

Berg, Kristoffer; Piacquadio, Paolo Giovanni

**Working Paper**

## The Equal-Sacrifice Social Welfare Function with an Application to Optimal Income Taxation

CESifo Working Paper, No. 8505

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Berg, Kristoffer; Piacquadio, Paolo Giovanni (2020) : The Equal-Sacrifice Social Welfare Function with an Application to Optimal Income Taxation, CESifo Working Paper, No. 8505, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/223577>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The Equal-Sacrifice Social Welfare Function with an Application to Optimal Income Taxation

*Kristoffer Berg, Paolo G. Piacquadio*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: <https://www.cesifo.org/en/wp>

# The Equal-Sacrifice Social Welfare Function with an Application to Optimal Income Taxation

## Abstract

We propose and axiomatically characterize a family of welfare criteria that prioritize individuals making larger sacrifices. By combining efficiency with a concern for equality of sacrifice, our criteria avoid serious shortcomings of utilitarianism. We illustrate our results within the Mirrleesian optimal taxation framework. Our simulated equal-sacrifice optimal tax schedule for the US has marginal tax rates in line with the US tax system and about 20 percentage points lower than the utilitarian recommendation.

JEL-Codes: D630, H210, I310.

Keywords: equal-sacrifice principle, optimal income taxation, welfare criterion.

*Kristoffer Berg*  
University of Oslo / Norway  
*kristoffer.berg@econ.uio.no*

*Paolo G. Piacquadio*  
University of Oslo / Norway  
*p.g.piacquadio@econ.uio.no*

August 9, 2020

The authors wish to thank Rolf Aaberge, Alberto Alesina, Geir Asheim, Marcus Berliant, Tomer Blumkin, Walter Bossert, Giacomo Brusco, Vidar Christiansen, Bård Harstad, Bas Jacobs, Claus Thustrup Kreiner, Etienne Lehmann, Ben Lockwood, Frikk Nesje, Sergio Ocampo-Diaz, Tone Ognedal, Andreas Peichl, Uwe Thuemmel, Dominik Sachs, Emmanuel Saez, Dirk Schindler, Kjetil Storesletten, Valer Suteu, Thor O. Thoresen, and Matt Weinzierl for comments, in addition to the audiences at the NTA 2019, Heidelberg, and the Norwegian-German tax conference. This project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme ERC Starting Grant VALURED (Grant agreement No. 804104) and from the Research Council of Norway (RCN) through the research center Oslo Fiscal Studies.

# 1 Introduction

In optimal income taxation, the first best is generally unachievable. Nevertheless, the first best identifies the ethical standard of perfection and defines the direction of social improvement.<sup>1</sup> Second-best policies then emerge as a compromise of the unavoidable tradeoff between equity and efficiency.

Optimal taxation is generally addressed through the lenses of utilitarianism.<sup>2</sup> With additively separable utility over consumption and labor and income effects, the utilitarian first-best tax policy is to equalize after-tax income. This egalitarian implication of utilitarianism did not pass unobserved (Mill, 1848; Musgrave, 1959; Feldstein, 1976; Saez and Stantcheva, 2016; Fleurbaey and Maniquet, 2018). As Feldstein (1976) clarifies, utilitarianism implicitly assumes that all differences in wages across individuals are undeserved: society jointly owns everyone’s potential earnings. At the same time, individuals ought to set their labor supply for the common (utilitarian) good, requiring high-skilled individuals to supply more labor and, thus, resulting in “slavery of the talented.”

Both the equalization of after-tax income and the slavery of the talented disappear at the second best (Mirrlees, 1971). When the government only observes income, labor supply responses prevent confiscatory income taxes. At the same time, this ensures the talented achieve a higher level of well-being. Nevertheless, utilitarianism remains problematic. The government would always accept some reduction of total consumption for the sake of more equality in after-tax income, leading to optimal marginal tax rates of over 70% (Mankiw, Weinzierl, and Yagan, 2009). In conflict with the utilitarian view, complete equality in after-tax income is generally considered confiscatory and unfair and, thus, not worth efficiency losses. In fact, most survey respondents seem to agree that, *ceteris paribus*, higher-income earners deserve a larger after-tax income (Schokkaert and Devooght, 2003; Weinzierl, 2014; Saez and Stantcheva, 2016).<sup>3</sup>

---

<sup>1</sup>As Mill (1848) writes: “the first object in every practical discussions should be to know what perfection is.” More recently, John Simmons (2010, p.34) argues that “we can hardly claim to know whether we are on the path to the ideal of justice until we can specify in what that ideal consists.”

<sup>2</sup>Some authors distinguish between utilitarianism—the sum of individuals’ utilities—and generalized utilitarianism—the sum of concave transformations of individuals’ utilities. With minor changes, the drawbacks highlighted next extend to generalized utilitarianism.

<sup>3</sup>To escape these objections, many authors resort to the weighted utilitarian criterion, where “Pareto weights” allow weighting individuals differently. However, when the Pareto weights are inversely correlated with productivities—as often suggested in the literature—the drawbacks of utilitarianism are aggravated. Moreover, constant Pareto weights are not flexible enough to avoid the shortcomings of utilitarianism (Fleurbaey and Maniquet, 2018). Finally, when the Pareto weights are left unspecified,

An alternative principle for optimal taxation—the equal-sacrifice principle—addresses these shortcomings of utilitarianism. Taxes ought to be set so that each tax-payer makes the same sacrifice. Since sacrifice is increasing in the tax payment, after-tax income equalization can never be first-best optimal. The intuitive appeal of the equal sacrifice principle spawned a large interest in the early economic literature (Mill, 1848; Sidgwick, 1883; Edgeworth, 1897; Pigou, 1928; Vickrey, 1947). Nevertheless, it remains unclear how to set the optimal second-best income tax schedule in accordance with the equal-sacrifice principle. Two issues are critical. First, it is unclear how to define sacrifice (Musgrave, 1959; Young, 1988). Second, equal sacrifice—as a standard of perfection—cannot be adopted for second-best analysis, as equal-sacrifice tax schedules are typically inefficient (Berliant and Gouveia, 1993; da Costa and Pereira, 2014).

In this paper, we address both issues. We propose and axiomatically characterize a family of welfare criteria inspired by the equal-sacrifice principle. These criteria trade off efficiency with equality of sacrifice and can therefore be used for second-best analysis. Each criterion in our family can be represented by an *equal-sacrifice social welfare function*, defined as the sum of specific indices of each individual’s well-being. These indices satisfy three intuitive properties.

1. They represent the preferences of individuals. Individuals are better off if and only if their index of well-being is larger, ensuring society abides by the Pareto principle and, thus, values efficiency.
2. They make comparisons in terms of sacrifice. Individuals’ levels of sacrifice are the only information society needs to optimally allocate marginal increases in consumption. As in the literature, we measure sacrifice with respect to the laissez-faire allocation (where no taxes are levied). The novelty is in our accounting for labor supply responses. To do so, we measure the sacrifice of an individual as the level of taxes that—at laissez-faire labor supply—she considers equally desirable as her consumption/leisure bundle.
3. They are strictly concave. Thus, the criteria are averse to inequality of sacrifice and prioritize individuals incurring a larger sacrifice.

The main difference between our equal-sacrifice social welfare function and utilitarianism is the way these criteria compare individuals. Utilitarianism compares individuals

---

the criterion only identifies the Pareto frontier (Negishi, 1960) and, thus, provides little guidance to optimal income taxation.

by their marginal utility of income. If  $i$  has a larger marginal utility than  $j$ ,  $i$  is considered more deserving than  $j$ : thus, social welfare increases with a small transfer of income from  $j$  to  $i$ . In essence, utilitarianism is averse to inequality in marginal utilities. In contrast, the equal-sacrifice social welfare function compares individuals by their level of sacrifice. If  $i$  makes a larger sacrifice than  $j$ ,  $i$  is more deserving than  $j$  and social welfare increases with a small transfer of income from  $j$  to  $i$ . The equal-sacrifice social welfare function is then averse to inequality in sacrifice.

Thus, when there is no asymmetric information, the tax scheme supported by the equal-sacrifice social welfare function is the one that equalizes the sacrifice incurred by each individual, exactly as advocated by proponents of the equal-sacrifice principle. In the second-best Mirrleesian setting—with asymmetric information—the equal-sacrifice social welfare function recommends taxation policies that are less redistributive than those of utilitarianism are. The concern for equality of sacrifice requires assigning a larger relative weight to high-skilled individuals than does utilitarianism and reduces redistribution.

We then apply our criterion to the US economy. We conduct a standard Mirrleesian optimal taxation simulation, following the exercise by Mankiw, Weinzierl, and Yagan (2009). While the utilitarian criterion recommends marginal tax rates above 60% (and up to 80%), our criterion recommends a rate that is about 20 percentage points lower. Furthermore, the utilitarian criterion subsidizes all workers in the bottom 35% of the income distribution, while our criterion does so only for the bottom 15%. In fact, the equal-sacrifice tax schedule is roughly in line with that of the current US system. An important implication emerges. The large difference between the utilitarian optimum and the real-world tax schedules might not necessarily be driven by the political influence of high-income individuals. Rather, the observed tax schedules might reflect the adoption of principles of distributive justice inspired by equal sacrifice, similar to the ones proposed here and in line with the survey evidence on ethical views on taxation (Weinzierl, 2014; Saez and Stantcheva, 2016).

Historically, equal sacrifice was conceived as a standard of perfection (Mill, 1848), and the debate was centered on the definition of sacrifice (Musgrave, 1959). One possibility is to measure sacrifice in terms of absolute or proportional loss of consumption due to taxation. However, this alternative disregards individuals' labor supply choices and, thus, can attribute an increase in sacrifice to an individual who is made better off. Another possibility is to measure sacrifice in terms of utility losses/gains with re-

spect to a reference level (Pigou, 1928). Interestingly, minimizing the sum of utility losses/gains turns out to be equivalent to utilitarianism. However, as in utilitarianism, this raises ethical concerns for the interpersonal comparability of utilities and empirical concerns for their real-world applicability. Closer to our contribution, Young (1988) has characterized a version of the equal-sacrifice principle that does not require utility information. In this case, the utility function adopted to evaluate the sacrifices of individuals is selected for the purpose of the evaluation and need not be related to individuals’ preferences.<sup>4</sup> However, doing so leads to a violation of the Pareto principle. Our approach extends Young’s in two directions. First, it respects the preferences of individuals: labor supply choices are evaluated through the preferences of each individual, while interpersonal comparisons of sacrifice are established through a counterfactual situation of *laissez-faire* labor supply. Second, it allows evaluating deviations from equal sacrifice by prioritizing individuals incurring a larger sacrifice. Thus, in a first-best setting, Young’s equal-sacrifice solution—corrected to respect preferences—emerges. In an incomplete information setting, equality of sacrifice is often too costly, and optimal taxation compromises between equality of sacrifice and efficiency.<sup>5</sup>

A very general approach to welfare criteria that can accommodate equal-sacrifice concerns is to let the social value of one more dollar for an individual—so called social marginal welfare weights—depend not only on the well-being of that individual, but also on some measure of her sacrifice (Saez and Stantcheva, 2016). Our characterization clarifies which social marginal welfare weights emerge when abiding by principles of fairness inspired by equal sacrifice. Moreover, our results identify these social marginal welfare weights for any allocation and not just in the vicinity of the local optimum.

The theory of justice developed here also provides a modern interpretation of “justice as mutual advantage” (Gauthier, 1986), which relates to the axiomatic work on Nash’s bargaining theory and its more recent extensions (Binmore, Rubinstein, and Wolinsky,

---

<sup>4</sup>Young (1990), Ok (1995), and Mitra and Ok (1996) build on this ordinal version of equal sacrifice and discuss the relationship between equal sacrifice and progressivity. Chambers and Moreno-Ternero (2017) introduce a concern for poverty. Stovall (2020) provides an improved characterization. He also suggests a generalization of equal sacrifice, allowing the utility functions to be person-specific.

<sup>5</sup>Recent proposals to capture equal sacrifice in a social welfare function are due to Weinzierl (2012) and Jessen, Metzing, and Rostam-Afschar (2019). Both criteria support policies that move away from an equal-sacrifice allocation, even in the absence of behavioral responses (unless one takes the limit case of infinite aversion to inequality in sacrifice). This drawback does not emerge in our setting. Alternative criteria for income taxation are recently reviewed in Fleurbaey and Maniquet (2018) (see also Saez and Stantcheva, 2016). The recent literature mostly addresses the issue of preference heterogeneity (such as Fleurbaey and Maniquet, 2006, and, in an abstract setting, Piacquadio, 2017), which, in the context of equal sacrifice, we leave to future work.



1986; Rubinstein, Safra, and Thomson, 1992). These theories unfold around two key ethical choices: the definition of equality and the comparative evaluation of inequalities. Crucially, bargaining theories of justice rely on cardinal and interpersonally comparable information about individuals' utilities to measure equality and evaluate inequalities. In contrast, here both the definition of equality and the evaluation of inequalities are endogenous. Finally, the characterizations of bargaining theories of justice generally build on scale invariance, while here we remain closer to the utilitarian tradition and require an additively separable representation of the welfare criterion.<sup>6</sup>

The rest of the paper is organized as follows. Section 2 contains an illustration of the criterion and a comparison with utilitarianism. Section 3 presents the formal model and the axioms. Section 4 discusses the characterization result. Section 5 explores the implications of the criterion with a simulation of the optimal tax system for the US economy. Section 6 briefly concludes. All the proofs are in the appendix.

