2\right]\frac{1}{\pi} + \frac{(w^n)^2 - w^u(w^u-q)}{(w^u)^2}\left\{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]\right\}^{-1/2}\frac{w^n}{\pi}. \tag{A18}$$

Once again, to evaluate when (A18) takes a negative sign, two cases need to be distinguished: i) $(w^u - q)w^u - (w^n)^2 < 0$; ii) $(w^u - q)w^u - (w^n)^2 > 0$. Under case i) we have that $\frac{\partial U^u(Y_B,B_B)}{\partial \overline{V}^n} < 0$ when

$$\frac{1}{\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}} < \frac{(1-\pi)(w^u)^2 - (w^n)^2}{w^n\left[(w^n)^2 - w^u(w^u-q)\right]}, \tag{A19}$$

whereas under case ii) we have that $\frac{\partial U^u(Y_B,B_B)}{\partial \overline{V}^n} < 0$ when

$$\frac{1}{\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2\overline{V}^n\right]}} > \frac{(1-\pi)(w^u)^2 - (w^n)^2}{w^n\left[(w^n)^2 - w^u(w^u-q)\right]}. \tag{A20}$$

For case i), when $\pi \geq \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$ (implying $(w^n)^2 - (1-\pi)(w^u)^2 > 0$), (A19) is never satisfied, ruling out the possibility that the bundle (Y_B, B_B) is second-best optimal. When instead $\pi < \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, (A19) is satisfied as long as

$$\overline{V}^n < \frac{(w^n)^2}{2}\left\{1 - \pi\frac{\left[(w^n)^2 - w^u(w^u-q)\right]^2}{\left[(1-\pi)(w^u)^2 - (w^n)^2\right]^2}\right\}, \tag{A21}$$

where the right hand side of (A21) is smaller than $\frac{(w^n)^2(1-\pi)}{2}$ for $\left(\frac{(w^u)^2-(w^n)^2}{(w^u)^2} >\right)\pi > \frac{(w^u)^2-(w^n)^2}{(w^u)^2} + \frac{w^u(w^u-q)-(w^n)^2}{(w^u)^2}$ and it is larger or equal than $\frac{(w^n)^2(1-\pi)}{2}$ for $\pi \leq \frac{(w^u)^2-(w^n)^2}{(w^u)^2} + \frac{w^u(w^u-q)-(w^n)^2}{(w^u)^2}$. Notice in particular that, when $\pi < \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$ but sufficiently close to $\frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, the right hand side of (A21) would define a value that is smaller than $-(w^n)^2/2$, in which case (Y_B, B_B), as defined by (A9)-(A10), does not represent a valid characterization of an incentive-compatible bundle offered to users.

For case ii), instead, we have that when $\pi \leq \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, (A20) is always satisfied. When instead $\pi > \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, (A20) is satisfied as long as

$$\overline{V}^n > \frac{(w^n)^2}{2}\left\{1 - \pi\left[\frac{w^u(w^u-q) - (w^n)^2}{(w^n)^2 - (1-\pi)(w^u)^2}\right]^2\right\}, \tag{A22}$$

where we notice in particular that the right hand side of (A22) is smaller than $-\frac{(w^n)^2}{2}$ when π is sufficiently close to $\frac{(w^u)^2-(w^n)^2}{(w^u)^2}$.[47]

Let's now compare $U^u(Y_A, B_A)$ and $U^u(Y_B, B_B)$ as given by (A11)-(A12). Simple algebra can be used to show that

$$U^u(Y_A, B_A) > (<) \, U^u(Y_B, B_B) \Longleftrightarrow (w^u - q) \, w^u < (>) \, (w^n)^2, \tag{A23}$$

or, equivalently, since $(w^u - q) \, w^u = Y^u_{LF}$ and $(w^n)^2 = Y^n_{LF}$,

$$U^u(Y_A, B_A) > (<) \, U^u(Y_B, B_B) \Longleftrightarrow Y^u_{LF} < (>) \, Y^n_{LF}.$$

The result stated in (A23), coupled with the results that we have obtained above analyzing $\frac{\partial U^u(Y_A,B_A)}{\partial \overline{V}^n}$ and $\frac{\partial U^u(Y_B,B_B)}{\partial \overline{V}^n}$, show that, when $(w^u - q) \, w^u > (w^n)^2$ and (A6) is violated, a second-best optimum will necessarily entail an upward distortion on the labor supply of users ($T'(Y^u_{SB}) < 0$); regarding the labor supply of non-users, this will either be left undistorted or, if π is sufficiently small and $\overline{V}^n$ sufficiently small, also non-users will face an upward distortion on their labor supply ($T'(Y^n_{SB}) \leq 0$). If the second-best optimum is such that both $T'(Y^u_{SB})$ and $T'(Y^n_{SB})$ are negative, the average tax rate on non-users is 100% and their consumption is pushed to its lower bound, i.e. zero.[48]

When instead $(w^u - q) \, w^u < (w^n)^2$ and (A6) is violated, a second-best optimum will entail a downward distortion on the labor supply of users ($T'(Y^u_{SB}) > 0$) and no-distortion on the labor supply of non-users ($T'(Y^n_{SB}) = 0$) as long as $\pi \geq \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$.[49] However, if $\pi < \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$ and $\overline{V}^n$ is sufficiently small, the second-best optimum may be characterized by an upward distortion on the labor supply of users ($T'(Y^u_{SB}) < 0$) and either a no-distortion or an upward distortion on the labor supply of non-users ($T'(Y^n_{SB}) \leq 0$).

In particular, this implies that when $(w^u - q) \, w^u < (w^n)^2$ and $\pi < \frac{(w^u)^2-(w^n)^2}{(w^u)^2}$, the second-best PF may feature a discontinuity at $U^n_{SB} = U^u_{SB} = \frac{(w^n)^2(1-\pi)}{2}$ (and this would

[47] In particular, when $\sqrt{\frac{2}{\pi}} \left[(w^n)^2 - (1-\pi) \, (w^u)^2 \right] < w^u \, (w^u - q) - (w^n)^2$.

[48] An additional point that is worth emphasizing is that, when $(w^u - q)^2 < (w^n)^2 < (w^u - q) \, w^u$ (so that $Y^u_{LF} > Y^n_{LF}$ but $U^u_{LF} < U^n_{LF}$), a max-min planner may succeed, if the proportion of users is sufficiently small, in equalizing the utility for the two groups despite the fact that users, who benefit from the redistribution enacted by the government, face a negative marginal tax rate.

