Vanderbilt University Department of Economics Working Paper Series
Nashville, TN 37235-1819

Vanderbilt University Department of Economics Working Papers 16-00005

Voting over Selfishly Optimal Nonlinear Income Tax Schedules with a Minimum-Utility Constraint

Craig Brett
Mount Allison University

John A Weymark
Vanderbilt University

Abstract

Pairwise majority voting over alternative nonlinear income tax schedules is considered when there is a continuum of individuals who differ in their labor productivities, which is private information, but share the same quasilinear-in-consumption preferences for labor and consumption. Voting is restricted to those schedules that are selfishly optimal for some individual. The analysis extends that of Brett and Weymark (Games and Economic Behavior, forthcoming) by adding a minimum-utility constraint to their incentive-compatibility and government budget constraints. It also extends the analysis of Röell (unpublished manuscript, 2012) and Bohn and Stuart (unpublished manuscript, 2013) by providing a complete characterization of the selfishly optimal tax schedules. It is shown that individuals have single-peaked preferences over the set of selfishly optimal tax schedules, and so the schedule proposed by the median skill type is a Condorcet winner.

Citation: Craig Brett and John A Weymark, (2016) "Voting over Selfishly Optimal Nonlinear Income Tax Schedules with a Minimum-Utility Constraint", *Vanderbilt University Department of Economics Working Papers*, VUECON-16-00005.
Contact: Craig Brett - cbrett@mta.ca, John A Weymark - john.weymark@vanderbilt.edu.
Submitted: February 11, 2016. **Published:** February 12, 2016.
URL:http://www.accessecon.com/Pubs/VUECON/VUECON-16-00005.pdf

Voting over Selfishly Optimal Nonlinear Income Tax Schedules with a Minimum-Utility Constraint

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`

February 2016

Abstract. Pairwise majority voting over alternative nonlinear income tax schedules is considered when there is a continuum of individuals who differ in their labor productivities, which is private information, but share the same quasilinear-in-consumption preferences for labor and consumption. Voting is restricted to those schedules that are selfishly optimal for some individual. The analysis extends that of Brett and Weymark (Games and Economic Behavior, forthcoming) by adding a minimum-utility constraint to their incentive-compatibility and government budget constraints. It also extends the analysis of Röell (unpublished manuscript, 2012) and Bohn and Stuart (unpublished manuscript, 2013) by providing a complete characterization of the selfishly optimal tax schedules. It is shown that individuals have single-peaked preferences over the set of selfishly optimal tax schedules, and so the schedule proposed by the median skill type is a Condorcet winner.

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

Keywords. Mirrlees tax problem; nonlinear income taxation; political economy of taxation; redistributive taxation; voting over tax schedules

---

1. Introduction

In democratic societies, tax policies are a prominent feature in political platforms and public forums. Some debates concern the level of taxation—on whether there should be tax increases or tax cuts. Others involve the distribution of tax payments. Is there a need for middle class tax relief? Should low income earners receive additional benefits? What is the appropriate tax rate for the highest earners? The diversity of issues that can be raised highlights the complexity of tax policy. Even fairly simple income tax schemes have several tax brackets, and there is a need to decide upon tax rates for each bracket and the positions of the brackets themselves. In short, tax policy is multidimensional. In contrast, economists' canonical model of the political determination of policies, the median voter model, works best in unidimensional settings.

In this article, we consider pairwise majority voting over alternative nonlinear income tax schedules when, as in Mirrlees (1971), there is a continuum of individuals who differ in their labor productivities, which is private information, but share the same preferences for labor and consumption.[1] A tax schedule is a Condorcet winner if a majority of voters weakly prefers it to any of the other tax schedules being considered. Given the potential complexity of a nonlinear income tax schedule, a Condorcet winner will only exist if some restrictions are placed on the set of permissible tax schedules. Here, we follow the lead of Röell (2012), Bohn and Stuart (2013), and our previous work, Brett and Weymark (2016), by restricting attention to selfishly optimal nonlinear income tax schedules. That is, voting is restricted to those nonlinear tax schedules that some individual would choose from among the feasible tax schedules if that person were a dictator.[2] What these schedules are depends on the feasibility constraints that are considered. In our companion article, the only constraints on a tax schedule are that it be incentive compatible and respect the government's budget constraint; the latter constraint is equivalent to the economy's material balance constraint. In this article, we further constrain the selfishly optimal schedules by requiring that they guarantee some minimum utility level to all individuals. This requirement prevents the adoption of tax schedules that extract excessive rents from the very poorest members of society.

Labor productivity is a unidimensional measure of an individual's skill. A selfishly optimal tax schedule depends on the skill level of the individual who proposes it. The set of tax schedules that are voted on consists of all of the schedules that are selfishly

[1] We do not explicitly model political competition. Bierbrauer and Boyer (2013) and Roemer (2012), among others, consider political candidates who compete for votes by proposing redistributive tax policies when labor productivities are privately known.

[2] Meltzer and Richard (1981) consider majority voting over selfishly optimal *linear* income tax schedules. Snyder and Kramer (1988) investigate majority voting over selfishly optimal nonlinear income tax schedules when 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. There is also an extensive literature that investigates the existence of a Condorcet winner when tax schedules that are not selfishly optimal are permitted. See, for example, Gans and Smart (1996) and Roberts (1977).

optimal for some skill type. Consequently, it is possible to index the set of admissible tax policies by the skill level. In effect, this index is a single dimension on which the policies can be ordered.

When preferences are quasilinear in consumption and there is a finite number of skill levels, Röell (2012) has shown that preferences over the selfishly optimal tax schedules is single-peaked provided that the minimum-utility constraint does not bind. Thus, the median voter theorem of Black (1948) applies, and so the median skill-type's most preferred tax schedule is a Condorcet winner. Röell (2012) only provides a partial characterization of the selfishly optimal tax schedules. In Brett and Weymark (2016), we provide a complete characterization of these schedules when there is a continuum of skill levels and the minimum-utility constraint is not imposed. Using this characterization, we were able to determine the utility that each type of individual obtains from the selfishly optimal schedule proposed by any other type and thereby identify how each skill type ranks the permissible tax schedules. This allowed us to provide a simple demonstration of Röell's result that these preferences are single-peaked in the proposers' skill levels, from which her median voter theorem follows.

Here, we extend our earlier results by providing a complete characterization for the continuum model of the selfishly optimal income tax schedules that satisfy the minimum-utility constraint in addition to the incentive-compatibility and government budget constraints when, as in our earlier article, preferences are quasilinear in consumption. Moreover, we show that individual preferences are single-peaked over these selfishly optimal tax schedules, and so the median skill-type's preferred schedule is a Condorcet winner. For a continuum of skill types, Bohn and Stuart (2013) also investigate majority voting over selfishly optimal tax schedules with the same constraints as are used here, but without our restriction that preferences are quasilinear. They show the existence of a Condorcet winner in their model without appealing to single-peakedness or Black's median voter theorem. As in Röell (2012), Bohn and Stuart (2013) only provide a partial characterization of the selfishly optimal tax schedules. The complete characterization of these schedules plays an important role in establishing our median voter theorem.

A selfishly optimal income tax schedule can be identified from a schedule that shows how the optimal before-tax income varies with the skill level. A proposer prefers to redistribute resources from other skill types towards himself. In effect, he uses a maxi-max social welfare function for types with lower skills and a maxi-min social welfare function for those with higher skills. In our companion article, we showed that if a proposer simply allocated the maxi-max incomes to all lower types and the maxi-min incomes to all higher types, then the second-order incentive constraint would be violated. In order to satisfy this constraint, the selfishly optimal before-tax income schedules must instead consist of three regions. In the lower part of the skill distribution, an individual receives his maxi-max income, whereas in the upper part of the skill distribution, an individual receives his maxi-min income. For intermediate skill levels, including the skill level of the proposer, everybody receives the same before-tax income. This region provides a "bridge" between the maxi-max and maxi-min parts of the schedule. As in the utilitarian optimum (Mirrlees, 1971), everybody on the maxi-min part of the schedule faces a positive marginal tax rate except for the most highly skilled, whose marginal tax rate is zero. On the maxi-max part of the schedule, everybody faces a negative marginal tax rate (a marginal wage subsidy) except for the least skilled, whose marginal tax rate is zero.

We show that with the addition of the minimum-utility constraint, the before-tax income schedules that are selfishly optimal also have three regions. Because the resources that can be extracted from the lowest skilled are now more limited, the first region no longer tracks the maxi-max solution. Instead, the before-tax incomes of the lowest types lie strictly between the maxi-min and maxi-max incomes when the minimum-utility constraint binds. In effect, the minimum-utility constraint gives rise to a countervailing incentive to transfer resources towards the lowest type, which pushes a selfishly optimal tax schedule in the direction of the maxi-min schedule. Because preferences are quasilinear in consumption, the introduction of the minimum-utility constraint does not affect the qualitative features of the other two regions of the before-tax income schedule.

In our companion article, we were able to determine the before-tax incomes for each skill level in a selfishly optimal schedule point-wise, and this facilitated our demonstration that individuals have single-peaked preferences over the set of selfishly optimal schedules. The minimum-utility constraint precludes us from characterizing these schedules point-wise, which adds to the complexity of the analysis. Nevertheless, in spite of this added complexity, we are able to show that individuals have single-peaked preferences over the selfishly optimal schedules and, hence, the schedule proposed by the median skill type is a Condorcet winner. If the minimum-utility constraint does not bind for the median type's schedule, then the resulting tax system is as described in Brett and Weymark (2016).

The remainder of this article is organized as follows. The next section describes the model economy. Section 3 contains a detailed analysis of the selfishly optimal schedules. The existence of a Condorcet winner is established in Section 4. Section 5 contains concluding remarks. The proofs of our results are given in the Appendix.

2. The Model

There is a continuum of individuals that differ in skill w. The skill parameter w is an individual's type. It measures an individual's constant marginal productivity of labor. The cumulative distribution function $F(\cdot)$ for this parameter is continuous with support $[\underline{w}, \bar{w}]$, where $0 < \underline{w} < \bar{w}$. The density $f(w)$ is assumed to be positive for all w in the support of F. Labor markets are perfectly competitive, so an individual's before-tax income is given by

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

where $l \geq 0$ is the amount of labor supplied. Thus, w is this type's wage rate. Income can also be thought of as being labor in efficiency units. There is a single consumption good which serves as the numeraire in this economy. The amount consumed is $x \geq 0$.

All individuals have the same quasilinear-in-consumption preferences over labor and consumption represented by the utility function

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

on $\mathbb{R}^2_+$, where the function h is increasing, strictly convex, and three-times continuously differentiable on $\mathbb{R}_+$. Because the number of hours needed to achieve a given level of income is decreasing in the skill level, individuals with different skills differ in their preferences over income and consumption. In light of (1), these preferences can be represented by the parametrized utility function

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

The standard Mirrlees (1971) single-crossing property of preferences is satisfied with respect to income and consumption because the marginal rate of substitution at any bundle (y, x) is decreasing in w when $y > 0$.