## 2 A simple illustration

We illustrate our approach in a Mirrleesian model with quasi-linear utilities. Individuals' preferences over consumption  $c$  and labor supply  $\ell$  are represented by a utility function  $u(c, \ell) = c - v(\ell)$  with  $v', v'' > 0$ . Individuals differ in their labor market productivity: each individual  $i$  is characterized by the wage rate  $w_i > 0$ .

### 2.1 The welfare criterion

We next introduce a simple version of our welfare criterion. The first step is to define how to measure and compare the sacrifice of any two individuals  $i$  and  $j$ . A natural starting point is the **laissez-faire allocation**. At the laissez-faire allocation, no taxes are levied and each individual  $i$  maximizes her utility over the budget set  $B_i \equiv \{(c_i, \ell_i) \mid c_i \leq w_i \ell_i\}$ . Let  $(\bar{c}_i, \bar{\ell}_i)$  denote the **laissez-faire bundle** of  $i$ . Clearly, at the reference laissez-faire allocation no individual makes any sacrifice.<sup>7</sup>

---

<sup>6</sup>Scale invariance requires the ranking of alternatives remain unchanged when rescaling alternatives. Scale invariance is logically independent of additive separability. Interestingly, the Nash and the Kalai–Smorodinsky bargaining solutions are characterized based on scale invariance, but also satisfy separability (on a subdomain of alternatives).

<sup>7</sup>Piketty and Saez (2013) have emphasized how utilitarianism fails to ensure laissez-faire prevails even when all agents have the same productivity level (see also Jacquet and Van de Gaer, 2011, and Fleurbaey and Maniquet, 2018).

Assume instead individual  $i$  consumes  $c_i$  and works the laissez-faire labor supply  $\bar{\ell}_i$ . Then, her **(proportional) sacrifice** is  $(\bar{c}_i - c_i) / \bar{c}_i$ , that is, the ratio of the (implicit) tax contribution  $\bar{c}_i - c_i$  to the before-tax income  $\bar{c}_i = w_i \bar{\ell}_i$ .<sup>8</sup>

The key fairness idea is the following. When two individuals incur the same level of sacrifice, society ought to be indifferent between assigning a marginal increase in consumption to either of them. This ethical stand leads to the **proportional-sacrifice social welfare function**, formally characterized as a special case in Section 4. For each individual  $i$ , define the **equivalent consumption** at  $(c_i, \ell_i)$  as the level of consumption  $k$  that makes the individual indifferent between the bundle  $(c_i, \ell_i)$  and consuming  $k$  while working the laissez-faire labor supply  $\bar{\ell}_i$ . Formally,  $e_i(c_i, \ell_i) = k$  if and only if  $u(k, \bar{\ell}_i) = u(c_i, \ell_i)$ . The proportional-sacrifice social welfare function is defined as

$$W^p \equiv \sum_i \bar{c}_i^\gamma \frac{[e_i(c_i, \ell_i)]^{1-\gamma}}{1-\gamma}, \text{ with } \gamma > 0.$$

The parameter  $\gamma$  measures the willingness of society to avoid inequalities in the level of sacrifice incurred by individuals. At the limit for  $\gamma = 0$ , society is indifferent to such inequalities and social welfare simplifies to the simple sum of individuals' utilities (social marginal welfare weights are constant). As  $\gamma$  increases, society is less and less willing to trade off inequalities in sacrifice against a larger sum of consumption. At the limit for  $\gamma \rightarrow \infty$ , society attributes full priority to the individual with the largest sacrifice.

The equivalent consumption function  $e_i$  is a representation of the preferences of individual  $i$ :  $e_i(c_i, \ell_i) = u(c_i, \ell_i) + v(\bar{\ell}_i)$ . Thus, society maximizes the sum of weighted and transformed equivalent consumptions of individuals.

The weight attached to the equivalent consumption of each individual depends, through the laissez-faire bundle, on her skill level. This dependence is crucial to ensure equal consideration for all individuals when they incur the same sacrifice. To see this, note that the social marginal welfare weight (Saez and Stantcheva, 2016) of an individual at bundle  $(c_i, \bar{\ell}_i)$  is

$$\frac{\partial W^p}{\partial c_i} = \frac{\partial}{\partial c_i} \left( \bar{c}_i^\gamma \frac{[e_i(c_i, \ell_i)]^{1-\gamma}}{1-\gamma} \right) = \left( \frac{c_i}{\bar{c}_i} \right)^{-\gamma}.$$

---

<sup>8</sup>If this definition of sacrifice was extended to any level of labor supply, it would necessarily be independent of the utility cost of working and thus would lead to violations of the Pareto principle.

The factor  $\bar{c}_i^\gamma$ —placing a larger weight on the utilities of high-skilled individuals—is key in achieving equal concern for sacrifice. When two individuals  $i$  and  $j$  incur the same sacrifice  $(\bar{c}_i - c_i)/\bar{c}_i = (\bar{c}_j - c_j)/\bar{c}_j$ , also  $c_i/\bar{c}_i = c_j/\bar{c}_j$  and society is indifferent between allocating a marginal increase in consumption to either  $i$  or  $j$ .

## 2.2 A comparison with utilitarianism: first best

We first assume away the asymmetric information problem: the government covers the budget requirement  $R$  by levying an individual-specific lump-sum tax  $T_i$ . By the quasi-linear utility function and lump-sum taxation, the labor supply is at the laissez-faire level. Thus, the maximization problem of a (generalized) utilitarian society (with isoelastic inequality aversion  $\rho \geq 0$ ) simplifies to

$$\begin{aligned} \max_{\{T_i\}} \quad & \sum \frac{[u(w_i \bar{\ell}_i - T_i, \bar{\ell}_i)]^{1-\rho}}{1-\rho} \\ \text{s.t.} \quad & \sum_i T_i \geq R. \end{aligned}$$

Similarly, the first-best maximization problem for the proportional-sacrifice social welfare function simplifies to

$$\begin{aligned} \max_{\{T_i\}} \quad & \frac{(w_i \bar{\ell}_i)^\gamma (w_i \bar{\ell}_i - T_i)^{1-\gamma}}{1-\gamma} \\ \text{s.t.} \quad & \sum_i T_i \geq R. \end{aligned}$$

The first-best optimum for the utilitarian criterion is instructive. If  $\rho = 0$ , the distribution of consumption is irrelevant and the lump-sum taxes are undefined. However, when  $\rho > 0$  and small (formally at the limit for  $\rho \rightarrow 0$ ), the optimal lump-sum taxes are set to equalize the levels of consumption. This redistribution is extreme: at the optimum, higher-skilled individuals will achieve a lower level of utility. Utilitarianism “forces” higher-skilled individuals to produce for the sake of providing more consumption to lower-skilled individuals. Only at the limit for  $\rho \rightarrow \infty$ , when the criterion is “Rawlsian,” are utilities equalized.

In contrast, the first-best optimum for the proportional-sacrifice social welfare function requires the lump-sum tax to be a fixed proportion of the gross income, independently of  $\gamma$ .<sup>9</sup> Combining the first-order conditions on the lump-sum taxes of  $i$  and  $j$

---

<sup>9</sup>Note that  $\gamma = 0$  is excluded and emerges only as a limit case. The reason is technical. When  $\gamma = 0$ , the criterion is insensitive to the distribution of individuals’ sacrifice and thus the notion of sacrifice itself cannot be singled out from the axioms.

leads to

$$\frac{T_i^*}{T_j^*} = \frac{w_i \bar{\ell}_i}{w_j \bar{\ell}_j} = \frac{\bar{c}_i}{\bar{c}_j} = \frac{c_i^*}{c_j^*}.$$

Therefore, the higher-skilled individuals combine a larger labor supply with a larger consumption. In contrast, lower-skilled individuals work less and consume less. This correlation between consumption and labor supply emerges from the proportional-sacrifice social welfare function attributing relatively more weight to the higher-skilled individuals. All individuals must contribute to the tax burden so as to incur the same level of sacrifice, here measured as a proportion of laissez-faire consumption.

Nevertheless, the proportional-sacrifice optimum does not ensure that the utility of higher-skilled individuals be higher than that of lower-skilled individuals. This should not come as a surprise, because it depends on the budget of the government. To illustrate, consider the extreme case of the budget of the government being equal to the laissez-faire income ( $R = \sum_i w_i \bar{\ell}_i$ ). At the optimum, each individual's tax burden is her laissez-faire income. Then, consumption is zero and equal across individuals, while the labor supply is unchanged and penalizes (in terms of utility) those individuals with a higher skill level (who supply more labor). Consider instead the other extreme, when the government need not raise money ( $R = 0$ ). The utilitarian optimum equalizes consumptions, while the proportional-sacrifice optimum is to avoid any taxation. Then, higher-skilled individuals are better off than lower-skilled ones.

## 2.3 A comparison with utilitarianism: second best

Assume there are two types of individuals,  $h$  and  $l$ , with  $w_h > w_l$ . Types are private information of the individuals. The government sets a tax schedule  $T$  associating a level of taxes  $T(y)$  to each level of income  $y$ . Let  $y_i \equiv w_i \ell_i$  and let  $P_i$  be a real-valued, strictly increasing and concave function for  $i = h, l$ , named the **Pareto function**. Then, both the utilitarian and sacrifice-based views are captured by the following sum-of-utility social welfare function

$$W = \sum_{i=h,l} P_i \left( u \left( c_i, \frac{y_i}{w_i} \right) \right).$$

The utilitarian criterion emerges when the Pareto functions  $P_h$  and  $P_l$  are equal across individuals, increasing, and concave. The proportional-sacrifice social welfare function emerges when the Pareto functions take the form  $P_i(z) = \bar{c}_i^\gamma \frac{(z + v(\bar{\ell}_i))^{1-\gamma}}{1-\gamma}$  for each  $z \in \mathbb{R}$  and  $i = h, l$ . Since  $W$  can represent any criterion that respects Pareto efficiency, the

government's problem has the same structure as in Stiglitz (1982).

The government maximizes the social welfare function  $W$  subject to the budget revenue requirement

$$\sum_{i=h,l} T(y_i) = \sum_{i=h,l} (y_i - c_i) \leq R,$$

and subject to the incentive compatibility constraints

$$ICC_h : u\left(c_h, \frac{y_h}{w_h}\right) \geq u\left(c_l, \frac{y_l}{w_h}\right),$$

$$ICC_l : u\left(c_l, \frac{y_l}{w_l}\right) \geq u\left(c_h, \frac{y_h}{w_l}\right).$$

For both the utilitarian and proportional-sacrifice criteria, we can exclude that  $ICC_l$  binds.

**Case 1.  $ICC_h$  binds.** When the incentive compatibility constraint of the skilled type is binding, the optimal tax schedule requires  $T'(y_h) = 0$  and  $T'(y_l) > 0$ . The labor supply choice of the skilled type is undistorted, while the labor supply choice of the low-skill type is distorted downward. The government trades off the efficiency cost of labor supply distortions against the information rent of the skilled type.

**Case 2. Neither  $ICC_h$  nor  $ICC_l$  binds.** The optimal tax schedule requires  $T'(y_h) = T'(y_l) = 0$ . Labor supply choices are not distorted and the first-best allocation is implemented.

With a utilitarian government, only case 1 is possible. The government would like the high-skill type to achieve a lower utility than the low-skill type. However, the skilled can always mimic the low skilled and achieve a higher utility (the utility cost of earning  $y_l$  is smaller). Thus, the incentive compatibility constraint of the high-skill type is always binding.

With the proportional equal-sacrifice government, both case 1 and case 2 are possible. There exists a threshold budget requirement level  $\bar{R} > 0$ . When the government budget requirement is small (when  $R \leq \bar{R}$ ), the incentive compatibility constraints are not binding and the first-best equal-sacrifice allocation can be implemented. When the government budget requirement is large (when  $R > \bar{R}$ ), the incentive compatibility constraint of the high-skill type binds.

Thus, with asymmetric information, the incentive compatibility constraints significantly limit the extent of redistribution the government can implement. Then, the

first-best goal of “favoring” the lower-skilled individuals is unachievable and is only indirectly reflected in the second-best optimum. Thus, asymmetric information reduces the gap between the implications of utilitarianism and common views on the optimal level of income redistribution.

Relatedly, the differences in second-best policies between welfare criteria might not be very large. Whether they are not depends on a combination of factors. First, a sufficiently large government budget requirement ensures that the incentive compatibility constraint of the high-skill type is binding for both criteria. Second, a large elasticity of labor supply limits the space of feasible redistributive policies for the government and thus forces the policies to be similar. Third, the inequality attitudes toward utility (for the utilitarian criterion) and toward sacrifice (for the proportional-sacrifice criterion) are not directly comparable and interact with both the budget requirement and labor supply elasticities.

## 3 Model and axioms

### 3.1 Model

The set of individuals is  $I \subset \mathbb{N}$ ; it is finite and satisfies  $|I| \geq 3$ . Individuals differ by their labor skills, reflected in their wage rates: for each  $i \in I$ , let  $w_i > 0$  denote the wage rate of individual  $i$ .