[49] Notice that, when $(w^u - q) \, w^u < (w^n)^2$, a max-min planner would like to redistribute from non-users to users. In this case, if $\pi < q/w^u$, a max-min planner would equalize the utility of both groups by distorting the labor supply of users to the point where $Y^u_{SB} = 0$. If instead $\pi \geq q/w^u$ a max-min planner would not equalize the utility for the two groups, and we would have that $U^n_{SB} > U^u_{SB}$. This is because, when $\pi \geq q/w^u$, the right hand side of (A16) defines a value that is weakly larger than $\frac{(w^n)^2(1-\pi)}{2}$. But for $\overline{V}^n > \frac{(w^n)^2(1-\pi)}{2}$ and $\pi \geq q/w^u$ the value of $U^u(Y_A, B_A)$ defined by (A11) is lower than $\frac{(w^n)^2(1-\pi)}{2}$. Thus, when $(w^u - q) \, w^u < (w^n)^2$ a max-min planner will never equalize the utility for the two groups unless $Y^u_{SB} = 0$ (which will be optimal only when $\pi < q/w^u$).

certainly happen if $(w^u - q)\, w^u$ is sufficiently close to $(w^n)^2$). Starting from this point it may be possible to further increase U^u_{SB} but this would require a discrete downward jump in U^n_{SB}. If it is feasible to raise U^u_{SB} above $U^u_{SB} = \frac{(w^n)^2(1-\pi)}{2}$, the corresponding second-best optimum would switch from an equilibrium where $T'(Y^u_{SB}) > 0$ and $T'(Y^n_{SB}) = 0$ to an equilibrium where $T'(Y^u_{SB}) < 0$ and $T'(Y^n_{SB}) \leq 0$.

Finally, let's consider the case when $(w^u - q)\, w^u = (w^n)^2$ so that $Y^u_{LF} = Y^n_{LF}$. In this case the right hand side of inequality (A6) simplifies to $(w^n)^2/2$, which is the utility achieved by non-users under laissez-faire. This shows that, when $(w^n)^2 = (w^u)(w^u - q)$, it is never possible to redistribute from non-users to users without distorting the labor supply of the latter. In order not to violate the incentive-compatibility constraint for non-users, users can either be offered the distorted bundle characterized by (A7)-(A8) or the distorted bundle characterized by (A9)-(A10). But when $(w^n)^2 = (w^u)(w^u - q)$, users are indifferent between the two bundles. Thus, as long as users prefer these bundles to their laissez-faire bundle, there will be two equivalent second-best optima, one entailing a downward distortion on the labor supply of users ($T'(Y^u_{SB}) > 0$), and one entailing an upward distortion on their labor supply ($T'(Y^u_{SB}) < 0$). If, however, users are better off under laissez-faire, no redistribution is feasible and the second-best optimum coincides with the laissez-faire (pooling) equilibrium.[50] Formally, this happens when $U^u_{LF} = \frac{(w^u-q)^2}{2} > U^u(Y_B, B_B) = U^u(Y_A, B_A)$, namely (taking into account that $(w^n)^2 = (w^u)(w^u - q)$):

$$\begin{aligned}\frac{(w^u-q)^2}{2} > \; & \frac{1+\pi}{2\pi}(w^u)(w^u-q) - \frac{1-\pi}{\pi}\overline{V}^n \\ & + \sqrt{(w^u)(w^u-q)}\sqrt{\frac{1}{\pi}\left[(w^u)(w^u-q) - 2\overline{V}^n\right]} \\ & - \frac{q}{w^u}\left\{(w^u)(w^u-q) + \sqrt{(w^u)(w^u-q)}\sqrt{\frac{1}{\pi}\left[(w^u)(w^u-q) - 2\overline{V}^n\right]}\right\} \\ & - \frac{\left\{(w^u)(w^u-q) + \sqrt{(w^u)(w^u-q)}\sqrt{\frac{1}{\pi}\left[(w^u)(w^u-q) - 2\overline{V}^n\right]}\right\}^2}{2(w^u)^2},\end{aligned}$$

which, after simplifying and collecting terms, can be restated as

$$w^u - \frac{q}{\pi} > \left(w^u - \frac{q}{\pi}\right)\frac{2}{(w^u-q)\, w^u}\overline{V}^n. \tag{A24}$$

When $\pi > q/w^u$, so that $w^u - q/\pi > 0$, (A24) holds when $(w^u - q)\, w^u/2 > \overline{V}^n$, namely

[50] In this case both the λ-constraint and the ϕ-constraint in the government's problem are binding at the second-best optimum.

whenever $\overline{V}^n$ falls below its laissez-faire value.[51] Thus, when $(w^n)^2 = (w^u)(w^u - q)$ and $\pi > q/w^u$, users' utility cannot be raised beyond its laissez-faire value (no redistribution from non-users to users is feasible and the second-best optimum coincides with laissez-faire).[52] When instead $\pi < q/w^u$, so that $w^u - q/\pi < 0$, (A24) holds when $(w^u - q)\, w^u/2 < \overline{V}^n$. Thus, when $\pi < q/w^u$ redistribution from non-users to users is feasible and users will face a non-zero marginal tax rate at a second-best optimum.

Notice also that when $(w^u - q)\, w^u = (w^n)^2$ and $\pi < q/w^u$, the second-best PF does not feature a discontinuity at $U^n_{SB} = U^u_{SB} = \frac{(w^n)^2(1-\pi)}{2}$. Starting from this point it would be possible to further increase U^u_{SB} without a discrete downward jump in U^n_{SB}. However, to increase U^u_{SB} beyond $\frac{(w^n)^2(1-\pi)}{2}$, the second-best optimum would necessarily entail $T'(Y^u_{SB}) < 0$ and $T'(Y^n_{SB}) \leq 0$.

Appendix B

Proof of Proposition 2: Assume now that the intended direction of redistribution is from users to non-users. This implies that the optimal bundles offered by the government will entail $Y^n - B^n < 0$ and $Y^u - B^u > 0$. Users will therefore obtain a utility that is lower than the utility they would obtain under laissez-faire. Denoting their laissez-faire utility by U^u_{LF} and their utility at a second-best optimum by U^u_{SB}, we have $U^u_{LF} = \frac{(w^u - q)^2}{2} > U^u_{SB}$.

With income tax revenue collected from each user being equal to $Y^u - B^u$, the revenue that can be transferred to each non-user is equal to $\frac{\pi}{1-\pi}[Y^u - B^u]$. With users being offered a bundle on their indifference curve with associated value U^u_{SB}, the maximum revenue that the government can collect from each of them is obtained at the bundle where their labor supply is undistorted (implying a zero implicit marginal income tax rate for users). In our setting with no income effects on labor supply this implies that, independently on the value of U^u_{SB}, we will have that $Y^u = (w^u - q)\, w^u$ (at least as long as $B^u > (w^u - q)\, q$, which implies $c^u > 0$). Thus, when the utility obtained by users at a second-best optimum is $U^u_{SB} < U^u_{LF}$ and their labor supply is left undistorted, the government collects from each user an amount of revenue equal to $Y^u - B^u = (w^u - q)\, w^u - \left[U^u_{SB} + \frac{1}{2}(w^u - q)^2 + (w^u - q)\, q\right] = \frac{1}{2}(w^u - q)^2 - U^u_{SB}$. This implies that the revenue that can be transferred to each user is equal to $\frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U^u_{SB}\right]$, which in turn implies that non-users will be offered a bundle on the virtual budget line:

$$B = \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U^u_{SB}\right] + Y \tag{B1}$$

[51] Notice that $U^n_{LF} = (w^u - q)\, w^u/2$ when $(w^n)^2 = (w^u)(w^u - q)$. Notice also that when $(w^n)^2 = (w^u)(w^u - q)$, we have that $q/w^u = \left[(w^u)^2 - (w^n)^2\right] / (w^u)^2$.

