Voting Over Selfishly Optimal Nonlinear Income Tax Schedules

Craig Brett[a], **John A. Weymark**[b]

[a]*Department of Economics, Mount Allison University, 144 Main Street, Sackville NB E4L 1A7, Canada. E-mail:* `cbrett@mta.ca`

[b]*Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA. E-mail:* `john.weymark@vanderbilt.edu`

September 2015

Abstract. Majority voting over the nonlinear income tax schedules proposed by a continuum of individuals with different labor productivities who have quasilinear-in-consumption preferences is considered. Each individual proposes the tax schedule that is selfishly optimal for him. Röell (unpublished manuscript, 2012) has identified some of the qualitative properties of these schedules. She has also shown that the individual preferences over these schedules are single-peaked and so Black's Median Voter Theorem applies. In this article, it is shown that each of the selfishly optimal schedules (i) has a region of bunching in a neighborhood of the proposer's skill type, (ii) coincides with the maxi-max schedule below this region, and (ii) coincides with the maxi-min schedule above it. Techniques introduced by Vincent and Mason (1967, NASA Contractor Report CR-744) are used to show that the bunching region can be identified by solving a simple unconstrained optimization problem. The characterization of the selfishly optimal schedules is used to provide a relatively simple proof that individuals have single-peaked preferences over them. In the majority rule equilibrium, marginal tax rates are negative on the maxi-max part of the schedule and positive on the maxi-min part except at the endpoints of the skill distribution where they are zero.

Journal of Economic Literature classification numbers. D72, D82, H21

Keywords. bunching, ironing, majority voting, nonlinear income taxation, redistributive taxation

---

1. Introduction

An alternative is a Condorcet winner if it does as well as any other alternative under consideration in a pairwise majority vote. Black's Median Voter Theorem (Black, 1948) shows that if individuals have single-peaked preferences, then any most-preferred alternative of a voter with a median preference peak is a Condorcet winner.[1] See, for example, Austen-Smith and Banks (1999) and Persson and Tabellini (2000). A problem that arises when applying Black's Median Voter Theorem to redistributive income tax policy is that individual preferences over feasible tax schedules may fail to be single peaked. For example, Itsumi (1974) and Romer (1975) show that even if the tax schedules are restricted to be linear and to satisfy the government's budget constraint, quite restrictive assumptions on the preferences and skill distribution are needed to ensure single-peakedness.

Using a finite type version of the Mirrlees (1971) model of nonlinear income taxation as in Guesnerie and Seade (1982), Röell (2012) proves that single-peakedness obtains if voting is restricted to the feasible tax schedules that are selfishly optimal for some individual when preferences are quasilinear in consumption.[2] Thus, when voting is restricted in this way, the tax schedule proposed by an individual with the median skill level is the Condorcet winner. Röell also identifies some of the qualitative properties of the selfishly optimal tax schedules without requiring quasilinearity.

While Röell has identified some of the properties of the selfishly optimal tax schedules, she has not provided a complete characterization of them. Using a continuum-of-types version of her problem, we show that it is possible to provide a *complete* characterization of these tax schedules in the quasilinear case and that this characterization can be used to provide a relatively simple proof that individuals have single-peaked preferences over them. Specifically, we show that each of the selfishly optimal schedules (i) has a region of bunching in a neighborhood of the proposer's skill type, (ii) coincides with the maxi-max schedule below this region, and (ii) coincides with the maxi-min schedule above it. Moreover, using techniques introduced by Vincent and Mason (1967, 1968), we show that the endpoints of this bunching region can be identified by solving a simple unconstrained optimization problem.

As in Röell (2012), we also show that in the majority rule equilibrium, marginal tax rates are negative on the maxi-max part of the schedule and positive on the maxi-min part except at the endpoints of the skill distribution where they are zero.. This finding provides support for Director's Law (see Stigler, 1970), which identifies a tendency for democratic governments to redistribute from both the poor and the rich toward the middle class. It also provides some support for why effective marginal tax rates in the United States are negative for low incomes and positive for higher incomes (Congressional

[1] When we refer to preferences as being single peaked, we employ the weak definition of single-peakedness of Austen-Smith and Banks (1999, p. 98). In this definition, an individual may have more than one most-preferred alternative (a plateau) and preferences need only be weakly decreasing as one moves farther from this plateau in either direction.

[2] An earlier version of Röell's article was written in 1996.

Budget Office, 2012).

The extensive literature on redistributive income taxation that builds on the seminal work of Mirrlees (1971) has primarily been normative. In the Mirrlees model, everybody has the same preferences for consumption and labor supply, but they differ in skill levels (their "types") as measured by their labor productivities. While the distribution of these productivities is common knowledge, the value of any individual's productivity is only known to himself. The government chooses a nonlinear income tax schedule to maximize a social welfare function subject to the constraints that (i) each individual optimally chooses his consumption and labor supply given the tax schedule and (ii) the resulting allocation satisfies the government's budget constraint.

In addition to replacing Mirrlees' welfare-maximizing approach with one based on voting, Röell (2012) supplements his two constraints with an additional constraint that guarantees each person a minimum utility. Imposing the minimal-utility constraint limits the degree to which an individual's selfishly optimal tax schedule can exploit low-skilled individuals in order to further his own interests.[3] In the absence of this constraint, it is optimal for any individual to chose a tax schedule for which the adjacent upward incentive constraints bind for all individuals with lower skill levels than his own. In order to show that individual preferences over the selfishly optimal tax schedules are single-peaked when the minimal-utility constraint is also accounted for, Röell assumes that this constraint is slack enough so that this pattern of binding incentive constraints is optimal. As a consequence, the minimum-utility constraint plays a limited role in her analysis. In fact, the qualitative properties of the selfishly optimal tax schedules are the same regardless of whether this constraint is considered. The inclusion of the minimum-utility constraint greatly complicates the analysis, so in order to present our results as simply as possible, we do not impose it. In Brett and Weymark (2015), we show how the selfishly optimal tax schedules must be modified if the minimum-utility constraint is also taken into account. This way of proceeding allows us to precisely identify the role that is played by the minimum-utility constraint.

Single-peakedness of the individual preferences over the set of income tax schedules being voted on is a sufficient condition for the existence of a Condorcet winner; it is not necessary. Provided that individual preferences for consumption and income satisfy the standard single-crossing property introduced by Mirrlees (1971), Gans and Smart (1996) show that if any two of the tax schedules under consideration only cross once, then pairwise majority voting generates a quasitransitive "social" preference on the set of these tax schedules.[4] When there are a finite number of tax schedules being voted on, quasitransitivity is sufficient for the existence of a Condorcet winner.[5] Because there

[3]Consumption must be nonnegative, so even the low-skilled individuals can afford to pay their income taxes. Even with a positive revenue requirement, a minimal-utility constraint simply limits the amount of redistribution that can take place.

[4]A weak preference relation is quasitransitive if the strict preferences are transitive

[5]The Gans–Smart result generalizes related results in Roberts (1977) for linear income taxes. The single-crossing property of linear income tax schedules is illustrated in Austen-Smith and Banks (1999, pp. 113–115) and Persson and Tabellini (2000, pp. 118–121) using specific functional forms. As Gans

is a continuum of tax schedules in our problem, demonstrating that the tax schedules satisfy the Gans–Smart single-crossing condition would not be sufficient to establish the existence of a Condorcet winner.

A different kind of single-crossing property has been used by Bohn and Stuart (2013).[6] Using a continuum version of the Röell model in which preferences are not required to be quasilinear, they have investigated voting over selfishly optimal income tax schedules and shown that the schedule proposed by the median skill type is a Condorcet winner when the minimum-utility constraint is satisfied. They do not appeal to Black's Median Voter Theorem. Associated with each each selfishly optimal income tax schedule is a utility curve that specifies the utility obtained with that schedule as a function of skill type. Bonn and Stuart show that any pair of these utility curves cross only once, which is a single-crossing property first investigated by Matthews and Moore (1987) in the context of a general adverse selection problem.[7] Bonn and Stuart's median voter result follows from this observation. They do not provide a complete characterization of the selfishly optimal tax schedules. Moreover, their analysis is quite technical.[8] By assuming that individual preferences are quasilinear in consumption and by not considering a minimum-utility constraint, we are not only able to provide a complete characterization of the selfishly optimal tax schedules, we are able to do so using elementary calculus.

In addition to Röell (2012) and Bohn and Stuart (2013), voting over selfishly optimal tax schedules has also been analyzed by Meltzer and Richard (1981), Snyder and Kramer (1988), and De Donder and Hindricks (2003). Meltzer and Richard (1981) consider voting rules for which a decisive individual (e.g., a dictator or a median voter) choses his most-preferred feasible linear income tax schedule. Snyder and Kramer (1988) investigate majority voting over selfishly optimal nonlinear income tax schedules in a model in which individuals allocate a fixed amount of labor between the taxable and underground sectors. De Donder and Hindricks (2003) use simulations to investigate the existence of a Condorcet winner among the set of selfishly optimal quadratic income tax schedules.[9]

and Smart (1996) note, the single-crossing tax schedule condition is equivalent to the schedules being completely ordered in terms of their progressivity and to the requirement that the individuals choose incomes that are nondecreasing in the skill level regardless of what tax schedule they face, a property that Roberts (1977) calls Hierarchical Adherence. Berliant and Gouveia (2001) have developed sufficient conditions for single-crossing income tax schedules in a model in which the set of skill types is a finite sample from a known distribution, the government's revenue requirement depends on the realized distribution, and voting takes place before the voters know what distribution is realized.

[6]Bohn and Stuart (2013) is a revised version of Bohn and Stuart (2002), and so both their analysis and that of Röell predate this article by many years.

[7]Unlike the Mirrlees (1971) single-crossing condition for preferences, the Matthews–Moore single-crossing condition does not require there to be only two goods.

[8]They make extensive use of an assumption about the curvature properties of an optimized value function. It is not clear what restrictions this assumption imposes on the primitives of the model.

[9]Other approaches to voting over nonlinear income tax schedules have also been considered. Blomquist and Christiansen (1999), Chen (2000), Roemer (2012), and Bierbrauer and Boyer (2013) suppose that candidates have some form of vote maximizing objective. Ledyard (2006, Sec. 3.6) briefly describes a model in which two candidates propose levels of public goods and nonlinear income tax

Our characterization of the selfishly optimal tax schedule for an individual with skill level k is indirect. We identify the income schedule that specifies how much before-tax income each skill type would receive with the selfishly optimal income tax schedule that is proposed by someone of this type. With quasilinear-in-consumption preferences, incentive compatibility of an allocation is preserved if everybody's income is changed by a common amount. As a consequence, once type k's optimal income schedule has been determined, the corresponding schedule showing how consumption varies with the skill level is easily computed using the government budget constraint (which must bind). The income tax an individual pays is the difference between his consumption and before-tax income.