Each individual  $i \in I$  supplies labor  $\ell_i \geq 0$ , earns income  $y_i \equiv w_i \ell_i$ , and consumes  $c_i \geq 0$ . Her preferences are represented by a utility function  $u(c_i, \ell_i)$ , which is continuous, increasing in  $c_i$ , decreasing in  $\ell_i$ , and strictly concave. We assume consumption is an essential good, that is,  $\lim_{c \rightarrow 0} u_c = \infty$ .

An allocation  $a \equiv (\{c_i, \ell_i\}_{i \in I})$  specifies a bundle  $(c_i, \ell_i)$  for each individual  $i \in I$ . Let  $A$  be the set of all allocations.

The government collects taxes to cover an exogenous budget  $R$ . Since skills/wages are private information of individuals, taxes can depend only on individuals' incomes. Let the tax function be denoted by  $T : \mathbb{R} \rightarrow \mathbb{R}$ . For each  $i \in I$ , after-tax income is  $y_i - T(y_i)$ .

The goal of the government is to set the tax function to maximize social preferences. Social preferences are a complete, transitive, and continuous preference relation  $\succsim$  on the set of allocations  $A$ . For each pair  $a, a' \in A$ ,  $a \succsim a'$  means that  $a$  is socially at

least as desirable as  $a'$ . The asymmetric and symmetric counterparts of  $\succsim$  are denoted  $\succ$  and  $\sim$ . Social welfare can be represented by a continuous social welfare function  $W : A \rightarrow \mathbb{R}$ . Thus, for each pair  $a, a' \in A$ ,  $a \succsim a'$  if and only if  $W(a) \geq W(a')$ .

As standard, we require social preferences to satisfy the Pareto principle. In other words, if individuals are made better off, social welfare cannot decrease.

**Efficiency:** *For each pair  $a, a' \in A$ , if  $u(c_i, \ell_i) \geq u(c'_i, \ell'_i)$  for each  $i \in I$  and  $u(c_i, \ell_i) > u(c'_i, \ell'_i)$  for some  $i \in I$ , then  $a \succ a'$ .*

Next, we impose inequality aversion on social preferences by imposing strict convexity.<sup>10</sup>

**Inequality aversion:** *For each pair  $a, a' \in A$  and each  $\beta \in (0, 1)$ ,  $a \sim a'$  implies  $\beta a + (1 - \beta) a' \succ a$ .*

Finally, we impose that social welfare comparisons do not depend on the bundle assigned to an unconcerned individual. For the sake of simplicity, denote by  $(a_i, a_{-i})$  the allocation  $a \in A$  that assigns  $a_i \equiv (c_i, \ell_i)$  to individual  $i$  and  $a_{-i} \equiv (c_j, \ell_j)_{j \in I \setminus \{i\}}$  to the other individuals.

**Separability:** *For each  $a, a' \in A$ , each  $i \in I$ , and each  $\bar{a}_i = (\bar{c}_i, \bar{\ell}_i)$ ,  $(a_i, a_{-i}) \succsim (a_i, a'_{-i})$  if and only if  $(\bar{a}_i, a_{-i}) \succsim (\bar{a}_i, a'_{-i})$ .*

*Efficiency*, *inequality aversion*, and *separability* imply the social welfare function belongs to a very general class of criteria. By *efficiency*, society evaluates individuals through their own preferences:  $W$  can be written as a function of the utilities achieved by each individual. By *inequality aversion*, social preferences are strictly convex with respect to the allocation and, thus,  $W$  is strictly concave in its arguments. By *separability*, the assignment of individual  $i$  does not matter for how society trades off the utility of individuals  $j$  and  $k$ ; thus,  $W$  is additively separable.

Let the **Pareto functions**  $(P_i)_{i \in I}$  be real-valued, continuous, and strictly increasing functions such that for each  $i \in I$ ,  $P_i(u(c_i, \ell_i))$  is strictly concave. A social welfare function  $W : A \rightarrow \mathbb{R}$  is “sum of utilities” if for each  $a \in A$ ,

---

<sup>10</sup>Convexity is significantly weaker than what is widely assumed in the literature, where this condition is supplemented with some form of symmetry or anonymity. In fact, most social welfare functions satisfy convexity. Here, strict convexity avoids a technical issue: when social preferences are linear, inequalities are irrelevant and the axioms cannot identify how to measure inequalities in sacrifice. Convex social preferences then emerge as a limit case.

$$W(a) = \sum_{i \in I} P_i(u(c_i, \ell_i)). \quad (1)$$

This family of social welfare functions is significantly more general than usual. Unlike in Mirrlees (1971), the functions  $(P_i)_{i \in I}$  need not be equal across individuals. Unlike in weighted utilitarianism (d'Aspremont and Gevers, 1977; Maskin, 1978), the functions  $(P_i)_{i \in I}$  need not be increasing affine transformations (thus, specifying Pareto weights). As we discuss in the following, this degree of freedom is necessary to incorporate principles of justice inspired by equality of sacrifice. Before doing so, we formalize the implication of the above axioms. The proof is in Appendix A.

**Lemma 1.** *Social preferences  $\succsim$  satisfy efficiency, inequality aversion, and separability if and only if  $\succsim$  can be represented by a sum-of-utilities social welfare function.*

### 3.2 Averting unequal sacrifice

In this subsection, we introduce axioms that ultimately (i) restrict the admissible definitions of sacrifice, (ii) discipline how social preferences ought to compare distributions of sacrifice, and thus (iii) characterize the Pareto functions  $(P_i)_{i \in I}$  consistent with social aversion to unequal sacrifice.<sup>11</sup>

To start with, define the **laissez-faire allocation**  $\bar{a}$ . At the laissez-faire allocation, each individual freely chooses how much labor to supply and consumes her entire income. Formally, the laissez-faire bundle of each individual  $i \in I$  is  $(\bar{c}_i, \bar{\ell}_i)$  such that  $u(\bar{c}_i, \bar{\ell}_i) \geq u(c_i, \ell_i)$  for each  $(c_i, \ell_i)$  with  $c_i \leq y_i \equiv \ell_i w_i$ . Crucially, since individuals entirely appropriate the returns from their own work, individuals make no sacrifice.<sup>12</sup>

Let individual  $i$ 's **(implicit) tax burden** at the bundle  $(c_i, \bar{\ell}_i)$  be measured by the difference in consumption with the laissez-faire allocation, that is,  $b_i \equiv \bar{c}_i - c_i$ . By

---

<sup>11</sup>A different approach is to start with a cardinally measurable and interpersonally comparable index of sacrifice of individuals. With such rich information, one could just require that social welfare decreases when sacrifice is transferred from a low-sacrifice individual to a high-sacrifice individual. However, without a theory of how to measure sacrifice at each allocation, the corresponding welfare criterion would not be applicable. Our approach is more ambitious. Here, the index of sacrifice emerges endogenously from the axioms as a way to represent the social ranking of allocations.

<sup>12</sup>This is a natural choice in the Mirrleesian framework. In general, however, the no-sacrifice allocation is a more controversial choice. For example, with general equilibrium effects on wages, the level of taxation can affect the relative productivity of individuals. More drastically, the government might not be able to ensure property rights without taxation, making the laissez-faire allocation undefined. Importantly, our results extend to alternative choices of no-sacrifice allocations when the axioms are modified accordingly.



definition of laissez-faire, an individual working  $\bar{\ell}_i$  has a gross income of  $\bar{y}_i \equiv \bar{\ell}_i w_i$ . At the laissez-faire allocation, individual  $i$  would consume the entire income  $\bar{c}_i = \bar{y}_i$ . At the bundle  $(c_i, \bar{\ell}_i)$  instead, individual  $i$  works the same time  $\bar{\ell}_i$  but consumes  $c_i$ . Then, the tax burden is the difference between these consumption levels.<sup>13</sup> Intuitively, each individual's sacrifice increases the larger her tax burden is.

The first principle of equal sacrifice tells us that society should avert situations whereby one individual makes a sacrifice while another individual does not. In other words, individuals should solidarily bear the cost of taxation. We state this ideal in the form of a transfer principle. More precisely, assume that at an allocation  $a \in A$ , individual  $i$  has a positive tax burden  $b_i > 0$  ( $i$  consumes less than at the laissez-faire bundle), while individual  $j$  has a negative tax burden  $b_j < 0$  ( $j$  consumes more than at the laissez-faire bundle). This distribution of the tax burden is unfair according to the equal-sacrifice principle. Then, *ceteris paribus*, a transfer of consumption from  $i$  to  $j$  increases further the tax burden of  $i$  while decreasing that of  $j$ . This distribution of the tax burden is even more unfair and thus social welfare cannot be higher.

**Tax solidarity:** *For each pair  $a, a' \in A$ , each pair  $i, j \in I$ , and each  $\varepsilon > 0$ , such that:*

- $b'_i - \varepsilon = b_i \geq 0 \geq b_j = b'_j + \varepsilon$ ;
- $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and
- $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ ;

*then,  $a \succsim a'$ .*

*Tax solidarity* implies that (together with *efficiency*) no taxation is optimal when the government's budget is  $R = 0$ . The intuition is immediate. The  $R = 0$  budget condition means taxation is not needed, but can be introduced for the sake of redistribution. However, *tax solidarity* tells us that redistribution away from the laissez-faire allocation cannot improve social welfare. Thus, when  $R = 0$ , the laissez-faire allocation is optimal and no taxation should be adopted. Moreover, no individual makes any sacrifice and equal sacrifice is achieved.<sup>14</sup>

---

<sup>13</sup>The fixed labor supply can be interpreted as a *ceteris paribus* assumption. Its importance is easily explained: when paying a positive income tax, individuals may adjust labor supply upward to compensate for the lost income. If labor supply is allowed to vary, the extent of this income effect will matter for measuring the tax burden. In some situations, the individual might end up consuming more than at the laissez-faire (when leisure is a Giffen good), but is still worse off.

<sup>14</sup>This axiom rules out redistribution motives. However, it captures the ideal of equal sacrifice.

The next ethical principle deals with a different type of unfairness. Without loss of generality, assume individual  $i$ 's consumption at the laissez-faire allocation is larger than  $j$ 's, that is,  $\bar{c}_i \geq \bar{c}_j$ . At an allocation  $a \in A$ , individual  $i$  has a smaller tax burden than  $j$  does, that is,  $0 \leq b_i < b_j$ ; labor supply is that of the laissez-faire allocation. Individuals earn incomes  $w_i \bar{\ell}_i = \bar{c}_i \geq \bar{c}_j = w_j \bar{\ell}_j$  and consume  $c_i > c_j$ . Crucially,  $b_i < b_j$  implies the difference in earnings is smaller than the difference in consumption: the tax burden imposed on individuals exacerbates inequality. Consider now increasing further the tax burden of  $j$ , while further reducing that of  $i$ . This transfer of consumption makes the allocation more unfair and thus cannot improve social welfare.

**Fair burden:** *For each pair  $a, a' \in A$ , each pair  $i, j \in I$  with  $\bar{c}_i \geq \bar{c}_j$ , and each  $\varepsilon > 0$ , such that:*

- $0 \leq b'_i + \varepsilon = b_i < b_j = b'_j - \varepsilon$ ;
- $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and
- $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ ;

*then,  $a \succsim a'$ .*

*Fair burden* deals with situations whereby the sacrifice of individual  $i$  (who is better off at the laissez-faire allocation) is too small relative to that of some other individual. Next, we discipline how social welfare deals with situations whereby the sacrifice of individual  $i$  is too large.

As before, assume individual  $i$ 's consumption at the laissez-faire allocation is larger than  $j$ 's, that is,  $\bar{c}_i \geq \bar{c}_j$ . At allocation  $a \in A$ , individual  $i$ 's consumption is smaller than  $j$ 's, that is,  $c_i < c_j$ ; labor supply is that of the laissez-faire allocation. Individuals earn incomes  $w_i \bar{\ell}_i = \bar{c}_i \geq \bar{c}_j = w_j \bar{\ell}_j$  and consume  $c_i \leq c_j$ . The sacrifice of  $i$  is so large that, net of the sacrifice, the consumption of  $i$  is smaller than that of  $j$ . Consider now making the sacrifice of  $i$  even harsher by reducing her consumption for the benefit of  $j$ . This change makes the allocation more unfair and cannot improve social welfare.

---

The interpretation is that individuals entirely deserve the wage rate their labor supply gives, not more nor less. In contrast, utilitarianism suggests that individuals do not deserve their wage rates, pushing for redistribution (Feldstein, 1976). The fathers of the equal-sacrifice principles suggest to deal with poverty by ensuring individuals have a subsistence consumption. Within our framework, a possible solution is to introduce a threshold of consumption  $s$  and measure the tax burden by  $b_i \equiv \max \{\bar{c}_i, s\} - c_i$ . Then, at laissez-faire allocation, lower-skilled individuals have a positive "burden" (when  $s > \bar{c}_i$ ) and redistribution is optimal, even when  $R = 0$ .

**Fair reward:** For each pair  $a, a' \in A$ , each pair  $i, j \in I$  with  $\bar{c}_i \geq \bar{c}_j$ , and each  $\varepsilon > 0$ , such that:

- $c'_i + \varepsilon = c_i < c_j = c'_j - \varepsilon$ ;
- $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and
- $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ ;

then,  $a \succsim a'$ .

We represent the three axioms introduced above in figure 1. On the Cartesian plane, the consumptions of individuals  $i$  and  $j$  are represented on the axes. Consumption levels  $\bar{c}_i$  and  $\bar{c}_j$  are those of the laissez-faire allocation.<sup>15</sup> Without loss of generality, here  $i$  is again the individual consuming more at the laissez-faire allocation, that is,  $\bar{c}_i \geq \bar{c}_j$ . Not represented, the labor supply choices are those of the laissez-faire allocation.

