Procedural fairness and redistributive proportional tax

Citation for published version (APA):

Herings, P. J. J., & Predtetchinski, A. (2015). Procedural fairness and redistributive proportional tax. *Economic Theory*, *59*(2), 333-354. https://doi.org/10.1007/s00199-014-0852-9

Document status and date:
Published: 01/01/2015

DOI:
10.1007/s00199-014-0852-9

Document Version:
Publisher's PDF, also known as Version of record

Document license:
Taverne

Econ Theory (2015) 59:333–354
DOI 10.1007/s00199-014-0852-9

RESEARCH ARTICLE

Procedural fairness and redistributive proportional tax

P. Jean-Jacques Herings · Arkadi Predtetchinski

Received: 12 February 2014 / Accepted: 19 November 2014 / Published online: 30 November 2014

Abstract We study the implications of procedural fairness on the share of income that should be redistributed. We formulate procedural fairness as a particular non-cooperative bargaining game and examine the stationary subgame perfect equilibria of the game. The equilibrium outcome is called tax equilibrium and is shown to be unique. The procedurally fair tax rate is defined as the tax rate that results in the limit of tax equilibria when the probability that negotiations break down converges to zero. The procedurally fair tax rate is shown to be unique. We also provide a characterization of the procedurally fair tax rate that involves the probability mass of below average income citizens and a particular measure of the citizens' boldness. This characterization is then used to show that in a number of interesting cases, the procedurally fair tax rate equals the probability mass of below average income citizens.

Keywords Procedural fairness · Tax rate · Bargaining

JEL Classification C78 · D63 · H24

The authors would like to thank the Netherlands Organization for Scientific Research (NWO) for financial support.

P. J.-J. Herings (✉) · A. Predtetchinski
Department of Economics, Maastricht University, P.O. Box 616,
6200 MD Maastricht, The Netherlands
e-mail: P.Herings@maastrichtuniversity.nl

A. Predtetchinski
e-mail: A.Predtetchinski@maastrichtuniversity.nl

1 Introduction

Modern societies redistribute significant shares of their total income to improve upon the fairness of the income distribution. To determine what fraction of total income should be redistributed to obtain a fair income distribution, the standard approach is to follow Mirrlees (1971) and consists of specifying a social welfare function that has to be optimized.

One may distinguish fair procedures from fair allocations to decide upon the allocation. Bolton et al. (2005) argue that procedural fairness is conceptually distinct from allocation fairness, although the two are linked in important ways. This paper is concerned with the application of procedural fairness to the redistribution of income. Procedures are deemed fair if they create equal chances for persons involved in the procedure. It is equal opportunities that matters rather than an equal allocation that results.

The fair procedure that we want to examine is the unanimous approval procedure. Under the unanimous approval procedure, a proposal is only carried out if it is approved by everyone. The unanimous approval procedure follows as a special case of the bargaining models of collective choice considered in Banks and Duggan (2000), when we restrict recognition probabilities to be uniform and the collection of decisive coalitions to consist only of the grand coalition. Contrary to the approach in Banks and Duggan (2000) or the case with an evolving status quo in Battaglini and Palfrey (2012), we intend to use bargaining models of collective choice as normative tools here. The unanimous approval procedure is motivated by the political theory of the legitimacy of political authority called contractarianism, which claims that the legitimate authority of government must derive from the consent of the governed.

We apply the unanimous approval procedure in order to study what fraction of income should be redistributed on the grounds of procedural fairness. We consider a society with citizens that are characterized by their pre-tax income and their utility function, where the distribution of characteristics is given by some probability measure. A proposal of a citizen therefore specifies a tax rate between zero and one, and a particular proposal is only implemented if all citizens approve of it. A tax rate of zero corresponds to the situation of laissez-faire and a tax rate of one to complete redistribution.

Citizens with below average pre-tax income would prefer complete redistribution, whereas citizens with above average pre-tax income have a tax rate of zero as their most preferred tax policy. Citizens with average pre-tax income are indifferent as far as the tax policy is concerned. Since tax rates are chosen in the unit interval, our analysis leads to the analysis of one-dimensional bargaining problems. In the bargaining and political economy literature, such problems are also studied in Banks and Duggan (2000), Imai and Salonen (2000), Cho and Duggan (2003), Kalandrakis (2006), Cardona and Ponsatí (2007, 2011), Herings and Predtetchinski (2010), Predtetchinski (2011).

So far, the literature has only studied societies with finitely many agents. In contrast, we allow for the case where there is a measure space of agents, a modeling assumption that is very helpful for the analysis of taxation in large populations. The extension to a measure space of agents makes that we cannot rely on continuity properties in the proofs of the main results, and we therefore have to resort to different techniques as for instance demonstrated by the use of Tarski's fixed point theorem rather than Brouwer's when showing equilibrium existence.

We define a tax equilibrium as the outcomes induced by stationary subgame perfect equilibria of the unanimous approval procedure for a given breakdown probability. A tax equilibrium is characterized by the unique proposal made by all above average income citizens and the unique proposal made by all below average income citizens. It is shown that in equilibrium the proposals made are accepted unanimously without delay and that tax equilibria are unique. Below average income citizens propose a strictly higher tax rate than above average income citizens.

Next, we define the procedurally fair tax rate as the limit of the rates proposed in a tax equilibrium when the breakdown probability converges to zero. We show that every society has a unique procedurally fair tax rate.

The procedurally fair tax rate admits a characterization in terms of the probability mass of citizens with below average income and an appropriate measure of boldness of the citizens. As shown in Roth (1989), boldness at a certain consumption level corresponds to the maximum probability by which a citizen prefers a particular gamble over getting the consumption level for sure. Boldness equals the first derivative of the utility function divided by utility itself. Tax rate boldness is defined as boldness applied to the indirect utility function in terms of tax rates. We argue that the bargaining power of the below average income citizens at a particular tax rate is equal to the supremum of tax rate boldness among them multiplied by their probability mass. The bargaining power of the above average income citizens is defined similarly. We demonstrate that the procedurally fair tax rate is the unique tax rate for which the bargaining power of the below average income citizens is equal to that of the above average income citizens.

Our characterization of the procedurally fair tax rate is extremely helpful in computations. For instance, in societies where pre-tax incomes have unbounded support and all citizens have the same preferences exhibiting constant relative risk aversion, the procedurally fair tax rate is given by the probability mass of below average income citizens.

We also consider heterogeneous societies, where the only assumption on preferences is that, both among below average income citizens and among above average income citizens, there are citizens that are close to being risk neutral. We obtain a simple expression for the procedurally fair tax rate, irrespective of the distribution of pre-tax incomes. When this distribution has unbounded support, we find again that the procedurally fair tax rate is given by the probability mass of below average income citizens.

The remainder of this paper is organized as follows. Section 2 introduces the primitives of our society, and Section 3 discusses the unanimous approval procedure. Section 4 defines the notion of tax equilibrium, studies the relation to stationary subgame perfect equilibria of the unanimous approval procedure, and derives some properties of tax equilibria. Section 5 discusses the relationship between boldness and bargaining power and introduces an excess bargaining power function, which is shown to have a unique zero point. Section 6 demonstrates that this zero point corresponds to the procedurally fair tax rate in the society. Section 7 uses this result to show that in societies with citizens having constant relative risk aversion utility functions, the procedurally fair tax rate is equal to the probability mass of below average income citizens. The same result is shown for sufficiently heterogeneous societies Section 8 concludes.

