Citizen Candidates and Voting Over Incentive-Compatible Nonlinear Income Tax Schedules

Craig Brett\textsuperscript{a}, John A. Weymark\textsuperscript{b}

\begin{itemize}
\item \textsuperscript{a}Department of Economics, Mount Allison University, 144 Main Street, Sackville NB E4L 1A7, Canada. E-mail: cbrett@mta.ca
\item \textsuperscript{b}Department of Economics, Vanderbilt University, VU Station B #351819, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA. E-mail: john.weymark@vanderbilt.edu
\end{itemize}

March 2014

Preliminary version. Please do not cite.

\textbf{Abstract.} Majority voting over the nonlinear tax schedules proposed by a continuum of citizen candidates is considered. The analysis extends the finite-individual model of Röell (unpublished manuscript, 2012). Each candidate proposes the tax schedule that is utility maximal for him subject to budget and incentive constraints. Each of these schedules is a combination of the maxi-min and maxi-max schedules along with a region of bunching in a neighborhood of the proposer’s type. Techniques introduced by Vincent and Mason (1967, NASA Contractor Report CR-744) are used to indentify the bunching region. As in Röell’s model, it is shown that individual preferences over these schedules is single-peaked, so the median voter theorem applies. In the majority rule equilibrium, marginal tax rates are typically negative for low-skilled individuals and positive for high-skilled individuals. When a minimum-utility constraint is added, the maxi-max region is modified in a way that increases the marginal tax rates for the low skilled.

\textit{Journal of Economic Literature} classification numbers. D72, D82, H21

\textit{Keywords.} bunching, citizen candidates, ironing, majority voting, nonlinear income taxation
1. Introduction

Concern over widening income inequality and the struggles of the middle class can lead to calls for higher taxes rates on the very rich. These calls highlight two features of income taxation that we wish to investigate in this paper. First, it is possible to vary marginal tax rates with income; that is, income taxes are nonlinear. Second, the shape of the income tax function is a matter of political discussion and, ultimately, determined by political processes. The design of the tax system must weigh the concerns of policy makers, be they motivated by politics or by a benevolent concern for social welfare, against the work incentives and induced levels of labor supply implied by the tax system.

In this paper, we consider a model economy like the one described in Mirrlees (1971). Workers differ in their respective skill levels and, for reasons of information asymmetry, the policy maker is constrained to set an anonymous income tax schedule. Workers then decide on their respective intensities of labor supply, which, in turn, determines their before- and after-tax incomes. While Mirrlees and much of the subsequent literature on optimal nonlinear income taxes study the choice of a benevolent policy maker among all possible anonymous income tax schedules, we aim to model which of these schedules arises as the outcome of a political process.

Our model of politics is the citizen-candidate framework of Osborne and Slivinski (1996). Individuals decide whether to run for election. The winner of the election is then able to enact his most favored policy. Thus, the elected candidate selects the tax schedule that maximizes his own utility. Voters then choose their preferred candidate, fully aware of the candidates’ intentions. We assume that entry of candidates is costless, so that the winning candidate, if there is one, is a Condorcet winner. We show that a candidate of the median skill type is a Condorcet winner. Thus, the tax schedule that maximizes the utility of the median type emerges as the outcome of the political process.

