Preference Responsibility versus Poverty Reduction in the Taxation of Labor Incomes*

Lancelot Henry de Frahan† François Maniquet‡

January 2021

Abstract

We study the tax schedules that maximize social welfare functions built on axioms of responsibility for one's preferences (the requirement that the social welfare function should treat identically agents with the same wage, independently of their preferences) and poverty reduction. We find zero and negative marginal tax rates on low incomes at the optimum and bunching at the income level of the most hardworking minimum wage households. When preferences are iso-elastic, we derive the optimal tax formula, which we calibrate to the US economy. Our formula approximates the shape of the current US tax function for households with at least one child. This result suggests that a fairness-based approach, and these axioms in particular, can help close the gap between the recommendations of optimal tax theory and actual policies.

JEL Classification: H21, D63.

Keywords: optimal taxation, responsibility, poverty reduction.

---

*François Maniquet's work was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC n° 269831. This research advanced while Maniquet was visiting the department of economics of University College London. Generous support from the Leverhulme Foundation is gratefully acknowledged. Declarations of interest: none.

†Becker Friedman Institute at the University of Chicago, Chicago, 60637, USA. Email: lancelot@uchicago.edu

‡CORE, Université catholique de Louvain, Louvain-La-Neuve, 1348, Belgium. Email: francois.maniquet@uclouvain.be.

> [...] a pivotal part of this economic plan is increasing the earned-income tax credit which, more than anything else we could do, will reward work and family and responsibility [...] this will be the first time in the history of our country when we will be able to say that if you work 40 hours a week and you have children in your home, you will be lifted out of poverty."

President Bill Clinton, July 29, 1993 (shortly before expanding the Earned Income Tax Credit)

1 Introduction

Income taxation in general, and labor income taxation in particular, is the ultimate policy instrument to reduce income inequality. Since Mirrlees (1971)'s seminal contribution, most of the literature has embraced the view that full income equality is not a legitimate objective. The literature on optimal taxation is mainly welfarist: the utilitarian objective, especially when individual utility functions are concave, has long been seen appropriate to reduce inequality without eliminating it. Several authors, however, have recently expressed dissatisfaction with the utilitarian objective, in particular when preference heterogeneity is taken into account (see, for instance, Boadway, 2012, or Piketty and Saez, 2013). In addition, as reflected in the above quotation of President Bill Clinton, public tax policy discussions rarely revolve around utilitarian principles and often invoke some notion of "fairness" or other non-welfarist considerations. Finally, Weinzierl (2014) also finds survey-based evidence that most Americans reject important implications of utilitarianism.

An alternative approach to utilitarianism has been recently proposed, after the introduction of the ethics of responsibility into normative debates (see detailed accounts in Roemer, 1998, and Fleurbaey, 2008). This ethics recently emerged in political philosophy and yielded theories of equality of resources and equality of opportunities, to list just two main examples (see Roemer, 1998, and Fleurbaey, 2008 for discussions and references). In a nutshell, the ethics of responsibility is grounded on the assumption that not all inequalities are unjust. As a consequence, the social objective should be to identify and eliminate the unjust inequalities and to remain neutral towards the other inequalities. Several authors have studied the consequences on income taxation of considering that income inequalities due to differences in labor time are not unjust (see Fleurbaey and Maniquet, 2006 and 2007, and Lockwood and Weinzierl, 2015). To put it differently, under the latter view, agents with the same wage rate should be free to choose their labor time and redistributing incomes among them is not legitimate. In addition to existing references to this view by politicians (see, for instance, the quotation of President Bill Clinton above), it seems to reflect views of citizens. Using questionnaire-based experiments, Konow (2001) presents hypothetical redistribution situations to American subjects. Redistributing production among equally skilled individuals having exerted different levels of efforts is largely viewed as unfair, whereas redistributing from high-skill to low-skill individuals is viewed as fair.[1]

One may argue, however, that freeness to choose may lead to freeness to lose, if there is not enough redistribution to lift agents out of poverty. Also using questionnaire experiments, Weinzierl (2014) investigates the popular support to redistributing towards the poor. Subjects have to express their opinion over actual and counterfactual tax-transfer schemes in the US. Weinzierl observes that a majority of subjects agree to subsidize the poor but less than what utilitarian or maximin welfarist preferences would recommend. That raises the question of the compatibility between the fairness principles of responsibility and poverty reduction.

In this paper, we derive the formula of the optimal labor income tax scheme consistent with the fairness principles of responsibility and poverty reduction and we calibrate the formula to the US economy. More specifically, we first build a family of social welfare functions that combine neutrality towards inequalities caused by different labor time choices with the goal of reducing poverty. Second, we derive an optimal tax formula from the maximization of these unconventional social welfare functions under incentive and budget constraints.[2] Third, using Current Population Survey (CPS) data, in combination with an estimate of the labor supply elasticity from Chetty (2012), we calibrate the formula to the US economy. Finally, we compare

[1]Schokkaert and Devooght (2003) confirm this finding for different countries. See Gaertner and Schokkaert (2012), chapter 4, for a survey. Saez and Stantcheva (2016, Online Appendix C) also report survey evidence according to which citizens are more enclined to be generous towards those who work hard at a low wage than towards those who work less at a higher wage, even if they earn the same income.

[2]To be clear, we consider the optimal tax function under the assumption that the planner only observe agents' incomes. Because of the latter informational constraint, and despite the normative goal of responsibility, agents with the same wage rate but different incomes may still end up being taxed differently at the solution. This would not be the case in a first-best context.

the resulting calibrations with the current system.

There are three main lessons to draw from our paper. First, it is possible to define a social objective combining the three goals of poverty alleviation, responsibility and efficiency and to derive the resulting optimal tax. Surprisingly enough, the most basic tension among these three goals is the conflict between poverty reduction and Pareto efficiency. If being poor means consuming a bundle of goods below some poverty line, the main issue comes from the fact that agents may prefer bundles below the poverty line to bundles above it if the labor time associated to the latter is too large. This issue echoes analyses of optimal labor income tax that reduces poverty when poverty is merely defined in terms of income (see, for instance, Kanbur, Keen and Tuomala, 1994a and b, and Wane, 2001), and the resulting optimal tax schemes are Pareto inefficient. We propose a definition of "being poor" that is consistent with Pareto efficiency. Somewhat surprisingly, it can easily be further adjusted to also be compatible with the requirement of responsibility (that is the requirement that the social welfare function should be such that, absent any informational constraint, no redistribution take place among agents with the same wage). Being poor, under this definition, means consuming a bundle on an indifference curve that lies everywhere below the poverty line.

The second main lesson of our paper is that it is possible to fully characterize the optimal income tax derived from an unconventional social objective that embodies notions of fairness. Indeed, on the one hand, the literature on optimal taxation has successfully obtained and calibrated expressions detailing the optimal tax as a function of behavioral responses, distribution of income (or types) and a set of weights capturing the normative preferences of society. It has provided very little guidance, however, on how to choose those weights.[3] When the objective is utilitarian, for instance, weights depend on an arbitrary cardinalization of the utility function used to describe behavioral responses.

On the other hand, a long tradition in the social choice literature has highlighted how different fairness principles can be called for to characterize social welfare functions (see, among many others, the surveys of Roemer 1998 or Fleurbaey and Maniquet 2011). Though some criteria for the taxation of labor income have been pre-

[3] An exception is given by the so-called Rawlsian (or maximin) objective, in which all weights are set to zero except those of the lowest-skill agents. This Rawlsian objective has no straightforward generalization when preferences differ.

viously derived from these social welfare functions, this literature has not derived the precise formula of the tax function, leading to somewhat of a disconnect with the optimal tax literature (in addition to not allowing for precise calibrations). As a result, we view this study as an example of how to bridge the divide between the (essentially axiomatic) fairness approach to optimal tax and a classical social welfare function based approach.

The third main lesson of our paper is that the optimal tax function that is consistent with our fairness principles shares some salient features of the current income tax schedule in the US. As can be seen on Figure 1, the optimal tax corresponding to the combination of Pareto efficiency, responsibility and poverty reduction exhibits zero and negative marginal tax rates on low-incomes as well as a discrete jump to positive marginal tax rates around $16,000. These features are somewhat unconventional in the optimal taxation literature and yet approximate some important characteristics of the current US income tax. The jump mimics the phaseout of the Earned Income Tax Credit (EITC) and of various welfare programs. The existence of non-positive marginal tax rates on low incomes and the steep jump towards positive rates is a very robust result under our normative objective and does not depend on the specific parameterizations of the poverty line, the agents' preferences or the distribution of types. Finally, depending on the specific parametrization of the poverty line, optimal marginal tax rates on high-incomes may be lower than would be implied by setting the marginal welfare weights to zero in the upper tail (a common normative assumption in the optimal tax literature under utilitarianism).[4] We discuss more precisely the similarities and differences between our calibrations and the current US tax schedule in Section 5.

To be precise, we don't claim that the current US tax schedule is the outcome of maximizing a normative objective similar to ours. Of course a tax schedule is the outcome of a series of reforms inspired by different normative and electoral objectives. What we do claim is that the tax schedule that emerges as optimal in our study *is* reasonable, given its similarity with an existing one. In addition, we believe that our analysis shows that switching from the traditional utilitarian normative assumption to a fairness-based approach can help close the gap between the recommendations of optimal tax models and the reality of policy.

[4]Lockwood and Weinzierl (2016) have previously noted that existing tax policies are consistent with less redistributive preferences that is traditionally assumed in the literature.

The rest of the paper is organized as follows. In Section 2, we define the model and the basic properties. In Section 3, we introduce a definition of poverty reduction that is compatible, as a normative goal, with Pareto efficiency and responsibility for agents' preferences and we define a prominent family of social welfare functions combining the three goals we are interested in as well as a fourth, auxiliary, property. In Section 4, assuming that preferences are quasi-linear and iso-elastic (customary assumptions in the literature), we derive the formula of the second-best tax schemes that maximize our social objective. In Section 5, we estimate the distribution of types using Current Population Survey (CPS) microdata and calibrate the tax formula to the U.S. economy. We also use the CPS data and NBER's TAXSIM to obtain a description of the current tax system and put our results into perspective.[5] Finally, in Section 6, we compare the ability of our results to mimic the U.S. system with that of other optimal tax results in the literature. In Section 7, we give some concluding comments.

2 The Model and the Basic Axioms

There are two goods, labor time, denoted ℓ, and consumption, denoted c. A bundle (of goods) is a pair $z = (\ell, c) \in X = \mathbb{R}^2_+$. There is a finite set N of agents. Each agent $i \in N$ is characterized by their wage, $w_i \in [\underline{w}, \infty)$, and their preferences $R_i \in \mathcal{R}$ over bundles. Wages are assumed to be bounded from below. Preferences are assumed to be continuous, increasing in consumption, decreasing in labor time, convex, and such that a best bundle exists in each budget.[6] An economy is a list of characteristics $E = (w_N, R_N) \in \mathcal{E} = ([\underline{w}, \infty)_+ \times \mathcal{R})^N$, one pair of characteristics for each agent. An allocation is a list of bundles $z_N = (z_i)_{i \in N}$, one bundle for each agent.

We are interested in ranking allocations as a function of the parameters of the economy. Formally, a social welfare function, in short a SWF, is a function $\mathbf{R}$, such that for each economy $E \in \mathcal{E}$, $\mathbf{R}(E)$ is an ordering on the set of allocations X^N. For two allocations $z_N, z'_N \in X^N$, we write $z_N \,\mathbf{R}(E)\, z'_N$, resp. $z_N \,\mathbf{P}(E)\, z'_N$, $z_N \,\mathbf{I}(E)\, z'_N$, to denote that z_N is socially as good as z'_N, resp. strictly better, indifferent.

The following notation will prove useful. For any preferences $R \in \mathcal{R}$ and set

[5] Regarding TAXSIM, see Feenberg and Coutts (1993).

[6] Formally, the latter condition requires that for all $w_i \in [\underline{w}, \infty)$, $R_i \in \mathcal{R}$ and $b \in \mathbb{R}$, there exists $z^* = (\ell^*, c^*) \in X$ such that $c^* \leq b + w_i \ell^*$ and $z^* \, R_i \, z$ for all $z = (\ell, c)$ such that $c \leq b + w_i \ell$.

$A \subset X$, we write $m(R, A)$ to denote the set of best bundles in A according to R:

$$m(R, A) = \{z \in A \mid \forall z' \in A : z \, R \, z'\}.$$

In case $m(R, A)$ contains more than one bundle, all bundles of $m(R, A)$ are indifferent for R. For any $z = (\ell, c) \in X$ and $w \in [\underline{w}, \infty)$, we write $B(z, w)$ to denote the budget of slope w going through z:

$$B(z, w) = \{z' = (\ell', c') \in X \mid c' - w\ell' \leq c - w\ell\}.$$

We write $B(z, w) + \Delta$ to denote the budget set obtained by translating $B(z, w)$ by an amount Δ, that is $\{z' = (\ell', c') \in X \mid c' - w\ell' \leq c - w\ell + \Delta\}$.

Finally, for any $z \in X$, $R \in \mathcal{R}$ and $w \in [\underline{w}, \infty)$, we write $IB(z, R, w)$ to denote the implicit budget of an agent (R, w) at z, that is the budget of slope w that leaves this agent indifferent between z and maximizing over that budget:

$$IB(z, R, w) = B(z', w) \text{ such that } z' \, I \, z \text{ and } z' \in m(R, IB(z, R, w)).$$

We complete this section with the definition of two well-known axioms that we impose on our SWFs. Strong Pareto is the requirement that if everyone considers his bundle in z_N at least as good as in z'_N, then z_N must be socially as good as z'_N. If, in addition, at least one agent strictly prefers his bundle in z_N, z_N must be deemed a strict social improvement.

Axiom 1 *strong Pareto*
For all $E = (w_N, R_N) \in \mathcal{E}$, *all* $z_N, z'_N \in X^N$,

$$[z_i R_i z'_i, \ \forall i \in N] \Rightarrow [z_N \mathbf{R}(E) z'_N]$$

$$\textit{and } [z_i R_i z'_i, \ \forall i \in N, \ \textit{and } \exists j \in N : \ z_j P_j z'_j] \Rightarrow [z_N \mathbf{P}(E) z'_N]$$

We now define the axiom capturing the fairness principle of responsibility. It captures the idea that income differences merely due to labor time differences should not be deemed unjust. Another way of stating this principle is by reference to full income (that is, income computed when leisure is valued at wage rate): two agents with the same wage rates should obtain the same full income. In terms of tax and transfer, it means that these two agents should ideally receive the same transfer or pay the same tax. Fleurbaey and Maniquet (2005, 2006) have proposed to formalize this principle as follows: take two agents with the same wage rate. Ideally, they should have the same full income, that is the bundles they consume should be their preferred bundle in the same budget. Assume, to the contrary, that the bundles they consume do not arise from the same budget. The following axiom says that such a *budget* inequality is unjust. Consequently, a budget inequality reducing transfer (as opposed to a bundle inequality reducing transfer) among them is a social improvement.

Axiom 2 *equal-wage transfer*
For all $E = (w_N, R_N) \in \mathcal{E}$, $z_N, z'_N \in X^N$, if there exist $j, k \in N$ and $\Delta \in \mathbb{R}_+$ such that $w_j = w_k$,

$$z'_j \in m\left(R_j, B(z'_j, w_j)\right), z_j \in m\left(R_j, B(z_j, w_j)\right),$$

$$z'_k \in m\left(R_k, B(z'_k, w_k)\right), z_k \in m\left(R_k, B(z_k, w_k)\right),$$

and

$$B(z'_j, w_j) = B(z_j, w_j) - \Delta \supseteq B(z'_k, w_k) = B(z_k, w_k) + \Delta,$$

and for all $i \neq j, k$: $z_i = z'_i$, then $z'_N \, \mathbf{P}(E) \, z_N$.

In the complete absence of taxation, that is at a laisser-faire allocation, all agents are treated identically by the tax system, as they do not receive nor pay anything. Moreover, the resulting allocation is Pareto efficient. This means that there exist SWFs, which consider laisser-faire allocations as social optima, satisfying both *strong Pareto* and *equal-wage transfer*, implying that these two axioms are compatible with each other. The drawback of laisser-faire, however, is that some agents may end up in material deprivation, and the more so if they have small wages. The main goal of this paper is to study the consequences of combining *strong Pareto* and *equal-wage transfer* with the idea that society should not let agents end up at too low of a material standard. We introduce such an axiom of poverty reduction in the next section.

3 Poverty Reduction

Studying poverty reduction first requires introducing a poverty line in the model. We indeed assume that there is a thin, connected and strictly increasing set of bundles in the consumption set that captures the desire by society to let all agents live a materially decent life. Let $PL \subset X$ satisfy the following properties:

- for all $z = (\ell, c), z' = (\ell', c') \in PL$, either $z = z'$ or $\ell < \ell' \Leftrightarrow c < c'$,
- for all $\ell \in \mathbb{R}$, there exists $c \in \mathbb{R}$ such that $(\ell, c) \in PL$,
- the set of bundles above PL is convex.

One comment is in order. Our poverty line is strictly increasing, that is, the minimal consumption level that makes a life materially decent is assumed to increase with the amount of labor performed by the agent. This can be interpreted in two ways. First, as labor time rises the agent's basic needs increase: more food is required by the organism to afford those efforts, clothes wear out faster, etc. Second, leisure can be considered as a basic need so that less leisure needs to be compensated by more income (we may think for instance of the cost parents incur for hiring someone to take care of their children when they work full time). All results derived in the paper hold for poverty lines arbitrarily close to an horizontal line so that our conclusions are not a byproduct of the assumption that PL is increasing in labor time. To further illustrate this point, we include the limiting case of a flat poverty line in the calibrations of online appendix E. Finally, note that, besides its shape, how high PL lies in the consumption set of the agents also determines the generosity of society towards the poor.

We write $z = (\ell, c) > PL$, for any $z \in X$, to describe the situation in which there is $(\ell, c') \in PL$ such that $c > c'$, and $A > PL$, for any $A \subset X$, to describe the situation in which $z > PL$ for all $z \in A$. We refer to indifference curves as to $I(z, R)$, that is, for any $z \in X$, $R \in \mathcal{R}$,

$$I(z, R) = \{z' \in X \mid z' \, I \, z\}.$$

We now define an axiom reflecting the objective of poverty reduction. Basically, we state that a lump-sum transfer of consumption from a non-poor to a poor agent must be a social improvement. This raises the question of who is (non-)poor. As we look for an axiom of poverty reduction that is compatible with *strong Pareto* and *equal-wage transfer*, we face constraints on the way we can define poverty. The main difficulty is this one: if poverty is defined as consuming a bundle below the poverty line, then this contradicts *strong Pareto* because an agent can be indifferent between a bundle below (in which case she is poor) and a bundle above (in which case she is non-poor) the poverty line.

The difficulty to define poverty reduction in a way that is consistent with *strong Pareto* also echoes the literature initiated by Fleurbaey and Trannoy (2003). This literature has explored the conditions under which a bundle transfer axiom is compatible with strong Pareto. Applying the results from this literature (see, for instance, the survey in Fleurbaey and Maniquet, 2011, chapter 2), we assume that an agent is poor (resp. non-poor) if her indifference curve lies below (resp. above) the poverty line, and we require that a bundle inequality reducing transfer of consumption from a non-poor to a poor agents that have the same labor time is a social improvement.

Axiom 3 *poverty-reduction transfer*
For all $E = (w_N, R_N) \in \mathcal{E}$, *all* $\Delta \in \mathbb{R}_+$, *all* $j, k \in N$, *all* $z_N = (l_N, c_N)$, $z'_N = (l'_N, c'_N) \in X^N$ *such that* $l_j = l'_j = l_k = l'_k$, $c'_j = c_j - \Delta$, $c'_k = c_k + \Delta$ *and for all* $i \neq j, k : z_i = z'_i$,

$$[I(z'_j, R_j) > PL > I(z'_k, R_k)] \Rightarrow [z'_N \mathbf{R}(E) z_N].$$

This axiom requires a number of comments. First, the definition of being poor conveyed by this axiom is much weaker than a definition in terms of consuming a bundle below the poverty line. As a result, an agent is neither poor nor non-poor if her indifference curve crosses the poverty line. Depending on the preferences of the agent, it may be the case that she is neither poor nor non-poor in a very large part of her indifference map. This does not imply, however, that only a tiny fraction of the agents will end up being poor. Our definition, indeed, can still be made as demanding as one wishes, by setting the poverty line higher in the consumption set. In the calibrations below, we make both the slope and the intercept of the poverty line vary.

Second, this definition can be given a welfarist flavor. Given that poverty is not defined based on the bundle one consumes but based on her indifference curve, we can think of being poor as having a utility level below some threshold. In such a welfarist frame, one should bear in mind that the utility threshold below which one qualifies as poor and the utility threshold above which one qualifies as non-poor typically do not coincide. As a consequence, this definition does not equip us with much comparability among agents. The welfare of one agent can be declared larger than that of another agent only if the former agent's utility level is above her individual non-poverty threshold whereas the latter agent's utility level is below her individual poverty threshold.

Third, the tension between the objectives of responsibility and poverty reduction, as declared in the title of the paper, is illustrated by the weakness of the notion of poverty embedded in this axiom. For the sake of completeness, we discuss in Appendix A stronger and maybe more natural axioms of poverty reduction that fail to be compatible with *strong Pareto* and *equal-wage transfer*.

Finally, on a more technical note, this allows us to give an additional justification for our somehow unconventional assumption of a poverty line that is strictly increasing in labor time. As it is customary in the optimal taxation literature, we assume that there is no bound on labor time. Finiteness of individually optimal labor times is then deduced from the assumption of an increasing disutility of labor. With a flat poverty line, therefore, we would have faced the difficulty that nobody ever has her indifference curve entirely below the poverty line (except for agents with flat indifference curves), so that no agent would have been declared poor according to our definition. Having a strictly increasing poverty line is, therefore, a natural assumption to study our weak axiom of poverty reduction in a model without a natural bound on the labor supply.

The combination of *strong Pareto*, *equal-wage transfer*, *poverty reduction transfer* and an auxiliary axiom of *separability*, which we formally define in Appendix B, leads us to consider the family of SWFs which we call $\mathbf{R}^{\tilde{R}lex}$. A $\mathbf{R}^{\tilde{R}lex}$ SWF first computes a well-being level for each agent at their assigned bundle, and then it applies the leximin criterion $\geq_{lex}$ to vectors of well-being levels.[7] The well-being measure associated to a $\mathbf{R}^{\tilde{R}lex}$ SWF is parameterized by some individual reference preferences $\tilde{R} \in \mathcal{R}$. These reference preferences need to be chosen among the preferences that contain the poverty line as an indifference curve: for all $z = (\ell, c), z' = (\ell', c') \in PL$,

$$z \, \tilde{I} \, z'.$$

[7]The leximin criterion consists in applying the maximin criterion lexicographically, that is a vector weakly dominates another vector according to the leximin if the minimal component of the latter vector is larger than the minimal component of the former vector, or they are equal and the second minimal component of the latter vector is larger than the second minimal component of the former vector, etc., or they are equal.

Figure 2 illustrates how well-being levels are computed. The well-being level of an agent (w_i, R_i) consuming bundle z_i is computed by looking at the bundle that $\tilde{R}$ prefers in the implicit budget of (w_i, R_i) at z_i. Formally, Let $\tilde{u} : X \to \mathbb{R}$ be a utility representation of $\tilde{R}$. Then, for $z_N, z'_N \in X^N$,

$$z_N \, \mathbf{R}^{\tilde{R}lex} \, z'_N \Leftrightarrow$$

$$(b(z_i, w_i, R_i))_{i \in N} \geq_{lex} (b(z'_i, w_i, R_i))_{i \in N},$$

where for all $z \in X$, $w \in [\underline{w}, \infty)$, and $R \in \mathcal{R}$,

$$b(z, w, R) = \tilde{u}(m(\tilde{R}, IB(z, w, R))).$$

The fact that $\mathbf{R}^{\tilde{R}lex}$ satisfies *strong Pareto* follows from $b(\cdot, w, R)$ being a utility representation of R and the leximin criterion being strictly increasing in all its arguments. The fact that $\mathbf{R}^{\tilde{R}lex}$ satisfies *equal-wage transfer* comes from the fact that the well-being index $b(\cdot, w, R)$ is computed using implicit budgets. As a result, if $w_j = w_k$, as soon as $z_j \in m\left(R_j, B(z_j, w_j)\right)$ and $z_k \in m\left(R_k, B(z_k, w_k)\right)$, we have

$$b(z_j, w_j, R_j) \geq b(z_k, w_k, R_k) \Leftrightarrow B(z_j, w_j) \supset B(z_k, w_k),$$

and the leximin criterion will give priority to the agent with the smallest budget.

The fact that $\mathbf{R}^{\tilde{R}lex}$ satisfies *poverty-reduction transfer* is related to the specific property satisfied by $\tilde{R}$. Indeed, let us assume that $I(z_j, R_j) > PL > I(z_k, R_k)$. It is sufficient to prove that $b(z_j, w_j, R_j) > b(z_k, w_k, R_k)$ because the leximin criterion gives priority to the lowest well-being level. First, $I(z_j, R_j) > PL$ implies that $IB(z_j, w_j, R_j)$ contains a part of PL. As a result, maximizing $\tilde{R}$ over $IB(z_j, w_j, R_j)$ must yield a utility level larger than at PL (remember PL is one of the indifference curves of $\tilde{R}$). On the contrary, $PL > I(z_k, R_k)$ implies that $IB(z_k, w_k, R_k)$ is everywhere below PL, so that maximizing $\tilde{R}$ over $IB(z_k, w_k, R_k)$ must yield a utility level lower than at PL, the desired result.

This proves that the $\mathbf{R}^{\tilde{R}lex}$ SWFs satisfy the three axioms we have defined. We can also prove that all SWFs satisfying these three axioms and *separability* rank allocations around the poverty line in the same way as $\mathbf{R}^{\tilde{R}lex}$ SWFs, that is they all rank allocations in the same way as $R^{\tilde{R}lex}$ when an increase in well-being of an agent whose implicit budget lies below the poverty line is accompanied by a decrease in that of an agent above the poverty line. In other words, the combination of the aforementioned axioms *locally* characterize social preferences around the poverty line. Given that this proof is similar to existing proofs in the literature, we send it to Section B in the appendix. This result, though, calls for several remarks.

First, the *separability* axiom that we use for this result formalizes the idea that indifferent agents should not influence the social ranking. This is a classical axiom in social choice theory and it is now well established that combining *strong Pareto*, *separability* and some inequality reduction axiom similar to *poverty reduction transfer* leads to SWFs of the leximin type, that is, the combination of these axioms strengthens a requirement of inequality aversion (like our axiom of *poverty reduction transfer*) into a requirement of infinite aversion to this inequality (see, for instance, the survey in Fleurbaey and Maniquet, 2011, chapter 3).

Second, per this local characterization, the definition of poor and non-poor agents has changed, to encompass many more situations. It is now sufficient to look at the implicit budget of the agents. As soon as the frontier of the implicit budget of an agent is everywhere below the poverty line, this agent qualifies as poor, whether or not this agent's entire indifference curve lies below it. As soon as the frontier of the implicit budget of an agent is somewhere above the poverty line, this agent qualifies as non-poor, whether or not this agent's entire indifference curve lies above the poverty line. Observe that with this new definition, agents are either poor or non-poor at an allocation; that is there is no subset of consumption bundles at which an agent is neither non-poor nor poor.

Third, all $\mathbf{R}^{\tilde{R}lex}$ SWFs identify in exactly the same way poor and non-poor agents because this only depends on whether their implicit budget lies below the poverty line or not. Consequently, we can deduce from this result that any allocation in which poverty is eliminated is socially preferable to any allocation in which at least one agent's implicit budget is still below the poverty line. This teaches us what the program of the planner has to be: designing a tax scheme that eliminates poverty in the sense that the implicit budget of no agent lies below the poverty line. This is formally stated and proven in the appendix. Of course, poverty being eliminated in this sense does not guarantee that all agents' *bundles* are above the poverty line.

This conclusion is illustrated in Figure 3. Agent j consumes z_j, which lies above the poverty line, and her indifference curve through z_j crosses the poverty line. If the wage of j is low, for instance if $w_j = w$, then her implicit budget, drawn in the figure, is everywhere below the poverty line (slopes of implicit budget lines are mentioned within parentheses below the lines). Agent j qualifies as poor, and a transfer from a non-poor agent to agent j is a social improvement. If, on the other hand, her wage is high, for instance if $w_j = w'$, then her implicit budget crosses the poverty line and she qualifies as non-poor. That illustrates the fact that well-being indices depend on wage. In the case of agent j, at the same bundle, her well-being is an increasing function of wage.

The contrary happens for agent k, when she consumes z_k, below the poverty line. We have $b(z_k, w, R_k) > b(z_k, w', R_k)$ because her implicit budget lies below the poverty line when $w_k = w'$ but not when $w_k = w$. The difference between agents j and k comes from the fact that agent j has preferences that are more work averse than the reference preferences (illustrated by the fact that her indifference curve crosses the poverty line from below) and the contrary is true for agent k. That implies that high wage agents may have low well-being levels, if their preferences are sufficiently less directed towards leisure than the reference preferences. This will have potentially drastic consequences on the shape of the optimal tax on high incomes. In our calibrations, for poverty lines that are steep enough, marginal tax rates in the right-tail are lower than the rate that maximizes revenue raised from high incomes.

4 Optimal Tax Functions

In the remaining of the paper, we study the properties of the tax schemes that maximize the SWFs we have justified in the previous sections. We adopt the classical Mirrlees setting. The planner does not observe the characteristics of the agents. She only observes their income. She has to design a redistributive tax scheme $\tau : \mathbb{R}_+ \to \mathbb{R}$. As a result, an agent with wage $w \in [\underline{w}, \infty)$ and a labor time of ℓ earns income $y = w\ell$, pays income tax $\tau(y)$ and consumes $c = y - \tau(y)$.

Given the complexity of the problem, we assume that agents' preferences are quasi-linear and iso-elastic: they are represented by

$$u_i(c, \ell) = c - \frac{\epsilon}{1+\epsilon}\left(\frac{\ell}{\theta_i}\right)^{\frac{1+\epsilon}{\epsilon}}. \tag{1}$$

Agents differ in their taste for leisure only through parameter $\theta_i \in [\underline{\theta}, \overline{\theta}]$. We define $n_1 = \underline{w}\overline{\theta}$, the type of the agent with the lowest skill and the lowest taste for leisure, that is, n_1 is the parameter of all agents who are behaviorally identical to the hard-working poor agent. Their earning level is denoted by y_1. We further assume that the poverty line itself, PL, has the shape of an iso-elastic indifference curve.

$$PL = \left\{ z = (c, \ell) \in X = \mathbb{R}^2_+ \mid \bar{P}_0 = c - \frac{\epsilon}{1+\epsilon} (\ell/\tilde{\theta})^{\frac{1+\epsilon}{\epsilon}} \right\}.$$

Studying tax schemes forces us to study the income/consumption space instead of the labor time/consumption space. That requires some rewriting. Abusing notation, however, we keep u_i to denote the utility function of agent i with wage w_i in the income/consumption space, and we write

$$u_i(c, y) = c - \frac{\epsilon}{1+\epsilon} \left(\frac{y}{w_i \theta_i} \right)^{\frac{1+\epsilon}{\epsilon}}. \tag{2}$$

This assumption, also used by Lockwood and Weinzierl (2015), reduces the two dimensions of heterogeneity to a single index in the income consumption space: $n_i = w_i \theta_i$.[8] All agents with the same parameter n_i will be behaviorally equivalent (that is they all have the same preferences in the income/consumption space) even if they differ in their wage.

The economies we are interested in are economies with a finite but large number of agents. As is customary in the optimal tax literature, we capture the assumption of a large economy by assuming that there is a continuum of types. We then assume that $n_i \sim F(n_i)$ in the population and the associated density is denoted by $f(n_i)$. Notice that $F(n_i)$ is a marginal distribution generated by the joint distribution of θ_i and w_i. We assume, in addition, that the joint probability density of (w_i, θ_i) is strictly positive everywhere on $[\underline{w}, \infty) \times \left[\underline{\theta}, \overline{\theta}\right]$. An important implication is that $\forall n \in [\underline{n}, n_1]$, there exist a non-zero density of agents earning the minimum wage $\underline{w}$.[9]

The constraints of the SWF maximization are well-known and we review them quickly. We begin with the incentive compatibility constraints. The planner chooses income-consumption bundles $(y(n), c(n))$ for each "behavioral" type n. We write $u(n)$ the utility of agents of type n at bundles $(y(n), c(n))$. Incentive compatibility

[8]More generally, collapsing multiple dimensions of heterogeneity to one parameter in the context of optimal taxation has been suggested by Mirrlees (1976) and used by Brett and Weymark (2003) and Choné and Laroque (2010).

[9]Similarly, $\forall n \in [n_1, \infty)$, there exist some agents with $\theta_i = \overline{\theta}$ and $w_i = n_i/\overline{\theta}$.

constraints require:

$$u(n) = \max_{n' \in [\underline{n}, \infty)} c(n') - \frac{\epsilon}{1+\epsilon} \left(\frac{y(n')}{n} \right)^{\frac{1+\epsilon}{\epsilon}} = c(n) - \frac{\epsilon}{1+\epsilon} \left(\frac{y(n)}{n} \right)^{\frac{1+\epsilon}{\epsilon}}. \tag{3}$$

In addition to incentive compatibility, the planner maximizes the SWF under the constraint that at least some revenue S be raised to finance public goods, which, using Eq. (2), can be written

$$\int_{\underline{n}}^{\infty} \left(y(n) - \left(u(n) + \frac{\epsilon}{1+\epsilon} \left(\frac{y(n)}{n} \right)^{\frac{1+\epsilon}{\epsilon}} \right) \right) dF(n) \geq S. \tag{4}$$

We now turn to the definition of the social objective. Agents with the same n may not all have the same well-being index $b(z_i, w_i, R_i)$, because it depends on their underlying wage w_i and preference parameter θ_i. For any $\mathbf{R}^{\tilde{R}lex}$, however, we can restrict our attention to the minimum well-being among agents of type n at utility level $u(n)$, which we denote $\underline{b}(u(n); n)$.

4.1 Quasi-linear $\tilde{R}$

We know from Section 3 that $\mathbf{R}^{\tilde{R}lex}$ is such that the reference preferences $\tilde{R}$ contain the poverty line as an indifference curve. While $\tilde{R}$ are thus characterized at the poverty line, the remaining indifference curves need to be drawn. For now, we consider the case where $\tilde{R}$ are quasi-linear. We will see in Section 4.2 how the solution to this case can be amended so as to provide a solution to two other polar choices of $\mathbf{R}^{\tilde{R}lex}$ as well. All three polar choices are considered in the calibrations of Section 5.

Under quasi-linear reference preferences, $\tilde{R}$ can be represented by the following utility function:

$$\tilde{u}(c, \ell) = c - \frac{\epsilon}{1+\epsilon} \left(\frac{l}{\tilde{\theta}} \right)^{\frac{1+\epsilon}{\epsilon}}. \tag{5}$$

The problem of the planner consists in maximizing $\mathbf{R}^{\tilde{R}lex}$, in which $\tilde{R}$ is represented by Eq. (5), under incentive constraints (3) and resource constraint (4). The following proposition, illustrated in Figure 4, completely characterizes the optimal tax scheme under the assumption that the distribution of types has an upper tail Pareto index α.

Proposition 1 *If $F(n)$ has a Pareto index α in the upper tail and if the lower-bound on utilities binds on one interval above y_1^*, then there exist $n_u \leq n_{\underline{b}} \leq n_{\bar{b}} \leq n_l \leq n' \in [\underline{n}, \infty)$, such that $\tau'(y(n))$ satisfies*

1. $\tau'(y(n)) = 0 \ \forall n \in [\underline{n}, n_u]$

2. $\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left(\frac{1+\epsilon}{\epsilon}\right) \frac{F(n_u)-F(n)}{nf(n)} \quad \forall n \in [n_u, n_{\underline{b}}]$

3. $y(n) = y_1^* \quad \forall n \in [n_{\underline{b}}, n_{\bar{b}}]$

4. $\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left(\frac{1+\epsilon}{\epsilon}\right) \frac{[F(n_l)-F(n)] + \frac{1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}}{\left(\frac{1+\epsilon}{\epsilon}\right)\left[1-\left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}} n_l f(n_l)}{nf(n)} \quad \forall n \in [n_{\bar{b}}, n_l]$

5. *If* $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} \geq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ *: then*

   $\tau'(y(n)) = 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \quad \forall n \in [n_l, \infty]$

6. *Otherwise,*

   $\tau'(y(n)) = 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \quad \forall n \in [n_l, n']$

   and

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = \frac{1-F(n)}{nf(n)} \left(\frac{1+\epsilon}{\epsilon}\right) \quad \forall n \in [n', \infty)$$

   and the marginal tax rate converges to $\tau'(y(n)) \to \frac{1+\epsilon}{1+\epsilon+\alpha\epsilon}$ *as* $n \to \infty$.

This proposition, whose proof is relegated to appendix D, gives the formula that will be used in the next section for the calibrations. Per the leximin nature of $\mathbf{R}^{\tilde{R}lex}$, the formula features intervals of n, $[\underline{n}, n_u]$ and $[n_l, n']$, over which the well-being index $\underline{b}(u(n); n)$ is equalized. On the remaining intervals, efficiency considerations govern the optimal tax rate. We give a more detailed explanation of why these different intervals appear at the solution and how the values of their endpoints are determined in online appendix Section I.2. Several comments are in order to interpret the result.

1. Very low incomes, that is incomes in the range $[y(\underline{n}), y(n_u)]$, are taxed at a zero marginal rate. First, it is important to note that this result does not depend

on the assumption that preferences are quasi-linear and iso-elastic. As it is transparent from the proof, this result is fully general, and holds even if both the extensive and the intensive reaction to taxation are taken into account.[10] Second, if it turns out that $y(\underline{n}) = 0$ (which is the relevant case), there may be some bunching at this earning level.

2. Incomes in the range $[y(n_u), y(n_{\underline{b}})]$ face a negative income tax rate. The result that marginal tax rates should be zero or negative below $y^*(n_1)$ is more related to the fact that the social welfare function satisfies *equal-wage transfer* than *poverty reduction transfer*. Indeed, it is the combination of *strong Pareto* and *equal-wage transfer* that forces us to measure well-being by using implicit budgets, with the result that well-being is equalized among minimal wage agents if they all maximize their utility over the same real (first-best) budget, as it is the case with a zero marginal tax rate. Surprisingly, the fact that the social welfare function also satisfies *poverty-reduction transfer* is not seen in the shape of the marginal tax rates on low incomes.[11]

3. The next feature of the optimal tax is that there is bunching at the earning level of agent 1. Indeed, all types of agents in the range $[n_{\underline{b}}, n_{\bar{b}}]$ earn the same income. The solution shares some similarities with the "selfishly optimal non-linear income tax schedules" studied by Brett and Weymark (2017). In both cases, the tax schedule is one that maximizes redistribution towards agents with some skill (n_1 under our notation) in the middle of the distribution (under the constraint that no other agent end up with a level of well-being below that of n_1). Ignoring the constraint that $y(n)$ be increasing in n_1, this leads to negative marginal tax rates to the left and positive marginal tax rates to the right of n_1. Bunching occurs at n_1 to maintain the monotonicity of $y(n)$. On the other hand, the tax rates we obtain in the $[\underline{n}, n_u]$ and $[n_{\bar{b}}, n_l]$ intervals differ from what Brett and Weymark (2017) obtain because they directly follow from our constraint that the well-being of no agent should be lower than that of agent 1.