2 The society

We consider a society composed of citizens in a set N, which may contain finitely or infinitely many elements. The characteristics of citizens are described by $(w_t, u_t)_{t \in N}$ and a probability measure μ. A citizen can work up to one unit of time and derives no disutility from labor. A citizen of type $t \in N$ who works one unit of time has pre-tax income $w_t \in \mathbb{R}_+$ and a utility function $u_t : [0, +\infty) \to \mathbb{R}$.

The triple $(N, \mathcal{A}, \mu)$ denotes a probability space, where N is the set of citizens, $\mathcal{A}$ a sigma–algebra of subsets of N, and μ a probability measure that represents the distribution of types within the population. If N is finite, then we take $\mathcal{A}$ to be the discrete sigma–algebra. We assume w_t to be measurable and integrable, and we let

$$\bar{w} = \int_{t \in N} w_t d\mu(t).$$

denote the average income in society. Furthermore, we denote the types with below average, average, and above average pre-tax income by N^-, N^0, and N^+, respectively. We make the regularity assumption that the probability is zero that a citizen has pre-tax income exactly equal to the average level, $\mu(N^0) = 0$. Summarizing, our assumptions are as follows.

Assumption 2.1 It is assumed that w_t is measurable and integrable. Moreover, $\mu(N^0) = 0$, $\mu(N^-) > 0$, and $\mu(N^+) > 0$. If N is finite, we take $\mathcal{A}$ to be the discrete sigma–algebra.

We let $m^- = \mu(N^-)$ and $m^+ = \mu(N^+)$ be the probability mass of citizens with below and above average income, respectively.

We address the question of what fraction $\beta \in B = [0, 1]$ of a citizen's income should be taxed in order to reach a procedurally fair income distribution. Under a tax rate of β, total tax revenues are equal to $\beta\bar{w}$. Under the requirement of a balanced budget and equal distribution of tax revenues across citizens, we find that after-tax income of citizen t is equal to

$$\beta\bar{w} + (1 - \beta)w_t = w_t + \beta(\bar{w} - w_t).$$

When $\beta = 0$, we obtain the situation of laissez-faire, where after-tax income is equal to pre-tax income. The case $\beta = 1$ corresponds to complete redistribution. Since we will assume utility functions to be strictly increasing, a citizen's preferred point is laissez-faire when his income is above $\bar{w}$ and complete redistribution for income below $\bar{w}$. A household with average income is indifferent with regard to the tax policy chosen.

A citizen of type t evaluates the after-tax income using a von Neumann–Morgenstern utility function $u_t : [0, +\infty) \to \mathbb{R}$.

Assumption 2.2 For each $t \in N$, the utility function u_t is continuous, concave, strictly increasing, and $u_t(0) = 0$.

If the tax rate $\beta \in B$ is agreed upon, citizen t enjoys a utility of

$$v_t(\beta) = u_t(w_t + \beta(\bar{w} - w_t)).$$

The function $v_t : B \to \mathbb{R}$ defined by the equation above is the *indirect utility function* of type t. The indirect utility function v_t is nonnegative, continuous, and concave. Moreover, it holds that $v_t(\beta) > 0$ whenever $0 < \beta < 1$. For each $t \in N^0$, v_t is a positive constant. For each $t \in N^-$, v_t is strictly increasing, while for each $t \in N^+$, it is strictly decreasing.

Consider some $t \in N^- \cup N^+$. We denote the inverse of the utility function u_t by x_t, so for v a feasible utility level of citizen t, $x_t(v)$ corresponds to the amount of consumption leading to this utility level. We define

$$h_t(v) = \frac{w_t - x_t(v)}{w_t - \bar{w}}.$$

The function h_t is the inverse of the indirect utility function v_t and corresponds to the tax rate that is needed to achieve a particular utility level. For each $t \in N^-$, the function h_t is strictly increasing and convex, while for each $t \in N^+$, it is strictly decreasing and concave.

3 The unanimous approval procedure

The *unanimous approval procedure* is defined as follows. In every period, each citizen has an equal chance to be selected as the proposer. Since the distribution of types is given by the probability measure μ, we have that a citizen of type t is selected as the proposer according to the probability measure μ. If in period τ a citizen of type t is selected as the proposer, he makes a proposal $p_t \in B$. After observing p_t, citizens decide simultaneously whether to accept or to reject the proposal. If all citizens accept, p_t is implemented, and the utility of a citizen of type i is given by $v_i(p_t)$. If one or more citizens reject, the procedure breaks down with probability $1 - \delta$ leading to the "no agreement position" and a consumption of 0 for all citizens.[1] With probability δ, the procedure is repeated in period $\tau + 1$ and starts with the selection of a proposer according to the probability measure μ. We are interested in the proposals that are made in the limit when δ converges to one.

The unanimous approval procedure leads to a well-defined game in extensive form. A *stationary strategy* of a citizen of type t, $\sigma_t = (p_t, A_t)$, consists of a *proposal* $p_t \in B$ and an *acceptance set* $A_t \subset B$. The acceptance set A_t consists of those proposals that are accepted by a citizen of type t. This specification results in a stationary strategy

[1] A procedurally fair procedure should offer equal opportunities to all involved parties. This implies in particular that, in the absence of agreement, all citizens consume the same amount. We choose this amount to be zero to ensure that all citizens prefer any positive tax rate to the disagreement point.

because p_t and A_t are time and history independent. The *social acceptance set* consists of the proposals that are accepted by all citizens and is given by $A = \cap_{t \in N} A_t$.

Under appropriate measurability conditions, a strategy profile $\sigma = (\sigma_t)_{t \in N}$ determines a unique probability measure over the tax rates that are implemented and thereby the expected utility $U_t(\sigma)$ for each citizen of type t as evaluated at the beginning of the game.[2] Since strategies are stationary, $U_t(\sigma)$ is also the continuation utility of a citizen of type t, the expected utility as evaluated at the beginning of any time period τ.

A strategy profile $\sigma = (p_t, A_t)_{t \in N}$ is an *equilibrium* if in any subgame the proposal p_t satisfies *sequential rationality* and the acceptance set A_t satisfies *weak dominance.* Sequential rationality means that in all subgames, for all $t \in N$, p_t is optimal given the strategies of all citizens. An acceptance set satisfies weak dominance when a citizen accepts a proposal if and only if the utility of the proposal is at least as large as the continuation utility, so for all $t \in N$, $\beta \in A_t$ if and only if $v_t(\beta) \geq \delta U_t(\sigma)$. The restrictions on the acceptance sets are standard in the bargaining literature. We will define the *procedurally fair tax rates* as those that are proposed in an equilibrium when the continuation probability δ tends to one.

4 Tax equilibrium

Consider a strategy profile $\sigma = (p_t, A_t)_{t \in N}$ which is an equilibrium. Let $N^{\mathrm{a}} = \{t \in N \mid p_t \in A\}$ denote the set of citizens whose proposal is accepted. Citizen i's expected utility $U_i(\sigma)$ of the strategy profile σ satisfies the equation

$$U_i(\sigma) = \int_{t \in N^{\mathrm{a}}} v_i(p_t) d\mu(t) + (1 - \mu(N^{\mathrm{a}}))\delta U_i(\sigma), \quad i \in N. \tag{4.1}$$

Moreover, we have that

$$A_i = \{\beta \in [0, 1] \mid v_i(\beta) \geq \delta U_i(\sigma)\}, \quad i \in N. \tag{4.2}$$

Theorem 4.1 *Consider a strategy profile $\sigma = (p_t, A_t)_{t \in N}$ and the induced social acceptance set A. If σ is an equilibrium strategy profile, then*

[A] *For every $i \in N$, $p_i \in A$.*
[B] *The interval $[\delta \bar{p}, 1 - \delta + \delta \bar{p}]$ is contained in A, where $\bar{p}$ denotes the expected equilibrium proposal:*

$$\bar{p} = \int_{t \in N} p_t d\mu(t).$$

[C] *The set A is a non-empty, compact interval.*

[2] Suitable measurability conditions are that $p : N \to [0, 1]$ be $\mathcal{A}$-measurable, and $\cap_{t \in N} A_t$ be Borel-measurable.