Individuals face an anonymous tax schedule $T\colon \mathbb{R}_+ \to \mathbb{R}$ that specifies the tax $T(y)$ paid, which could be negative, by someone with income y. The choice of this schedule is determined by majority voting, as described below. The maximum consumption of any individual is his after-tax income. Thus, an individual's budget set consists of all bundles $(y, x) \in \mathbb{R}^2_+$ for which $x \leq y - T(y)$. Because everybody faces the same income tax schedule, they share a common budget set.

An allocation $(y(\cdot), x(\cdot))$ specifies the income $y(w)$ and consumption $x(w)$ for each type w. Admissible allocations are assumed to be integrable. The first requirement for an allocation $(y(\cdot), x(\cdot))$ to be feasible—the incentive constraint—is that there is an anonymous tax schedule for which $(y(w), x(w))$ is utility maximal for type w in the budget set corresponding to this tax schedule. The maximized level of utility for type w with the allocation $(y(\cdot), x(\cdot))$ is

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

Because everybody faces the same budget set, it follows from (3) that $V(w, y(\cdot), x(\cdot))$ is nondecreasing in w.

We assume that taxation is purely redistributive, so in order for a tax schedule to be feasible, the allocation $(y(\cdot), x(\cdot))$ that it generates must also satisfy the government budget constraint

$$\int_{\underline{w}}^{\bar{w}} [y(w) - x(w)] f(w)\, dw \geq 0. \tag{5}$$

Because preferences are quasilinear, the qualitative features of our analysis are unaffected if the government instead requires a fixed positive amount of revenue. By Walras' Law, the constraint (5) is equivalent to the economy's materials balance constraint.

In Brett and Weymark (2016), an income tax schedule is feasible if the allocation that results from its adoption satisfies the incentive and government budget constraints described above. Each type w proposes the feasible tax schedule that is utility maximal for him. It is these selfishly optimal tax schedules that are voted on.

Because each type is only concerned with promoting his own interests and only takes account of the interests of other types to the extent that is necessary to satisfy the incentive constraint, each type will propose a tax schedule that redistributes resources from those less skilled to himself. This redistribution could result in the least skilled working for little or no reward. To preclude this kind of extreme exploitation, here, as in Röell (2012) and Bohn and Stuart (2013), we suppose that the feasibility of a tax schedule also requires that it provide the opportunity for everybody to achieve some minimal level of utility, say u_0. If individuals have the option of emigrating, this constraint can be thought of as being an individual participation constraint. Alternatively, it may be the case that there is a social convention forbidding anybody to obtain utility below this threshold. Because the incentive constraint implies that utility is nondecreasing in w, it is sufficient to impose the minimum-utility constraint for the lowest type; that is, to require that

$$V(\underline{w}, y(\cdot), x(\cdot)) \geq u_0, \tag{6}$$

where $(y(\cdot), x(\cdot))$ is the allocation generated by the tax schedule being considered. To make matters non-trivial, we assume that this constraint binds for the allocation obtained with the maxi-max tax schedule proposed by the highest skilled type.

Because this is a static economy, the taxation principle (see Hammond, 1979; Guesnerie, 1995) applies. According to this principle, any allocation that can be achieved by individuals maximizing utility given some anonymous income tax schedule can also be obtained by directly specifying the consumption bundle $(y(w), x(w))$ for each type w provided that this allocation satisfies the incentive-compatibility condition that

$$V(w, y(\cdot), x(\cdot)) = \max_{w' \in [\underline{w}, \bar{w}]} x(w') - h\left(\frac{y(w')}{w}\right), \quad \forall w \in [\underline{w}, \bar{w}]. \tag{7}$$

Because it is more convenient to work with the consumption and before-tax income schedules, $x(\cdot)$ and $y(\cdot)$, than with the underlying tax schedule $T(\cdot)$, henceforth, we assume that $x(\cdot)$ and $y(\cdot)$ are chosen directly rather than indirectly through the intermediation of $T(\cdot)$.

Because the single-crossing property is satisfied, it follows from Mirrlees (1976) that the first-order (envelope) condition for an allocation $(y(\cdot), x(\cdot))$ to be incentive compatibile is

$$V_w(w, y(\cdot), x(\cdot)) = h'\left(\frac{y(w)}{w}\right)\frac{y(w)}{w^2}, \quad \forall w \in [\underline{w}, \bar{w}], \tag{8}$$

and the second-order incentive-compatibility condition is

$$y'(w) \geq 0, \quad \forall w \in [\underline{w}, \bar{w}].[3] \tag{9}$$

Because h is increasing, (8) not only implies that utility is nondecreasing in w whenever the incentive-compatibility constraint is satisfied, it also implies that utility is strictly

[3]The expressions in (8) and (9) are required to hold at all points for which $y(\cdot)$ is differentiable.

increasing for all w for which $y(w) > 0$. Consumption must also be nondecreasing in w. Indeed, consumption is strictly increasing in w whenever income is also strictly increasing in w, and any two types that have the same income also have the same consumption, in which case these types are said to be bunched (see Laffont and Martimort, 2002, sec. 3.1).

In summary, for an allocation $(y(\cdot), x(\cdot))$ (and the underlying tax schedule) to be feasible, it must satisfy the first- and second-order incentive-compatibility constraints (8) and (9), the government budget constraint (5), and the minimum-utility constraint (6), where the maximized utility $V(w, y(\cdot), x(\cdot))$ that appears in these constraints is defined by (4). Type k's selfishly-optimal tax schedule is implicitly defined by the allocation $(y(\cdot, k), x(\cdot, k))$ that maximizes his utility subject to these feasibility constraints, where now the income and consumption schedules are indexed by the type k that proposes them. Formally, *type k's problem* is

$$\max_{x(\cdot,k),y(\cdot,k)} V(k) \quad \text{subject to (4), (5), (6), (8), and (9).} \tag{10}$$

The set of alternatives that are voted on consists of the allocations that solve (10) for some type $k \in [\underline{w}, \bar{w}]$.

In addition to determining the allocation that solves type k's problem, we are interested in determining the marginal tax rates implied by his proposal. For the allocation $(x(\cdot, k), y(\cdot, k))$ that solves type k's problem, the marginal tax rate for an individual of type w is given by

$$\tau(w, x(\cdot, k), y(\cdot, k)) = 1 - h'\left(\frac{y(w,k)}{w}\right)\frac{1}{w}. \tag{11}$$

The marginal tax rates do not depend on consumption because utility is quasilinear in consumption. For this reason, we henceforth omit $x(\cdot, k)$ as a argument of the function τ. If the tax function is differentiable in y at $y(w, k)$ (or, equivalently, the boundary of the budget set is differentiable at $(y(w, k), x(w, k))$), then (11) simply expresses the fact that an individual of type w chooses his consumption bundle so that the marginal rate of substitution between income and consumption equals the retention rate, which is one minus the marginal tax rate. If, however, the tax function is not differentiable at $y(w, k)$, then there is a tax function that is locally differentiable at $y(w, k)$ that would induce the same behavior, in which case (11) is the implicit marginal tax rate faced by type w.

3. Selfishly Optimal Tax Schedules

Our analysis of type k's problem begins by characterizing the solution to what we call *type k's relaxed problem*, which is (10) with the second-order incentive-compatibility constraint (9) removed.[4] While the solution to type k's problem does not coincide with

Because incentive compatibility implies that income is nondecreasing in w, $y(\cdot)$ is differentiable almost everywhere.

[4]Our plan of attack thus makes initial use of the first-order approach commonly used in screening problems. The second-order incentive-compatibility conditions in optimal tax problems have been explicitly taken into account by Brito and Oakland (1977) and Ebert (1992).

the solution to his relaxed problem, the latter is a useful building block in constructing the former. A useful benchmark is the solution to (10) without either the second-order incentive-compatibility constraint (9) or the minimum-utility constraint (6). We refer to this problem as *type k's doubly relaxed problem.* In this section, except for the occasional use of the optimal schedules of types $\underline{w}$ and $\bar{w}$ as points of comparison, we hold the proposer's type fixed, and so suppress the use of k to index his proposal.

We adapt the procedure of Lollivier and Rochet (1983) to our model in order to derive a reduced-form problem whose solution is the before-tax income schedule that solves type k's relaxed problem. The key steps in this procedure are to use (4), (5), and (8) to derive an expression for $V(w, y(\cdot), x(\cdot))$ in which the consumption schedule $x(\cdot)$ does not appear. In this derivation, we make use of the fact that the government budget constraint binds at a solution to the relaxed problem.[5] Because the optimal marginal tax rates do not depend on consumption, type k's optimal consumption schedule can be computed from the optimal income schedule using the incentive-compatibility and binding government budget constraints, so we do not consider it explicitly.

Proposition 1. *The optimal schedule of before-tax incomes $y(\cdot)$ for type k's relaxed problem is obtained by solving*

$$\max_{y(\cdot)} \int_{\underline{w}}^{k} G^M(w, y(w)) + \int_{k}^{\bar{w}} G^R(w, y(w)) \quad \textit{subject to} \quad \int_{\underline{w}}^{\bar{w}} G^R(w, y(w))dw \geq u_0, \quad (12)$$

where

$$G^M(w, y) = \left[y - h\left(\frac{y}{w}\right)\right] f(w) + \frac{y}{w^2}h'\left(\frac{y}{w}\right) F(w) \quad (13)$$

and

$$G^R(w, y) = \left[y - h\left(\frac{y}{w}\right)\right] f(w) - \frac{y}{w^2}h'\left(\frac{y}{w}\right) [1 - F(w)] \, . \quad (14)$$

The optimal schedule of before-tax incomes $y(\cdot)$ for type k's doubly relaxed problem is obtained by maximizing the objective function in (12) without the constraint. For the doubly relaxed case, we let $y^R(\cdot)$ and $y^M(\cdot)$ denote the optimal income schedules for types $\underline{w}$ and $\bar{w}$, respectively. They are the optimal income schedules that are obtained using the maxi-min and maxi-max social welfare functions when the only constraints are those of a doubly relaxed problem.

When $k = \bar{w}$, it is the highest skilled's utility that is being maximized, so $G^M(w, y(w))$ is the integrand in the reduced-form maxi-max problem. Analogously, when $k = \underline{w}$, it is the lowest skilled's utility that is being maximized, so $G^R(w, y(w))$ is the integrand in the reduced-form maxi-min problem. Thus, in (12), the objective function that type k uses to determine his optimal income schedule employs the maxi-max utility objective for types that are less skilled than himself and the maxi-min utility objective for types that are more

[5]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 both the utility of type k individuals (whose utility is being maximized) and of type $\underline{w}$ individuals (to whom the minimum-utility constraint applies).

skilled. In deriving this objective function, all of the constraints in the relaxed problem have been accounted for except for the minimum-utility constraint. The constraint in (12) is the reduced-form version of it. As is shown in the proof of Proposition 1, with the income schedule $y(\cdot)$ chosen by type k, the lowest skilled obtain utility equal to the value of the integral in the reduced-form constraint. This integral computes the value of a maxi-min objective function using the income schedule that is optimal for type k. When the proposer's type k differs from $\underline{w}$, this utility is not what type $\underline{w}$ would obtain with his own proposed income schedule as that schedule differs from the one proposed by type k.