[10] In order to underscore this point, we formally derive this generalization in online appendix J.

[11] This also explains the similarity with the optimal tax rates on low incomes derived in Fleurbaey and Maniquet (2007), which studies a social objective combining *equal-wage transfer* with an axiom of equality in utility among agents with the same preferences. Under the assumption that labor time is bounded above, they obtain a zero marginal tax rate below $y^*(n_1)$. Proposition 1 generalizes their result, making intervals $\left[y(n_u), y(n_{\underline{b}})\right]$ and $[n_{\bar{b}}, n_l]$ appear.

4. Because the SWF is a leximin, the optimal tax rate above $y_1^* = y(n_1)$ is the one that equalizes the well-being index of agents with $\theta_i = \overline{\theta}$ (they are the worst-off within each n) unless there exists a Pareto improvement. The possibility of a Pareto improvement is the reason why the revenue-maximizing tax rate plays a key role in the proposition. The revenue-maximizing tax rate is described by $\frac{\tau'(y(n))}{1-\tau'(y(n))} = \frac{1-F(n)}{nf(n)}\left(\frac{1+\epsilon}{\epsilon}\right)$ which implies $\tau'(n) \to \frac{1+\epsilon}{1+\epsilon+\alpha\epsilon}$ in the upper-tail when $F(n)$ has Pareto index α. Let us first consider large incomes: we compare the revenue-maximizing tax rate in the tail to the well-being equalizing tax rate, $1-\left[1-\left(\frac{\tilde{\theta}}{\overline{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$. If the former is lower than the latter, taxing revenue in the upper-tail at $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon}$ is a Pareto improvement over equalizing well-being since it raises more revenue while agents in that part of the distribution also face a lower tax rate. A similar reasoning applies to incomes right above $y(n_1)$. Recall that agents with $n \in (n_1, n_{\overline{b}}]$ all choose income $y(n_1)$ and have well-being index strictly larger than 1's. As a result, the lower-bound on utilities derived from the leximin nature of the social objective is not binding at $n_{\overline{b}}$.

5. We can intuitively sum up the influence of each axiom embedded in the social objective on the optimal tax as follows. First, *equal-wage transfer* implies that the marginal rate is zero on very low incomes and negative on the following interval. Second, *poverty reduction transfer* is key in 1) determining the amount of transfer to the lowest income earners, and a higher poverty line implies a larger transfer, and 2) determining the marginal tax rates on incomes above $y(n_{\overline{b}})$, where a steeper poverty line yields a lower tax rate. Finally, *strong Pareto* is responsible for the shape of the tax function around $y(n_1)$ and on very large incomes.

4.2 Three Polar Choices for $\tilde{R}$

Proposition 3 in the appendix characterizes social preferences locally around the poverty line. The fact that the characterization is only local is reflected in that the reference preferences $\tilde{R}$ must have an indifference curve that coincides with the poverty line but the rest of the indifference map is not characterized. A corollary of Proposition 3 is that, under *strong Pareto*, *equal-wage transfer*, *poverty reduction transfer* and *separability*, the tax schedule must prioritize lifting all implicit budgets above the poverty line (see Corollary 1 in appendix Section C).

Assume that we have identified the allocation that just achieves this. It is typically the case that some additional surplus can still be produced. The question becomes how do we allocate this additional surplus? Maximizing a SWF from the $\mathbf{R}^{\tilde{R}lex}$ family ensures that such potential surplus is redistributed in a way that is consistent with the aforementioned four axioms. On the other hand, the $\mathbf{R}^{\tilde{R}lex}$ family encompasses many choices of reference preferences above the poverty line. The shape of these indifference curves govern how the surplus beyond poverty eradication is to be redistributed across the distribution of n. We consider three polar cases.

The most redistributive $\mathbf{R}^{\tilde{R}lex}$ consists in maximizing the surplus that is redistributed to agents with low n while the least redistributive $\mathbf{R}^{\tilde{R}lex}$ minimizes redistribution above what is necessary to lift all implicit budgets above the poverty line. The third, intermediary, way of allocating the surplus consists in distributing it equally among agents. When preferences are quasi-linear, equal distribution means that we equalize the utility gain above the allocation at which poverty is eliminated. This is accomplished by choosing the $\tilde{R}$ that is itself quasi-linear. Proposition 1 was derived under the latter assumption. In appendix H, we provide a detailed derivation of the optimal tax function in the other two cases. We briefly describe the corresponding solutions in the next two paragraphs. In addition, figure 5 presents the shape of the indifference curves above the poverty line for the least- and most redistributive $\tilde{R}$. Finally, and most importantly, note that in the calibrations, all three polar cases end up corresponding to surprisingly similar optimal tax schemes.

From Proposition 1, we know that steeper reference preferences (i.e. a steeper poverty line) lead to less redistribution. As a result, the least redistributive reference preferences consistent with our axioms are such that the indifference curves become labor averse as quickly as possible above the poverty line. In order to obtain such preferences, we rotate the relevant indifference curves around $(c, \ell) = \left(\overline{P}_0, 0\right)$ by letting $\tilde{\theta}$ become smaller. Under these reference preferences, the formula in Proposition 1 describes the optimal tax by replacing $\tilde{\theta}$ with $\tilde{\theta}'$. The latter is the lowest $\tilde{\theta}$ so that the constraint that everybody (and in particular agents with $n = n_1$) be lifted out of poverty is satisfied.

The opposite reasoning applies to the most redistributive $\mathbf{R}^{\tilde{R}lex}$: we let the reference preferences become as hardworking as possible above the poverty line. Given that there is no upper bound on labor time, there is no natural bundle around which to rotate the reference indifference curves, unlike in the least redistributive case. Therefore, we fix a sufficiently large ℓ'' and we define the reference preferences in such a way that indifference curves rotate around bundle (c'', ℓ'') on the poverty line. As a result, there is a kink in the preferences at ℓ'' that can be seen on figure 5. The resulting optimal tax function resembles the one described in Proposition 1 in that there are intervals over which the minimal well-being is equalized across n and intervals over which the lower-bound is not binding. The lower-bound changes in the following ways. First, it remains the same as the allocation that just eliminates poverty for agents with $n \geq n''$ where $\frac{n''}{\tilde{\theta}}$ is the wage level at which preferences characterized by $\tilde{\theta}$ would choose a labor time of ℓ''. In other words, ℓ'' and n'' are defined so that, along the lower-bound, the utility of agents with $n \geq n''$ is at the poverty line and they do not reap any share of the surplus above the allocation that eradicates poverty. Second, there is a portion $[n_0, n'']$ along which the marginal tax rate is decreasing. This corresponds to the wage levels at which the reference preferences are tangent to the budget set at ℓ''. Third, let $\tilde{\theta}''$ be the value associated to the highest reference indifference curve for which the resource constraint is binding, for all agents with $n \in [n_1, n_0]$, the marginal tax rate along the lower-bound is $1 - \left[1 - \left(\frac{\tilde{\theta}''}{\tilde{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$.[12] As far as the optimal allocation is concerned, Proposition 1 applies, $\forall n \in [\underline{n}, n_0)$, with $\tilde{\theta} = \tilde{\theta}''$. For $[n_0, \infty)$, the solution is the well-being equalizing allocation previously described. In some cases, and similar to Proposition 1, there exists some $n' \in [n_0, \infty)$ after which applying the revenue maximizing tax rate generates a Pareto improvement.

5 Calibration to the US Economy

The optimal tax formula that we obtained in the previous section leaves us with a number of questions that cannot be answered theoretically and that require some calibrations. In general, we are interested in the magnitudes implied by our optimal formula as well as the differences between the three polar choices of $\tilde{R}$. More specifically, we focus on four questions - introduced below - that calibrations help us answer. Before proceeding to the analysis, let us briefly describe the data sources and

[12] More precisely, as $\tilde{\theta}''$ increases, so does the marginal tax rate along the lower-bound over $[n_1, n'')$. Holding the utility of agents with $n = n''$ at the poverty line, the utility of all agents with $n < n''$ is increased as $\tilde{\theta}''$ is raised. A larger $\tilde{\theta}''$ therefore requires larger aggregate resources in the economy. At the solution, $\tilde{\theta}''$ is such that the resource constraint is just binding.

methods used to calibrate.

The distribution of types n in the economy is estimated via Maximum Likelihood using CPS microdata. We estimate a four parameter double Pareto Lognormal distribution (DPLn). For each household in the data, we simulate the marginal federal tax rate they face using NBER's TAXSIM and estimate the implicit marginal tax rate from the phaseout of government assistance. Conditional on a calibrated labor supply elasticity parameter, we can recover agents' types n_i by inverting the first-order conditions of the agents' problem. The parameters of the DPLn distribution are then estimated by maximizing the log-likelihood of this sample of n_i. We pick a labor supply elasticity $\epsilon = 0.33$ which is the preferred estimate in Chetty (2012). We use information about hours worked and hourly wage rates in our sample of CPS data to discipline $\overline{\theta}$. Finally, we need to calibrate the level and slope of the poverty line. These are normative choices. For this reason, we present the results for many different choices of these parameters in appendix E. In the main text, we start by assuming that $\overline{P}_0 = \$18,850$. According to the 2015 CPS (March Supplement), the average household size in the US is 2.54. The weighted average of the 2014 poverty lines for households of size 3 is $\$18,850$ (DeNavas-Walt and Proctor 2015). It is slightly more challenging to calibrate the average slope of the poverty line but the federal minimum wage ($\$7.25/h$) may be a useful benchmark. We start the calibrations with a value of $\$6/h$. More details about the calibrations are provided in appendix F.

Figure 6 and the middle panel of table 1 show the calibrated optimal tax rates and consumption schedules according to all three choices of $\tilde{R}$. We focus on four characteristics of the calibrated optimal allocations. First, the optimal consumption of agents with zero earnings range from $\$11,009 - \$12,025$. As previously noted, poverty can be eradicated in the sense that all implicit budgets are at or above the poverty line, yet some agents consume a bundle that is located below the line. In these calibrations, this is the case for all agents earning no income which illustrates the trade-off between poverty and responsibility. Second, the formula recommends a sharp increase in the marginal taxation rates around the income earned by those who work full time at the minimum wage (income y_1^* in Proposition 1). The second question we would like to answer is: how big should this increase be? It is very similar across the three polar choices of $\tilde{R}$. The tax rates jump from around -0.29 to around 0.58. Third, the formula also requires marginal tax rates to be decreasing in intervals just before (between the income levels earned by n_u and $n_{\underline{b}}$ in the proposition) and just after (between the income levels earned by $n_{\underline{b}}$ and n_l) this sharp increase. How large should these intervals be and how decreasing should the rates be over this interval is our third question. Across all three $\tilde{R}$, these two intervals are located around $\$13,900 - \$16,300$ and $\$16,300 - \$19,160$.

Consistently with what we said when deriving this formula, we do not think, however, that how marginal rates evolve in these two intervals is essential because it crucially depends on the specific shape of the utility function that we assumed. More important is the fourth question. The formula recommends that, quickly after the decrease in marginal tax rates that follows the sharp increase around the minimum wage, rates remain stable on a rather large interval (interval $[n_l, n']$ in the proposition). Around which level should rates stabilize and how large should this interval be? The answer is remarkably consistent across the three polar cases: $0.47 - 0.50$. Notice that these rates apply from $y(n_l)$ to all larger incomes. In these calibrations, tax rates on large incomes are determined by equalizing the well-being $\underline{b}(u(n), n)$ across agents with $n \in [n_l, \infty)$. This is mainly due to the average slope of the poverty line being set at $\$6/h$. For flatter poverty lines, tax rates on the largest incomes are sometimes determined so as to maximize revenue collected to be re-distributed.

An important lesson from these calibrations is that, while in theory there may be large differences between the three polar cases of $\tilde{R}$, empirically, at these values of $\overline{P}_0$ and $\tilde{\theta}$, optimal tax rates and consumption levels are not very sensitive to the choice of $\tilde{R}$. This conclusion is supported for instance by the small differences in $c(0)$ or in the tax rate on large incomes. Intuitively, if the poverty line is sufficiently high relative to aggregate resources in the economy, the focus on poverty reduction and responsibility almost completely pins down the optimal allocation even though social preferences are only fully characterized in the neighborhood of the poverty line. As we can see in the top panel of table 1, choices of lower poverty lines leave more room for an additional normative choice to be made about how to redistribute the surplus above the allocation that just eliminates poverty. On the contrary, if the poverty line is so high that not everyone is lifted out of poverty at the optimum, the three cases collapse to a unique solution. This is the case in the bottom panel of table 1.

Figure 1 plots the calibrated optimal tax rates alongside the current marginal tax rates for single filers with 0, 1 and 2 or more children respectively. We show the marginal tax rates for joint filers in appendix F. In the US, households with at least one child face negative marginal tax rates on incomes below $\$6,000$ to $\$13,000$

depending on the number of children and filing status. This is driven by the Earned Income Tax Credit - though some of the work incentives from the EITC are undone by the implicit positive marginal tax rate generated by the phaseout of welfare programs such as SSI, TANF and SNAP. Marginal rates tend to increase steeply roughly around $\$10,000$ and reach up to 60% for households with more than 2 children.[13] Marginal tax rates then tend to decline (after a plateau) and stabilize around 40% for high incomes. Households with no children do not face negative marginal tax rates, nor do they face as high positive marginal tax rates around $\$10,000$ to $\$45,000$ because both the EITC and the welfare system are less generous towards families without children.

Interestingly, our calibrated optimal tax rates resemble the tax schedule faced by households with at least one child. Our formula recommends a sharp increase in marginal rates around the income earned by those working full time at the minimal wage and this is what we observe for households with children.[14] Our formula also recommends that marginal rates stay around zero before this sharp increase and stay more or less constant afterwards. This is again consistent with the way households with children are taxed in the US. Besides these similarities, there are four major differences. First, we cannot rationalize negative marginal rates in the neighborhood of zero earning. The rates should be zero according to our formula. This is not a huge difference, but the current scheme is better explained in this interval by arguments in terms of extensive margin effects (people deciding to move from unemployment to employment), like in Saez (2002), or in terms of incentives to go beyond overestimated costs of finding jobs, such as in Lockwood (Forthcoming).

The second difference between our optimal tax scheme and the current one in the US is that it is hard to rationalize a brutal decrease in marginal tax rates after the phasing out of the EITC, followed by a plateau. Our results push towards a smoothing of this part of the tax system, where the constraint that the well-being of agents earning incomes in this bracket should not be lower than the well-being of the hardworking poor is likely to be binding.

Third, in the calibrations, tax rates on larger incomes $(0.47 - 0.49)$ are about

[13]In appendix F, we show, for each household type, the break down of the marginal tax rate between the federal income tax and the phaseout of welfare payments.

[14]The exact income, y_1^*, at which the jump happens is determined endogeneously at the solution and depends on the calibrated value of $n_1 = \underline{w}\overline{\theta}$. We pick $\underline{w} = \$7.25$ (the federal minimum wage). We calibrate $\overline{\theta}$ using the $95th$ percentile of the empirical distribution of θ_i in a sample of hourly workers in the CPS. More details are provided in online appendix F.

10 percent higher than the current US federal schedule implying more redistribution towards everybody else. On the other hand, they remain lower than those that would maximize revenue raised at the top (0.54).[15] It is worth noting that, unlike the other differences and similarities mentioned here, these values are somewhat sensitive to the particular choice of poverty line and reference preferences chosen for the calibration.

Finally, the fourth difference is that the US tax system for childless households does not look at all like the one our axioms rationalize. It is therefore clear that responsibility and poverty reduction may be seen as related to the tax system only for what concerns households with children. Furthermore, this observation is consistent with Bill Clinton's quote at the beginning of this article which only mentions families with children.

6 Relationship to the literature

In this section, we compare the shape of our optimal tax scheme with optimal tax schemes derived from other SWFs.

First, when agents are assumed to have identical preferences, in which case the SWF is often written as utilitarian, marginal tax rates are typically large on low incomes and decreasing after, contrary to what we obtain (see Diamond, 1998, for a survey of the consequences of the utilitarian objective). This is the case under the maximin objective in incomes as well. This maximin objective also leads to different conclusions than ours regarding the taxation of large incomes, when our lower bound on utilities is binding. At the risk of underlining something obvious, let us recall that a utilitarian optimal tax depends 1) on the cardinalization of the utility functions representing preferences, whereas our result does not, and 2) on the shape of the density function of types, whereas our result does not, except on intervals in which the optimal tax is determined by efficiency only (intervals $[n_u, n_{\underline{b}}]$, $[n_{\bar{b}}, n_l]$ and $([n', \infty))$.

Many authors, however, have reached similar conclusions as ours about the optimality of non-positive tax rates on low earning levels from different viewpoints. First, negative marginal tax rates are optimal when labor responses on the extensive margin

[15] The revenue maximizing rate is likely an underestimate. This is because our estimate, in CPS data, of the upper-tail Pareto parameter implies a thinner tail of the income distribution than studies using administrative data. We discuss this point more in depth in appendix G and, importantly, we show that the results of our main calibrations are not sensitive to assuming a thicker tail.

are important either because of a fixed cost of labor participation (see Saez, 2002, Diamond, 1980 or Blundell and Shephard, 2012), or a present bias (see Lockwood, Forthcoming).

Second, negative marginal tax rates can be justified under the maximization of a utilitarian objective when preferences differ and social weights are a function of these preferences (see Boadway, Marchand, Pestieau and Racionero, 2002).

Third, many studies have already identified a connection between poverty reduction and negative marginal tax rates on low incomes. Saez and Stantcheva (2016) reach that conclusion from a SWF that minimizes the poverty gap when it is defined in a way that is compatible with Pareto efficiency. Maniquet and Neumann (Forthcoming) find negative marginal rates on earnings that lead to an after-tax income below the poverty line (but they do not derive the formula of the optimal tax), under the requirement that redistributing from rich to poor is desirable only if they have the same labor time. In both cases, negative marginal rates are immediately related to the objective of poverty reduction, whereas it is more related to the objective of responsibility in our case.

Two rather different SWFs, capturing a goal of poverty reduction that conflicts with Pareto efficiency, lead to negative tax rates on low incomes. Wane (2001) models aggregate poverty as an externality that individuals have a willingness to pay to reduce. As a result, society's concern for poverty reduction can be addressed through pigouvian earning subsidies to the poor. At the optimal tax, the poorest are worse-off than if the negative externality caused by poverty is ignored in the SWF. In Kanbur, Keen and Tuomala (1994a), the social planner minimizes an Atkinson index of poverty that only depends on incomes. As a result, individuals at the bottom of the income distribution are incentivized to work longer than required by efficiency so as to increase their income and decrease poverty.

Related to our approach, Fleurbaey and Maniquet (2006, 2007) obtain non-positive marginal tax rates from social objectives that are also justified by fairness principles. The SWF used in Fleurbaey and Maniquet (2006) does not satisfy our responsibility axiom but a weaker axiom based on the idea that no redistribution should take place in economies in which all agents would have the same wage rate. As a result, the objective of the planner should be to maximize the average income subsidy rate over low incomes and to maximize the tax return on all incomes above the one of the hardworking poor agent. The SWFs used in Fleurbaey and Maniquet (2007) are closer to ours, as they satisfy the same responsibility axiom as ours (our axiom of *equal-wage transfer*). The axiom with which it is combined is logically unrelated to our axiom of poverty reduction, but the resulting SWFs are similar to ours. Their result that marginal tax rates should be zero over an interval of low incomes is similar to ours, as it also derived from their SWFs satisfying responsibility (see remark 2 after Proposition 1 above). Importantly, Fleurbaey and Maniquet (2006, 2007) do not derive the precise formula of the optimal tax function, but we believe the method we use to solve our problem could be used to provide the formula of the optimal tax for their SWF too. On a more technical note, the results they obtained are grounded on the simplifying assumption that there is an upper bound on the labor supply, a (natural) assumption that is not typical in the optimal tax literature and that we don't impose in this paper.

Some authors have also studied SWF based on responsibility. Saez and Stantcheva (2016) present the results of their income weights approach when the objective consists of the equality of opportunity objective of Roemer et al. (2003). By assumptions, incomes are determined by agents' socio-economic background and by their efforts. The objective is then to equalize as much as possible the incomes of agents exerting the same level of effort. As a consequence, income weights are proportional to the share of the population coming from low background and earning that income level. As this share decreases with income but does not converge to zero, the optimal tax is U-shaped, tax rates are lower than what the classical utilitarianism would yield and tax rates on very high incomes are lower than under the income maximin objective. The latter feature of their result is the only one similar to ours.

Lockwood and Weinzierl (2015) study a similar objective of responsibility as ours. They do not axiomatize their SWF, but they show that if a fixed heterogeneity of types is more and more due to differences in preferences, which do not call for redistribution, then the optimal tax is less progressive, with lower but yet typically positive tax rates on low incomes.

To conclude this section, we can add that none of the optimal formulas derived in this literature would reproduce the US labor income tax of households with children as closely as our formula does. The closest to the current shape of the US tax is given by the formula of Saez and Stantcheva (2016), according to which tax rates should be negative until the income of those agents whose consumption is equal to the poverty line, after which taxe rates jump to a very high level. However, the earning threshold at which the jump takes place is lower than under the current US tax, and tax rates beyond this threshold are those that maximize tax revenues, much larger than the current ones.

7 Conclusion

This paper starts an analysis of optimal taxation at the level of axioms that the social welfare function should satisfy and it completes the analysis at the level of the full characterization of the optimal tax formula, a formula that is then calibrated to the US economy.

The axioms that we started with embed fairness principles. For this reason, this paper follows a current trend of the literature in which the emphasis shifts from the classical utilitarian objective, in which the social optimum is defined in terms of allocation of subjective utility, to the fairness of the resulting allocation of resources, in this case labor time and consumption.

The two main fairness norms that we have studied are responsibility for one's preferences and poverty reduction. The former is related to the requirement that, absent any information constraint, no redistribution should take place among agents having the same ability to earn income, a requirement that seems to be central to the many approaches to fairness in labor income taxation (see Fleurbaey and Maniquet, 2018, for a survey). The latter, poverty reduction, has already received attention in the literature, but in a way that is incompatible with Pareto efficiency, with the consequence that, contrary to what we obtain here, an increase in the social objective can be concomitant with a decrease in the utility of the poor. The axiom of poverty reduction that we have studied here avoids this paradox, but at the price of weakening the definition of being poor. Rather than defining poverty as consuming a bundle below the poverty line, we defined it as consuming a bundle on an indifference curve entirely below the poverty line.

As a result, a tax scheme may be optimal in our sense and, yet, leave some agents consume bundles below the poverty line. What the optimal tax scheme accomplishes, though, is to avoid redistribution among the low-wage agents. As a result, the marginal tax rates are zero or even negative up to the earning level of the low-wage agent choosing the longest labor time. Consequently, the budget of these agents crosses the poverty line from below, so that agents end up being lifted out of poverty, that is consuming bundles above the poverty line, provided their labor time is sufficient.

Another consequence of our axioms is that, in spite of our social preference taking the shape of an extremely egalitarian one, the optimal tax rates on high incomes may differ from the ones that would maximize the tax revenue (the so-called Rawlsian tax scheme). This is even more true as we choose a steeper or lower poverty line. This has to do with the way we define the well-being index which the social preference tries to maximin. Indeed, it is typically not true that an agent with a larger wage ends up with a larger well-being level (whereas in the classical framework an agent with a larger wage always ends up with a higher utility). This comes from the responsibility axiom, which forces us to look at implicit budgets rather than consumption bundles.

Finally, we compare our calibrated tax rates to the current US tax system. Despite a few differences, it is surprisingly close to the one that we rationalize with our axioms, at least as far as households with one or more children are concerned. This shows that switching from the traditional utilitarian normative assumption to a fairness-based approach can help close the gap between the recommendations of optimal tax models and the reality of policy.

A recent development - and attractive feature - of the modern optimal tax literature has been to derive formulae that are somewhat robust to model specification and are expressed in terms of sufficient statistics. Though our results do not share this feature, we view it as an unavoidable consequence of the normative stance taken in this paper: we assume that society aims at correcting for certain sources of inequalities, the ones generating poverty, but remain neutral towards others, the ones emerging from different labor time choices. This requires a *model* of what the sources of inequality are and it is therefore not surprising that the corresponding tax formula is *model dependent*. This relates to a long tradition in ethics positing that normative judgments must be context dependent.

Acknowledgements

We thank two anonymous referees, the editor, Marianne Bertrand, Stéphane Bonhomme, Richard Blundell, Nathaniel Hendren, Hilary Hoynes, Guy Laroque, Casey Mulligan, Doron Ravid, Raaj Sah, Johannes Spinnewyn, Matthew Weinzierl, Glen Weyl and John Weymark as well as seminar participants at the LSE, IFS, LISER, the Universities of Chicago, Cologne, Manchester, Oslo, the Delhi School of Economics and conference attendees at the 2016 and 2018 Social Choice and Welfare meetings at Lund and Seoul, 2017 RIDGE conference in public finance, Montevideo, for their comments.

References

[1] Blundell, Richard and Shephard, Andrew. (2012). "Employment, hours of work and the optimal taxation of low-income families," *The Review of Economic Studies* 79(2), 481–510.

[2] Boadway, R. 2012, *From Optimal Tax Theory to Tax Policy*, Cambridge and New York: MIT Press.

[3] Boadway, R. and L. Jacquet. (2008), "Optimal marginal and average income taxation under maximin." *Journal of Economic Theory* 143, 425–441.

[4] Boadway, Robin, Marchand, Maurice, Pestieau, Pierre and Racionero, Maria del Mar. (2002). "Optimal redistribution with heterogenous preferences for leisure," *Journal of Public Economic Theory*, 4(4) 475–498.

[5] Brett, Craig, and John A. Weymark. (2003). "Financing Education Using Optimal Redistributive Taxation." *Journal of Public Economics*, 87: 2549–2569.

[6] Brett, Craig, and John A. Weymark. (2017). "Voting over selfishly optimal nonlinear income tax schedules, " *Games and Economic Behavior*, 101: 172–188.

[7] Chetty, Raj. 2012. "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply, *Econometrica, Econometric Society,* vol. 80(3), pages 969-1018.

[8] Choné, Philippe, and Guy Laroque. 2010. "Negative marginal tax rates and heterogeneity." *American Economic Review* 100(5): 2532–2547.

[9] DeNavas-Walt, Carmen and Bernadette D. Proctor, U.S. Census Bureau, Current Population Reports, P60-252, *Income and Poverty in the United States: 2014*, U.S. Government Printing Office, Washington, DC, 2015.

[10] Diamond, Peter. (1980). "Income taxation with fixed hours of work," *Journal of Public Economics* 13(1), 101–110.

[11] Diamond A. 1998, "Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. " *The American Economic Review*, 88(1):83-95.

[12] Feenberg Daniel and Elisabeth Coutts 1993, "An Introduction to the TAXSIM Model", *Journal of Policy Analysis and Management*, Vol. 12, No.1, pp.189–194

[13] Fleurbaey M. 2008, *Fairness, responsibility, and welfare*, Oxford: Oxford University Press.

[14] Fleurbaey M. and F. Maniquet 2005, "Fair social orderings with unequal production skills", *Social Choice and Welfare* 24: 1–35.

[15] Fleurbaey, Marc, and François Maniquet. 2006. "Fair Income Tax." *Review of Economic Studies*, 73: 55–83.

[16] Fleurbaey, Marc, and François Maniquet. 2007. "Help the Low-Skilled, of Let the Hardworking Thrive? A Study of Fairness in Optimal Income Taxation." *Journal of Public Economic Theory*, 9: 467–500.

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

[18] Fleurbaey M. and F. Maniquet (2018), Optimal income taxation theory and principles of fairness, *Journal of Economic Literature*, forthcoming.

[19] Fleurbaey M. and A. Trannoy 2003, "The impossibility of a Paretian Egalitarian", *Social Choice and Welfare* 21: 243–464.

[20] Flood S., King M., Rodgers R., Ruggles S., and J.R. Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: IPUMS, 2018. https://doi.org/10.18128/D030.V6.0

[21] Gaertner, W. and E. Schokkaert 2012, *Empirical Social Choice*, Cambridge: Cambridge University Press.

[22] Kanbur, R. M. Keen and M. Tuomala 1994a, "Optimal non-linear income taxation for the alleviation of income-poverty," *European Economic Review* 38: 1613-1632.

[23] Kanbur, R. M. Keen and M. Tuomala 1994b, "Labor supply and targeting in poverty alleviation programs.," *The World Bank Economic Review* 8, 191–211.

[24] Konow, J. 2001, "Fair and square: the four sides of distributive justice," *Journal of Economic Behavior and Organization* 46, 137–164.

[25] Lockwood, Benjamin B. (Fothcoming), "Optimal income taxation with present bias.", *American Economic Journal: Economic Policy.*

[26] Lockwood, Benjamin B., and Matthew Weinzierl. 2015. "De Gustibus Non Est Taxandum: Heterogeneity in Preferences and Optimal Redistribution." *Journal of Public Economics*, 124: 74–80.

[27] Lockwood, Benjamin B., and Matthew Weinzierl. 2016. "Positive and normative judgments implicit in U.S. tax policy, and the costs of unequal growth and recessions.", *Journal of Monetary Economics* 77(2016):30-47.

[28] Maniquet, Francois and Neumann, Dirk. (Forthcoming) "Well-Being, Poverty and Labor Income Taxation: Theory and Application to Europe and the U.S.", *American Economic Journal: Microeconomics.*

[29] Mirrlees, James A. 1971. "An exploration in the theory of optimum income taxation." *The Review of Economic Studies* 38(2) 175–208.

[30] Mirrlees, James A. 1976. "Optimal Tax Theory: A Synthesis." *Journal of Public Economics*, 6: 327–358.

[31] Piketty, T. and E. Saez 2013, "Optimal Labor Income Taxation." In *Handbook of Public Economics*, Volume 5, edited by Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, 391–474, Amsterdam: Elsevier, North Holland.

[32] Roemer, John E. 1998, *Equality of Opportunity*, Cambridge: Harvard University Press.

[33] Roemer, John E., Rolf Aaberge, Ugo Colombino, Johan Fritzell, Stephen P. Jenkins, Arnaud Lefranc, Ive Marx, Marianne Page, Evert Pommer, Javier Ruiz-Castillo, Maria Jesus San Segundo, Torben Tranaes, Alain Trannoy, Gert G. Wagner, Ignacio Zubiri 2003, "To what extent do fiscal regimes equalize opportunities for income acquisition among citizens?" *Journal of Public Economics*, 87(3–4) 539–565.

[34] Saez, Emmanuel. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses?, *The Quarterly Journal of Economics* 117(3), 1039–1073.

[35] Saez, E. and S. Stantcheva 2016, "Generalized social marginal welfare weights for optimal tax theory," *American Economic Review* 106, 24–45.

[36] Samuelson, Paul A. 1974 "Complementarity – An Essay on the 40th anniversary of the Hicks-Allen revolution in demand theory," Journal of Economic Literature 12, 1255–1289.

[37] Schokkaeert, E. and K. Devooght 2003, "Responsibility-sensitive fair compensation in different cultures," *Social Choice and Welfare* 21, 207–242.

[38] Wane, W. 2001, "The optimal income tax when poverty is a public bad," *Journal of Public Economics* 82: 271-299.

[39] Weinzierl, Matthew. 2014. "The Promise of Positive Optimal Taxation: Normative Diversity and a Role for Equal Sacrifice." *Journal of Public Economics*, 118: 128–142.

Main Tables

Table 1: Calibrated Optimal Tax Schedule

$\overline{P}_0 = \$15,379$	$c(0)$ (\$)	$y(n_u)$ (\$)	y_1^* (\$)	$y(n_l)$ (\$)	$T'(y_1^* - \varepsilon)$	$T'(y_1^* + \varepsilon)$	$T'(y(n_l))$	$T'(y \to \infty)$
Most Red.	12,778	13,326	16,015	19,228	−0.34	0.63	0.53	0.49
Quasi-linear	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	0.49
Least Red.	8,632	14,951	16,859	19,022	−0.21	0.48	0.39	0.39
$\overline{P}_0 = \$18,850$								
Most Red.	12,025	13,711	16,223	19,187	−0.30	0.59	0.50	0.49
Quasi-linear	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	0.49
Least Red.	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	0.47
$\overline{P}_0 \geq \$20,036$								
All	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	0.49

Note: Each panel presents the calibrated optimal tax according to each choice of $\tilde{R}$ (most redistributive, quasi-linear and least redistributive) for poverty lines at, respectively, \$15,379, \$18,850, and \$24,230. These values correspond to weighted averages of the 2015 official US poverty lines for households of size 2, 3, and 4. The average slope of the poverty line is calibrated at $\$6/h$. The second column, $c(0)$, is the level of consumption for agents with zero income. The fourth column, y_1^*, is the income level at which there is a discrete jump in marginal tax rates. The third, fourth, and fifth columns indicate the endpoints of the intervals over which rates are decreasing. The sixth and seventh column report the marginal tax rates just before and right after y_1^*. The tax rate on incomes larger than $y(n_l)$ is reported in column eight. The last column indicates the marginal tax rate in the tail.

Main Figures

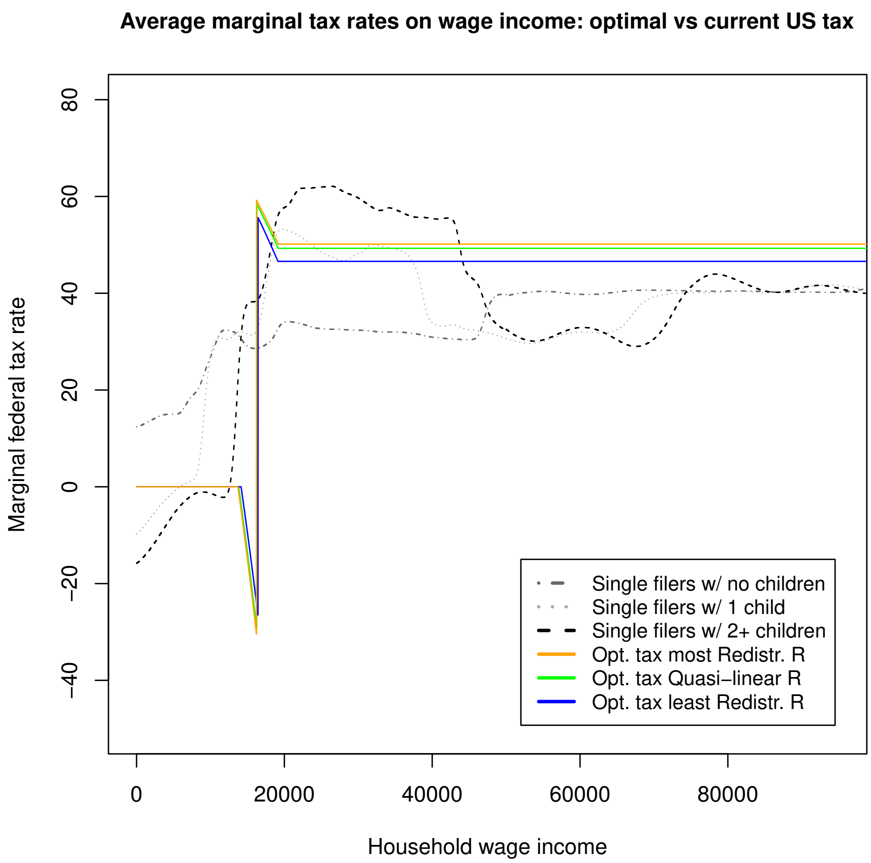

Figure 1: The solid lines depict the optimal labor income tax according to three versions of our social objective when the poverty line is set at \$18,850 with an average slope of \$6/h. The dashed and dotted lines depict the actual marginal tax rates in the US for different household structures. The current average marginal tax rate includes both the actual federal rate and the implicit tax rate from the phaseout of various welfare programs. Details on the calibration and estimation of the current tax system using CPS data can be found in Appendix F.

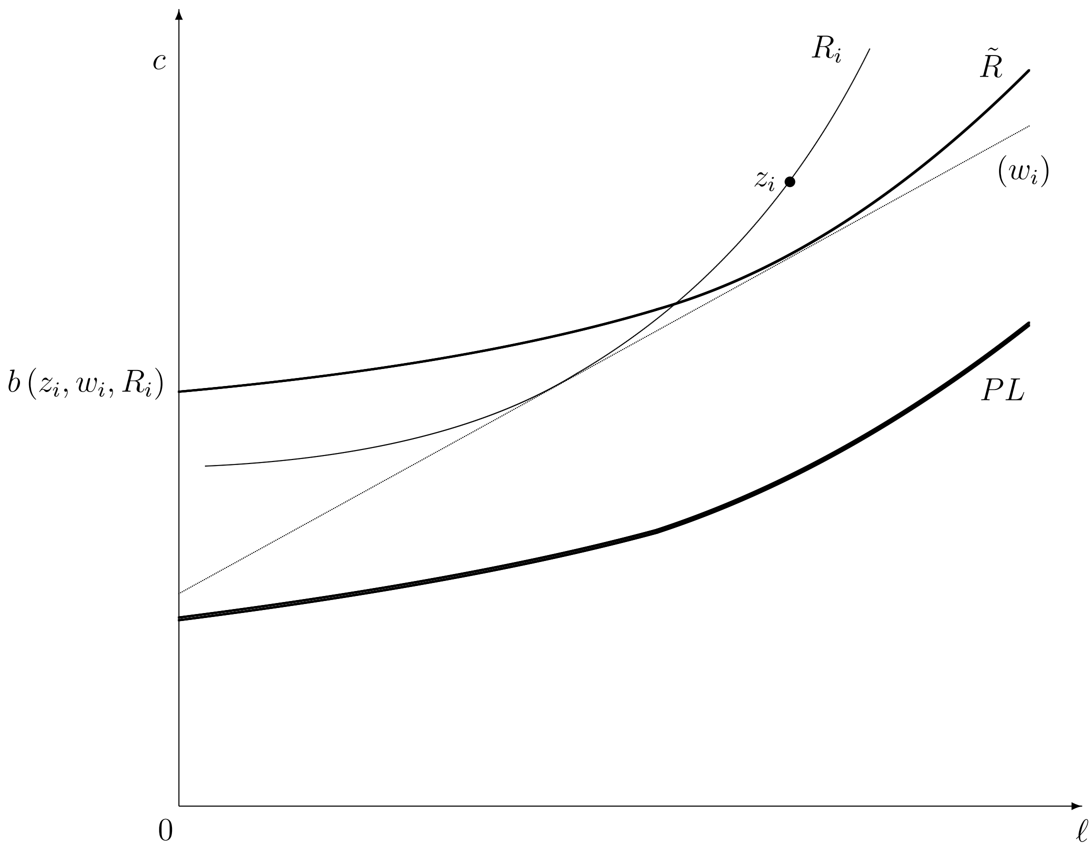

Figure 2: Illustration of the SWF $\mathbf{R}^{\tilde{R}lex}$. Agent i consumes bundle z_i. Her implicit budget is the budget of slope w_i that is tangent to her indifference curve going through z_i. The reference preferences $\tilde{R}$ contain PL as an indifference curve. The utility associated with the best bundle for $\tilde{R}$ in i's implicit budget gives the well-being index for i: $b\,(z_i, w_i, R_i)$. Finally, $\mathbf{R}^{\tilde{R}lex}$ applies the leximin to vectors of $b(.)$.