Proof We must have $\mu(N^{\mathrm{a}}) > 0$, for otherwise $U_t(\sigma) = 0$ from Eq. (4.1) for all $t \in N$, and citizens would have a profitable deviation by proposing any tax rate in $(0, 1)$. Let $\bar{p}^{\mathrm{a}}$ denote the average proposal of the citizens in N^{a}, that is,

$$\bar{p}^{\mathrm{a}} = \frac{1}{\mu(N^{\mathrm{a}})} \int_{t \in N^{\mathrm{a}}} p_t d\mu(t).$$

This expression is well defined since we have argued $\mu(N^{\mathrm{a}})$ to be strictly positive. We have for every $i \in N$,

$$v_i(\bar{p}^{\mathrm{a}}) \geq \frac{1}{\mu(N^{\mathrm{a}})} \int_{t \in N^{\mathrm{a}}} v_i(p_t) d\mu(t) = \frac{1 - \delta + \delta\mu(N^{\mathrm{a}})}{\mu(N^{\mathrm{a}})} U_i(\sigma) \geq U_i(\sigma),$$

where the first inequality follows from the concavity of v_i and the equality from Eq. (4.1).

Now consider the interval $[\delta\bar{p}^{\mathrm{a}}, 1 - \delta + \delta\bar{p}^{\mathrm{a}}]$. Each of its points can be written in the form $(1 - \delta)\beta + \delta\bar{p}^{\mathrm{a}}$ for some $\beta \in [0, 1]$. Since for every $i \in N$,

$$v_i((1 - \delta)\beta + \delta\bar{p}^{\mathrm{a}}) \geq (1 - \delta)v_i(\beta) + \delta v_i(\bar{p}^{\mathrm{a}}) \geq \delta v_i(\bar{p}^{\mathrm{a}}) \geq \delta U_i(\sigma),$$

each point of the interval $[\delta\bar{p}^{\mathrm{a}}, 1 - \delta + \delta\bar{p}^{\mathrm{a}}]$ is unanimously accepted.

We conclude the proof of the theorem by demonstrating that each citizen in $N \setminus N^{\mathrm{a}}$ has a profitable deviation from σ, thereby obtaining a contradiction, and showing that $N^{\mathrm{a}} = N$. Thus, take a citizen $t \in N$ such that p_t is not an element of A. Citizen t's equilibrium proposal p_t is rejected and leads to utility $\delta U_t(\sigma)$ for citizen t. To obtain the desired contradiction, it is sufficient to show that there is a point $\beta \in A$ such that $v_t(\beta) > \delta U_t(\sigma)$, for then proposing β, rather than p_t, would be a profitable deviation for citizen t. Suppose first that $t \in N^- \cup N^+$. As we have seen above, $v_t(\beta) \geq \delta U_t(\sigma)$ for each β such that $\delta\bar{p}^{\mathrm{a}} \leq \beta \leq 1 - \delta + \delta\bar{p}^{\mathrm{a}}$. Since v_t is strictly increasing for $t \in N^-$ and strictly decreasing for $t \in N^+$, we must have $v_t(\beta) > \delta U_t(\sigma)$ whenever $\delta\bar{p}^{\mathrm{a}} < \beta < 1 - \delta + \delta\bar{p}^{\mathrm{a}}$. Suppose next $t \in N^0$. Then, the function v_t is a positive constant, so it follows that any $\beta \in [0, 1]$, and in particular any $\beta \in A$, has the property that $v_t(\beta) > \delta v_t(\beta) \geq \delta U_t(\sigma)$.

From Eq. (4.2) and the fact the function v_i is continuous and concave, it follows that citizen i's acceptance set A_i is a compact interval. It follows that also the social acceptance set is a compact interval. □

Henceforth, we denote the interval A by $[\beta^-, \beta^+]$. By the preceding theorem, the set A contains an interval of length $1 - \delta$, so it follows that

$$\beta^+ - \beta^- \geq 1 - \delta. \tag{4.3}$$

The proposal of a citizen of type t is the point in A closest to his most preferred point. Therefore, if $t \in N^+$ then $p_t = \beta^-$ and if $t \in N^-$ then $p_t = \beta^+$. For $t \in N^0$, the indirect utility function is a constant, so citizen t's proposal can be an arbitrary element of the social acceptance set. Hence, Equation (4.1) simplifies to

$$U_i(\sigma) = m^- v_i(\beta^+) + m^+ v_i(\beta^-), \quad i \in N.$$

Thus, the social acceptance set $[\beta^-, \beta^+]$ and the individual acceptance sets in an equilibrium strategy profile are such that

$$[\beta^-, \beta^+] = \bigcap_{i \in N} A_i, \tag{4.4}$$

$$A_i = \{\beta \in [0, 1] \mid v_i(\beta) \geq \delta[m^- v_i(\beta^+) + m^+ v_i(\beta^-)]\}, \quad i \in N. \tag{4.5}$$

In what follows, we provide a characterization of the endpoints of the social acceptance set in an equilibrium strategy profile as a fixed point of an appropriately defined function. Recall that h_i has been defined for each $i \in N^- \cup N^+$ as the inverse of the function v_i. For each $i \in N^- \cup N^+$, we define

$$f_{1i}(\beta^-) = h_i(\alpha^+ v_i(\beta^-)), \quad \text{where} \quad \alpha^+ = \delta m^+/(1 - \delta + \delta m^+), \tag{4.6}$$

$$f_{2i}(\beta^+) = h_i(\alpha^- v_i(\beta^+)), \quad \text{where} \quad \alpha^- = \delta m^-/(1 - \delta + \delta m^-). \tag{4.7}$$

The definition of $f_{1i}(\beta^-)$ ensures that a citizen i is indifferent between a tax rate equal to $f_{1i}(\beta^-)$ and a gamble where the tax rate is $f_{1i}(\beta^-)$ with probability δm^-, a tax rate β^- with probability δm^+, and zero consumption with probability $1 - \delta$. Similarly, the definition of $f_{2i}(\beta^+)$ ensures that a citizen i is indifferent between a tax rate equal to $f_{2i}(\beta^+)$ and a gamble where the tax rate is $f_{2i}(\beta^+)$ with probability δm^+, a tax rate β^+ with probability δm^-, and zero consumption with probability $1 - \delta$.