The northwest and southeast areas from the laissez-faire consumptions are those where *tax solidarity* applies. These areas are characterized by one individual making a sacrifice, while the other does not. The arrow pointing toward the laissez-faire consumptions  $(\bar{c}_i, \bar{c}_j)$  represents the direction of increasing social welfare.

Southwest of the laissez-faire consumptions is the area where both individuals make a sacrifice. *Fair burden* applies below the 45 degree line from the laissez-faire consumptions. In this area, the tax burden of  $j$  is larger than that of  $i$ . Again, the arrow suggests that transferring consumption from  $i$  to  $j$  increases social welfare.

Finally, the portion of the area above the 45 degree line from the origin is such that the consumption of  $j$  is larger than that of  $i$ . In this area, *fair reward* applies. In this area, the tax burden of  $i$  is so large that her consumption is now smaller than  $j$ 's. The arrow points to reducing  $j$ 's consumption for the benefit of  $i$ 's. *Fair reward* suggests this transfer increases social welfare.

To further illustrate our axioms, it is helpful to contrast them to the Pigou-Dalton transfer principle, satisfied by the utilitarian criterion. The Pigou-Dalton transfer principle requires that a transfer from a poorer to a richer individual reduces social welfare.

---

<sup>15</sup>These consumption levels can be interpreted as legitimate claims of individuals. In a “claims problem,” the objective is to fairly allocate an endowment that is not sufficient to cover the claims (Thomson, 2019). Interestingly, the “path of awards of a rule”—that is, the locus of assignments associated to each level of endowment—is related to our sacrifice function—that is, the locus of allocations where individuals incur the same sacrifice (a similar point emerges also in Chambers and Moreno-Ternero, 2017 and Stovall, 2020). Our focus on complete rankings and the framework (multidimensional and with individuals’ choices) make the axiomatic analysis very different.

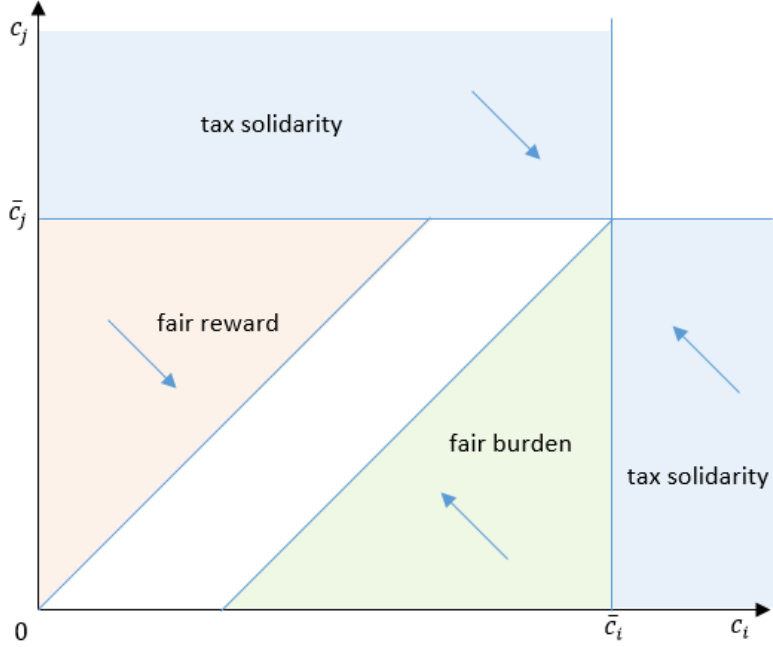


Figure 1: Equal-sacrifice principles.

Let utility be additively separable between consumption and leisure. Then in the above graph, the arrow of increasing social welfare for the Pigou-Dalton principle always points to the 45 degree line, independently of the laissez-faire income of the individuals. Thus, the standard Pigou-Dalton transfer principle agrees with the implication of both *fair burden* and *fair reward*, while these remain weaker.<sup>16</sup> Such a weakening is necessary to accommodate the view expressed by survey respondents (see Saez and Stantcheva, 2016). They overwhelmingly support the view that a family earning 50,000\$ and paying 15,000\$ of taxes is more deserving than an (otherwise identical) family earning 40,000\$ and paying 5,000\$. This means that—in contrast to the Pigou-Dalton transfer principle and utilitarianism—the arrow of social improvement for the survey respondents points away from an equal after-tax income of 35,000\$ to the benefit of the higher-income earner. Similar survey evidence is shown in Schokkaert and Devooght (2003) and Weinzierl (2014).

<sup>16</sup>As the graphical representation suggests, the axiom that conflicts with the Pigou-Dalton transfer principle and, thus, rules out utilitarianism is *tax solidarity*.

## 4 The sacrifice-based welfare criteria

### 4.1 Comparisons of sacrifice

First, for each allocation  $a \in A$  and each individual  $i \in I$ , let the **equivalent consumption of  $i$  at  $a$**  be the level of consumption  $e_i(c_i, \ell_i)$  such that

$$e_i(c_i, \ell_i) = k \iff u(c_i, \ell_i) = u(k, \bar{\ell}_i),$$

where  $\bar{\ell}_i$  is the labor supply at the laissez-faire allocation. The function  $e_i(c_i, \ell_i)$  tells us the level of consumption that, when combined with the laissez-faire labor supply, gives the same utility as that of the bundle  $(c_i, \ell_i)$ .

Next, we define the **sacrifice function**  $S : \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$ . This function measures the sacrifice made by each individual in an interpersonally comparable way. Let individual  $i \in I$  be assigned the bundle  $(c_i, \bar{\ell}_i)$ ; then,  $i$ 's sacrifice is given by  $S(c_i; \bar{\ell}_i)$ . More generally,  $i$ 's sacrifice at bundle  $(c_i, \ell_i)$  is given by  $S(e(c_i, \ell_i); \bar{\ell}_i)$ . Let the sacrifice function  $S$  be decreasing in the first argument, increasing in the second argument, and continuous. Furthermore, it satisfies the following restrictions:

- [zero sacrifice normalization]  $x = y$  implies  $S(x; y) = 0$ ;
- [slope bound for positive sacrifice] whenever  $S(x; y) = S(x'; y') > 0$ , then  $|x - x'| \leq |y - y'|$ .

Let  $\mathcal{S}$  be the domain of these functions. Importantly, the sacrifice function is ordinal, because it represents only an ordering of levels of sacrifice incurred by any two individuals. In figure 2, we illustrate how the sacrifice function works. On the left part of the Cartesian plane, individual  $i$  faces the wage level  $w_i$  and chooses the utility-maximizing bundle  $(\bar{c}_i, \bar{\ell}_i)$ . Similarly, individual  $j$  with wage level  $w_j$  chooses the utility-maximizing bundle  $(\bar{c}_j, \bar{\ell}_j)$ . The bundles  $(\bar{c}_i, \bar{\ell}_i)$  and  $(\bar{c}_j, \bar{\ell}_j)$  constitute the laissez-faire allocation. The corresponding levels of consumption  $\bar{c}_i$  and  $\bar{c}_j$  are reported by the horizontal axis of the right part of the Cartesian plane, where we represent the sacrifice function through its **isosacrifice curves**. At the 45 degree line, the level of sacrifice is 0. The level of sacrifice decreases with the assigned consumption and increases with the laissez-faire consumption, making the isosacrifice curves increasing. The slope bound implies that, for positive levels of sacrifice, the slope of the sacrifice function cannot exceed 1.

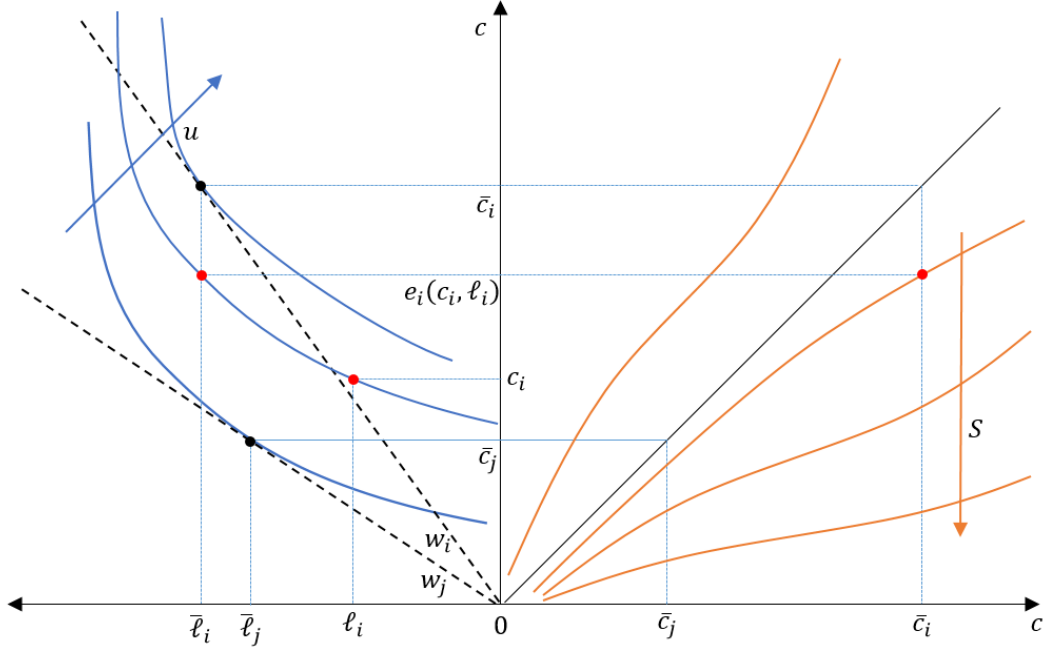


Figure 2: Isosacrifice curves and comparability in terms of sacrifice.

Let individual  $i$  be assigned the bundle  $(c_i, \ell_i)$ . Her level of utility is the same as if she was assigned her equivalent consumption  $e_i(c_i, \ell_i)$  and the laissez-faire labor supply  $\bar{\ell}_i$ . The implicit tax burden of  $i$  is given by the difference between  $\bar{c}_i$  and  $e_i(c_i, \ell_i)$ . Interpersonal comparisons of sacrifice are made through the isosacrifice curve of level  $S(e_i(c_i, \ell_i); \bar{c}_i)$ . Individual  $i$  makes a larger sacrifice than  $j$  whenever  $S(e_i(c_i, \ell_i); \bar{c}_i) \geq S(e_j(c_j, \ell_j); \bar{c}_j)$ .

## 4.2 The measurement of each individual's contribution to social welfare

For each individual  $i$ , let the **Pareto function of  $i$**  be denoted by  $P_i : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ . Then,  $P_i(u(c_i, \ell_i))$  is  $i$ 's **contribution to social welfare** at  $(c_i, \ell_i)$ . As we shall discuss, society maximizes the sum over individuals of these contributions. Before that, we introduce the restrictions on the Pareto functions.

For each  $i$  and each bundle  $(c_i, \ell_i)$ , denote by  $\beta_i(c_i, \ell_i)$  the **social marginal welfare weight** of  $i$  at bundle  $(c_i, \ell_i)$ ; formally:

$$\beta_i(c_i, \ell_i) \equiv \frac{\partial}{\partial c_i} P_i(u(c_i, \ell_i)).$$

Then, we impose there exists a sacrifice function  $S \in \mathcal{S}$  and a real-valued increasing function  $g$  such that for each  $i$  and each  $c_i \in \mathbb{R}_+$ :

1.  $\beta_i(c_i, \bar{\ell}_i) = g(S(c_i; \bar{c}_i)) > 0$ ; and
2.  $P_i(u(c_i, \ell_i))$  is strictly concave in its arguments.

To explain, since the social marginal welfare weights are positive, the Pareto functions are increasing. Thus, the contributions to social welfare of individuals are representations of their preferences. Condition 1 also imposes equality of social marginal welfare weights when the level of sacrifice incurred by individuals is the same, as identified by the sacrifice function  $S$ . Furthermore, since  $g$  is increasing, the social marginal welfare weights are higher for individuals incurring a larger level of sacrifice. Condition 2 ensures concavity of individuals' contributions to social welfare and thus aversion to inequality in sacrifice.

When the above conditions are satisfied, we say a set of Pareto functions  $(P_i)_{i \in I}$  is **consistent** with a sacrifice function  $S$ .

### 4.3 The welfare criterion

The **equal-sacrifice social welfare function**  $W : A \rightarrow \mathbb{R}$  is defined by setting for each  $a \in A$ ,

$$W(a) \equiv \sum_{i \in I} P_i(u(c_i, \ell_i)),$$

where  $(P_i)_{i \in I}$  are Pareto functions consistent with a sacrifice function  $S \in \mathcal{S}$ .

Our main result shows the above ethical principles characterize the family of equal-sacrifice social welfare functions.

**Theorem 1.** *Social welfare  $\succsim$  satisfies efficiency, inequality aversion, separability, tax solidarity, fair burden, and fair reward if and only if it can be represented by an equal-sacrifice social welfare function.*

### 4.4 The proportional-sacrifice social welfare function

In Section 2, we introduced the proportional-sacrifice social welfare functions, a special case of the family of equal-sacrifice social welfare functions. Two ethical choices characterize these criteria. First, the sacrifice function  $S$  is proportional, that is,

$S(c; \bar{c}) = (\bar{c} - c) / \bar{c}$ . Second, the social attitude toward inequality in sacrifice is captured by a unique parameter  $\gamma > 0$ . In the following, we provide a characterization of these criteria.