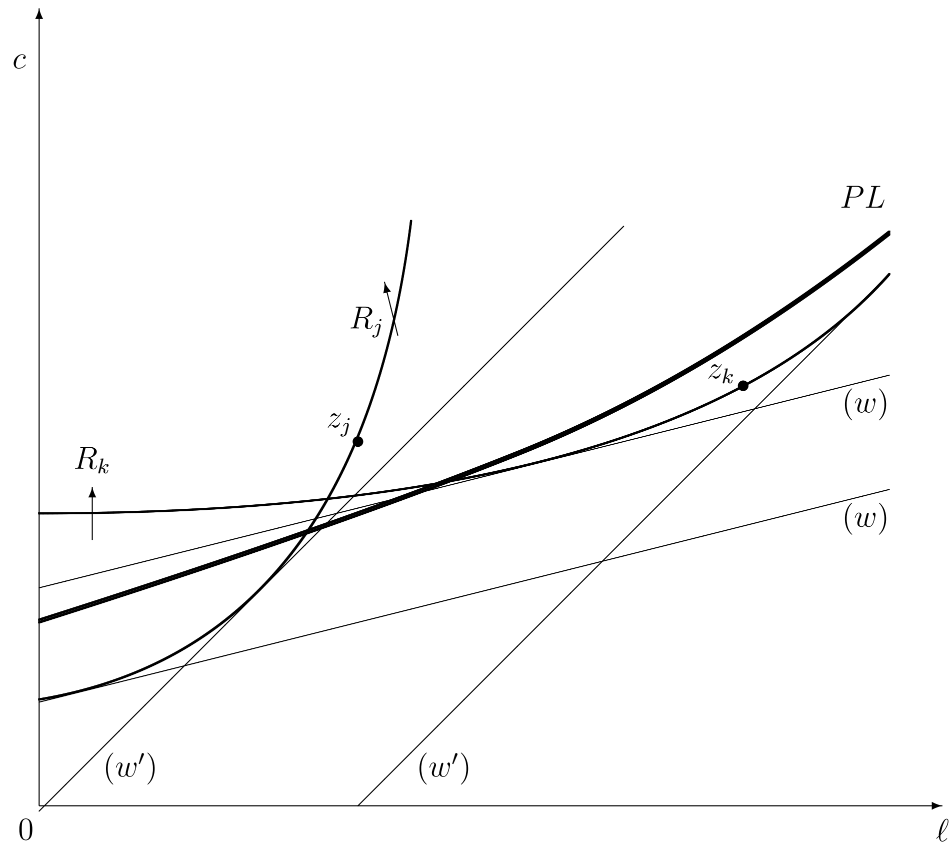

Figure 3: Who is (non-)poor? Agent j consumes bundle z_j above the poverty line. She is poor when her wage is $w_j = w$ but non-poor when $w_j = w'$. Agent k consumes bundle z_k below the poverty line. She is poor when her wage $w_k = w'$ but non-poor when $w_k = w$. This figure illustrates two important points. First, agents may consume a bundle under the poverty line, yet be non-poor according to our well-being index. Second, well-being indices depend on wages.

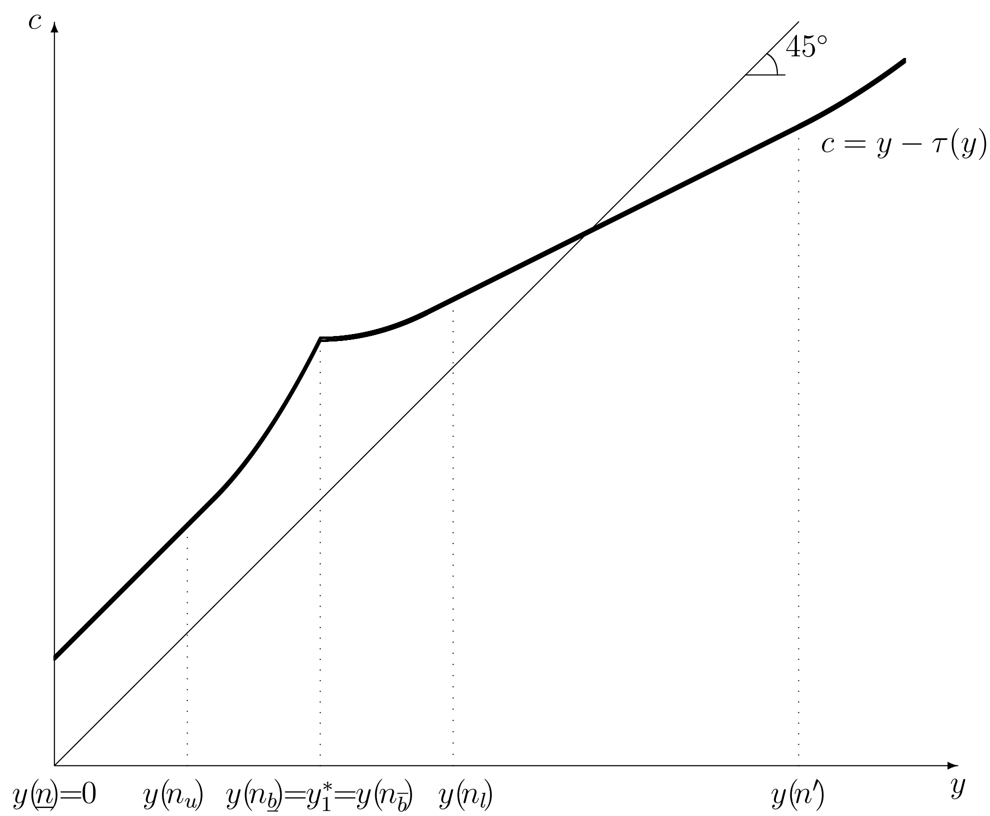

Figure 4: Illustration of Proposition 1. The income range is divided in 5 intervals. Marginal tax rates are zero in the first interval $[y(\underline{n}), y(n_u)]$, negative and decreasing in the second one $[y(n_u), y(n_{\underline{b}})]$, positive and decreasing in the third one $[y(n_{\overline{b}}), y(n_l)]$, constant in the fourth one $[y(n_l), y(n')]$, and decreasing in the last one $[y(n'), \infty]$.

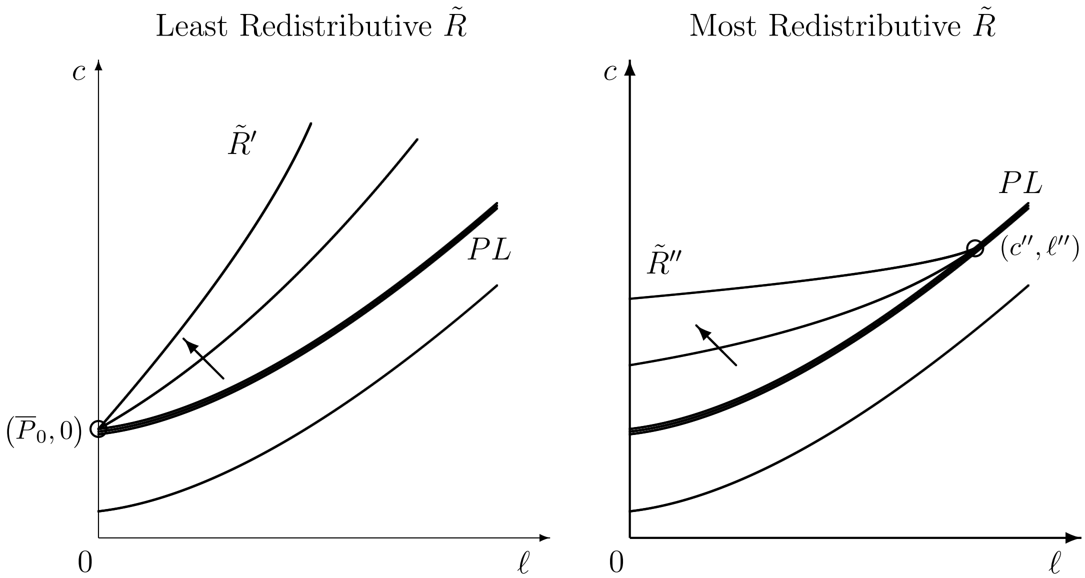

Figure 5: Illustration of the least- and most redistributive reference preferences. The left panel shows some indifference curves of the least redistributive $\tilde{R}$. Above the poverty line, the curves rotate around $(\overline{P}_0, 0)$. The right panel shows some indifference curves of the most redistributive $\tilde{R}$. Above the poverty line, the portion of each curve over $[0, \ell'')$ rotate around (c'', ℓ''). There is a kink at (c'', ℓ'') and the remaining portion (over $[\ell'', \infty)$) follows PL. In both cases, $\tilde{R}$ are quasi-linear under PL.

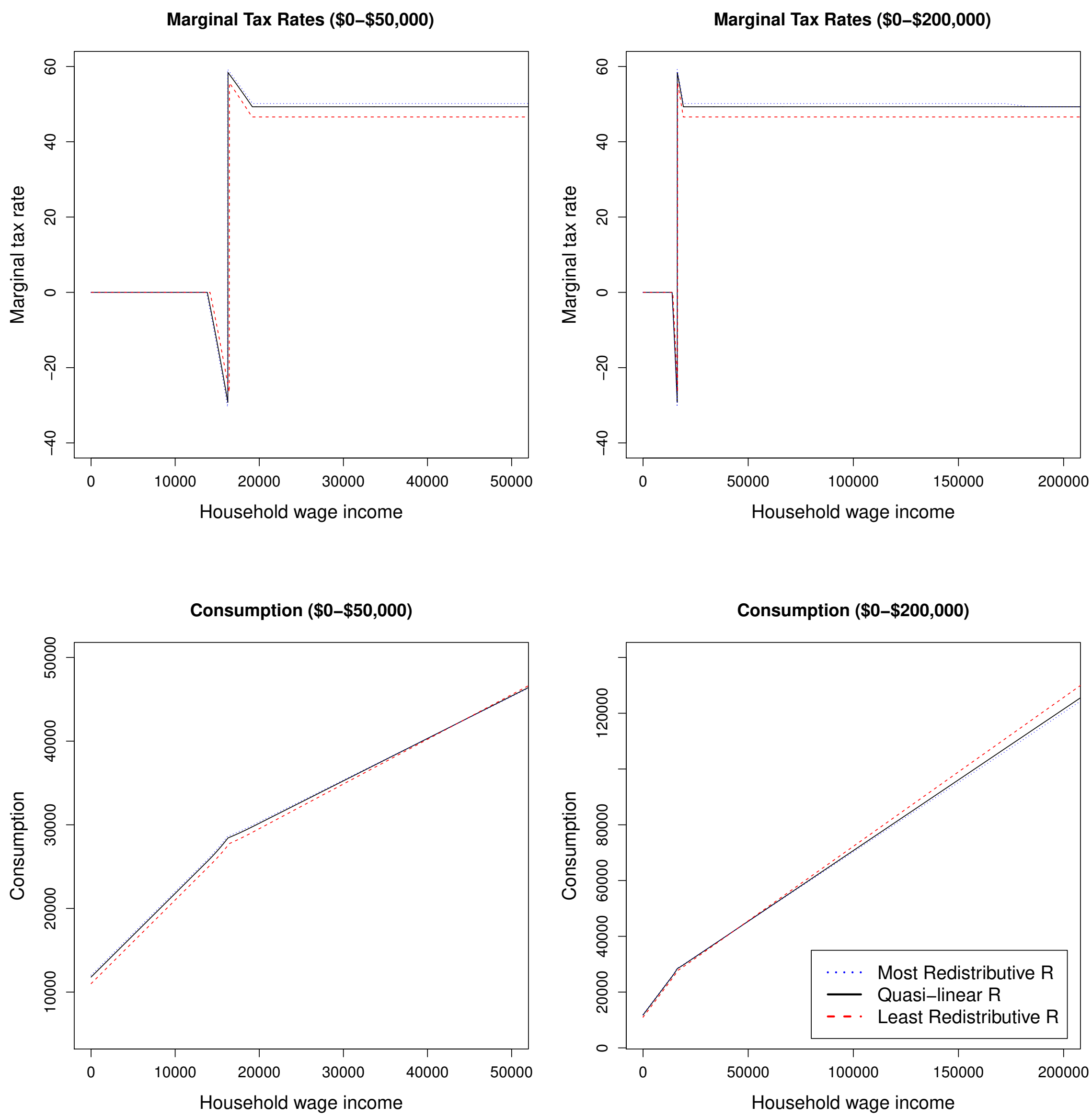

Figure 6: Comparing optimal tax rates and consumption under the 3 polar cases ($PL = \$18,850$; Slope $= \$6/h$). The overall amount of redistribution is similar across the 3 versions of the SWF. Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn).

Appendix for Publication

A Axioms of poverty reduction

Axiom 4 *poverty-reduction transfer 1*
For all $E = (w_N, R_N) \in \mathcal{E}$, *all* $\Delta \in \mathbb{R}_+$, *all* $j, k \in N$, *all* $z_N = (l_N, c_N)$, $z'_N = (l'_N, c'_N) \in X^N$ *such that* $l_j = l'_j = l_k = l'_k$, $c'_j = c_j - \Delta$, $c'_k = c_k + \Delta$ *and for all* $i \neq j, k : z_i = z'_i$,

$$[z'_j > PL > z'_k] \Rightarrow [z'_N \mathbf{R}(E) z_N].$$

Axiom 5 *poverty-reduction transfer 2*
For all $E = (w_N, R_N) \in \mathcal{E}$, *all* $\Delta \in \mathbb{R}_+$, *all* $j, k \in N$, *all* $z_N = (l_N, c_N)$, $z'_N = (l'_N, c'_N) \in X^N$ *such that* $l_j = l'_j = l_k = l'_k$, $c'_j = c_j - \Delta$, $c'_k = c_k + \Delta$ *and for all* $i \neq j, k : z_i = z'_i$,

$$[I(z'_j, R_j) > PL > z'_k] \Rightarrow [z'_N \mathbf{R}(E) z_N].$$

These three axioms are illustrated in Figure 7. It should be transparent that the first version of the axiom is logically stronger than the second one, which is logically stronger than Axiom 3 defined in Section 3 (henceforth referred to as *poverty-reduction transfer 3*). Moreover, we have the following (in)compatibilities with the other axioms. To state these (in)compatibilities, we recall a well-known implication of strong Pareto, Pareto indifference. It is the requisite that if all agents are indifferent between two allocations, then social preferences must also be indifferent.

Axiom 6 *Pareto indifference*
For all $E = (w_N, R_N) \in \mathcal{E}$, *and* $z_N = (l_N, c_N), z'_N = (l'_N, c'_N) \in X^N$, *if* $z_i I_i z'_i$ *for all* $i \in N$, *then* $z_N \mathbf{I}(E) z'_N$

Proposition 2
1. *No SWF satisfies* Pareto indifference *and* poverty-reduction transfer 1*.*
2. *There are SWFs satisfying* Pareto indifference *and* poverty-reduction transfer 2*.*

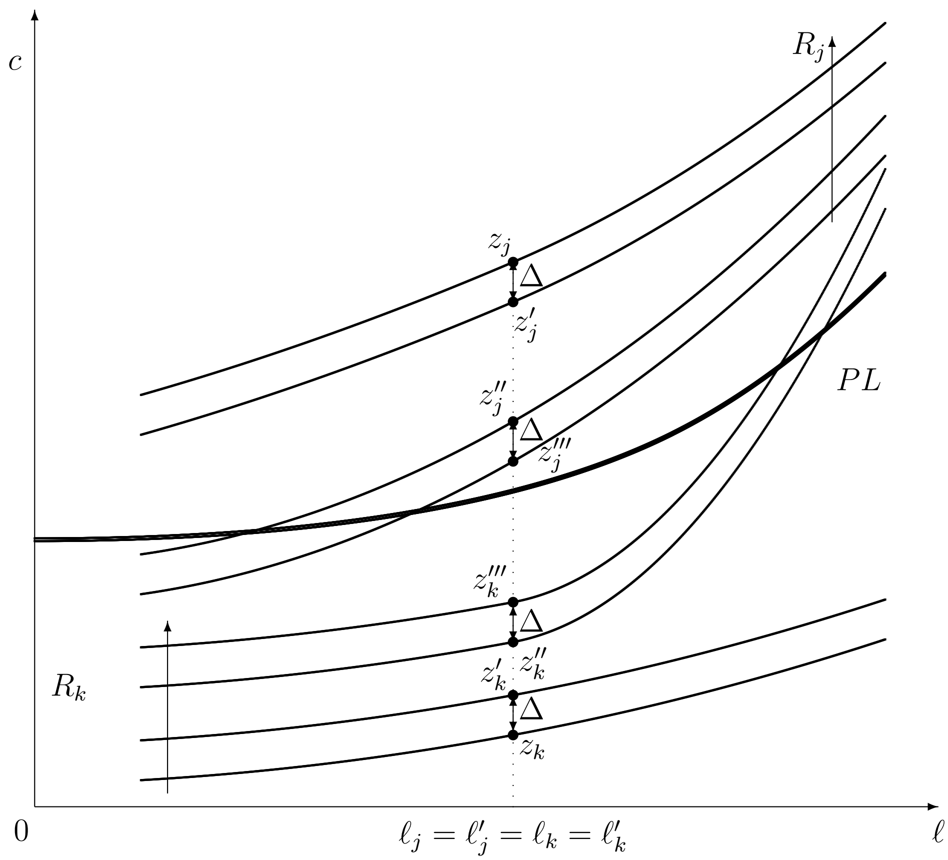

Figure 7: Illustration of the axioms of *poverty-reduction transfer* 1, 2 and 3: according to *poverty-reduction transfer* 1, any transfer of Δ from either z_j or z''_j to either z_k or z''_k is a strict social improvement; according to *poverty-reduction transfer* 2, any transfer of Δ from z_j to either z_k or z''_k is a strict social improvement; according to *poverty-reduction transfer* 3, only a transfer of Δ from z_j to z_k is a strict social improvement.

3. *No SWF satisfies* Pareto indifference*,* equal-wage transfer *and* poverty-reduction transfer 2*.*

4. *There are SWFs satisfying* strong Pareto*,* equal-wage transfer *and* poverty-reduction transfer 3*.*

The proof of the four statements of Proposition 2 is rather intuitive. We present it in the text. Statement 1 is reminiscent of the proof of the impossibility of a Paretian egalitarian in Fleurbaey and Trannoy (2003). It is illustrated in Figure 8. The key feature of the figure is that both agents j and k have preferences that cross the poverty line, but in opposite direction. As a result, when they both choose a low labor time, say $\ell_j = \ell_k$, then agent j has a consumption level above the poverty line

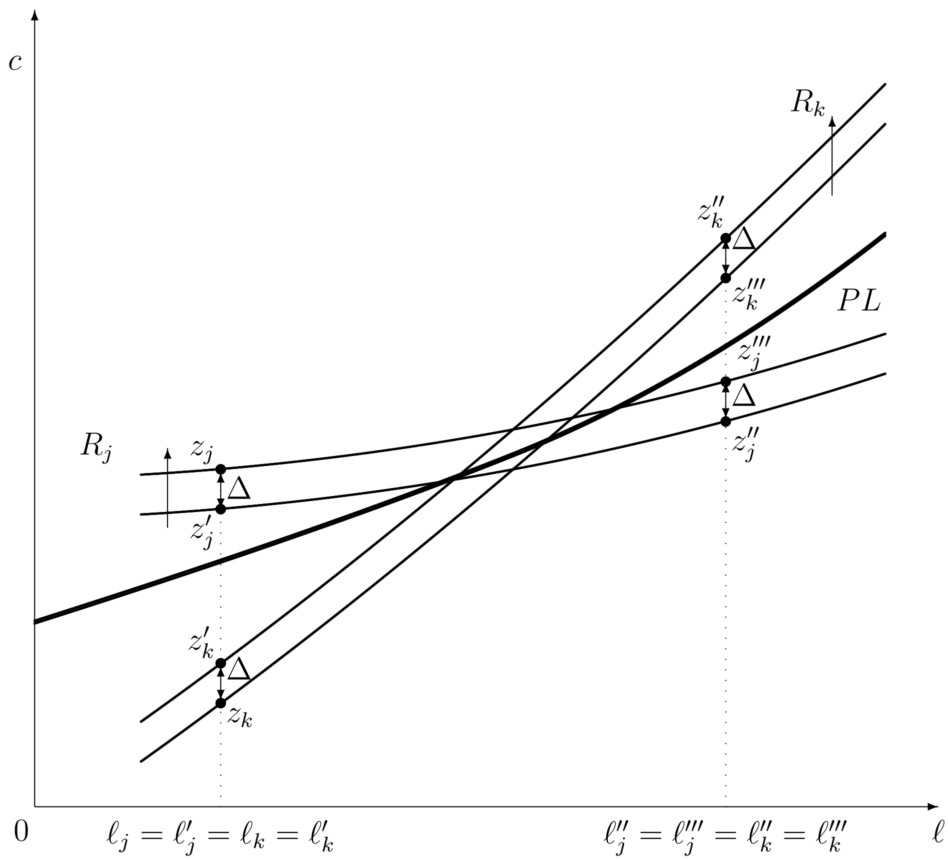

Figure 8: Illustration of Proposition 2, statement 1. According to *poverty-reduction transfer 1*, (z_j', z_k') is socially preferred to (z_j, z_k) and (z_j''', z_k''') is preferred to (z_j'', z_k''). By *Pareto indifference*, (z_j, z_k) is socially indifferent to (z_j''', z_k''') and (z_j', z_k') is indifferent to (z_j'', z_k''). This leads to an intransitivity showing that no SWF satisfies *Pareto indifference* and *poverty-reduction transfer 1*.

whereas k's consumption is below the poverty line. As a result, (z_j', z_k') is socially preferred to (z_j, z_k). When they both choose a large labor time, say $\ell_j'' = \ell_k''$, at the same satisfaction levels, k's consumption level is above the poverty line contrary to j's consumption level. As a result, (z_j''', z_k''') is socially preferred to (z_j'', z_k''). By *Pareto indifference*, which is implied by *strong Pareto*, (z_j, z_k) is socially indifferent to (z_j''', z_k''') and (z_j', z_k') is socially indifferent to (z_j'', z_k''), creating an intransitivity.

The proof of statement 3 is illustrated in Figure 9. *Poverty-reduction transfer 2* requires that (z_j', z_k') be socially preferred to (z_j, z_k). *Equal-wage transfer* requires that (z_j''', z_k) be socially preferred to (z_j'', z_k'). Finally, *Pareto indifference* requires that (z_j, z_k) be socially indifferent to (z_j''', z_k) and (z_j', z_k') to (z_j'', z_k'), creating an intransitivity.

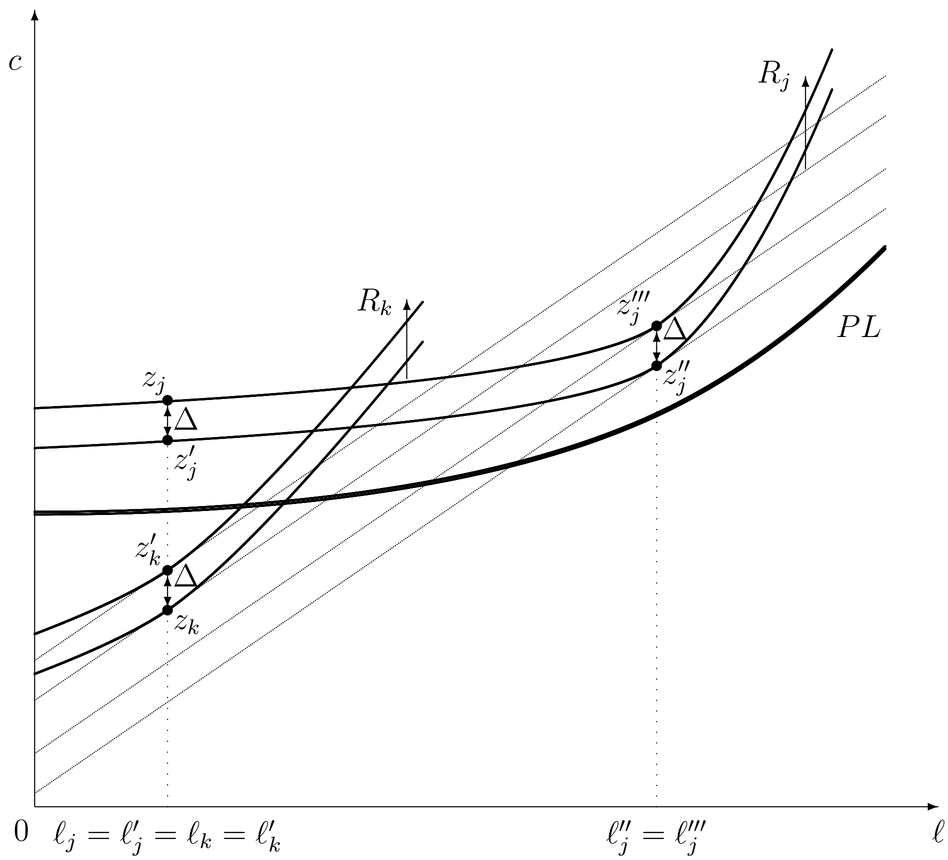

Figure 9: Illustration of Proposition 2, statement 3. By *poverty-reduction transfer 2*, $\left(z_j', z_k'\right)$ is socially preferred to (z_j, z_k). By *Pareto indifference*, (z_j, z_k) is socially indifferent to $\left(z_j''', z_k\right)$ and $\left(z_j', z_k'\right)$ is indifferent to $\left(z_j'', z_k'\right)$. Finally, by *equal-wage transfer*, $\left(z_j''', z_k\right)$ is socially preferred to $\left(z_j'', z_k'\right)$. This leads to an intransitivity showing that no SWF satisfies *Pareto indifference*, *equal-wage transfer* and *poverty-reduction transfer 2*.

Statement 3 is at the heart of the ethical dilemma that this paper addresses. Agents j and k have the same wage, represented in the figure by the slope of their budgets. The responsibility goal pushes towards no redistribution among them two. However, agent k enjoys a bundle below the poverty line, whereas agent j's indifference curve is everywhere above the poverty line. This is why even the weak *poverty-reduction transfer 2* requires to prefer a redistribution between the two agents when they don't work hard, that is, when their labor times are $\ell_j = \ell_k$. The two transfers go in opposite directions, whereas agent j is indifferent between the two involved bundles, hence the incompatibility.[16]

[16]Observe that this incompatibility does not depend on our assumption that the poverty line is strictly increasing. Even a flat poverty line would display the same basic problem.

Statements 2 and 4 are proven by way of examples. Statement 4 is proven in the main text in Section 3 ($\mathbf{R}^{\tilde{R}lex}$ satisfies *strong Pareto*, *equal-wage transfer* and *poverty-reduction transfer 3*). We now prove statement 2. We first need to introduce some terminology. Let $\mathcal{U}$ denote the set of utility functions representing preferences in $\mathcal{R}$. This example works by choosing a utility representation for each individual preference relation and apply the leximin criterion to the utility vector corresponding to the evaluated allocations. Let $U : \mathcal{R} \to \mathcal{U}$ be a representation function satisfying the property that for each $R \in \mathcal{R}$, $U(R)$ is a utility function representing R and if we write $u = U(R)$ then for all $z \in X$

$$u(z) = 1 \Leftrightarrow z \, I \, m(R, PL), \tag{6}$$

that is, all agents have the same utility level at their preferred bundle on the poverty line (and this utility level is equal to 1). Social preferences R^U are defined as follows: for all $E = (w_N, R_N) \in \mathcal{E}$, all $z_N, z'_N \in X^N$,

$$z_N \, R^U \, z'_N \Leftrightarrow (u_i(z_i))_{i \in N} \geq_{lex} (u_i(z'_i))_{i \in N}$$

where for all $i \in N$, $u_i = U(R_i)$. The fact that R^U satisfies *strong Pareto* follows from u_i being a utility representation of R_i and the leximin criterion being strictly increasing in all its arguments. The fact that R^U satisfies *poverty-reduction transfer 2* follows Eq. (6) satisfied by the utility representation. Indeed, as long as $I(z'_j, R_j) > PL > z'_k$, we have $u_j(z'_j) > 1 > u_k(z'_k)$, so that the leximin criterion gives priority to agent k, with the consequence that z'_N is strictly preferred to z_N.

B Proof of Proposition 3 on SWF's

We begin by defining *separability* formally. Assume the social preference relation is set between two allocations, z_N and z'_N, whereas each agent in a subset M of the population is allocated exactly the same bundle in both allocations. The requirement is that agents in M should not matter in the social welfare function, which is captured by the requirement that the social preference relation should remain unaffected if the preferences of agents in M change and if their bundles change in such a way that they are still exactly the same in the two resulting allocations.

Axiom 7 *separability*
For all $E=(w_N,R_N), E'=(w'_N,R'_N)\in\mathcal{E}$, *all* z_N, z'_N, z''_N *and* $z'''_N\in X^N$, *if* $w_i=w'_i$ *for all* $i\in N$ *and there exists* $M\subset N$ *such that*

$$\begin{aligned}\forall i\in M: z_i=z'_i \quad \textit{and} \quad z''_i=z'''_i,\\ \forall i\in N\setminus M: R_i=R'_i,\\ \forall i\in N\setminus M: z_i=z''_i \quad \textit{and} \quad z'_i=z'''_i,\end{aligned}$$

then

$$z_N\mathbf{R}(E)z'_N\Leftrightarrow z''_N\mathbf{R}(E')z'''_N.$$

We need to introduce the following terminology, which corresponds to the implicit budget terminology of Samuelson (1974). The equivalent income is the intercept of an implicit budget. Formally, for (z,w,R), we let $t(z,w,R)$ denote the intercept of $IB(z,w,R)$, that is,

$$t\left(z,w,R_i\right)=\min\left\{t\in\mathbb{R}\mid\exists\left(\ell,c\right)\in X,\ c=t+w\ell,\ \left(\ell,c\right)R\,z\right\}.$$

With the same abuse of notation as above, we define

$$t\left(w,PL\right)=\min\left\{t\in\mathbb{R}\mid\exists\left(\ell,c\right)\in PL,\ c=t+w\ell\right\}.$$

Slightly abusing on notation, we write $IB\left(w,PL\right)$ to denote the implicit budget of slope w that is tangent from below to the poverty line.

Proposition 3 *If a SWF satisfies* strong Pareto, equal-wage transfer, poverty-reduction transfer *and* separability, *then for all* $E=(w_N,R_N)\in\mathcal{E}$ *such that* $\mid N\mid\geq 4$, *all* $z_N, z'_N\in X^N$, *if there exist* $j,k\in N$ *such that*

- $IB\left(z_j,w_j,R_j\right)\supset IB\left(z'_j,w_j,R_j\right)\supset IB\left(w_j,PL\right)$,
- $IB\left(w_k,PL\right)\supset IB\left(z'_k,w_k,R_k\right)\supset IB\left(z_k,w_k,R_k\right)$,

and for all $i\neq j,k$: $z_i=z'_i$, *then* $z'_N\,\mathbf{P}(E)\,z_N$.

Proof: In the proof, we make use of *Pareto indifference*, previously defined as axiom 6. It is a well-known implication of strong Pareto. We prove the theorem in three steps. The first step is a preliminary result about the existence of specific preference relations. Take any two $w, w' \in [\underline{w}, \infty)$, and any $t, t', t'', t''' \in \mathbb{R}$ such that

- $t < t' < t(w, PL)$,
- $t(w', PL) < t'' < t'''$.

There necessarily exists $R^\circ \in \mathcal{R}$ and $z = (\ell, c), z' = (\ell', c'), z'' = (\ell'', c''), z''' = (\ell''', c''') \in X$ such that

- $\ell = \ell' = \ell'' = \ell'''$,
- $c + c''' = c' + c''$,
- $t(z, w, R^\circ) = t$,
- $t(z', w', R^\circ) = t'$,
- $t(z'', w', R^\circ) = t''$,
- $t(z''', w', R^\circ) = t'''$.

Instead of proving the claim algebraically, we illustrate its simple intuition in the following figure. The distance between the two budgets of slope $\tilde{w}'$ is much larger than that between the budgets of slope $\tilde{w}$. In spite of that, it is possible to construct indifference curves that are tangent to the former budgets and yet become arbitrarily close to each other.

We now turn to the second step. We prove the theorem in a specific case. Let $E = (w_N, R_N) \in \mathcal{E}$, $z_N, z'_N \in X^N$, and $j, k \in N$ satisfy the conditions of the statement. Assume, moreover, that there exist $t_j > t(w_j, PL)$ and $t_k < t(w_k, PL)$ such that

$$\begin{aligned} t\left(z'_j, w_j, R_j\right) - t_j &= t(z_j, w_j, R_j) - t\left(z'_j, w_j, R_j\right), \\ t_k - t(z'_k, w_k, R_k) &= t(z'_k, w_k, R_k) - t(z_k, w_k, R_k). \end{aligned}$$

As $\mid N \mid \geq 4$, we assume $\mid N \mid = 4$ to save on notation. Let $N = \{j, k, a, b\}$. We need to prove that $(z'_j, z'_k, z_a, z_b)\, \mathbf{P}(E)\, (z_j, z_k, z_a, z_b)$. By *Pareto indifference*, we can do as if z_j and z'_j (resp. z_k and z'_k) were the best bundles of j (resp. k) in the corresponding implicit budget. If it is not the case, indeed, then we can identify those best bundles and replace z_j, z'_j, z_k and z'_k by those bundles. Let $R^\circ \in \mathcal{R}$ and $z = (\ell, c), z' = (\ell', c'), z'' = (\ell'', c''), z''' = (\ell''', c''') \in X$ be such that

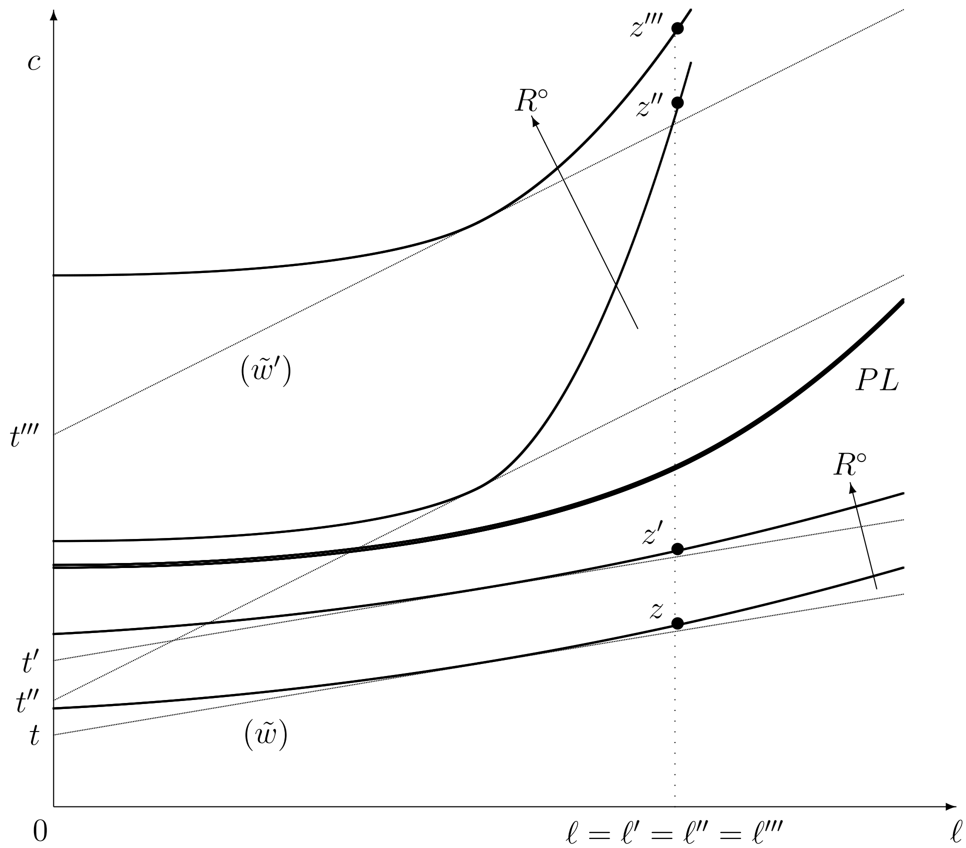

Figure 10: Illustration of the existence of R°. Four indifference curves of R° are represented. The two indifference curves above the poverty line are such that despite the distance between implicit budgets at wage $\tilde{w}'$ being large, they are very close to each other at $\ell = \ell' = \ell'' = \ell'''$. In fact, $c''' - c'' = c' - c$. For any given distance between budgets of the same slope, there always exists such pair of indifference curves that are tangent to both budgets, yet become arbitrarily close to each other. The existence of such preferences is key in proving proposition 3.

- $\ell = \ell' = \ell'' = \ell'''$,
- $c + c''' = c' + c''$,
- $t\left(z, w_k, R^\circ\right) = t\left(z'_k, w_k, R_k\right)$,
- $t\left(z', w_k, R^\circ\right) = t_k$,
- $t\left(z'', w_j, R^\circ\right) = t_j$,
- $t\left(z''', w_j, R^\circ\right) = t\left(z'_j, w_j, R_j\right)$.

Let $E' \in \mathcal{E}$ be defined by

$$E' = \left((w_j, w_k, w_j, w_k), (R_j, R_k, R^\circ, R^\circ) \right).$$

By *equal-wage transfer* applied twice, first between agents j and a and then between agents k and b,

$$(z'_j, z'_k, z''', z) \, \mathbf{P}(E') \, (z_j, z_k, z'', z').$$

By *poverty-reduction* applied between agents a and b,

$$(z'_j, z'_k, z'', z') \, \mathbf{P}(E') \, (z'_j, z'_k, z''', z).$$

By transitivity,

$$(z'_j, z'_k, z'', z') \, \mathbf{P}(E') \, (z_j, z_k, z'', z').$$

By *separability*,

$$(z'_j, z'_k, z_a, z_b) \, \mathbf{P}(E) \, (z_j, z_k, z_a, z_b),$$

the desired outcome. Of course, it is unlikely that such t_j and t_k exist, as $t\left(z'_j, w_j, R_j\right)$ can be arbitrarily close to $t\left(w_j, PL\right)$ and $t\left(z'_k, w_k, R_k\right)$ arbitrarily close to $t\left(w_k, PL\right)$.

As a third and final step, we deal with the general case. We can find $t_j > t\left(w_j, PL\right)$ and $t_k < t\left(w_k, PL\right)$ and an integer n such that

$$\begin{aligned} t\left(z'_j, w_j, R_j\right) - t_j &= \frac{1}{n}\left(t\left(z_j, w_j, R_j\right) - t\left(z'_j, w_j, R_j\right)\right), \\ t_k - t\left(z'_k, w_k, R_k\right) &= \frac{1}{n}\left(t\left(z'_k, w_k, R_k\right) - t\left(z_k, w_k, R_k\right)\right). \end{aligned}$$

Then, we can define two series of $n-1$ bundles $z_j^1, \ldots, z_j^{n-1}$ and $z_k^1, \ldots, z_k^{n-1}$ such that for all $p \in \{1, \ldots, n-1\}$,

$$t\left(z_j^p, w_j, R_j\right) = t\left(z_j, w_j, R_j\right) - p\left(t\left(z'_j, w_j, R_j\right) - t_j\right), \tag{7}$$

$$t\left(z_k^p, w_k, R_k\right) = t\left(z_k, w_k, R_k\right) + p\left(t_k - t\left(z'_k, w_k, R_k\right)\right). \tag{8}$$

For all $p \in \{1, \ldots, n\}$, we can apply step 2 and prove that

$$(z_j^p, z_k^p, z_a, z_b) \, \mathbf{P}(E) \, (z_j^{p-1}, z_k^{p-1}, z_a, z_b),$$

with the obvious convention that $z_j^0 = z_j, z_j^n = z_j'$, and $z_k^0 = z_k, z_k^n = z_k'$. The proof then goes on by transitivity applied to those n steps, to reach the final outcome that

$$(z_j', z_k', z_a, z_b)\, \mathbf{P}(E)\, (z_j, z_k, z_a, z_b).$$

□

C Proof of Corollary 1 on the planner's objective

The following corollary proves that the objective of any planner interested in maximizing $\mathcal{R}^{\tilde{R}lex}$ should be to lift all agents out of poverty in the sense that the implicit budget of no agent should lie below the poverty line.