Define the functions $f_1 : [0, 1] \to [0, 1]$ and $f_2 : [0, 1] \to [0, 1]$ by letting

$$f_1(\beta^-) = \min\left\{1, \inf_{i \in N^+} f_{1i}(\beta^-)\right\},$$

$$f_2(\beta^+) = \max\left\{0, \sup_{i \in N^-} f_{2i}(\beta^+)\right\}.$$

Lemma 4.2 *Let $[\beta^-, \beta^+]$ be the social acceptance set induced by an equilibrium strategy profile. Then,*

$$\beta^+ = f_1(\beta^-),$$

$$\beta^- = f_2(\beta^+).$$

Proof We show that the second of these equations holds. The proof of the first one is similar. Since β^- is in the social acceptance set, it belongs to every citizen's individual acceptance set A_i. Hence, by (4.5)

$$v_i(\beta^-) \geq \delta m^- v_i(\beta^+) + \delta m^+ v_i(\beta^-), \quad i \in N.$$

Using the fact that $m^- + m^+ = 1$ and rearranging, we find that

$$v_i(\beta^-) \geq \alpha^- v_i(\beta^+), \quad i \in N.$$

Now consider a citizen $i \in N^-$. For each such citizen i, the function h_i is strictly increasing. Hence, applying h_i to the preceding equation yields

$$\beta^- = h_i(v_i(\beta^-)) \geq h_i(\alpha^- v_i(\beta^+)) = f_{2i}(\beta^+), \quad i \in N^-.$$

We conclude that $\beta^- \geq f_2(\beta^+)$. Now suppose $\beta^- \geq f_2(\beta^+) + \varepsilon$ for some $\varepsilon > 0$. Then,

$$\beta^- - \varepsilon \geq h_i(\alpha^- v_i(\beta^+)), \quad i \in N^-.$$

Applying the increasing function v_i to the above inequality yields

$$v_i(\beta^- - \varepsilon) \geq v_i(h_i(\alpha^- v_i(\beta^+))) = \alpha^- v_i(\beta^+),\ i \in N^-.$$

But, then the point

$$\beta = (1 - \delta m^+)(\beta^- - \varepsilon) + \delta m^+ \beta^-$$

is in the acceptance set of each citizen $i \in N^-$, because

$$\begin{aligned} v_i(\beta) &\geq (1 - \delta m^+) v_i(\beta^- - \varepsilon) + \delta m^+ v_i(\beta^-) \\ &\geq (1 - \delta m^+) \alpha^- v_i(\beta^+) + \delta m^+ v_i(\beta^-) \\ &= \delta m^- v_i(\beta^+) + \delta m^+ v_i(\beta^-) \\ &= \delta U_i(\sigma), \end{aligned}$$

where the inequality in the first line follows by concavity of v_i. Since $\beta < \beta^-$ is clearly in the acceptance set of each citizen in $i \in N^0 \cup N^+$, it is in the social acceptance set. This clearly contradicts the fact that β^- is the left endpoint of the social acceptance set. □

The first equality in the preceding lemma expresses that β^+ should be sufficiently low to make it acceptable for all above average income citizens, and similarly, the second equality specifies that β^- should be sufficiently high to be acceptable for all below average income citizens. It motivates the following definition.

Definition 4.3 A *tax equilibrium* is pair (β^-, β^+) satisfying

$$\begin{aligned} \beta^+ &= f_1(\beta^-), \\ \beta^- &= f_2(\beta^+). \end{aligned}$$

As it is immediate from the definition, every equilibrium strategy profile induces a tax equilibrium. We now establish the converse, stating every tax equilibrium corresponds to an equilibrium strategy profile, showing that the concepts are equivalent.

Theorem 4.4 *Given a tax equilibrium* (β^-, β^+)*, there is an equilibrium strategy profile* σ *with social acceptance set equal to* $[\beta^-, \beta^+]$.

Proof Define p_t to be β^+ for each $t \in N^0 \cup N^-$ and β^- for each $t \in N^+$. Define the individual acceptance sets by Eq. (4.5), so they satisfy weak dominance by definition. An argument similar to that in the proof of Lemma 4.2 can be used to show that Eq. (4.4) holds.

It remains to verify sequential rationality of the proposals. It is a routine excercise to demonstrate that no type has a profitable one-shot deviation from σ when being the proposer. Here, a one-shot deviation in a subgame is a single deviation by a player at the root of the subgame. It follows from a standard argument, see for instance Fudenberg and Tirole (1991), that if there is a subgame where a type has some profitable deviation from σ, then there must also be a subgame where this type has a profitable one-shot deviation. □

Define the function $f : [0, 1]^2 \to [0, 1]^2$ by letting

$$f(\beta^+, \beta^-) = (f_1(\beta^-), f_2(\beta^+)).$$

Clearly, the tax equilibria are exactly the fixed points of the function f.

Theorem 4.5 *A tax equilibrium exists.*

Proof The set $[0, 1]^2$ is a complete lattice, and the function f has the property that $f_1(\beta) \le f_1(\beta')$ and $f_2(\beta) \le f_2(\beta')$ whenever $\beta \le \beta'$. Hence, Tarski's fixed point theorem implies that f has a fixed point. □

Theorem 4.6 *A tax equilibrium is unique.*

Proof Consider the function $g : [0, 1] \to \mathbb{R}$ defined by $g(\beta) = f_1(f_2(\beta)) - \beta$. Obviously, a zero point β of g is in a one-to-one relationship with a tax equilibrium $(f_2(\beta), \beta)$, and it follows from the previous paragraph that g has at least one zero point. We argue that g is strictly decreasing, thereby showing that it has a unique zero point.

We write

$$g(\beta) = f_1(f_2(\beta)) - f_2(\beta) + f_2(\beta) - \beta.$$

We show first that $f_2(\beta) - \beta$ is strictly decreasing in β. Consider some $i \in N^-$ and β, β' with $\beta < \beta'$. Since h_i is a convex and increasing function, we have

$$\begin{aligned} f_{2i}(\beta) - f_{2i}(\beta') &= h_i(\alpha^- v_i(\beta)) - h_i(\alpha^- v_i(\beta')) \\ &\ge \alpha^-[h_i(v_i(\beta)) - h_i(v_i(\beta'))] \\ &= \alpha^-(\beta - \beta'), \end{aligned}$$

hence

$$f_{2i}(\beta) \ge \alpha^-(\beta - \beta') + f_{2i}(\beta').$$

Taking the supremum with respect to $i \in N^-$ on both sides of the preceding equation yields

$$\sup_{i \in N^-} f_{2i}(\beta) \geq \alpha^-[\beta - \beta'] + \sup_{i \in N^-} f_{2i}(\beta').$$

Hence,

$$\sup_{i \in N^-} f_{2i}(\beta) - \beta \geq \sup_{i \in N^-} f_{2i}(\beta') - \beta' + (1 - \alpha^-)(\beta' - \beta) > \sup_{i \in N^-} f_{2i}(\beta') - \beta',$$

which implies that

$$\sup_{i \in N^-} f_{2i}(\beta) - \beta$$

is strictly decreasing in β. Notice that

$$f_2(\beta) - \beta = \max\Big\{ - \beta, \sup_{i \in N^-} f_{2i}(\beta) - \beta\Big\},$$

so, it is the maximum of two expressions strictly decreasing in β, and therefore strictly decreasing itself. A completely symmetric argument shows that $f_1(\beta) - \beta$ is strictly decreasing in β. Since the function f_2 is increasing, $f_1(f_2(\beta)) - f_2(\beta)$ is decreasing in β. Since $f_2(\beta) - \beta$ is strictly decreasing in β, we have that g is a strictly decreasing function. □

The unanimous approval procedure results in a unique tax equilibrium (β^-, β^+) with $\beta^- < \beta^+$. The tax rate proposed by an above average income citizen is β^-, whereas β^+ is proposed by below average income citizens. A proposal β strictly smaller than β^- would be rejected by at least one below average income citizen, and a proposal β strictly greater than β^+ by at least one above average income citizen. At equilibrium, rejections do not occur.

5 Boldness and bargaining power

Aumann and Kurz (1977) observe that the Nash bargaining solution for the two player case can be characterized as selecting a division of the potential surplus at which the players are equally bold. Assuming a differentiable utility function, the *boldness* of a citizen of type t at consumption c_t equals $u_t'(c_t)/u_t(c_t)$. In this section, we study the relation between tax equilibria and the boldness of citizens.