[52] Users' utility cannot be raised beyond its laissez-faire value even when $\pi = q/w^u$.

On this virtual budget line, however, some bundles cannot be offered since they would induce mimicking by users. To find the set of incentive-compatible bundles on the virtual budget line (B1), one has to identify the two values for Y at which the relevant indifference curve for users (i.e. the one associated with utility U^u_{SB}) intersects the virtual budget line.

Taking into account that the relevant indifference curve for users has equation

$$B = U^u_{SB} + \frac{1}{2}\left(\frac{Y}{w^u}\right)^2 + q\frac{Y}{w^u}, \tag{B2}$$

by equating (B1) and (B2) we can find the two relevant values for Y. These are given by

$$\begin{aligned} Y &= (w^u)^2 \left\{1 - \frac{q}{w^u} \pm \sqrt{1 + \frac{q^2}{(w^u)^2} - 2\frac{q}{w^u} - \frac{2}{(w^u)^2}\frac{1}{1-\pi}\left[U^u_{SB} - \pi\frac{1}{2}(w^u - q)^2\right]}\right\} \\ &= (w^u)^2 \left\{1 - \frac{q}{w^u} \pm \sqrt{\frac{1}{1-\pi}\left[\frac{q^2}{(w^u)^2} - 2\frac{q}{w^u} + 1 - \frac{2}{(w^u)^2}U^u_{SB}\right]}\right\} \\ &= (w^u)^2\left(1 - \frac{q}{w^u}\right) \pm w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U^u_{SB}\right]} \\ &= (w^u)(w^u - q) \pm w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U^u_{SB}\right]}, \end{aligned}$$

where the term within square root is positive due to our assumption that $U^u_{SB} < U^u_{LF} = \frac{(w^u - q)^2}{2}$.

On the virtual budget line (B1), the incentive-compatible bundles (which do not induce users to behave as mimickers) are those satisfying either of the following two conditions:

$$Y \leq (w^u)(w^u - q) - w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U^u_{SB}\right]},$$
$$Y \geq (w^u)(w^u - q) + w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U^u_{SB}\right]}.$$

If incentive-compatibility considerations were not an issue, non-users could be offered on the virtual budget line (B1) the undistorted bundle

$$(Y, B) = \left((w^n)^2, (w^n)^2 + \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U^u_{SB}\right]\right).$$

Thus, if it is either the case that

$$(w^n)^2 \geq (w^u)(w^u - q) + w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U^u_{SB}\right]}, \tag{B3}$$

or that

$$(w^n)^2 \leq (w^u)(w^u - q) - w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U^u_{SB}\right]}, \tag{B4}$$

the second-best optimum would entail no distortion on the labor supply of non-users $(T'(Y_{SB}^n) = 0)$. Solving (B3) and (B4) for U_{SB}^u, one finds that $T'(Y_{SB}^n) = 0$ when

$$U_{SB}^u \geq \frac{(w^u - q)^2}{2} - \frac{(1-\pi)\left[(w^u - q)w^u - (w^n)^2\right]^2}{2(w^u)^2}, \tag{B5}$$

where the right hand side of (B5) is strictly lower than $(w^u - q)^2/2 = U_{LF}^u$ as long as $(w^u - q)w^u \neq (w^n)^2$. (The case when $(w^u - q)w^u = (w^n)^2$ will be considered later.)

Taking into account that when non-users are offered an undistorted bundle, their utility is

$$\begin{aligned} U^n &= (w^n)^2 + \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U_{SB}^u\right] - \frac{(w^n)^2}{2} \\ &= \frac{(w^n)^2}{2} + \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U_{SB}^u\right], \end{aligned} \tag{B6}$$

and substituting for U_{SB}^u in (B6) the value provided by the right hand side of (B5), one gets the maximum utility that can be enjoyed by non-users without resorting to distort their labor supply:

$$U^n = \frac{(w^n)^2}{2} + \pi\frac{\left[(w^n)^2 - (w^u)(w^u - q)\right]^2}{2(w^u)^2}.$$

Suppose now that the socially optimal degree of redistribution from users to non-users is sufficiently large, so that inequality (B5) does not hold. Offering non-users an undistorted bundle along the virtual budget line (B1) would then violate the incentive-compatibility constraint for users. This implies that non-users will either be offered the bundle (Y_C, B_C) where

$$Y_C = (w^u)(w^u - q) - w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U_{SB}^u\right]} \tag{B7}$$

$$B_C = \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U_{SB}^u\right] + w^u\left\{w^u - q - \sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U_{SB}^u\right]}\right\} \tag{B8}$$

and the labor supply of non-users is distorted downwards $(T'(Y_{SB}^n) > 0)$, or the bundle (Y_D, B_D) where

$$Y_D = (w^u)(w^u - q) + w^u\sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U_{SB}^u\right]} \tag{B9}$$

$$B_D = \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u - q)^2 - U_{SB}^u\right] + w^u\left\{w^u - q + \sqrt{\frac{1}{1-\pi}\left[(w^u - q)^2 - 2U_{SB}^u\right]}\right\} \tag{B10}$$

and the labor supply of non-users is distorted upwards $(T'(Y_{SB}^n) < 0)$.

For later purposes, notice that from (B7), since Y_C cannot take negative values, U^u can never fall below $\frac{(w^u-q)^2\pi}{2}$ when non-users are offered the bundle (Y_C, B_C). Notice also that (B9)-(B10) represents a valid characterization of an incentive-compatible bundle offered to non-users as long as $U^u \geq -(w^u-q)^2/2$. The reason is that in deriving (B9)-(B10) we have assumed that the labor supply of users was left undistorted; this implies that, since their consumption must be non-negative, $U^u \geq -(w^u-q)^2/2$. This does not mean that it is not possible to push the utility of users below $-(w^u-q)^2/2$ in an incentive-compatible way; it only means that to do that it is necessary to distort upwards their labor supply, which in turn would imply a different characterization than (B9)-(B10) for the incentive-compatible bundle offered to non-users.[53]