The following principle strengthens *fair burden* to deal with situations in which the sacrifice of individual  $i$  (who is better off at the laissez-faire allocation) is relatively too small as opposed to that of some other individual  $j$ . In these cases, a regressive transfer from  $j$  to  $i$  cannot improve social welfare.

**Fair relative burden:** *For each pair  $a, a' \in A$ , each pair  $i, j \in I$  with  $\bar{c}_i \geq \bar{c}_j$ , and each pair  $\alpha, \varepsilon > 0$ , such that:*

- $c'_i + \varepsilon = c_i > \alpha \bar{c}_i$  and  $\alpha \bar{c}_j > c_j = c_j - \varepsilon$ ;
- $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and
- $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ ;

*then,  $a \succsim a'$ .*

The next principle strengthens *fair reward* and disciplines situations in which the sacrifice of individual  $i$  (who is better off at the laissez-faire allocation) is relatively too large as opposed to that of some other individual  $j$ . In these cases, a regressive transfer from  $i$  to  $j$  cannot improve social welfare.

**Fair relative reward:** *For each pair  $a, a' \in A$ , each pair  $i, j \in I$  with  $\bar{c}_i \geq \bar{c}_j$ , and each pair  $\alpha, \varepsilon > 0$ , such that:*

- $c'_i + \varepsilon = c_i < \alpha \bar{c}_i$  and  $\alpha \bar{c}_j < c_j = c_j - \varepsilon$ ;
- $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and
- $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ ;

*then,  $a \succsim a'$ .*

We represent these principles in figure 3.

Finally, we introduce a weak form of scale invariance with respect to the tax consumption of individuals. Assume all individuals work at the laissez-faire labor supply. Then, proportional changes in consumptions do not affect how society ranks two allocations.

**Scale invariance:** *For each  $a, a', a'', a''' \in A$  and each  $\kappa > 0$  such that:*



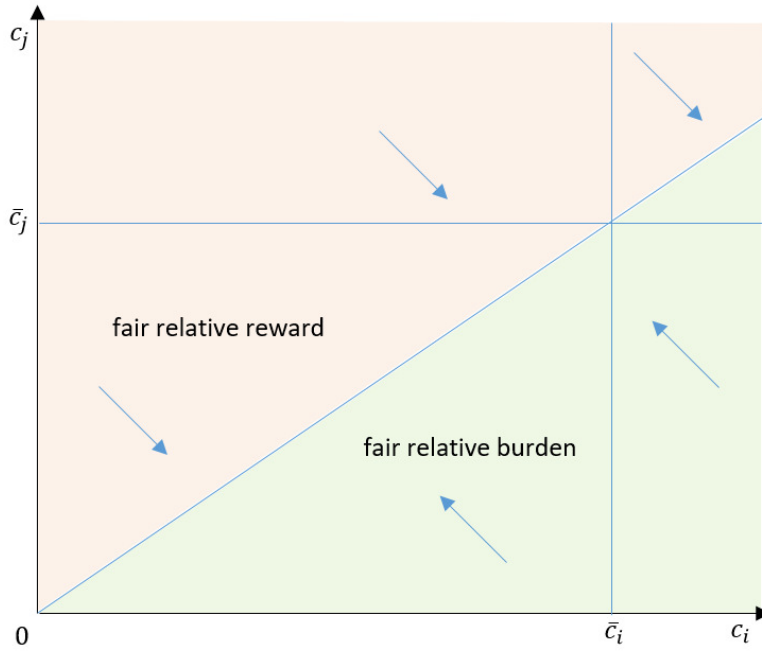


Figure 3: Relative equal-sacrifice principles.

- $c_i = \kappa c_i''$  and  $c_i' = \kappa c_i'''$  for each  $i \in I$ ;
- $\ell_i = \ell_i' = \ell_i'' = \ell_i''' = \bar{\ell}_i$  for each  $i \in I$ ;

then,  $a \succsim a'$  if and only if  $a'' \succsim a'''$ .

The following lemma summarizes the relationships between the above axioms.

**Lemma 2.** For a social welfare  $\succsim$ :

1. fair relative burden implies fair burden;
2. fair relative reward implies fair reward;
3. fair relative burden and fair relative reward imply tax solidarity;
4. fair burden and scale invariance imply fair relative burden;
5. fair reward and scale invariance imply fair relative reward;
6. tax solidarity and scale invariance imply fair relative burden and fair relative reward.

We can now characterize the proportional-sacrifice social welfare function. The proof is similar to that in Piacquadio (2020) and is omitted.

**Theorem 2.** *Social welfare  $\succsim$  satisfies efficiency, inequality aversion, separability, tax solidarity and scale invariance if and only if it can be represented by a social welfare function  $W^p$  such that for each  $a \in A$ ,*

$$W^p(a) \equiv \sum_{i \in I} P_i(u(c_i, \ell_i)),$$

where, for each  $i \in I$ ,  $P_i$  satisfies

$$\frac{\partial}{\partial c_i} P_i(u(c_i, \bar{\ell}_i)) \equiv \beta_i(c_i, \bar{\ell}_i) = \left(\frac{c_i}{\bar{c}_i}\right)^{-\gamma},$$

for some  $\gamma > 0$ .

This result states that, for the proportional-sacrifice social welfare function, the Pareto functions  $(P_i)_{i \in I}$ : (i) need to be consistent with the proportional-sacrifice function  $S(c, \bar{c}) = (\bar{c} - c) / \bar{c}$ ; and that (ii) their derivatives at the laissez-faire labor supply (defining the social marginal welfare weights) need to be a power transformation of  $1 - S(c, \bar{c})$ .

## 5 A simulation exercise

Next, we turn to the continuous version of the Mirrlees model for a simulation exercise. Individuals' wage rates are now continuously distributed according to  $f(w)$  on  $w \in [w^b, w^t]$ .<sup>17</sup> We assume an additively separable utility function  $U(c_i, l_i) = u(c_i) - v(l_i)$  with  $u'(c) > 0, u''(c) < 0$ . Following Saez (2001), the first-order condition for optimal marginal tax rates  $T'(y(w))$  is given by

$$\frac{T'(y(w))}{1 - T'(y(w))} = \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)} \frac{U_c(w)}{w f(w)} \int_w^{w^t} \left( \left( 1 - \frac{P'_\theta(U(\theta)) U_c(\theta)}{\lambda} \right) \frac{1}{U_c(\theta)} f(\theta) d\theta \right),$$

where  $\varepsilon^u(w)$  and  $\varepsilon^c(w)$  are, respectively, the uncompensated and compensated labor supply elasticities. The marginal utility of consumption is  $U_c(w)$  and the marginal social

---

<sup>17</sup>The continuity of the welfare criterion with respect to the types of individuals—here identified by their wage rate—ensures that the continuous version of our criterion is well-defined.

welfare weight of increasing consumption for the agent with skill  $w$  is  $P'_w(U(w))U_c(w)$ . Since the structure of the optimal tax problem is unaffected by the Pareto functions  $P_w$  chosen for each individual, all the standard results apply, including no negative tax rates and the zero marginal tax rate at the upper limit. Moreover, the necessary condition is also sufficient when the single-crossing condition is satisfied ( $y(w)$  non-decreasing).

Unfortunately, little more can be said analytically about optimal tax rates from the Mirrlees problem. Hence, we turn to numerical simulations. Our objective is to highlight the difference between the optimal tax schedules under the utilitarian and the equal-sacrifice social welfare functions. We apply the standard additively separable utility function used in many simulations (e.g., Mankiw et al., 2009, and Saez, 2001):  $U(c_i, l_i) = (c_i^{1-\rho} - 1) / (1 - \rho) - \alpha l_i^\sigma / \sigma$ . Then, equivalent consumption is given by  $e_i(c_i, l_i) = ((1 - \rho) (U(c_i, l_i) + v(\bar{l}_i)) + 1)^{1/(1-\rho)}$ . The marginal social welfare weights for the proportional-sacrifice criterion are

$$\frac{\partial}{\partial c_i} W^p(c_i, l_i) = \left( \frac{e(c_i, l_i)}{\bar{c}_i} \right)^{-\gamma} \frac{\partial e(c_i, l_i)}{\partial c_i} = \frac{1}{1 - \rho} \left( \frac{e(c_i, l_i)}{\bar{c}_i} \right)^{-\gamma} \left( \frac{e(c_i, l_i)}{c_i} \right)^\rho.$$

These marginal social welfare weights differ from those obtained in Section 2. Here, the additional term  $\partial e(c_i, l_i) / \partial c_i$  is here different than 1 when  $\ell_i \neq \bar{\ell}_i$  due to income effects, which were excluded in Section 2. When individuals distort their labor supply downwards ( $l_i < \bar{l}_i$ ), the income effects amplify the well-being change of an increase in consumption (the indifference curves are “closer” to each other), which society accepts by *efficiency*. We set the same utility parameters as in Mankiw et al. (2009):  $\rho = 1.5$ ,  $\alpha = 2.55$ , and  $\sigma = 3$ .<sup>18</sup> We also use the same income distribution parameters (for the US in 2007). However, we deviate from their study by including an exogenous revenue requirement,  $R$ , equal to 30% of total laissez-faire income.<sup>19</sup>

Our exercise consists in comparing the *log utilitarian social welfare function*

$$W^U = \int_w^{w^t} \ln [u(c_\theta, \ell_\theta)] f(\theta) d\theta$$

with the *log proportional-sacrifice social welfare function* (emerging when  $\gamma = 1$ )

---

<sup>18</sup>We thank Gregory Mankiw, Matthew Weinzierl, and Danny Yagan for making their data and code available.

<sup>19</sup>As a reminder, in the absence of a revenue requirement, the equal-sacrifice social welfare function would optimally select the laissez-faire outcome.

$$W^S = \int_w^{w^t} \bar{c}_i \ln [e(c_\theta, \ell_\theta)] f(\theta) d\theta.$$

A few preliminary remarks are in order. First, the goal of this exercise is simply to illustrate the applicability of our results and roughly argue that the policy implications of our proposal are reasonable. This ensures that one cannot reject our axiom and equal-sacrifice criterion by Rawls' reflective equilibrium argument. Second, the concern for the different types of inequality is set to logarithmic. The logarithm is recognized as a middle ground level of concavity, between linear (no concern for inequality) and infinite (full priority to the worst-off). We leave to future research the analysis of the inverse optimal taxation problem, whereby the ethical parameters of the welfare criteria are set to match the observed tax system (see Bourguignon and Spadaro, 2012).

The results are summarized by the graphs in figure 4, representing the marginal tax rate and the average tax rate emerging for the two criteria. The optimal tax system derived with the proportional-sacrifice social welfare function is less redistributive than the one derived with the utilitarian criterion. Crucially, this does not imply that the equal-sacrifice criterion is insensitive to the well-being of low-income individuals. Rather, the proportional-sacrifice criterion is characterized by a lower willingness of society to transfer to the worst-off individuals.

This lower concern for redistribution appears to be consistent with real-world income taxes. In particular, the utilitarian second-best policy supports marginal tax rates above 60% for all individuals (and up to 80% for high-income individuals). In contrast, the proportional-sacrifice criterion supports marginal tax rates that are about 20 percentage points lower.<sup>20</sup> These marginal tax rates are similar to those observed in the US tax system. Here, we plot the 2007 marginal tax rates for the combined federal and California state taxes on the income of singles (since California sets the highest state income taxes, it provides the strongest case for the tax schedule implied by utilitarianism compared to the one implied by proportional sacrifice). Note that the difference of marginal tax rates for low incomes can be explained by the absence of an extensive labor supply margin (see the discussion in Diamond and Saez, 2011). When the extensive margin is accounted for, optimal marginal tax rates at the lowest income levels are

---

<sup>20</sup>Note that the optimal tax schedule for the utilitarian criterion is not very sensitive to inequality aversion in utilities. At the extreme of no inequality aversion, the utilitarian social welfare function is the sum of (untransformed) utilities. The corresponding optimal tax schedule has marginal tax rates that are about 5 percentage points lower than with the log-utilitarian criterion.