Consider a gamble where a citizen of type t receives consumption 0 with probability q_t and $c_t + \varepsilon$ with probability $1 - q_t$, where $\varepsilon > 0$. Let $q_t(c_t, \varepsilon)$ be the maximum probability for which a citizen of type t weakly prefers the gamble over consuming c_t for sure. As pointed out in Roth (1989), boldness corresponds to the maximum probability for which type t is willing to accept the gamble, per dollar of additional gains, when ε tends to zero. That is, the boldness of a citizen of type t at consumption c_t is equal to $\lim_{\varepsilon \downarrow 0} q_t(c_t, \varepsilon)/\varepsilon$.

Since we do not assume that utility functions are differentiable, we extend the notion of boldness in the following way. The *boldness* of a citizen of type t at $c_t > 0$ is defined as

$$b_t(c_t) = \partial_+ u_t(c_t)/u_t(c_t),$$

where $\partial_+ u_t$ denotes the right derivative of u_t. Left derivatives will be denoted by ∂_-. We define $b_t(0) = +\infty$. Since utility functions are assumed to be concave, the right derivative of the utility function is well defined. Moreover, boldness is a strictly decreasing function of consumption, as the numerator in the definition of boldness is decreasing by concavity of u_t and the denominator is strictly increasing since u_t is strictly increasing in consumption.

The next result shows that also in the absence of differentiability, boldness still admits an interpretation in terms of gambles.

Theorem 5.1 *For $c_t > 0$, it holds that $b_t(c_t) = \lim_{\varepsilon \downarrow 0} q_t(c_t, \varepsilon)/\varepsilon$.*

Proof By continuity of u_t and for ε sufficiently small, it holds that

$$u_t(c_t) = q_t(c_t, \varepsilon)u_t(0) + (1 - q_t(c_t, \varepsilon))u_t(c_t + \varepsilon).$$

We use that $u_t(0) = 0$ and rearrange terms to find that

$$\frac{q_t(c_t, \varepsilon)}{\varepsilon} = \frac{u_t(c_t + \varepsilon) - u_t(c_t)}{\varepsilon} \frac{1}{u_t(c_t + \varepsilon)}.$$

When ε tends to zero, we find that the first term in the product converges to the right derivative of u_t at c_t due to concavity of u_t and the second term converges to $u_t(c_t)$ because of continuity, so

$$\lim_{\varepsilon \downarrow 0} \frac{q_t(c_t, \varepsilon)}{\varepsilon} = \frac{\partial_+ u_t(c_t)}{u_t(c_t)}.$$

□

It follows immediately from Theorem 5.1 and the concavity of u_t that an alternative characterization of $b_t(c_t)$ can be given as the limit inferior of difference quotients of u_t at c_t. We observe that b_t is a strictly decreasing function, which may have points of discontinuity.

For each $\beta \in (0, 1)$, define

$$d_{1t}(\beta) = -\frac{\partial_- v_t(\beta)}{v_t(\beta)} = (w_t - \bar{w})b_t(w_t + \beta(\bar{w} - w_t)),$$
$$d_{2t}(\beta) = \frac{\partial_+ v_t(\beta)}{v_t(\beta)} = (\bar{w} - w_t)b_t(w_t + \beta(\bar{w} - w_t)).$$

The above expressions are well defined since $v_t(\beta) > 0$ for each $\beta \in (0, 1)$.

We will show that these functions yield the appropriate measure of boldness of a citizen of type t at a tax rate equal to β, where d_{1t} applies to types with above average income and d_{2t} to those with below average income. We refer to this measure of boldness as *tax rate boldness*. Indeed, for a type t in N^+, tax rate boldness $d_{1t}(\beta)$ is equal to the usual definition of boldness applied to the (increasing) indirect utility function $-v_t$, which in turn is equal to his boldness at the consumption $w_t + \beta(\bar{w} - w_t)$ induced by the tax rate β, multiplied by $w_t - \bar{w}$, which equals the marginal change in consumption due to a change in the tax rate. Similarly, $d_{2t}(\beta)$ is the appropriate measure for boldness corresponding to tax rate proposals for below average income types.

Since boldness is strictly decreasing in consumption, it holds that tax rate boldness d_{1t} is strictly increasing in β, whereas d_{2t} is strictly decreasing in β. What will matter in the end is to convince the boldest citizen to accept a proposal, which motivates the following definitions. For $\beta \in (0, 1)$, let

$$d_1(\beta) = \sup_{t \in N^+} d_{1t}(\beta),$$
$$d_2(\beta) = \sup_{t \in N^-} d_{2t}(\beta).$$

At this point, it is useful to provide some intuition as to why the tax rate boldness functions d_1 and d_2 are useful in characterizing tax equilibria. Suppose indirect utility functions are differentiable and suppose that we can replace the indirect utility functions by their first-order approximations around the expected equilibrium offer $\bar{\beta}$:

$$v_t(\beta) \approx v_t(\bar{\beta}) + v_t'(\bar{\beta})(\beta - \bar{\beta}).$$

A type $t \in N$ accepts a proposal β if and only if $v_t(\bar{\beta}) + v_t'(\bar{\beta})(\beta - \bar{\beta}) \geq \delta v_t(\bar{\beta})$. Using that for $t \in N^+$, $v_t'(\bar{\beta}) < 0$ and for $t \in N^-$, $v_t'(\bar{\beta}) > 0$, we find that a type $t \in N^+$ accepts all proposals β satisfying

$$\beta \leq \bar{\beta} - (1 - \delta)\frac{v_t(\bar{\beta})}{v_t'(\bar{\beta})},$$

whereas types $t \in N^-$ accept all proposals β satisfying

$$\beta \geq \bar{\beta} - (1 - \delta)\frac{v_t(\bar{\beta})}{v_t'(\bar{\beta})}.$$

The endpoints of the social acceptance set are therefore given by

$$\beta^- = \bar{\beta} - (1 - \delta) \inf_{t \in N^-} \left\{\frac{v_t(\bar{\beta})}{v_t'(\bar{\beta})}\right\} = \bar{\beta} - \frac{1 - \delta}{d_2(\bar{\beta})},$$
$$\beta^+ = \bar{\beta} - (1 - \delta) \sup_{t \in N^+} \left\{\frac{v_t(\bar{\beta})}{v_t'(\bar{\beta})}\right\} = \bar{\beta} + \frac{1 - \delta}{d_1(\bar{\beta})}.$$

Since the expected equilibrium offer is given by $\bar{\beta} = m^+\beta^- + m^-\beta^+$, we find that

$$\bar{\beta} = m^+\bar{\beta} - m^+\frac{1-\delta}{d_2(\bar{\beta})} + m^-\bar{\beta} + m^-\frac{1-\delta}{d_1(\bar{\beta})}.$$

Since $m^+ + m^- = 1$, we can rearrange terms and find that $m^+d_1(\bar{\beta}) = m^-d_2(\bar{\beta})$. We will demonstrate in Sect. 6 that the intuition coming from first-order approximations can be made precise.

For each $\beta \in (0, 1)$, we define

$$z(\beta) = m^-d_2(\beta) - m^+d_1(\beta).$$

It will be convenient to extend the function z to the whole interval [0, 1] by letting $z(0) = +\infty$ and $z(1) = -\infty$. The function z is the difference between the supremum of tax rate boldness among below average income types multiplied by their probability mass m^- and the supremum of tax rate boldness among above average income types multiplied by their probability mass m^+. The function z can be interpreted as an "excess bargaining power function," where positive values of $z(\beta)$ indicate that below average income types have more bargaining power than above average income types, causing increases in the tax rate, and the reverse when $z(\beta)$ is negative. Indeed, $m^+d_1(\beta)$ represents the bargaining power of the rich, and $m^-d_2(\beta)$ the bargaining power of the poor. We can think of $z(\beta)$ as the direction and magnitude by which β would be adjusted as a consequence of the prevailing bargaining forces.