Corollary 1 *If a SWF satisfies* strong Pareto*,* equal-wage transfer*,* poverty-reduction transfer *and* separability*, then for all $E = (w_N, R_N) \in \mathcal{E}$ such that $\mid N \mid \geq 4$, all $z_N, z_N' \in X^N$, if there exists $j \in N$ such that*

$$IB\,(w_j, PL) \supset IB\,(z_j, w_j, R_j)$$

whereas for all $i \in N$

$$IB\,(z_i', w_i, R_i) \supset IB\,(w_i, PL)\,,$$

then $z_N'\, \mathbf{P}(E)\, z_N$.

Proof. Let $E = (w_N, R_N) \in \mathcal{E}$, $z_N, z_N' \in X^N$, and $j \in N$ satisfy the conditions of the corollary. Let us assume, contrary to the claim, that $z_N\, \mathbf{R}(E)\, z_N'$. Let $M \subset N$ be defined by

$$M = \{i \in N | i \neq j, z_i\, P_i\, z_i'\}.$$

If $M = \emptyset$, then, by *strong Pareto*, we have reached a contradiction. Let us then assume that $M \neq \emptyset$. Let $z_N'' \in X^N$ be such that

$$\begin{array}{rl} \forall i \in M, & IB(w_i, z_i, R_i), IB(w_i, z_i', R_i) \supset IB(w_i, z_i'', R_i) \supset IB(w_i, PL), \\ \forall i \notin M, i \neq j & z_i'' = z_i, \\ & IB(w_j, PL) \supset IB(w_j, z_j'', R_j) \supset IB(w_j, z_j, R_j). \end{array}$$

Existence of such a z_N'' is guaranteed by the conditions imposed on z_N and z_N' in the statement of the corollary.

Let $r > 0$ and $z_j^0, z_j^1, \ldots, z_j^m \in X$, $m = |M|$ be such that

$$\begin{aligned} IB(w_j, z_j^0, R_j) &= IB(w_j, z_j, R_j), \\ IB(w_j, z_j^k, R_j) &= IB(w_j, z_j^{k-1}, R_j) + (0, r), \forall k \in \{1, \ldots, m\}, \\ IB(w_j, z_j^m, R_j) &= IB(w_j, z_j'', R_j). \end{aligned}$$

By applying Proposition 3 m times, we can prove that $z_N'' \, \mathbf{P}(E) \, z_N$. Indeed, at each step, z_j^{k-1} is replaced by z_j^k, which means that an agent whose implicit budget is strictly below the poverty line is made better-off, and z_i is replaced with z_i'', which means that an agent strictly above the poverty line is made worse-off, which implies a strict social preference. By transitivity, $z_N'' \, \mathbf{P}(E) \, z_N'$, which contradicts *strong Pareto*, because $z_i' \, R_i \, z_i''$ for all $i \in N$ and the preference is strict for some agents, including j. ■

D Proof of Proposition 1

Overview

We prove Proposition 1 in four steps described below. We first provide a sketch of the proof describing each step and how they relate to each other. A formal statement and derivation of each step follows.

In the first step, we quickly prove a useful property: for each n, those with the lowest wage are the worst-off among all agents with $n_i = n$. Step 1 is useful because, by the leximin property of $\mathcal{R}^{\tilde{R}lex}$, we can focus on agents with $w_i = \underline{w}$ $\forall n \in [\underline{n}, \underline{w}\overline{\theta}]$ and with $w_i = n_i/\overline{\theta}$ $\forall n_i \in [\underline{w}\overline{\theta}, \infty)$. We define the well-being index of the worst-off agents at n_i as $\underline{b}\,(u\,(n_i)\,; n_i)$.

In the second step, we prove that agents with $\theta_i = \overline{\theta}$ and $w_i = \underline{w}$ (the most-hardworking minimum wage agents) are among the worst-off at the second-best optimum. We define $n_1 = \underline{w}\overline{\theta}$. To be clear, step 1 is a comparison within n while step 2 compares across n. It is also important to note that step 2 proves that agents with $\theta_i = \overline{\theta}$ and $w_i = \underline{w}$ are among the worst-off at the optimum but other types $n \neq n_1$ may very well have just as low a well-being index at the solution. The result derived in step 2 is important because it allows to re-write the maximization problem in a more tractable way. The original maximization problem consists in maximizing the minimum well-being across n_i under incentive constraints (3) and the resource constraint (4). Assume that the maximized objective takes value b^*. From step 2, we know that $b^* = \underline{b}(u^*(n_1);n_1)$ where $u^*(n)$ denotes the utility for type n at the solution. We can consider, instead of the original maximization problem, its dual which consists in maximizing a budget surplus under the constraint that $\underline{b}(u(n);n) \geq b^*$ for all n (in addition to incentive compatibility constraints). As $\underline{b}(u(n);n)$ is a utility representation for lowest wage agents within type n, it is strictly increasing in the first argument. As a result, the constraint can be rewritten directly in terms of a lower-bound on utilities:

$$u(n_1) = \underline{u}_1 \tag{9}$$

$$u(n) \geq \underline{u}(n) \quad \forall n \in [\underline{n}, \infty) \tag{10}$$

where $\underline{u}_1 = u^*(n_1)$ and $\underline{u}(n)$ is the utility assignment as a function of n such that $\underline{b}(u(n);n) = b^*$ for all n. Notice that (9) is a consequence of the result proven in step 2. Indeed, from the point of view of the dual problem, we established in that step that the lower-bound on utilities is always binding for agents with $n = n_1$.

In the third step, we set up an Hamiltonian corresponding to the (dual) problem described above and derive the first-order conditions. Constraint (9) provides the initial condition: the state variable $u(n)$ is anchored at n_1. Since typically $\underline{n} < n_1 < \infty$, we divide the problem in two subproblems: the lower sub-problem which consists in solving the optimal trajectory, $\{y^*(n), u^*(n)\}$, on $[\underline{n}, n_1]$ (with initial condition (9) at the upper-end) and the upper sub-problem which consists in solving the optimal trajectory on $[n_1, \infty)$ (with initial condition (9) at the lower-end). The second-order conditions of the incentive compatibility constraints, which are equivalent to $y(n)$ being monotonic everywhere, are binding at n_1 leading to some bunching.[17] The solution within each sub-problem is characterized by intervals over which the lower-bound on utilities (10) is binding and intervals over which it is not. The former intervals are delimited by threshold values of n: $\underline{n}$, n_u, n_l and n'. Consequently, these

[17] Bunching at $y^*(n_1)$ happens on interval $\left[n_{\underline{b}}, n_{\bar{b}}\right]$. It occurs for reasons similar to Brett and Weymark (2017). The optimal tax problem described in this section shares some similarities with the maximization problem they consider with the important difference that we face an additional constraint: the lower-bound on utilities (10).

threshold values correspond to the values of n at which the lower-bound on utilities becomes binding. By definition of the lower-bound, over these intervals, the marginal tax rate is such that the well-being index of the worst-off agent at each n is equalized across n. In the lower-subproblem, that is $\forall n \in [\underline{n}, n_1]$, this consists in equalizing the well-being of agents earning the minimum wage $\underline{w}$ but having different preferences θ_i. For $\mathcal{R}^{\tilde{R}lex}$, the incentive compatible allocation that equalizes the well-being of agents with the same wage corresponds to a zero marginal tax rate. As a result, the optimal tax on low-incomes under $\mathcal{R}^{\tilde{R}lex}$ is always such that there is, for some $n_u \in [\underline{n}, n_1]$, an interval $[\underline{n}, n_u]$ with zero marginal tax rates followed by an interval $[n_u, n_{\underline{b}})$ over which marginal tax rates are negative.[18] In the upper sub-problem, that is $\forall n \in [n_1, \infty)$, characterizing the exact shape of the incentive compatible allocation that equalizes well-being across the worst-off within each n (agents with the same preference parameter $\overline{\theta}$ but different wages) requires parameterizing agents' (and reference) preferences.

This brings us to step 4. The assumption of iso-elastic and quasi-linear preferences really only is needed in this final step. In step 4, we first derive Lemma 2, an algebraic expression for the well-being index of the worst-off within each n: $\underline{b}(u(n); n)$. Using this algebraic expression, we are able to calculate the optimal marginal tax rate on intervals for which the lower-bound (10) is binding. Under our assumptions on preferences, the marginal tax rate is constant on these intervals. We also derive an expression for the solution on intervals where the lower-bound is not binding.

After introducing some notation, we formally prove steps 1-4. A few additional details and more technical derivations are left for online appendix I. In online appendix J, we also present some generalizations of steps 1-3 and the consequent robustness of certain key features of the optimal tax schedule to allowing for more general preferences.

Introducing some notation

Studying tax schemes forces us to study the income/consumption space instead of the labor time/consumption space. As a general rule, we underline variables when they are converted from the former to the latter space. Let $E = (w_N, R_N)$ be an economy.

[18] The remaining types $[n_{\underline{b}}, n_1]$ are bunched with n_1 at income y_1^*.

For each $i \in N$, we redefine preferences R_i as $\underline{R}_i$ as follows:

$$(y, c)\, \underline{R}_i\, (y', c') \Leftrightarrow (\frac{y}{w_i}, c)\, R_i\, (\frac{y'}{w_i}, c).$$

We now refer to (second-best) budgets as $\underline{B}(\tau)$, defined as

$$\underline{B}(\tau) = \{\underline{z} = (y, c) \in \underline{X} \mid c \leq y - \tau(y)\}.$$

All implicit budgets in the (y, c) space have slopes equal to 1. We now refer to them as follows:

$$\underline{IB}(\underline{z}_i, \underline{R}_i) = \{\underline{z} = (y, c) \in \underline{X} \mid \exists t \in \mathbb{R} : c - y \leq t \text{ and } \underline{z}_i\, \underline{I}_i\, \underline{m}(\underline{R}_i, \underline{IB}(\underline{z}_i, \underline{R}_i))\}.$$

Let τ be a tax scheme. Let $\underline{z}_N$ be the allocation generated by the tax scheme, that is for all $i \in N$,

$$\underline{z}_i \in \underline{m}(\underline{R}_i, \underline{B}(\tau)).$$

We say that τ is feasible if it generates $\underline{z}_N = (y_i, c_i)_{i \in N}$ such that

$$\sum_{i \in N} c_i \leq \sum_{i \in N} y_i.$$

We say that a feasible tax scheme τ is the optimal tax scheme in economy $E = (w_N, R_N)$ if τ generates allocation $\underline{z}_N$ corresponding to allocation z_N in the labor time/consumption space, and for all other feasible tax scheme τ', generating allocation $\underline{z}'_N$ corresponding to allocation z'_N in the labor time/consumption space we have

$$z_N\, \mathbf{R}^{\tilde{R}lex}\, z'_N.$$

D.1 Step 1: The worst-off within each n

Lemma 1 *Among agents of type n, agents with the lowest wage are always the worst-off according to $b(z_i, w_i, \theta_i)$. That is, $\forall j, k$ with $n_j = n_k$, $\underline{z}_j = \underline{z}_k$, $w_j < w_k$,*

$$b(z_j, w_j, \theta_j) \leq b(z_k, w_k, \theta_k)$$

Proof. Indeed, let us consider two agents, j and k with $n_j = n_k$, such that $\underline{R}_j = \underline{R}_k$, but $w_j < w_k$. Because they have the same preferences in the income-consumption space, they choose the same $\underline{z}_j = (y_j, c_j) = \underline{z}_k = (y_k, c_k)$, with the resulting property that $\underline{IB}(\underline{z}_j, \underline{R}_j) = \underline{IB}(\underline{z}_k, \underline{R}_k)$. Therefore, shifting our attention from the (y, c)-space to the (ℓ, c)-space, we note that $w_j < w_k$ implies $IB(z_j, w_j, R_j) \subset IB(z_k, w_k, R_k)$, where $z_j = (y_j/\theta_j, c_j)$ and $z_k = (y_k/\theta_k, c_k)$. Consequently, $b(z_j, w_j, R_j) \leq b(z_k, w_k, R_k)$. The argument is illustrated in figure 11.

■

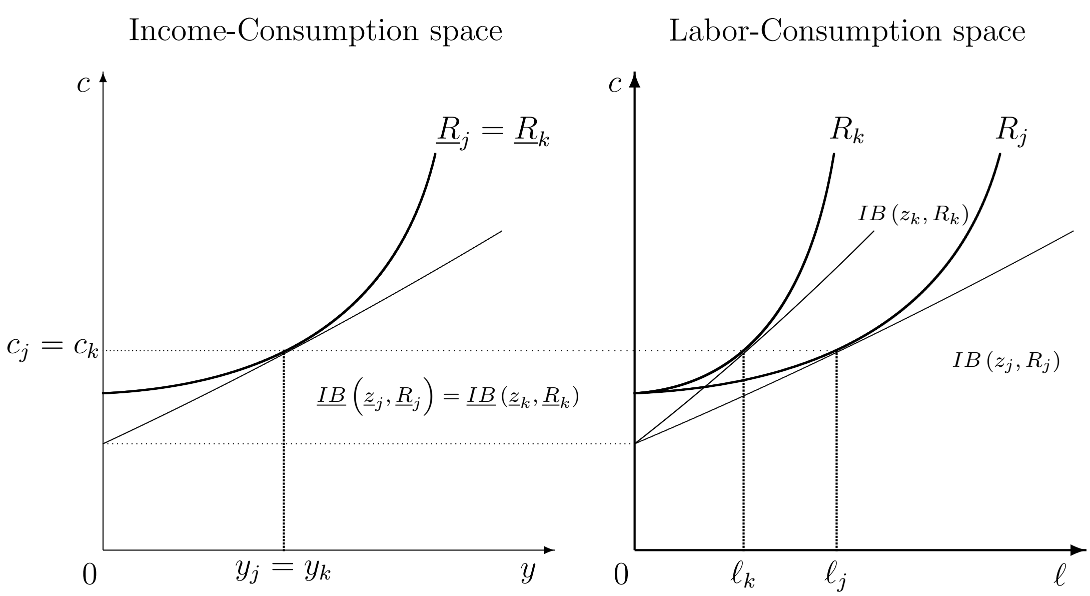

Figure 11: Illustration of Lemma 1. In the (y, c) space, agents j and k have the same preferences $\underline{R}$, choose the same bundle $\underline{z}$ and have the same implicit budget $\underline{IB}$. They have different wages $w_j < w_k$. To be represented in the (ℓ, c) space, each agent's indifference curve R, bundle z and implicit budget IB needs to be re-scaled by their wage. Indeed, $\ell_k = y_k/w_k < \ell_j = y_j/w_j$. Clearly, $IB\left(z_j, w_j, R_j\right) \subset IB\left(z_k, w_k, R_k\right)$. Therefore, $b\left(z_j, w_j, R_j\right) \leq b\left(z_k, w_k, R_k\right)$.

D.2 Step 2: Who is the worst-off across n?

Proposition 4 *If tax scheme τ is optimal in economy $E = (w_N, R_N)$ for a social welfare function $\mathbf{R}^{\tilde{R}lex}$, then allocation $\underline{z}_N$ generated by τ has the feature that the corresponding allocation z_N in the labor time/consumption space is such that agent 1* *has the lowest well-being level:*

$$b(z_1, w_1, R_1) \leq b(z_i, w_i, R_i), \forall i \in N.$$

Proof. Figure 12 illustrates the reasoning. Let $E = (w_N, R_N)$. Let τ be optimal in E for $\mathbf{R}^{\tilde{R}lex}$. Let $\underline{z}_N$ be generated by τ. Let z_N corresponds to $\underline{z}_N$ in the (ℓ, c) space. Without loss of generality, we restrict our attention to functions $y - \tau(y)$ that follow the lower envelope of the union of the upper contours sets at all bundles $\underline{z}_i$. Formally, we assume that $\forall y \in \mathcal{R}_+$, there exists $i \in N$ such that $(y, y - \tau(y)) \, \underline{I}_i \, \underline{z}_i$. If it were not the case, then at least some interval of incomes would not be chosen by any agents. Modifying the tax scheme, on that latter interval, so that it follows the lower envelope of the agents' upper contours sets does not modify any agent's behavior nor well-being so that the change is irrelevant.

Let us assume, contrary to the claim, that

$$b(z_1, w_1, R_1) > b(z_j, w_j, R_j) = b$$

for some $j \in N$. Let $\hat{\underline{z}}_1 = (\hat{y}_1, \hat{c}_1) \in \underline{X}$ be defined by:

$$b\left((\frac{\hat{y}_1}{w_1}, \hat{c}_1), w_1, R_1 \right) = b$$

and

$$\hat{\underline{z}}_1 \in \underline{IB}(\hat{\underline{z}}_1, \underline{R}_1).$$

Note that by assumption, $\underline{z}_1 \, \underline{P}_1 \, \hat{\underline{z}}_1$. Later on in the proof, we will use the following fact:

$$\forall y \leq \hat{y}_1, \tau(y) \leq \hat{y}_1 - \hat{c}_1. \tag{11}$$

Let us prove this fact. Assume it is not true. Then there exists $y^* < \hat{y}_1$, such that $\tau(y^*) = \max_{y < \hat{y}_1} \tau(y)$ and $\tau(y^*) > \tau(\hat{y}_1)$. By the assumption made in the beginning of the proof, there exists $i \in N$ such that $\underline{z}_i \, \underline{I}_i \, (y^*, y^* - \tau(y^*))$. Note that among all the feasible bundles that leave i indifferent to $(y^*, y^* - \tau(y^*))$, this bundle itself maximizes the paid tax (or minimizes the received transfer), so that, if τ is supposed to be optimal, it is i's bundle: $\underline{z}_i = (y^*, y^* - \tau(y^*))$. Indeed, if it were not the case, by replacing $\underline{z}_i$ with $(y^*, y^* - \tau(y^*))$, the planner would not change anything in the well-being of the agents, all the incentive constraints would remain satisfied and there would be a budget surplus, which could be redistributed. As a consequence, and because we have assumed that $\forall n \in [\underline{n}, n_1]$ there exist agents with the minimum wage, there exists $j \in N$ such that $w_j = w_1$ and $\underline{z}_j = (y^*, y^* - \tau(y^*))$. As a result,

$$\underline{IB}(\underline{z}_j, \underline{R}_j) = \{\underline{z} = (y, c) \in \underline{X} \mid c = y - \tau(y^*)\}.$$

That implies

$$\underline{IB}(\underline{z}_j, \underline{R}_j) \subset \underline{IB}(\hat{\underline{z}}_1, \underline{R}_1),$$

with the consequence that

$$b\left((\frac{y^*}{w_j}, y^* - \tau(y^*)), w_j, R_j \right) < b,$$

the desired contradiction, proving the fact stated above.

We now define an income threshold $\hat{y}$ by distinguishing two cases.

1. $\hat{c}_1 > \hat{y}_1 - \tau(\hat{y}_1)$: $\hat{y}$ is defined by

$$(\hat{y}, \hat{y} - \tau(\hat{y})) \, \underline{I}_1 \, \hat{\underline{z}}_1,$$

   and in case several such $\hat{y}$ exist, we take the largest;

2. $\hat{c}_1 > \hat{y}_1 - \tau(\hat{y}_1)$: $\hat{y}$ is defined by

$$\tau(\hat{y}) = \hat{y}_1 - \hat{c}_1,$$

   and in case several such $\hat{y}$ exist, we take the largest.

Let τ' be defined by: for all $y \geq 0$

$$\tau'(y) = \max\{\tau(y), \tau(\hat{y})\}. \tag{12}$$

and let $\underline{z}'_N$ be the allocation generated by τ'.

Claim 1: τ' collects a budget surplus: $\sum_{i \in N} \tau'(y'_i) > 0$. By construction, $\underline{B}(\tau') \subset \underline{B}(\tau)$. Therefore, if $\tau'(y_i) = \tau(y_i)$, that is, τ' and τ coincide at y_i, then $\underline{z}'_i = \underline{z}_i$, that is agent i does not change her choice of bundle. If $\tau'(y_i) \neq \tau(y_i)$, because $\tau(y_i) < \tau(\hat{y})$, then $\underline{z}'_i \neq \underline{z}_i$ but by construction of τ', $\tau'(y'_i) \geq \tau(y_i)$. In any case, $\tau'(y'_1) > \tau(y_1)$, which proves the claim.

Claim 2: $\min_{i \in N} b(z_i', w_i, R_i) = b$. Let $k \in N$ be such that

$$b(z_k', w_k, R_k) = \min_{i \in N} b(z_i', w_i, R_i).$$

We consider two cases in turn. Case 1: $w_k = w_1$. Because $R_k \leq R_1$ since $\left(\theta_k \leq \overline{\theta} = \theta_1\right)$, $y_k' \leq y_1'$. Gathering Eq. (12), and the fact that $\tau(\hat{y}) \leq \hat{y}_1 - \hat{c}_1$, we obtain

$$\forall y \leq \hat{y}_1, \tau'(y) \leq \hat{y}_1 - \hat{c}_1.$$

This implies

$$\underline{IB}(\underline{z}_k, \underline{R}_k) \subseteq \underline{IB}(\underline{\hat{z}}_1, \underline{R}_1),$$

with the immediate consequence that $b(z_k', w_k, R_k) \geq b$, the desired outcome. Case 2: $w_k > w_1$. If $y_k' \geq \hat{y}$, then $y_k' = y_k$, $\underline{z}_k' = \underline{z}_k$, nothing changes in the well-being levels and the claim is proven. If $y_k' < \hat{y}$, then

$$\underline{IB}(\underline{z}_k, \underline{R}_k) \supseteq \underline{IB}(\underline{\hat{z}}_1, \underline{R}_1),$$

so that, because $w_k > w_1$,

$$IB(z_k', R_k, w_k) \supset IB(z_1', R_1, w_1),$$

so that

$$b(z_k', R_k, w_k) \geq b(z_1', R_1, w_1),$$

the desired outcome that completes the proof of claim 2.

Gathering both claims, τ' does not decrease the minimal well-being level whereas it collects a budget surplus. Redistributing that surplus among all agents would strictly increase the minimal well-being level, contradicting the assumption that τ is optimal.

■

D.3 Step 3: Setting-up the maximization problem and deriving the first-order conditions

We consider the "dual" of the original problem:

$$V \equiv \max_{\{u(n),y(n)\}} \int_{\underline{n}}^{\infty} \left(y(n) - \left(u(n) + \frac{\epsilon}{1+\epsilon} \left(\frac{y(n)}{n} \right)^{\frac{1+\epsilon}{\epsilon}} \right) \right) dF(n) \tag{13}$$

subject to (3),(9) and (10)

Constraints (9) and (10) are the lower-bound on utilities at the solution presented in the overview of the proof and constraint (3) is the incentive compatibility constraint. The first-order conditions of (3) being satisfied at $n' = n$ implies:

$$u'(n) = \left(\frac{y\,(n)}{n} \right)^{\frac{1+\epsilon}{\epsilon}} \left(\frac{1}{n} \right) \tag{14}$$

The second-order conditions of (3) are equivalent to (see Mirrlees 1976):

$$y'(n) \geq 0 \ \forall n \in [\underline{n}, \infty) \tag{15}$$

We divide (13) into two independent subproblems:

$$V^L(y_1) \equiv \max_{\{u(n),y(n)\}} \int_{\underline{n}}^{n_1} \left(y(n) - \left(u(n) + \frac{\epsilon}{1+\epsilon} \left(\frac{y(n)}{n} \right)^{\frac{1+\epsilon}{\epsilon}} \right) \right) dF(n)$$

subject to (9), (10), (14), (15) and

$$y(n) \leq y_1 \ \forall n$$

and

$$V^U(y_1) \equiv \max_{\{u(n),y(n)\}} \int_{n_1}^{\infty} \left(y(n) - \left(u(n) + \frac{\epsilon}{1+\epsilon} \left(\frac{y(n)}{n} \right)^{\frac{1+\epsilon}{\epsilon}} \right) \right) dF(n)$$

subject to (9), (10), (14), (15) and

$$y(n) \geq y_1 \ \forall n$$

We denote by $y^*(n_1)$ the income of agent 1 at the solution. The requirement that $y(n) \leq y^*(n_1) \ \forall n \in [\underline{n}, n_1]$ in the lower subproblem and $y(n) \geq y^*(n_1) \ \forall n \in [n_1, \infty)$ in the upper subproblem follows from the constraint that $y(n)$ be everywhere increasing in n (including at n_1). This last constraint is likely to be binding at n_1 (see appendix I.1 for a detailed explanation).

In practice, both $V^L(y)$ and $V^U(y)$ can be solved independently for a given y_1 using Hamiltonians. By definition of $y^*(n_1)$ being optimal, it solves:

$$\frac{\partial V^L}{\partial y_1}\left(y^*(n_1)\right) + \frac{\partial V^U}{\partial y_1}\left(y^*(n_1)\right) = 0 \tag{16}$$

where the derivatives of $V^L(y)$ and $V^U(y)$ can be found using the envelope theorem.

In online appendix I.3, we define the Hamiltonians corresponding to each sub-problem and derive the first-order conditions which can be re-arranged as:

$$\frac{\tau'\left(y(n)\right)}{1-\tau'\left(y(n)\right)} = A(n)D(n)C(n) \qquad \forall n \in [\underline{n}, n_{\underline{b}}] \tag{17}$$

$$\frac{\tau'\left(y(n)\right)}{1-\tau'\left(y(n)\right)} = A(n)B(n)C(n) \qquad \forall n \in [n_{\overline{b}}, \infty) \tag{18}$$

with

$$\begin{aligned}
A(n) &= \left[\frac{1+\epsilon}{\epsilon}\right] && (19)\\
B(n) &= 1 - F\left(n\right) - \int_n^\infty \mu(t)dt && (20)\\
C(n) &= \frac{1}{nf(n)} && (21)\\
D(n) &= -F\left(n\right) + \int_{\underline{n}}^n \mu(t)dt && (22)
\end{aligned}$$

where $\mu\left(n\right)$ is the lagrange multiplier associated with the lower-bound on utilities (constraint (10)) at n. In addition, $n_{\underline{b}} \in [\underline{n}, n_1]$ and $n_{\overline{b}} \in [n_1, \infty)$ are the end-points of the interval of types over which agents are bunched and earn $y^*\left(n_1\right)$. For any n at which the constraint $u(n) \geq \underline{u}(n)$ is binding, $\mu\left(n\right) > 0$. Otherwise, $\mu\left(n\right) = 0$.

D.4 Step 4: Optimal tax formula as a function of economic and normative primitives

Lemma 2 *Under the assumption that reference preferences can be represented by the utility function in (5), the worst-off agent at type n_i has well-being index*

$$\underline{b}\left(u(n_i);n_i\right) = u(n_i) + \frac{1}{1+\epsilon}\left[\left(n_i\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon} - (n_i)^{1+\epsilon}\right] \text{ if } n_i \geq n_1 \tag{23}$$

$$= u(n_i) + \frac{1}{1+\epsilon}\left[\left(\underline{w}\tilde{\theta}\right)^{1+\epsilon} - (n_i)^{1+\epsilon}\right] \text{ if } n_i \leq n_1. \tag{24}$$

Proof.

The intercept of the implicit budget of agent i is the smallest T_i such that

$$T_i + \max_{\ell_i}\left\{w_i\ell_i - \frac{\epsilon}{1+\epsilon}\left(\frac{\ell_i}{\theta_i}\right)^{\frac{1+\epsilon}{\epsilon}}\right\} \geq u_i$$

Therefore:

$$T_i = u_i - \frac{1}{1+\epsilon}\left(w_i\theta_i\right)^{1+\epsilon}$$

Maximizing utility function (5) (the reference preferences) over i's implicit budget, we obtain the following indirect utility:

$$b\left(u_i, w_i, \theta_i\right) = u_i + \frac{1}{1+\epsilon}\left[\left(w_i\tilde{\theta}\right)^{1+\epsilon} - \left(w_i\theta_i\right)^{1+\epsilon}\right]$$

Finally, per Lemma 1, the minimum well-being within each n_i is:

$$\underline{b}\left(u(n_i);n_i\right) = u(n_i) + \frac{1}{1+\epsilon}\left[\left(n_i\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon} - (n_i)^{1+\epsilon}\right] \text{ if } n_i \geq n_1$$

$$= u(n_i) + \frac{1}{1+\epsilon}\left[\left(\underline{w}\tilde{\theta}\right)^{1+\epsilon} - (n_i)^{1+\epsilon}\right] \text{ if } n_i \leq n_1$$

■

Using the expression for the well-being index in Lemma 2, we first derive the marginal tax rate on intervals where the lower-bound on utilities (10) is binding.

Equalizing the well-being of the worst-off for each n implies $\frac{d}{dn}\underline{b}\left(u(n);n\right)=0 \Rightarrow$:

$$u'(n) = \qquad n^{\epsilon} \qquad \forall n \in [\underline{n}, n_1) \tag{25}$$

$$u'(n) = \left[1-\left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right] n^{\epsilon} \quad \forall n \in [n_1, \infty) \tag{26}$$

Combining Eq. (25) and (26) with the first-order conditions of the incentive compatibility constraint, we find:

$$y(n) = n^{1+\epsilon} \qquad \forall n \in [\underline{n}, n_1) \tag{27}$$

$$y(n) = \left[1-\left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{\epsilon}{1+\epsilon}} n^{1+\epsilon} \qquad \forall n \in [n_1, \infty) \tag{28}$$

Finally, using the fact that $1-\tau'\left(y(n)\right)=\left(\frac{y(n)}{n}\right)^{\frac{1}{\epsilon}}\left(\frac{1}{n}\right)$:

$$\tau'\left(y(n)\right) = 0 \qquad \forall n \in [\underline{n}, n_1) \tag{29}$$

$$\tau'\left(y(n)\right) = 1-\left[1-\left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \qquad \forall n \in [n_1, \infty) \tag{30}$$

Let us first focus on $[\underline{n}, n_1)$ and fully characterize the solution on that interval. From (15), $y(n)$ is increasing. Therefore, $y^*(n) \leq y_1 = y^*(n_1)$ can only bind (if at all) on $[n_{\underline{b}}, n_1]$ for some $n_{\underline{b}} \leq n_1$. If such an interval exists, then $y^*(n) = y^*(n_1)$ $\forall n \in [n_{\underline{b}}, n_1]$.

Re-arranging the first-order conditions (17), we have

$$\int_{\underline{n}}^{n} \mu(t)dt = \frac{1+\epsilon^u}{\epsilon^c}\frac{\tau'\left(y(n)\right)}{1-\tau'\left(y(n)\right)} n f\left(n\right) + \int_{\underline{n}}^{n} f\left(t\right) dt$$

As a result, we can write

$$\mu(n) = \frac{d}{dn}\left[\frac{1+\epsilon^u}{\epsilon^c}\frac{\tau'\left(y(n)\right)}{1-\tau'\left(y(n)\right)} n f\left(n\right)\right] + f\left(n\right)$$

Whenever $\mu(n) > 0$, we have $\tau'(y(n)) = 0$ and $\tau''(y(n)) = 0$ because of (29). As a result, either $\mu(n) = 0$ or $\mu(n) = f(n)$. This implies that $\int_{\underline{n}}^{n} f(t)\,dt \geq \int_{\underline{n}}^{n} \mu(t)dt$ which, combined with (17), means that tax rates are (weakly) negative on $[\underline{n}, n_1]$. Notice that if marginal tax rates are everywhere zero or negative, the lower-bound (along which the marginal tax rate is zero) can be binding on at most one interval from $\underline{n}$ to some n_u (this proves claim 1 in Proposition 1). Finally, because the lower-bound is only binding on interval $[\underline{n}, n_u]$, we have that $\int_{\underline{n}}^{n} \mu(t)dt = \int_{\underline{n}}^{n_u} \mu(t)dt = \int_{\underline{n}}^{n_u} f(t)dt\ \forall n \in [n_u, n_1]$. Plugging back into (17) proves claim 2 of Proposition 1.

We now turn our attention to $[n_1, \infty)$. First, because of (15), the constraint that $y(n) \geq y_1 = y^*(n_1)$ can only bind (if at all) on interval $[n_1, n_{\bar{b}})$ for some $n_{\bar{b}} \in [n_1, \infty)$. This directly proves claim 3 of Proposition 1. Second, by assumption there is at most one interval above n_1 where (10) binds. We call that interval $[n_l, n']$ with $n_l \leq n' \in (n_{\bar{b}}, \infty)$. The marginal tax rate on that interval is given by (30).

Regarding the optimal tax rate in the upper tail, notice that when $\int_n^\infty \mu(t)dt = 0$, the optimal tax rate is the revenue-maximizing one defined by:

$$\frac{\tau'(y(n))}{1 - \tau'(y(n))} = \frac{1 - F(n)}{nf(n)} \left(\frac{1+\epsilon}{\epsilon}\right) \tag{31}$$

and converges to $\tau'(y(n)) \to \frac{1+\epsilon}{1+\epsilon+\alpha\epsilon}$ as $n \to \infty$.[19] Denote by $\tilde{\tau}'(y(n))$ the marginal tax rate defined by (31). Two scenarii are possible. If $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} \leq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\theta}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$, there exist some n' such that $\tilde{\tau}'(n') = 1 - \left[1 - \left(\frac{\tilde{\theta}}{\theta}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ and $\tilde{\tau}'(n) \leq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\theta}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ $\forall n \in [n', \infty)$. In that case, applying $\tau(y(n)) = \tilde{\tau}(y(n))$ $\forall n \in [n', \infty)$ is a Pareto improvement compared to applying (30). This proves claim 6 in Proposition 1. In that case, we also have:

$$\int_n^\infty \mu(t)dt = \int_{n_l}^{n'} \mu(t)dt = 1 - F(n_l) - \frac{1 - \left[1 - \left(\frac{\tilde{\theta}}{\theta}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}}{\left(\frac{1+\epsilon}{\epsilon}\right)\left[1 - \left(\frac{\tilde{\theta}}{\theta}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}} n_l f(n_l)$$

[19] See Boadway and Jacquet (2008) for a discussion of the Rawlsian marginal tax rates under quasi-linear preferences and various distributions of wages.

$\forall n \in [n_1, n_l]$.[20] Combining the latter with Eq.(18) yields an expression for marginal tax rates over $[n_{\bar{b}}, n_l]$ (claim 4 in Proposition 1).

If $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} \geq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$, the lower-bound on utilities remains binding in the tail. In that case, the marginal tax rate over $[n_{\bar{b}}, n_l]$ is the same as in the previous case but the tax rate over interval $[n_l, \infty)$ is also given by Eq.(30).[21] This demonstrates claim 5 in Proposition 1 and concludes the proof.

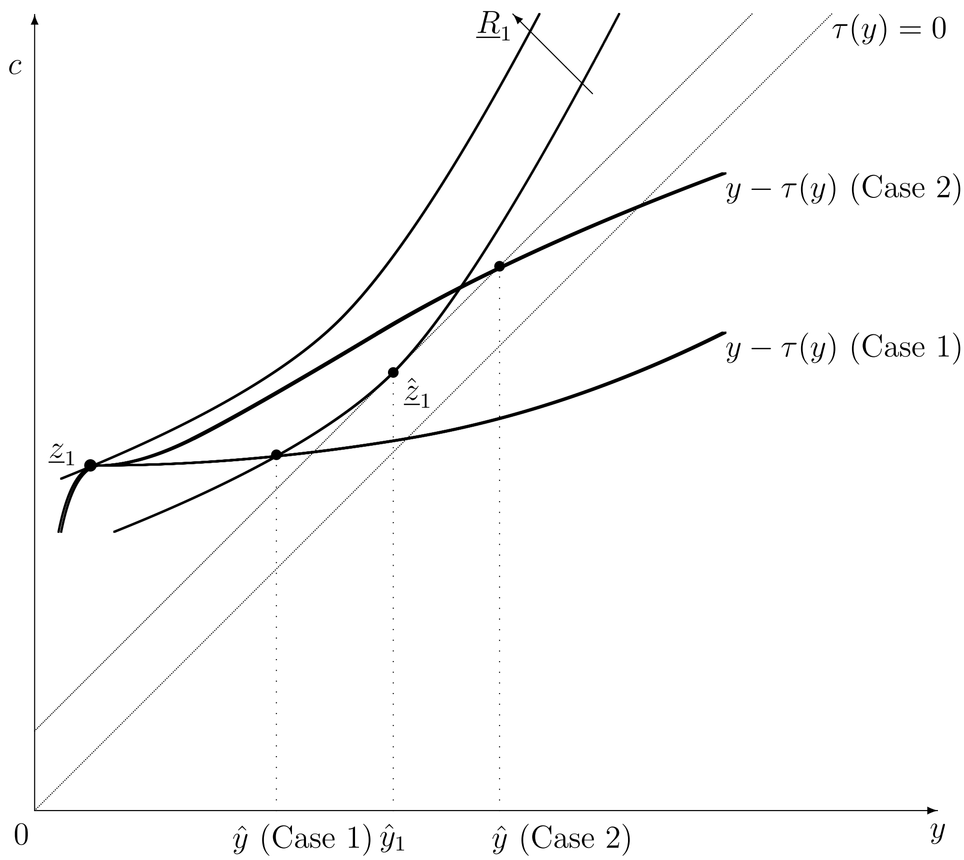

Figure 12: Illustration of the proof of Proposition 4. The result is proven by contradiction. Contrary to the claim, assume that τ is optimal for $\mathbf{R}^{\tilde{R}lex}$ but some agent $j \neq 1$ has a well-being index $b(z_j, w_j, R_j) = b < b(z_1, w_1, R_1)$. The proof involves bundle $\hat{\underline{z}}_1$ at which 1's well-being index is b and $\hat{\underline{z}}_1$ is on the frontier of 1's implicit budget at that level of well-being. There are two possible cases: either τ crosses 1's indifference curve (at well-being b) below $\hat{\underline{z}}_1$ (case 1) or it crosses above it (case 2). In each case we can define some $\hat{y}$ such that alternative tax scheme $\tau'(y) = \max\{\tau(y), \tau(\hat{y})\}$ generates a budget surplus and no agents has well-being below b. This contradicts the optimality of τ.

[20]For more details, see paragraph "**Case 2**" in online appendix I.3.4.

[21]For more details, see paragraph "**Case 1**" in online appendix I.3.4.

Online Appendix

E Other Calibrated Optimal Tax Rates and Comparative Statics

In this section, we show the results of different calibration exercises.

Comparing the three polar choices of $\tilde{R}$ under different assumptions on the poverty line

- Figure 13 presents the optimal tax and consumption schedule for the three polar choices of $\tilde{R}$ under the assumption that $\overline{P}_0 = \$15,379$ and the average slope is $\$6/h$. Compared to the case where the poverty line is set at $\$18,850$, more divergence appears between the three polar cases. For instance, the tax rate on large incomes ranges from 0.39 to 0.49 and the consumption of agents with zero income varies between $\$8,632$ and $\$12,778$. When the poverty line is set lower, there is more surplus (above the allocation that just eradicates poverty) to be redistributed. As a result, the choice between least-, most-redistributive and quasi-linear $\tilde{R}$ matters more.