A voter’s assessment of a candidate is derived from the utility he can gain when facing the tax schedule that candidate would choose if elected. Thus, a ranking of candidates is tantamount to a ranking of tax schedules. It is, therefore, necessary to describe the preferred tax schedule of each potential candidate as part of the development of the theory of political equilibrium. Any successful candidate will wish to redistribute income from all other types toward his own type. For skill-types above the winner, the desire is to redistribute income downward, and the part of the tax schedule these types face is qualitatively similar to the optimal nonlinear tax schemes in most normative models. In particular, the highly skilled face positive (or zero, in the case of the very highest type should one exist) marginal income tax rates. But the winner also wishes to redistribute income upward from types below his own. This redistribution is constrained by upward incentive-compatibility constraints and gives rise to negative (or zero for the very lowest type, in some circumstances) marginal income tax rates. The optimal tax schedule, therefore, features an abrupt change from negative to positive marginal tax rates near the winner’s type. As a result, there must be a kink in the optimal tax schedule and some bunching of types near the winner’s.
We assume that preferences are quasilinear in consumption. Under this assumption, the optimal allocation of before-tax income is independent of the optimal consumption allocation, making it possible to give a precise description of the before-tax income schedule a candidate would enact. We start by describing the optimal tax schedule in a relaxed problem in which the second-order incentive compatibility constraint is ignored. We show that for types greater than its own, the candidate optimally designs the before-tax income schedule that a social planner with a maxi-min utilities objective would design. Both the candidate and the maxi-min social planner wish to extract as much revenue as incentives allow from the higher types. Because the optimal before-tax income schedule does not depend on the distribution of consumption, the candidate’s wish to give that revenue to himself rather than to the least-skilled is of no consequence to the before-tax income schedule.

On the other hand, the candidate selects a before-tax income schedule that corresponds to the maxi-max utilities solution for types below its own. We show that, in contrast to the well-studied maxi-min solution [cf. Boadway and Jacquet (2008)], the maxi-max solution can be quite ill-behaved. In particular, the second-order conditions for the candidate’s optimization problem can fail, so that the optimal schedule induces a range of individuals to supply the maximum possible amount of labor. The reason for this result is that a utility-compensated increase in before-tax income for a type lower than the candidate’s serves to slacken the binding upward incentive constraints, allowing the candidate to extract income from still lower types. Because the mass of these lower types increases with the candidate’s type, the second-order conditions are more often violated for candidates of a relatively high type.

Even when the maxi-max solution is well-behaved, the relaxed solution cannot be the solution to the candidate’s problem. With quasilinear-in-consumption utilities, the abrupt change from negative to positive marginal tax rates at the candidate’s type requires a downward jump in before-tax income at the candidate’s type. Such a downward jump violates the second-order conditions for incentive compatibility (Lollivier and Rochet, 1983; Boadway et al., 2000). The full solution can be constructed by building a level “bridge” that connects the two components of the relaxed solution. The solution coincides with the maxi-max solution up to the point of entry onto the bridge, which is necessarily below the candidate’s type. The solution then leaves the bridge at some type above the candidate’s, whereupon it coincides with the maxi-min solution. The types “on the bridge” are exactly those that are optimally bunched with the candidate.

A voter compares candidates on the basis of the tax schedules they choose. That comparison is facilitated by a comparative static result on how the before-tax income schedule changes as the candidate’s type increases. By establishing that the ends of the bridge (bunching interval) move to the right as the candidate’s type increases, we

---

1For candidate types sufficiently close to the lowest, it is possible that the solution starts on the bridge, in which case it provides a case of bunching at the bottom similar to that studied by Ebert (1992). If the distribution of types is bounded above, it is possible to have bunching at the top, as well. This corresponds to the before-tax income schedule ending on the bridge.
show that no-one’s before-tax income decreases when the type of the winning candidate increases. We then show that this property of the candidates’ choices implies that voter preferences are single-peaked in the skill level of the candidate. Consequently, the median type is a Condorcet winner in the electoral competition.

This paper is closely related to work by Röell (2012). In an economy with a finite number of worker types, she derives conditions similar to ours under which voter preferences are single-peaked over the type of the winning candidate. She does not provide a detailed description of optimal tax schedules as we do here. Bohn and Stuart (2013) extend the Röell model to a continuum of types, again arriving at a median voter result. They impose a constraint on candidate behavior that we initially do not. Specifically, they require that candidates’ tax schedules guarantee a minimum utility level to the least-skilled type.\footnote{The Bohn-Stuart analysis also makes extensive use of an assumption about the curvature properties of an optimized value function. It is not clear what restrictions this assumption imposes on the primitives of the model.} We show that our analysis can be modified to allow for such a minimum-utility constraint, without drastically changing the basic structure of the optimal before-tax income schedule. Interestingly, the addition of such a constraint makes the assumption that the second-order conditions for a candidate’s optimum more plausible. The minimum-utility constraint induces a candidate to behave as if he is maximizing a weighted average of his own utility and the lowest-type utility. For types above the candidate, this has no effect on the optimal before-tax income schedule because the candidate continues to desire redistribution downward from these types. For types below the candidate, however, the weighted averaging pulls the before-tax income schedule toward the more well-behaved maxi-min schedule. Moreover, we show that the shadow value of the minimum-utility constraint increases in the type of the candidate, so that candidates of a relatively high type — exactly those that in our basic model are more prone to violations in second-order conditions — optimally choose schedules most shifted away from what they would choose in the absence of the minimum-utility constraint.