Theorem 5.2 *The excess bargaining power function z is strictly decreasing on* [0, 1].

The proof of Theorem 5.2 can be found in Appendix A. It consists of showing that d_1 is finite and strictly increasing and that d_2 is finite and strictly decreasing.

A point $\beta^* \in [0, 1]$ is called a *generalized zero point* of z if $z(\beta) > 0$ for $\beta < \beta^*$ and $z(\beta) < 0$ for $\beta > \beta^*$. When applied to the case $\beta^* = 0$, then being a generalized zero point means $z(\beta) < 0$ for all $\beta > 0$, and similarly $\beta^* = 1$ is a generalized zero point of z if $z(\beta) > 0$ for all $\beta < 1$. Since z is strictly decreasing by Theorem 5.2, it follows that z has a unique generalized zero point.

Corollary 5.3 *The excess bargaining power function z has a unique generalized zero point.*

At the generalized zero point β^* of z, the difference between $m^-d_2(\beta)$ and $m^+d_1(\beta)$ is minimized and is equal to zero if z is continuous at β^*. The point β^* is therefore the tax rate where the bargaining power of citizens with below average income is as close as possible to that of above average income citizens. We will show in Theorem 6.4 that all proposals made in tax equilibria converge to the unique generalized zero point of z when δ converges to 1.

When the set of citizens is finite, the excess bargaining function z as defined above coincides with the characteristic function ξ as defined in Predtetchinski (2011), who also proves that ξ is a strictly decreasing function. His proof however only applies when there are finitely many citizens.

6 Procedurally fair tax rates

In this section, we establish that the procedurally fair tax rate is equal to the generalized zero point of the function z.

We show first that along any convergent sequence of tax equilibria, as the continuation probability δ converges to one, the social acceptance set converges to a singleton set.

Theorem 6.1 *Let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence converging to 1 and, for $n \in \mathbb{N}$, let (β_n^-, β_n^+) be the tax equilibrium corresponding to δ_n. Suppose the sequence $(\beta_n^-)_{n\in\mathbb{N}}$ converges to a point β^- and the sequence $(\beta_n^+)_{n\in\mathbb{N}}$ converges to β^+. Then, $\beta^- = \beta^+$.*

Proof We know that $\beta_n^- \leq \beta_n^+$ for each n. Hence, $\beta^- \leq \beta^+$. The point β_n^-, being an element of the social acceptance set in a tax equilibrium, is accepted by all citizens. In particular, for $i \in N^-$, we have $v_i(\beta_n^-) \geq \delta_n(m^- v_i(\beta_n^+) + m^+ v_i(\beta_n^-))$. Taking the limit of both sides of the inequality as n goes to infinity, we obtain $v_i(\beta^-) \geq m^- v_i(\beta^+) + m^+ v_i(\beta^-)$. Rearranging yields $v_i(\beta^-) \geq v_i(\beta^+)$. Since v_i is an increasing function for $i \in N^-$, we have $\beta^- \geq \beta^+$. □

We define a procedurally fair tax rate as the limit of proposals in a tax equilibrium when the continuation probability δ converges to one.

Definition 6.2 The tax rate β is *procedurally fair* if it is the limit of a sequence $(\beta_{t,n})_{n\in\mathbb{N}}$, where $\beta_{t,n}$ is the proposal of a citizen t in the tax equilibrium corresponding to δ_n and $(\delta_n)_{n\in\mathbb{N}}$ is a sequence converging to one.

Since the first-order approximations used in Section 5 are not exact, it is not exactly true that the expected proposal in the tax equilibrium equals the zero point β^* of z, even if all indirect utility functions were differentiable. Theorem 6.3 demonstrates that nevertheless it is true that $\beta^- \leq \beta^* \leq \beta^+$.

Theorem 6.3 *The tax equilibrium (β^-, β^+) satisfies $\beta^- \leq \beta^* \leq \beta^+$, where β^* is the generalized zero point of z.*

The proof of Theorem 6.3 is relegated to Appendix B. Although our proof technique is similar to that in Predtetchinski (2011), we are able to simplify some of the derivations using the special structure of the payoff functions at hand, in particular the fact that v_t is increasing for each $t \in N^-$ and decreasing for each $t \in N^+$.

Combining Theorem 6.1 and 6.3, it follows that the procedurally fair tax rate is unique and is equal to the generalized zero point of the function z, so we obtain the following results.

Theorem 6.4 *Let $(\delta_n)_{n\in\mathbb{N}}$ be a sequence converging to 1 and, for $n \in \mathbb{N}$, let (β_n^-, β_n^+) be the tax equilibrium of the game corresponding to δ_n. Then, both sequences $(\beta_n^-)_{n\in\mathbb{N}}$ and $(\beta_n^+)_{n\in\mathbb{N}}$ converge to the generalized zero point β^* of z.*

Corollary 6.5 *A society has a unique procedurally fair tax rate.*

7 Two illustrative examples

In this section, we consider two special cases of the model. The case where the citizens exhibit constant relative risk aversion and the case where there are sufficiently many risk-neutral citizens. We conclude in both cases that the procedurally tax rate equals the total measure of below average income citizens, m^-.

7.1 The case of constant relative risk aversion

As an illustration, consider the case where all citizens $t \in N$ have constant relative risk aversion utility functions

$$u_t(c_t) = c_t^{1-\gamma}, \quad c_t \in \mathbb{R}_+,$$

where $\gamma \in [0, 1)$ is the coefficient of relative risk aversion. Moreover, we assume that

$$\inf_{t \in N} w_t = 0 \quad \text{and} \quad \sup_{t \in N} w_t = +\infty.$$

It is straightforward to derive that tax rate boldness of an above average income citizen $t \in N^+$ is given by

$$d_{1t}(\beta) = \frac{(1-\gamma)(w_t - \bar{w})}{(1-\beta)w_t + \beta\bar{w}}, \quad \beta \in [0, 1].$$

Since, for a given tax rate β, this expression is increasing in w_t, it attains its maximum value when t is equal to the highest income citizen. Since the support of μ is unbounded, we have that the bargaining power of the rich is given by

$$m^+ d_1(\beta) = m^+ \sup_{t \in N^+} d_{1t}(\beta) = m^+ \frac{1-\gamma}{1-\beta}, \quad \beta \in [0, 1).$$

The tax rate boldness of a below average income citizen $t \in N^-$ is equal to

$$d_{2t}(\beta) = \frac{(1-\gamma)(\bar{w} - w_t)}{(1-\beta)w_t + \beta\bar{w}}, \quad \beta \in (0, 1].$$

For given tax rate β, this expression is decreasing in w_t, and the maximum over $t \in N^-$ is therefore attained by the lowest income citizen. We have that the bargaining power of the poor is equal to

$$m^- d_2(\beta) = m^- \sup_{t \in N^-} d_{2t}(\beta) = m^- \frac{1-\gamma}{\beta}, \quad \beta \in (0, 1].$$

The unique zero point of the excess bargaining power function

$$z(\beta) = m^- d_2(\beta) - m^+ d_1(\beta) = m^- \frac{1-\gamma}{\beta} - m^+ \frac{1-\gamma}{1-\beta}, \quad \beta \in (0, 1),$$

is given by $\beta = m^-$. The unique procedurally fair tax rate of a society populated by constant relative risk aversion citizens is equal to the probability mass of below average income citizens, irrespective of the shape of the distribution μ of types within the population.