- Figure 14 presents the optimal tax and consumption schedule for the three polar choices of $\tilde{R}$ under the assumption that $\overline{P}_0 = \$18,850$ and the average slope is $\$4/h$. Compared to the case where the average slope is $\$6/h$, more divergence appears between the three polar cases. For instance, the tax rate in the tail varies between 0.47 and 0.54 while the consumption at zero income ranges from $\$11,009$ to $\$13,884$. Similar to a lower poverty line, when the poverty line is flatter, there is more surplus to be redistributed.

- **Comment:** figures 13 and 14 show calibrations for lower or flatter poverty lines. On the other hand, higher and steeper poverty lines lead to less discrepancies between the three polar choices. In particular, holding the average slope constant at $\$6/h$, for any poverty line above $\$20,036$, the solution according to all three polar cases collapse to a unique allocation. Similarly, holding the level of the poverty line at $\$18,850$, for any poverty line steeper than $\$6.98/h$, the solution according to all three choices of $\tilde{R}$ collapse to a unique optimal allocation.

Holding the choice of $\tilde{R}$ constant but varying the slope or level of the poverty line

- Figure 15 displays the optimal allocations when $\tilde{R}$ are quasi-linear and the average slope of the poverty line takes on different values. When $\tilde{R}$ is quasi-linear, the level of the poverty line does not matter. Lower slopes lead to more redistributive allocations and higher jump in marginal tax rates at y_1^*. In particular, $c(0)$ is larger with flatter poverty lines and tax rates on large incomes are also larger. We also notice that y_1^* becomes lower as the poverty line flattens. The intuition is that tax systems are more and more generous towards the most hardworking minimum wage agents as the slope decreases, but this goes with an increase in the incentives of higher types to mimic this agent who is then incentivized to earn a lower income. Finally, for flat enough poverty lines (below $\$1.51/h$) the lower-bound on the well-being of agents with $n > n_1$ is never binding: the optimal tax is the one that maximizes tax revenue from all agents with $n \geq n_{\bar{b}}$. When the average slope is above $\$1.51/h$ but at or below $\$4/h$, the lower-bound on the well-being of agents with $n > n_1$ is binding on some interval $[n_l, n')$ but marginal tax rates on all agents above n' are those that maximizes revenue raised at the top. In other words, for flat enough poverty lines (at least lower than $\$4/h$), tax rates on top incomes are determined by efficiency considerations alone.

- Figure 16 shows the optimal allocations for the least redistributive $\tilde{R}$ and for different levels of the poverty line. We maintain the assumption that the average slope of the poverty line is $\$6/h$. Recall that, under the least-redistributive $\tilde{R}$, the optimal allocation is characterized by Proposition 1 where $\tilde{\theta}$ is determined endogeneously as the lowest value of $\tilde{\theta}$ for which poverty is just eradicated. In some cases, no such value exists because there are not enough resources in the economy to lift everybody out of poverty (at that level and slope of the PL). In these cases, the solution according to all three polar choices of $\tilde{R}$ collapse to a unique allocation that can be found by applying Proposition 1 with the value of $\tilde{\theta}$ that corresponds to an average slope of PL of $\$6/h$. As a result, for any poverty line above $\$20,036$, the allocation is identical and at least some agents are not lifted out of poverty. For other values of the poverty line: lower PLs lead to less redistribution. This means, among other things: lower $c(0)$, lower marginal tax rates on incomes above y_1^*, smaller jump in tax rates at y_1^* and consequently a higher y_1^*.

- Figure 17 presents the optimal allocation for the most redistributive $\tilde{R}$ and for different slopes of the poverty line. We maintain the assumption that the PL is anchored at $\$18,850$. Flatter poverty lines lead to more redistributive tax schemes. As mentioned in Section 3, high wage agents may have low well-being levels if their preferences are sufficiently less directed towards leisure than the reference preferences. This is more likely to be the case for steeper poverty lines (they lead to more labor averse reference preferences). As a result, a steeper poverty line leads to a more binding constraint on the well-being level of agents with $n > n_1$ and, as a result, limits the amount of surplus that is directed towards agents with $n \leq n_1$. Also notice, relatedly, that in this case, when the average slope is $\$8/h$, poverty is not eradicated.

- Figure 18 displays the optimal allocations for the most redistributive $\tilde{R}$ and for different levels of the poverty line. We maintain the assumption that the average slope of the PL is $\$6/h$. The comparative statics are slightly counter-intuitive: lower poverty lines lead to more redistribution. This is because the most redistributive $\tilde{R}$ directs as much surplus as possible towards agents with a lower skill under the constraint that agents with $n \geq n''$ have their implicit budget at or above the poverty line (along the lower-bound on utilities). In other words, higher poverty lines "protect" the well-being of agents with $n \geq n''$ and limit the amount of revenue that can be raised from them to re-distribute to all other agents.

Flat poverty lines: comparative statics when the average slope is $\$0/h$

Poverty lines that are strictly increasing in labor time are somewhat unusual. Our analysis easily accommodates perfectly flat poverty lines. When the poverty line is flat, the well-being of agents with $n > n_1$ is always strictly greater than 1's well-being - they are "protected" by incentive compatibility constraints. It immediately follows that quasi-linear and the most-redistributive $\tilde{R}$ lead to the same optimal allocation. This allocation corresponds to maximizing the well-being of agents with $n = n_1$ under the constraint that no agent with $n < n_1$ be worse-off than 1. As a result, tax rates on incomes above y_1^* are those that maximize revenue raised to be re-distributed to everyone.

- Figure 19 compares the optimal allocation according to the most-redistributive (and quasi-linear) $\tilde{R}$ to the optimum under the least redistributive $\tilde{R}$. When the poverty line is set at $\$18,850$, as is the case in figure 19, there is a large difference between these two extremes.

- Figure 20 shows the optimal allocation for the least redistributive $\tilde{R}$ under different levels of the poverty line. Similar to figure 16, lower poverty lines lead to less redistributive tax schemes. An important difference is that, when the poverty line is flat, there is enough resources to lift everybody out of poverty even for poverty lines as high as $\$28,695$.

Comparative statics on $\overline{\theta}$

In figure 21 and 22, we show the sensitivity of our calibrations to $\overline{\theta}$ for, respectively, quasi-linear and the least redistributive $\tilde{R}$. In both cases, we assume that $\overline{P}_0$ and the average slope of the PL is $\$6/h$. In theory, $\overline{\theta}$ is an important parameter in the optimal tax formula. First, it is clear from Proposition 1 that larger values of $\overline{\theta}$ (more hardworking) lead to lower marginal tax rates. Actually, the ratio $\frac{\tilde{\theta}}{\overline{\theta}}$ enters the expressions. As a result, varying $\overline{\theta}$ has a similar effect as varying the slope of the poverty line with one important difference: $\overline{\theta}$ also directly affects y_1^*. Conceptually, we can think of these comparative statics this way: we hold the distribution of n fixed and vary the extent to which inequality in n (and thus income) are due to differences in preference parameter θ. As we increase $\overline{\theta}$, we increase the range of θ in the economy and conceptually load more of the income inequality on heterogeneity of preferences as opposed to differences in wages.

- Figure 21 displays the comparative statics on $\overline{\theta}$ when $\tilde{R}$ are quasi-linear. We can see that larger values of $\overline{\theta}$ lead to non-positive marginal tax rates on a larger interval of income and the jump in marginal tax rates happens at a higher income level. It also yields lower marginal tax rates above y_1^* and lower average tax rates as the interval over which rates are non-positive is also extended. Given the comment that we have made about the relationship between $\overline{\theta}$ and heterogeneity in preferences, this result is similar in spirit to Lockwood and Weinzierl (2015), in which more heterogeneity of preferences also leads to less redistribution. They however consider a different measure of heterogeneity of preferences. Because our social welfare function is of the leximin type, an increase in heterogeneity of preferences that leaves $\overline{\theta}$ unchanged also leaves the optimal tax unaffected. This is an important difference with Lockwood and Weinzierl (2015).

- Figure 22 shows the sensitivity of marginal tax rates and the optimal consumption schedule to different values of $\overline{\theta}$ when $\tilde{R}$ are the least redistributive. Contrary to figure 21, the overall level of redistribution is not very sensitive to this parameter. In this case, the amount of redistribution through the tax system is mainly pinned down by the level of the poverty line. The main difference is that the range of incomes over which rates are negative or zero is larger when $\overline{\theta}$ is larger.

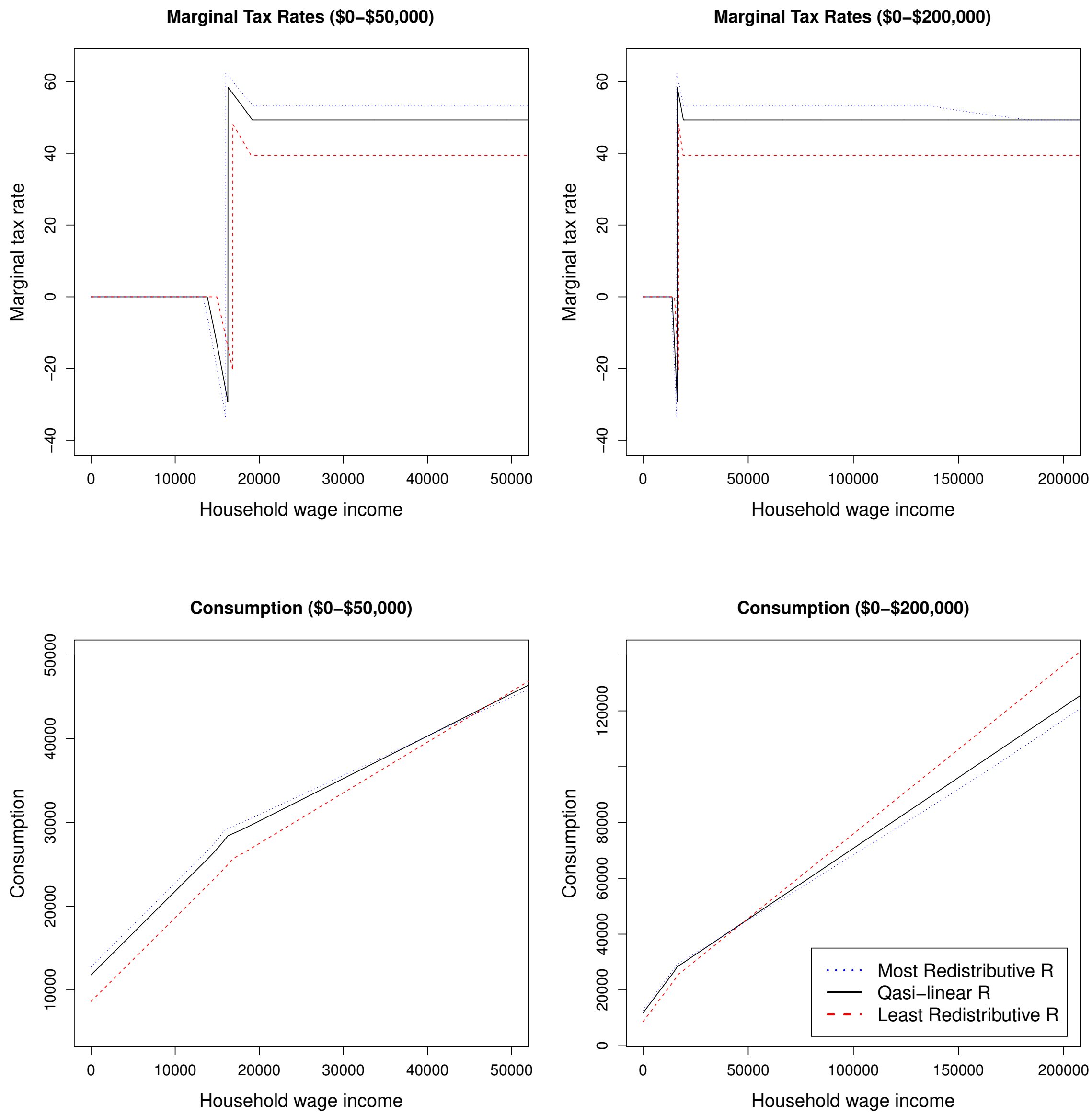

Figure 13: Comparing optimal tax rates and consumption under the 3 polar cases ($PL = \$15,379$; Slope $= \$6/h$). Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn).

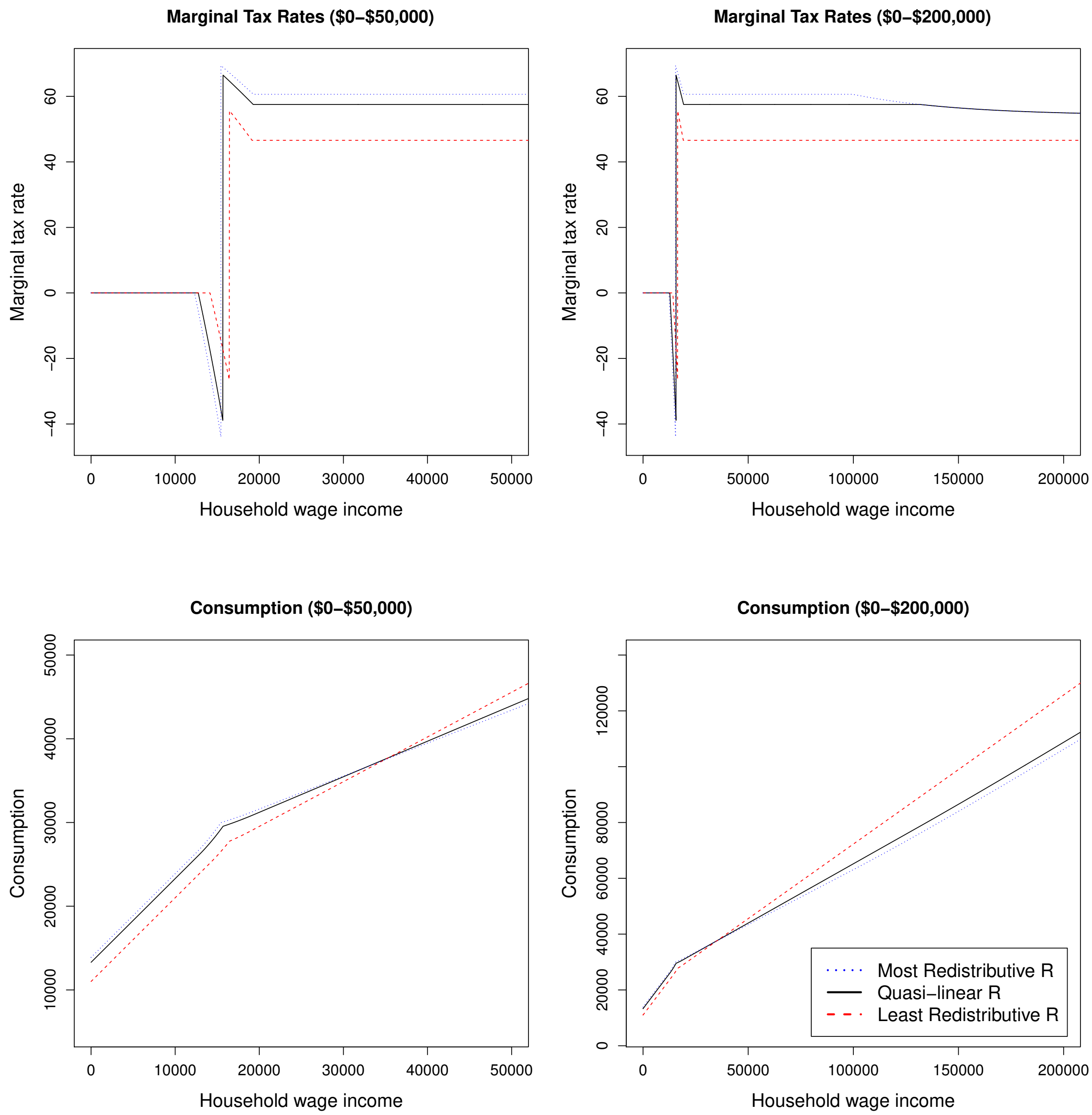

Figure 14: Comparing optimal tax rates and consumption under the 3 polar cases ($PL = \$18,850$; Slope $= \$4/h$). Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn).

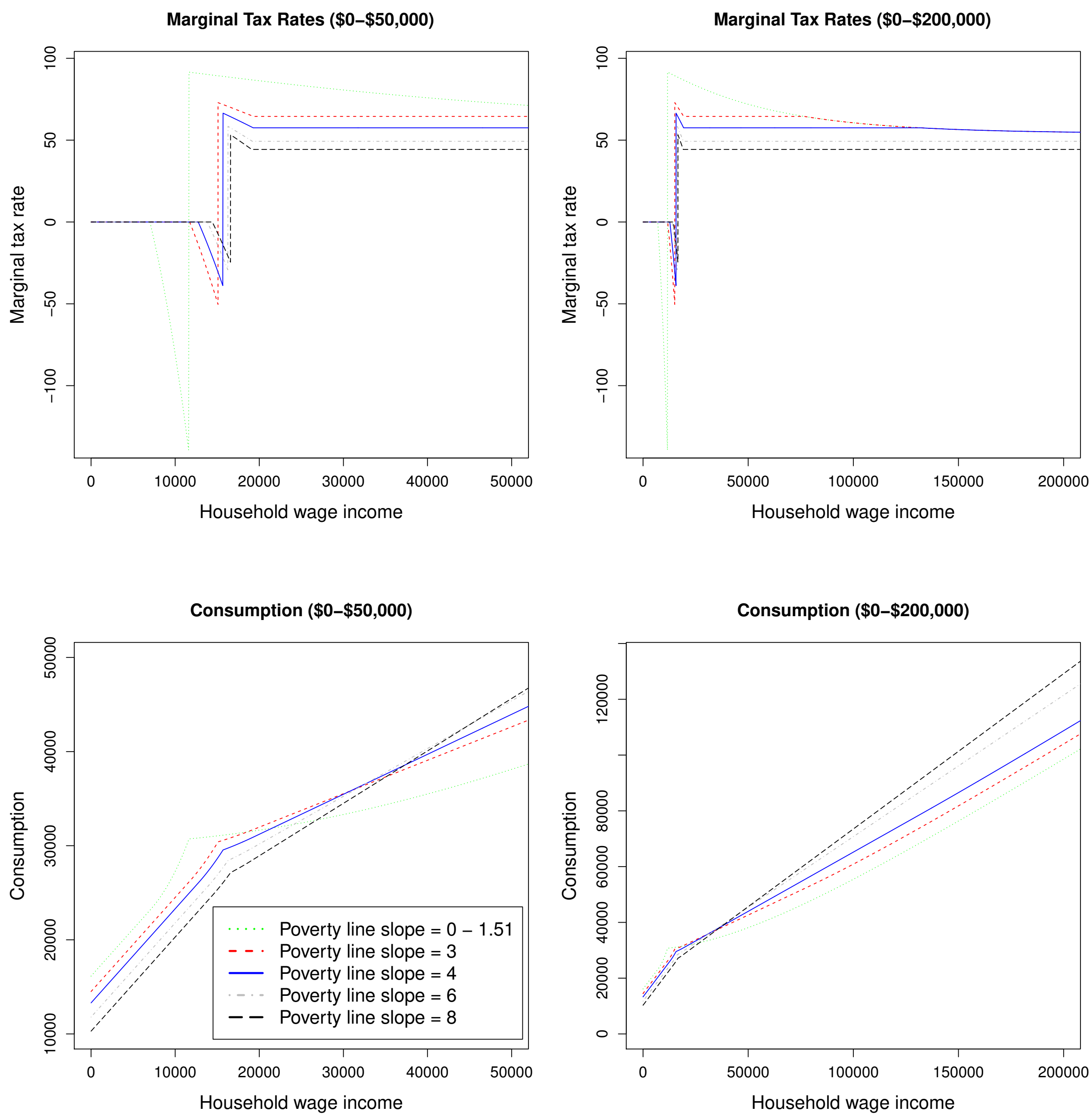

Figure 15: Quasi-linear R - comparative statics - different slopes of the poverty line. Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn). The reference preferences are assumed to be quasi-linear and the shape of indifference curves is chosen so that their average slope is x $\$/h$ with $x \in \{0, 3, 4, 6, 8\}$.

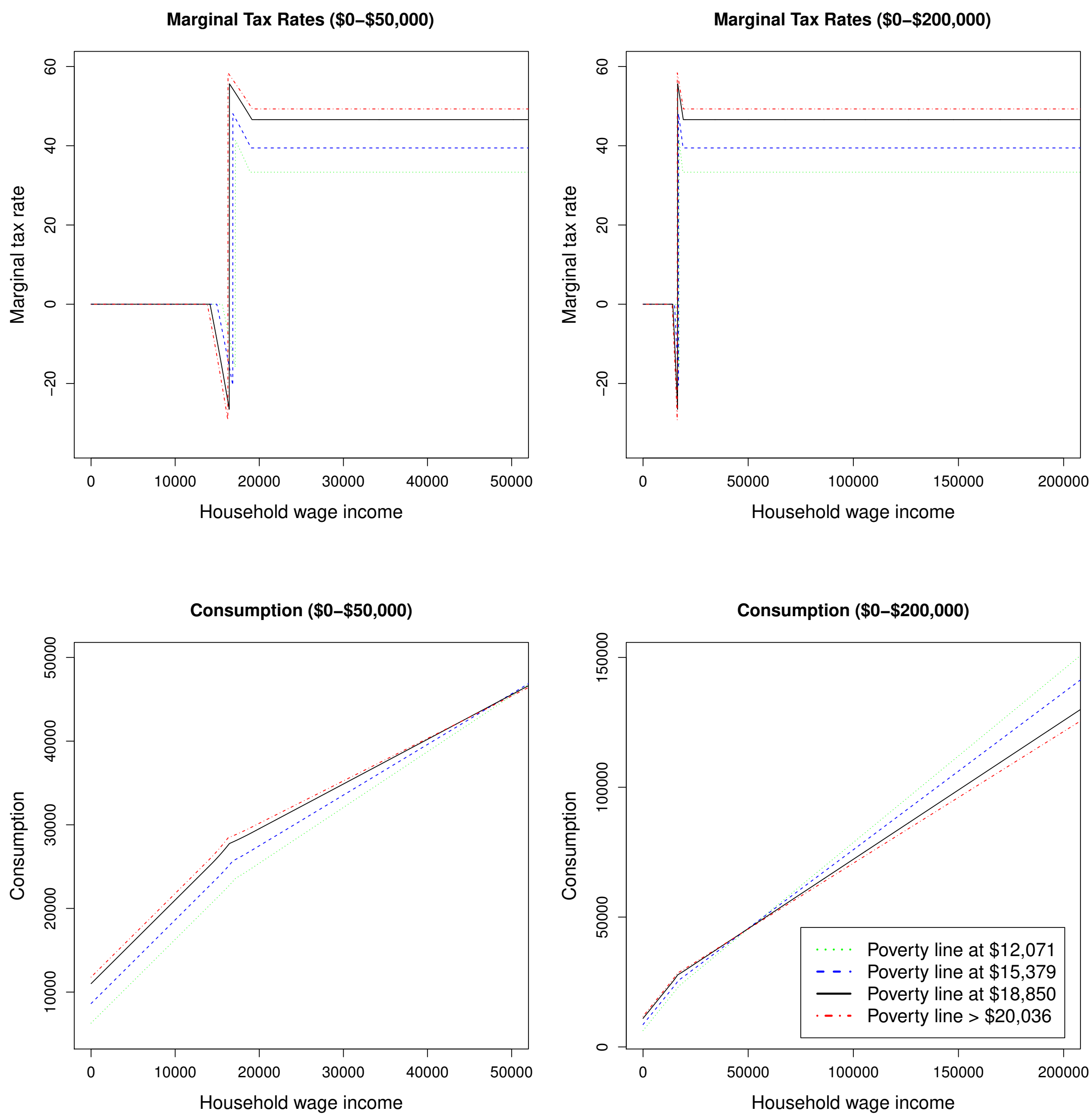

Figure 16: Least redistributive R - comparative statics - different levels of the poverty line (with an average slope = $\$6/h$). The values of the poverty line are the 2014 official poverty lines in the US for households of size one to three members respectively. Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn).

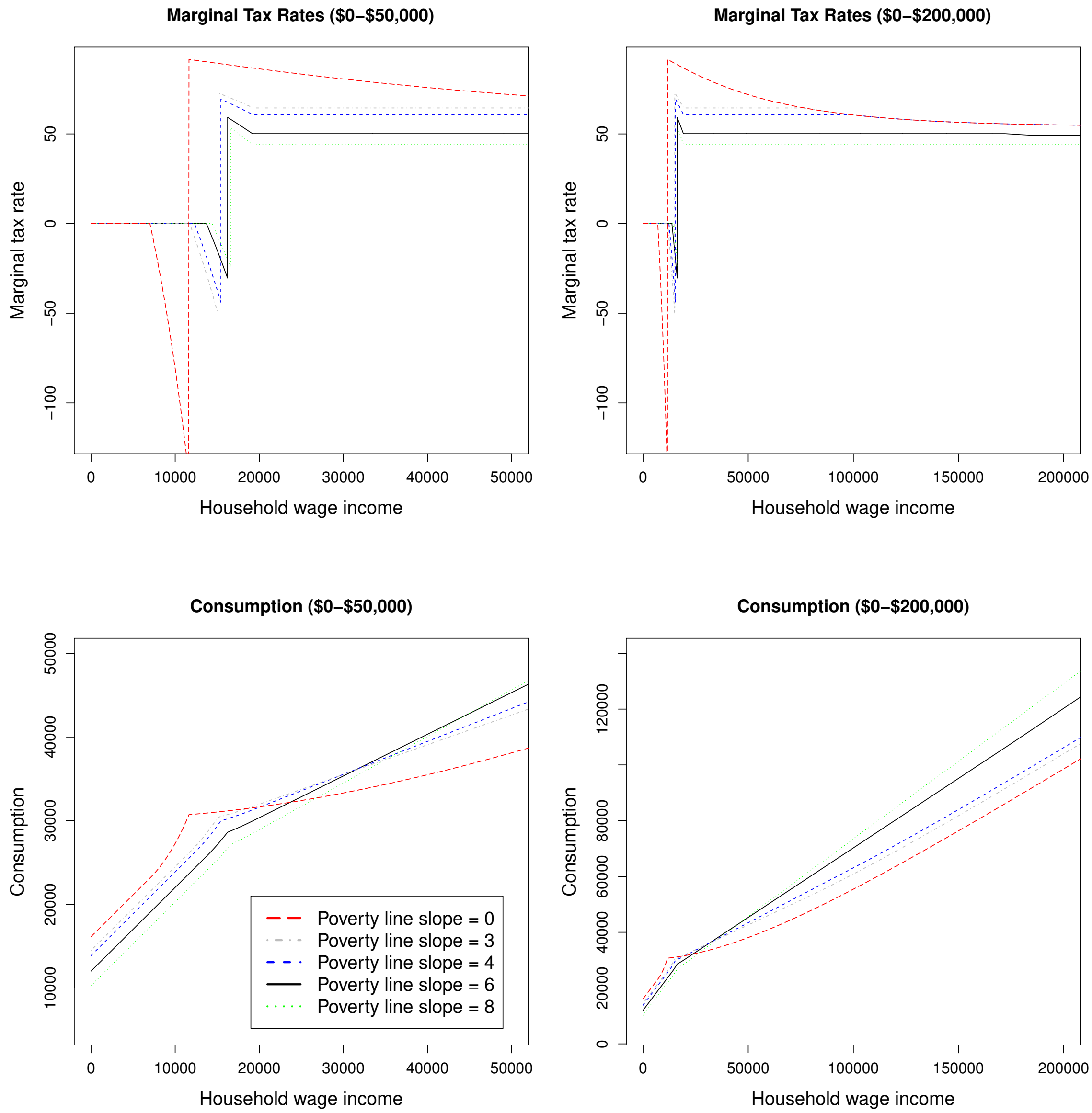

Figure 17: Most redistributive R - comparative statics - different slopes of the poverty line (with $\overline{P}_0 = \$18,850$). Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn). $\tilde{\theta}$ is chosen so that the average slope of PL is x \$/h with $x \in \{0, 3, 4, 6, 8\}$.

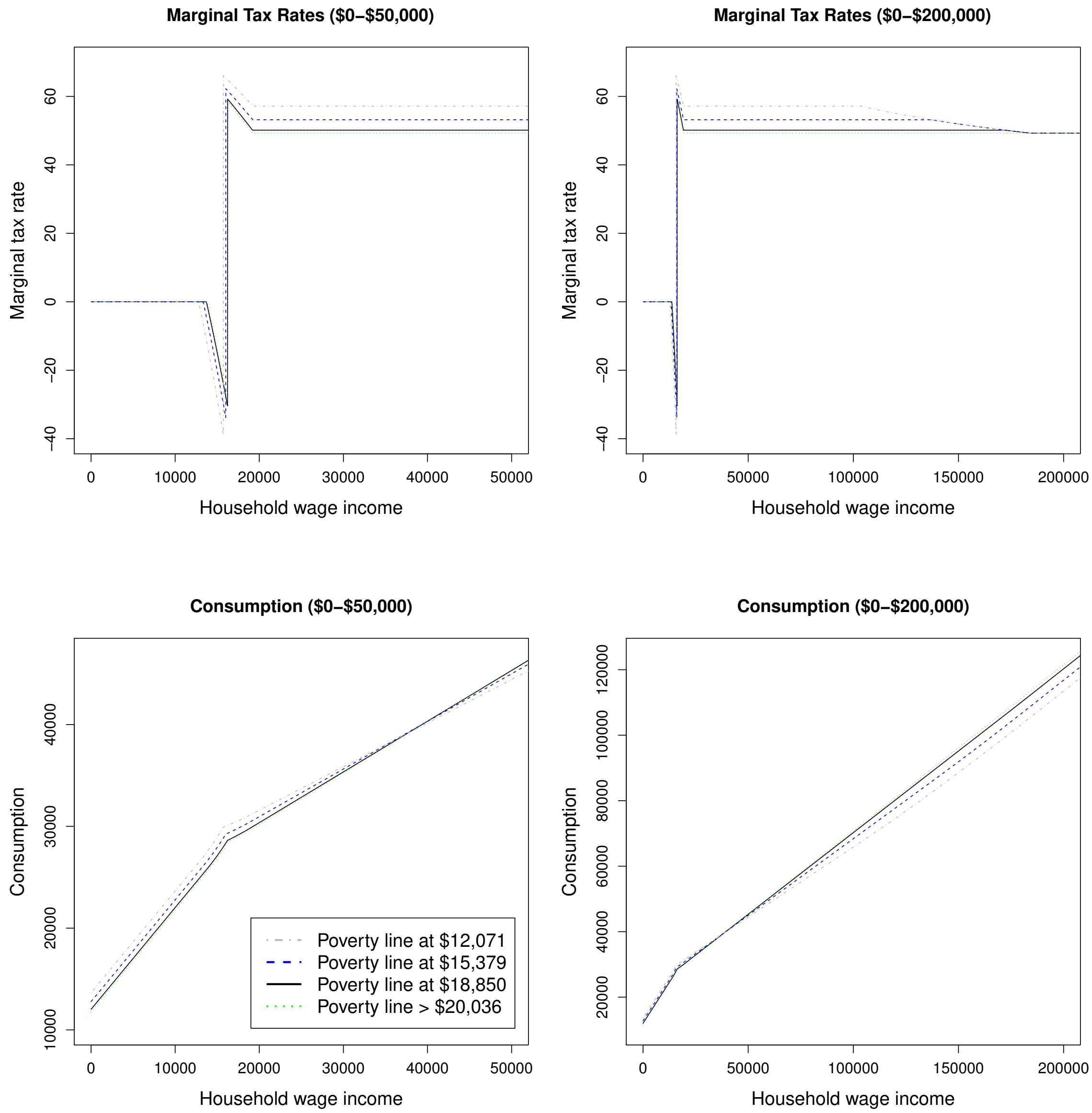

Figure 18: Most redistributive R - comparative statics - different levels of the poverty line (with an average slope of $\$6/h$). Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn). The values of the poverty line are the 2014 official poverty lines in the US for households of size one to three members respectively.

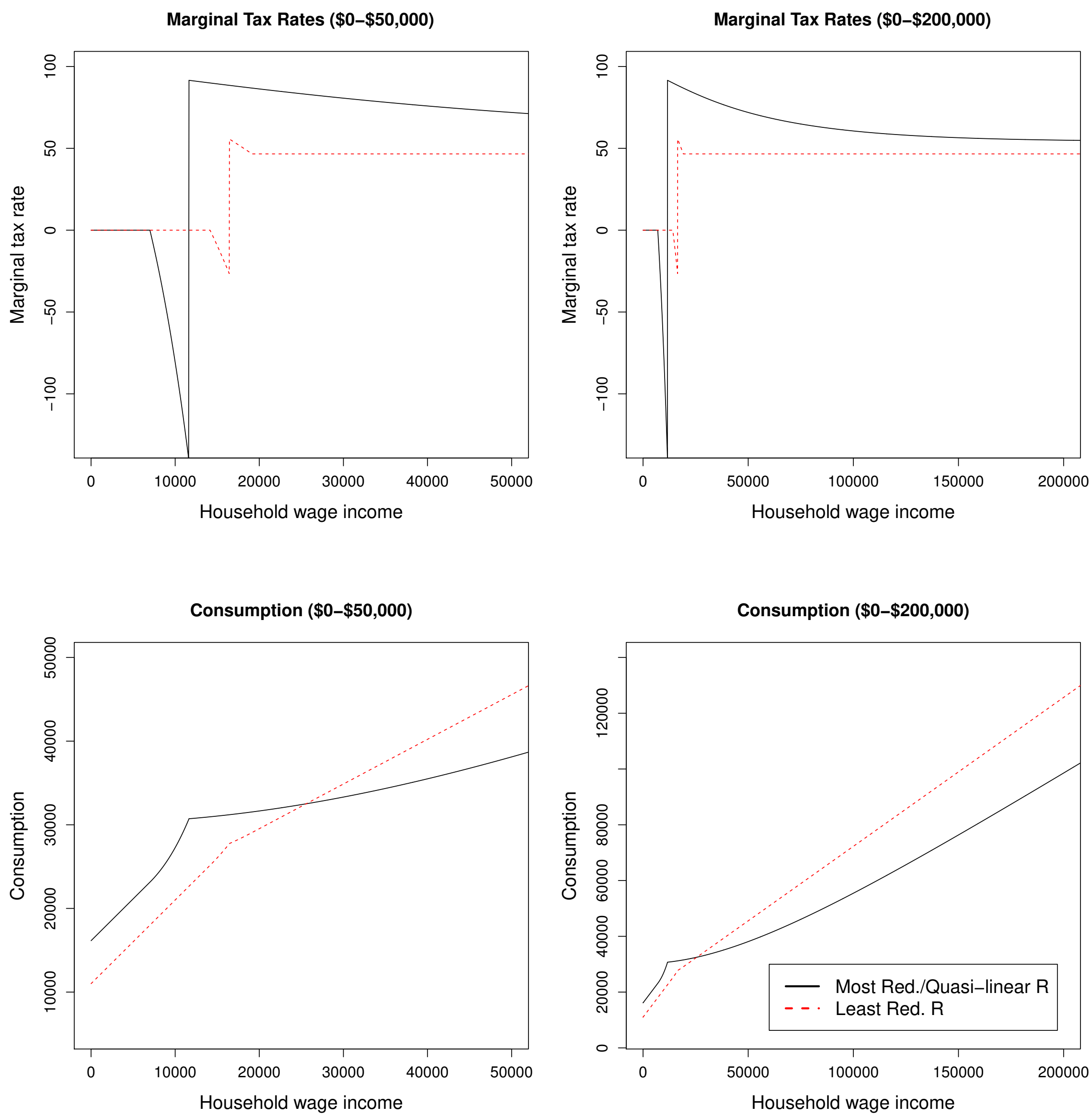

Figure 19: Flat poverty line ($0/h) - Comparing the most-, least-redistributive and quasi-linear $\tilde{R}$ (when $\overline{P}_0 = \$18,850$). Most redistributive and quasi-linear $\tilde{R}$ lead to identical solutions. Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\overline{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn).

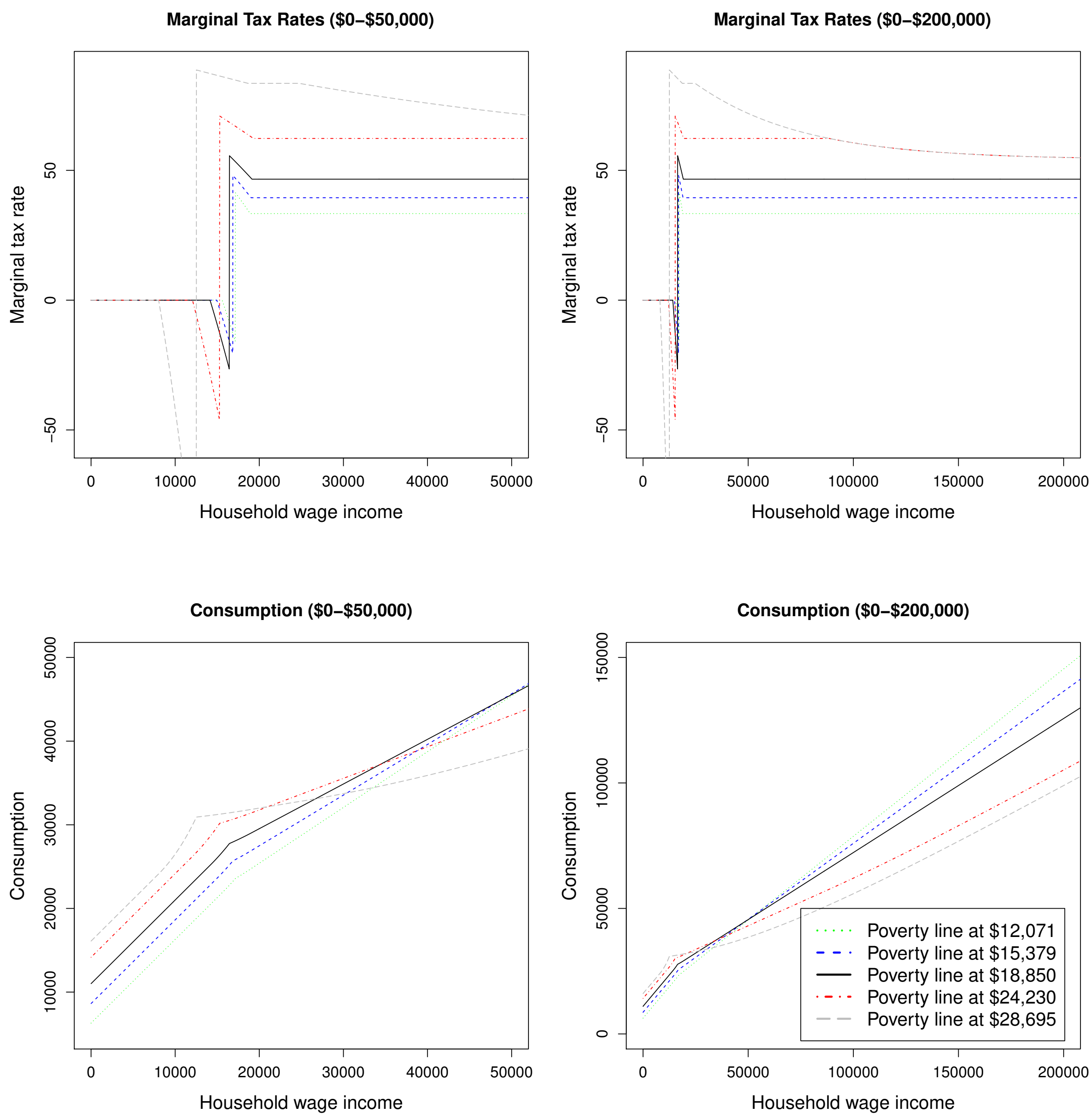

Figure 20: Flat poverty line ($0/$h$)- comparative statics - different levels of the poverty line. Labor supply elasticity is $\epsilon = 0.33$. The preference parameter of the most hardworking agents ($\bar{\theta} = 222.28$) is calibrated using hours worked in a year by hourly workers in the CPS. The distribution of types n is estimated (using CPS data) via Maximum likelihood and assuming a Double Pareto Lognormal distribution (DPLn). The values of the poverty line are the 2014 official poverty lines in the US for households of size one to five members respectively.