To the best of our knowledge, all other models of voting over nonlinear income taxes with variable labor supply are restricted to two-type worlds. Roemer (2012) derives the equilibrium positions of two partisan political parties that choose anonymous tax schedules for two types of workers. Bierbrauer and Boyer (2013) consider a similar problem under the assumption that parties wish to maximize votes, and they allow for the parties to differ in their efficiency in running government. Our analysis is complementary to these existing models in that it provides a look at a different political process.

The remainder of this paper is organized as follows. The next section provides an outline the model economy. Section 3 contains a detailed account of a candidate's choice of tax schedule. This is followed by an analysis of the voting equilibrium in Section 4. We discuss the implications of a minimum-utility constraint in Section 5. Section 6 provides concluding remarks. The proofs of our results are given in an Appendix.
2. The Model

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

$$y = wl,$$

where $l$ 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.

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

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

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

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

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

Individuals face an anonymous nonlinear income taxation schedule that specifies the tax paid as a function of income $T(y)$, subject to which individuals choose their most preferred combination of consumption and before-tax income (equivalently, after-tax income and labor supply). Admissible tax schedules are assumed to be piecewise continuously differentiable. By the taxation principle (see Hammond, 1979; Guesnerie, 1995), having individuals choose consumption and income subject to an anonymous tax schedule is equivalent to directly specifying these variables as functions of type subject to incentive-compatibility constraints. These schedules, $x(\cdot)$ and $y(\cdot)$, as well as the labor supply schedule $l(\cdot)$ corresponding to $y(\cdot)$, are also piecewise continuously differentiable.\(^3\)

Because there are no mass points in the distribution, $T(\cdot)$, $x(\cdot)$, $y(\cdot)$, and $l(\cdot)$ are all continuous (see Hellwig, 2010). The bundle allocated to individuals of type $w$ is thus $(y(w), x(w))$. The resulting utility level is

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

\(^3\)As shown by Hellwig (2010), it is only necessary to assume that these schedules are integrable. The stronger assumption of piecewise continuous differentiability is typically made to facilitate the use of standard control-theoretic arguments.
Incentive compatibility requires that

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

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

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

and the second-order condition is

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

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

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

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

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

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

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

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

---

4The expressions in (6) and (7) are required to hold at all points for which \( y(w) \) is differentiable. Because incentive compatibility implies that income is non-decreasing in type, \( y(w) \) is differentiable almost everywhere.
3. The Citizen Candidate’s Problem

Each individual proposes a tax schedule and then pairwise majority rule is used to choose which of these schedules is implemented. In this role, individuals are citizen candidates. Individuals of the same type propose the same tax schedule, so we can equivalently think of voting as taking place over types of individuals or over their proposed tax schedules.

A candidate has no power to commit to a tax policy. Therefore, if elected, a candidate selects the policy that maximizes his own utility subject to the resulting allocation satisfying the incentive-compatibility and government budget constraints. Voters know this, and this informs their choice of candidates when voting. The tax schedule chosen by a candidate is equivalent to the choice of an allocation schedule \((y(\cdot), x(\cdot))\) that specifies a bundle \((x(w), y(w))\) for each type \(w \in [w, \bar{w}]\).\(^5\) Formally, a candidate of type \(k\) solves

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

We refer to (10) as candidate \(k\)'s problem.

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

3.1. The First-Order Approach

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

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

We refer to (11) as candidate \(k\)'s relaxed problem.