Evaluating the utility of non-users at the bundle characterized by (B7)-(B8), we have:

$$\begin{aligned} U^n(Y_C, B_C) &= \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u-q)^2 - U^u_{SB}\right] \\ &\quad + (w^u)(w^u-q) - w^u\sqrt{\frac{1}{1-\pi}\left[(w^u-q)^2 - 2U^u_{SB}\right]} \\ &\quad - \frac{1}{2}\frac{1}{(w^n)^2}\left\{(w^u)(w^u-q) - w^u\sqrt{\frac{1}{1-\pi}\left[(w^u-q)^2 - 2U^u_{SB}\right]}\right\}^2, \end{aligned} \tag{B11}$$

whereas the utility of non-users at the bundle characterized by (B9)-(B10) is

$$\begin{aligned} U^n(Y_D, B_D) &= \frac{\pi}{1-\pi}\left[\frac{1}{2}(w^u-q)^2 - U^u_{SB}\right] \\ &\quad + (w^u)(w^u-q) + w^u\sqrt{\frac{1}{1-\pi}\left[(w^u-q)^2 - 2U^u_{SB}\right]} \\ &\quad - \frac{1}{2}\frac{1}{(w^n)^2}\left\{(w^u)(w^u-q) + w^u\sqrt{\frac{1}{1-\pi}\left[(w^u-q)^2 - 2U^u_{SB}\right]}\right\}^2. \end{aligned} \tag{B12}$$

Before comparing the utility of non-users at (Y_C, B_C) and (Y_D, B_D), notice that a necessary condition for (Y_C, B_C) to be part of the second-best PF is that $\frac{\partial U^n(Y_C,B_C)}{\partial U^u_{SB}} < 0$ (and similarly, a necessary condition for (Y_D, B_D) to be part of the second-best PF is that $\frac{\partial U^n(Y_D,B_D)}{\partial U^u_{SB}} < 0$).

[53] Notice instead that a similar caveat does not apply to (B7)-(B8). The reason is that the consumption of users is strictly non-negative (and in particular equal to $(w^u-q)^2(1+\pi)/2$) at the bundle where their labor supply is left undistorted and they are on the indifference curve associated with the utility level $(w^u-q)^2\pi/2$.

Consider first $\partial U^n(Y_C, B_C)/\partial U^u_{SB}$. This is given by:

$$\frac{\partial U^n(Y_C, B_C)}{\partial U^u_{SB}} = \left\{ -\pi + \frac{(w^u)^2}{(w^n)^2} + \frac{w^u - \frac{(w^u)^2(w^u-q)}{(w^n)^2}}{\left[\frac{(w^u-q)^2-2U^u_{SB}}{1-\pi}\right]^{1/2}} \right\} \frac{1}{1-\pi}. \tag{B13}$$

Thus, we have that $\frac{\partial U^n(Y_C,B_C)}{\partial U^u_{SB}} < 0$ when

$$(w^u)^2 - \pi (w^n)^2 + \frac{(w^n)^2 - w^u(w^u - q)}{\left[\frac{(w^u-q)^2-2U^u_{SB}}{1-\pi}\right]^{1/2}} w^u < 0. \tag{B14}$$

Condition (B14) is never satisfied for $(w^n)^2 - w^u(w^u - q) \geq 0$. For $(w^n)^2 - w^u(w^u - q) < 0$, instead, (B14) holds for

$$U^u_{SB} > \frac{(w^u - q)^2}{2} - \frac{(1-\pi)(w^u)^2 \left[w^u(w^u - q) - (w^n)^2\right]^2}{2\left[(w^u)^2 - \pi(w^n)^2\right]^2}, \tag{B15}$$

where the right hand side of (B15) defines a lower bound for U^u_{SB} along the second-best PF.

Substituting for U^u_{SB} into (B11) the value provided by the right hand side of (B15) allows deriving an upper bound for U^n_{SB}, and therefore $\overline{V}^n$, when $(w^n)^2 < w^u(w^u - q)$. After tedious calculations one gets:[54]

$$U^n_{SB} = \frac{(w^n)^2}{2} + \frac{\pi\left[(w^u - q)w^u - (w^n)^2\right]^2}{2\left[(w^u)^2 - \pi(w^n)^2\right]}. \tag{B16}$$

It is easy to verify that the right hand side of (B15) is larger than $\frac{1}{2}(w^u - q)^2 - \frac{1-\pi}{2}\left[\frac{(w^u-q)w^u-(w^n)^2}{(w^u)^2-(w^n)^2} w^u\right]^2$, which represents the value of U^u_{SB} that implies $Y_C = \Omega$ (where Y_C is defined by (B7) and $\Omega \equiv q\frac{(w^n)^2 w^u}{(w^u)^2-(w^n)^2}$ represents the threshold value for Y separating the bundles where $MRS^u_{YB} > MRS^n_{YB}$, i.e. those bundles where $Y < \Omega$, from the bundles where $MRS^u_{YB} < MRS^n_{YB}$, i.e. those bundles where $Y > \Omega$). This shows that it can never be optimal to discourage the labor supply of non-users to the point where $Y^n_{SB} = 0$.[55]

Consider now $\partial U^n(Y_D, B_D)/\partial U^u_{SB}$. This is given by:

$$\frac{\partial U^n(Y_D, B_D)}{\partial U^u_{SB}} = \left\{ -\pi + \frac{(w^u)^2}{(w^n)^2} - \frac{w^u - \frac{(w^u)^2(w^u-q)}{(w^n)^2}}{\left[\frac{(w^u-q)^2-2U^u_{SB}}{1-\pi}\right]^{1/2}} \right\} \frac{1}{1-\pi}. \tag{B17}$$

[54] Details of the calculations are available upon request.

[55] It is however worth noticing that this does not imply that, by redistributing from users to non-users, a max-min planner will never equalize the utility for the two groups. In fact, if $w^u - q > w^n$, so that $U^u_{LF} > U^n_{LF}$, a max-min planner will redistribute from users to non-users. In this case, if w^n is sufficiently close to $w^u - q$, it is straightforward to show that U^n_{SB}, as defined by (B16), is larger than the right hand side of (B15), which implies that, by distorting downwards the labor supply of non-users, it is possible to reverse the utility ranking prevailing under laissez-faire. In turn, this implies that, when w^n is sufficiently close to $w^u - q$, a max-min planner will equalize the utility of the two groups.

Thus, we have that $\frac{\partial U^n(Y_D,B_D)}{\partial U^u_{SB}} < 0$ when

$$\left(w^u\right)^2 - \pi \left(w^n\right)^2 - \frac{\left(w^n\right)^2 - w^u\left(w^u - q\right)}{\left[\frac{\left(w^u-q\right)^2 - 2U^u_{SB}}{1-\pi}\right]^{1/2}} w^u < 0. \tag{B18}$$

Condition (B18) is never satisfied for $\left(w^n\right)^2 - w^u\left(w^u - q\right) \leq 0$. For $\left(w^n\right)^2 - w^u\left(w^u - q\right) > 0$, instead, (B18) holds for values of U^u_{SB} that satisfy (B15).