An individual with the lowest skill type proposes the maxi-min income schedule, whereas one with the highest skill type proposes the maxi-max income schedule. We prove that the maxi-max schedule lies everywhere above the maxi-min schedule. For an individual of any type other than the lowest and highest, we show that he proposes an income schedule that (i) coincides with the maxi-max schedule for the lower part of the skill distribution, (ii) coincides with the maxi-min schedule for the upper part of the skill distribution, and (iii) "bridges" these two segments with a region of bunching that contains this individual's type.[10] The endpoints of this bunching region are nondecreasing in the proposer's type, from which it follows that everybody has a single-peaked preference over the proposed tax schedules, and so Black's Median Voter Theorem applies.

In order to obtain the characterization of the selfishly optimal income schedule of an individual of type k, we proceed by first considering a relaxed version of his optimization problem in which we ignore the second-order incentive-compatibility condition that income is nondecreasing in the skill level (Lollivier and Rochet, 1983). This is the first-order approach to this individual's problem. If the maxi-min and maxi-max income schedules are strictly increasing in the skill level, the solution to this relaxed problem is easy to describe. An individual of type k wishes to redistribute income from all other types toward his own type. To do this, for types greater than his own, he optimally employs the maxi-min income schedule, whereas for types smaller than his own, he optimally employs the maxi-max income schedule. In the maxi-min case, someone with skill level k and a maxi-min utility social planner both wish to extract as much revenue as incentives

schedules knowing that it is costly to vote so that not everybody votes. (Ledyard notes that his analysis is based on unpublished work with Marcus Berliant.) Voting has also been used to help determine non-linear tax schedules in dynamic models of taxation by Acemoglu et al. (2008) and Farhi et al. (2012). Acemoglu et al. (2008) hold politicians accountable by allowing them to be voted out of office. Farhi et al. (2012) assume that a political candidate wants to maximize his chances of being elected with the winner being determined using a probabilistic voting model. In this respect, their model is a dynamic extension of the probabilistic voting model considered by Chen (2000), but with the added possibility of taxing capital.

[10] For types sufficiently close to the lowest, it is possible that his optimal schedule starts on the bridge, in which case it provides a case of bunching at the bottom similar to that studied by Ebert (1992). If the distribution of types is bounded above, it is possible to have bunching at the top as well. This corresponds to having the income schedule ending on the bridge.

allow from the higher types. Because type k's optimal before-tax income schedule does not depend on the distribution of consumption, his desire to give that revenue to himself rather than to the least skilled is of no consequence for the specification of this part of the income schedule. The optimality of using the maxi-max solution for the rest of the skill distribution follows from similar reasoning.

Because the maxi-max schedule lies above the maxi-min schedule, the solution to type k's relaxed problem exhibits a downward discontinuity at his own skill level, and so violates the second-order monotonicity condition for incentive compatibility. To obtain the solution when this constraint is taken into account requires "ironing" the schedule described above by introducing a level bridge that connects the maxi-max and maxi-min components of the relaxed solution. The standard way of showing that it is optimal to iron a non-monotone schedule and to identify the endpoints of a bunching interval is to use the kind of control-theoretic techniques described in Guesnerie and Laffont (1984). We instead provide an elementary proof that it is optimal to link the maxi-max and maxi-min schedules with a bridge and employ the procedure developed by Vincent and Mason (1967, 1968) for smoothing discontinuous control trajectories. Applied to type k's problem, solving for the bridge endpoints using the Vincent–Mason approach is a simple unconstrained optimization problem.

In the maxi-min part of type k's income schedule, redistribution is constrained by downward incentive-compatibility constraints and gives rise to the familiar positive (or zero for the highest type) marginal income tax rates. Similarly, in the maxi-max part of this schedule, redistribution is constrained by upward incentive-compatibility constraints, which gives rise to negative (or zero for the lowest type) marginal income tax rates. As a consequence, there must be a kink in the optimal income tax schedule at the income chosen by type k and bunching of some of the types near him.

The rest of this article is organized as follows. Section 2 introduces the model of the economy. Section 3 contains a detailed account of each skill type's choice of a selfishly optimal income tax schedule. This is followed in Section 4 by an analysis of the voting equilibrium. Section 5 provides concluding remarks. The Appendix provides the proofs omitted from the main text.

2. The Model

The economy is populated by individuals that differ in labor productivity. Differences in skills are described by a parameter w which is continuously distributed with support $[\underline{w}, \bar{w}]$, density function $f(w) > 0$, and cumulative distribution function $F(w)$. It is assumed that $0 < \underline{w} < \bar{w}$. An individual with skill level w produces w units of a consumption good per unit of labor time in a perfectly competitive labor market and earns a (before-tax) income of

$$y = wl, \tag{1}$$

where l is the nonnegative amount of labor supplied.[11] Thus, w is this type's wage rate. Income can also be thought of as being labor in efficiency units.

An individual has nonnegative consumption x, which is also his after-tax income. Preferences over consumption and labor supply are represented by the quasilinear-in-consumption utility function

$$\tilde{u}(l, x) = x - h(l), \tag{2}$$

which is common to all individuals. The function h is increasing, strictly convex, and three-times continuously differentiable. The government can observe an individual's before- and after-tax incomes, but not his skill level or labor supply. Using (1), the utility function in terms of observable variables is

$$u(y, x; w) = x - h\left(\frac{y}{w}\right). \tag{3}$$

In terms of consumption and income, the marginal rate of substitution at any bundle (y, x) is decreasing in w when $y > 0$, so the standard Mirrlees (1971) single-crossing condition for preferences is satisfied.

Individuals face an anonymous nonlinear income taxation schedule that specifies the tax paid as a function of income $T(y)$, subject to which individuals choose their most preferred combination of consumption and before-tax income (equivalently, after-tax income and labor supply). Admissible tax schedules are assumed to be integrable. The Taxation Principle (see Hammond, 1979; Guesnerie, 1995) shows that having individuals choose consumption and income subject to an anonymous tax schedule is equivalent to directly specifying these variables as functions of type subject to incentive-compatibility constraints.[12] These schedules, $x(\cdot)$ and $y(\cdot)$, as well as the labor supply schedule $l(\cdot)$ corresponding to $y(\cdot)$, are also integrable.[13] The bundle allocated to individuals of type w is $(y(w), x(w))$. The resulting utility level is

$$V(w) = x(w) - h\left(\frac{y(w)}{w}\right), \quad \forall w \in [\underline{w}, \bar{w}]. \tag{4}$$

Incentive compatibility requires that

$$V(w) = \max_{w'} \; x(w') - h\left(\frac{y(w')}{w}\right), \quad \forall w, w' \in [\underline{w}, \bar{w}]. \tag{5}$$

[11] For simplicity, we do not explicitly consider an upper bound on labor supply. Weak assumptions on preferences can be introduced that ensure that any such bound is not binding.

[12] A direct mechanism is *simple* if each individual's allocation is a function of only his own reported type with the same function used for all individuals. By assuming that everybody faces the same tax schedule (as is the standard practice in the optimal income tax literature), the Taxation Principle implies that we can restrict attention to simple direct mechanisms. In contrast, with a general mechanism design approach, the restriction to a direct mechanism is justified by the Revelation Principle. A direct tax mechanism need not be simple as it could base an individual's tax on the reports of other individuals, not just his own. Bierbrauer (2011) has shown that if the individual types are random variables drawn from a known distribution, then a utilitarian optimal mechanism is simple.

[13] The stronger assumption of piecewise continuous differentiability is typically made to facilitate the use of standard control-theoretic arguments. See Hellwig (2010).

Because the Mirrlees single-crossing property is satisfied, it follows from Mirrlees (1976) that the first-order (envelope) condition for incentive compatibility is

$$V'(w) = h'\left(\frac{y(w)}{w}\right)\frac{y(w)}{w^2}, \quad \forall w \in [\underline{w}, \bar{w}], \tag{6}$$

and the second-order condition is

$$y'(w) \geq 0, \quad \forall w \in [\underline{w}, \bar{w}].^{14} \tag{7}$$

Consumption must also be nondecreasing in type. Moreover, the single-crossing property and incentive compatibility imply that two types either (i) differ in both income and consumption or (ii) have the same bundle, in which case they are said to be *bunched* (see Laffont and Martimort, 2002, sec. 3.1). Because h is increasing, (6) implies that utility is nondecreasing in w whenever incentive compatibility is satisfied and it is strictly increasing for all w for which $y(w) > 0$.

The income tax schedule must be differentiable almost everywhere. At any income for which it is not differentiable, the marginal tax rate $\tau(w)$ is not well-defined. At incomes for which it is well-defined as the derivative of the tax schedule, $\tau(w)$ is equal to one minus the marginal rate of substitution between consumption and income (i.e., between after-tax and before-tax incomes). As is standard, this expression can be used to define an implicit marginal tax rate for values of y for which $T(y)$ is not differentiable. Thus,

$$\tau(w) = 1 - h'\left(\frac{y(w)}{w}\right)\frac{1}{w}, \quad \forall w \in [\underline{w}, \bar{w}]. \tag{8}$$

Because utility is quasilinear in consumption, marginal tax rates do not depend on consumption.

The only purpose of taxation is to redistribute income, so the government budget constraint is

$$\int_{\underline{w}}^{\bar{w}} [y(w) - x(w)] f(w)\, dw \geq 0. \tag{9}$$

The qualitative features of our analysis are unaffected if the government instead requires a fixed positive amount of revenue.

3. Selfishly Optimal Income Tax Schedules

Each individual proposes an income tax schedule and then pairwise majority rule is used to choose which of these schedules is implemented. A tax schedule can only be proposed if the resulting allocation satisfies the incentive-compatibility conditions (6) and (7) and

[14] The expressions in (6) and (7) are required to hold at all points for which $y(w)$ is differentiable. Because incentive compatibility implies that income is nondecreasing in type, $y(w)$ is differentiable almost everywhere.

the government budget constraint (9). These conditions characterize the *feasible income tax schedules*. We suppose that the tax schedule proposed by an individual is his selfishly optimal tax schedule. That is, it is the feasible schedule that maximizes his own utility. Individuals of the same type propose the same tax schedule.