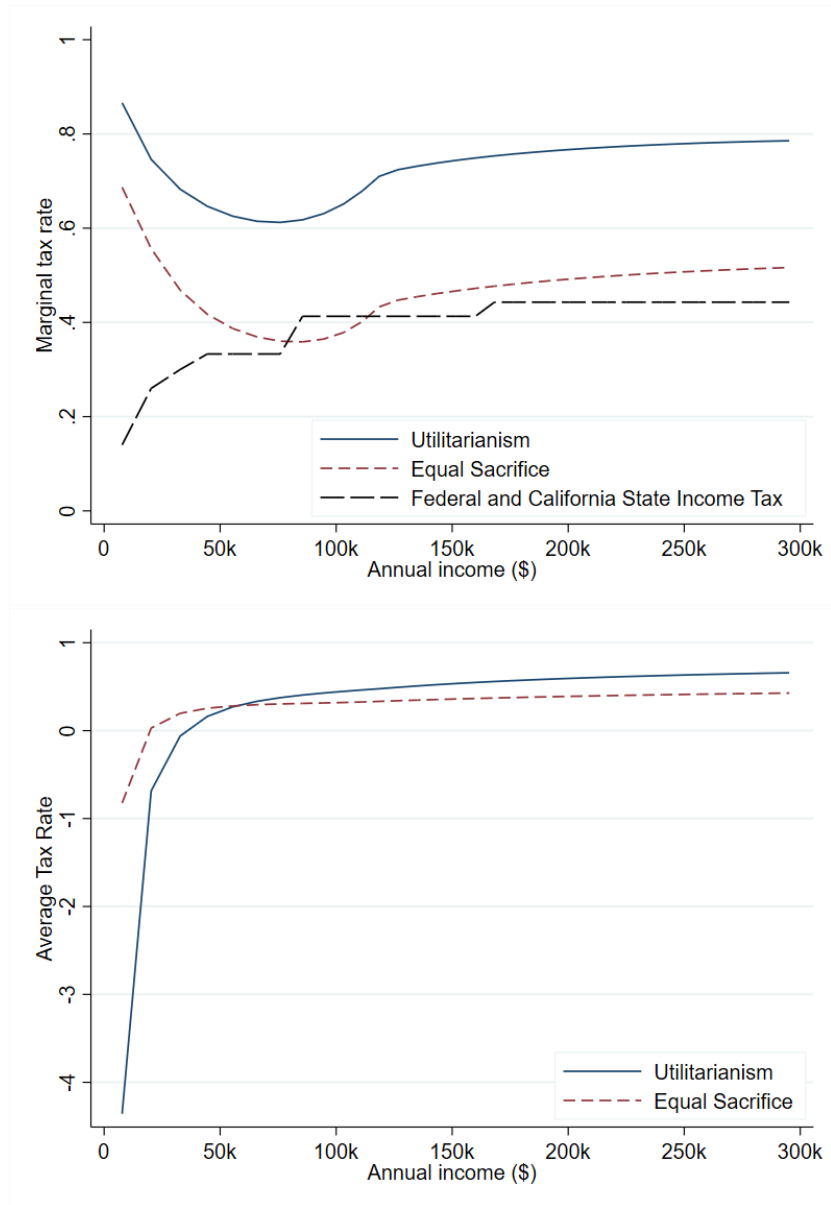


Figure 4: A comparison between the utilitarian and equal-sacrifice criteria: the marginal and average tax rates.

significantly lower (Saez, 2002; Jacquet, Lehman, and van der Linden, 2013).<sup>21</sup>

The average tax rates are also informative. Utilitarianism suggests subsidies (negative average taxes) ought to be distributed to the bottom 35% of the population, while the proportional-sacrifice criterion does so for only the bottom 15% of the population. In fact, utilitarianism implies a lump-sum transfer more than 5 times larger than that of the proportional-sacrifice criterion. Contrary to what one might expect, the tax system implied by our criterion redistributes on net to the lowest income levels. This is due to the presence of income effects. The incentive compatibility constraint forces downward labor supply distortions for the low-skilled workers. At such low labor supply, two contrasting forces define their second-best after-tax income. On the one hand, the negative sacrifice (due to the large lump-sum transfer) suggests society ought to increase their taxes for the benefit of higher income individuals. On the other hand, the income effects amplify the welfare effect of changes in their after-tax income. Thus, while the equity motive suggests an additional dollar be given to higher-skilled individuals, the efficiency motive dominates for the lowest income earners. This result disappears when society is infinitely averse to inequality in sacrifice ( $\gamma \rightarrow \infty$ ), at which point the criterion requires minimizing the largest sacrifice.

## 6 Conclusion

The optimal choice of income taxation is a key question in public economics. The answer requires combining a positive model of the economy—capturing behavioral choices of individuals—with normative aspects—reflecting ethical principles about how to compare benefits and losses of individuals. However, since the seminal contribution of Mirrlees (1971), the literature has mostly advanced by considering richer models of the economy, while the normative criterion was generally set to be utilitarian.

The utilitarian criterion is subject to a number of criticisms. Among those, Edgeworth (1897) highlights the utilitarian criterion imposes too strong a motivation to redistribute: with inelastic earnings, the optimal taxation policy is to tax income at 100% and redistribute the tax revenues equally across individuals. A partial solution has been to introduce different Pareto weights for different individuals. However, we share Piketty and Saez’s view (2013) that “the Pareto weight approach is too general

---

<sup>21</sup>We expect the introduction of an extensive margin to bring the equal-sacrifice optimal tax rates very close to the US tax system also for low incomes. We leave this exercise to future research.

to deliver practical policy prescriptions in most cases” (p. 393).

In this paper, we revisit an old and well-known theory of fairness in taxation, namely equal sacrifice. As a standard of perfection, equal sacrifice was unable to provide reasonable policy guidance: in second best settings, equalizing sacrifice leads to inefficient tax schedules. In contrast, our results establish it is possible to construct a social welfare function combining fairness considerations based on the equal-sacrifice principle with a concern for efficiency. The main result of the paper is the axiomatic characterization of a family of equal-sacrifice social welfare functions, which prioritizes individuals making a larger sacrifice. Our criterion redeems utilitarianism’s counterintuitive instances of a redistributive motive and thus can have large impacts on optimal tax policy.

To speak to those impacts, we show that second-best optimal tax policy differs most from the utilitarian one when the government’s budget requirement is not too large and when the labor-supply elasticity is small. Then, we demonstrate how to apply the criterion in a continuous Mirrlees model. In a numerical simulation for the US economy, we show that the proportional equal-sacrifice social welfare function implies lower tax rates and less redistribution than utilitarianism does, and is more in line with the observed US tax schedule.

We conclude with a remark about our methodology. We believe our axiomatic approach innovatively shows how to bridge the gap between approaches to first-best distribution of resources—as addressed in the theory of fair allocations—and fine-grained welfare criteria that trade off equity and efficiency considerations and, thus, can accommodate second best policy analysis. Such a methodology looks particularly promising for the study of optimal second-best policies where differences in marginal utilities are not enough information to accommodate widespread views on distributive justice. Examples include other (more general) optimal taxation settings, but also the provision of public goods, the allocation of health services, matching problems, and so on.

## References

- BERLIANT, M. AND M. GOUVEIA (1993): “Equal sacrifice and incentive compatible income taxation,” *The Journal of Public Economics*, 51, 219–240.
- BINMORE, K., A. RUBINSTEIN, AND A. WOLINSKY (1986): “The Nash bargaining solution in economic modelling,” *The RAND Journal of Economics*, 17, 176–188.

- BOURGUIGNON, F. AND A. SPADARO (2012): “Tax–benefit revealed social preferences,” *The Journal of Economic Inequality*, 10, 75–108.
- CHAMBERS, C. P. AND J. D. MORENO-TERNERO (2017): “Taxation and poverty,” *Social Choice and Welfare*, 48, 153–175.
- DA COSTA, C. E. AND T. PEREIRA (2014): “On the efficiency of equal sacrifice income tax schedules,” *The European Economic Review*, 70, 399–418.
- D’ASPREMONT, C. AND L. GEVERS (1977): “Equity and the informational basis of collective choice,” *The Review of Economic Studies*, 44, 199–209.
- DIAMOND, P. AND E. SAEZ (2011): “The case for a progressive tax: from basic research to policy recommendations,” *Journal of Economic Perspectives*, 25, 165–90.
- EDGEWORTH, F. Y. (1897): “The pure theory of taxation,” *The Economic Journal*, 7, 46–70.
- FELDSTEIN, M. (1976): “On the theory of tax reform,” *The Journal of Public Economics*, 6, 77–104.
- FLEURBAEY, M. AND F. MANIQUET (2006): “Fair income tax,” *The Review of Economic Studies*, 73, 55–83.
- (2018): “Optimal income taxation theory and principles of fairness,” *The Journal of Economic Literature*, 56, 1029–79.
- GAUTHIER, D. (1986): *Morals by agreement*, Oxford: Oxford University Press.
- GORMAN, W. M. (1968): “The structure of utility functions,” *The Review of Economic Studies*, 35, 367–390.
- JACQUET, L., E. LEHMANN, AND B. VAN DER LINDEN (2013): “Optimal redistributive taxation with both extensive and intensive responses,” *The Journal of Economic Theory*, 148, 1770–1805.
- JACQUET, L. AND D. VAN DE GAER (2011): “A comparison of optimal tax policies when compensation or responsibility matter,” *The Journal of Public Economics*, 95, 1248–1262.
- JESSEN, R., M. METZING, AND D. ROSTAM-AFSCHAR (2019): “Optimal taxation under different concepts of justness,” *Working Paper*.
- MANKIW, N. G., M. WEINZIERL, AND D. YAGAN (2009): “Optimal taxation in theory and practice,” *The Journal of Economic Perspectives*, 23, 147–74.



- MASKIN, E. (1978): “A theorem on utilitarianism,” *The Review of Economic Studies*, 45, 93–96.
- MILL, J. S. (1848): *Principles of political economy*, London: John W. Parker.
- MIRRLEES, J. (1971): “An exploration in the theory of optimum income taxation,” *The Review of Economic Studies*, 38, 175–208.
- MITRA, T. AND E. A. OK (1996): “Personal income taxation and the principle of equal sacrifice revisited,” *The International Economic Review*, 37, 925–948.
- MUSGRAVE, R. (1959): *Theory of public finance: A study in public economy*, New York: McGraw-Hill.
- NEGISHI, T. (1960): “Welfare economics and existence of an equilibrium for a competitive economy,” *Metroeconomica*, 12, 92–97.
- OK, E. A. (1995): “On the principle of equal sacrifice in income taxation,” *The Journal of Public Economics*, 58, 453–467.
- PIACQUADIO, P. G. (2017): “A fairness justification of utilitarianism,” *Econometrica*, 85, 1261–1276.
- (2020): “The ethics of intergenerational risk,” *Journal of Economic Theory*, 186, 104999.
- PIGOU, A. C. (1928): *A study in public finance*, Read Books Ltd.
- PIKETTY, T. AND E. SAEZ (2013): “Optimal labor income taxation,” in *Handbook of public economics*, Elsevier, vol. 5, 391–474.
- RUBINSTEIN, A., Z. SAFRA, AND W. THOMSON (1992): “On the interpretation of the Nash bargaining solution and its extension to non-expected utility preferences,” *Econometrica*, 60, 1171–1186.
- SAEZ, E. (2001): “Using elasticities to derive optimal income tax rates,” *The Review of Economic Studies*, 68, 205–229.
- (2002): “Optimal income transfer programs: intensive versus extensive labor supply responses,” *The Quarterly Journal of Economics*, 117, 1039–1073.
- SAEZ, E. AND S. STANTCHEVA (2016): “Generalized social marginal welfare weights for optimal tax theory,” *The American Economic Review*, 106, 24–45.

- SCHOKKAERT, E. AND K. DEVOOGHT (2003): “Responsibility-sensitive fair compensation in different cultures,” *Social Choice and Welfare*, 21, 207–242.
- SIDGWICK, H. (1883): *The principles of political economy*, London: Macmillan.
- SIMMONS, A. J. (2010): “Ideal and nonideal theory,” *Philosophy & public affairs*, 38, 5–36.
- STIGLITZ, J. E. (1982): “Self-selection and Pareto efficient taxation,” *The Journal of Public Economics*, 17, 213–240.
- STOVAL, J. E. (2020): “Equal sacrifice taxation,” *Games and Economic Behavior*, 121, 55–75.
- THOMSON, W. (2019): *How to divide when there isn’t enough: From Aristotle, the Talmud, and Maimonides to the axiomatics of resource allocation*, vol. 62, Cambridge University Press.
- VICKREY, W. (1947): *Agenda for progressive taxation*, New York: Ronald Press.
- WEINZIERL, M. (2012): “Why do we redistribute so much but tag so little? The principle of equal sacrifice and optimal taxation,” *Working Paper*.
- (2014): “The promise of positive optimal taxation: Normative diversity and a role for equal sacrifice,” *The Journal of Public Economics*, 118, 128–142.
- YOUNG, H. (1988): “Distributive justice in taxation,” *The Journal of Economic Theory*, 44, 321–335.
- YOUNG, H. P. (1990): “Progressive taxation and equal sacrifice,” *The American Economic Review*, 80, 253–266.

## A Proof of Lemma 1

*Inequality aversion* implies that any subset of individuals is “strictly essential:” for each  $I' \subseteq I$  and each  $\{a_i^*\}_{i \in I \setminus I'}$ , allocations  $a, a' \in A$  with  $a_i = a'_i = a_i^*$  for each  $i \in I \setminus I'$  are not all indifferent. By continuity of the social preferences, *separability*, and *strict essentiality*, Theorem 1 and Theorem 2 in Gorman (1968) apply and prove the existence of a representation  $W(a) = \sum_{i \in I} H_i(c_i, \ell_i)$ , where  $H_i$  is continuous for each  $i \in I$ . By *efficiency*,  $H_i(c_i, \ell_i)$  is an order preserving transformation of  $u(c_i, \ell_i)$ . Thus, there exist a continuous function  $P_i$  such that  $H_i(c_i, \ell_i) = P_i(u(c_i, \ell_i))$ . Substituting gives the result.