Notice also that, for $\left(w^n\right)^2 - w^u\left(w^u - q\right) > 0$, the right hand side of (B15) is larger or equal than $-\left(w^u - q\right)^2/2$ provided that the following condition holds:

$$\frac{1 - \pi\frac{\left(w^n\right)^2}{\left(w^u\right)^2}}{\sqrt{1-\pi}} \geq \frac{\left(w^n\right)^2 - \left(w^u - q\right) w^u}{\sqrt{2}\left(w^u - q\right) w^u}. \tag{B19}$$

This is important to bear in mind since (B9)-(B10) represents a valid characterization of an incentive-compatible bundle offered to non-users only insofar as $U^u_{SB} \geq -\left(w^u - q\right)^2/2$, which is the value of U^u_{SB} that implies $B^u = \left(w^u - q\right) q$, and therefore $c^u = 0$, when the labor supply of users is left undistorted. Thus, when (B19) is satisfied, one can never enter a region of the second-best PF where the labor supply of both users and non-users is distorted upwards (with redistribution from users to non-users).

Let's now compare $U^n\left(Y_C, B_C\right)$ and $U^n\left(Y_D, B_D\right)$ as given by (B11)-(B12). Simple algebra can be used to show that

$$U^n\left(Y_C, B_C\right) > (<)\, U^n\left(Y_D, B_D\right) \Longleftrightarrow \left(w^u - q\right) w^u > (<) \left(w^n\right)^2, \tag{B20}$$

or, equivalently, since $\left(w^u - q\right) w^u = Y^u_{LF}$ and $\left(w^n\right)^2 = Y^n_{LF}$,

$$U^n\left(Y_C, B_C\right) > (<)\, U^n\left(Y_D, B_D\right) \Longleftrightarrow Y^n_{LF} < (>)\, Y^u_{LF}.$$

The result stated in (B20), coupled with the results that we have obtained above analyzing $\partial U^n\left(Y_C, B_C\right)/\partial U^u_{SB}$ and $\partial U^n\left(Y_D, B_D\right)/\partial U^u_{SB}$, show that, when $\left(w^u - q\right) w^u < \left(w^n\right)^2$ and (B5) is violated, a second-best optimum will necessarily entail an upward distortion on the labor supply of non-users ($T'\left(Y^n_{SB}\right) < 0$); regarding the labor supply of users, this will either be left undistorted or, if (B19) is violated and $\overline{V}^n$ sufficiently large, also non-users will face an upward distortion on their labor supply ($T'\left(Y^u_{SB}\right) \leq 0$). If the second-best optimum is such that both $T'\left(Y^u_{SB}\right)$ and $T'\left(Y^n_{SB}\right)$ are negative, the average tax rate on users is 100% and their consumption is pushed to its lower bound, i.e. zero.

When instead $\left(w^u - q\right) w^u > \left(w^n\right)^2$ and (B5) is violated, a second-best optimum will entail a downward distortion on the labor supply of non-users ($T'\left(Y^n_{SB}\right) > 0$) and no-distortion on the labor supply of users ($T'\left(Y^u_{SB}\right) = 0$).

Finally, let's consider the case when $(w^u - q)\,w^u = (w^n)^2$ so that $Y^u_{LF} = Y^n_{LF}$. In this case the right hand side of inequality (B5) simplifies to $(w^u - q)^2/2$, which is the utility achieved by users under laissez-faire. This shows that, when $(w^n)^2 = (w^u)(w^u - q)$, it is never possible to redistribute from users to non-users without distorting the labor supply of the latter. In order not to violate the incentive-compatibility constraint for users, non-users can either be offered the distorted bundle characterized by (B7)-(B8) or the distorted bundle characterized by (B9)-(B10). With $(w^n)^2 = (w^u)(w^u - q)$, non-users are indifferent between the two bundles. However, from (B13) and (B17) we also have that, when $(w^n)^2 = (w^u)(w^u - q)$, $\frac{\partial U^n(Y_C,B_C)}{\partial U^u_{SB}} = \frac{\partial U^n(Y_D,B_D)}{\partial U^u_{SB}} = \left[-\pi + \frac{(w^u)^2}{(w^n)^2}\right]\frac{1}{1-\pi} > 0$, which implies that the laissez-faire equilibrium is second-best optimal when the intended direction of redistribution is from users to non-users.

Appendix C

Proof of the result that, with $\varphi(h) = \frac{Y}{w^u}q_1 + \left(\frac{Y}{w^u}\right)^3\frac{q_3}{3}$, it might be that $Y^u_{LF} > Y^n_{LF}$ while $Y^u_{SB} < Y^n_{SB}$: Let's first consider the labor supply choice of a user under laissez-faire:

$$\max_h \quad w^u h - q_1 h - \frac{q_3}{3}(h)^3 - \frac{1}{2}h^2.$$

The associated first order condition is given by:

$$w^u - q_1 - q_3 (h)^2 - h = 0,$$

which implies

$$h^u_{LF} = \frac{-1 + \sqrt{1 + 4(w^u - q_1)\,q_3}}{2q_3},$$

and therefore:[56]

$$Y^u_{LF} = \frac{-1 + \sqrt{1 + 4(w^u - q_1)\,q_3}}{2q_3}w^u.$$

Given that the income earned by non-users under laissez-faire is given by $(w^n)^2$, it follows that:

$$Y^u_{LF} > Y^n_{LF} \iff \frac{-1 + \sqrt{1 + 4(w^u - q_1)\,q_3}}{2q_3}w^u > (w^n)^2, \tag{C1}$$

or, equivalently:

$$(w^u - q_1)\,w^u - (w^n)^2 > (w^n)^4\,\frac{q_3}{w^u}.$$

Assume that the socially desirable direction of redistribution is from non-users to users, so that $\overline{V}^n = U^n_{SB} < U^n_{LF} = (w^n)^2/2$. Following a procedure similar to that used in the proof of Proposition 1 (see Appendix A), we have that, when incentive-compatibility

[56] The equation $w^u - q_1 - q_3(h)^2 - h = 0$ has one solution for $h > 0$ and one solution for $h < 0$. The latter can be safely neglected.

considerations require to offer users a distorted bundle, users will either be offered the bundle (A7)-(A8), entailing a downward distortion on their labor supply, or the bundle (A9)-(A10), entailing an upward distortion on their labor supply. Evaluating the utility of users at each of these two bundles we have:

$$\begin{aligned} U^u(Y_A, B_A) &= \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}U_{SB}^n - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} \\ &-\frac{q_1}{w^u}\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\} \\ &-\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\}^2 \\ &-\frac{q_3}{3}\left(\frac{1}{w^u}\right)^3\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\}^3, \end{aligned} \tag{C2}$$

and

$$\begin{aligned} U^u(Y_B, B_B) &= \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}U_{SB}^n + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} \\ &-\frac{q_1}{w^u}\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\} \\ &-\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\}^2 \\ &-\frac{q_3}{3}\left(\frac{1}{w^u}\right)^3\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\}^3. \end{aligned} \tag{C3}$$