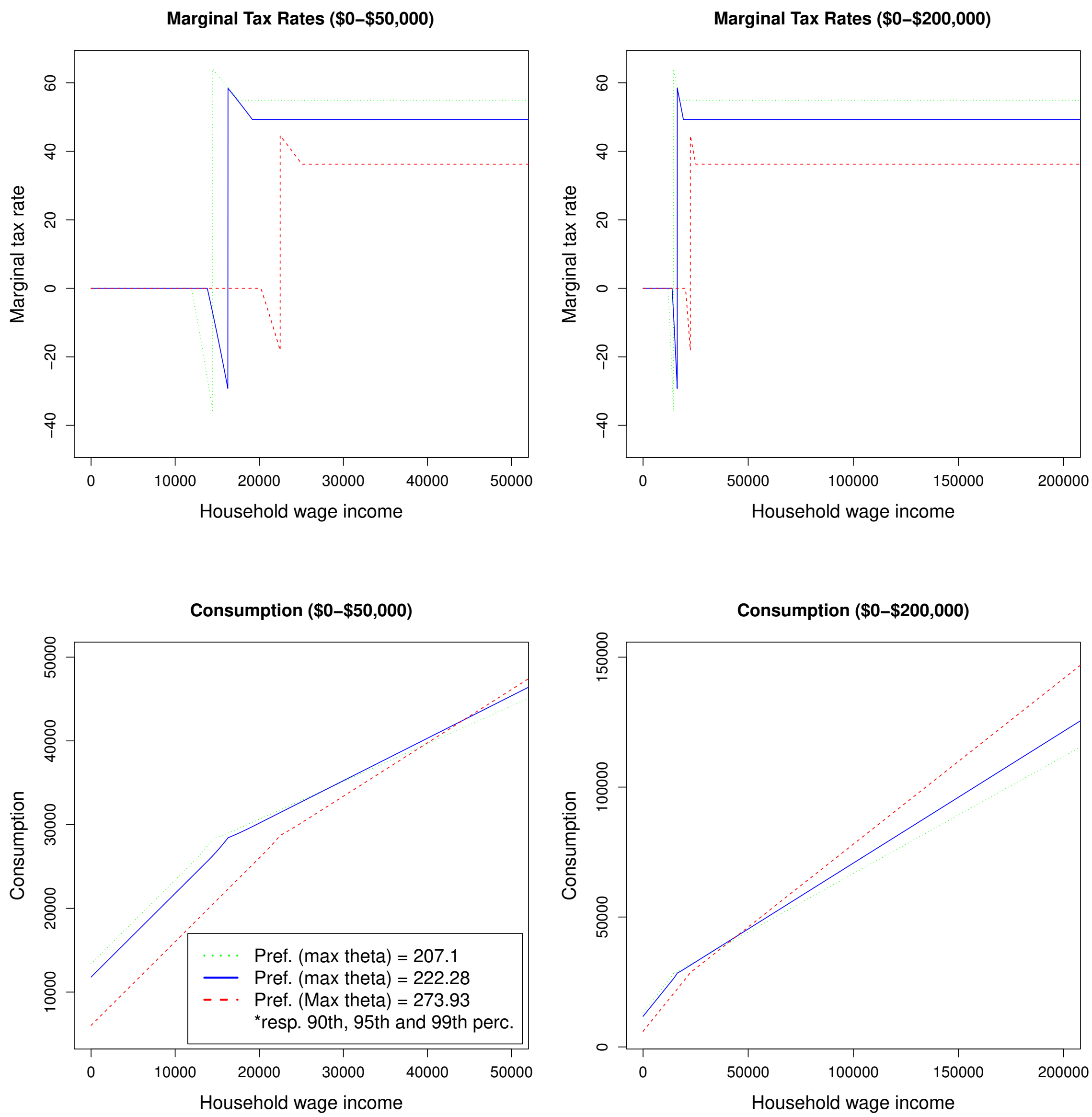

Figure 21: Quasi-linear $\tilde{R}$ - comparative statics - different values of $\overline{\theta}$ (the preference parameter for the most hardworking among minimum wage agents). We show $\overline{\theta} \in \{207.1, 222.28, 273.93\}$. These values are, respectively, the 90^{th}, 95^{th} and 99^{th} percentiles of the distribution of estimated θ in the CPS sample of hourly workers. θ is estimated by inverting the first-order conditions of agents' labor supply decision, taking into account their after-tax hourly wage rate (more details in appendix F). All calibrations in this figure are made under the assumption that the poverty line has an average slope of \$6/hour and the elasticity of labor supply is 0.33. The distribution of types n is estimated (using CPS data) via Maximum likelihood and we assume a Double Pareto Lognormal distribution (DPLn).

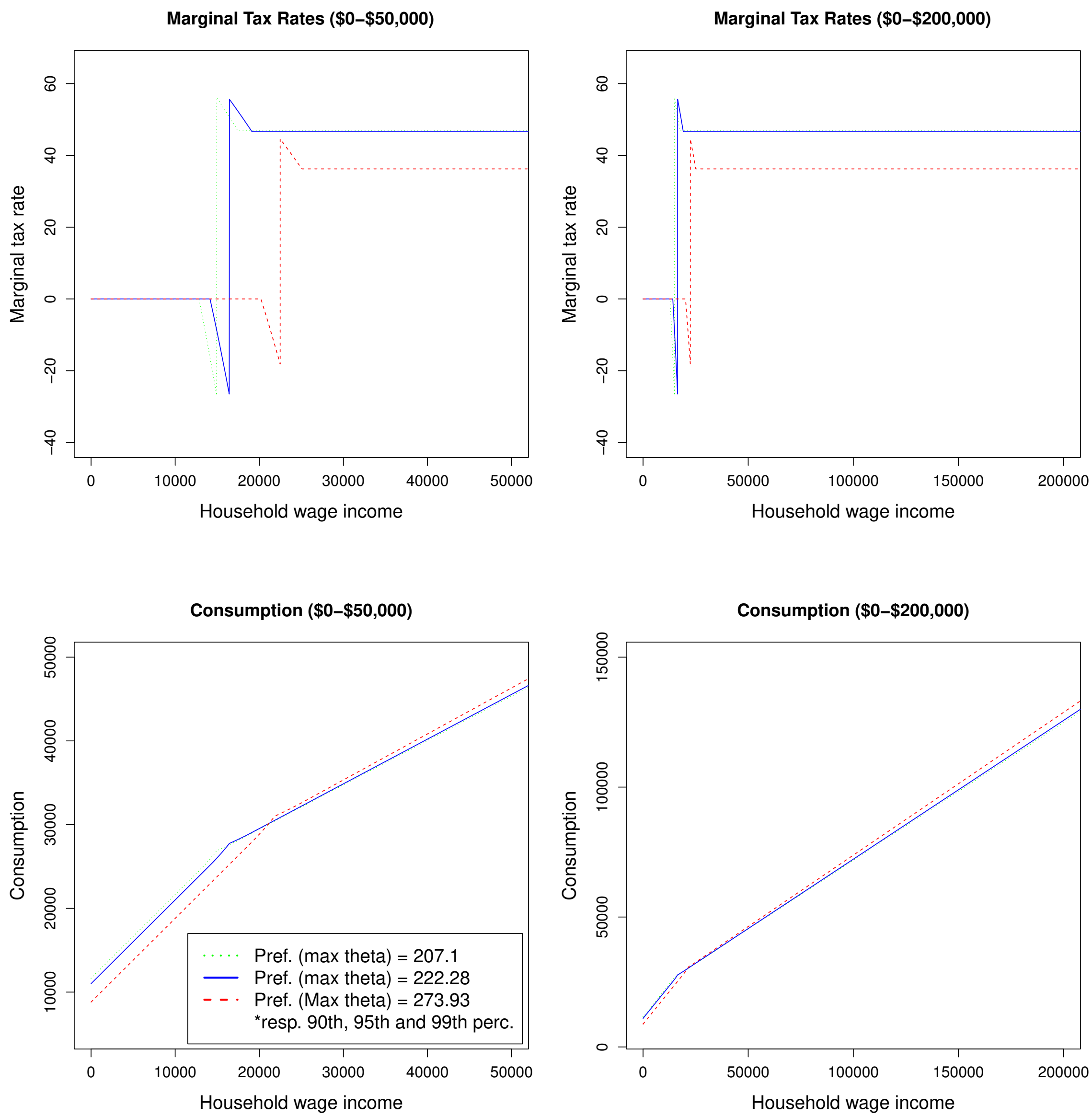

Figure 22: Least redistributive $\tilde{R}$ - comparative statics - different values of $\overline{\theta}$ (the preference parameter for the most hardworking among minimum wage agents). We show $\overline{\theta} \in \{207.1, 222.28, 273.93\}$. These values are, respectively, the 90^{th}, 95^{th} and 99^{th} percentiles of the distribution of estimated θ in the CPS sample of hourly workers. θ is estimated by inverting the first-order conditions of agents' labor supply decision, taking into account their after-tax hourly wage rate (more details in appendix F). All calibrations in this figure are made under the assumption that the poverty line has an average slope of \$6/hour and $\overline{P}_0 = \$18,850$. The elasticity of labor supply is 0.33. The distribution of types n is estimated (using CPS data) via Maximum likelihood and we assume a Double Pareto Lognormal distribution (DPLn).

Table 2: Calibrated Optimal Tax Schedule

	$c(0)$ (\$)	$y(n_u)$ (\$)	y_1^* (\$)	$y(n_l)$ (\$)	$T'(y_1^*-\varepsilon)$	$T'(y_1^*+\varepsilon)$	$T'(y(n_l))$	$y(n')$ (\$)	$T'(y\to\infty)$
$\overline{P}_0=\$18,850$; Slope $=\$6/h$									
Most Red. $\tilde{R}$	12,025	13,711	16,223	19,187	−0.30	0.59	0.50	∞	0.49
Quasi-linear $\tilde{R}$	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	∞	0.49
Least Red. $\tilde{R}$	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	∞	0.47
$\overline{P}_0=\$15,379$; Slope $=\$6/h$									
Most Red. $\tilde{R}$	12,778	13,326	16,015	19,228	−0.34	0.63	0.53	∞	0.49
Quasi-linear $\tilde{R}$	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	∞	0.49
Least Red. $\tilde{R}$	8,632	14,951	16,859	19,022	−0.21	0.48	0.39	∞	0.39
$\overline{P}_0=\$18,850$; Slope $=\$4/h$									
Most Red. $\tilde{R}$	13,884	12,300	15,438	19,302	−0.44	0.69	0.61	99,740	0.54
Quasi-linear $\tilde{R}$	13,297	12,743	15,692	19,277	−0.39	0.66	0.58	131,848	0.54
Least Red. $\tilde{R}$	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	∞	0.47
Quasilinear $\tilde{R}$; slope =									
$\$0-\$1.51/h$	16,144	6,997	11,643	.	−1.41	0.92	.	11,643	0.54
$\$3/h$	14,505	11,719	15,093	19,317	−0.50	0.73	0.64	76,804	0.54
$\$4/h$	13,297	12,743	15,692	19,277	−0.39	0.66	0.58	131,848	0.54
$\$6/h$	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	∞	0.49
$\$8/h$	10,293	14,410	16,587	19,100	−0.25	0.53	0.44	∞	0.44
Least Red. $\tilde{R}$; sl. $=\$6/h$; $\overline{P}_0=$									
\$12,071	6,295	15,587	17,169	18,923	−0.16	0.41	0.33	∞	0.33
\$15,379	8,632	14,951	16,859	19,022	−0.21	0.48	0.39	∞	0.39
\$18,850	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	∞	0.47
$>\$20,036$	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	∞	0.49
Most Red. $\tilde{R}$; $\overline{P}_0=\$18,850$; sl. =									
$\$0/h$	16,144	6,997	11,643	.	−1.41	0.92	.	11,643	0.54
$\$3/h$	15,178	10,860	14,558	19,295	−0.61	0.77	0.70	56,594	0.54
$\$4/h$	13,884	12,300	15,438	19,302	−0.44	0.69	0.61	99,740	0.54
$\$6/h$	12,025	13,711	16,223	19,187	−0.30	0.59	0.50	∞	0.49
$\$8/h$	10,293	14,410	16,587	19,100	−0.25	0.53	0.44	∞	0.44
Most Red. $\tilde{R}$; sl. $=\$6/h$; $\overline{P}_0=$									
\$12,071	13,645	12,795	15,721	19,273	−0.39	0.66	0.57	∞	0.49
\$15,379	12,778	13,326	16,015	19,228	−0.34	0.63	0.53	∞	0.49
\$18,850	12,025	13,711	16,223	19,187	−0.30	0.59	0.50	∞	0.49
$>\$20,036$	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	∞	0.49
$\overline{P}_0=\$18,850$; slope $=\$0/h$									
Most Red./Quasi-linear $\tilde{R}$	16,144	6,997	11,643	.	−1.41	0.92	.	11,643	0.54
Least Red. $\tilde{R}$	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	∞	0.47
Least Red. $\tilde{R}$; slope $=\$0/h$, $\overline{P}_0=$									
\$12,071	6,295	15,587	17,169	18,923	−0.16	0.41	0.33	∞	0.33
\$15,379	8,632	14,951	16,859	19,022	−0.21	0.48	0.39	∞	0.39
\$18,850	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	∞	0.47
\$24,230	14,162	12,058	15,296	19,311	−0.47	0.71	0.62	88,673	0.54
\$28,695	16,097	8,031	12,522	18,669	−1.13	0.89	0.83	24,749	0.54
Quasi-linear $\tilde{R}$; sl. $=\$6/h$, $\overline{P}_0=\$18,850$									
$\overline{\theta}=207.10$	13,338	11,921	14,463	17,514	−0.36	0.64	0.55	203,815	0.54
$\overline{\theta}=222.28$	11,794	13,819	16,280	19,175	−0.29	0.58	0.49	∞	0.49
$\overline{\theta}=273.93$	6,023	20,195	22,482	25,063	−0.18	0.45	0.36	∞	0.36
Least Red. $\tilde{R}$; sl. $=\$6/h$, $\overline{P}_0=\$18,850$									
$\overline{\theta}=207.10$	11,656	12,828	14,949	17,420	−0.27	0.56	0.47	∞	0.47
$\overline{\theta}=222.28$	11,009	14,144	16,450	19,135	−0.27	0.56	0.47	∞	0.47
$\overline{\theta}=273.93$	8,806	18,926	21,849	25,265	−0.25	0.54	0.45	∞	0.45

Note: Each panel presents different comparative statics corresponding to figures 6 and 13-22. The second column, $c(0)$, is the level of consumption for agents with zero income. The fourth column, y_1^*, is the income level at which there is a discrete jump in marginal tax rates. The third, fourth, and fifth columns indicate the endpoints of the intervals over which rates are decreasing. The sixth and seventh column report the marginal tax rates just before and right after y_1^*. The fifth and ninth columns report the endpoints of the interval over which $\underline{b}(u(n);n)$ is equalized and the marginal tax rate is constant. The latter is reported in column eight. Finally, the ninth column is the income level at which - if applicable - marginal tax rates are determined by efficiency considerations alone. The last column indicates tax rates in the upper tail.

F Calibration Details and Description of the Current US Income Tax Schedule

The same sample is used and the same sample restrictions are applied to estimate the parameters of the distribution of types, calibrate $\overline{\theta}$, estimate the net revenue raised by the government at the status-quo and the current total marginal tax rate faced by households of different structures along the income distribution.

We pool CPS data from 2013-2015 and drop all households in group quarters and households with negative or zero income. We also drop households with any self-employment or farming income because of measurement error concerns. We drop households with any retirement income or whose head is above 60. Finally, we also drop married couples that are not of legal age to be married (because it might create issues in the TAXSIM simulator).

Marginal federal tax rate and welfare receipt We consider that married couples (with spouse present) are joint filers while other households are single filers. We feed this information along with the age, wage and salary income of each spouse, state of residence and year into NBER's TAXSIM. We also infer the number of dependent exemptions claimed by the head of households from the family relationship and household structure. We obtain estimates, for each household, of the federal tax liability, federal marginal tax rate, payroll tax liability (fica) and marginal payroll tax rate. The federal marginal tax rate is defined as the sum of these two tax rates and, of course, includes the EITC and Child Tax Credit. A smooth tax schedule - as a function of household income - is then estimated using a kernel estimator.

In order to obtain a more accurate description of the labor supply incentives that households face, we also include a measure of average cash welfare receipt at each point in the wage distribution. We sum the supplemental security income (SSI), Temporary Assistance for Needy Families (TANF) and value of food stamps (SNAP) reported by each household in the CPS. We adjust for underreporting of welfare receipt in the CPS using the dollar reporting rate for each program (and specific to the CPS) documented in Meyer, Mok and Sullivan (2015). The reporting rates can be found in appendix tables 3,4 and 8 of their study. We average the values for the four most recent years available (2009-2012). As a result, each respondent's reported welfare receipt is divided by 0.8758 for SSI, 0.5706 for SNAP and 0.573 for TANF/welfare income. Finally, we use a local kernel regression to estimate the average relationship between welfare receipt and wage income. We take the derivative to estimate the phaseout of these programs and obtain an additional implicit marginal tax that we add to the marginal federal tax rate in our descriptions of the current tax schedule.

Below are graphs depicting the resulting estimates of marginal tax rates along the income distribution.

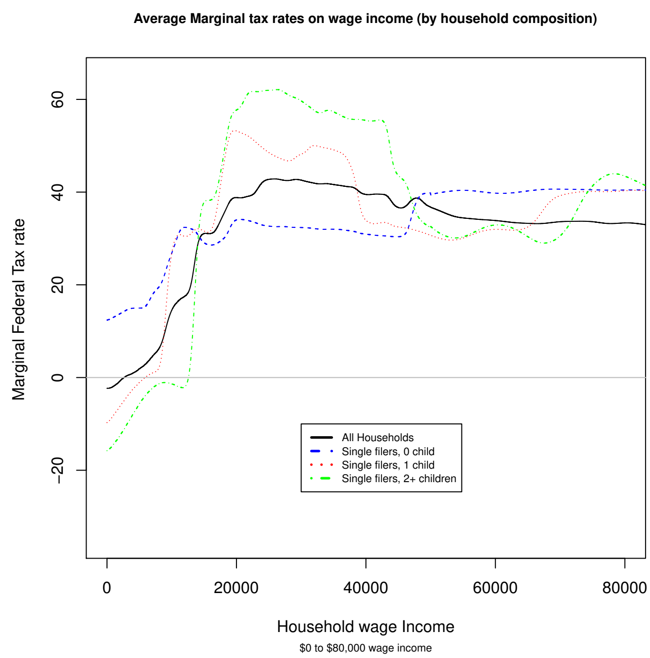

Figure 23: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under "Marginal federal tax rate and welfare receipt".

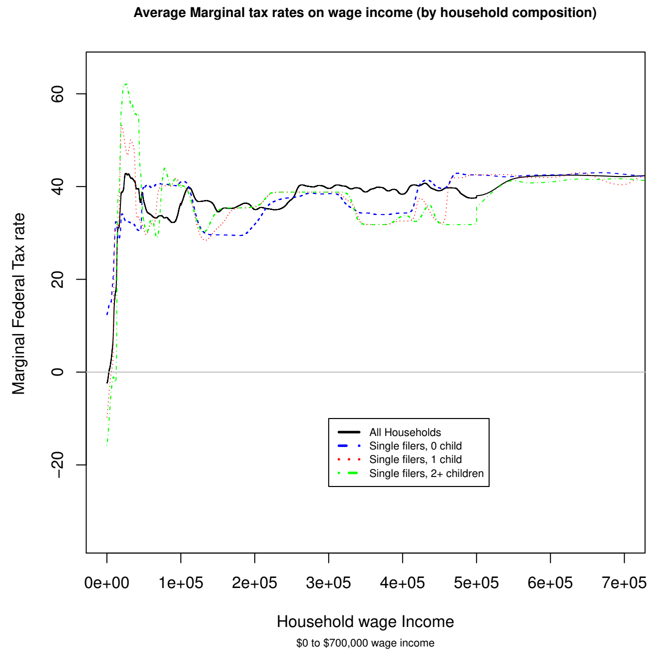

Figure 24: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under “Marginal federal tax rate and welfare receipt”.

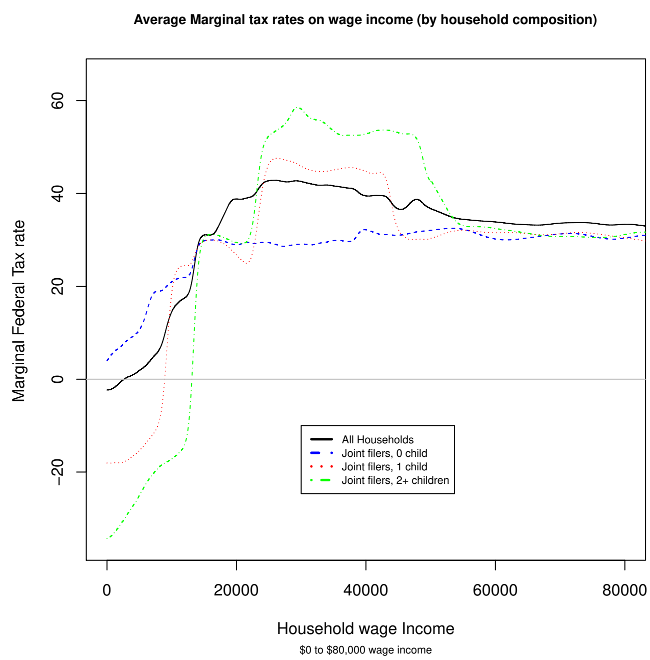

Figure 25: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under “Marginal federal tax rate and welfare receipt”.

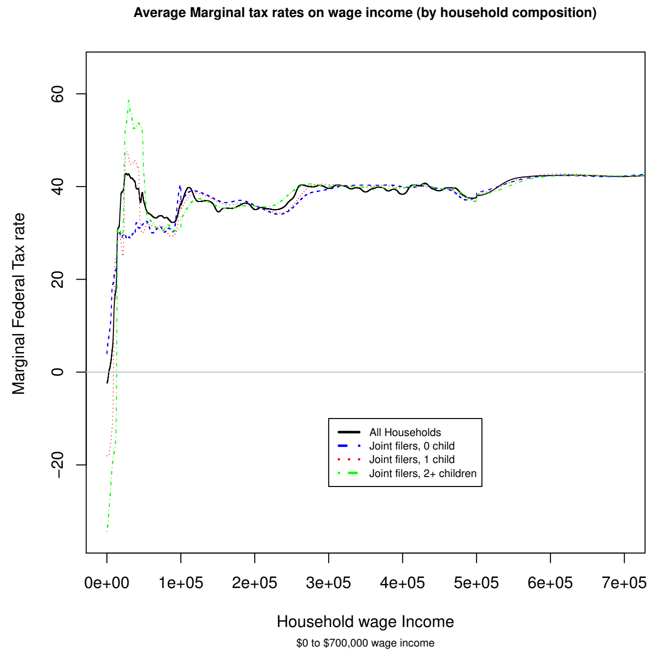

Figure 26: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under "Marginal federal tax rate and welfare receipt".

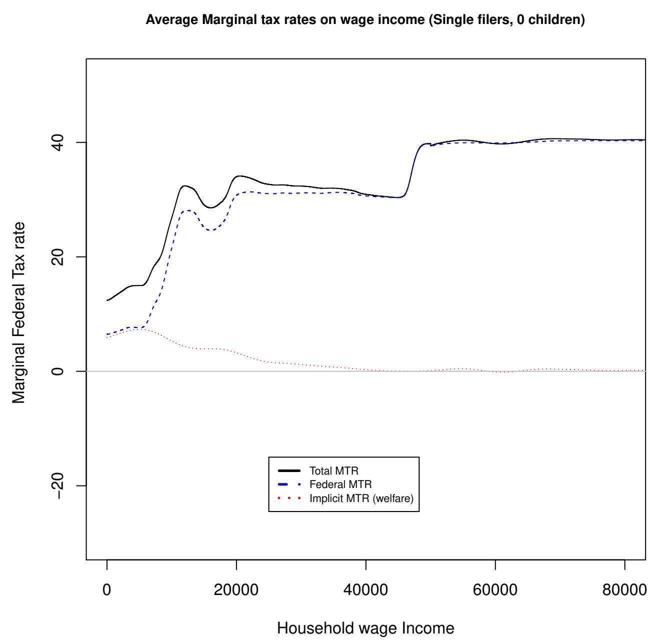

Figure 27: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under "Marginal federal tax rate and welfare receipt".

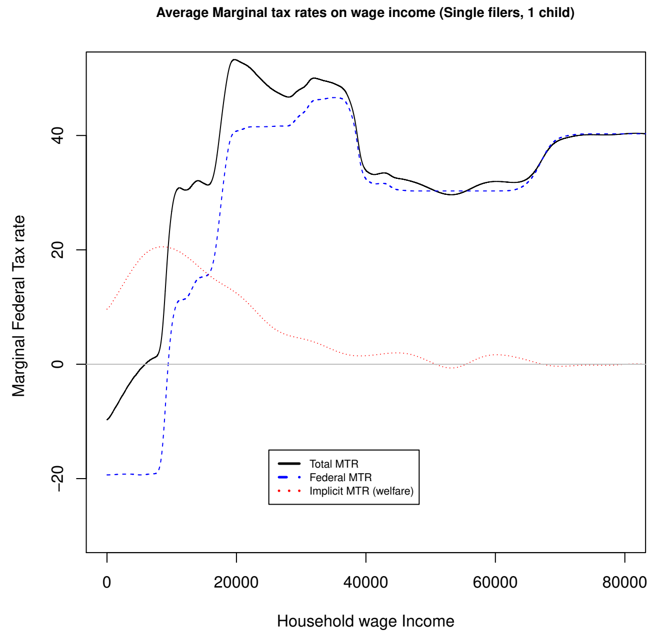

Figure 28: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under "Marginal federal tax rate and welfare receipt".

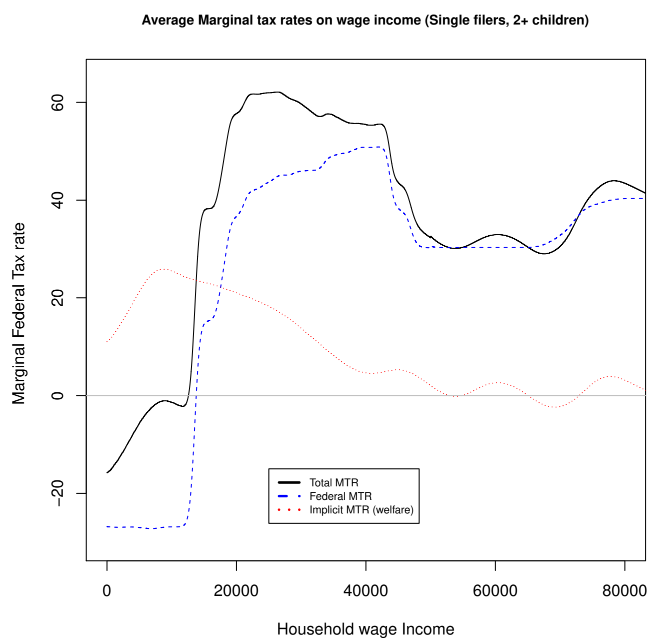

Figure 29: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under "Marginal federal tax rate and welfare receipt".

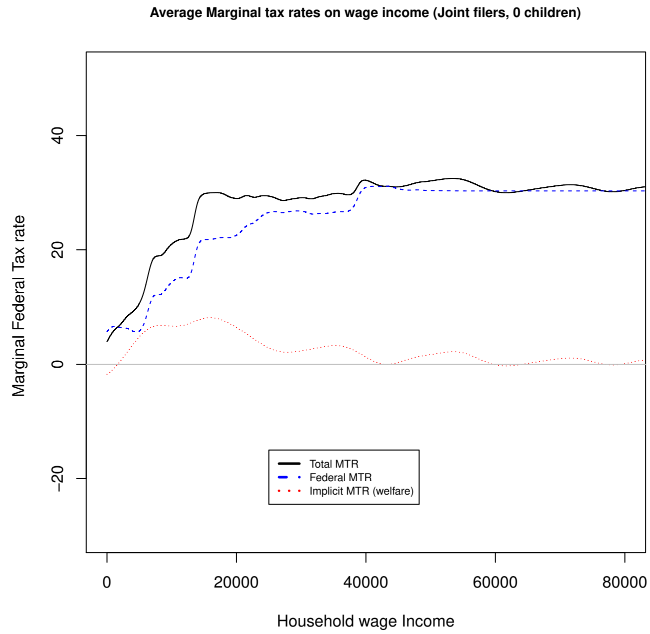

Figure 30: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under “Marginal federal tax rate and welfare receipt”.

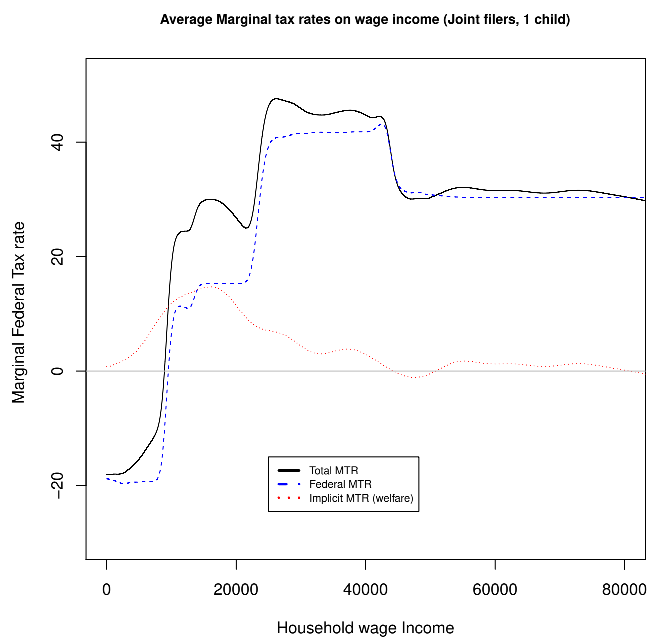

Figure 31: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under “Marginal federal tax rate and welfare receipt”.

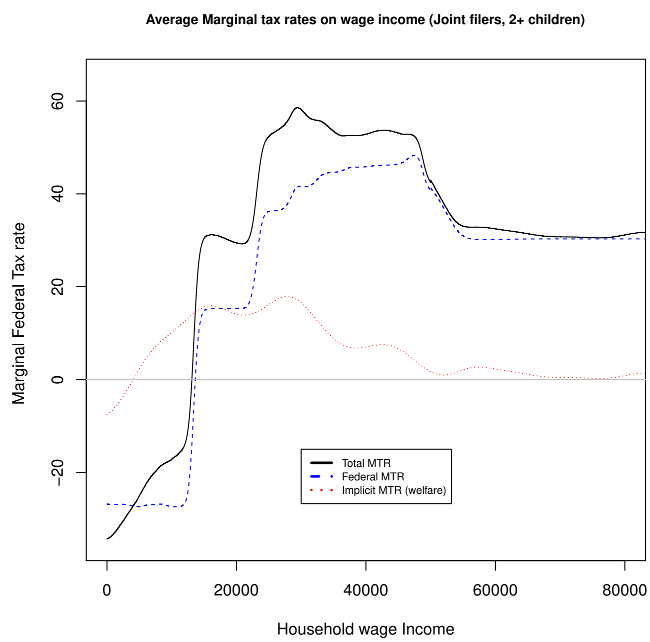

Figure 32: Authors own calculation. Source of data: CPS 2013-2015 and NBER TAXSIM. More details about the calculation are provided in appendix F under "Marginal federal tax rate and welfare receipt".

Estimation of the distribution of types Because 14.9% of households in the original sample have an income of \$0, we assume that %14.9 of households have $n = 0$. For the remaining %85.1, the distribution of types is estimated via Maximum Likelihood using the same sample as for the description of the current tax schedule. This guarantees consistency in our comparisons between optimal and current tax rates. We estimate a four parameter Double Pareto Lognormal distribution (DPLn).[22] As shown by Reed (2003), the DPLn distribution can be derived from a stochastic model for the generation of observed income distributions, and more importantly, fits observed income distributions well in many countries (including in the US).[23]

For each household in the data, we simulate the total marginal federal tax rate using the procedure described above. Conditional on a labor supply elasticity parameter (we choose $\epsilon = 0.33$), the marginal tax rate and wage income of the household, we can recover type n_i by inverting the first-order condition:

$$n_i = (1 - \tau_i)^{-\frac{\epsilon}{1+\epsilon}} y_i^{\frac{1}{1+\epsilon}}$$

Inverting the first-order condition for each observation in the sample, we obtain the empirical distribution of types n_i. The parametrized distribution is then estimated by maximizing the log-likelihood of this distribution of n_i. Using the notation of Hajargasht and Griffiths (2013), the estimated parameters are: $m = 8.56$ (s.e. 0.003), $\sigma = 0.34$ (s.e. 0.003), $\alpha = 3.42$ (s.e. 0.033) and $\beta = 1.55$ (s.e. 0.006).[24] As evidenced by the small standard errors, the four parameters are precisely estimated. In addition, the DPLn distribution fits the data quite well. Figures 33 and 34 compare the calibrated cumulative distribution and probability distribution functions, respectively, to their empirical counterparts. Figure 35 compares the calibrated local

[22]The Double Pareto Lognormal distribution was introduced by Reed (2003). Also see Hajargasht and Griffiths (2013) for an empirical application.

[23]In our case, we estimate the distribution of types, not income. However, the distribution of incomes will inherit important features of the distribution of types. For instance, an upper-tail Pareto parameter α of the distribution of types will lead to a Pareto parameter $\alpha_{income} = \frac{\alpha}{1+\epsilon}$ of the distribution of incomes when the marginal tax rate is constant on large incomes (as is the case both under current policy and in our calibrations). Importantly, we show below that the DPLn distribution fits our sample distribution of n_i well.

[24]The reported standard errors are asymptotic standard errors. They are estimated by inverting the Hessian of the log-likelihood function evaluated at the solution. See p.479-480 of Wooldridge (2010). Intuitively, α governs the behavior of the distribution in the upper-tail and β governs the lower-tail. When $(\alpha, \beta) \to \infty$, the DPLn converges to a lognormal with location parameter m and scale parameter σ.

Pareto parameter to its empirical counterpart. The local Pareto parameter of the distribution of n is an important component of the so-called "Rawlsian tax rate", the tax rate that maximizes revenue raised by the government. Recall that, according to proposition 1, the Rawlsian tax rate applies on large incomes if the lower-bound on utilities (constraint (10)) is not binding in that part of the distribution. Under a DPLn distribution, the local Pareto parameter, $\frac{1-F(n)}{nf(n)} \to \frac{1}{\alpha}$ as $n \to \infty$. Consequently, parameter α governs the thickness of the upper-tail of the distribution of types n.

Despite the good empirical fit in our sample of CPS data, the calibrated upper-tail of the income distribution (under current actual tax rates) is slightly thinner than estimated in previous studies. We discuss the reasons for this discrepancy in Section G in the appendix and perform a robustness check to verify the sensitivity of our main results to assuming a thicker upper-tail.

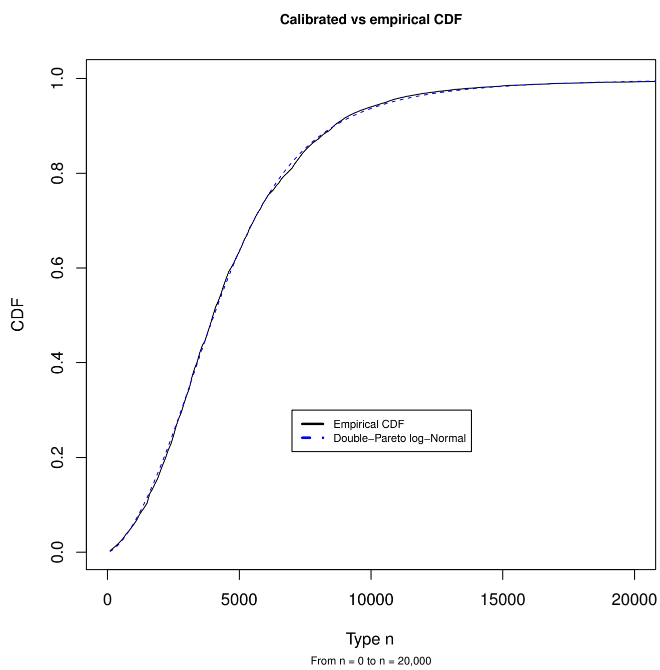

Figure 33: Double Pareto Lognormal estimated on CPS data (2013-2015) versus the empirical CDF. For each head of household i in the sample, n_i is obtained by inverting the first-order conditions (conditioning on the estimated actual marginal tax rate at y_i). We fit a Double Pareto Lognormal distribution to the sample distribution of n_i. The calibrated parameters are $m = 8.56$, $\sigma = 0.34$, $\alpha = 3.42$ and $\beta = 1.55$.

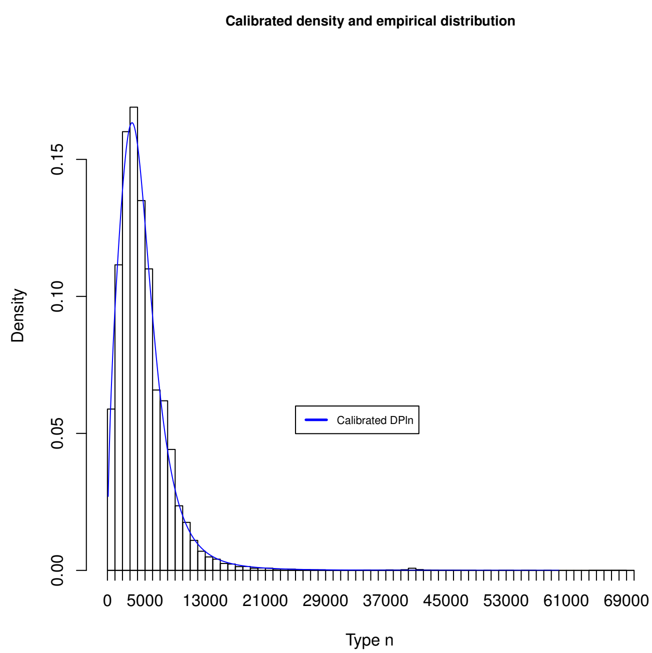

Figure 34: Double Pareto Lognormal estimated on CPS data (2013-2015) versus empirical histogram. For each head of household i in the sample, n_i is obtained by inverting the first-order conditions (conditioning on the estimated actual marginal tax rate at y_i). We fit a Double Pareto Lognormal distribution to the sample distribution of n_i. The calibrated parameters are $m = 8.56$, $\sigma = 0.34$, $\alpha = 3.42$ and $\beta = 1.55$.