Rather than thinking of an individual as proposing an income tax schedule, it is more convenient to appeal to the Taxation Principle and think of him as choosing an allocation schedule $(y(\cdot), x(\cdot))$ that specifies a bundle $(x(w), y(w))$ for each type $w \in [\underline{w}, \bar{w}]$.[15] Formally, an individual of type k (the "proposer") determines his optimal allocation schedule by solving

$$\max_{x(\cdot),y(\cdot)} V(k) \quad \text{subject to } (4), (6), (7), \text{and } (9). \tag{10}$$

We refer to (10) as *type k's problem*.

Two characteristics of type k's problem distinguish it from the standard Mirrlees (1971) problem: the form of the objective function and the explicit inclusion of the second-order incentive-compatibility constraint. Mirrlees used a utilitarian objective function, whereas here the utility of a particular type of individual is being maximized When this type is $\underline{w}$, the objective is simply the maxi-min criterion, which has been studied in detail by Boadway and Jacquet (2008). For reasons of tractability, Mirrlees and most subsequent authors only considered the first-order conditions for incentive compatibility, what is known as the *first-order approach*. The second-order incentive-compatibility conditions have been explicitly taken into account by Brito and Oakland (1977) and Ebert (1992). The complete solution to type k's problem can be determined from the solution to the relaxed problem in which the second-order monotonicity constraint (7) is ignored, so we begin by considering it. An analysis of the relaxed problem also yields useful insights into the nature of type k's optimization problem.

3.1. The First-Order Approach

The relaxed problem in which the monotonicity constraint on before-tax income is suppressed provides a picture of how the form of the objective function in type k's problem helps to create a solution to the optimal tax problem that is very different from those found in other optimal nonlinear income tax problems. Formally, this problem is

$$\max_{x(\cdot),y(\cdot)} V(k) \quad \text{subject to } (4), (6), \text{and } (9). \tag{11}$$

We refer to (11) as *type k's relaxed problem*.

By modifying the arguments found in Lollivier and Rochet (1983), we show in Proposition 1 that it is possible to formulate an unconstrained optimization problem that provides the before-tax income schedule that solves type k's relaxed problem. As is

[15] In this section, the identity of the skill type proposing the schedule is fixed, so we do not index the schedules by his type.

standard in a nonlinear income tax problem, it is optimal for the government budget constraint to bind.[16] Once type k's optimal before-tax income schedule has been determined, the corresponding consumption schedule (and, hence, the income tax schedule) can be derived using the incentive-compatibility and binding government budget constraints. With quasilinear-in-consumption utility, the relevant properties of the optimal bundles for each type can be inferred from the before-tax income schedule, so we do not consider the consumption schedule that solves type k's relaxed problem explicitly.

Proposition 1. *For any $k \in [\underline{w}, \bar{w}]$, the optimal schedule of before-tax incomes $y(\cdot)$ for type k's relaxed problem is obtained by solving*

$$\begin{aligned}\max_{y(\cdot)} \int_{\underline{w}}^{k} &\left\{ \left[(y(w) - h\left(\frac{y(w)}{w}\right) \right] f(w) + \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) F(w) \right\} dw \\ + \int_{k}^{\bar{w}} &\left\{ \left[(y(w) - h\left(\frac{y(w)}{w}\right) \right] f(w) - \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) [1 - F(w)] \right\} dw.\end{aligned} \tag{12}$$

For ease of exposition, we suppose for now that the solutions to (12) for $k = \underline{w}$ and $k = \bar{w}$ are strictly increasing in w (so there is no bunching) and, hence, that both of these solutions satisfy the monotonicity constraint (7). Later, we shall relax this assumption. Thus, when $k = \underline{w}$, the solution to (12) is the *maxi-min income schedule*, which we denote by $y^R(\cdot)$, and when $k = \bar{w}$, the solution is the *maxi-max income schedule*, which we denote by $y^M(\cdot)$.[17] From (12), we see that the income schedule that solves type k's relaxed problem coincides with the maxi-max solution for individuals with skill types smaller than that of the proposer and coincides with the maxi-min solution for individuals with skill types larger than that of the proposer.

Not only is the optimization problem (12) unconstrained, it can be solved point-wise. Thus, simple differentiation with respect to $y(w)$ provides the first-order conditions for type k's relaxed problem, which we write in the implicit form as

$$\begin{aligned}\theta^M(w, y(w)) = 0, &\quad \forall w \in [\underline{w}, k), \\ \theta^R(w, y(w)) = 0, &\quad \forall w \in (k, \bar{w}],\end{aligned} \tag{13}$$

where

$$\theta^M(w, y) = \left[1 - h'\left(\frac{y}{w}\right) \frac{1}{w} \right] f(w) + \left[h''\left(\frac{y}{w}\right) \frac{y}{w^3} + h'\left(\frac{y}{w}\right) \frac{1}{w^2} \right] F(w) \tag{14}$$

and

$$\theta^R(w, y) = \left[1 - h'\left(\frac{y}{w}\right) \frac{1}{w} \right] f(w) - \left[h''\left(\frac{y}{w}\right) \frac{y}{w^3} + h'\left(\frac{y}{w}\right) \frac{1}{w^2} \right] [1 - F(w)]. \tag{15}$$

[16] If the budget constraint does not bind, because preferences are quasilinear in consumption, each person's consumption can be increased by a common small amount without violating incentive compatibility, thereby increasing the utility of type k individuals.

[17] It is commonplace to call the maxi-min objective "Rawlsian" even though Rawls (1971) used an index of primary goods rather than utility in his criterion. Our notation reflects this usage.

For $k = \underline{w}$ (resp. $k = \bar{w}$), the second (resp. first) first-order condition in (13) applies for all $w \in [\underline{w}, \bar{w}]$. For $k \neq \underline{w}, \bar{w}$, there is an indeterminacy in the solution to the relaxed problem at $w = k$ and the value of the integrand we maximize in (12) is the same no matter how we resolve it.[18]

Using (8), (13), and (15), the optimal maxi-min marginal tax rates are

$$\tau^R(y(w)) = \frac{1 - F(w)}{f(w)} \left[h''\left(\frac{y(w)}{w}\right) \frac{y(w)}{w^3} + h'\left(\frac{y(w)}{w}\right) \frac{1}{w} \right], \quad \forall w \in [\underline{w}, \bar{w}].^{19} \tag{16}$$

Hence, the marginal tax rate is zero for the highest skilled and positive for all other types with the maxi-min objective. In the absence of the incentive constraints, personalized lump-sum taxes would be used for redistribution. Given our quasilinearity assumption, it then follows that compared to the full-information benchmark, every type except for the highest has his income and labor supply distorted downwards, whereas the highest skilled have the same income and labor supply as in the benchmark case. This pattern of distortions coincides with those found using a utilitarian objective function except that it is optimal in the utilitarian case for the lowest skilled to face a zero marginal tax rate provided that it is not bunched with any other type (see Sadka, 1976; Seade, 1977).

From (8), (13), and (14), the optimal maxi-max marginal tax rates are

$$\tau^M(y(w)) = -\frac{F(w)}{f(w)} \left[h''\left(\frac{y(w)}{w}\right) \frac{y(w)}{w^3} + h'\left(\frac{y(w)}{w}\right) \frac{1}{w} \right], \quad \forall w \in [\underline{w}, \bar{w}]. \tag{17}$$

Therefore, with the maxi-max objective, the marginal tax rate is zero for the lowest skilled and negative for all other types. Compared to the full-information benchmark, all types except the lowest skilled (who are not distorted) have their incomes and labor supply distorted upwards.

Using these observations, some intuition can be provided for the first-order conditions (13) for type k's relaxed problem. For ease of exposition, it is useful to think of types as being discrete, but with skill levels arbitrarily close to each other. Type k wishes to maximize the utility of individuals of his own type. The function $\theta^M(w, y)$ captures the additional consumption (hence, utility) that individuals of this type can gain by increasing $y(w)$ by one unit for some $w < k$. At the solution, this value must be zero. In the first instance, increasing $y(w)$ by one unit makes available $f(w)$ extra units of consumption that can be diverted to the type k individuals. But appropriate adjustments must also be made in order to ensure that incentive compatibility is re-established after this increase. Type k wishes to redistribute resources away from lower types towards his own type. Individuals of lower types are distorted upwards, so this type of redistribution is constrained by upward incentive compatibility conditions that prevent individuals of

[18] We write all first-order conditions for the optimal incomes as equalities, thereby implicitly assuming that the nonnegativity constraints on incomes are not binding. The qualitative features of our analysis are unaffected if these constraints are taken into account.

[19] Boadway and Jacquet (2008, eqn. (21), p. 435) state this condition in an implicit form.

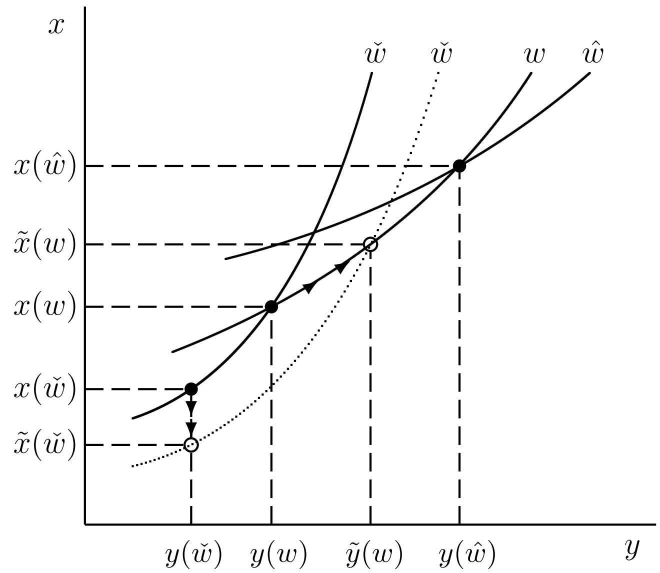

Figure 1: Adjustments to restore incentive compatibility following an increase in $y(w)$ when $w < k$

lower types from mimicking types above them. Thus, any increase in $y(w)$ for a $w < k$ must be accompanied by adjustments that ensure that the upward incentive constraints are satisfied. These adjustments are illustrated in Figure 1.