By modifying the arguments found in Lollivier and Rochet (1983), we show in Proposition 1 that it is possible to formulate an unconstrained optimization problem that

\(^5\text{In this section, the identity of the candidate is fixed, so we do not index the schedules by the candidate’s type.}\)
provides the before-tax income schedule that solves candidate $k$’s relaxed problem. As is standard in a nonlinear income tax problem, it is optimal for the government budget constraint to bind. Once candidate $k$’s optimal before-tax income schedule has been determined, the corresponding consumption schedule can be derived using the incentive-compatibility and binding government budget constraints. With quasilinear-in-consumption utility, the relevant properties of the optimal bundles for each type can be inferred from the before-tax income schedule, so we do not consider the consumption schedule that solves candidate $k$’s relaxed problem explicitly.

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

$$
\max_{y(\cdot)} \int_{w}^{k} \left\{ \left[ (y(w) - h\left(\frac{y(w)}{w}\right)) f(w) + \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) F(w) \right] \right\} dw
$$

$$
+ \int_{\bar{w}}^{k} \left\{ \left[ (y(w) - h\left(\frac{y(w)}{w}\right)) f(w) - \frac{y(w)}{w^2} h'\left(\frac{y(w)}{w}\right) [1 - F(w)] \right] \right\} dw.
$$

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

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

$$
\theta^M(w, y(w)) = 0, \quad \forall w \in [w, k),
$$

$$
\theta^R(w, y(w)) = 0, \quad \forall w \in [k, \bar{w}],
$$

where

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

\footnote{If the budget constraint does not bind, because preferences are quasilinear in consumption, each person’s consumption can be increased by a common small amount without violating incentive compatibility, thereby increasing the utility of type $k$ individuals.}

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

\footnote{It is commonplace to call the maxi-min objective “Rawlsian” even though Rawls (1971) used an index of primary goods rather than utility in his criterion. Our notation reflects this common usage.}
and

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

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

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

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

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

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

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

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

---

8We write all first-order conditions for the optimal incomes as equalities, thereby implicitly assuming that the non-negativity constraints on incomes are not binding. The qualitative features of our analysis are unaffected if these constraints are taken into account.

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

First, each individual of type \( w \) can be given \( h'(\frac{y(w)}{w}) \frac{1}{w} \) additional units of consumption to place him on his initial indifference curve, thereby ensuring that he has no incentive to mimic any other type. This is shown by the adjustment from \((y(w), x(w))\) to \((\tilde{y}(w), \tilde{x}(w))\) in Figure 1. Moreover, this change does not affect the incentives of any types above \( w \).

These units of consumption must be subtracted from the \( f(w) \) units that can be diverted to type \( k \) individuals. This accounts for the second term in the first bracket in (14). Moving individuals of type \( w \) upward along their indifference curves in this way slackens the upward incentive constraint for the type \( \hat{w} \) immediately below \( w \). Because preferences are quasilinear in consumption, by reducing the consumption of everybody whose type is smaller than \( w \) by the amount in the final bracket in (14) restores incentive compatibility. This is illustrated by the adjustment from \((y(\hat{w}), x(\hat{w}))\) to \((\hat{y}(\hat{w}), \hat{x}(\hat{w}))\) for type \( \hat{w} \) in Figure 1. There are \( F(w) \) individuals whose types are smaller than \( w \), so the second term in (14) is the total amount of consumption that type \( k \) individuals can re-claim from these types in this way.