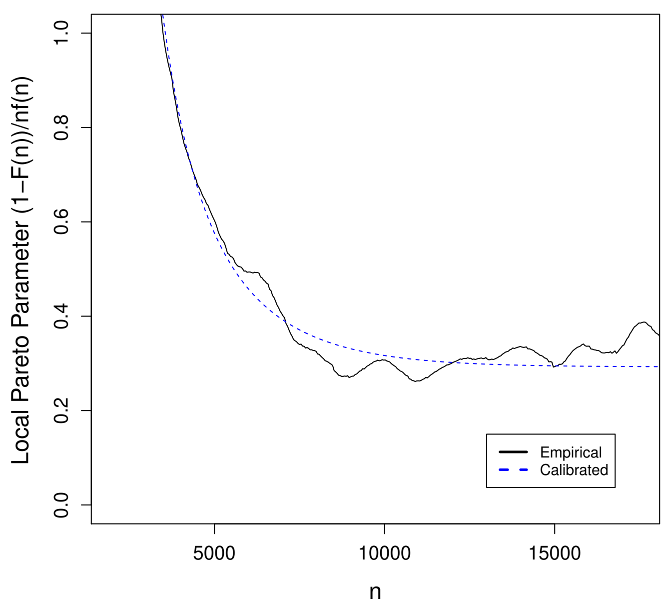

Figure 35: Comparison of the calibrated local Pareto parameter of the distribution of n, $\frac{1-F(n)}{nf(n)}$, with its empirical counterpart. For each head of household i in the sample, n_i is obtained by inverting the first-order conditions (conditioning on the estimated actual marginal tax rate at y_i). The empirical local Pareto parameter is computed using the empirical CDF of n_i and a kernel density estimator of its PDF. The calibrated distribution of n is obtained by fitting a Double Pareto Lognormal distribution to the sample distribution of n_i. The calibrated parameters are $m = 8.56$, $\sigma = 0.34$, $\alpha = 3.42$ and $\beta = 1.55$. The value of the local Pareto parameter is above 1 for values below $n = 3487$.

Calibrating the preference parameters We pick a labor supply elasticity of $\epsilon = 0.33$ which is the preferred estimate in Chetty (2012). We assume that the minimum wage in the economy is equal to the federal minimum wage $\underline{w} = \$7.25$. Finally, we also need to calibrate $\overline{\theta}$. We restrict the CPS sample to workers that were asked whether they work hourly, answered yes and reported their hourly wage (9.51% of the sample).[25] We use the reported hours usually worked per week at all jobs and the number of weeks worked last year to infer labor supply.[26]

Conditional on a labor supply elasticity of $\epsilon = 0.33$, we can infer θ_i for everyone in the sample using the following (inverted) first-order condition:

$$\theta_i = [(1 - \tau_i)\, w_i]^{-\frac{\epsilon}{1+\epsilon}}\, h_i^{\frac{1}{1+\epsilon}}$$

We use for each agent in the sample their projected total number of hours worked over the year (h_i) and reported hourly wage (w_i). We also impute the marginal tax rate (τ_i) faced by each household using the method described above. The target parameter is $\overline{\theta}$ - the preferences of the minimum wage agent choosing the longest hours. Because the highest order statistics is especially sensitive to measurement error, we pick the value corresponding to a percentile close to the top. Our main results use the 95^{th} percentile of the sample distribution of θ_i to calibrate $\overline{\theta} = 222.28$. We have also considered the sensitivity of our results to using the 90^{th} ($\overline{\theta} = 207.1$) or 99^{th} ($\overline{\theta} = 273.93$) percentile of the sample distribution (see Fig. 21 and 22). The sample distribution of the estimated preference parameter (θ_i) is shown in Fig. 36.

Calibrating government revenue We assume that the government must raise $\$12,152$ on average per working household to finance public goods. This was estimated again in our sample of CPS data so that our calibrated optimal tax schedule raises the same amount of government revenue as current policy in the US. $\$12,152$ is the average net simulated federal tax (Federal tax + payroll tax - welfare receipt) paid in the sample. This estimate does not include disability payments (because we also exclude them from our estimates of current marginal tax rates in the US) and excludes households with zero income. Recall that the sample also excludes households

[25] About 17.46% of the sample of households used for the calibration of the income distribution indicate whether they are paid hourly or not. 9.51% of the sample responds yes and proceed to reporting their hourly wage rate. We keep only those workers who indicate being paid hourly and whose hourly wage is reported. Sample size is $17,747$.

[26] The results are very similar when we use reported "hours usually worked per week last year".

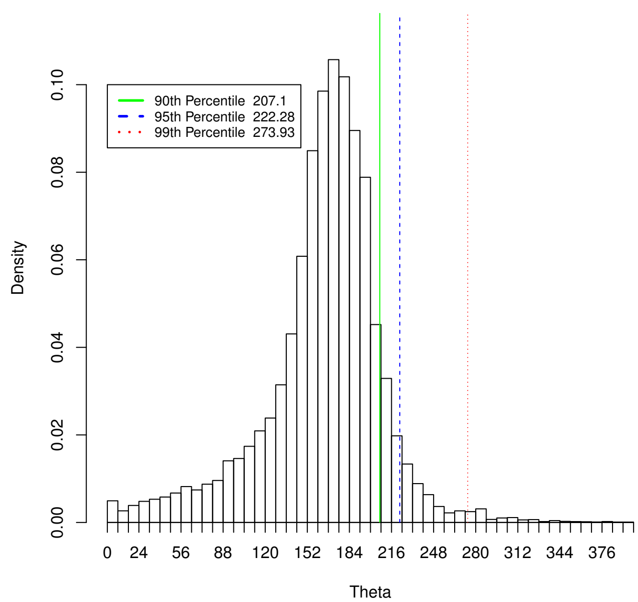

Figure 36: Histogram of the estimated preference parameter (θ_i) among hourly workers (Source: CPS 2013-2015). We estimate θ_i by inverting the first-order conditions of each hourly worker's labor supply decision (taking the after-tax wage into account). Most calibrations use $\overline{\theta} = 222.28$ to calibrate the preference parameter of the most hardworking minimum wage household. We also consider the sensitivity of our results to choosing $\overline{\theta} \in \{207.1, 273.93\}$.

whose head is above 60 years old.

Because our model does not include a fixed cost (or disutility) of working, there is no extensive margin. Yet, our calibrations have to be consistent with the fact that 14.9% of households in the original sample have an income of \$0. In the CPS sample, they receive an average of \$4,086 in welfare (this number includes only food stamps, SSI, TANF and disability benefits). In all calibrations, we assume that there is a mass of %14.9 of households with $n = 0$. They therefore have zero income in all calibrations. We assume that they receive $-\tau(0) = -\lim_{n \to 0^+} \tau(n)$. Across all calibrations considered, the minimum value of $-\tau(0)$ is \$6,295 which is larger than average welfare receipt for non-working households in the CPS data (\$4,086).

Calibrating the "normative" parameters (slope and level of the poverty line) $\tilde{\theta}$ governs the slope of the poverty line. We pick $\tilde{\theta}$ by choosing an average slope for the poverty line between a labor time of 0 and $2250 = 50 \times 45$ hours. To calibrate the level of the poverty line, we consider five different values (all evaluated at a labor time of zero): \$12,071, \$15,379, \$18,850, \$24,230 and \$28,695. These are the weighted average 2014 official poverty line thresholds in the US for families of one to five people respectively.[27]

G Discussion of the estimated upper-tail Pareto index and robustness to a thicker tail

It is worth discussing our estimate of α. Because the marginal tax rate (under current US policy) in the upper-tail is constant and given the first-order conditions of agents' behavior, the estimated $\alpha = 3.42$ translates into a Pareto tail for the income distribution of $\alpha_{income} = \frac{\alpha}{1+\epsilon} = 2.57$. This is a slightly larger value (hence a thinner tail) than conventional estimates using administrative data for the US. For instance, Atkinson, Piketty and Saez (2011) estimate a Pareto index of 1.5 in the US. As a result, our estimated revenue-maximizing tax rate on top incomes (0.54) is slightly lower than conventional estimates. The revenue-maximizing rate on top incomes corresponding to a Pareto index (for the income distribution) of 1.5 and a labor supply elasticity of 0.33 is 0.67. The discrepancy is likely due to two factors. First, consistent

[27] Source: DeNavas-Walt and Proctor (2015).

with our model, we invert types n based on wage and salary income only (we exclude other sources of income such as dividends, capital gains,... that likely make the tail thicker). Second, the CPS data is sparse in the upper-tail. Moreover, Bollinger, Hirsch, Hokayem and Ziliak (2019) find evidence of under-reporting of earnings in the upper-tail in the CPS.

This raises the question of how our results are affected by this potential under-estimate of the thickness of the upper-tail. To answer this question, we have performed the following robustness check. We have fixed α (for the distribution of n) to a value of 2 (which translates to an upper-tail Pareto index of 1.5 for the income distribution). We have re-estimated the three other parameters of the distribution of types (m, σ and β) by maximizing the log-likelihood of the CPS sample of n_i (obtained by inversion of the first-order conditions). In addition, we have re-estimated the average net revenue raised per working household (S in the resource constraint equation (4)) to account for the fact that we are missing high income agents in the CPS sample. Indeed, we calibrate S so that average net revenue raised per household in the calibrations is equal to current policy in the US (as measured in our CPS sample). If we assume that large earnings are missing in the CPS sample, a corollary is that we also under-estimate the average net revenue raised by the current tax system. In order to re-calibrate S, we have calculated the average net tax paid by each household as a function of type n under current policy in the US. Using the newly estimated distribution of n (with tail parameter fixed at $\alpha = 2$), we average the net tax paid by each household. Finally, using the new "corrected" distribution of n and estimate of S, we have re-calibrated our optimal tax formulas.

Table 3 below shows the results of this exercise. It compares to table 1 in the main text. We can see that our main calibrations are remarkably robust to assuming a thicker Pareto tail of the distribution of types. There are mainly two reasons for this. First, in the main calibrations, the slope of the poverty line ($\$6/h$) is such that the tax rate in the upper-tail is not the revenue-maximizing rate but is the one that equalizes well-being across n (0.49 for quasi-linear preferences). As a result, the tax rate at the top is not affected by the change in Pareto index. Second, even though there is added mass in the upper-tail under the new calibration, and hence additional resources in the economy to be potentially re-distributed, we have also re-adjusted the estimated average net revenue raised per household under the status-quo (S) to account for the missing mass in the upper-tail (in the CPS sample). This is important, for the sake of consistency, as we compare potential reforms to the status-quo. As a result, some of the additional revenue that is raised by the tax scheme under the new calibration of types is spent on a larger S instead of being redistributed. In the end, conform to intuition, the new calibration accounting for a thicker tail does lead to a larger surplus, beyond poverty eradication, to be redistributed. In turn, it leads to a slightly larger discrepancy between the three polar choices of $\tilde{R}$. But the difference is small. It is likely that larger discrepancies between the alternative calibrations would appear if we consider flatter poverty lines (especially if the poverty line is flat enough that the tax rate at the top is the revenue-maximizing one).

Table 3: Robustness - Calibrated Optimal Tax Schedule under $\alpha = 2$

$\overline{P}_0 = \$15,379$	$c(0)$ (\$)	$y(n_u)$ (\$)	y_1^* (\$)	$y(n_l)$ (\$)	$T'(y_1^* - \varepsilon)$	$T'(y_1^* + \varepsilon)$	$T'(y(n_l))$	$T'(y \to \infty)$
Most Red.	13,987	13,195	15,943	19,243	−0.35	0.63	0.54	0.49
Quasi-linear	12,681	13,821	16,281	19,178	−0.29	0.58	0.49	0.49
Least Red.	8,518	14,873	16,820	19,037	−0.21	0.49	0.40	0.40
$\overline{P}_0 = \$18,850$								
Most Red.	13,133	13,623	16,176	19,200	−0.31	0.60	0.51	0.49
Quasi-linear	12,681	13,821	16,281	19,178	−0.29	0.58	0.49	0.49
Least Red.	11,146	14,253	16,506	19,124	−0.26	0.55	0.46	0.46
$\overline{P}_0 \geq \$20,923$								
All	12,681	13,821	16,281	19,178	−0.29	0.58	0.49	0.49

Note: Each panel presents the calibrated optimal tax according to each choice of $\tilde{R}$ (most redistributive, quasi-linear and least redistributive) for poverty lines at, respectively, \$15,379, \$18,850, and \$24,230. These values correspond to weighted averages of the 2015 official US poverty lines for households of size 2, 3, and 4. The average slope of the poverty line is calibrated at \$6/h. The second column, $c(0)$, is the level of consumption for agents with zero income. The fourth column, y_1^*, is the income level at which there is a discrete jump in marginal tax rates. The third, fourth, and fifth columns indicate the endpoints of the intervals over which rates are decreasing. The sixth and seventh column report the marginal tax rates just before and right after y_1^*. The tax rate on incomes larger than $y(n_l)$ is reported in column eight. The last column indicates the marginal tax rate in the tail.

H Solving for the optimal tax function under the least- and most-redistributive $\tilde{R}$

The shape of the reference preferences affects the shape of the lower-bound on utilities (constraint (10) in appendix D). As the surplus over the allocation that just eradicates poverty increases, so does $\underline{b}(u(n_1); n_1)$ at the optimum. Graphically, it means that agent 1's implicit budget is tangent to a higher reference indifference curve. Figure 5 sketches the indifference curves above PL for the least- and most-redistributive $\tilde{R}$.

Let us analyze each case, least- and most-redistributive, separately below.

H.1 Least Redistributive Preferences

In the least redistributive case, higher indifference curves are characterized by a lower $\tilde{\theta}$ (they are steeper). As $\underline{b}(u(n_1); n_1)$ increases, $\tilde{\theta}$ decreases. As a result, the marginal tax rate along the lower-bound becomes:

$$1 - \tau'(y(n)) = 1 \qquad \forall n \in [\underline{n}, n_1) \tag{32}$$

$$1 - \tau'(y(n)) = \left[1 - \left(\frac{\tilde{\theta}'}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \qquad \forall n \in [n_1, \infty) \tag{33}$$

for some $\tilde{\theta}' < \tilde{\theta}$ (if there is a surplus to be redistributed). More precisely, $\tilde{\theta}'$ is defined such that

$$u(n_1) + \frac{1}{1+\epsilon}\left[\left(\underline{w}\tilde{\theta}'\right)^{1+\epsilon} - (n_1)^{1+\epsilon}\right] = \overline{P}_0. \tag{34}$$

Notice that $u(n_1)$ is, at the optimum, implicitly a function of the aggregate resources in the economy. Tax rates along the lower-bound (equations (32) and (33)) are derived as in Section D.4 in the appendix. Formulae (17) and (18) still describe the optimal tax function. The formula of Proposition 1 still applies and describes the optimal tax function but replacing $\tilde{\theta}$ with $\tilde{\theta}'$.
As the resource constraint is increased (if revenue to be raised S decreases for instance), there is more surplus to be redistributed to everyone, including n_1, so $u(n_1)$ increases. We can see from equation (34) that it leads to a lower $\tilde{\theta}'$. The same holds for a lower level of the poverty line $\overline{P}_0$. Recall that a lower $\tilde{\theta}$ leads to lower-marginal tax rate along the lower-bound. This result is illustrated by the calibrations, for the least redistributive $\tilde{R}$, in online appendix figure 16, where a lower level of the poverty line leads to lower levels of redistribution and lower tax rates.

H.2 Most Redistributive Preferences

In the most redistributive case, higher indifference curves (above the PL) rotate around (c'', ℓ''). The portion of the indifference curves over $[0, \ell'']$ are characterized by a higher $\tilde{\theta}'' > \tilde{\theta}$ (which leads to flatter indifference curves). There is a kink at (c'', ℓ''), then the remaining portion of the indifference curve follows the poverty line. This is illustrated in figure 5. As $\underline{b}(u(n_1); n_1)$ increases, $\tilde{\theta}''$ increases. The marginal tax rate that equalizes well-being becomes:

$$1 - \tau'(y(n)) = 1 \qquad \forall n \in [\underline{n}, n_1) \tag{35}$$

$$1 - \tau'(y(n)) = \left[1 - \left(\frac{\tilde{\theta}''}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \qquad \forall n \in [n_1, n_0] \tag{36}$$

$$1 - \tau'(y(n)) = \left[1 - \frac{\ell''}{\overline{\overline{\theta}} n^{\epsilon}}\right]^{\frac{1}{1+\epsilon}} \qquad \forall n \in [n_0, n''] \tag{37}$$

$$1 - \tau'(y(n)) = \left[1 - \left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \qquad \forall n \in [n'', \infty) \tag{38}$$

Though the formula is defined piecewise on $[n_1, \infty)$, the marginal tax rate along the lower-bound is continuous over that interval because:

$$\left[1 - \left(\frac{\tilde{\theta}''}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} = \left[1 - \frac{\ell''}{\overline{\overline{\theta}} n_0^{\epsilon}}\right]^{\frac{1}{1+\epsilon}} \tag{39}$$

and

$$\left[1 - \frac{\ell''}{\overline{\overline{\theta}} n''^{\epsilon}}\right]^{\frac{1}{1+\epsilon}} = \left[1 - \left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}} \tag{40}$$

These two equalities define the values of n_0 and n''. Intuitively, for all budget lines with slope $w \leq \frac{n_0}{\overline{\overline{\theta}}}$, the budget is tangent to the reference indifference curve on its $[0, \ell'')$ portion (where the marginal rate of substitution is a function of $\tilde{\theta}''$). For all slopes $\frac{n_0}{\overline{\overline{\theta}}} < w < \frac{n''}{\overline{\overline{\theta}}}$, the budget is tangent to the reference indifference curve at the kink in ℓ''. Finally, all implicit budgets with a slope $w \geq \frac{n''}{\overline{\overline{\theta}}}$ are evaluated by $\tilde{R}$ on its $[\ell'', \infty)$ portion where the marginal rate of substitution is a function of $\tilde{\theta}$. For this reason, marginal tax rates (along the lower-bound) over $[n_1, n_0]$ and $[n'', \infty)$ look familiar. In order to obtain the tax rate over interval $[n_0, n'']$, notice that the reference indifference curves essentially compare the level of consumption at ℓ'' when evaluating the implicit budgets. As a result, $\frac{d\,\underline{b}(u(n);n)}{dn} = u'(n) - n^{\epsilon} + \frac{\ell''}{\overline{\theta}}$. For the sake of completeness, setting $\frac{d\,\underline{b}(u(n);n)}{dn} = 0$, we obtain the derivative of utility with respect to n:

$$u'(n) = n^{\epsilon} \qquad \forall n \in [\underline{n}, n_1) \tag{41}$$

$$u'(n) = \left[1 - \left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\epsilon} \qquad \forall n \in [n_1, n_0) \tag{42}$$

$$u'(n) = n^{\epsilon} - \frac{\ell''}{\overline{\theta}} \qquad \forall n \in [n_0, n'') \tag{43}$$

$$u'(n) = \left[1 - \left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\epsilon} \qquad \forall n \in [n'', \infty) \tag{44}$$

The marginal tax rate along the lower-bound is found by plugging the first-order conditions of the incentive compatibility constraint in equations (41)-(44). Finally, $\tilde{\theta}''$ is defined such that

$$\overline{P}_0 = \underline{u\,(n'')} + \frac{1}{1+\epsilon}\left[\left(\underline{w}\tilde{\theta}''\right)^{1+\epsilon} - (n'')^{1+\epsilon}\right] \tag{45}$$

$$= u(n_1) + \int_{n_1}^{n_0}\left[1 - \left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\epsilon} dx + \int_{n_0}^{n''}\left\{x^{\epsilon} - \frac{\ell''}{\overline{\theta}}\right\} dx$$
$$+ \frac{1}{1+\epsilon}\left[\left(\underline{w}\tilde{\theta}''\right)^{1+\epsilon} - (n'')^{1+\epsilon}\right] \tag{46}$$

where $\underline{u\,(n'')}$ is the well-being index, along the lower-bound, for agents with $n = n''$. Equation (45) comes from the fact that the portion of the reference indifference curve over $[n'', \infty)$ coincides with the poverty line. As a result, the utility of agents with n'' (along the lower-bound) must be equal to their utility at the poverty line even as the (societal) surplus over the allocation that eradicates poverty increases. In practice, at the optimum, the actual well-being of agents with $n = n''$ may be larger if the lower-bound is not binding at the solution. Indeed, the full solution to the optimal tax problem is described by formulae (17) and (18). Whenever $\mu(n) > 0$, the lower-bound on utilities is binding and marginal tax rates are described by formulae

(35)-(38). Similar to Proposition 1, there are intervals where $\mu(n) = 0$. Assuming, as in Proposition 1, that $F(n)$ has a Pareto index α in the upper tail and that the lower-bound on utilities binds on (at most) one interval above y_1^*, then for $n_u \leq n_{\underline{b}} \leq n_l \leq n' \in [\underline{n}, \infty)$, $\tau'(y(n))$ is such that

1. $\tau'(y(n)) = 0\ \forall n \in [\underline{n}, n_u]$
2. $\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left(\frac{1+\epsilon}{\epsilon}\right) \frac{F(n_u)-F(n)}{nf(n)} \quad \forall n \in [n_u, n_{\underline{b}}]$
3. $y(n) = y_1^* \quad \forall n \in [n_{\underline{b}}, n_{\bar{b}}]$
4. $\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left(\frac{1+\epsilon}{\epsilon}\right) \left[\frac{F(n_l)-F(n)}{nf(n)}\right] + \frac{\tau'(y(n_l))}{1-\tau'(y(n_l))} \frac{n_l f(n_l)}{nf(n)}\ \forall n \in [n_{\bar{b}}, n_l]$
5. If $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} \geq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$: then
   The marginal tax rate is defined by equations (36)-(38) $\forall n \in [n_l, \infty)$ (with n_0, n'' defined by equations (39)-(40))
6. Otherwise,
   The marginal tax rate is defined by equations (36)-(38) $\forall n \in [n_l, n']$ (with n_0, n'' defined by equations (39)-(40)) and

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = \frac{1-F(n)}{nf(n)} \left(\frac{1+\epsilon}{\epsilon}\right) \quad \forall n \in [n', \infty)$$

   and the marginal tax rate converges to $\tau'(y(n)) \rightarrow \frac{1+\epsilon}{1+\epsilon+\alpha\epsilon}$ as $n \rightarrow \infty$.

As the surplus increases (or the poverty line decreases), $u(n_1)$ increases, and $\tilde{\theta}''$ increases as well leading to larger marginal tax rates (on $[n_1, n'')$). Hence there is more redistribution from higher skill agents towards everyone else. This result is illustrated by the calibrations, for the most redistributive $\tilde{R}$, in figure 18, where a lower level of the poverty line leads to higher levels of redistribution and higher tax rates. Finally, in practice, in the calibrations, we choose n'' as the 95^{th} percentile of the calibrated distribution of n. We find the corresponding ℓ'' and n_0 using equations (39) and (40).

I Details and Technical Derivations for the Proof of Proposition 1

I.1 Why is there bunching at n_1?

In this subsection, we explain why the constraint that $y(n)$ be increasing is likely to be binding at n_1. The reasoning is illustrated in Figure 37. Let us assume as a preliminary step, that we solve V^L and V^U without taking care of the last constraint on $y(n)$. Then, the solution to V^L would be that $\underline{b}(u(n);n) = b^*$ and $\tau'(y(n)) = 0$ for all $n \leq n_1$, and $y(n_1) = y_1^L$. Indeed, a zero marginal tax rate below y_1^L guarantees that $\underline{IB}(\underline{z}_j, \underline{R}_j) = \underline{IB}(\underline{z}_k, \underline{R}_k)$ for all j,k such that $w_j = w_k = \underline{w}$, so that $b(z_j, w_j, R_j) = b(z_k, w_k, R_k) = b(z_1, w_1, R_1)$. Moreover, all agents getting the same $\tau(y(n))$ is clearly the cheapest way of providing them with well-being level b^*.

However, this tax cannot be optimal if one adds the constraint that $y(n) \leq y^*(n_1)$. Indeed, if we look now at the V^U program without the constraint on $y(n)$, the solution would be that from some y_1^U the optimal τ would guarantee that $\underline{b}(u(n);n) = b^*$ for all n in some interval $[n_1, n']$.[28] As illustrated in the figure, as soon as $y_1^U \neq y_1^L$, that cannot be the global solution.

The optimal $y^*(n_1)$ that links the solution to both subproblems must be such that $y_1^U \leq y^*(n_1) \leq y_1^L$ and the optimal τ is like in the figure: all agents of type $n \in [\underline{n}, n_u]$ face a zero marginal tax rate, and, among them, all agents with the minimal wage have exactly the same well-being as agent 1. All agents of type $n \in [n_u, n_{\underline{b}}]$ face a negative marginal tax rate, and they all have a strictly larger well-being level than agent 1. All agents of type $n \in [n_{\underline{b}}, n_1]$ earn the same income as agent 1 but all have a strictly larger well-being.

[28]The upper bound of this interval would be the limit above which the well-being constraint is no longer binding and the optimal tax is derived from efficiency considerations only or $n' = \infty$ if the marginal tax rate along the well-being constraint is lower in the tail than the revenue maximizing tax rate (see Proposition 1).

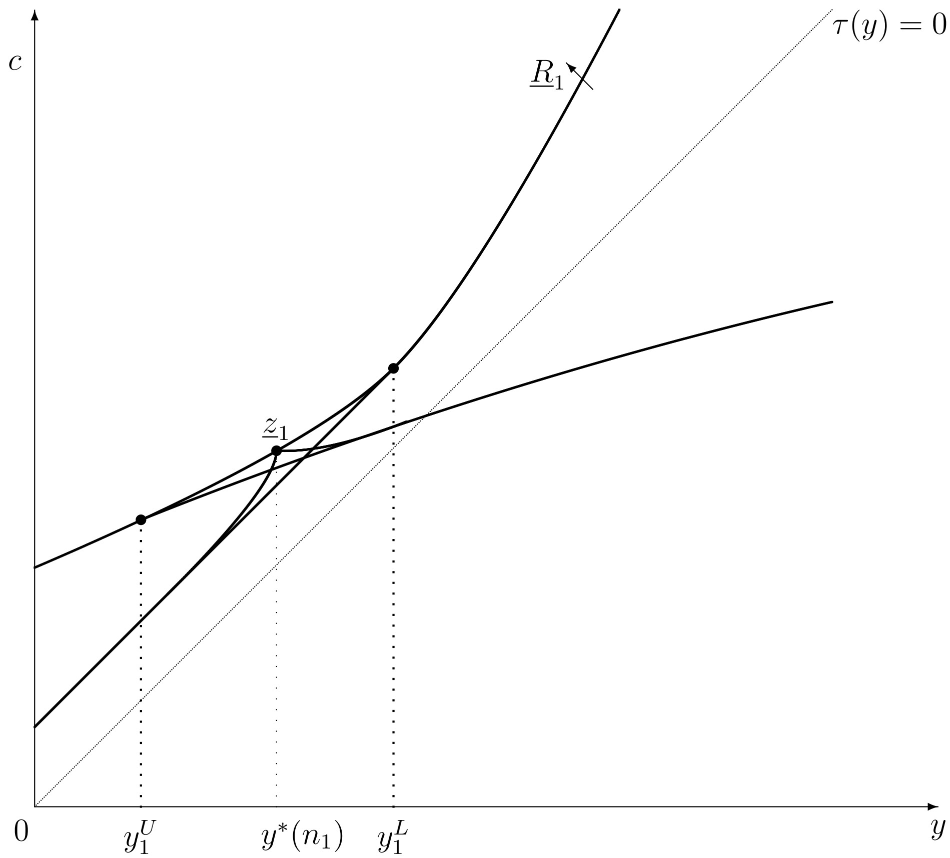

Figure 37: Illustration of bunching at $y^*(n_1)$ in Proposition 1. The optimal path of the state variable $u(n)$ is anchored at n_1 which is in the interior of $[\underline{n}, \infty)$. The full problem can be separated into two *subproblems*: one solving for the optimal income and utility path below n_1 and one solving for the optimal path above n_1. Ignoring the constraint that $y(n)$ must be increasing in n at n_1, the solution of the first *subproblem* is a zero marginal tax rate everywhere in $[\underline{n}, n_1)$ and the solution of the second *subproblem* implies some positive (in general non-constant) marginal tax rate for all $[n_1, \infty)$. But this implies a discrete jump in the marginal tax rate at n_1 with $y_1^L > y_1^U$ which violates the constraint that $y(n)$ be increasing at n_1. The optimal $y^*(n_1)$ that links the solutions to both subproblems must be between y_1^L and y_1^U and it generates some bunching at $y^*(n_1)$ for agents with $n \in [n_{\underline{b}}, n_{\bar{b}}]$ where $n_{\underline{b}} < n_1 < n_{\bar{b}}$. The solution also implies zero marginal tax rates for $n \in [\underline{n}, n_u]$ and negative marginal tax rates for $[n_u, n_{\underline{b}}]$.

I.2 A description of thresholds n_u, $n_{\underline{b}}$, $n_{\bar{b}}$, n_l and n'

In proposition 1, the solution to the optimal tax problem is presented. It is characterized "piecewise": different marginal tax rates apply to different intervals. In this section, we attempt to give an intuitive explanation for why these intervals appear and how their endpoints (the "thresholds") are determined. In the online appendix subsections I.3.2 and I.3.4, we show how to formally solve for their values.

The SWF is a leximin in well-being index $b(z_i, w_i, R_i)$. As a result, the social objective is to equalize $b(z_i, w_i, R_i)$. In a first-best context (when incentive compatibility constraints do not need to be satisfied), the solution is such that $b(z_i, w_i, R_i)$ is indeed equalized across everyone. In second-best, the solution will be such that $\underline{b}(u(n); n)$ is equalized unless either i) it is incompatible with incentive constraints or ii) there exists a Pareto improvement (agents with the lowest b are made better-off under an alternative distribution where b is not equalized).[29] In proposition 1, we essentially show that, at the optimum, there are a few intervals, delimited by the above thresholds, where either i) or ii) applies.

i) applies to interval $[n_{\underline{b}}, n_{\bar{b}}]$. Indeed, if we ignore the second-order conditions of the incentive compatibility constraints, the marginal tax rates that equalize $\underline{b}(u(n); n)$ across n would exhibit a discrete jump at n_1, leading agents with $n = n_1 - \varepsilon$ to choose a higher income than agents with $n = n_1 + \varepsilon$ which is a violation of the second-order conditions of the incentive compatibility constraints.[30] As a result, at the complete solution where the second-order conditions of incentive compatibility are imposed, there is an interval around n_1 where individuals are "bunched" and all choose the same income. The thresholds $n_{\underline{b}}$ and $n_{\bar{b}}$ simply delimit this interval and are endogenous to the solution.

Now, at the optimum, individuals with $n = n_{\underline{b}}$ and $n = n_{\bar{b}}$ have a strictly higher well-being index $\underline{b}(u(n); n)$ than agents with $n = n_1$. Therefore, by an argument of continuity, there are two intervals $(n_u; n_{\underline{b}}]$ and $[n_{\bar{b}}; n_l)$ over which the agents have a well-being index strictly larger than $\underline{b}(u(n_1); n_1)$. The threshold values that were not previously defined, n_u and n_l are simply the endpoints of these intervals. By the leximin nature of the SWF, the solution on these intervals consists in maximizing, under three constraints, revenue to be raised (on the agents in these intervals) to be redistributed lump-sum to everyone. The three constraints are incentive compatibility, $\underline{b}(u(n_u); n_u) = \underline{b}(u(n_1); n_1)$ and $\underline{b}(u(n_l); n_l) = \underline{b}(u(n_1); n_1)$. The values of n_u and n_l are endogenous to the solution and determined using the latter two constraints.

Finally, in some cases (depending on the slope of the poverty line and the Pareto parameter of the distribution of n), ii) applies in the upper-tail: relaxing marginal

[29] Recall that all agents with the same n are behaviorally equivalent in the income-consumption space, regardless of their respective w and θ. All agents with the same n choose the same basket and have the same utility. We have defined $\underline{b}(u(n); n)$ as the lowest well-being index among individuals with n at utility $u(n)$.

[30] See online subsection I.1 for more details.

tax rates on incomes larger than some n' (compared to the tax rate that equalizes $\underline{b}(u(n);n)$) is a Pareto improvement. Indeed, in those cases a lower tax rate after n' raises more revenue (that can be re-distributed lump-sum to everyone) while individuals on that interval are clearly better-off since they also face a lower marginal tax rate.

I.2.1 Solving for the thresholds n_u, $n_{\underline{b}}$, $n_{\overline{b}}$, n_l and n'

We explain formally, in online appendix I.3.2 and I.3.4, how to solve for the values of all thresholds n_u, $n_{\underline{b}}$, $n_{\overline{b}}$, n_l and n'. In this section, we attempt to give a slightly less formal and more intuitive description of their solution. In practice, we solve for n_u, n_l, y_1, $n_{\underline{b}}$ and $n_{\overline{b}}$ simultaneously. We have a system of 5 simultaneous equations in five unknowns and the value of n' is solved outside of that system. We describe the key equation/condition to solve for each parameter below. Finally, note that in that system of simultaneous equations, (16) is the fifth equation (and is key in determining y_1 at the solution).

- **Solving for $n_{\underline{b}}$ and $n_{\overline{b}}$**: Suppose that we know the values of y_1, n_u and n_l. Then, 2. and 4. in Proposition 1 give the marginal tax rates for types just outside of interval $[n_{\underline{b}}; n_{\overline{b}}]$. By the incentive compatibility constraints, there is some income for each type that corresponds to these tax rates (call such incomes $\underline{y}(n)$ for the tax rates defined in 2. and $\overline{y}(n)$ for the tax rates defined in 4.). The values of thresholds $n_{\underline{b}}$ and $n_{\overline{b}}$ are such that $\underline{y}(n_{\underline{b}}) = y_1$ and $\overline{y}(n_{\overline{b}}) = y_1$.

- **Solving for n_l:** Intuitively, we can solve for n_l by first guessing its value. Then, 4. in Proposition 1 gives the formula for the marginal tax rate as a function of n_l (which at this stage is still only a guess). Then we compute the minimum level of well-being at n_l.[31] If it is equal to $\underline{b}(u(n_1;n_1))$ then we have found n_l. Otherwise, we take another guess until the value of n_l is found.

- **Solving for n_u:** Intuitively, we can solve for n_u by first guessing its value. Then, 2. in Proposition 1 gives the formula for the marginal tax rate as a function of n_u (which at this stage is still only a guess). Then we compute the

[31] In order to do this, we need to know $n_{\overline{b}}$ (reason why n_l and $n_{\overline{b}}$ are to be solved as part of the system of simultaneous equations). From (9), we know $\underline{b}(u(n_1);n_1)$ at the solution. From the solution over $[n_1, n_{\overline{b}}]$ and $[n_{\overline{b}}, n_l]$, we can calculate $\int_{n_1}^{n_l} u'(x)\,dx$. So, conditioning on $n_{\overline{b}}$ and our "guess" of n_l, we can compute $u(n_l)$ and the corresponding $\underline{b}(u(n_l);n_l)$.

minimum level of well-being at n_u.[32] If it is equal to $\underline{b}(u(n_1; n_1))$ then we have found n_u. Otherwise, we take another guess until the value of n_u is found.

- **Solving for n':** In order to solve for n', let us define the Rawlsian marginal tax rate (the one which maximizes revenue raised): $\tau^R(n)$ such that $\frac{\tau^R(n)}{1-\tau^R(n)} = \frac{1-F(n)}{nf(n)}\left(\frac{1+\epsilon}{\epsilon}\right)$. Recall that the marginal tax rate that equalizes $\underline{b}(u(n); n)$ across n is equal to $1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$. The threshold value n' is defined such that $\tau^R(n') = 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ and $\tau^R(n') \leq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ $\forall n \in [n', \infty)$. If no such n' exists, then the well-being equalizing tax rate applies on all types above n_l.

I.2.2 Some comparative statics on the thresholds

In this section, we use the comparative statics from online appendix section E to comment on the effect of the level and slope of the poverty line on the values of the thresholds n_u, $n_{\underline{b}}$, $n_{\bar{b}}$, n_u, n' and their associated level of income.

A higher poverty line leads to a more generous tax scheme towards low incomes under the least redistributive $\tilde{R}$.[33] In addition, under all three reference preferences, a flatter poverty line also leads to a more redistributive tax scheme. When the tax schedule is more generous towards low-incomes (and in particular towards agents with n_1) the incentive for higher types to mimic n_1 increases. As a result, it is optimal to have n_1 choose a lower income in order to relax the incentive compatibility constraint on higher types. As a result, $y(n_1) = y_1$ is lower. Both because $y(n_1)$ is lower, and because the marginal tax rate on incomes above $n_{\bar{b}}$ is higher (i.e. the tax scheme is more generous towards low incomes), the discrete jump in marginal tax rates at n_1 is larger. As a result, the set of types around n_1 that is "bunched" is larger. Hence, $n_{\underline{b}}$ is lower and $n_{\bar{b}}$ is higher. In addition, with a generous poverty line (higher level or flatter slope), because $y(n_1)$ is lower, the income at which the lower-bound (10)

[32] In order to do this, we need to know $n_{\underline{b}}$ (reason why n_u and $n_{\underline{b}}$ are to be solved as part of the system of simultaneous equations). From (9), we know $\underline{b}(u(n_1); n_1)$ at the solution. From the solution over $[n_{\underline{b}}, n_1]$ and $[n_u, n_{\underline{b}}]$, we can calculate $\int_{n_1}^{n_u} u'(x)\,dx$. So, conditioning on $n_{\underline{b}}$ and our "guess" of n_u, we can compute $u(n_u)$ and the corresponding $\underline{b}(u(n_u); n_u)$.

[33] The optimal tax is invariant to the level of the poverty line under quasilinear reference preferences while a higher poverty line leads to a *less* generous tax scheme towards low incomes under the most redistributive preferences (see the comments made about figure 18 in online appendix section E).

(temporarily) stops being binding, $y(n_u)$, is also lower. In the calibration, we also find that $y(n_l)$ is increasing in the level of the poverty line and decreasing in its slope (though the magnitudes are very small). Finally, the marginal tax rate that equalizes well-being, namely $1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$, is larger under a higher or a flatter poverty line. As a result, n' is lower (notice from figure 35 that the calibrated local Pareto parameter, and hence the Rawlsian tax rate τ^R, is decreasing on most of the support of n until it stabilizes at $\frac{1}{\alpha}$ in the tail).

I.3 Hamiltonians and Deriving the First-Order Conditions

We set-up the Hamiltonian for the lower and upper subproblems under more general assumptions than in Section 4. In particular, we assume that preferences are additively separable in consumption and leisure. At the end, we plug in the functional form assumptions from Section 4 (quasilinear and iso-elastic preferences) to directly obtain formulae (17) and (18). We present the derivations under the more general assumption of additive separability in order to support the discussion in online appendix J.2 whose goal is to generalize the characterization of the optimal tax on low-incomes.