## B Proof of Theorem 1

**Part 1.** We first show the equal-sacrifice social welfare function satisfies the axioms. Let the sacrifice function be  $S \in \mathcal{S}$  and let the individual Pareto functions  $(P_i)_{i \in I}$  be consistent with  $S$ . Then, the equal-sacrifice social welfare function is

$$W(a) \equiv \sum_{i \in I} P_i(u(c_i, \ell_i)).$$

Efficiency. Since social marginal welfare weights are positive, the Pareto functions are increasing. Then, for each  $i \in I$  and each pair  $(c_i, \ell_i), (c'_i, \ell'_i)$ ,  $u(c_i, \ell_i) \geq u(c'_i, \ell'_i)$  if and only if  $P_i(u(c_i, \ell_i)) \geq P_i(u(c'_i, \ell'_i))$ . Consider a pair of allocations  $a, a' \in A$  such that  $u(c_i, \ell_i) \geq u(c'_i, \ell'_i)$  for each  $i \in I$  and  $u(c_i, \ell_i) > u(c'_i, \ell'_i)$  for some  $i \in I$ . Thus, also  $P_i(u(c_i, \ell_i)) \geq P_i(u(c'_i, \ell'_i))$  for each  $i \in I$  and  $P_i(u(c_i, \ell_i)) > P_i(u(c'_i, \ell'_i))$  for some  $i \in I$ . Then,  $\sum_{i \in I} P_i(u(c_i, \ell_i)) = W(a) > W(a') = \sum_{i \in I} P_i(u(c'_i, \ell'_i))$  and  $a \succ a'$ . This proves the equal-sacrifice social welfare function satisfies *efficiency*.

Inequality aversion. By construction, for each  $i \in I$ ,  $P_i(u(c_i, \ell_i))$  is strictly concave in its arguments. It follows that  $W(a)$  is strictly concave in its arguments and *inequality aversion* holds.

Separability. *Separability* follows from the additivity of the function  $W$ : the bundle of an unconcerned individual is irrelevant for the ranking of two allocations.

Tax solidarity. Consider a pair of allocations  $a, a' \in A$  satisfying the requirements in the definition of *tax solidarity*. These allocations are such that for some pair of individuals  $i, j \in I$  and some  $\varepsilon > 0$ :  $b'_i - \varepsilon = b_i \geq 0 \geq b_j = b'_j + \varepsilon$ ;  $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and  $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ . By definition,  $b_i \equiv \bar{c}_i - c_i$  and  $b'_i \equiv \bar{c}_i - c'_i$ . Thus,  $c'_i = \bar{c}_i - b'_i = \bar{c}_i - b_i - \varepsilon$  and, substituting for  $b_i$ ,  $c'_i = c_i - \varepsilon$ . Similarly,  $c'_j = c_j + \varepsilon$ . Substituting and since individuals  $k \in I / \{i, j\}$  are unaffected, we may write

$$W(a) - W(a') = P_i(u(c_i, \bar{\ell}_i)) - P_i(u(c_i - \varepsilon, \bar{\ell}_i)) + P_j(u(c_j, \bar{\ell}_j)) - P_j(u(c_j + \varepsilon, \bar{\ell}_j)).$$

Now, by first-degree Taylor expansion and concavity of individuals' contributions to social welfare,

$$P_i(u(c_i - \varepsilon, \bar{\ell}_i)) \leq P_i(u(c_i, \bar{\ell}_i)) - \varepsilon \beta_i(c_i, \bar{\ell}_i),$$

where, for memory,  $\beta_i(c_i, \bar{\ell}_i) = \frac{\partial}{\partial c_i} P_i(u(c_i, \bar{\ell}_i))$ ; and

$$P_j(u(c_j + \varepsilon, \bar{\ell}_j)) \leq P_j(u(c_j, \bar{\ell}_j)) + \varepsilon \beta_j(c_j, \bar{\ell}_j),$$

where  $\beta_j(c_j, \bar{\ell}_j) = \frac{\partial}{\partial c_j} P_j(u(c_j, \bar{\ell}_j))$ .

Thus,

$$W(a) - W(a') \geq \varepsilon [\beta_i(c_i, \bar{\ell}_i) - \beta_j(c_j, \bar{\ell}_j)].$$

Finally, since  $S(c_i; \bar{c}_i) > 0 > S(c_j; \bar{c}_j)$  and since  $g$  is increasing,  $\beta_i(c_i, \bar{\ell}_i) = g(S(c_i; \bar{c}_i)) > g(S(c_j; \bar{c}_j)) = \beta_j(c_j, \bar{\ell}_j)$ . Thus,  $W(a) \geq W(a')$  and  $a \succsim a'$ . This proves *tax solidarity* holds.

Fair burden. Consider a pair of allocations  $a, a' \in A$  satisfying the requirements in the definition of *fair burden*. These allocations are such that for some pair of individuals  $i, j \in I$  with  $\bar{c}_i \geq \bar{c}_j$  and some  $\varepsilon > 0$ :  $0 \leq b'_i + \varepsilon = b_i < b_j = b'_j - \varepsilon$ ;  $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and  $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ . Substituting for  $b_i, b'_i, b_j$ , and  $b'_j$ ,  $c'_i = c_i + \varepsilon$  and  $c'_j = c_j - \varepsilon$ . Substituting and since individuals  $k \in I / \{i, j\}$  are unaffected, we may write

$$W(a) - W(a') = P_i(u(c_i, \bar{\ell}_i)) - P_i(u(c_i + \varepsilon, \bar{\ell}_i)) + P_j(u(c_j, \bar{\ell}_j)) - P_j(u(c_j - \varepsilon, \bar{\ell}_j)).$$

Now, by first-degree Taylor expansion and concavity of individuals' contributions to social welfare,

$$P_i(u(c_i + \varepsilon, \bar{\ell}_i)) \leq P_i(u(c_i, \bar{\ell}_i)) + \varepsilon \beta_i(c_i, \bar{\ell}_i),$$

and

$$P_j(u(c_j - \varepsilon, \bar{\ell}_j)) \leq P_j(u(c_j, \bar{\ell}_j)) - \varepsilon \beta_j(c_j, \bar{\ell}_j).$$

Thus,

$$W(a) - W(a') \geq \varepsilon [\beta_j(c_j, \bar{\ell}_j) - \beta_i(c_i, \bar{\ell}_i)].$$

Finally, since  $c_i \leq \bar{c}_i$ ,  $S(c_i; \bar{c}_i) \geq 0$ . Furthermore,  $c_i - c_j \geq \bar{c}_i - \bar{c}_j$ . Thus, by the slope restriction on  $S$ ,  $S(c_i; \bar{c}_i) \leq S(c_j; \bar{c}_j)$ . Then, since  $g$  is increasing,  $\beta_i(c_i, \bar{\ell}_i) = g(S(c_i; \bar{c}_i)) \leq g(S(c_j; \bar{c}_j)) = \beta_j(c_j, \bar{\ell}_j)$ . Thus,  $W(a) \geq W(a')$  and  $a \succsim a'$ . This proves *fair burden* holds.

Fair reward. Consider a pair of allocations  $a, a' \in A$  satisfying the requirements in the definition of *fair reward*. These allocations are such that for some pair of individuals

$i, j \in I$  with  $\bar{c}_i \geq \bar{c}_j$  and some  $\varepsilon > 0$ :  $c'_i + \varepsilon = c_i < c_j = c'_j - \varepsilon$ ;  $\ell_i = \ell'_i = \bar{\ell}_i$  and  $\ell_j = \ell'_j = \bar{\ell}_j$ ; and  $(c_k, \ell_k) = (c'_k, \ell'_k) = (\bar{c}_k, \bar{\ell}_k)$  for each  $k \in I / \{i, j\}$ . Since individuals  $k \in I / \{i, j\}$  are unaffected, we may write

$$W(a) - W(a') = P_i(u(c_i, \bar{\ell}_i)) - P_i(u(c_i - \varepsilon, \bar{\ell}_i)) + P_j(u(c_j, \bar{\ell}_j)) - P_j(u(c_j + \varepsilon, \bar{\ell}_j)).$$

Now, by first-degree Taylor expansion and concavity of individuals' contributions to social welfare,

$$P_i(u(c_i - \varepsilon, \bar{\ell}_i)) \leq P_i(u(c_i, \bar{\ell}_i)) - \varepsilon \beta_i(c_i, \bar{\ell}_i),$$

and

$$P_j(u(c_j + \varepsilon, \bar{\ell}_j)) \leq P_j(u(c_j, \bar{\ell}_j)) + \varepsilon \beta_j(c_j, \bar{\ell}_j).$$

Thus,

$$W(a) - W(a') \geq \varepsilon [\beta_i(c_i, \bar{\ell}_i) - \beta_j(c_j, \bar{\ell}_j)].$$

Finally,  $S$  is decreasing in the first argument and increasing in the second:  $c_i < c_j$  and  $\bar{c}_i \geq \bar{c}_j$  imply that  $S(c_i; \bar{c}_i) > S(c_j; \bar{c}_j)$ . Thus, since  $g$  is increasing,  $\beta_i(c_i, \bar{\ell}_i) = g(S(c_i; \bar{c}_i)) > g(S(c_j; \bar{c}_j)) = \beta_j(c_j, \bar{\ell}_j)$ . Thus,  $W(a) \geq W(a')$  and  $a \succsim a'$ . This proves *fair reward* holds.

**Part 2.** We now show social preferences satisfying the axioms admit a representation by means of an equal-sacrifice social welfare function.

The proof is divided in several steps.

Step 1. *Assume social preferences  $\succsim$  satisfy the axioms. Then, there exists real-valued increasing and strictly concave functions  $(h_i)_{i \in I}$  such that social welfare  $W$  representing  $\succsim$  is defined by setting for each  $a \in A$ ,*

$$W(a) = \sum_{i \in I} h_i(e_i(c_i, \ell_i)).$$

*Proof.* Lemma 1 establishes social preferences  $\succsim$  may be represented by a social welfare function of the sum-of-utilities type. Formally, there exist real-valued increasing functions  $(f_i)_{i \in I}$  such that  $f_i(u(c_i, \ell_i))$  is strictly concave for each  $i \in I$  and such that, for each pair  $a, a' \in A$ ,  $a \succsim a'$  if and only if

$$W(a) = \sum_{i \in I} f_i(u(c_i, \ell_i)) \geq \sum_{i \in I} f_i(u(c'_i, \ell'_i)) = W(a').$$

Next, for each  $i \in I$ ,  $e_i(c_i, \ell_i)$  is the consumption-equivalent representation of preferences of  $i$ . Thus, there exists a real-valued increasing function  $h_i$  such that  $h_i(e_i(c_i, \ell_i)) = f_i(u(c_i, \ell_i))$  for each  $(c_i, \ell_i)$ . This shows social preferences may be represented by the social welfare function  $W$ , as defined above.

It remains to show the functions  $(h_i)_{i \in I}$  are strictly concave. By definition of the consumption-equivalent representation of preferences, for each  $i \in I$  and each  $c_i \in \mathbb{R}_+$ ,  $h_i(c_i) = f_i(u(c_i, \bar{\ell}_i))$ . Since  $f_i(u(c_i, \ell_i))$  is strictly concave, also  $h_i$  is strictly concave.  $\square$

Next, for each  $c_i \in \mathbb{R}_+$ , denote  $h'_i(c_i^-)$  and  $h'_i(c_i^+)$  the left and right first-order derivatives, respectively, of  $h_i$  at  $c_i$ . Let  $\bar{A}$  be the set of allocations  $a \in A$  such that  $\ell_i = \bar{\ell}_i$  for each  $i \in I$ . Then, for each  $a \in \bar{A}$ ,  $W(a) = \sum_{i \in I} h_i(c_i)$ . Let the choice correspondence  $C$  be defined as follows: for each  $k \geq 0$ ,  $C(k)$  is the set of consumption vectors  $(c_i)_{i \in I}$  with  $\sum_{i \in I} c_i \leq k$  that maximize  $W$ . Let  $\bar{k} \equiv \sum_{i \in I} \bar{c}_i$ . The following steps characterize the properties of  $C$  (with a slight abuse of notation, we shall use  $C$  also to denote the choice function, after showing the correspondence  $C$  is single-valued).

Step 2. *The choice correspondence  $C$  satisfies the following properties:*

1. *it is non-empty, single-valued, and continuous with respect to  $k$ ;*
2. *it is strictly monotonic,  $k > k'$  implies  $C(k) \gg C(k')$ ;*
3.  $C(\bar{k}) = (\bar{c}_i)_{i \in I}$ ;
4.  $(c_i)_{i \in I} = C(k)$  implies  $c_i > c_j \iff \bar{c}_i > \bar{c}_j$  for each  $i, j \in I$ ;
5. *for  $k \leq \bar{k}$ ,  $(c_i)_{i \in I} = C(k)$  implies  $c_i - c_j < \bar{c}_i - \bar{c}_j$  for each  $i, j \in I$ .*

*Proof.* 1.  $W$  is increasing, continuous, and strictly concave, and so is  $\sum_{i \in I} h_i(c_i)$ . Thus, the choice correspondence  $C$  is non-empty, single-valued, and continuous with respect to  $k$ .