First, each individual of type w can be given $h'(\frac{y(w)}{w})\frac{1}{w}$ additional units of consumption to place him on his initial indifference curve, thereby ensuring that he has no incentive to mimic any other type. This is shown by the adjustment from $(y(w), x(w))$ to $(\tilde{y}(w), \tilde{x}(w))$ in Figure 1. Moreover, this change does not affect the incentives of any types above w.[20] These units of consumption must be subtracted from the $f(w)$ units that can be diverted to type k individuals. This accounts for the second term in the first bracket in (14). Moving individuals of type w upward along their indifference curves in this way slackens the upward incentive constraint for the type $\breve{w}$ immediately below w. Because preferences are quasilinear in consumption, by reducing the consumption of everybody whose type is smaller than w by the amount in the final bracket in (14) restores incentive compatibility. This is illustrated by the adjustment from $(y(\breve{w}), x(\breve{w}))$ to $(y(\breve{w}), \tilde{x}(\breve{w}))$ for type $\breve{w}$ in Figure 1. There are $F(w)$ individuals whose types are smaller than w, so the second term in (14) is the total amount of consumption that type k individuals can re-claim from these types in this way.

Type k also wishes to move resources away from types higher than himself towards individuals of his own type. These types are downward distorted, so this kind of re-

[20]Recall that we are assuming that the maxi-min and maxi-max solutions exhibit no bunching, so $y(w) \neq y(\hat{w})$, where $\hat{w}$ is the next highest type above w. We are implicitly assuming that units of income are sufficiently small so that $\tilde{y}(w) < y(\hat{w})$.

distribution is constrained by downward incentive compatibility constraints. The function $\theta^R(w, y)$ shows the additional consumption that type k individuals can secure for themselves through a one unit increase in $y(w)$ for some $w > k$. The only difference between $\theta^R(w, y)$ and $\theta^M(w, y)$ is in the final term. This difference arises because it is the downward incentive constraints that bind for types above k. Moving individuals of type $w > k$ upward along their indifference curves in the manner described in the preceding paragraph would lead to a violation of the downward incentive constraint for the next highest type. Because preferences are quasilinear in consumption, satisfaction of these constraints can be re-established by giving these individuals and everyone of a higher type more consumption in the amount given in the final bracket in (15). There are $1 - F(w)$ such individuals. Because this consumption must be given to individuals of types different from that of the proposer, these resources are subtracted from the amount available to the type k individuals.

As we have noted, with the maxi-min income schedule, everyone has his income distorted downward compared to the full-information solution except for the highest type who is undistorted, whereas with the maxi-max income schedule, everyone has his income distorted upward compared to the full-information solution except for the lowest type who is undistorted. Thus, the maxi-max schedule lies everywhere above the maxi-min schedule. As a consequence, for any type $k \neq \underline{w}, \bar{w}$, his optimal income schedule has a downward discontinuity at his skill type, as illustrated in Figure 2.[21] We summarize our main findings in Proposition 2.

Proposition 2. *For any $k \in (\underline{w}, \bar{w})$, the optimal schedule of before-tax incomes $y(\cdot)$ for type k's relaxed problem is given by*

$$y(w) = \begin{cases} y^M(w), & \forall w \in [\underline{w}, k), \\ y^R(w), & \forall w \in (k, \bar{w}]. \end{cases} \tag{18}$$

There is a downward discontinuity in this schedule at $w = k$. For $k = \underline{w}, \bar{w}$, the optimal schedules are $y^R(\cdot)$ and $y^M(\cdot)$, respectively.

Because the solution to type k's relaxed problem for $k \neq \underline{w}, \bar{w}$ features a jump from the maxi-max to the maxi-min tax schedule at skill level k, there is a discontinuity in the associated marginal tax rates, which are given by

$$\tau(w) = \begin{cases} \tau^M(w), & \forall w \in [\underline{w}, k), \\ \tau^R(w), & \forall w \in (k, \bar{w}]. \end{cases} \tag{19}$$

We thus have a switch from negative marginal tax rates for types just below type k to positive marginal tax rates for types just above this type.

The discontinuities in the income schedule and in the marginal tax rates are intertwined. As we move from types just below type k to types just above it, the upward

[21] There is also a downward discontinuity in type k's optimal consumption schedule at his type.

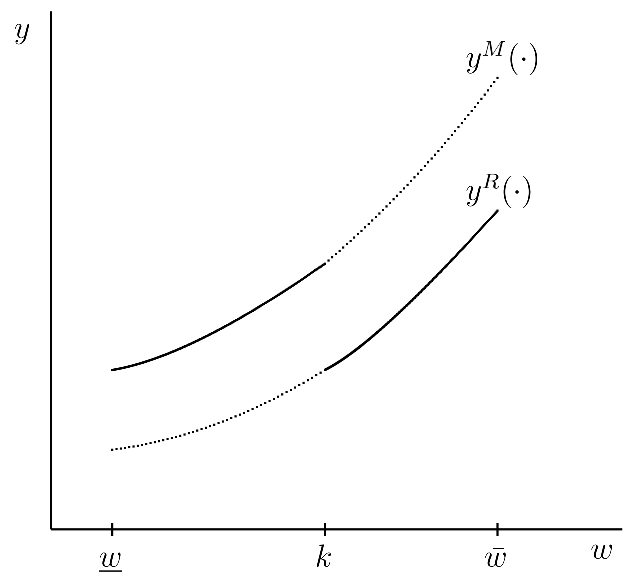

Figure 2: The optimal income schedule for type k's relaxed problem

distortions in incomes switch to downward distortions and the signs of the marginal tax rates change from negative to positive. Because the maxi-max income schedule lies strictly above the maxi-min schedule at $w = k$, it is impossible to reconcile these competing distortions without a downward jump in the income schedule and a change in sign in the marginal tax rates at k.

The downward jump in the solution to type k's relaxed problem for the income schedule clearly violates the second-order incentive compatibility conditions. So even if the maxi-min and maxi-max income schedules do satisfy these second-order conditions, we have not found a solution to type k's problem in (10). Nevertheless, as we show below, some elements of these solutions feature in the complete solution to his problem.

3.2. The Complete Solution

If either the maxi-min or maxi-max income schedule obtained using the first-order approach fails to satisfy the second-order incentive-compatibility condition (7) (i.e., the requirement that the schedule must be nondecreasing), then it is necessary to bunch all types in a decreasing part of the schedule with some types who are in an increasing part, what is known as *ironing*. Any bunching region must be a closed interval. Its endpoints can be determined using the approach described by Guesnerie and Laffont (1984).[22] Because ironing in this kind of situation is well understood and we do not need to know

[22]The first analysis of ironing in economics appears to have been by Arrow (1968). Arrow was concerned with devising an optimal capital policy with irreversible investment. The irreversability of investment imposes a monotonicity constraint analogous to the one on incomes found here.

where the endpoints of these bunching regions are for our results, we shall simply suppose that these schedules have been ironed. We let $y^{R*}(\cdot)$ and $y^{M*}(\cdot)$ denote the optimal maxi-min and maxi-max income schedules when the second-order incentive-compatibility constraint has been taken into account.[23]

Once the bunching regions for $y^{R*}(\cdot)$ and $y^{M*}(\cdot)$ have been determined, it is straightforward to modify the objective function (12) in type k's relaxed problem for $k = \underline{w}, \bar{w}$ so as to take account of the second-order incentive-compatibility condition (7). Doing so will facilitate the analysis of the other types' problems. Let $\mathcal{B}^M$ and $\mathcal{B}^R$ denote the types that are bunched with some other type in the complete solution to the maxi-max and maxi-min problems, respectively. When w is bunched, we let $[w_-, w_+]$ denote the set of types bunched with w.

In the maxi-max case, only the first integral in (12) applies. Its integrand is replaced by $G^{M*}(w, y(w))$, where

$$G^{M*}(w, y(w)) = \begin{cases} \left[(y(w) - h\left(\frac{y(w)}{w}\right)\right] f(w) + \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) F(w), & \forall w \notin \mathcal{B}^M, \\ \left[(y(w) - h\left(\frac{y(w)}{w}\right)\right] \left[\int_{w_-}^{w_+} f(t)dt\right] + \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) F(w_-), & \forall w \in \mathcal{B}^M. \end{cases} \tag{20}$$

Similarly, in the maxi-min case, only the second integral in (12) applies. Its integrand is replaced by $G^{R*}(w, y)(w)$, where

$$G^{R*}(w, y(w)) = \begin{cases} \left[(y(w) - h\left(\frac{y(w)}{w}\right)\right] f(w) - \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) [1 - F(w)], & \forall w \notin \mathcal{B}^R, \\ \left[(y(w) - h\left(\frac{y(w)}{w}\right)\right] \left[\int_{w_-}^{w_+} f(t)dt\right] - \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) [1 - F(w_+)], & \forall w \in \mathcal{B}^R. \end{cases} \tag{21}$$

Ironing does not affect the solution outside a bunching region, so no modifications to the integrands in (12) are needed for types that are not bunched. The intuition for these expressions when there is bunching is similar to that provided above for type k's relaxed problem. Now, if an extra unit of consumption is given to type w individuals, it must be given to all individuals who are bunched with them, whose mass is $\int_{w_-}^{w_+} f(t)dt$. For this reason, the $f(w)$ that appears in the first cases of both (20) and (21) is replaced by this integral in the second cases. When w is bunched, in the maxi-max case, some of this extra consumption can be reclaimed from individuals of lower type than those bunched with w, whose mass is $F(w_-)$. The corresponding individuals in the maxi-min case are those individuals of higher type than those bunched with w, whose mass is $[1 - F(w_+)]$. In the second cases of (20) and (21), these expressions are used to replace the $F(w)$ and $[1 - F(w)]$ that appear in the first cases.

[23] In an earlier version of our article (Working Paper No. 14-00010, Department of Economics, Vanderbilt University), we showed that only relatively mild assumptions are needed to ensure that the maxi-min income schedule is increasing without the need for ironing, but that much more stringent conditions are needed to ensure that this is the case for the maxi-max schedule.

Before characterizing the optimal income schedule that type k would propose, we first show that this schedule is continuous.[24]

Proposition 3. *For any $k \in [\underline{w}, \bar{w}]$, the optimal schedule of before-tax incomes $y^*(\cdot)$ for type k's problem is continuous.*

Proof. Type k's problem is obtained from his relaxed problem by adding the constraint (7) that income is nondecreasing in type, and this implies that there can be no downward jumps in $y^*(\cdot)$. Also note that any set of types bunched with each other is an interval.