The functions $G^M(\cdot)$ and $G^R(\cdot)$ can also be given virtual surplus interpretations. The term in square brackets on the left-hand sides of both (13) and (14) is a type w individual's output less the utility cost of producing that output. The second terms on the left-hand sides of (13) and (14) measure the information rent afforded to type w individuals when the minimum-utility constraint is ignored. The exact form of the information rents is different for types below k than it is for types above k. This is because the optimal schedule for type k extracts as many resources as possible from all other types and redistributes them to individuals of type k. When $w > k$, this is a typical downward redistribution of income, which is constrained by downward binding incentive constraints. For this reason, the information rent term in (14) has a conventional form. The less familiar form of the information rents in (13) arise because of the motive for upward redistribution of income from types with $w < k$. We provide a detailed interpretation of this term in Brett and Weymark (2016).

The optimization problem (12) admits a point-wise solution. After associating a nonnegative Lagrange multiplier λ (which depends on k) with the minimum-utility constraint in type k's relaxed problem, simple differentiation with respect to $y(w)$ yields the following first-order conditions:

$$\begin{aligned} \theta^M(w, y(w)) + \lambda\theta^R(w, y(w)) = 0, &\quad \forall w \in [\underline{w}, k), \\ \theta^R(w, y(w)) = 0, &\quad \forall w \in (k, \bar{w}], \end{aligned} \tag{15}$$

where

$$\theta^M(w, y) = \frac{\partial G^M(w, y)}{\partial y} = \left[1 - h'\left(\frac{y}{w}\right)\frac{1}{w}\right] f(w) + \left[h''\left(\frac{y}{w}\right)\frac{y}{w^3} + h'\left(\frac{y}{w}\right)\frac{1}{w^2}\right] F(w) \tag{16}$$

and

$$\theta^R(w, y) = \frac{\partial G^R(w, y)}{\partial y} = \left[1 - h'\left(\frac{y}{w}\right)\frac{1}{w}\right] f(w) - \left[h''\left(\frac{y}{w}\right)\frac{y}{w^3} + h'\left(\frac{y}{w}\right)\frac{1}{w^2}\right] [1 - F(w)].^{6} \tag{17}$$

The corresponding first-order conditions for the doubly relaxed problem are obtained by

[6] 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.

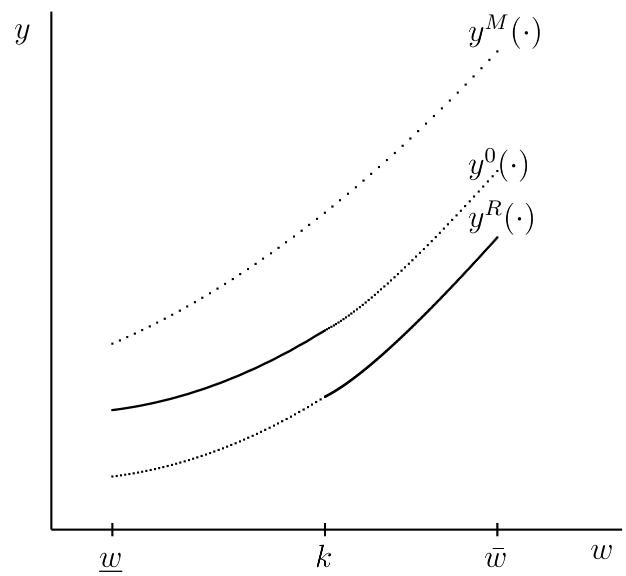

Figure 1: The optimal income schedule for type k's relaxed problem

setting $\lambda = 0$ in (15). We assume that the Lagrangians for both the relaxed and doubly relaxed problems are strictly concave in income. That is, we assume that

$$\theta_y^M(w,y) < 0,\ \theta_y^R(w,y) < 0,\ \text{and}\ \theta_y^M(w,y) + \lambda\theta_y^R(w,y) < 0,\ \forall(w,y) \in [\underline{w}, \bar{w}] \times \mathbb{R}_+. \quad (18)$$

The solution to the first-order conditions (15) is type k's optimal income schedule for the relaxed problem. Similarly, the solution to (15) with $\lambda = 0$ is type k's optimal income schedule for the doubly relaxed problem. The solution $y^0(\cdot)$ to type $\bar{w}$'s relaxed problem is obtained by solving the first-order condition in the first line of (15) for all $w \in [\underline{w}, \bar{w}]$. When $\lambda = 0$, this solution coincides with the solution $y^M(\cdot)$ to the corresponding doubly relaxed problem, but if $\lambda > 0$ (and, therefore, the minimum-utility constraint binds), it does not. The solution to type $\underline{w}$'s relaxed problem is obtained by solving the first-order condition in the second line of (15) for all $w \in [\underline{w}, \bar{w}]$. Because λ does not appear in this expression, the solutions to this type's relaxed and doubly relaxed problems are the same, namely, $y^R(\cdot)$.[7] As we shall see, the $y^M(\cdot)$ schedule lies above the $y^R(\cdot)$ schedule. The schedule $y^0(\cdot)$ is a weighted combination of these two schedules. Thus, the solution to type k's relaxed problem at w is given by $y^0(w)$ for types less skilled than himself and by $y^R(w)$ for those who are more skilled, as illustrated in Figure 1.

By the Implicit Function Theorem (Sundaram, 1996, Theorem 1.77), it follows from (18) that $y^M(\cdot)$, $y^R(\cdot)$, and $y^0(\cdot)$ are continuously differentiable functions. An implication of (18) is that the second-order condition for the choice of $y(w)$ holds strictly for type

[7]When the minimum-utility constraint binds, type $\underline{w}$ takes this into account by adjusting the maxi-min consumption schedule, not the maxi-min income schedule.

k's reduced-form optimization problem:

$$\begin{aligned} \theta_y^M(w, y(w)) + \lambda \theta_y^R(w, y(w)) < 0, \quad &\forall w \in [\underline{w}, k), \\ \theta_y^R(w, y(w)) < 0, \quad &\forall w \in (k, \bar{w}]. \end{aligned} \tag{19}$$

These inequalities also imply that corresponding second-order conditions for the doubly relaxed problems of types $\bar{w}$ and $\underline{w}$ hold strictly:

$$\begin{aligned} \theta_y^M(w, y(w)) < 0, \quad &\forall w \in [\underline{w}, \bar{w}], \\ \theta_y^R(w, y(w)) < 0, \quad &\forall w \in [\underline{w}, \bar{w}]. \end{aligned} \tag{20}$$

We now confirm that the $y^M(\cdot)$ schedule lies above the $y^R(\cdot)$ schedule, with the $y^0(\cdot)$ schedule lying between them when $\lambda > 0$. For all points on the graphs of $y^R(\cdot)$, $y^M(\cdot)$, and $y^0(\cdot)$, we have $\theta^R(w, y(w)) = 0$, $\theta^M(w, y(w)) = 0$, and $\theta^M(w, y(w)) + \lambda\theta^R(w, y(w)) = 0$, respectively. Thus, by the second line in (20), $\theta^R(w, y(w))$ is negative (resp. positive) above (resp. below) the graph of $y^R(\cdot)$. Similarly, by the first line in (20), $\theta^M(w, y(w))$ is negative (resp. positive) above (resp. below) the graph of $y^M(\cdot)$. Thus, the graph of $y^M(\cdot)$ lies above that of $y^R(\cdot)$. On or below the graph of $y^R(\cdot)$, $\theta^M(w, y(w)) + \lambda\theta^R(w, y(w)) > 0$, so the graph of $y^0(\cdot)$ must lie above $y^R(\cdot)$. Moreover, if $\lambda > 0$, on or above the graph of $y^M(\cdot)$, $\theta^M(w, y(w)) + \lambda\theta^R(w, y(w)) < 0$, so the graph of $y^0(\cdot)$ must lie below $y^M(\cdot)$ when $\lambda > 0$.

For simplicity, we assume that $y^R(\cdot)$ and $y^M(\cdot)$ are nondecreasing and, hence, satisfy the second-order incentive-compatibility condition (9). If this assumption is not satisfied, then in the subsequent analysis, $y^R(\cdot)$ and $y^M(\cdot)$ are replaced by the maxi-min and maxi-max solutions to the (unrelaxed) problems of types $\underline{w}$ and $\bar{w}$, respectively, ignoring the minimum-utility constraint. These modified schedules can be obtained from $y^R(\cdot)$ and $y^M(\cdot)$ using the ironing technique described in Guesnerie and Laffont (1984).

For a type w below k, the addition of the minimum-utility constraint (6) effectively turns type k's problem into one of maximizing a weighted sum of his own utility and that of the lowest type, where the weight given to the lowest type is the endogenous shadow value of the minimum-utility constraint. This gives rise to a problem of countervailing incentives, as in Jullien (2000). On the one hand, type k wants to redistribute resources upwards toward his own type. In the absence of a minimum-utility constraint, this would give rise to the maxi-max utility solution $y^M(w)$. On the other, he needs to move resources downward to the very lowest skilled in order to satisfy the minimum-utility constraint. The before-tax income that reconciles these motives naturally lies between their maxi-max and maxi-min values.

For a type w above k, as we have seen, the solution is exactly the same as the one obtained with the relaxed problem in the absence of a minimum-utility constraint; that is, it is given by the maxi-min utility solution $y^R(w)$. A type k individual wants to redistribute resources downward from types greater than his own, and redistribution from higher types is limited by downward incentive constraints. Because preferences are quasilinear, type k proposes the same income schedule for the upper part of the skill distribution as would type $\underline{w}$. Because these two types propose different incomes for the rest of the skill distribution, in order to ensure that the minimum-utility constraint is satisfied, they propose different consumption schedules.

Using (11), (15), (16), and (17), the associated marginal tax rates for the $y^R(\cdot)$, $y^M(\cdot)$ and $y^0(\cdot)$ income schedules are

$$\tau^R(y(w,k) = \frac{1-F(w)}{f(w)}\left[h''\left(\frac{y(w)}{w}\right)\frac{y(w)}{w^3} + h'\left(\frac{y(w)}{w}\right)\frac{1}{w}\right], \quad \forall w \in [\underline{w}, \bar{w}],^{8} \tag{21}$$

$$\tau^M(y(w),k) = -F(w)\frac{1}{f(w)}\left[h''\left(\frac{y(w)}{w}\right)\frac{y(w)}{w^3} + h'\left(\frac{y(w)}{w}\right)\frac{1}{w}\right], \quad \forall w \in [\underline{w}, \bar{w}] \tag{22}$$

and

$$\tau^0(y(w),k) = \left[\frac{\lambda}{1+\lambda} - F(w)\right]\frac{1}{f(w)}\left[h''\left(\frac{y(w)}{w}\right)\frac{y(w)}{w^3} + h'\left(\frac{y(w)}{w}\right)\frac{1}{w}\right], \quad \forall w \in [\underline{w}, \bar{w}] \tag{23}$$

In (21), the term is square brackets is positive, so the marginal tax rate is positive except for $w = \bar{w}$, where it is it is zero. This is the familiar pattern of marginal tax rates for a maxi-min social welfare function. Similarly, $\tau^M(y(w))$ is negative except at the very lowest skill level, which is the pattern of marginal tax rates for a maxi-max social welfare function when the minimum-utility constraint does not bind. The motive for upward redistribution for a maxi-max social welfare function provides a reason to provide these marginal wage subsidies. However, a binding minimum utility constraint produces countervailing incentives that serve to increase optimal marginal tax rates from the maxi-max levels. Nevertheless, because $\lambda/(1+\lambda) < 1$, even if $\tau^0(y(w))$ is positive for some w, it is less than the corresponding maxi-min optimal tax rate. Thus, the desire for upward redistribution is never completely overpowered by the countervailing incentive.