Therefore, we have that $U^u(Y_A, B_A) > U^u(Y_B, B_B)$ when the following condition holds:

$$\begin{aligned} &-2w^n\sqrt{\frac{1}{\pi}\left((w^n)^2 - 2U_{SB}^n\right)} + \frac{q_1}{w^u}2w^n\sqrt{\frac{1}{\pi}\left((w^n)^2 - 2U_{SB}^n\right)} \\ &+\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{\begin{array}{c}\left[(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left((w^n)^2 - 2U_{SB}^n\right)}\right]^2 \\ -\left[(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left((w^n)^2 - 2U_{SB}^n\right)}\right]^2\end{array}\right\} \\ &> \\ &\frac{q_3}{3}\left(\frac{1}{w^u}\right)^3\left\{\begin{array}{c}\left[(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left((w^n)^2 - 2U_{SB}^n\right)}\right]^3 \\ -\left[(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left((w^n)^2 - 2U_{SB}^n\right)}\right]^3\end{array}\right\}, \end{aligned}$$

or, equivalently:

$$\begin{aligned} & -2w^{n}\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}+2\frac{q_{1}}{w^{u}}w^{n}\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)} \\ & +2\left(\frac{1}{w^{u}}\right)^{2}(w^{n})^{3}\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)} \\ > & \\ & \frac{q_{3}}{3}\left(\frac{1}{w^{u}}\right)^{3}\left\{\begin{array}{c}\left[(w^{n})^{2}-w^{n}\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\right]^{3} \\ -\left[(w^{n})^{2}+w^{n}\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\right]^{3}\end{array}\right\}, \end{aligned}$$

and therefore:

$$w^{u}\left(w^{u}-q_{1}\right)-(w^{n})^{2} \quad < \quad \frac{(w^{n})^{2}q_{3}}{6w^{u}\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}}\left\{\begin{array}{c}\left[w^{n}+\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\right]^{3} \\ -\left[w^{n}-\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\right]^{3}\end{array}\right\}. \tag{C4}$$

Finally, since we have:[57]

$$\begin{aligned} & \left[w^{n}+\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\right]^{3}-\left[w^{n}-\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\right]^{3} \\ = \; & 2\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}\left[3(w^{n})^{2}+\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)\right], \end{aligned}$$

we can restate (C4) as

$$w^{u}\left(w^{u}-q_{1}\right)-(w^{n})^{2}<(w^{n})^{4}\frac{q_{3}}{w^{u}}\left[1+\frac{1}{3\pi}\frac{1}{(w^{n})^{2}}\left((w^{n})^{2}-2U_{SB}^{n}\right)\right]. \tag{C5}$$

When $Y_{LF}^{u} > Y_{LF}^{n}$ and $U^{u}(Y_{A}, B_{A}) > U^{u}(Y_{B}, B_{B})$ it will follow that a second-best optimum entails a downward distortion on the labor supply of users and therefore $Y_{SB}^{u} < Y_{SB}^{n}$. Thus, what we need to ascertain is whether it is possible that $Y_{LF}^{u} > Y_{LF}^{n}$ while $U^{u}(Y_{A}, B_{A}) > U^{u}(Y_{B}, B_{B})$. Putting together (C1) and (C5), the required condition is:

$$(w^{n})^{4}\frac{q_{3}}{w^{u}}<w^{u}\left(w^{u}-q_{1}\right)-(w^{n})^{2}<(w^{n})^{4}\frac{q_{3}}{w^{u}}\left[1+\frac{1}{3\pi}\frac{1}{(w^{n})^{2}}\left((w^{n})^{2}-2U_{SB}^{n}\right)\right].$$

Appendix D

Proof of the result that, with $\varphi(h)=\varphi(h)=\frac{Y}{w^{u}}q_{1}+\left(\frac{Y}{w^{u}}\right)^{1/2}2q_{2}$, it might be that $Y_{LF}^{u} > Y_{LF}^{n}$ while $Y_{SB}^{u} < Y_{SB}^{n}$: Consider a user whose earned income under laissez-faire

[57]We are applying the formula $(a+b)^{3}-(a-b)^{3}=\left(b^{2}+3a^{2}\right)2b$ with $a=w^{n}$ and $b=\sqrt{\frac{1}{\pi}\left((w^{n})^{2}-2U_{SB}^{n}\right)}$.

is slightly higher than the income earned by a non-user (which is equal to $(w^n)^2$). In particular, assume that $Y_{LF}^u = (w^n + \epsilon)^2$, where $\epsilon > 0$. Since the first order condition for an optimal labor supply (h) for users is

$$w^u - q_1 - h = q_2 h^{-1/2}, \tag{D1}$$

and since $h = Y_{LF}^u / w^u$, we have that

$$(w^u - q_1) w^u - (w^n + \epsilon)^2 = \frac{(w^u)^{3/2} q_2}{w^n + \epsilon},$$

or, equivalently:

$$(w^u - q_1) w^u - (w^n)^2 = \frac{(w^u)^{3/2} q_2}{w^n + \epsilon} + 2w^n\epsilon + \epsilon^2. \tag{D2}$$

Notice that the first order condition (D1) has at most two solutions for $Y > 0$. To capture a local maximum maximum rather than a minimum, it must be the case that, at $h = (w^n + \epsilon)^2 / w^u$ the slope of the left hand side of (D1) is lower than the slope of its the right hand side, namely that

$$-1 < -\frac{1}{2} q_2 \left(\frac{(w^n + \epsilon)^2}{w^u} \right)^{-3/2},$$

or, equivalently:

$$w^u < (w^n + \epsilon)^2 \left(\frac{2}{q_2} \right)^{2/3}. \tag{D3}$$

Assuming that (D3) is satisfied, it also follows that the right hand side of (D2) defines a value that is larger than $(w^u)^{3/2} q_2 / w^n$.[58] From (D3) it also follows that

$$\frac{\partial \left(\frac{(w^u)^{3/2} q_2}{w^n + \epsilon} + 2w^n\epsilon + \epsilon^2 \right)}{\partial \epsilon} = 2(w^n + \epsilon) - \frac{(w^u)^{3/2} q_2}{(w^n + \epsilon)^2} > 0,$$

so that (D2) can be rewritten as

$$(w^u - q_1) w^u - (w^n)^2 = \frac{(w^u)^{3/2} q_2}{w^n} + \delta, \tag{D4}$$

where $\delta > 0$ and $\delta \longrightarrow 0$ for $\epsilon \longrightarrow 0$. Assume that the socially desirable direction of redistribution is from non-users to users, so that $\overline{V}^n = U_{SB}^n < U_{LF}^n = (w^n)^2 / 2$. Following a procedure similar to that used in the proof of Proposition 1 (see Appendix A), we have that, when incentive-compatibility considerations require to offer users a distorted