(a) We first show that $y^*(\cdot)$ is continuous at $\bar{k} > k$. Suppose that it is not. Thus, there is an upward jump in $y^*(\cdot)$ at $\bar{k}$. There are two cases.

Case 1. Type $\bar{k}$ is not bunched with anybody of higher skill. Single-crossing and satisfaction of the self-selection constraints imply that everybody with a type larger than $\bar{k}$ strictly prefers his bundle to that of types no larger than $\bar{k}$ because the set of types more skilled than $\bar{k}$ is an open set. Because preferences are quasilinear in consumption, it is possible to reduce the consumption of everybody more skilled than $\bar{k}$ by a small positive amount without violating the self-selection constraints. If the total amount of consumption saved is now shared equally by everybody, the government budget constraint again binds and, by quasilinearity, the self-selection constraints are still satisfied. Hence, the resulting allocation is feasible. Type k only cares about his own utility and has been made better off. Thus, the original schedule could not be optimal for type k. Hence, type k's optimal schedule must be continuous at $\bar{k}$.

Case 2. There is a higher skilled type bunched with $\bar{k}$. If there is also a lower skilled type bunched with $\bar{k}$, then $y^*(\cdot)$ is constant at $\bar{k}$ and, hence, continuous at this type. If no lower skilled type is bunched with $\bar{k}$, then everybody with a type at least $\bar{k}$ strictly prefers his bundle to that of types smaller than $\bar{k}$. The argument is the same as in Case 1 except that type $\bar{k}$ is now included in the types whose consumptions are initially reduced.

(b) We now show that $y^*(\cdot)$ is continuous at $\bar{k} < k$. Suppose that it is not. If the consumption of $\underline{w}$ is positive, then everybody has positive consumption. Hence, by first adjusting consumptions for the low skilled relative to $\bar{k}$, arguments analogous to those used in (a) show that type k's optimal schedule must be continuous at $\bar{k}$.

If the consumption of $\underline{w}$ is zero, this argument does not apply because it is not possible to reduce this type's consumption or that of anybody bunched with this type. In this case, instead of first reducing consumptions for the low skilled relative to $\bar{k}$, their incomes are instead increased. Because preferences are not quasilinear in income, the amount of this increase needs to be type specific so as not to violate the self-selection constraints. There is now additional consumption that can be distributed as in (a) so as to make type k better off. Hence, type k's optimal schedule must also be continuous at $\bar{k}$ for $\bar{k} < k$.

(c) To show that $y^*(\cdot)$ is continuous, it remains to show that it is continuous at k. If type k is bunched with any higher type, $y^*(\cdot)$ is locally constant on an interval of the form $[k, \bar{k})$ and, hence, is continuous from the right at k. If type k is not bunched with any higher type, then the argument in Case 1 of (a) with the initial adjustments made

[24] It follows from this result that $T^*(\cdot)$, $x^*(\cdot)$ and $\ell^*(\cdot)$ are also continuous.

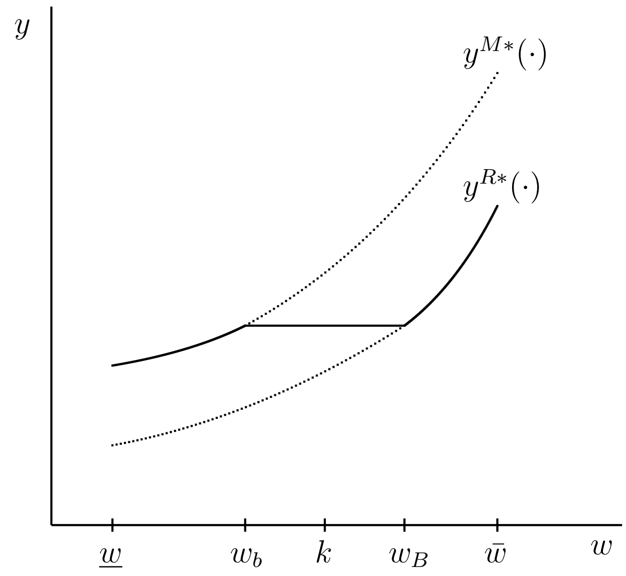

Figure 3: A bridge

to everybody more skilled than k shows that $y(\cdot)$ is continuous from the right at k. An analogous argument using (b) shows that $y^*(\cdot)$ is continuous from the left at k. $\square$

Using $y^{R*}(\cdot)$ and $y^{M*}(\cdot)$ instead of $y^R(\cdot)$ and $y^M(\cdot)$ in (18) for $k \neq \underline{w}, \bar{w}$, we obtain the income schedule for type k's relaxed problem that must be modified in order to determine the complete solution to his problem. The only decreasing part of this schedule is the downward discontinuity at k's type. We now show that it is optimal for k to build a *bridge* that includes his own type between the maxi-max and maxi-min parts of this schedule, as illustrated in Figure 3. That is, it is optimal to iron the relaxed solution by adding a single new bunching interval of the form $[w_b, w_B]$ with $w_b < w_B$ that includes k. It is possible that $w_b = \underline{w}$, in which case the optimal income schedule starts with the bridge and then tracks the maxi-min solution. It is also possible that $w_B = \bar{w}$, in which case the optimal income schedule first tracks the maxi-max solution and then ends with the bridge.[25]

Proposition 4. *For any $k \in (\underline{w}, \bar{w})$, the optimal schedule of before-tax incomes $y^*(\cdot)$ for*

[25]It is conceivable that $w_b = \underline{w}$ and $w_B = \bar{w}$, in which case the bridge is the whole income schedule. This possibility is so unlikely that we do not consider it explicitly when we identify the optimal endpoints.

type k's problem is given by

$$y^*(w) = \begin{cases} y^{M*}(w), & \forall w \in [\underline{w}, w_b), \\ y^{M*}(w_b), & \forall w \in [w_b, w_B] \text{ if } w_b > \underline{w}, \\ y^{R*}(w_B), & \forall w \in [w_b, w_B] \text{ if } w_B < \bar{w}, \\ y^{R*}(w), & \forall w \in (w_B, \bar{w}], \end{cases} \tag{22}$$

for some $w_b, w_B \in [\underline{w}, \bar{w}]$ for which $w_b < w_B$ and $k \in [w_b, w_B]$.

Proof. (a) Suppose that there exists a type $\bar{k} > k$ for which $y^*(\bar{k})$ is not the maximin income. The monotonicity constraint (7) must therefore bind at $\bar{k}$, so the slope of $y^*(\cdot)$ is zero at $\bar{k}$. Because $y^*(\cdot)$ is continuous by Propostion 3, this implies that there exists a $w_B > \bar{k}$ such that $y^*(\cdot)$ is (i) constant on $[k, w_B]$ and (ii) coincides with the maximin schedule on $[w_B, \bar{w}]$.

(b) If there exists a type $\bar{k} < k$ for which $y^*(\bar{k})$ is not the maximax income, then the argument used in (a) shows that there exists a $w_b < \bar{k}$ such that $y^*(\cdot)$ is (i) constant on $[w_b, k]$ and (ii) coincides with the maximax schedule on $[\underline{w}, k)$. □

The shape of type k's optimal income schedule $y^*(\cdot)$ is given by (22). As we have already noted, for types smaller than the lower endpoint of the bridge, type k's optimal income schedule coincides with the maxi-max income schedule $y^{M*}(\cdot)$, whereas for types larger than the upper endpoint of the bridge, it coincides with the maxi-min income schedule $y^{R*}(\cdot)$. Provided that $w_B < \bar{w}$, the income for skill type w_B is $y^{R*}(w_B)$. Consequently, all individuals on the bridge receive this income. Analogously, if $w_b > \underline{w}$, then all individuals on the bridge receive $y^{M*}(w_b)$. If both $w_b > \underline{w}$ and $w_B < \bar{w}$, then $y^{M*}(w_b) = y^{R*}(w_B)$.

The values of the bunching interval endpoints w_b and w_B are determined optimally so as to minimize the loss in type k's utility that results from deviating from his relaxed solution. To identify these endpoints, we employ a procedure that was introduced by Vincent and Mason (1967, 1968) to smooth discontinuous control trajectories. First, the optimal schedule is selected for each fixed pair of values of the bridge endpoints w_b and w_B. Then, among these schedules, the one that maximizes type k's utility is selected. This is a simple unconstrained optimization problem.[26]

The derivatives of $G^{M*}(w, y)$ and $G^{R*}(w, y)$ with respect to income are denoted by $\theta^{M*}(w, y)$ and $\theta^{R*}(w, y)$, respectively. The first-order conditions that characterize the optimal bridge endpoints in the solution to type k's problem are given in Proposition 5.

[26] Guesnerie and Laffont (1984) use control theory to show that it is optimal to iron and to identify the endpoints of bunching regions. Our approach is much simpler and does not use control theory. Moreover, we cannot use the Guesnerie–Laffont proof strategy because their differentiability assumptions are not satisfied here.

Proposition 5. *For any $k \in (\underline{w}, \bar{w})$, the optimal values of the bridge endpoints w_b and w_B are determined by the first-order condition*

$$\int_{w_b}^{k} \theta^{M*}(w, y^{M*}(w_b))dw + \int_{k}^{w_B} \theta^{R*}(w, y^{M*}(w_b))dw = 0 \tag{23}$$

if $w_b > \underline{w}$ and by the first-order condition

$$\int_{w_b}^{k} \theta^{M*}(w, y^{R*}(w_B))dw + \int_{k}^{w_B} \theta^{R*}(w, y^{R*}(w_B))dw = 0 \tag{24}$$

if $w_B < \bar{w}$.

When both $w_b > \underline{w}$ and $w_B < \bar{w}$ hold, (23) and (24) are equivalent conditions. These two equations are similar to the standard ironing condition found in Guesnerie and Laffont (1984, eqn. (3.16), p. 347). Taken together, Propositions 4 and 5 provide a complete characterization of type k's selfishly optimal income schedule.