At the solution to type k's relaxed problem, (23) applies for $w \leq k$ and (21) for $w > k$. Thus, there is a downward jump discontinuity in the marginal tax rate at k as it switches from negative to positive values. As illustrated in Figure 1, there is a corresponding downward jump discontinuity in the solution to type k's relaxed problem at k. Because the graph of $y^M(\cdot)$ lies above that of $y^R(\cdot)$ and that of $y^0(\cdot)$ lies between these two curves, this discontinuity is a downward jump.

The negative marginal tax rates associated with $y^M(\cdot)$ imply that incomes are distorted upwards relative to the full-information solution. Similarly, the positive marginal tax rates associated with $y^R(\cdot)$ imply that incomes are distorted downwards relative to this benchmark. It is for this reason that the graph of $y^M(\cdot)$ lies above that of $y^R(\cdot)$.

We have seen that the before-tax income schedule that solves type k's relaxed problem satisfies the second-order incentive compatibility condition everywhere except possibly at k, where this schedule has a downward jump. Because of this downward discontinuity,

[8] This equation is equivalent to the condition stated in Boadway and Jacquet (2008, eqn. (21), p. 435) for optimal tax rates with a maxi-min objective.

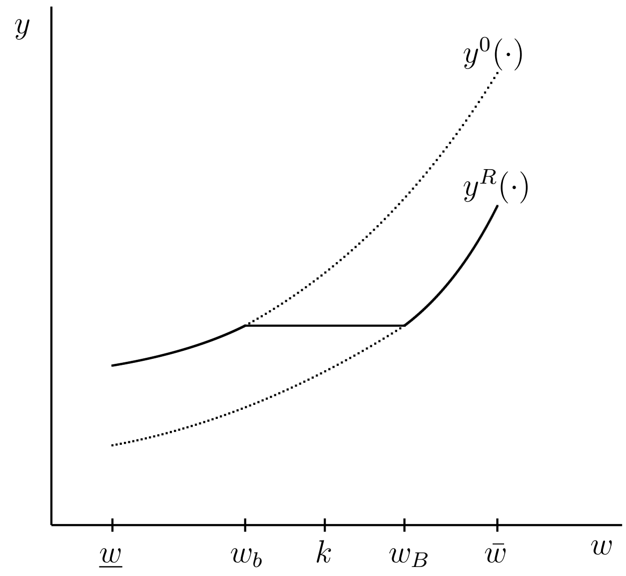

Figure 2: A bridge

the solution to the relaxed problem is not incentive compatible. The standard remedy for dealing with a single decreasing portion of an otherwise optimal allocation is to replace this income schedule by a continuous schedule that has a single bunching region which removes the decreasing segment using the ironing technique of Guesnerie and Laffont (1984). Here, this involves connecting the lower part of $y^0(\cdot)$ to the upper part of $y^R(\cdot)$ using a horizontal segment, as illustrated in Figure 2. In order to avoid some technical complexities, we assume that the solution to type k's reduced-form problem has this form without specifying what further restrictions are sufficient in order to ensure that this is the case.[9] The types who are bunched together consist of a closed interval $[w_b, w_B]$. We refer to this interval as a *bridge* and to w_b and w_B as the lower and upper bridge endpoints, respectively. The proposer's own type lies on the bridge. We assume that it is not optimal for everybody to be bunched; that is, we assume that the bridge is not all of $[\underline{w}, \bar{w}]$.

In Brett and Weymark (2016), we use a relatively simple procedure that was introduced by Vincent and Mason (1968) to identify the bridge endpoints when the minimum utility constraint does not apply. The procedure consists of two steps. First, the op-

[9]As shown in Brett and Weymark (2016), in the absence of the minimum-utility constraint, it is optimal for type k to choose an income schedule of this form without the need for any further restrictions on the model. Their arguments apply here as well for the part of k's optimal income schedule that concerns types with skill levels at least k, but they do not apply for the rest of the skill distribution if the minimum-utility constraint binds. See Guesnerie and Laffont (1984), Jullien (2000), Nöldeke and Samuelson (2007), and Hellwig (2010) for analyses of bunching and continuity for quasilinear adverse selection problems with a minimum-utility (participation) constraint but no analogue of our government budget constraint.

timal 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 chosen. Because the minimum-utility constraint (12) contains an integral involving the entire before-tax income schedule, including its value along the bridge, it is impossible to use this two-step procedure to find the bridge endpoints. Instead, here, the bridge endpoints are determined jointly with the before-tax incomes and the Lagrange multiplier λ. The value of the Lagrange multiplier helps to determine the value of $y^0(w)$ (but not the value of $y^R(w)$). It is, nevertheless, possible to determine first-order conditions that characterize the simultaneous choice of the Lagrange multiplier and the bridge endpoints. These conditions, and a summary of the preceding discussion, are given in Proposition 2.

Proposition 2. *The optimal schedule of before-tax incomes $y(\cdot)$ for type k's problem is given by*

$$y(w) = \begin{cases} y^0(w), & \forall w \in [\underline{w}, w_b), \\ y^0(w_b), & \forall w \in [w_b, w_B] \text{ if } w_b > \underline{w}, \\ y^R(w_B), & \forall w \in [w_b, w_B] \text{ if } w_B < \bar{w}, \\ y^R(w), & \forall w \in (w_B, \bar{w}]. \end{cases} \tag{24}$$

The optimal values of the Lagrange multiplier λ and the bridge endpoints w_b and w_B are determined by solving the complementary slackness condition for the minimum-utility constraint

$$\lambda \left[\int_{\underline{w}}^{\bar{w}} G^R(w, y(w))dw - u_0 \right] = 0 \tag{25}$$

together with either the first-order condition

$$\int_{w_b}^{k} \theta^M(w, y^0(w_b))dw + \lambda \int_{w_b}^{k} \theta^R(w, y^0(w_b))dw + (1+\lambda) \int_{k}^{w_B} \theta^R(w, y^0(w_b))dw = 0 \tag{26}$$

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

$$\int_{w_b}^{k} \theta^M(w, y^R(w_B))dw + \lambda \int_{w_b}^{k} \theta^R(w, y^R(w_B))dw + (1+\lambda) \int_{k}^{w_B} \theta^R(w, y^R(w_B))dw = 0 \tag{27}$$

if $w_B < \bar{w}$.

In light of (24), conditions (26) and (27) are equivalent when both bridge endpoints are in the interior of the type space. These bridging first-order conditions have the standard interpretation that the average of the marginal virtual surpluses on a bunching region must be zero (see, for example, Nöldeke and Samuelson, 2007, p. 416). Whenever (26) or (27) are satisfied, a marginal change in the level of income on the bridge has no effect on the objective function in (12). The income level on the bridge pertains to all types in $[w_b, w_B]$, so the marginal effects that the incomes of all these types have on the objective and constraint in type k's reduced-form problem must be accounted for. This explains why integrals appear in the first-order conditions (26) and (27).

When $\lambda = 0$, $y^0(w) = y^M(w)$. Thus, when the minimum-utility constraint does not bind, type k's optimal income schedule has the form shown in Figure 2 with $y^0(\cdot)$ replaced by the maxi-max income schedule $y^M(\cdot)$. That is, $y(\cdot)$ is obtained by inserting a bridge between the maxi-max and maxi-min schedules. The solution for this special case was identified in Brett and Weymark (2016).

4. The Voting Equilibrium

We now turn to the voting over the selfishly optimal income tax schedules. We show that preferences over the selfishly optimal tax policies are single-peaked with respect to the skill level of the types for whom these schedules are optimal. As a consequence, with pairwise majority voting, the selfishly optimal income tax schedule that maximizes the utility of the median type is chosen; it is a Condorcet winner.

Because we are now considering comparisons across tax schedules that are optimal for different types, henceforth, we index allocations by the proposer's type k. Recall that $(x(w,k), y(w,k))$ denotes the optimal allocation assigned to an individual of type w in the solution to type k's problem. Let

$$V^0(w,k) = x(w,k) - h\left(\frac{y(w,k)}{w}\right) \tag{28}$$

denote type w's optimized utility if type k's proposal is accepted. An individual of type w evaluates the schedules chosen by each of the types according to the function $V^0(w,\cdot)$. Let $\lambda(k)$, $w_b(k)$, and $w_B(k)$ respectively denote the values of the Lagrange multiplier associated with minimum-utility constraint and the two bridge endpoints at the solution to type k's problem. By the Implicit Function Theorem, $\lambda(\cdot)$, $w_b(\cdot)$, and $w_B(\cdot)$ are continuously differentiable functions. Identifying the structure of $V^0(w,\cdot)$ is necessary in order to establish single-peakedness. This structure, in turn, is determined by the comparative statics of the functions $\lambda(\cdot)$, $w_b(\cdot)$, and $w_B(\cdot)$ with respect to k. These comparative statics are presented in Proposition 3.

Proposition 3. *For all $k \in [\underline{w}, \bar{w})$, a marginal increase in k results in*

1. *an increase in $\lambda(k)$ if $\lambda(k) > 0$;*
2. *an increase in both $w_b(k)$ and $w_B(k)$ if $\underline{w} < w_b(k) < w_B(k) < \bar{w}$;*
3. *an increase in $w_b(k)$ and no change in $w_B(k)$ if $\underline{w} < w_b(k) < w_B(k) = \bar{w}$;*
4. *an increase in $w_B(k)$ if $\underline{w} = w_b(k) < w_B(k) < \bar{w}$ if $\lambda(k) = 0$;*
5. *no change in $w_B(k)$ if $\underline{w} = w_b(k) < w_B(k) < \bar{w}$ if $\lambda(k) > 0$;*

An implication of Part 1 of this proposition is that the set of types who propose a selfishly-optimal tax schedule for which the minimum-utility constraint does not bind consists of an interval at the bottom of the skill distribution.

In Brett and Weymark (2016), we show that the endpoints of this bridge are non-decreasing in the type of proposer when the minimum-utility constraint is not binding. Parts (2)–(4) of Proposition 3 provide a somewhat sharper statement of our earlier findings. Specifically, the upper endpoint of the bridge is increasing in k unless it cannot be increased at all because it is already at $\bar{w}$. The lower endpoint is also increasing in k, except possibly when the bridge starts at $\underline{w}$. Intuitively, when the proposer's type increases, there are more types below him. He wants to distort their incomes upwards, so he applies the maxi-max schedule to more (or at least, no fewer) types. Similarly, there are fewer types above him and he wants to distort their incomes downwards, so he applies the maxi-min schedule to fewer (or at least, no more) types.