7.2 Heterogeneous societies

In this section, we analyze the procedurally fair tax rate for societies with sufficiently heterogeneous citizens. What we have in mind is that there is sufficient dispersion in preferences among citizens, in particular in terms of boldness. From a technical point of view, we make the assumption that there are risk-neutral types having very high and risk-neutral types having very low income.

Let S be the set of citizens $t \in N$ such that $u_t(x) = x$, $x \in \mathbb{R}_+$. We assume that

$$\inf_{t \in S} w_t = 0 \quad \text{and} \quad \sup_{t \in S} w_t = +\infty.$$

The importance of risk-neutral individuals stems from the fact that these individuals display the greatest boldness in the society. This implies that it is effectively the risk-neutral individuals who determine the endpoints of the social acceptance set.

We derive in Appendix C that

$$d_1(\beta) = \frac{1}{1-\beta}, \tag{7.1}$$

$$d_2(\beta) = \frac{1}{\beta}. \tag{7.2}$$

The excess bargaining power function z is then given by

$$z(\beta) = \frac{m^-}{\beta} - \frac{m^+}{1-\beta}.$$

Clearly, $z(m^-) = 0$, so m^- is the procedurally fair tax rate.

The result that emerges from this paper is that under a wide variety of circumstances, m^- appears as the procedurally fair tax rate. This result is in sharp contrast to the literature on fairness that often argues in favor of complete redistribution, implying a tax rate equal to one. The result is also not out of line of what is observed in reality. Mankiw et al. (2009) report that the average top marginal tax wedge in 2007, which combines the top marginal income tax rate with the rate of value-added tax, is just above 60 percent in OECD countries and was in fact nearly 80 percent in 1984. We remark that this finding is subject to a number of qualifications, including our assumption that pre-tax incomes are not affected by the tax rate.

8 Conclusion

We study the implications of procedural fairness for the share of income that should be redistributed. Procedural fairness is modeled as the outcome of a bargaining procedure where all citizens have an equal chance to propose, and all have to agree to a proposal in order for it to be implemented. Societies are shown to have unique procedurally fair tax rates. We also provide a characterization of the procedurally fair tax rate, which can be used to demonstrate that the procedurally fair tax rate is equal to the probability mass m^- of below average income citizens in a variety of circumstances.

This paper has confined itself to a rather simple economic environment in order to obtain sharp results. Many extensions of the model are worthwhile to investigate. We have started our analysis from exogenously given pre-tax incomes. An interesting avenue for research is to make labor supply endogenous, the main channel through which complete redistribution is avoided in the traditional optimal taxation literature. One would expect that also in our set-up, the incorporation of elastic labor supply leads to a reduction of the procedurally fair tax rate below m^-.

In this paper, we study the share of income that should be redistributed on the grounds of procedural fairness. In terms of tax schedules, this implies a limitation to affine tax schedules, consisting of a fixed subsidy to all citizens and constant marginal tax rates. Mirrlees (1971) argued that affine tax schedules are nearly optimal in the context of the traditional optimal taxation literature. Though this view has been challenged many times, Mankiw et al. (2009) claim that proposals for a flat tax are not inherently unreasonable. It is an open issue whether flat tax schedules are also nearly optimal in our framework.

Appendix A The proof of theorem 5.2

Theorem 5.2 follows immediately from Lemmas 8.1 and 8.2 below.

Lemma 8.1 *The function d_1 is a strictly increasing function on $(0, 1)$ assuming only finite values.*

Proof Consider some $t \in N^+$ and some $\beta \in (0, 1)$. Using $v_t(1) \geq 0$ and the concavity of v_t, we find that

$$-(1-\beta)\partial_- v_t(\beta) \leq v_t(\beta).$$

Rearranging yields the inequality

$$d_{1t}(\beta) \leq \frac{1}{1-\beta}.$$

This establishes the finiteness of d_1 on $(0, 1)$. We show that d_1 is strictly increasing. Let β and β' be elements of $(0, 1)$ such that $\beta < \beta'$. For $t \in N^+$ we have

$$-\partial_- v_t(\beta)(\beta' - \beta) \leq v_t(\beta) - v_t(\beta')$$

We therefore obtain

$$\begin{aligned}-\partial_- v_t(\beta)(v_t(\beta') - \partial_- v_t(\beta')(\beta' - \beta)) &\leq -\partial_- v_t(\beta) v_t(\beta') - \partial_- v_t(\beta')(v_t(\beta) - v_t(\beta')) \\ &= (\partial_- v_t(\beta') - \partial_- v_t(\beta)) v_t(\beta') - \partial_- v_t(\beta') v_t(\beta) \\ &\leq -\partial_- v_t(\beta') v_t(\beta),\end{aligned}$$

where the last inequality uses $v_t(\beta') \geq 0$ and concavity. Dividing both sides of the inequality by $v_t(\beta) v_t(\beta')$ gives

$$d_{1t}(\beta)(1 + (\beta' - \beta) d_{1t}(\beta')) \leq d_{1t}(\beta').$$

Dividing by $d_{1t}(\beta')$ yields the inequality

$$d_{1t}(\beta) \leq \frac{d_{1t}(\beta')}{1 + (\beta' - \beta) d_{1t}(\beta')}.$$

Taking the supremum with respect to $t \in N^+$ yields

$$d_1(\beta) = \sup_{t \in N^+} d_{1t}(\beta) \leq \sup_{t \in N^+} \frac{d_{1t}(\beta')}{1 + (\beta' - \beta) d_{1t}(\beta')} \leq \frac{d_1(\beta')}{1 + (\beta' - \beta) d_1(\beta')} < d_1(\beta').$$

This completes the proof. □

The proof that the function d_2 is strictly decreasing follows by a similar argument and is therefore omitted.

Lemma 8.2 *The function d_2 is a strictly decreasing function on* $(0, 1)$ *assuming only finite values.*

Appendix B The proof of theorem 6.3

Consider the bargaining equilibrium (β^-, β^+) and consider the bargaining power of the poor and the rich, respectively, at the upper bound β^+ of the social acceptance set. Lemma 8.3 states that the size of the social acceptance set $\beta^+ - \beta^-$ multiplied by the bargaining power of the poor $m^- d_2(\beta^+)$ is bounded from above by $(1 - \delta)/\delta$, and is greater than or equal to this number when multiplied by the bargaining power $m^+ d_1(\beta^-)$ of the rich.

Lemma 8.3 *Consider the tax equilibrium* (β^-, β^+). *If* $\beta^+ < 1$, *then*

1. $\delta m^- d_2(\beta^+)(\beta^+ - \beta^-) \leq (1 - \delta)$.

2. $\delta m^+ d_1(\beta^+)(\beta^+ - \beta^-) \geq (1 - \delta)$.

Proof For $t \in N^-$, it holds that the proposal β^- is accepted, so $\delta(m^- v_t(\beta^+) + m^+ v_t(\beta^-)) \leq v_t(\beta^-)$. Rewriting this inequality results in

$$\delta m^-(v_t(\beta^+) - v_t(\beta^-)) \leq (1 - \delta)v_t(\beta^-) \leq (1 - \delta)v_t(\beta^+).$$

Since v_t is concave, we find that

$$\delta m^- \partial_+ v_t(\beta^+)(\beta^+ - \beta^-) \leq (1 - \delta)v_t(\beta^+),$$

and therefore $\delta m^- d_{2t}(\beta^+)(\beta^+ - \beta^-) \leq 1 - \delta$. Since this inequality holds for all $t \in N^-$, we obtain

$$\delta m^- d_2(\beta^+)(\beta^+ - \beta^-) = \sup_{t \in N^-} \delta m^- d_{2t}(\beta^+)(\beta^+ - \beta^-) \leq 1 - \delta.$$

This proves Lemma 8.3.1.