4. The Political Equilibrium

Majority rule is used to determine the income tax schedule that is implemented. As we observed in the previous section, we can equivalently think of voting as taking place over the types of individuals, with a type k candidate implementing his selfishly optimal income tax schedule if elected.[27] The advantage of this way of formulating the problem is that a type is simply a skill level, and so voting takes place over a one-dimensional issue space. We show that individual preferences over the types are single-peaked with respect to the skill level. It then follows from Black's Median Voter Theorem (Black, 1948) that there exists a Condorcet winner and it is the type with the median skill level.[28] Thus, the selfishly optimal tax schedule for an individual with the median skill level (weakly) beats the selfishly optimal tax schedule for any other type of individual in a pairwise majority vote.

It is now necessary to distinguish allocation schedules by the types that propose them. Let $(x^*(w,k), y^*(w,k))$ denote the optimal allocation assigned to an individual of type w by type k's selfishly optimal tax schedule. When $k = \bar{w}$, $y^*(\cdot, k)$ is the maxi-max income schedule $y^{M*}(\cdot)$ and when $k = \underline{w}$, it is the maxi-min income schedule $y^{R*}(\cdot)$. The utility obtained by an individual with skill level w with the schedule proposed by type k is

$$V(w,k) = x^*(w,k) - h\left(\frac{y^*(w,k)}{w}\right). \tag{25}$$

The bridge in type k's selfishly optimal income schedule is now denoted by $[w_b(k), w_B(k)]$.

[27] Bohn and Stuart (2013) require any individual who is indifferent between two types to either vote for only one of them or to abstain from voting on this pair.

[28] Our assumption that $f(w) > 0$ for all $w \in [\underline{w}, \bar{w}]$ implies that there is a unique median skill level.

While it is only necessary to introduce a bridge between the maxi-max and maxi-min income schedules for types other than $\underline{w}$ and $\bar{w}$, it will simplify the statement of our next result if we define bridges for these two types as well. These bridges are defined to be the set of types that are bunched with the relevant endpoint of the skill distribution. Thus, $w_b(\underline{w}) = \underline{w}$ and $w_B(\bar{w}) = \bar{w}$. In either case, there need not be any bunching, in which case the endpoints of the bridge are the same. In contrast, a bridge must be a nondegenerate interval for any other type.

One determinant of an individual's preferences for the schedules proposed by the different skill types is the amount of before-tax income he will obtain with them. As a preliminary to identifying how an individual's income varies with the type of the proposer, we first show that the endpoints of a bridge are nondecreasing in the proposer's type.

Proposition 6. *The bridge endpoints $w_b(k)$ and $w_B(k)$ are nondecreasing in k for all $k \in [\underline{w}, \bar{w}]$.*

Consider two types k_1 and k_2 with $k_1 < k_2$. Because (i) their proposed income schedules coincide with the maxi-max schedule for types below the lower endpoint of the bridge and coincide with the maxi-min schedule for types above the upper endpoint of the bridge and (ii) the endpoints of a bridge are nondecreasing in the type of the proposer, any individual's income is nondecreasing in the proposer's type.

Proposition 7. *For all $w, k_1, k_2 \in [\underline{w}, \bar{w}]$ for which $k_1 < k_2$, $y^*(w, k_1) \leq y^*(w, k_2)$.*[29]

Some intuition for the comparative static results described in Propositions 6 and 7 may be obtained by considering the income distortions for different types. For individuals with skill types below the lower endpoint of a bridge, their incomes are determined by the maxi-max schedule. These incomes are distorted upwards relative to the full-information solution. The higher the proposer's type, the more types there are below him, and so it is beneficial for him to distort the incomes (and labor supplies) of more of these types upwards. In contrast, for individuals with skill types above the upper endpoint of a bridge, their incomes are determined by the maxi-min schedule and, so, are distorted downwards. In this case, the higher the proposer's type, the fewer the types there are above him, and so it is beneficial for him to distort the incomes (and labor supplies) of fewer of these types downwards.

All proposers face the same government budget and incentive constraints. As a consequence, because each type proposes a feasible income schedule that is best for him, he must weakly prefer what he obtains with his own schedule to what anybody else proposes for him. Formally,

$$V(w, w) \geq V(w, k), \quad \forall w, k \in [\underline{w}, \bar{w}]. \tag{26}$$

Thus, each individual's preferences on the set of skill types has a peak at his own type. An individual of type w has a (weakly) *single-peaked preference* on the set of types if

$$V(w, w) \geq V(w, k_1) \geq V(w, k_2) \text{ if } w < k_1 < k_2 \tag{27}$$

[29] The second inequality is in fact strict for all $w \in (w_b(k_1), w_B(k_2))$ if $w_b(k_1) < w_b(k_2)$ and $w_B(k_1) < w_B(k_2)$.

and

$$V(w,w) \geq V(w,k_1) \geq V(w,k_2) \text{ if } w > k_1 > k_2. \tag{28}$$

We show in Proposition 8 that everybody's preferences on the set of types are single-peaked.

Proposition 8. *Individual preferences are single-peaked on the set of skill types.*

Whereas Proposition 7 only considers how an individual's income varies as the proposer's type changes, Proposition 8 compares how his utility varies, and so takes account of both income and consumption changes. Even if two types of proposer offer an individual the same income, it does not follow that they offer the same consumption as well. For example, suppose that $w < k_1 < k_2$ and that $y^*(w,k_1) = y^*(w,k_2)$ because w is on the maxi-max part of the income schedules proposed by both types k_1 and k_2. Further suppose that both endpoints of the bridge are larger with the k_2 schedule than with the k_1 schedule. The latter assumption implies that aggregate before-tax income increases when k_2's schedule is used instead of k_1's and, therefore, there is an additional amount of consumption that can be distributed. Suppose that individual incomes are first adjusted to the levels stipulated by type k_2's income schedule holding consumptions at the levels for type k_1's consumption schedule. Because preferences are quasilinear in consumption, if this allocation satisfies the incentive-compatibility constraints, it is then optimal to distribute the surplus consumption to everybody equally as this will preserve incentive compatibility. However, it would then be the case that a type w individual would prefer type k_2 to type k_1 because he receives more consumption with k_2's proposal than with that of k_1 without any change in income, contradicting Proposition 8. Thus, the first stage of this adjustment procedure violates incentive compatibility. In moving to type k_2's schedule, the individuals with skill types in $(w_b(k_1), w_b(k_2))$ have their labor supplies further upward distorted. To restore incentive compatibility, the consumption of everybody of lower skill, including individuals of type w, must have their consumptions reduced even after the additional consumption generated by the increase in the aggregate income has been distributed.[30]

When preferences over selfishly optimal income tax schedules are single-peaked, by Black's Median Voter Theorem, the tax schedule proposed by the median skill type does at least as well as any other proposed tax schedule in a majority vote. Hence, the following proposition is a direct consequence of Proposition 8 and Black's Theorem.[31]

[30]If $w > k_2 > k_1$, $y^*(w,k_1) = y^*(w,k_2)$ because w is on the maxi-min part of both income schedules, and the bridge endpoints shift as above, then fewer types have their labor supplies downward distorted in moving from type k_1's schedule to that of type k_2. For this reason, some of the newly generated consumption is allocated to type w individuals to help restore incentive compatibility.

[31]Black (1948) assumes that each voter has a unique most-preferred alternative. The reasoning used to establish his Median Voter Theorem extends to the weakly single-peaked preferences considered here. We conjecture that each skill type has a unique seflishly optimal income tax schedule, but that in some circumstances an interval of types may prefer the same schedule.

Proposition 9. *The selfishly optimal income tax schedule for the median skill type is a Condorcet winner when majority voting is restricted to the income tax schedules that are selfishly optimal for some skill type.*

The main insight of Röell (2012) is that by restricting attention to the nonlinear optimal income tax schedules that are selfishly optimal from some individual's perspective, there is a Condorcet winner provided that preferences are quasilinear in consumption and there is a finite number of skill types. Proposition 9 confirms that this insight extends to an economy with a continuum of skills. In the typical case in which this schedule has three segments, it features marginal wage subsidies for individuals with incomes below the bridge except for the least skilled who are undistorted, a bunching region that includes the median type, and positive marginal tax rates for individuals with incomes above the bridge except for the most highly skilled who are undistorted. In particular, there must be a kink in the tax schedule at the income of an individual with the median skill.

Given the complexity of the tax problem being considered, it is striking that a Condorcet winner exists. Indeed, if voting were over all feasible nonlinear income tax schedules, the schedule identified in Proposition 9 would not be a Condorcet winner. To see why, consider the nonlinear budget set associated with the selfishly optimal income tax schedule for the median skill type. Choose a bundle $(\bar{y}, \bar{x})$ on the upper boundary of this set so that only a small group of types near the top of the skill distribution choose bundles with incomes larger than $\bar{y}$ and so that the median skill type chooses a bundle with an income smaller than it. (Recall that both income and consumption are nondecreasing in type.) Introduce a kink in the budget set at $(\bar{y}, \bar{x})$ and pivot the upper boundary of the budget set down slightly for incomes above $\bar{y}$. This can be done so that only the types with initial incomes greater than $\bar{y}$ choose a new bundle, so that these new bundles are close to the original ones, and so that the government budget constraint no longer binds. This group is made worse off, but they are a small group. Next, pivot the upper boundary of the budget set up slightly for incomes below $\bar{y}$. All types with initial incomes less than $\bar{y}$ are made better off and they constitute a majority of the population. By making the change in the budget set sufficiently small in this second step, the new bundles are close to the original ones and the budget surplus generated in the first stage is not completely dissipated, so the change is feasible. Hence, the tax schedule corresponding to the final budget set would defeat the median selfishly optimal schedule in a pairwise majority contest. By Proposition 9, this new tax schedule is not selfishly optimal for anyone.

5. Conclusion

We have fully characterized the feasible income tax schedule that is selfishly optimal for each skill type when preferences are quasilinear in consumption. The desire of an individual to redistribute resources towards himself both from those who are less skilled and from those who are more skilled has been shown to result in a downward discontinuity in the selfishly optimal income schedule identified using the first-order approach, which is a novel kind of violation of a second-order incentive-compatibility constraint in the optimal income tax literature. Moreover, the upward redistribution from the less skilled results in a non-standard pattern of distortions reminiscent of those found in the screening literature with participation constraints and countervailing incentives, as studied in detail by Jullien (2000). Nevertheless, we have shown that the selfishly optimal income (and allocation) schedules can be quite simply described.

Our comparative static result describing how the selfishly optimal income schedules vary with the type of proposer is easy to state in terms of bridge endpoints. This simple structure is what gives rise to single-peaked preferences over the types of proposer (and their tax schedules), from which our median voter result follows. Our comparative static result can also be used to analyze the equilibria of other voting rules, such as various forms of plurality rule, or as a building block in understanding the positions that different constituencies might hold and advocate in more institutionally rich models of political decision-making.