Part 1 of Proposition 3 shows that the shadow value of the minimum-utility constraint is increasing in the type of proposer whenever this constraint is binding. When the type of proposer increases, it becomes relatively more attractive to extract information rents from types less skilled than the proposer than from those types who are more skilled because there are now more of the former and less of the latter. However, the minimum-utility constraint acts to limit the upward transfers from the first of these groups. As a consequence, a relaxation of the minimum-utility constraint has a greater value for higher type proposers.

The reasoning about the placement of the upper bridge endpoint that was used for the case in which the minimum-utility constraint does not bind also applies when it does. Thus, if the initial upper bridge endpoint is not at $\bar{w}$, it is increased when the proposer's type increases, which necessitates increasing the income of the types on the bridge to preserve the form of income schedule shown in Figure 2 .

Provided that the income schedule does not start on the bridge, increasing the proposer's type increases the lower bridge endpoint when the minimum-utility constraint binds, just as it does when it does not. Because $\lambda(k)$ is increasing in k when the minimum-utility constraint binds, as can be seen from the first-order conditions (26) and (27), the $y^0(\cdot)$ income schedule that is used to determine incomes below the bridge shifts down closer to the maxi-min schedule $y^R(\cdot)$ when the proposer's type increases.[10] This change in $y^0(\cdot)$ necessitates increasing the lower endpoint of the bridge so that the lower part of the new $y^0(\cdot)$ income schedule connects to the bridge. Figure 3 illustrates how the selfishly optimal income schedule changes when the proposer's type increases in the presence of a binding minimum-utility constraint.

Type k's selfishly optimal income schedule $y(\cdot, k)$ starts on the bridge when $y^0(\cdot, k)$ lies everywhere above $y(\cdot, k)$. In this case, it is not optimal to change the lower bridge endpoint when the proposer's type increases if the minimum-utility constraint binds. If the proposer's type is increased marginally, as we have seen, the $y^0(\cdot)$ schedule shifts

[10]Because the minimum-utility constraint continues to bind, both the the consumption and income of the lowest skilled type are decreased so that this type's utility remains at u_0.

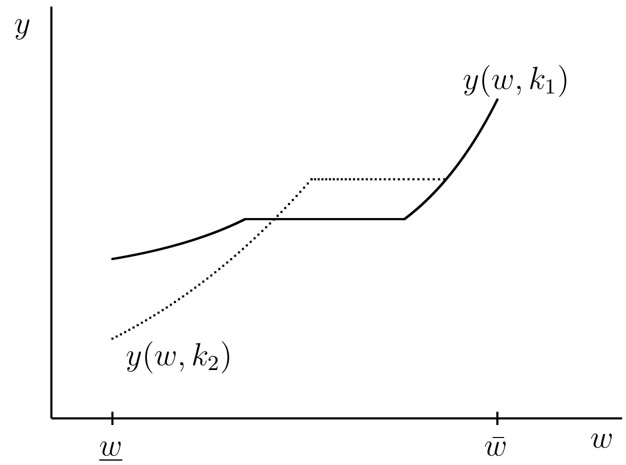

Figure 3: Selfishly optimal income schedules for $k_1 < k_2$

down and the graph of the bridge shifts up, but the former remains above the latter. It is only possible to increase the lower bridge endpoint, but if this is done, the types who are removed from the bridge will then have their incomes determined by the $y^0(\cdot)$ schedule. This would result in a downward discontinuity in the selfishly optimal income schedule, thereby violating the second-order optimality condition.

We now turn our attention to how voters rank the selfishly optimal income schedules that are proposed. These schedules are indexed by the type of proposer, so a preference over them can be regarded as being a preference on the set of skill types. All proposers face the same constraints on the choice of an income schedule. Hence, because each type proposes a feasible income schedule that is selfishly optimal for him, he must weakly prefer what he obtains with his own schedule to what anybody else proposes for him. That is,

$$V(w, w) \geq V(w, k), \quad \forall w, k \in [\underline{w}, \bar{w}]. \tag{29}$$

An individual of type w has a (weakly) *single-peaked preference* on the set of types if

$$V(w, w) \geq V(w, k_1) \geq V(w, k_2) \text{ if } w < k_1 < k_2 \tag{30}$$

and

$$V(w, w) \geq V(w, k_1) \geq V(w, k_2) \text{ if } w > k_1 > k_2. \tag{31}$$

This definition of a single-peaked preference does not require an individual's preference to be strictly monotonic on each side of his peak. In particular, an individual's own proposal need not be uniquely best for himself, so the "peak" may in fact be a "plateau". We show in Proposition 4 that everybody's preferences on the set of types are single-peaked.

Proposition 4. *Individual preferences are single-peaked on the set of skill types.*

In Brett and Weymark (2016), we show that the same result holds in the absence of a minimum-utility constraint. In that case, each proposer chooses where to bridge between the maxi-max and maxi-min income schedules. The situation is more complex when the minimum-utility constraint binds because the schedule used to determine incomes for types below the lower endpoint of the bridge depends on the proposer's type. This type dependence greatly complicates the proof of the single-peakedness result.

Proposition 4 is established by showing that $\partial V^0(w,k)/\partial k$ is nonnegative when $k < w$ and nonpositive when $k > w$. When the minimum-utility constraint does not bind (which, as we have shown, can only occur for an interval of types at the bottom of the skill distribution), these conclusions follow from Brett and Weymark (2016, Proposition 8). So, we only need to consider the case in which the minimum-utility constraint binds when type k is the proposer. In this case, some insight about Proposition 4 can be gained from noting that (8) implies that

$$V^0(w,k) = u_0 + \int_{\underline{w}}^{w} h'\left(\frac{y(t,k)}{t}\right)\frac{y(t,k)}{t^2}\,dt, \quad \forall (w,k) \in [\underline{w}, \bar{w}]. \tag{32}$$

The partial derivative of this function with respect to k is

$$\frac{\partial V^0(w,k)}{\partial k} = \int_{\underline{w}}^{w} \left[h''\left(\frac{y(t,k)}{t}\right)\frac{y(t,k)}{t^3} + h'\left(\frac{y(t,k)}{t}\right)\frac{1}{t^2}\right]\frac{\partial y(t,k)}{\partial k}\,dt, \quad \forall (w,k) \in [\underline{w}, \bar{w}]. \tag{33}$$

The term in the square bracket in (33) is always positive. How type t's income responds to a marginal increase in the type k of the proposer depends on whether the bridge starts or ends at one of the endpoints of the skill distribution.

Consider a marginal increase in the proposer's type from k_1 to k_2. First, suppose that type k_1's bridge does not include either of the endpoints of the skill distribution. In this case, the type space is partitioned into three intervals. As illustrated in Figure 3, except at a boundary between two regions, in the lowest (resp. middle, highest) skilled of these regions, incomes are decreased (resp. increased, unchanged) when type k_2 replaces type k_1 as the proposer. The increase in incomes in the middle region is due to the rising bridge. The boundary between the first two regions occurs at a type on k_1's bridge who is lower skilled than k_2. Thus, the income partial derivatives in the integrand in (33) are negative in the first region, positive in the second, and zero on the third. If $w < k$, all the terms in the integral in (33) are negative if w is in the first region. If w is in the second region, some of the terms in this integral are positive, but there is a preponderance of negative terms, so they dominate. Thus, $\partial V^0(w,k)/\partial k < 0$ if $w < k$. Once k is large enough so that the positive terms in (33) dominate for some type $\hat{w}$, $\partial V^0(w,k)/\partial k > 0$ for any larger value of w because only positive or zero terms are added to the integral for $\hat{w}$. In the case we are considering, type k_2's proposed schedule is uniquely optimal for him, so his utility is increased when he replaces type k_1 as the proposer. Type k_2 can be chosen to be arbitrarily close to type k_1, so by setting $k = k_1$ and $w = k_2$, we have $\partial V^0(w,k)/\partial k > 0$ for w arbitrarily close to k from above. But then, as we have seen, this implies that $\partial V^0(w,k)/\partial k > 0$ if $w > k$. The same reasoning applies if the upper end of the bridge is at $\bar{w}$ in type k's proposal. In this case, the third of the regions described above does not exist. If, however, type k's bridge starts at $\underline{w}$ (so there is no first region), marginally increasing k has no effect on the schedule being proposed, so for all w, $\partial V^0(w,k)/\partial k = 0$.

The median voter theorem (Black, 1948) applies even if some of the inequalities in the definition of a single-peaked preference are not strict. Hence, an immediate consequence of Proposition 4 is that the median skill type's proposed tax schedule does at least as well in a pairwise majority vote as the tax schedule proposed by any other type.

Proposition 5. *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 qualitative features of the income tax schedule chosen by pairwise majority rule depend on whether the minimum-utility constraint binds for the median skill type or not. If it does not bind, then the Condorcet-winning tax schedule is as described in Brett and Weymark (2016). Specifically, when the bridge does not include either of the endpoints of the skill distribution, this schedule 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 skill 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. If the schedule starts or ends on the bridge, the same qualitative features are exhibited on the two regions that do exist. This is the pattern of marginal tax rates for the Condorcet-winning income tax schedule found by Röell (2012) for a discrete population.

When the minimum-utility constraint binds, the Condorcet-winning tax schedule exhibits the same pattern of marginal tax rates for skill types on or above the bridge as when this constraint does not bind because the maxi-min income schedule is used for types above the bridge in both of these situations. However, except in neighborhoods of $\underline{w}$ and $w_b(k^*)$, the sign of the marginal tax rate is now ambiguous below the bunching region because it depends on the exact value of the multiplier $\lambda(k^*)$ for the minimum-utility constraint for the type k^* whose proposal is the Condorcet winner. From (23), it follows that there is an interval of types that includes $\underline{w}$ that face a positive marginal tax rate. If $\lambda(k^*)$ is close to zero, then the $y^0(\cdot, k^*)$ schedule lies near the maxi-max schedule $y^M(\cdot)$, and so the marginal tax rate is negative for some types below the bridge. In this case, there also exists a undistorted type in the interior of the interval below the bridge. If $\lambda(k^*)$ is relatively large, then $y^0(\cdot, k^*)$ lies near the maxi-min schedule $y^R(\cdot)$, and marginal tax rates are everywhere positive below the bridge. Bohn and Stuart (2013) show that a similar pattern of marginal tax rates is exhibited by the Condorcet-winning tax schedule when preferences are not restricted to be quasilinear in consumption and the minimum-utility constraint binds.

5. Conclusion

We have extended the analysis of Brett and Weymark (2016) to allow for a minimum-utility constraint. Many of the qualitative features of the Condorcet-winning income tax schedule are invariant to the introduction of this constraint. However, the maxi-max segment of the schedule that applies below the bunching region when the minimum-utility constraint does not bind is replaced by one that is less extractive towards the low skilled when this constraint binds. A binding minimum-utility constraint moves the marginal wage subsidies below the bridge identified in our companion article closer to the positive marginal tax rates found with the maxi-min schedule. For those individuals near the bottom of the skill distribution, this countervailing force is sufficiently strong that they now face positive marginal taxes.