[58] We have that $\frac{(w^u)^{3/2} q_2}{w^n + \epsilon} + 2w^n\epsilon + \epsilon^2 > \frac{(w^u)^{3/2} q_2}{w^n}$ when $(w^u)^{3/2} q_2 < (2w^n + \epsilon)(w^n + \epsilon) w^n$. Since (D3) can be restated as $(w^u)^{3/2} q_2 < 2(w^n + \epsilon)^3$, and given that $2(w^n + \epsilon)^3 < (2w^n + \epsilon)(w^n + \epsilon) w^n$, it follows that the right hand side of (D2) defines a value that is larger than $(w^u)^{3/2} q_2 / w^n$.

bundle, users will either be offered the bundle (A7)-(A8), entailing a downward distortion on their labor supply, or the bundle (A9)-(A10), entailing an upward distortion on their labor supply. Evaluating the utility of users at each of these two bundles we have:

$$
\begin{aligned}
U^u(Y_A, B_A) = & \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}U_{SB}^n - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} \\
& -\frac{q_1}{w^u}\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\} \\
& -\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\}^2 \\
& -2\frac{q_2}{(w^u)^{1/2}}\left((w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right)^{1/2},
\end{aligned}
$$

and

$$
\begin{aligned}
U^u(Y_B, B_B) = & \frac{1+\pi}{2\pi}(w^n)^2 - \frac{1-\pi}{\pi}U_{SB}^n + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} \\
& -\frac{q_1}{w^u}\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\} \\
& -\frac{1}{2}\left(\frac{1}{w^u}\right)^2\left\{(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right\}^2 \\
& -2\frac{q_2}{(w^u)^{1/2}}\left((w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right)^{1/2}.
\end{aligned}
$$

Therefore, we have that $U^u(Y_A, B_A) > U^u(Y_B, B_B)$ when the following condition holds:

$$
\begin{aligned}
& 2w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} - 2\frac{q_1}{w^u}w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} \\
& -2\left(\frac{1}{w^u}\right)^2(w^n)^3\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]} \\
< & \\
& 2\frac{q_2}{(w^u)^{1/2}}\left\{\begin{array}{c}\left[(w^n)^2 + w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right]^{1/2} \\ -\left[(w^n)^2 - w^n\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right]^{1/2}\end{array}\right\},
\end{aligned}
$$

or, equivalently:

$$
(w^u - q_1)w^u - (w^n)^2 < \frac{(w^u)^{3/2}q_2}{(w^n)^{1/2}\sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}}\left\{\begin{array}{c}\left[w^n + \sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right]^{1/2} \\ -\left[w^n - \sqrt{\frac{1}{\pi}\left[(w^n)^2 - 2U_{SB}^n\right]}\right]^{1/2}\end{array}\right\}.
$$

Applying the binomial expansion $(a+b)^{1/2}-(a-b)^{1/2}=ba^{-1/2}+b^3a^{-5/2}/8+...$ with $a=w^n$ and $b=\sqrt{\frac{1}{\pi}\left((w^n)^2-2U^n_{SB}\right)}$, we can rewrite the condition above as follows:

$$(w^u-q_1)\,w^u-(w^n)^2<$$
$$\frac{(w^u)^{3/2}q_2}{(w^n)^{1/2}\left[\frac{1}{\pi}\left((w^n)^2-2U^n_{SB}\right)\right]^{1/2}}\left\{\frac{\left[\frac{1}{\pi}\left((w^n)^2-2U^n_{SB}\right)\right]^{1/2}}{(w^n)^{1/2}}+\frac{\left[\frac{1}{\pi}\left((w^n)^2-2U^n_{SB}\right)\right]^{3/2}}{8\,(w^n)^{5/2}}+...\right\},$$

or, equivalently:

$$(w^u-q_1)\,w^u-(w^n)^2<\frac{(w^u)^{3/2}q_2}{w^n}\left\{1+\frac{1}{8\pi}\frac{(w^n)^2-2U^n_{SB}}{(w^n)^2}+...\right\}. \tag{D5}$$

When $Y^u_{LF}>Y^n_{LF}$ and $U^u(Y_A,B_A)>U^u(Y_B,B_B)$ it follows that a second-best optimum entails a downward distortion on the labor supply of users and therefore $Y^u_{SB}<Y^n_{SB}$. Putting together (D4), which provides a condition such that Y^u_{LF} is only slightly larger than Y^n_{LF}, and (D5), we have that a sufficient condition to have $Y^u_{LF}>Y^n_{LF}$ while $Y^u_{SB}<Y^n_{SB}$ is

$$\frac{(w^u)^{3/2}q_2}{w^n}<w^u(w^u-q_1)-(w^n)^2<\frac{(w^u)^{3/2}q_2}{w^n}\left[1+\frac{1}{8\pi}\frac{(w^n)^2-2U^n_{SB}}{(w^n)^2}+...\right].$$

Appendix E

Numerical example showing the possibility that a distortion arises even though no self-selection constraint is binding at a second-best optimum: Set $w^u=12.87$, $w^n=10$, $q_1=5$, $q_2=0.25$ and $\pi=1/5$. Under laissez-faire we have that $Y^u_{LF}=100.13$ and $Y^n_{LF}=(w^n)^2=100$, with $U^u_{LF}=29.57$ and $U^n_{LF}=50$. Assume that in the Pareto efficient tax problem $\overline{V}^n$ is set equal to 40.01. At a second-best optimum we get that $Y^u_{SB}=0$, so that the labor supply of users is distorted downwards, $Y^n_{SB}=100$ (no distortion on the labor supply of non-users), $U^n_{SB}=40.01$ and $U^u_{SB}=39.96$.[59] However, since the utility for a non-user choosing the bundle intended for users would be equal to 39.96, and the utility for a user choosing the bundle intended for non-users would be equal to 19.58, it follows that no self-selection constraint is binding at the second-best optimum. Nonetheless, observe that without a self-selection constraint requiring non-users not to be tempted to mimic users, the latter could have been offered an undistorted bundle (in our example, the bundle $(Y,B)=(100.13,140.09)$).

Appendix F

[59] We also have $B^u_{SB}=39.96$ and $B^n_{SB}=90.01$. Notice also that the second-best optimum features income re-ranking with respect to the laissez-faire equilibrium.