Our proof strategy for showing that the selfishly optimal incomes schedules are continuous and the tools we use to identify the endpoints of a bridge are of potential interest in the study of other screening problems in which the first-order approach results in discontinuities in the solution that violate the second-order incentive-compatibility conditions. For example, a policy design problem in which some group is explicitly excluded from the planner's objective but is part of the set of contributors to the policy's finances may well give rise to such a discontinuity.

In Brett and Weymark (2015), we show that our analysis can be modified to take account of the minimum-utility constraint employed by Röell (2012) and Bohn and Stuart (2013) without fundamentally changing the basic structure of a type's selfishly optimal income schedule. The minimum-utility constraint induces a proposer to behave as if he is maximizing a weighted average of his own utility and the utility of the least skilled. For types above his own, this modified objective has no effect on this proposer's optimal income schedule because he still wants to redistribute resources downward from these types. For types below him, however, the weighted averaging pulls the income schedule toward the maxi-min schedule. Moreover, the shadow value of the minimum-utility constraint increases with the proposer's type, so that relatively high-skilled proposers optimally choose income schedules that are the most shifted away from what they would choose in the absence of the minimum-utility constraint.

While the income tax schedule implemented by a median voter in our model differs in a number of respects from the schedules that are observed in practice, it does exhibit redistribution from the endpoints of the income distribution towards the middle, and so is consistent with Director's Law. Exploring whether this feature of income tax policy is preserved when the model is enriched to allow for labor market decisions at the extensive margin or for other policy instruments, such as workfare, is a natural topic for future research.

Acknowledgements

A previous version of this article was circulated as "Citizen Candidates and Voting over Incentive-Compatible Nonlinear Income Tax Schedules." Craig Brett's research has been generously supported by the Canada Research Chair Program. We are grateful to Siddharth Kothari for his research assistance. We have benefited from comments from and discussions with Niels Anthonisen, Marcus Berliant, Henning Bohn, Pierre Boyer, Marc Fleurbaey, Thomas Groll, Guy Laroque, Emanuel Hansen, and Laurent Simula, as well as from the comments of our referees and the Editor, Tom Palfrey. We have also benefited from comments received at various conferences and seminars.

Appendix

Proof of Proposition 1. By (6),

$$V(w) = V(\underline{w}) + \int_{\underline{w}}^{w} \frac{y(t)}{t^2} h'\left(\frac{y(t)}{t}\right) dt. \tag{A.1}$$

Integrating (A.1) over the support of the distribution of types yields

$$\int_{\underline{w}}^{\bar{w}} V(w) f(w) dw = \int_{\underline{w}}^{\bar{w}} V(\underline{w}) f(w) dw + \int_{\underline{w}}^{\bar{w}} \int_{\underline{w}}^{w} \frac{y(t)}{t^2} h'\left(\frac{y(t)}{t}\right) f(w) dt dw. \tag{A.2}$$

Reversing the order of integration in (A.2), we obtain

$$\begin{aligned} \int_{\underline{w}}^{\bar{w}} V(w) f(w) dw &= V(\underline{w}) + \int_{\underline{w}}^{\bar{w}} \frac{y(t)}{t^2} h'\left(\frac{y(t)}{t}\right) \left[\int_{t}^{\bar{w}} f(w) dw\right] dt \\ &= V(\underline{w}) + \int_{\underline{w}}^{\bar{w}} \frac{y(t)}{t^2} h'\left(\frac{y(t)}{t}\right) [1 - F(t)]\, dt. \end{aligned} \tag{A.3}$$

On the other hand, by (4),

$$\int_{\underline{w}}^{\bar{w}} V(w) f(w) dw = \int_{\underline{w}}^{\bar{w}} x(w) f(w) dw - \int_{\underline{w}}^{\bar{w}} h\left(\frac{y(w)}{w}\right) f(w) dw. \tag{A.4}$$

As previously noted, it is optimal for the government budget constraint (9) to bind. Substituting the equality form of (9) into (A.4) yields

$$\int_{\underline{w}}^{\bar{w}} V(w) f(w) dw = \int_{\underline{w}}^{\bar{w}} y(w) f(w) dw - \int_{\underline{w}}^{\bar{w}} h\left(\frac{y(w)}{w}\right) f(w) dw. \tag{A.5}$$

Combining (A.3) and (A.5) implies that

$$\begin{aligned} V(\underline{w}) = &\int_{\underline{w}}^{\bar{w}} y(w) f(w) dw - \int_{\underline{w}}^{\bar{w}} h\left(\frac{y(w)}{w}\right) f(w) dw \\ &- \int_{\underline{w}}^{\bar{w}} \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) [1 - F(w)]\, dw. \end{aligned} \tag{A.6}$$

The maximand in (12) is obtained by substituting (A.6) into (A.1) and setting $w = k$. The preceding calculations have accounted for all the constraints in (11), and so the proof is complete. $\square$

Proof of Proposition 4. First, fix the bridge endpoints w_b and w_B and let $y^*(w_b, w_B)$ denote the optimal income on the bridge $[w_b, w_B]$. A bridge cannot begin in the interior of a bunching interval of $y^{M*}(\cdot)$, nor can it end in the interior of a bunching interval of $y^{R*}(\cdot)$, so it is supposed in the rest of this proof that w_b and w_B satisfy these restrictions. The rest of type k's optimal income schedule is obtained by solving

$$\begin{aligned} \max_{y(\cdot)} \left[\int_{\underline{w}}^{w_b} G^{M*}(w, y(w))dw + \int_{w_B}^{\bar{w}} G^{R*}(w, y(w))dw \right] \\ \text{subject to } y(w) = y^*(w_b, w_B), \ \forall w \in [w_b, w_B]. \end{aligned} \tag{A.7}$$

This problem can be solved point-wise. Its solution is given implicitly by the first-order conditions

$$\begin{aligned} \theta^{M*}(w, y(w)) = 0, &\quad \forall w \in [\underline{w}, w_b), \\ \theta^{R*}(w, y(w)) = 0, &\quad \forall w \in (w_B, \bar{w}]. \end{aligned} \tag{A.8}$$

Thus,

$$y^*(w) = y^{M*}(w), \quad \forall w \in [\underline{w}, w_b), \tag{A.9}$$

and

$$y^*(w) = y^{R*}(w), \quad \forall w \in (w_B, \bar{w}]. \tag{A.10}$$

We now need to determine the optimal endpoints of the bridge and the optimal income on the bridge. There are two cases to consider.

Case 1: $w_B < \bar{w}$. By the continuity of $y^*(\cdot)$, $y^*(w_B) = y^{R*}(w_B)$. Hence, $y^*(w_B) = y^*(w_b)$ because income is a constant on the bridge. If $w_b > \underline{w}$, then by continuity, $y^*(w_b) = y^{M*}(w_b)$. Let

$$\phi(w_B) = \begin{cases} {y^{M*}}^{-1}(y^{R*}(w_B)), & \text{if } w_b > \underline{w}, \\ w_b, & \text{if } w_b = \underline{w}. \end{cases} \tag{A.11}$$

Because $y^{M*}(\cdot)$ is continuous and integrable, so is $\phi(\cdot)$. Type k's choice of w_B is therefore the solution to

$$\begin{aligned} \max_{w_B} B(w_B, k) \equiv \int_{\underline{w}}^{\phi(w_B)} G^{M*}(w, y^{M*}(w))dw + \int_{\phi(w_B)}^{k} G^{M*}(w, y^{R*}(w_B))dw \\ + \int_{k}^{w_B} G^{R*}(w, y^{R*}(w_B))dw + \int_{w_B}^{\bar{w}} G^{R*}(w, y^{R*}(w))dw, \end{aligned} \tag{A.12}$$

where use has been made of (A.9) and (A.10) in (A.12). Setting the derivative of $B(w, k)$ with respect to w_B equal to zero, we obtain the first-order condition

$$
\begin{aligned}
&G^{M*}(w_b, y^{M*}(w_b))\frac{d\phi(w_B)}{dw_B} - G^{M*}(w_b, y^{M*}(w_b))\frac{d\phi(w_B)}{dw_B} \\
&\quad + \int_{w_b}^{k} \theta^{M*}(w, y^{R*}(w_B))\frac{dy^{R*}(w_B)}{dw_B}dw + \int_{k}^{w_B} \theta^{R*}(w, y^{R*}(w_B))\frac{dy^{R*}(w_B)}{dw_B}dw \\
&\qquad + G^{R*}(y^{R*}(w_b, y(w_B)) - G^{R*}(y^{R*}(w_b, y(w_B)) = 0.
\end{aligned} \tag{A.13}
$$

Simplifying (A.13) yields

$$
\frac{dy^{R*}(w_B)}{dw_B}\left[\int_{w_b}^{k} \theta^{M*}(w, y^{R*}(w_B))dw + \int_{k}^{w_B} \theta^{R*}(w, y^{R*}(w_B))dw\right] = 0. \tag{A.14}
$$

Because a bridge cannot end in the middle of a bunching region of y^{R*}, it must be the case that the derivative in this expression is positive. Hence, (24) holds.