Our model differs from that of Bohn and Stuart (2013) by supposing that preferences are quasilinear in consumption. This preference restriction is what permits us to provide a complete characterization of the selfishly optimal income schedules and to show that individuals have single-peaked preferences over them. Bohn and Stuart only provide partial characterizations of their selfishly optimal schedules and do not appeal to single-peakedness to show that there is a Condorcet winner.

While imposing a minimum-utility constraint is a reasonable way to introduce a concern for the low skilled, and thereby to limit the extent to which they are disadvantaged through income taxation, it is not the only way. One could instead imagine that individuals care not just for their own utility, but also for the utility of the least fortunate. Our analysis can be easily extended to model the possibility that individuals have objective functions of the form

$$V(k) + \lambda V(\underline{w}) \tag{34}$$

but do not face a minimum-utility constraint. Indeed, the first-order conditions for the problem faced by such an individual are exactly (15). However, λ is fixed in this variant of our model and so the analysis of our companion paper, Brett and Weymark (2016), is needed to characterize the voting equilibrium, even though the outcome is qualitatively similar to the voting equilibrium described here.[11]

In general, without some restrictions on the set of feasible tax instruments, no voting equilibrium exists. While we restrict voting to the set of selfishly optimal income tax schedules, we place no *a priori* restrictions on the form of the tax functions themselves. In so doing, we have uncovered the features of a Condorcet-winning tax system that would carry out redistribution toward the middle of the income distribution. Actual tax systems might deviate from this benchmark due to other restrictions on the set of feasible policies or because of the details of political institutions from which we have abstracted.

[11] Boadway, Brett, and Jacquet (2015) provide a detailed description of optimal marginal tax rates for an objective function similar to (34) that arises in a normative tax model in which individuals differ in both skills and preferences for leisure.

Appendix

Proof of Proposition 1. By (8),

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

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

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

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

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

On the other hand, by (4),

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

As we have shown, it is optimal for the government budget constraint (5) to bind. Substituting the equality form of this constraint into (A.4) yields

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

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

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

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 type k's relaxed problem except for the minimum-utility constraint. Substituting (A.6) into the minimum-utility constraint (6) yields the constraint in (12). $\square$

Proof of Proposition 2. It is convenient to solve type k's reduced-form problem in two steps. In the first step, optimal incomes are chosen for fixed values of the multiplier λ associated with the minimum-utility constraint and the bridge endpoints w_b and w_B. In the second step, the optimal values of w_b, w_B, and λ are determined using the parameterized schedules obtained in the first step.

Fix λ, w_b, and w_B. Outside of the bridge, the before-tax income schedule satisfies the first-order conditions (15) and, therefore, the optimal income for any type not on the bridge does not depend on w_b or w_B. Let $\tilde{y}(w, \lambda)$ denote the solution to these first-order conditions for a fixed value of λ. For $w > w_B$, $\tilde{y}(w, \lambda) = y^R(w)$, which does not depend on λ. Let $\bar{y}$ denote the before-tax income allocated to everyone on the bridge.

The Lagrangian associated with the second step is

$$\begin{aligned}\int_{\underline{w}}^{w_b} G^M(w, \tilde{y}(w,\lambda))dw + \int_{w_b}^{k} G^M(w, \bar{y})dw + \int_{k}^{w_B} G^R(w, \bar{y})dw + \int_{w_B}^{\bar{w}} G^R(w, y^R(w))dw \\ + \lambda \biggl(\int_{\underline{w}}^{w_b} G^R(w, \tilde{y}(w, \lambda))dw + \int_{w_b}^{k} G^R(w, \bar{y})dw \\ + \int_{k}^{w_B} G^R(w, \bar{y})dw + \int_{w_B}^{\bar{w}} G^R(w, y^R(w))dw - u_0 \biggr). \end{aligned} \tag{A.7}$$

There are three cases.

Case 1: $\underline{w} < w_b < w_B < \bar{w}$. The assumed continuity of the optimal income schedule implies that

$$\bar{y} = \tilde{y}(w_b, \lambda) = y^R(w_B). \tag{A.8}$$

Hence, the choice of w_b, w_B, and λ is determined by simultaneously solving

$$\tilde{y}(w_b, \lambda) - y^R(w_B) = 0, \tag{A.9}$$

$$\lambda \left[\int_{\underline{w}}^{w_b} G^R(w, \tilde{y}(w, \lambda))dw + \int_{w_b}^{w_B} G^R(w, y^R(w_B))dw + \int_{w_B}^{\bar{w}} G^R(w, y^R(w))dw - u_0 \right] = 0, \tag{A.10}$$

and

$$\int_{w_b}^{k} \theta^M(w, y^R(w_B))dw + \lambda \int_{w_b}^{k} \theta^R(w, y^R(w_B))dw + (1 + \lambda) \int_{k}^{w_B} \theta^R(w, y^R(w_B))dw = 0. \tag{A.11}$$

Equation (A.9) is the requirement that the before-tax income function be continuous at the two endpoints of the bridge. Equation (A.10) is the complementary slackness condition for the minimum-utility constraint with the value of the before-tax income function substituted therein. Equation (A.11) is the first-order condition for the placement of the bridge endpoints. Its left-hand side is the derivative of the Lagrangian (A.7) with respect to the level of before-tax income on the bridge evaluated at $y^R(w_B)$.

When λ is chosen optimally, by definition, $\tilde{y}(w,\lambda) = y^0(w)$ for $w < w_b$. We have thus shown that type k's optimal income schedule $y(\cdot)$ is given by (24). Observe that (A.11) is equivalent to (27) and that (25) and (A.10) are equivalent for the optimal income schedule for this case.

Case 2: $\underline{w} < w_b < w_B = \bar{w}$. In this case, the value of before-tax income for the types on the bridge is not, as in Case 1, $y^R(w_B)$. Instead, the types in $[w_b, \bar{w}]$ receive $\tilde{y}(w_b, \lambda)$. The two variables w_b and λ are determined by the following modified versions of (A.10) and (A.11):

$$\lambda \left[\int_{\underline{w}}^{w_b} G^R(w, \tilde{y}(w,\lambda))dw + \int_{w_b}^{\bar{w}} G^R(w, \tilde{y}(w_b,\lambda))dw - u_0 \right] = 0 \tag{A.12}$$

and

$$\int_{w_b}^{k} \theta^M(w, \tilde{y}(w_b,\lambda))dw + \lambda \int_{w_b}^{k} \theta^R(w, \tilde{y}(w_b,\lambda))dw + (1+\lambda) \int_{k}^{\bar{w}} \theta^R(w, \tilde{y}(w_b,\lambda))dw = 0. \tag{A.13}$$

Noting that $\tilde{y}(w,\lambda) = y^0(w)$ for $w \leq w_b$, the proof for this case is completed as in Case 1 by observing that (A.13) is equivalent to (26).

Case 3: $\underline{w} = w_b < w_B < \bar{w}$. In this case, the value of before-tax income for the types on the bridge is $y^R(w_B)$, as in Case 1. The two variables w_B and λ are determined by solving

$$\lambda \left[\int_{\underline{w}}^{w_B} G^R(w, y^R(w_B))dw + \int_{w_B}^{\bar{w}} G^R(w, y^R(w))dw - u_0 \right] = 0 \tag{A.14}$$

and (A.11). The proof for this case is completed as in the previous two cases. □

Proof of Proposition 3. Recall that $\tilde{y}(w,\lambda)$ is the solution to the first-order conditions (15) for a fixed value of λ. Outside of the bridge on the income schedule proposed by type k, $y(w,k) = \tilde{y}(w,\lambda(k))$. It follows from (15), (19), and the Implicit Function Theorem that for all $k \in [\underline{w}, \bar{w}]$,

$$\frac{\partial \tilde{y}(w,\lambda(k))}{\partial \lambda} = -\frac{\theta^R(w, y(w,k))}{\theta^M_y(w, y(w,k)) + \lambda \theta^R_y(w, y(w,k))} < 0, \quad \forall w < w_b(k) \tag{A.15}$$

and

$$\frac{\partial \tilde{y}(w,\lambda(k))}{\partial \lambda} = \frac{\partial y^R(w,k)}{\partial \lambda} = 0, \quad \forall w > w_B(k). \tag{A.16}$$

The inequality in (A.15) follows from the observations immediately below (20).

The same three cases arise as in the proof of Proposition 2.

Case 1: $\underline{w} < w_b < w_B < \bar{w}$. We note that the left-hand sides of (A.9)–(A.11) are functions of w_b, w_B, λ, and k. We call these functions $\Upsilon(w_b, w_B, \lambda, k)$, $\Psi(w_b, w_B, \lambda, k)$, and $\Phi(w_b, w_B, \lambda, k)$, respectively. Let $\Psi^*(w_b, w_B, \lambda, k)$ denote the term in square brackets on the left-hand side of (A.10). We employ the Implicit Function Theorem to determine how the endogenous variables in (A.9)–(A.11) respond to a change in k. Implicit differentiation of these three equations yields

$$\begin{bmatrix} \Upsilon_{w_b} & \Upsilon_{w_B} & \Upsilon_\lambda \\ \Psi_{w_b} & \Psi_{w_B} & \Psi_\lambda \\ \Phi_{w_b} & \Phi_{w_B} & \Phi_\lambda \end{bmatrix} \begin{bmatrix} dw_b \\ dw_B \\ d\lambda \end{bmatrix} = \begin{bmatrix} -\Upsilon_k \\ -\Psi_k \\ -\Phi_k \end{bmatrix} dk. \tag{A.17}$$

We now compute the entries in (A.17) and evaluate them at the solution to type k's problem. From (A.9),

$$\Upsilon_{w_b} = \frac{\partial \tilde{y}(w_b(k), \lambda(k))}{\partial w_b} > 0; \tag{A.18}$$

$$\Upsilon_{w_B} = -\frac{\partial y^R(w_B(k))}{\partial w_B} < 0; \tag{A.19}$$

$$\Upsilon_\lambda = \frac{\partial \tilde{y}(w_b(k), \lambda(k))}{\partial \lambda} < 0; \tag{A.20}$$

$$\Upsilon_k = 0. \tag{A.21}$$

The inequalities in (A.18) and (A.19) hold because an optimal income schedule is increasing outside of the bridge. Because of the continuity of an optimal income schedule, the equality in (A.15) also holds for $w = w_b(k)$. The inequality in (A.20) follows from this observation and the argument used to establish the inequality (A.15).