2. Let  $(c_i)_{i \in I} = C(k)$  and  $(c'_i)_{i \in I} = C(k')$ . By contradiction of strict monotonicity, assume  $k > k'$  and  $C(k) \not\gg C(k')$ . Then, there exists a pair of individuals  $i, j \in I$  such that  $c'_i \leq c_i$  and  $c'_j > c_j$ . At the optima,  $h'_i(c_i^-) \geq h'_j(c_j^+)$  and  $h'_i(c_i^+) \leq h'_j(c_j^-)$  and, similarly,  $h'_i(c_i^-) \geq h'_j(c_j^+)$  and  $h'_i(c_i^+) \leq h'_j(c_j^-)$ . By strict concavity,  $h'_i(c_i^-) \geq h'_i(c_i^+) \geq h'_i(c_i^-) \geq h'_i(c_i^+)$  and  $h'_j(c_j^-) \geq h'_j(c_j^+) > h'_j(c_j^-) \geq h'_j(c_j^+)$ . Combining these conditions leads to the following contradiction:

$$h'_i(c_i^-) \geq h'_j(c_j^+) > h'_j(c_j^-) \geq h'_i(c_i^+) \geq h'_i(c_i^-).$$

3. By contradiction of  $C(\bar{k}) = (\bar{c}_i)_{i \in I}$ , assume  $(\bar{c}_i)_{i \in I} \neq C(\bar{k}) \equiv (c_i)_{i \in I}$ . Then,  $\sum_{i \in I} h_i(c_i) > \sum_{i \in I} h_i(\bar{c}_i)$ . At  $(c_i)_{i \in I}$ , the tax burden of each individual  $i \in I$  is  $b_i \equiv \bar{c}_i - c_i$ . Since  $\bar{k} = \sum_i \bar{c}_i$ ,  $\sum_{i \in I} b_i = 0$ . Let  $\vec{b}$  be the reordered vector of tax burdens of individuals:  $\vec{b} \equiv (b_{(1)}, \dots, b_{(|I|)})$  is such that  $b_{(1)} \leq b_{(2)} \leq \dots \leq b_{(|I|)}$ , where  $(i)$  is the individual that, after permutation, occupies the  $i$ 'th place in the order of tax burdens. Since  $(\bar{c}_i)_{i \in I} \neq (c_i)_{i \in I}$ ,  $\vec{b} \neq 0$  and thus  $b_{(1)} < b_{(|I|)}$ .

We next apply *tax solidarity* a finite number of times to show that  $(\bar{c}_i)_{i \in I}$  is socially at least as desirable as  $(c_i)_{i \in I}$ , leading to a contradiction. The process is iterative and indexed by  $t$ . Let  $c^t \equiv (c_i^t)_{i \in I}$  and let  $\vec{b}^t$  be the corresponding reordered vector of tax burdens. Let  $c^1 \equiv (c_i)_{i \in I}$ . At each step  $t$ , three cases emerge.

*Case (i).*  $|b_{(1)}^t| < |b_{(|I|)}^t|$ . Let  $c^{t+1}$  be such that  $c_{(1)}^{t+1} = \bar{c}_{(1)}$ ,  $c_{(|I|)}^{t+1} = c_{(|I|)}^t + b_{(1)}^t$ , and  $c_i^{t+1} = c_i^t$  for each  $i$  such that  $(i) \neq (1), (|I|)$ . Since  $b_{(1)}^t < 0$ , this is a transfer from  $(|I|)$  to  $(1)$ .

*Case (ii).*  $|b_{(1)}^t| > |b_{(|I|)}^t|$ . Let  $c^{t+1}$  be such that  $c_{(1)}^{t+1} = c_{(1)}^t - b_{(|I|)}^t$ ,  $c_{(|I|)}^{t+1} = \bar{c}_{(|I|)}$ , and  $c_i^{t+1} = c_i^t$  for each  $i$  such that  $(i) \neq (1), (|I|)$ . Since  $b_{(|I|)}^t > 0$ , this is again a transfer from  $(|I|)$  to  $(1)$ .

*Case (iii).*  $|b_{(1)}^t| = |b_{(|I|)}^t| = 0$ . Let  $c^{t+1} = c^t$ .

The process converges in a finite number of iterations:  $c^{(|I|)} = (\bar{c}_i)_{i \in I}$ . For each  $t = 1, \dots, (|I|)$ , let the allocation  $a^t \in A$  assign to each individual the bundle  $(c_i^t, \bar{\ell}_i)$ . Then, for each iteration  $t$ , if  $a^t \neq a^{t-1}$ , cases (i) or (ii) apply and, by *tax solidarity*,  $a^t \succsim a^{t-1}$ . Otherwise,  $a^t = a^{t-1}$  and thus  $a^t \sim a^{t-1}$ . By the representation result in Step 1, this implies  $\sum_{i \in I} h_i(c_i^t) \geq \sum_{i \in I} h_i(c_i^{t-1})$ . Thus, also  $\sum_{i \in I} h_i(\bar{c}_i) \geq \sum_{i \in I} h_i(c_i)$ . This is a contradiction.

4. The proof is similar to that of 3, where *fair reward* is applied.

5. The proof is similar to that of 3, where *fair burden* is applied.  $\square$

We next construct the function  $S : \mathbb{R}_+ \times \mathbb{R}_{++} \rightarrow \mathbb{R}$  and then verify  $S$  is a sacrifice function.

First, for each  $i \in I$  and each  $k \geq 0$ , let  $S(c_i; \bar{c}_i) = \bar{k} - k$  if and only if  $(c_i)_{i \in I} = C(k)$ .

Second, we complete the sacrifice function *linearly* for non-observed levels of laissez-faire consumption (when  $y \neq \bar{c}_i$  for each  $i$ ). Reorder individuals in increasing order of laissez-faire consumption, that is,  $(i) \leq (j)$  if  $\bar{c}_i \leq \bar{c}_j$ . Let  $\bar{c}_0 = 0$ . Set  $y \in \mathbb{R}_{++}$ . Then, two cases emerge: either (1) there exists  $i \in I$  such that  $\bar{c}_{(i-1)} \leq y \leq \bar{c}_{(i)}$  or (2)  $y > \bar{c}_{(|I|)}$ .

Case 1. Let  $i \in I$  be such that  $\bar{c}_{(i-1)} \leq y \leq \bar{c}_{(i)}$  and let  $\alpha \in [0, 1]$  be such that  $y = \alpha \bar{c}_{(i-1)} + (1 - \alpha) \bar{c}_{(i)}$ . Then, for each  $x \in \mathbb{R}_+$ ,  $S(x; y) = \bar{k} - k$  if and only if

$x = \alpha c_{(i-1)} + (1 - \alpha) c_{(i)}$  where  $S(c_{(i-1)}; \bar{c}_{(i-1)}) = S(c_{(i)}; \bar{c}_{(i)}) = \bar{k} - k$ .

Case 2. Let  $(i)$  be such that  $\bar{c}_{(|I|)} - \bar{c}_{(i)}$  is positive and smallest for all  $i \in I$ .<sup>22</sup> Let  $\alpha > 1$  be such that  $(y - \bar{c}_{(i)}) = \alpha (\bar{c}_{(|I|)} - \bar{c}_{(i)})$ . Then, for each  $x \in \mathbb{R}_+$ ,  $S(x; y) = \bar{k} - k$  if and only if  $(x - c_{(i)}) = \alpha (c_{(|I|)} - c_{(i)})$  where  $S(c_{(i)}; \bar{c}_{(i)}) = S(c_{(|I|)}; \bar{c}_{(|I|)}) = \bar{k} - k$ .

Step 3. *The function  $S$  is a sacrifice function. That is,  $S$  satisfies the following conditions:*

1. a) *decreasing in the first argument, b) increasing in the second argument, and c) continuous;*
2.  *$x = y$  implies  $S(x; y) = 0$ ; and*
3.  *$S(x; y) = S(x'; y') > 0$  implies  $|x - x'| \leq |y - y'|$ .*

*Proof.* 1a) For each  $i$ , the function  $S(c_i; \bar{c}_i)$  is decreasing in  $c_i$  by strict monotonicity of  $C(k)$ : more precisely, let  $k < k'$ ; then,  $(c_i)_{i \in I} = C(k) \ll C(k') = (c'_i)_{i \in I}$ ,  $c_i < c'_i$ , and  $S(c_i; \bar{c}_i) = \bar{k} - k > S(c'_i; \bar{c}_i) = \bar{k} - k'$ . For each  $y \in \mathbb{R}_+$ ,  $S(x; y)$  is decreasing in  $x$  as it is constructed as a linear combination of functions  $(S(c_i; \bar{c}_i))_{i \in I}$  which are decreasing in the first variable.

1b) Property 4 of Step 3 states that:  $(c_i)_{i \in I} = C(k)$  implies  $c_i > c_j \iff \bar{c}_i > \bar{c}_j$  for each  $i, j \in I$ . By construction of  $S$ , this implies that  $S(x, y) = S(x', y')$  with  $y < y'$  if and only if  $x < x'$ . Since  $S$  is decreasing in the first argument,  $S(x, y) < S(x, y')$ .

1c) Since  $C(k)$  is continuous in  $k$ , for each  $i$ , the function  $S(c_i; \bar{c}_i)$  is continuous in  $c_i$ . Continuity of  $S$  then follows by construction.

2) By construction,  $S(\bar{c}_i, \bar{c}_i) = \bar{k} - \bar{k} = 0$  for each  $i \in I$ . Now, for each  $y \in \mathbb{R}_{++}$ , either there exists  $i \in I$  such that  $\bar{c}_{(i-1)} \leq y \leq \bar{c}_{(i)}$  or  $y > \bar{c}_{(|I|)}$ . In the first case,  $S(y, y) = 0$  since, by definition of  $S$ ,  $S(\bar{c}_{(i-1)}; \bar{c}_{(i-1)}) = S(\bar{c}_{(i)}; \bar{c}_{(i)}) = 0$  and  $y = \alpha \bar{c}_{(i-1)} + (1 - \alpha) \bar{c}_{(i)}$  for some  $\alpha \in [0, 1]$ . In the second case, let  $(i)$  be such that  $\bar{c}_{(|I|)} - \bar{c}_{(i)}$  is positive and smallest for all  $i \in I$ . Then,  $S(y; y) = 0$  since, by definition of  $S$ ,  $S(\bar{c}_{(i)}; \bar{c}_{(i)}) = S(\bar{c}_{(|I|)}; \bar{c}_{(|I|)}) = 0$  and  $(y - \bar{c}_{(i)}) = \alpha (\bar{c}_{(|I|)} - \bar{c}_{(i)})$  for some  $\alpha > 1$ .

3) By contradiction, let  $k \equiv S(x; y) = S(x'; y') > 0$  and  $|x - x'| > |y - y'|$ . Without loss of generality, let  $x > x'$  and  $y > y'$ . By construction, the implicit function  $S(x, y) = k$  is piecewise linear: it may change slope only in correspondence to  $y = \bar{c}_i$  with  $(i) = 2, \dots, (|I|) - 1$ . By the mean value theorem,  $x - x' > y - y'$  implies there exists a pair

---

<sup>22</sup>When  $\bar{c}_{(|I|)} = \bar{c}_{(|I|-1)}$ , such individual  $(i)$  differs from  $(|I| - 1)$ .



$i, j \in I$  such that  $c_i - c_j > \bar{c}_i - \bar{c}_j$  with  $S(c_i; \bar{c}_i) = S(c_j; \bar{c}_j) = k$ . Clearly,  $c_i$  and  $c_j$  belong to  $(c_m)_{m \in I} = C(k)$ . Thus,  $c_i - c_j > \bar{c}_i - \bar{c}_j$  is a violation of *fair reward* (as shown above).  $\square$

The proof is completed by the following step, which shows that the Pareto functions  $P_i = f_i$  are consistent with the sacrifice function  $S$ .

Step 4. *For each  $i \in I$ , let the Pareto function of  $i$  be  $P_i$  such that  $P_i(u(c_i, \ell_i)) \equiv h_i(e_i(c_i, \ell_i))$ . The Pareto functions  $(P_i)_{i \in I}$  are consistent with  $S$ . That is, for each  $i \in I$ , the social marginal welfare weights  $\beta_i$  satisfy  $\beta_i(c_i, \bar{\ell}_i) = g(S(c_i; \bar{c}_i)) > 0$ , where  $g$  is a real-valued increasing function, equal across individuals, and such that individuals' contributions to social welfare  $(P_i(u(\cdot, \cdot)))_{i \in I}$  are strictly concave in their arguments.*

*Proof.* Strict concavity of  $P_i(u(\cdot, \cdot))$  immediately follows from Step 1. Strict concavity and efficiency also imply  $g$  is increasing and social marginal welfare weights are positive.

Finally, we show  $g$  is equal across individuals. By contradiction, assume not. Then, there exists a pair  $i, j \in I$  and  $c_i, c_j \in \mathbb{R}_+$  with  $S(c_i; \bar{c}_i) = S(c_j; \bar{c}_j)$  such that  $\beta_i(c_i, \bar{\ell}_i) \neq \beta_j(c_j, \bar{\ell}_j)$ . By construction,  $c_i$  and  $c_j$  belong to  $(c_m)_{m \in I} = C(k)$  for some  $k \geq 0$ . However,  $\beta_i(c_i, \bar{\ell}_i) \neq \beta_j(c_j, \bar{\ell}_j)$  imply  $\frac{\partial h_i(c_i, \bar{\ell}_i)}{\partial c_i} \neq \frac{\partial h_j(c_j, \bar{\ell}_j)}{\partial c_j}$ . This contradicts that  $C(k)$  maximizes social welfare  $W$  (from Step 1) among the vectors  $(c'_m)_{m \in I}$  such that  $\sum_{m \in I} c'_m \leq k$ .  $\square$