Case 2: $w_b > \underline{w}$. Continuity of the solution now implies that $y^*(w_b) = y^{M*}(w_b)$. Let

$$
\psi(w_b) = \begin{cases} y^{R*^{-1}}(y^{M*}(w_b)), & \text{if } w_B < \bar{w}, \\ w_B, & \text{if } w_B = \bar{w}. \end{cases} \tag{A.15}
$$

Type k's choice of w_b is the solution to

$$
\begin{aligned}
\max_{w_b} b(k, w_b) \equiv &\int_{\underline{w}}^{w_b} G^{M*}(w, y^{M*}(w))dw + \int_{w_b}^{k} G^{M*}(w, y^{M*}(w_b))dw \\
&+ \int_{k}^{\psi(w_b)} G^{R*}(w, y^{M*}(w_b))dw + \int_{\psi(w_b)}^{\bar{w}} G^{R*}(w, y^{R*}(w))dw,
\end{aligned} \tag{A.16}
$$

The analogue of (A.14) is

$$
\frac{dy^{M*}(w_b)}{dw_b}\left[\int_{w_b}^{k} \theta^{M*}(w, y^{M*}(w_b))dw + \int_{k}^{w_B} \theta^{R*}(w, y^{M*}(w_b))dw\right] = 0, \tag{A.17}
$$

from which (24) follows because a bridge cannot begin in the middle of a bunching region of y^{M*}. □

Proof of Proposition 6. We first show that $w_B(k)$ is nondecreasing in k. Consider any value of k for which $w_B(k) < \bar{w}$. Differentiating the left-hand side of (A.14) with respect to k, we obtain

$$
\begin{aligned}
\frac{\partial^2 B(w_B, k)}{\partial w_B \partial k} &= \frac{dy^{R*}(w_B)}{dw_B}\left[\theta^{M*}(k, y^{R*}(w_B)) - \theta^{R*}(k, y^{R*}(w_B))\right] \\
&= \frac{dy^{R*}(w_B)}{dw_B}\left[h''\left(\frac{y^{R*}(w_B)}{k}\right)\frac{y^{R*}(w_B)}{k^3} + h'\left(\frac{y^{R}(w_B)}{k}\right)\frac{1}{k^2}\right],
\end{aligned} \tag{A.18}
$$

where the second equality follows from (14) and (15). Because $h(\cdot)$ is a strictly convex function, the cross-partial derivative in (A.18) is positive. Hence, by Theorems 10.3 and 10.4 in Sundaram (1996), the function B is supermodular and therefore exhibits increasing differences in (w, k). Thus, by Topkis' Theorem (Topkis, 1978, Theorem 6.1), it follows that $w_B(k)$ is nondecreasing in k for values of k for which $w_B(k) < \bar{w}$.

If $w_B(k) = \bar{w}$, the proof proceeds as above except that $y^*(\bar{w})$ is used instead of $y^{R*}(w_B)$ in (A.11).

The proof that $w_b(k)$ is nondecreasing in k is almost the same as that for the other endpoint of the bridge. In this case, the left-hand side of (A.17) is differentiated with respect to k to obtain

$$\begin{aligned}\frac{\partial^2 b(w_b, k)}{\partial w_b \partial k} &= \frac{dy^{M*}(w_b)}{dw_b}\left[\theta^{M*}(k, y^{M*}(w_b)) - \theta^{R*}(k, y^{M*}(w_b))\right] \\ &= \frac{dy^{M*}(w_b)}{dw_b}\left[h''\left(\frac{y^{M*}(w_b)}{k}\right)\frac{y^{M*}(w_b)}{k^3} + h'\left(\frac{y^{M*}(w_b)}{k}\right)\frac{1}{k^2}\right],\end{aligned} \tag{A.19}$$

and $y^*(\underline{w})$ is used instead of $y^{M*}(w_b)$ in (A.15). □

Proof of Proposition 8. We only show (27) as the proof of (28) analogous. Consider three types w, k_1, and k_2 for which $w < k_1 < k_2$. The first inequality in (27) follows immediately from (26).

To show that the second inequality holds as well, first note that by (6),

$$V(w, k_1) = V(k_1, k_1) - \int_w^{k_1} h'\left(\frac{y^*(t, k_1)}{t}\right)\frac{y^*(t, k_1)}{t^2}dt. \tag{A.20}$$

Therefore, using (26),

$$V(w, k_1) \geq V(k_1, k_2) - \int_w^{k_1} h'\left(\frac{y^*(t, k_1)}{t}\right)\frac{y^*(t, k_1)}{t^2}dt. \tag{A.21}$$

It also follows from (6) that

$$V(w, k_2) = V(k_1, k_2) - \int_w^{k_1} h'\left(\frac{y^*(t, k_2)}{t}\right)\frac{y^*(t, k_2)}{t^2}dt. \tag{A.22}$$

Subtracting (A.22) from (A.21) yields

$$V(w, k_1) - V(w, k_2) = \int_w^{k_1}\left[h'\left(\frac{y^*(t, k_2)}{t}\right)\frac{y^*(t, k_2)}{t^2} - h'\left(\frac{y^*(t, k_1)}{t}\right)\frac{y^*(t, k_1)}{t^2}\right]dt. \tag{A.23}$$

By Proposition 7, $y^*(t, k_2) \geq y^*(t, k_1)$ for all t. Therefore, by the convexity of $h(\cdot)$, the expression in (A.23) is nonnegative, as was to be shown. □

References

Acemoglu, D., Golosov, M., Tsyvisnki, A., 2008. Markets versus governments. Journal of Monetary Economics 55, 159–189.

Arrow, K. J., 1968. Optimal capital policy with irreversible investment. In: Wolfe, J. N. (Ed.), Value, Capital, and Growth: Papers in Honour of Sir John Hicks. Edinburgh University Press, Edinburgh, pp. 1–19.

Austen-Smith, D., Banks, J. S., 1999. Positive Political Theory I: Collective Preference. University of Michigan Press, Ann Arbor.

Berliant, M., Gouveia, M., 2001. On the political economy of income taxation, unpublished manuscript, Department of Economics, Washington University at St. Louis and School of Business and Economics, Universidade Católica Portuguesa.

Bierbrauer, F. J., 2011. On the optimality of optimal income taxation. Journal of Economic Theory 146, 2105–2116.

Bierbrauer, F. J., Boyer, P., 2013. Political competition and Mirrleesian income taxation: A first pass. Journal of Public Economics 103, 1–14.

Black, D., 1948. On the rationale of group decision making. Journal of Political Economy 56, 23–34.

Blomquist, S., Christiansen, V., 1999. The political economy of publicly provided private goods. Journal of Public Economics 73, 31–54.

Boadway, R., Jacquet, L., 2008. Optimal marginal and average income taxation under maximin. Journal of Economic Theory 143, 425–441.

Bohn, H., Stuart, C., 2002. Voting over non-linear taxes in a stylized representative democracy, unpublished manuscript, Department of Economics, University of California, Santa Barbara.

Bohn, H., Stuart, C., 2013. Revenue extraction by median voters, unpublished manuscript, Department of Economics, University of California, Santa Barbara.

Brett, C., Weymark, J. A., 2015. Voting over nonlinear income tax schedules with a minimum-utility constraint, in preparation.

Brito, D. L., Oakland, W. H., 1977. Some properties of the optimal income-tax. International Economic Review 18, 407–423.

Chen, Y., 2000. Electoral systems, legislative process, and income taxation. Journal of Public Economic Theory 2, 71–100.

Congressional Budget Office, 2012. Effective marginal tax rates for low- and moderate-income workers. Publication No. 4149, Congressional Budget Office, Congress of the United States.

De Donder, P., Hindricks, J., 2003. The politics of progressive income taxation with incentive effects. Journal of Public Economics 87, 2491–2505.

Ebert, U., 1992. A reexamination of the optimal nonlinear income tax. Journal of Public Economics 49, 47–73.

Farhi, E., Sleet, C., Werning, I., Yeltekin, S., 2012. Non-linear capital taxation without commitment. Review of Economic Studies 79, 1469–1493.

Gans, J. S., Smart, M., 1996. Majority voting with single-crossing preferences. Journal of Public Economics 59, 219–237.

Guesnerie, R., 1995. A Contribution to the Pure Theory of Taxation. Cambridge University Press, Cambridge.

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

Guesnerie, R., Seade, J., 1982. Nonlinear pricing in a finite economy. Journal of Public Economics 17, 157–179.

Hammond, P. J., 1979. Straightforward incentive compatibility in large economies. Review of Economic Studies 46, 263–282.

Hellwig, M. F., 2010. Incentive problems with unidimensional characteristics: A unified approach. Econometrica 78, 1201–1237.

Itsumi, Y., 1974. Distributional effects of linear income tax schedules. Review of Economic Studies 41, 371–381.

Jullien, B., 2000. Participation constraints in adverse selection problems. Journal of Economic Theory 93, 1–47.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives: The Principle-Agent Problem. Princeton University Press, Princeton, NJ.

Ledyard, J., 2006. Voting and efficient public good mechanisms. In: Weingast, B. R., Wittman, D. A. (Eds.), The Oxford Handbook of Political Economy. Oxford University Press, Oxford, pp. 479–501.

Lollivier, S., Rochet, J.-C., 1983. Bunching and second-order conditions: A note on optimal tax theory. Journal of Economic Theory 31, 392–400.

Matthews, S., Moore, J., 1987. Monopoly provision of quality and warranties: An exploration in thetheory of multi-dimensional screening. Econometrica 55, 441–467.

Meltzer, A. H., Richard, S. F., 1981. A rational theory of the size of government. Journal of Political Economy 89, 914–927.

Mirrlees, J. A., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175–208.

Mirrlees, J. A., 1976. Optimum tax theory: A synthesis. Journal of Public Economics 6, 327–358.

Persson, T., Tabellini, G., 2000. Political Economics: Explaining Economic Policy. MIT Press, Cambridge, MA.

Rawls, J., 1971. A Theory of Justice. Harvard University Press, Cambridge, MA.

Roberts, K. W. S., 1977. Voting over income tax schedules. Journal of Public Economics 8, 329–340.

Röell, A., 2012. Voting over nonlinear income tax schedules, unpublished manuscript, School of International and Public Affairs, Columbia University. An earlier version was written in 1996.

Roemer, J. E., 2012. The political economy of income taxation under asymmetric information: The two-type case. SERIEs: Journal of the Spanish Economic Association 3,

181–199.
Romer, T., 1975. Individual welfare, majority voting, and the properties of a linear income tax. Journal of Public Economics 4, 163–185.
Sadka, E., 1976. On income distribution, incentive effects and optimal income taxation. Review of Economic Studies 43, 261–267.
Seade, J. K., 1977. On the shape of optimal tax schedules. Journal of Public Economics 7, 203–235.
Snyder, J. M., Kramer, G. H., 1988. Fairness, self-interest, and the politics of the progressive income tax. Journal of Public Economics 36, 197–230.
Stigler, G. J., 1970. Director's Law of public income redistribution. Journal of Law and Economics 13, 1–10.
Sundaram, R. K., 1996. A First Course in Optimization Theory. Cambridge University Press, Cambridge.
Topkis, D. M., 1978. Minimizing a submodular function on a lattice. Operations Research 26, 305–321.
Vincent, T. L., Mason, J. D., 1967. A new method for treating optimal trajectories with restricted segments. Contractor Report CR-744, NASA.
Vincent, T. L., Mason, J. D., 1968. Disconnected optimal trajectories. Journal of Optimization Theory and Applications 3, 263–281.