Proof of the result that (18) can be rewritten as (19): Rewrite (18) as

$$\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]-(w^n)^2+\frac{1}{2}(w^u-q)^2+q\frac{(w^n)^2}{w^u}+\frac{1}{2}\left(\frac{(w^n)^2}{w^u}\right)^2\geq qs\frac{(w^n)^2}{w^u}.$$

Multiplying both sides by $w^u/(w^n)^2$ gives

$$\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\frac{w^u}{(w^n)^2}-(w^u-q)+\frac{1}{2}(w^u-q)^2\frac{w^u}{(w^n)^2}+\frac{1}{2}\frac{(w^n)^2}{w^u}\geq qs.$$

Substituting for s the value provided by (17) gives:

$$\begin{aligned}&\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\frac{w^u}{(w^n)^2}-(w^u-q)+\frac{1}{2}(w^u-q)^2\frac{w^u}{(w^n)^2}+\frac{1}{2}\frac{(w^n)^2}{w^u}\\ \geq\ &\frac{(w^u-q)\,w^u+\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]-\frac{1}{2}\frac{(w^n)^4+[(w^u-q)w^u]^2}{(w^n)^2}}{w^u-q}.\end{aligned}$$

Multiplying both sides by w^u-q gives

$$\begin{aligned}&\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\frac{w^u(w^u-q)}{(w^n)^2}-(w^u-q)^2+\frac{1}{2}\frac{(w^u-q)^3w^u}{(w^n)^2}+\frac{1}{2}\frac{(w^n)^2(w^u-q)}{w^u}\\ \geq\ &\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]-\frac{1}{2}\frac{(w^n)^4+[(w^u-q)\,w^u]^2}{(w^n)^2}+w^u(w^u-q),\end{aligned}$$

which can be rewritten as

$$\begin{aligned}&\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\frac{w^u(w^u-q)-(w^n)^2}{(w^n)^2}\\ \geq\ &\frac{2(w^u)^2(w^n)^2(w^u-q)+2(w^u-q)^2(w^n)^2w^u}{2(w^n)^2w^u}\\ &-\frac{(w^u-q)^3(w^u)^2+(w^n)^4(w^u-q)+(w^n)^4w^u+(w^u-q)^2(w^u)^3}{2(w^n)^2w^u}.\end{aligned}$$

Multiplying both sides by $2(w^n)^2w^u$ gives

$$\begin{aligned}&\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\left[w^u(w^u-q)-(w^n)^2\right]2w^u\\ \geq\ &2(w^u)^2(w^n)^2(w^u-q)+2(w^u-q)^2(w^n)^2w^u-(w^u-q)^3(w^u)^2\\ &-(w^n)^4(w^u-q)-(w^n)^4w^u-(w^u-q)^2(w^u)^3,\end{aligned}$$

or, equivalently:

$$\begin{aligned}&\frac{1}{\pi}\left[\frac{1}{2}(w^n)^2-\overline{V}^n\right]\left[w^u(w^u-q)-(w^n)^2\right]2w^u\\ \geq\ &2(w^u-q)(w^n)^2w^u(2w^u-q)-(w^u-q)^2(w^u)^2(2w^u-q)\\ &-(w^n)^4(2w^u-q).\end{aligned}$$

Collecting terms we can write

$$\begin{aligned} & \frac{1}{\pi}\left[\frac{1}{2}\left(w^n\right)^2-\overline{V}^n\right]\left[w^u\left(w^u-q\right)-\left(w^n\right)^2\right]2w^u \\ \geq \quad & -\left(2w^u-q\right)\left[-2\left(w^u-q\right)\left(w^n\right)^2w^u+\left(w^u-q\right)^2\left(w^u\right)^2+\left(w^n\right)^4\right], \end{aligned}$$

or, equivalently:

$$\frac{2}{\pi}\left[\frac{1}{2}\left(w^n\right)^2-\overline{V}^n\right]\left[w^u\left(w^u-q\right)-\left(w^n\right)^2\right]w^u \geq -\left(2w^u-q\right)\left[w^u\left(w^u-q\right)-\left(w^n\right)^2\right]^2.$$

References

[1] Anderberg, D. and A. Balestrino (2000) Household production and the design of the tax structure. *International Tax and Public Finance* **7**, 563-584.

[2] Bastani, S., Blomquist, S. and L. Micheletto (2013) The welfare gains of age-related optimal income taxation. *International Economic Review* **54**, 1219-1249.

[3] Bastani, S., Blomquist, S. and L. Micheletto (2017) Child care subsidies, quality, and optimal income taxation. CESifo Working Paper no. 6533.

[4] Bierbrauer, F. J. and P. C. Boyer (2014) The Pareto-frontier in a simple Mirrleesian model of income taxation. *Annals of Economics and Statistics* **113/114**, 185-206.

[5] Blomquist, S., Christiansen, V. and L. Micheletto (2010) Public provision of private goods and nondistortionary marginal tax rates. *American Economic Journal: Economic Policy* **2**, 1-27.

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

[7] Choné, P. and G. Laroque (2010) Negative marginal tax rates and heterogeneity. *American Economic Review* **100**, 2532-2547.

[8] Cremer, H. and F. Gahvari (2002) Nonlinear pricing, redistribution and optimal tax policy. *Journal of Public Economic Theory* **4**, 139-161.

[9] Cremer, H., Gahvari, F. and N. Ladoux (1998) Externalities and optimal taxation. *Journal of Public Economics* **70**, 343-364.

[10] Cremer, H., Pestieau, P. and J. C. Rochet (2001) Direct versus indirect taxation: the design of the tax structure revisited. *International Economic Review* **42**, 781-799.

[11] Gahvari, F. (2007) On optimal commodity taxes when consumption is time consuming. *Journal of Public Economic Theory* **9**, 1-27.

[12] Golosov, M., Troshkin, M., Tsyvinski, A., and M. Weinzierl (2013) Preference heterogeneity and optimal capital income taxation. *Journal of Public Economics* **97**, 160-175.

[13] Ho, C. and N. Pavoni (2018) Efficient child care subsidies. Mimeo.

[14] Jacobs, B. and R. Boadway (2014) Optimal linear commodity taxation under optimal non-linear income taxation. *Journal of Public Economics* **117**, 201-210.

[15] Jacquet, L., Lehmann, E. and B. Van der Linden (2013) Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory* **148**, 1770-1805.

[16] Judd, K., Ma, D., Saunders, M. A. and C. L. Su (2018) Optimal income taxation with multidimensional taxpayer types. Mimeo.

[17] Kleven, H. J., Kreiner, C. T. and E. Saez (2009) The optimal income taxation of couples. *Econometrica* **77**, 537-560.

[18] Koehne S. and D. Sachs (2017) Pareto efficient tax breaks. Mimeo.

[19] Lockwood, B. B. and M. Weinzierl (2015) De gustibus non est taxandum: heterogeneity in preferences and optimal redistribution. *Journal of Public Economics* **124**, 74-80.

[20] Micheletto, L. (2008) Redistribution and optimal mixed taxation in the presence of consumption externalities. *Journal of Public Economics* **92**, 2262-2274.

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

[22] Rothschild, C. and F. Scheuer (2014) A theory of income taxation under multidimensional skill heterogeneity. NBER Working Paper No. 19822.

[23] Scheuer, F. (2014) Entrepreneurial taxation with endogenous entry. *American Economic Journal: Economic Policy* **6**, 126-163.