From (A.10),

$$\Psi_{w_b} = \lambda(k)[G^R(w_b(k), \tilde{y}(w_b(k), \lambda(k))) - G^R(w_b(k), y^R(w_B(k)))] = 0; \tag{A.22}$$

$$\begin{aligned} \Psi_{w_B} &= \lambda(k)\Bigg[G^R(w_B(k), y^R(w_B(k))) \\ &\qquad + \int_{w_b(k)}^{w_B(k)} \theta^R(w, y^R(w_B(k)))dw - G^R(w_b(k), y^R(w_B(k)))\Bigg] \\ &= \lambda(k) \int_{w_b(k)}^{w_B(k)} \theta^R(w, y^R(w_B(k)))dw \begin{cases} < 0 & \text{if } \lambda(k) > 0 \\ = 0 & \text{if } \lambda(k) = 0 \end{cases}; \end{aligned} \tag{A.23}$$

$$\Psi_\lambda = \lambda(k) \int_{\underline{w}}^{w_b(k)} \theta^R(w, \tilde{y}(w, \lambda(k))) \frac{\partial \tilde{y}(w, \lambda(k))}{\partial \lambda} dw + \Psi^*(w_b, w_B, \lambda, k) > 0; \tag{A.24}$$

$$\Psi_k = 0. \tag{A.25}$$

The second equality in (A.22) follows from (A.9). When $\lambda(k) > 0$, the inequalities in (A.23) and (A.24) follow from the fact that $\theta^R(y, w) < 0$ for all (y, w) above the $y^R(w)$

locus and, in the latter case, from (A.10) and (A.15). When $\lambda(k) = 0$, the corresponding equality in (A.23) trivially holds and the inequality in (A.24) follows because then $\Psi^*(w_b, w_B, \lambda, k) > 0$.[12]

From (A.11),

$$\Phi_{w_b} = -\theta^M(w_b(k), y^R(w_B(k))) - \lambda(k)\theta^R(w_b(k), y^R(w_B(k))) = 0; \tag{A.26}$$

$$\begin{aligned}\Phi_{w_B} = \Bigg(&\int_{w_b(k)}^{k} \left[\theta_y^M(w, y^R(w_B(k))) + \lambda(k)\theta_y^R(w, y^R(w_B(k)))\right] dw \\ &+ (1+\lambda(k)) \int_{k}^{w_B(k)} \theta_y^R(w, y^R(w_B(k)))dw \Bigg) \frac{\partial y^R(w_B(k))}{\partial w_B} < 0;\end{aligned} \tag{A.27}$$

$$\Phi_\lambda = \int_{w_b(k)}^{w_B(k)} \theta^R(w, y^R(w_B(k)))dw = \Psi_{w_B} \begin{cases} < 0 & \text{if } \lambda(k) > 0 \\ = 0 & \text{if } \lambda(k) = 0 \end{cases}; \tag{A.28}$$

$$\Phi_k = \theta^M(k, y^R(w_B(k))) - \theta^R(k, y^R(w_B(k))) > 0. \tag{A.29}$$

The second equality in (A.26) follows from (15) and (A.9). The inequality in (A.27) follows from the second-order condition (19) and the increasingness of an optimal income schedule outside of the bridge. The sign of Φ_λ in (A.28) follows from (A.23). The inequality in (A.29) follows from (16) and (17).

Let A be the matrix on the left-hand side of (A.17). The determinant $|A|$ is given by

$$|A| = \Upsilon_{w_b}\left[\Psi_{w_B}\Phi_\lambda - \Psi_\lambda \Phi_{w_B}\right] = (+)\left[(\leq 0)(\leq 0) - (+)(-)\right] > 0. \tag{A.30}$$

The sign of $|A|$ in (A.30) follows from (A.18), (A.23), (A.28), (A.24), and (A.27).

Now, by Cramer's Rule, and making use of (A.18)–(A.30),

$$\begin{aligned}\frac{dw_b(k)}{dk} &= \frac{\begin{vmatrix} 0 & \Upsilon_{w_B} & \Upsilon_\lambda \\ 0 & \Psi_{w_B} & \Psi_\lambda \\ -\Phi_k & \Phi_{w_B} & \Phi_\lambda \end{vmatrix}}{|A|} = \frac{-\Phi_k[\Upsilon_{w_B}\Psi_\lambda - \Upsilon_\lambda \Psi_{w_B}]}{|A|} \\ &= \frac{-(+)\overbrace{[(-)(+) - (-)(\leq 0)]}^{(-)}}{(+)} > 0;\end{aligned} \tag{A.31}$$

$$\frac{dw_B(k)}{dk} = \frac{\begin{vmatrix} \Upsilon_{w_b} & 0 & \Upsilon_\lambda \\ 0 & 0 & \Psi_\lambda \\ 0 & -\Phi_k & \Phi_\lambda \end{vmatrix}}{|A|} = \frac{\Phi_k \Upsilon_{w_b} \Psi_\lambda}{|A|} = \frac{(+)(+)(+)}{(+)} > 0; \tag{A.32}$$

[12] Strictly speaking, it is possible to have $\lambda(k) = \Psi^*(w_b, w_B, \lambda, k) = 0$. We ignore this knife-edge case because once Proposition 3 has been established for k for which this is not the case, continuity can be used to establish the proposition for this case as well.

$$\frac{d\lambda(k)}{dk} = \frac{\begin{vmatrix} \Upsilon_{w_b} & \Upsilon_{w_B} & 0 \\ 0 & \Psi_{w_B} & 0 \\ 0 & \Phi_{w_B} & -\Phi_k \end{vmatrix}}{|A|} = \frac{-\Phi_k \Upsilon_{w_b} \Psi_{w_B}}{|A|} = \frac{-(+)(+)(-)}{(+)} > 0 \text{ if } \lambda(k) > 0\ . \quad \text{(A.33)}$$

This completes the proof for the first case.

Case 2: $\underline{w} < w_b < w_B = \bar{w}$. Let $\overline{\Psi}(w_b, \lambda, k)$ and $\overline{\Phi}(w_b, \lambda, k)$ denote the expressions on the left-hand sides of (A.12) and (A.13), respectively, and let $\overline{\Psi}^*(w_b, \lambda, k)$ denote the term in square brackets in (A.12). The calculations needed to apply the Implicit Function Theorem are as follows:

$$\begin{aligned} \overline{\Psi}_{w_b} = \lambda(k)\Bigg[& G^R(w_b(k), \tilde{y}(w_b(k), \lambda(k))) - G^R(w_b(k), \tilde{y}(w_b(k), \lambda(k))) \\ & + \left(\int_{w_b(k)}^{\bar{w}} \theta^R(w, \tilde{y}(w_b(k), \lambda(k)))dw \right) \frac{\partial \tilde{y}(w_b(k), \lambda(k))}{\partial w_b} \Bigg] \begin{cases} < 0 & \text{if } \lambda(k) > 0 \\ = 0 & \text{if } \lambda(k) = 0 \end{cases}; \end{aligned} \quad \text{(A.34)}$$

$$\begin{aligned} \overline{\Psi}_\lambda = \lambda(k)\Bigg[& \int_{\underline{w}}^{w_b(k)} \theta^R(w, \tilde{y}(w_b(k), \lambda(k))) \frac{\partial \tilde{y}(w, \lambda(k))}{\partial \lambda} dw \\ & + \int_{w_b(k)}^{\bar{w}} \theta^R(w, \tilde{y}(w_b(k), \lambda(k))) \frac{\partial \tilde{y}(w_b(k), \lambda(k))}{\partial \lambda} dw \Bigg] + \overline{\Psi}^*(w_b, \lambda, k) > 0; \end{aligned} \quad \text{(A.35)}$$

$$\overline{\Psi}_k = 0; \quad \text{(A.36)}$$

$$\begin{aligned} \overline{\Phi}_{w_b} = & -\theta^M(w_b(k), \tilde{y}(w_b(k), \lambda(k))) - \lambda(k)\theta^R(w_b(k), \tilde{y}(w_b(k), \lambda(k))) \\ & + \Bigg(\int_{w_b(k)}^{k} \left[\theta_y^M(w, \tilde{y}(w_b(k), \lambda(k))) + \lambda(k)\theta_y^R(w, \tilde{y}(w_b(k), \lambda(k))) \right] dw \\ & + (1 + \lambda(k)) \int_k^{\bar{w}} \theta_y^R(w, \tilde{y}(w_b(k), \lambda(k)))dw \Bigg) \frac{\partial \tilde{y}(w_b(k), \lambda(k))}{w_b} < 0; \end{aligned} \quad \text{(A.37)}$$

$$\overline{\Phi}_\lambda = \int_{w_b(k)}^{\bar{w}} \theta^R(w, \tilde{y}(w_b(k), \lambda)(k)) < 0; \quad \text{(A.38)}$$

$$\begin{aligned} \overline{\Phi}_k &= \theta^M(k, \tilde{y}(w_b(k), \lambda(k))) + \lambda(k)\theta^R(k, \tilde{y}(w_b(k), \lambda(k))) - (1 + \lambda(k))\theta^R(k, \tilde{y}(w_b(k), \lambda(k))) \\ &= \theta^M(k, \tilde{y}(w_b(k), \lambda(k))) - \theta^R(k, \tilde{y}(w_b(k), \lambda(k))) > 0. \end{aligned} \quad \text{(A.39)}$$

The arguments establishing the signs of these derivatives are very similar to those used in the previous case. One notable additional argument is that the top line on the right-hand side of (A.37) vanishes by (15).

Using notation and a style of argument similar to that used in the proof of Case 1, we consider the matrix equation:

$$\begin{bmatrix} \overline{\Psi}_{w_b} & \overline{\Psi}_\lambda \\ \overline{\Phi}_{w_b} & \overline{\Phi}_\lambda \end{bmatrix} \begin{bmatrix} dw_b \\ d\lambda \end{bmatrix} = \begin{bmatrix} 0 \\ -\overline{\Phi}_k \end{bmatrix} dk. \quad \text{(A.40)}$$

Using (A.34)–(A.39) together with (A.40), we conclude that

$$|\bar{A}| = \overline{\Psi}_{w_b}\overline{\Phi}_\lambda - \overline{\Psi}_\lambda\overline{\Phi}_{w_b} = (\leq 0)(-) - (+)(-) > 0; \tag{A.41}$$

$$\frac{dw_b(k)}{dk} = \frac{\begin{vmatrix} 0 & \overline{\Psi}_\lambda \\ -\overline{\Phi}_k & \overline{\Phi}_\lambda \end{vmatrix}}{|\bar{A}|} = \frac{\overline{\Psi}_\lambda\overline{\Phi}_k}{|\bar{A}|} = \frac{(+)(+)}{(+)} > 0; \tag{A.42}$$

$$\frac{d\lambda(k)}{dk} = \frac{\begin{vmatrix} \overline{\Psi}_{w_b} & 0 \\ \overline{\Phi}_{w_b} & -\overline{\Phi}_k \end{vmatrix}}{|\bar{A}|} = \frac{-\overline{\Psi}_{w_b}\overline{\Phi}_k}{|\bar{A}|} = \frac{-(-)(+)}{(+)} > 0 \text{ if } \lambda(k) > 0 \ . \tag{A.43}$$

It remains to show that $w_B(k)$ remains at $\bar{w}$ if k is marginally increased. If, on the contrary, it decreased, this would contradict the finding in Case 1 that $w_B(\cdot)$ is increasing at k when $\underline{w} < w_b < w_B < \bar{w}$. This completes the proof for the second case.