Consider some $t \in N^+$. By concavity of v_t we have that $-\partial_- v_t(\beta^+)(\beta^+ - \beta^-) \geq v_t(\beta^-) - v_t(\beta^+)$. It then follows that

$$d_{1t}(\beta^+)(\beta^+ - \beta^-) \geq \frac{v_t(\beta^-)}{v_t(\beta^+)} - 1.$$

We take the supremum over all $t \in N^+$ and find that

$$d_1(\beta^+)(\beta^+ - \beta^-) \geq \sup_{t \in N^+} \frac{v_t(\beta^-)}{v_t(\beta^+)} - 1.$$

We complete the proof of Lemma 8.3.2 by showing that

$$\delta m^+ \left(\sup_{t \in N^+} \frac{v_t(\beta^-)}{v_t(\beta^+)} - 1 \right) = 1 - \delta,$$

or equivalently

$$\sup_{t \in N^+} \frac{\delta(m^- v_t(\beta^+) + m^+ v_t(\beta^-))}{v_t(\beta^+)} = 1.$$

Since all $t \in N^+$ accept the proposal β^+, we have

$$\frac{\delta(m^- v_t(\beta^+) + m^+ v_t(\beta^-))}{v_t(\beta^+)} \leq 1, \quad t \in N^+.$$

Suppose there is an $\varepsilon > 0$ with the property

$$\frac{\delta(m^- v_t(\beta^+) + m^+ v_t(\beta^-))}{v_t(\beta^+)} \leq 1 - \varepsilon$$

for all $t \in N^+$. Then

$$\begin{aligned} v_t((1-\varepsilon)\beta^+ + \varepsilon) &\geq (1-\varepsilon)v_t(\beta^+) + \varepsilon v_t(1) \geq (1-\varepsilon)v_t(\beta^+) \\ &\geq \delta(m^- v_t(\beta^+) + m^+ v_t(\beta^-)). \end{aligned}$$

All citizens $t \in N^+$ (as well as citizens in $N^- \cup N^0$) therefore accept the proposal $(1-\varepsilon)\beta^+ + \varepsilon > \beta^+$ thereby contradicting that β^+ is the upper bound of the social acceptance set. □

We are now in a position to prove the first half of Theorem 6.3.

Lemma 8.4 *The tax equilibrium* (β^-, β^+) *satisfies* $\beta^* \leq \beta^+$*, where* β^* *is the generalized zero point of* z*.*

Proof The result is obviously true when $\beta^+ = 1$. In particular, if $\delta = 0$ then $\beta^+ = 1$ and $\beta^- = 0$ by Equation (4.3). Suppose now $\delta > 0$ and $\beta^+ < 1$. Subtracting the second inequality of Lemma 8.3 from the first one gives $\delta z(\beta^+)(\beta^+ - \beta^-) \leq 0$. Since by (4.3) $0 < \beta^+ - \beta^-$, it follows that $z(\beta^+) \leq 0$. The result follows since z is a decreasing function. □

The proof of the second half of Theorem 6.3 follows from an analogous argument and is therefore omitted.

Appendix C The derivation of Equations (7.1)–(7.2)

We provide a derivation of Equations (7.1)–(7.2).

For each $t \in N$ and each $c_t > 0$, we have $0 = u_t(0) \leq u_t(c_t) - c_t \partial_+ u_t(c_t)$, where the inequality holds by the concavity of u_t. Rearranging this expression, we obtain that

$$b_t(c_t) \leq \frac{1}{c_t}, \quad c_t > 0. \tag{8.1}$$

For each $t \in N^+ \cap S$, we have

$$d_{1t}(\beta) = \frac{w_t - \bar{w}}{w_t + \beta(\bar{w} - w_t)}, \quad \beta \in (0, 1)$$

This expression is non–decreasing in w_t on $[\bar{w}, +\infty)$, and it converges to $1/(1-\beta)$ as w_t approaches infinity. As the set $\{w_t \mid t \in S\}$ is unbounded from above, it holds that $d_1(\beta) \geq \frac{1}{1-\beta}$. On the other hand, for each $t \in N^+$ it holds that

$$d_{1t}(\beta) = (w_t - \bar{w})b_t(w_t + \beta(\bar{w} - w_t)) \leq \frac{w_t - \bar{w}}{w_t + \beta(\bar{w} - w_t)} \leq \frac{1}{1-\beta}, \quad \beta \in (0, 1),$$

where the first inequality is implied by (8.1). It follows that $d_1(\beta) \leq 1/(1-\beta)$, which proves Eq. (7.1).

Similarly for each citizen $t \in N^- \cap S$, we have

$$d_{2t}(\beta) = \frac{\bar{w} - w_t}{w_t + \beta(\bar{w} - w_t)}, \quad \beta \in (0, 1).$$

Since the infimum of $\{w_t \mid t \in S\}$ is zero, it holds that $d_2(\beta) \geq 1/\beta$. On the other hand, for each $t \in N^-$ it holds that

$$d_{2t}(\beta) = (\bar{w} - w_t)b_t(w_t + \beta(\bar{w} - w_t)) \leq \frac{\bar{w} - w_t}{w_t + \beta(\bar{w} - w_t)} \leq \frac{1}{\beta}, \quad \beta \in (0, 1),$$

where the first inequality follows by (8.1). It follows that $d_2(\beta) \leq 1/\beta$, which proves Eq. (7.2).

References

Aumann, R.J., Kurz, M.: Power and taxes. Econometrica **45**, 1137–1161 (1977)

Banks, J., Duggan, J.: A bargaining model of collective choice. Am. Polit. Sci. Rev. **94**, 73–88 (2000)

Battaglini, M., Palfrey, T.R.: The dynamics of distributive politics. Econ. Theory **49**, 739–777 (2012)

Bolton, G.E., Brandts, J., Ockenfels, A.: Fair procedures: evidence from games involving lotteries. Econ. J. **115**, 1054–1076 (2005)

Cardona, D., Ponsatí, C.: Bargaining one-dimensional social choices. J. Econ. Theory **137**, 627–651 (2007)

Cardona, D., Ponsatí, C.: Uniqueness of stationary equilibria in bargaining one-dimensional policies under (super) majority rules. Games Econ. Behav. **73**, 65–75 (2011)

Cho, S.J., Duggan, J.: Uniqueness of stationary equilibria in a one-dimensional model of bargaining. J. Econ. Theory **113**, 118–130 (2003)

Fudenberg, D., Tirole, J.: Game theory. MIT Press, Cambridge, Massachusetts (1991)

Herings, P.J.J., Predtetchinski, A.: One-dimensional bargaining with Markov recognition probabilities. J. Econ. Theory **145**, 189–215 (2010)

Imai, H., Salonen, H.: The representative Nash solution for two-sided bargaining problems. Math. Soc. Sci. **39**, 349–365 (2000)

Kalandrakis, T.: Regularity of pure strategy equilibrium points in a class of bargaining games. Econ. Theory **28**, 309–329 (2006)

Mankiw, N.G., Weinzierl, M., Yagan, D.: Optimal taxation in theory and practice. J. Econ. Perspect. **23**, 147–174 (2009)

Mirrlees, J.A.: An exploration in the theory of optimum income taxation. Rev. Econ. Stud. **38**, 175–208 (1971)

Predtetchinski, A.: One-dimensional bargaining. Games Econ. Behav. **72**, 526–543 (2011)

Roth, A.E.: Risk aversion and the relationship between Nash's solution and subgame perfect equilibrium of sequential bargaining. J. Risk Uncertain. **2**, 353–365 (1989)