We assume that agents have additively separable utility functions:

$$u_i(c,\ell)=v(c)-h\left(\frac{\ell}{\theta_i}\right) \tag{47}$$

with $v'(.)>0, v''(.)\leq 0$ and $h'(.)>0, h''(.)\geq 0$.[34] We have

$$\underline{u_i}(c,y)=v(c)-h\left(\frac{y}{w_i\theta_i}\right)=v(c)-h\left(\frac{y}{n_i}\right)$$

The income and consumption profiles will define an implicit marginal tax rate:

$$\tau'(y(n))=1-\frac{h'\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)}{v'(c)} \tag{48}$$

[34]In Section 4, we assume $v(c)=c$ and $h\left(\frac{\ell}{\theta_i}\right)=\frac{\epsilon}{1+\epsilon}\left(\frac{\ell}{\theta_i}\right)^{\frac{1+\epsilon}{\epsilon}}$.

Incentive compatibility constraints require:

$$u(n) = \max_{n' \in [\underline{n}, \infty)} v\left(c(n')\right) - h\left(\frac{y(n')}{n}\right) = v\left(c(n)\right) - h\left(\frac{y(n)}{n}\right)$$

The first-order conditions of the latter maximization problem being satisfied at $n' = n$ implies:

$$u'(n) = h'\left(\frac{y(n)}{n}\right)\left(\frac{y(n)}{n^2}\right) \;\; \forall n \in [\underline{n}, \infty) \tag{49}$$

As before, the second-order condition is given by (15):

$$y'(n) \geq 0 \;\; \forall n \in [\underline{n}, \infty)$$

The government's budget constraint now writes

$$\int_{\underline{n}}^{\infty} \left(y(n) - v^{-1}\left(u(n) + h\left(\frac{y(n)}{n}\right)\right)\right) dF(n) \geq S$$

The two subproblems can be written as:

$$V^L(y_1) \equiv \max_{\{u(n), y(n)\}} \int_{\underline{n}}^{n_1} \left(y(n) - v^{-1}\left(u(n) + h\left(\frac{y(n)}{n}\right)\right)\right) dF(n)$$

subject to (9), (10), (49), (15) and

$$y(n) \leq y_1 \;\; \forall n$$

and

$$V^U(y_1) \equiv \max_{\{u(n), y(n)\}} \int_{n_1}^{\infty} \left(y(n) - v^{-1}\left(u(n) + h\left(\frac{y(n)}{n}\right)\right)\right) dF(n)$$

subject to (9), (10), (49), (15) and

$$y(n) \geq y_1 \;\; \forall n$$

I.3.1 Lower Subproblem: Deriving (17)

We solve the problem by defining the following Hamiltonian:

$$\begin{aligned}\mathcal{H}\left(u, y, n, \lambda, \mu, \nu^L\right) \equiv & \quad \left(y(n) - v^{-1}\left(u(n) + h\left(\frac{y(n)}{n}\right)\right)\right) f(n) \\ + & \quad \lambda(n) h'\left(\frac{y(n)}{n}\right)\left(\frac{y(n)}{n^2}\right) \\ + & \quad \mu(n)\left[u(n) - \underline{u(n)}\right] \\ + & \quad \nu^L(n)[y_1 - y(n)]\end{aligned}$$

The second line is the law of motion of utility given by the first-order conditions of the incentive compatibility constraints. The corresponding co-state variable is $\lambda(n)$. The third line is the lower-bound on utilities and $\mu(n)$ is the lagrange multiplier associated with this constraint. The fourth line is the constraint that all incomes be smaller or equal to some income y_1 (which is equal to $y^*(n_1)$ at the solution).[35]

Let us first deal with the constraint that $y(n) \leq y_1 \; \forall n$. Because of (15), $y(n)$ must be increasing. As a result, this constraint can only be binding on an interval $[n_{\underline{b}}, n_1]$. Over this interval, $\nu^L(n) \geq 0$ and $y(n) = y_1$. Solving for ν^L turns out to be important in determining $V^L_{y_1}(y_1)$ which in turn needs to be known in order to solve for $y^*(n_1)$. From the first-order conditions (evaluated at the solution $y(n) = y_1$ and corresponding $c(n) = c_1$):[36] $\frac{\partial}{\partial y}\mathcal{H} = 0 \Rightarrow$
$\nu^L(n) =$

$$\left[1 - \frac{1}{v'(c_1)} h'\left(\frac{y_1}{n}\right)\frac{1}{n}\right] f(n) + \lambda(n)\left[h''\left(\frac{y_1}{n}\right)\left(\frac{y_1}{n^3}\right) + h'\left(\frac{y_1}{n}\right)\left(\frac{1}{n^2}\right)\right]$$

Now, we focus on the solution over $[\underline{n}, n_{\underline{b}}]$. By assumption, the constraint that $y(n) \leq y_1$ is not binding and $\nu^L(n) = 0 \; \forall n \in [\underline{n}, n_{\underline{b}}]$. Under this assumption, the first-order conditions for $y(n)$ and $u(n)$ respectively yield:
$\frac{\partial}{\partial y}\mathcal{H} = 0 \Rightarrow$

[35] Notice that we ignore the constraint that $y(n)$ be increasing over all n other than n_1. In our calibrations, this constraint is only binding at n_1. But in case it was binding at some n for some distribution of types, then the solution would consist in "ironing" the income schedule. See Guesnerie and Laffont (1984).

[36] Because $u^*(n_1) = \underline{u_1}$ and y_1 are known, we can infer $c(n_1) = c_1$

$$\left[1-\frac{1}{v'(c(n))}h'\left(\frac{y(n)}{n}\right)\frac{1}{n}\right]f(n)=-\lambda(n)\left[h''\left(\frac{y(n)}{n}\right)\left(\frac{y(n)}{n^3}\right)+h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n^2}\right)\right] \tag{50}$$

and $-\frac{\partial}{\partial u}\mathcal{H}=\lambda'(n)\Rightarrow$

$$\lambda'(n)=\frac{1}{v'\left(c(n)\right)}f(n)-\mu(n)$$

Using the boundary (transversality) condition, $\lambda(\underline{n})=0$, we can write:

$$\lambda(n)=\int_{\underline{n}}^{n}\left\{\frac{f(t)}{v'\left(c(t)\right)}-\mu(t)\right\}dt \tag{51}$$

Using the fact that $\tau'\left(y(n)\right)=1-\frac{h'\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)}{v'(c)}$, we can re-arrange (50) into:

$$\tau'\left(y(n)\right)f(n)=-\lambda(n)\left[1-\tau'\left(y(n)\right)\right]\left[\frac{v'(c(n))}{h'\left(\frac{y(n)}{n}\right)\frac{1}{n}}h''\left(\frac{y(n)}{n}\right)\left(\frac{y(n)}{n^2}\right)+v'\left(c(n)\right)\left(\frac{1}{n}\right)\right]$$

The latter can be re-arranged into

$$\frac{\tau'\left(y(n)\right)}{1-\tau'\left(y(n)\right)}=-\frac{\lambda(n)}{nf(n)}v'\left(c(n)\right)\left[\frac{h''\left(\frac{y(n)}{n}\right)y(n)}{h'\left(\frac{y(n)}{n}\right)}+1\right]$$

Plugging in (51), we have

$$\frac{\tau'\left(y(n)\right)}{1-\tau'\left(y(n)\right)}=-\left[\frac{h''\left(\frac{y(n)}{n}\right)y(n)}{h'\left(\frac{y(n)}{n}\right)}+1\right]\frac{v'\left(c(n)\right)}{nf(n)}\int_{\underline{n}}^{n}\left[\frac{f(t)}{v'\left(c(t)\right)}-\mu(t)\right]dt \tag{52}$$

We can express (52) as a function of the compensated and uncompensated elasticities, ϵ^c and ϵ^u, where the latter are defined by

$$\epsilon^c = \frac{1-\tau'(y)}{y}\left.\frac{\partial y}{\partial(1-\tau'(y))}\right|_{\overline{u}} \qquad = \frac{h'\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)^2}{y\left[h''\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)-(1-\tau'(y))^2 v''(c)\right]}$$

$$\epsilon^u = \frac{1-\tau'(y)}{y}\frac{\partial y}{\partial(1-\tau'(y))} \qquad = \frac{h'\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)+(1-\tau'(y))^2 yv''(c)}{y\left[h''\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)^2-(1-\tau'(y))^2 v''(c)\right]}$$

Indeed, $\left[\frac{h''\left(\frac{y(n)}{n}\right)y(n)}{h'\left(\frac{y(n)}{n}\right)}+1\right]$

$$= \frac{h''\left(\frac{y(n)}{n}\right)y(n)\left(\frac{1}{n}\right)+h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n}\right)}{h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n}\right)}$$

$$= \frac{h''\left(\frac{y(n)}{n}\right)y(n)\left(\frac{1}{n}\right)+h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n}\right)+(1-\tau'(y(n)))^2 y(n)v''(c(n))-(1-\tau'(y(n)))^2 y(n)v''(c(n))}{h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n}\right)}$$

$$= \frac{\frac{h''\left(\frac{y(n)}{n}\right)y(n)\left(\frac{1}{n}\right)-(1-\tau'(y(n)))^2 y(n)v''(c(n))}{y\left[h''\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)-(1-\tau'(y))^2 v''(c)\right]}+\frac{h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n}\right)+(1-\tau'(y(n)))^2 y(n)v''(c(n))}{y\left[h''\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)-(1-\tau'(y))^2 v''(c)\right]}}{\frac{h'\left(\frac{y(n)}{n}\right)\left(\frac{1}{n}\right)}{y\left[h''\left(\frac{y}{n}\right)\left(\frac{1}{n}\right)-(1-\tau'(y))^2 v''(c)\right]}}$$

$$= \left[\frac{1+\epsilon^u}{\epsilon^c}\right]$$

As a result, we obtain:

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = -\left(\frac{1+\epsilon^u}{\epsilon^c}\right)\frac{v'(c(n))}{nf(n)}\int_{\underline{n}}^{n}\left[\frac{f(t)}{v'(c(t))}-\mu(t)\right]dt \qquad (53)$$

Finally, to obtain (17), notice that under the assumptions of Section 4, $\epsilon^u=\epsilon^c=\epsilon$ and $v^{'}(c(n))=1$.

I.3.2 Solving for $\mu(n)$, $V_{y_1}^U(y_1)$, $n_{\underline{b}}$ and n_u

Equation (53) holds for all $n \in [\underline{n}, n_{\underline{b}}]$. There are two possible cases. Either the lower-bound on utilities is not binding in n in which case $\mu(n) = 0$ and the marginal tax rate can be obtained from equation (53). Or the lower-bound is binding in n in which case $\tau'(y(n)) = 0$ and $\mu(n)$ can be solved using equation (53).

Notice that using the envelope theorem, we can find:

$$V_{y_1}^L(y_1) = \int_{n_{\underline{b}}}^{n_1} \nu^L(t)dt \tag{54}$$

Finally, we need to solve for n_u and $n_{\underline{b}}$. Recall that we defined utility of type n along the lower-bound as $\underline{u}(n)$. Utility of agents of type 1 is known and fixed at $\underline{u_1}$. n_u is the largest type for which the lower-bound is binding:

$$\underline{u_1} - \int_{n_u}^{n_1} u'(t)dt = \underline{u}(n_u) \tag{55}$$

$n_{\underline{b}}$ is the lowest type for which the constraint that $y(n) \leq y_1$ is binding. Let $\tilde{y}(n)$ be income implicitly defined by
$\frac{\tau'(y(n))}{1-\tau'(y(n))} = -\left[\frac{1+\epsilon^u}{\epsilon^c}\right] \frac{v'(c(n))}{nf(n)} \left[\int_{\underline{n}}^{n} \frac{f(t)}{v'(c(t))}dt - \int_{\underline{n}}^{n_u} \frac{f(t)}{v'(c(t))}dt\right]$. $n_{\underline{b}}$ is defined by the following equation:

$$\tilde{y}(n_{\underline{b}}) = y_1 \tag{56}$$

Equation (55) and (56) is the system of two equations in two unknowns used to solve for n_u and $n_{\underline{b}}$.

I.3.3 Upper Subproblem: Deriving (18)

We define the following Hamitonian:

$$
\begin{aligned}
\mathcal{H}\left(u, y, n, \lambda, \mu, \nu^U\right) \equiv & \left(y(n) - v^{-1}\left(u(n) + h\left(\frac{y(n)}{n}\right)\right)\right) f(n) \\
+ & \lambda(n) h'\left(\frac{y(n)}{n}\right)\left(\frac{y(n)}{n^2}\right) \\
+ & \mu(n)\left[u(n) - \underline{u(n)}\right] \\
+ & \nu^U(n)[y(n) - y_1]
\end{aligned}
$$

From constraint (15), we know that $y(n)$ is increasing. As a result, the constraint that $y(n) \geq y_1$ can only be binding (if at all) on an interval $[n_1, n_{\bar{b}}]$. On this interval, $y(n) = y_1$ and $\nu^U(n) \geq 0$. From the first-order conditions (evaluated at the solution $y(n) = y_1$ and corresponding $c(n) = c_1$): $\frac{\partial}{\partial y}\mathcal{H} = 0 \Rightarrow$
$\nu^U(n) =$

$$
-\left[1 - \tfrac{1}{v'(c_1)} h'\left(\tfrac{y_1}{n}\right)\tfrac{1}{n}\right] f(n) - \lambda(n)\left[h''\left(\tfrac{y_1}{n}\right)\left(\tfrac{y_1}{n^3}\right) + h'\left(\tfrac{y_1}{n}\right)\left(\tfrac{1}{n^2}\right)\right]
$$

We now focus on the solution over interval $[n_{\bar{b}}, \infty)$. By definition, $\nu^U(n) = 0$ over this interval. Under this assumption, the first-order conditions for $y(n)$ and $u(n)$ respectively yield:
$\frac{\partial}{\partial y}\mathcal{H} = 0 \Rightarrow$

$$
\left[1 - \tfrac{1}{v'(c(n))} h'\left(\tfrac{y(n)}{n}\right)\tfrac{1}{n}\right] f(n) = -\lambda(n)\left[h''\left(\tfrac{y(n)}{n}\right)\left(\tfrac{y(n)}{n^3}\right) + h'\left(\tfrac{y(n)}{n}\right)\left(\tfrac{1}{n^2}\right)\right] \tag{57}
$$

and $-\frac{\partial}{\partial u}\mathcal{H} = \lambda'(n) \Rightarrow$

$$
\lambda'(n) = \frac{1}{v'(c(n))} f(n) - \mu(n)
$$

Using the transversality condition, $\lim_{n\to\infty} \lambda(n) = 0$, we can write:

$$
\lambda(n) = -\int_n^\infty \left\{\frac{f(t)}{v'(c(t))} - \mu(t)\right\} dt \tag{58}
$$

Using the same reasoning as for the lower-subproblem, we can re-arrange (57) into

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = -\frac{\lambda(n)}{nf(n)} v'(c(n)) \left[\frac{h''\left(\frac{y(n)}{n}\right) y(n)}{h'\left(\frac{y(n)}{n}\right)} + 1 \right]$$

Plugging in (58), we have

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left[\frac{h''\left(\frac{y(n)}{n}\right) y(n)}{h'\left(\frac{y(n)}{n}\right)} + 1 \right] \frac{v'(c(n))}{nf(n)} \int_n^\infty \left[\frac{f(t)}{v'(c(t))} - \mu(t) \right] dt$$

Using a result derived in online appendix subsection I.3.1, we obtain equation (59):

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left[\frac{1+\epsilon^u}{\epsilon^c} \right] \frac{v'(c(n))}{nf(n)} \int_n^\infty \left[\frac{f(t)}{v'(c(t))} - \mu(t) \right] dt \qquad (59)$$

Finally, to obtain (18), notice that under the assumptions of Section 4, $\epsilon^u = \epsilon^c = \epsilon$ and $v^{'}(c(n)) = 1$.

I.3.4 Solving for $\mu(n)$, $V^L_{y_1}(y_1)$, $y^*(n_1)$, $n_{\bar{b}}$, n_l and n'

Using the envelope theorem, we can write

$$V^U_{y_1}(y_1) = \int_{n_1}^{n_{\bar{b}}} \nu^U(t) dt \qquad (60)$$

We can plug (54) and (60) into (16) to solve for $y^*(n_1)$. In order to solve for $\mu(n)$, $n_{\bar{b}}$, n_l and n', we impose the assumptions from Section 4 ($v(c) = c$ and $h(\ell/\theta_i) = \frac{\epsilon}{1+\epsilon}(\ell/\theta_i)^{\frac{1+\epsilon}{\epsilon}}$). Also recall that we assume in Proposition 1 that the lower-bound on utilities is binding on at most one interval above n_1. Let us define $\tau^R(n)$ as the tax rate implicitly defined by (18) when $\int_n^\infty \mu(t)\, dt = 0$. That is $\tau^R(n)$ is the optimal tax rate at n when the lower-bound on utilities (10) is never binding above n. As previously noted, it corresponds to the tax rate under a Rawlsian maximin. It is the tax rate that maximizes revenue raised under incentive compatibility constraints. We need to consider three cases.

Case 1 First, we need to check whether $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} \geq 1 - \left[1 - \left(\frac{\tilde{\theta}}{\bar{\theta}}\right)^{1+\epsilon} \right]^{\frac{1}{1+\epsilon}}$. If it is the case, we know that the lower-bound on utilities will be binding in the tail. The tax rate will be $1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ on some interval $[n_l, \infty)$.

Because $lim_{n\to\infty}F(n) = 1$ and $lim_{n\to\infty}\lambda(n) = 0$ by the transversality condition, we can derive $\int_n^\infty \mu(t)dt = 1 - F(n_l) - \frac{1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}}{\left(\frac{1+\epsilon}{\epsilon}\right)\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}} n_l f(n_l)$. This expression combined with Eq.(18) yields the marginal tax rate over $[n_{\bar{b}}, n_l]$:

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = \left(\frac{1+\epsilon}{\epsilon}\right) \frac{[F(n_l)-F(n)] + \frac{1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}}{\left(\frac{1+\epsilon}{\epsilon}\right)\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}} n_l f(n_l)}{n f(n)} \quad \forall n \in [n_{\bar{b}}, n_l] \tag{61}$$

Case 2 If $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} < 1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$, then there exists some n' such that $\tau^R(n') = 1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ and $\tau^R(n) \leq 1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$ $\forall n \in [n', \infty)$. The revenue maximizing tax rate $\tau^R(n)$ applies on the interval $[n', \infty)$. Assuming that the lower-bound is binding over interval $[n_l, n']$ only, we have:

$$\int_n^\infty \mu(t)dt = \int_{n_l}^{n'} \mu(t)dt = 1 - F(n_l) - \frac{1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}}{\left(\frac{1+\epsilon}{\epsilon}\right)\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}} n_l f(n_l)$$

$\forall n \in [n_1, n_l]$.[37]

Notice that this is the same expression as in case 1. As a result, marginal tax rates over $[n_{\bar{b}}, n_l]$ are also defined by (61).

[37] The derivation of this formula uses the definition of n' which implies $1 - F(n') = \frac{1-\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}}{\left(\frac{1+\epsilon}{\epsilon}\right)\left[1-\left(\frac{\tilde{\theta}}{\bar{\bar{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}} n' f(n')$.

Case 3 If $\frac{1+\epsilon}{1+\epsilon+\alpha\epsilon} < 1 - \left[1 - \left(\frac{\tilde{\theta}}{\overline{\overline{\theta}}}\right)^{1+\epsilon}\right]^{\frac{1}{1+\epsilon}}$, it is possible that the lower-bound is never binding above n_1. In the calibrations, this typically happens when the poverty line is flat (or high enough). Incentive compatibility "protects" higher types agents from ever being against the lower-bound. In that case, the revenue maximizing tax rate $\tau^R(n)$ applies everywhere on $[n_{\overline{b}}, \infty)$.

Finally, we need to solve for $n_{\overline{b}}$ and n_l (if in case 1 or 2). Recall that we defined utility of type n along the lower-bound as $\underline{u}(n)$. Utility of agents of type 1 is known and fixed at $\underline{u_1}$. n_l is the lowest type for which the lower-bound is binding:

$$\underline{u_1} + \int_{n_1}^{n_l} u'(t)dt = \underline{u}(n_l) \tag{62}$$

$n_{\overline{b}}$ is the largest type for which the constraint that $y(n) \geq y_1$ is binding. Let $\tilde{y}(n)$ be income implicitly defined by (61) (in case 1 and 2) and the fact, adapted from (48), that $1 - \tau'(y(n)) = \left(\frac{y(n)}{n}\right)^{\frac{1}{\epsilon}}\left(\frac{1}{n}\right)$. $n_{\overline{b}}$ is defined by the following equation:

$$\tilde{y}(n_{\overline{b}}) = y_1 \tag{63}$$

If we are in case 3, let $\tilde{y}(n)$ be defined by the Rawlsian tax rate $\tau^R(n)$ and $1 - \tau'(y(n)) = \left(\frac{y(n)}{n}\right)^{\frac{1}{\epsilon}}\left(\frac{1}{n}\right)$. Equation (62) and (63) is the system of two equations in two unknowns used to solve for n_l and $n_{\overline{b}}$.

One last comment is in order. The system above conditions on y_1. Finding a complete solution to the optimal tax problem, requires solving for n_u, $n_{\underline{b}}$, $n_{\overline{b}}$, n_l, and $y^*(n_1)$ simultaneously. The system of five equations in five unknowns combines (16), (55), (56), (62), and (63).[38]

J Some Partial Generalizations of Proposition 1

J.1 Generalizing steps 1 and 2: Negative Marginal Tax Rates on Low-Incomes

It is clear from the proofs of steps 1 and 2, that the assumptions of quasi-linearity, iso-elasticity or additive separability of preferences play no role in obtaining Lemma 1 and

[38] If we are in case 3, there is no n_l, and we solve a system of four equations in four unknowns: (16), (55), (56), and (63).

Proposition 4. As a matter of fact, these results can be proven in a much more general model allowing for non-separabilities between leisure and consumption, changes in labor supply on the extensive margin and true multidimensional heterogeneity (in the sense that heterogeneity in preferences and wages does not need to collapse into heterogeneity in a single index such as n_i). This is an important point since, as emphasized by corollary 2 below, it implies that the result that marginal tax rates are weakly negative below the income earned by agent 1 extends to such models. To formalize this point, we first define two (much more general) assumptions that can replace the assumptions made in Section 4 and under which Proposition 4 still holds. Next, we state and prove corollary 2.

We make the following two assumptions, which are (much) weaker than the assumptions of Section 4.

Assumption 1: There exists agent $1 \in N$ such that for all $i \in N$, $w_1 \leq w_i$ and $R_1 \geq R_i$, where we write $R \geq R'$ to denote that whatever the set of bundles among which to choose, R chooses a larger labor time than R'.

Assumption 2: Let y_1 be the earning level of agent 1. For all $y \leq y_1$, if there exists $i \in N$ such that $\underline{z}_i = (y, y - \tau(y))$, then there exists $j \in N$ such that $w_j = w_1$ and $\underline{z}_j = (y, y - \tau(y))$.

Proposition 5 *Let Assumptions 1 and 2 hold. If tax scheme τ is optimal in economy $E = (w_N, R_N)$ for a social welfare function $\mathbf{R}^{\tilde{R}lex}$, then allocation $\underline{z}_N$ generated by τ has the feature that the corresponding allocation z_N in the labor time/consumption space is such that agent 1 has the lowest well-being level:*

$$b(z_1, w_1, R_1) \leq b(z_i, w_i, R_i), \forall i \in N.$$

An immediate corollary of that proposition, which proof is essentially contained in the proof of the proposition, is that the optimal tax scheme τ satisfies the following property, the first part of which is illustrated in Fig. 38.

Corollary 2 *Let Assumptions 1 and 2 hold. A) If tax scheme τ is optimal in economy $E = (w_N, R_N)$ for a social welfare function $\mathbf{R}^{\tilde{R}lex}$, then allocation $\underline{z}_N$ generated by τ with $\underline{z}_1 = (y_1, c_1)$ is such that agent* 1 *receives the largest transfer, that is for all income $y \leq y_1$:*

$$\tau(y_1) \leq \tau(y) \leq t,$$

where $t \leq 0$ is the intercept of agent 1's implicit budget,

$$\underline{IB}(\underline{z}_1, \underline{R}_1) = \{(y, c) \in \underline{X} \mid c \leq y - t\},$$

that is, agent 1 receives the largest transfer and marginal tax rates are on average non-positive below y_1. B) Moreover, if another agent has the same wage as agent 1, different preferences but ends up with the same well-being level than 1, then this agent and all agents earning less than this agent receive a transfer exactly equal to t: if there exists $j \in N$ such that $w_j = w_1$, $\underline{IB}(\underline{z}_1, \underline{R}_1) = \underline{IB}(\underline{z}_j, \underline{R}_j)$ and $\underline{R}_j < \underline{R}_1$, then

$$\tau(y_j) = t,$$

and for all $y \leq y_j$,

$$\tau(y) = t.$$

In particular, this Corollary implies that all agents who do not work receive a transfer of t.

Proof. Let $E = (w_N, R_N)$. Let τ be optimal for $\mathbf{R}^{\tilde{R}lex}$. Let $\underline{z}_N$ be generated by τ with $\underline{z}_1 = (y_1, c_1)$. A) By Proposition 5, $b(z_1, w_1, R_1) \leq b(z_i, w_i, R_i)$, for all $i \in N$. Let $\underline{IB}(\underline{z}_1, \underline{R}_1) = \{(y, c) \in \underline{X} \mid c \leq y - t\}$. By replicating the proof of the fact appearing at the beginning of the proof of Proposition 4, we show that $\tau(y) \leq t$ for all $y \leq y_1$. Assume that, contrary to the claim of the Corollary, $\tau(y_1) > \tau(y^*)$ for some $y^* \leq y_1$. Then, we can replicate the proof of Claim 2 of the proof of Proposition 4, define τ' as $\tau'(y) = \max\{\tau(y), \tau(y^*)\}$, and show that τ' is better than τ according to $\mathbf{R}^{\tilde{R}lex}$, the desired contradiction. B) Let $j \in N$ be such that $w_j = w_1$, $\underline{IB}(\underline{z}_1, \underline{R}_1) = \underline{IB}(\underline{z}_j, \underline{R}_j)$ and $\underline{R}_j < \underline{R}_1$. Let $\underline{z}_j^* = (y_j^*, c_j^*) \in \underline{X}$ be such that $\underline{z}_j^* \, \underline{I}_j \, \underline{z}_j$ and $\underline{z}_j^* \in \underline{IB}(\underline{z}_j, \underline{R}_j)$. Note, in particular, that $\tau(y_j^*) = t$. We first claim that there is no loss of generality in assuming that $\underline{z}_j^* = \underline{z}_j$. Assume, on the contrary, that $\underline{z}_j^* \neq \underline{z}_j$. By construction, $\tau(y_j) < \tau(y_j^*)$. Therefore, $\underline{z}'_N$, defined by $\underline{z}'_i = \underline{z}_i$ for all $i \in N \setminus \{j\}$ and $\underline{z}'_j = \underline{z}_j^*$ is budget balanced. Moreover, it is incentive compatible. Indeed, assume not. There exists $k \in N$ such that $\underline{z}_j^* \, \underline{P}_k \, \underline{z}_k$. By Assumption 2, there exists $\ell \in N$ such that $\underline{z}_j^* \, \underline{P}_\ell \, \underline{z}_\ell$ and $w_\ell = w_j = w_1$. Consequently,

$$\underline{IB}(\underline{z}_\ell, \underline{R}_\ell) \subset \underline{IB}(\underline{z}_j^*, \underline{R}_j) = \underline{IB}(\underline{z}_j, \underline{R}_j),$$

so that

$$b(z_\ell, w_\ell, R_\ell) < b(z_1, w_1, R_1),$$

contradicting Proposition 5. This proves the claim. Now, it follows that for all $y \leq y_j$,

$$\tau(y) = t.$$

Indeed, if it were not the case, then, as in the proof of part A) above, it would be possible to define τ' as follows: for all $y \geq y_j$, $\tau'(y) = \tau(y)$ and for all $y \leq y_j$, $\tau'(y) = \tau(y_j)$, which leads to the same social welfare but collects a budget surplus, contradicting the optimality of τ. ■

J.2 Generalizing step 3: Characterizing the Income Tax on Low-Incomes under Additive Separable Preferences

When presenting the derivations of the first-order conditions in online appendix I.3, we used the more general assumption that preferences are additively separable. In other words, we allowed for income effects and for the elasticity of labor supply to vary with c and ℓ. As presented formally in Proposition 6, the implication is that we can fully characterize the optimal tax on low-incomes (incomes lower than $y^*(n_1)$) under these more general assumptions.

Proposition 6 *Under the additively separable preferences defined in* (47)*, if a planner maximizes some* $\mathbf{R}^{\tilde{R}lex}$*, there exist* $n_u \in [\underline{n}, n_1]$ *and* $n_{\underline{b}} \in [n_u, n_1]$ *such that*

1. $\tau'(y(n)) = 0 \ \forall n \in [\underline{n}, n_u]$
2. $\tau'(y(n)) \leq 0 \ \forall n \in [n_u, n_{\underline{b}}]$ *and*

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = -\left[\frac{1+\epsilon^u}{\epsilon^c}\right]\frac{v'(c(n))}{nf(n)}\left[\int_{\underline{n}}^{n}\frac{f(t)}{v'(c(t))}dt - \int_{\underline{n}}^{n_u}\frac{f(t)}{v'(c(t))}dt\right]$$

3. $y^*(n) = y^*(n_1) \ \forall n \in [n_{\underline{b}}, n_1]$

where ϵ^c *and* ϵ^u *are, respectively, the compensated and uncompensated elasticity of labor earnings.*

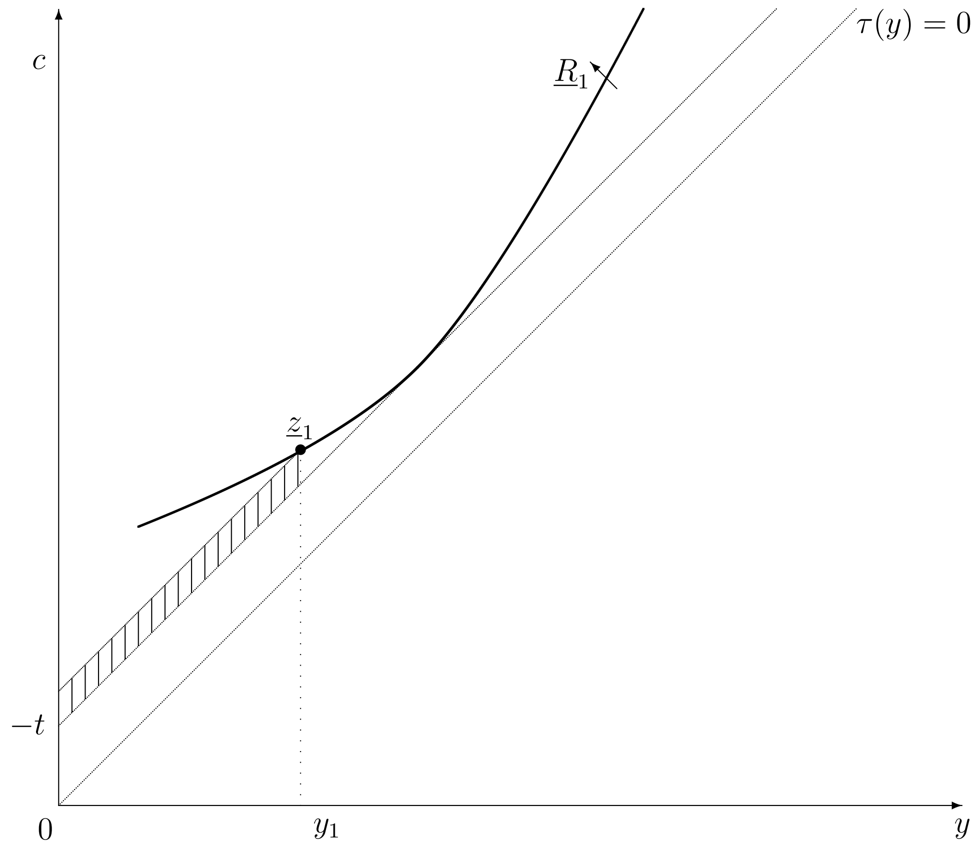

Figure 38: Illustration of Corollary 2. Agent 1's indifference curve at the optimal tax scheme τ is depicted on the graph. Her consumption bundle under τ is labeled as $\underline{z}_1$. The tax transfer to 1 under τ, $-\tau(y_1)$, is the vertical difference between $\underline{z}_1$ and the 45° line going through the origin. Corollary 2 states that the tax transfer received by all agents with income $y \leq y_1$ must be lower than $-\tau(y_1)$ but larger than $-t$. Hence, the graph of $y - \tau(y)$ must lie in the shaded area for all $y \leq y_1$.

Proof.

Writing the first-order conditions for the Hamiltonians corresponding to each subproblem and rearranging, we obtain the following formula for the optimal marginal tax rate:[39]

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = A(n)D(n)C(n) \qquad \forall n \in [\underline{n}, n_{\underline{b}}] \tag{64}$$

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = A(n)B(n)C(n) \qquad \forall n \in [n_{\bar{b}}, \infty) \tag{65}$$

[39] See derivations in online appendix I.3

with

$$
\begin{aligned}
A(n) &= \left[\frac{1+\epsilon^u}{\epsilon^c}\right] && (66)\\
B(n) &= v'(c(n))\int_n^\infty \left[\frac{f(t)}{v'(c(t))}-\mu(t)\right]dt && (67)\\
C(n) &= \frac{1}{nf(n)} && (68)\\
D(n) &= -v'(c(n))\int_{\underline{n}}^n \left[\frac{f(t)}{v'(c(t))}-\mu(t)\right]dt && (69)
\end{aligned}
$$

Notice that ϵ^c and ϵ^u are, in general, functions of the marginal tax rate $\tau'(y)$ and the agent's type n. We now turn to the three claims in Proposition 6. The third claim is rather obvious. By incentive compatibility, $y^*(n)$ is increasing. Therefore, the constraint that $y^*(n) \le y_1 = y^*(n_1)$ can only bind (if at all) on an interval $[n_{\underline{b}}, n_1]$ for some $n_{\underline{b}}$. If the constraint is binding on this interval, then $y^*(n) = y^*(n_1)$ $\forall n \in [n_{\underline{b}}, n_1]$. Notice that we do not rule out the case where $n_{\underline{b}} = n_1$.

We will prove the rest of Proposition 6 in three steps. The first step is intuitive and was actually proven for more general assumptions in the previous section. Formally, we want to show that if $\exists a,b \in [\underline{n}, n_1]$, $a<b$, such that $\mu(n) > 0$ $\forall n \in [a,b]$, then

$$\tau(y(n)) = \overline{\tau},\ \forall n \in [a,b)$$

where $\overline{\tau}$ is such that $\max_{y\in[0,\infty)}\left(v(y-\overline{\tau}) - h\left(\frac{y}{b}\right)\right) = \underline{u}(b)$. The statement also implies that $\tau'(y(n)) = \tau''(y(n)) = 0$ $\forall n \in [a,b)$.

Clearly, letting agents with $n \in [a,b]$ choose over a segment where $\tau(y(n)) = \overline{\tau}$ equalizes the implicit budget of all agents with $n \in [a,b]$ and $w_i = \underline{w}$. As a result, all agents in this interval, with $w_i = \underline{w}$, are against the lower-bound on their utilities: $u(n) = \underline{u}(n)$ $\forall n \in [a,b]$. Also, for any other $(y'(n), c'(n))$ such that $u'(n) \ge \underline{u}(n)$, we would have

$$\int_a^b (y'(n) - c'(n))f(n)dn \le \int_a^b (y(n)-c(n))f(n)dn = \overline{\tau}(F(b)-F(a)).$$

The above reasoning implies that $\mu(n) > 0$ $\forall n \in [a,b] \Rightarrow \tau'(y(n)) = 0$ and $\tau''(y(n)) = 0$ $\forall n \in [a,b]$.

For the second step, we show that $\tau'(y(n)) \leq 0 \ \forall n \in [\underline{n}, n_{\underline{b}}]$. From equation (64), we have

$$\frac{\tau'(y(n))}{1-\tau'(y(n))} = A(n)D(n)C(n)$$

This equation holds for all $n \in [\underline{n}, n_{\underline{b}}]$. We want to show that $A(n)D(n)C(n) \leq 0$. We have $A(n) = \frac{1+\epsilon^u}{\epsilon^c} = 1 + \frac{h''\left(\frac{y(n)}{n}\right)y(n)}{h'\left(\frac{y(n)}{n}\right)} > 0$ and $C(n) = \frac{1}{nf(n)} > 0$. The proof boils down to showing that $\int_{\underline{n}}^{n} \frac{f(t)}{v'(c(t))}dt \geq \int_{\underline{n}}^{n} \mu(t)dt$ in which case $D(n) \leq 0$. Rearranging (64), we have

$$\int_{\underline{n}}^{n} \mu(t)dt = \frac{1+\epsilon^u}{\epsilon^c}\frac{nf(n)}{v'(c(n))}\frac{\tau'(y(n))}{1-\tau'(y(n))} + \int_{\underline{n}}^{n} \frac{f(t)}{v'(c(t))}dt$$

As a result, we can write

$$\mu(n) = \frac{d}{dn}\left[\frac{1+\epsilon^u}{\epsilon^c}\frac{nf(n)}{v'(c(n))}\frac{\tau'(y(n))}{1-\tau'(y(n))}\right] + \frac{f(n)}{v'(c(n))}$$

Whenever $\mu(n) > 0$, we have $\tau'(y(n)) = 0$ and $\tau''(y(n)) = 0$ because the marginal tax rate along the lower-bound is constant at zero (as discussed in the first step). As a result, either $\mu(n) = 0$ or $\mu(n) = \frac{f(n)}{v'(c(n))}$. This implies that $\int_{\underline{n}}^{n} \frac{f(t)}{v'(c(t))}dt \geq \int_{\underline{n}}^{n} \mu(t)dt$.

The third step of the proof consist in noticing that if marginal tax rates are everywhere zero or negative, the lower-bound (along which the marginal tax rate is zero) can be binding on at most one interval from $\underline{n}$ to some n_u.

Finally, because the lower-bound is only binding on interval $[\underline{n}, n_u]$, we have that $\int_{\underline{n}}^{n} \mu(t)dt = \int_{\underline{n}}^{n_u} \mu(t)dt = \int_{\underline{n}}^{n_u} \frac{f(t)}{v'(c(t))}dt \ \forall n \in [n_u, n_1]$. Claim 2 follows directly from this and equation (64).

■

References

[1] Atkinson, Anthony B., Piketty, Thomas, and Emmanuel Saez (2011). "Top Incomes in the Long Run of History," *Journal of Economic Literature* 49(1):3–71.

[2] Bollinger, Barry T., Hirsch, Charles M. and James P. Ziliak (2019). "Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith and Welch," *Journal of Political Economy*, 127(5): 2143–2185

[3] Guesnerie, R. and J.-J. Laffont 1984, “A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm.,” *Journal of Public Economics* 25, 329-369.

[4] Hajargasht Gholamreza and William E. Griffiths 2013, “Pareto-lognormal distributions: Inequality, poverty and estimation from grouped income data,” *Economic Modelling* 33, 593–604.

[5] Meyer Bruce D., Mok Wallace K.C. and James X. Sullivan 2015, ”The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences,” Mimeo.

[6] Reed, William J. 2003, “The Pareto law of incomes - an explanation and an extension,” *Physica A* 319, 469–486

[7] Wooldridge, Jeffrey M. 2010, “Econometric Analysis of Cross Section and Panel Data”, *MIT Press.*