Case 3: $\underline{w} = w_b < w_B < \bar{w}$. This case is similar to Case 1 except that (A.9) is not used and the functions Φ and Ψ are evaluated with $w_b = \underline{w}$. Notably, this results in λ not appearing in the term in square brackets in the Ψ function, so $\Psi_\lambda = \Psi^*(w_b, w_B, \lambda, k)$. The term Υ_{w_b} disappears from the analogs to (A.30), (A.32), and (A.33), but all other calculations remain essentially unchanged from Case 1. Specifically,

$$|A| = \Psi_{w_B}\Phi_\lambda - \Psi_\lambda\Phi_{w_B} > 0; \tag{A.44}$$

$$\frac{dw_B(k)}{dk} = \frac{\Phi_k\Psi_\lambda}{|A|} = \frac{(+)(0)}{(+)} = 0 \text{ if } \lambda(k) > 0 \ ; \tag{A.45}$$

$$\frac{dw_B(k)}{dk} = \frac{\Phi_k\Psi_\lambda}{|A|} = \frac{(+)(+)}{(+)} > 0 \text{ if } \lambda(k) = 0 \ ; \tag{A.46}$$

$$\frac{d\lambda(k)}{dk} = \frac{-\Phi_k\Psi_{w_B}}{|A|} = \frac{-(+)(-)}{(+)} > 0 \text{ if } \lambda(k) > 0 \ . \tag{A.47}$$

This completes the proof for the third case. Proposition 3 follows from what has been established in the three cases. □

Proof of Proposition 4. We need to show that the function $V^0(w, \cdot)$ is single-peaked. That is, we need to show that $V^0(w,k)$ is nonincreasing in k for $k > w$ and nondecreasing in k for $k < w$. This is done by determining the sign of the right-hand side of (33). We only need to consider the case in which $\lambda(k) > 0$ because single-peakedness for the case in which the the minimum-utility constraint does not bind has been established by Brett and Weymark (2016, Proposition 8). Recall that $k \in (w_b(k), w_B(k))$ (that is, type k is on the bridge of his selfishly optimal schedule) and that $\tilde{y}(w,\lambda)$ is the solution to the first-order conditions (15) for a fixed value of λ. As in the preceding two proofs, there are three cases to consider.

Case 1: $\underline{w} < w_b(k) < w_B(k) < \bar{w}$. We make use of the following inequalities:

$$\frac{\partial y(w,k)}{\partial k} = \frac{\partial \tilde{y}(w,\lambda(k))}{\partial \lambda}\frac{d\lambda(k)}{dk} < 0, \quad \forall w < w_b(k); \tag{A.48}$$

$$\frac{\partial y(w,k)}{\partial k} = \frac{\partial y^R(w_B(k))}{\partial w_B}\frac{dw_B(k)}{dk} > 0, \quad \forall w \in (w_b(k), w_B(k)); \tag{A.49}$$

$$\frac{\partial y(w,k)}{\partial k} = \frac{\partial y^R(w)}{\partial k} = 0, \quad \forall w > w_B(k); \tag{A.50}$$

The first of these inequalities follows from (A.15) and Proposition 3. The other two follow from (24) and, in the case of (A.49), from Part 2 of Proposition 3. The expression in the middle of (A.49) is the change in the income of types on the bridge due to a change in k, which is the same for all types in $(w_b(k), w_B(k))$.

(a) We first suppose that $k > w$ and show that $V^0(w,k)$ is decreasing in k. By (16), (17), and (33),

$$\frac{\partial V^0(w,k)}{\partial k} = \int_{\underline{w}}^{w} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt. \tag{A.51}$$

The term in square brackets (A.51) is equal to the term in square brackets in (33), which is positive. Thus, if $w \leq w_b(k)$, (A.48) implies that the integrand in (A.51) is everywhere negative. Therefore, $\partial V^0(w,k)/\partial k < 0$ when $k > w$ and $w \leq w_b(k)$.

On the other hand, if $w_b(k) < w < k$, substituting (15) into (A.51) yields

$$\begin{aligned}
\frac{\partial V^0(w,k)}{\partial k} &= -(1+\lambda(k))\int_{\underline{w}}^{w_b(k)} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\
&+ \int_{w_b(k)}^{k} \theta^M(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt - \int_{w_b(k)}^{k} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\
&- \int_{w}^{k} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.
\end{aligned} \tag{A.52}$$

Using the first-order condition (A.11) for the placement of the bridge endpoints and the observation that $\partial y(t,k)/\partial k$ is a constant for all $t \in (w_b, w_B)$ to re-express the second term on the right-hand side of (A.52), we obtain

$$\begin{aligned}
\frac{\partial V^0(w,k)}{\partial k} &= -(1+\lambda(k))\int_{\underline{w}}^{w_b(k)} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\
&-\lambda(k)\int_{w_b(k)}^{k} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt - (1+\lambda(k))\int_{k}^{w_B(k)} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\
&- \int_{w_b(k)}^{k} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt - \int_{w}^{k} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.
\end{aligned} \tag{A.53}$$

Collecting terms in (A.53), this equality reduces to

$$\begin{aligned}
\frac{\partial V^0(w,k)}{\partial k} &= -(1+\lambda(k))\int_{\underline{w}}^{w_B(k)} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\
&- \int_{w}^{k} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.
\end{aligned} \tag{A.54}$$

Differentiating the minimum utility-constraint (A.10) with respect to k and using (A.50) and the definition of $\theta^R(w,y)$ in (17), it follows that the first term on the right-hand side of (A.54) vanishes. By (A.49), the integrand in the second term of the right-hand side of (A.54) is always positive. Hence, $\partial V^0(w,k)/\partial k < 0$ when $w_b(k) < w < k$. We have thus shown that $V^0(w,k)$ is decreasing in k when $k > w$.

(b) We now show that $k < w$ implies that $V^0(w,k)$ is increasing in k. By (33),

$$\begin{aligned}\frac{\partial V^0(w,k)}{\partial k} = &\int_{\underline{w}}^{w_b(k)} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt \\ &+ \int_{w_b(k)}^{w_B(k)} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt \\ &+ \int_{w_B(k)}^{w} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.\end{aligned} \tag{A.55}$$

Using (15) in the first term on the right-hand side of (A.55) and rearranging yields

$$\begin{aligned}\frac{\partial V^0(w,k)}{\partial k} = &-(1+\lambda(k))\int_{\underline{w}}^{w_b(k)} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\ &+ \int_{w_b(k)}^{k} \theta^M(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt - \int_{w_b(k)}^{k} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\ &+ \int_{k}^{w_B(k)} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt \\ &+ \int_{w_B(k)}^{w} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.\end{aligned} \tag{A.56}$$

Using (A.11) in the second term on the right-hand side of (A.56) and our previous observation that the expression on the right-hand side of the equality in (A.49) is constant on the interval (w_b, w_B) yields, after some re-grouping of terms,

$$\begin{aligned}\frac{\partial V^0(w,k)}{\partial k} = &-(1+\lambda(k))\int_{\underline{w}}^{w_B(k)} \theta^R(t,y(t,k))\frac{\partial y(t,k)}{\partial k}dt \\ &+ \int_{k}^{w} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.\end{aligned} \tag{A.57}$$

As in the case of $k > w$, the first term on the right-hand side of (A.57) vanishes, so that

$$\frac{\partial V^0(w,k)}{\partial k} = \int_{k}^{w} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt. \tag{A.58}$$

If $w < w_B(k)$, then the integrand on the right-hand side of (A.58) is always positive by

(A.49). If $w > w_B(k)$, we can rewrite (A.58) as

$$\begin{aligned}\frac{\partial V^0(w,k)}{\partial k} = \int_k^{w_B(k)} & [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt \\ & + \int_{w_B(k)}^{w} [\theta^M(t,y(t,k)) - \theta^R(t,y(t,k))]\frac{\partial y(t,k)}{\partial k}dt.\end{aligned} \tag{A.59}$$

The first-term on the right-hand side of (A.59) is positive by (A.49) and the second term is zero by (A.50). Thus, the right-hand side of (A.58) is positive for all $w \neq w_B$ for which $k < w$. By continuity, the same conclusion also holds when $w = w_b$. Thus, $\partial V^0(w,k)/\partial k > 0$ when $k < w$. That is, $V^0(w,k)$ is increasing in k for $k < w$, which completes the proof that $V^0(w,\cdot)$ is single-peaked for the case in which $\underline{w} < w_b(k) < w_B(k) < \bar{w}$.

Case 2: $\underline{w} < w_b(k) < w_B(k) = \bar{w}$. For all $w < w_b(k)$, (A.48) applies. The income on the bridge is now $\tilde{y}(w_b(k),\lambda(k))$. Hence,

$$\frac{\partial \tilde{y}(w_b(k),\lambda(k))}{\partial k} = \frac{\partial \tilde{y}(w_b(k),\lambda(k))}{\partial w_b}\frac{dw_b(k)}{dk} + \frac{\partial \tilde{y}(w_b(k),\lambda(k))}{\partial \lambda}\frac{d\lambda(k)}{dk}. \tag{A.60}$$

Using (A.42) and (A.43) in (A.60), upon rearranging terms we obtain

$$\frac{\partial \tilde{y}(w_b(k),\lambda(k))}{\partial k} = \frac{\overline{\Phi}_k}{|\bar{A}|}\left[\frac{\partial \tilde{y}}{\partial w_b}\overline{\Psi}_\lambda - \frac{\partial \tilde{y}}{\partial \lambda}\overline{\Psi}_{w_b}\right] = \frac{(+)}{(+)}[(+)(+) - (\leq 0)(-)] > 0, \tag{A.61}$$

where we have suppressed the arguments of the terms on the right-hand side of this equation. The sign of $\partial \tilde{y}(w_b(k),\lambda(k))/\partial w_b$ is positive because when w_b increases, the original lower endpoint of the bridge now has its income determined by the $y^0(\cdot)$ schedule, which is increasing. The sign of $\partial \tilde{y}(w_b(k),\lambda(k))/\partial \lambda$ follows from (A.15) and the continuity of $y^0(\cdot)$. The other signs have been established in the proof of Proposition 3.

The rest of the proof for this case is almost identical to the proof in Case 1. Instead of appealing to (A.10) and (A.11), appeal is made to (A.12) and (A.13). Also, (A.61) is used to show that $\partial y(t,k)/\partial k > 0$ on the bridge in (A.54), (A.58), and (A.59) rather than (A.49).

Case 3: $\underline{w} = w_b(k) < w_B(k) < \bar{w}$. Only minor modifications are needed to the proof of Case 1. The terms in which $w < w_b(k)$ drop out as they are no longer relevant. The inequality in (A.49) is now an equality, so $\partial V^0(w,k)/\partial k = 0$ in (A.54), (A.58), and (A.59). Thus, a marginal increase in k does not affect the utility of any type w. □

References

Bierbrauer, F., 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.

Boadway, R., Brett, C., Jacquet, L., 2015. Optimal nonlinear income taxes with compensation. Working Paper No. 2015-15, THEMA, University of Cergy-Pontoise.

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

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., 2016. Voting over selfishly optimal nonlinear income tax schedules. Games and Economic Behavior, forthcoming.

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

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.

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.

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.

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.

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

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.

Nöldeke, G., Samuelson, L., 2007. Optimal bunching without optimal control. Journal of Economic Theory 134, 405–420.

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.
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.
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.
Sundaram, R. K., 1996. A First Course in Optimization Theory. Cambridge University Press, Cambridge.
Vincent, T. L., Mason, J. D., 1968. Disconnected optimal trajectories. Journal of Optimization Theory and Applications 3, 263–281.