Candidate \( k \) also wishes to move resources away from types higher than himself towards individuals of his own type. These types are downward distorted, so this kind

\(^{10}\)Recall that we are assuming that the maxi-min solution exhibits no bunching, so \( y(w) \neq y(\hat{w}) \), where \( \hat{w} \) is the next highest type above \( w \). We are implicitly assuming that units of income are sufficiently small so that \( \hat{y}(w) < y(\hat{w}) \).
of redistribution is constrained by downward incentive compatibelity constraints. The function $\theta^R(w, y)$ shows the additional consumption that type $k$ individuals can secure for themselves through a one unit increase in $y(w)$ for some $w > k$. The only difference between $\theta^R(w, y)$ and $\theta^M(w, y)$ is in the final term. This difference arises because it is the downward incentive constraints that bind for types above $k$. Moving individuals of type $w > k$ upward along their indifference curves in the manner described in the preceding paragraph would lead to a violation of the downward incentive constraint for the next highest type. Because preferences are quasilinear in consumption, satisfaction of these constraints can be re-established by giving these individuals and everyone of a higher type more consumption in the amount given in the final bracket in (15). There are $1 - F(w)$ such individuals. Because this consumption must be given to individuals of types different from that of the candidate, these resources are subtracted from the amount available to the type $k$ individuals.

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

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

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

(18)

For $k \neq \underline{w}, \bar{w}$, there is a downward discontinuity in this schedule at $w = k$.

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

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

(19)

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

The discontinuities in the income schedule and in the marginal tax rates are intertwined. As we move from types just below type $k$ to types just above it, the upward distortions in incomes switch to downward distortions and the signs of the marginal
tax rates change from negative to positive. Because the maxi-max income schedule lies strictly above the maxi-min schedule at $w = k$, it is impossible to reconcile these competing distortions without a downward jump in the income schedule and a change in sign in the marginal tax rates at $k$.

The downward jump in the solution to candidate $k$’s relaxed problem for the income schedule clearly violates the second-order incentive compatibility conditions. So even if the maxi-min and maxi-max income schedules do satisfy these second-order conditions, we have not found a solution to candidate $k$’s problem in (10). Nevertheless, as we show below, some elements of these solutions feature in the complete solution to his problem. Before we turn to that issue, we first need to consider the circumstances in which the maxi-min and maxi-max income schedules are increasing, as assumed in this subsection.

3.2. Monotonicity of the Maxi-min and Maxi-max Income Schedules

Of the two components of candidate $k$’s relaxed solution, the part that tracks the maxi-min income schedule $y^R(\cdot)$ is the more familiar. Increasingness of this schedule is equivalent to the locus of points for which $\theta^R(w, y) = 0$ being increasing. For fixed $w$, the second-order condition for $y(w)$ to solve the unconstrained optimization problem (12) for this type is that $\theta^R_y(w, y(w)) \leq 0$. Thus, increasingness of $y(\cdot)$ requires that $\theta^R_y(w, y) < 0$ and $\theta^R_w(w, y) > 0$ for all $w \neq \bar{w}, \bar{w}$. From (15), a sufficient condition for $\theta^R_y(w, y) < 0$ is that $h''(l) \geq 0$. This assumption is satisfied by the commonly-used iso-elastic form. Given that $\theta^R_y(w, y) < 0$, a sufficient condition for $\theta^R_w(w, y) > 0$ is that the labor supply
schedule $l(\cdot)$ is upward sloping.\footnote{We have $-\theta^R_w(w,y)/\theta^R_y(w,y) = y'(w) = w'l(w) + l(w)$, from which it follows that $\theta^R_w(w,y) > 0$ if $\theta^R_y(w,y) < 0$ and $l'(w) > 0$.} Thus, only relatively mild assumptions are needed to ensure that the maxi-min income schedule is increasing.

Similarly, increasingness of the maxi-max income schedule $y^M(\cdot)$ is equivalent to the locus of points for which $\theta^M(w,y) = 0$ being increasing. However, in the maxi-max case, satisfaction of the second-order condition $\theta^M_y(w,y(w)) \leq 0$ for $y^M(w)$ to solve (12) for this value of $w$ may be problematic. By (12), $y^M(w)$ is found by maximizing

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

with respect to $y(w)$. Because $h$ is convex, the first term on the right-hand side of (20) is concave. However, when $h''(l) > 0$, the second term might be convex. If the curvature of the second term dominates that of the first, then $G^M$ is convex in $y$, and so the second-order condition is violated. Specifically, the second-order condition $\theta^M_y(w,y(w)) \leq 0$ is equivalent to

$$\left[\frac{2F(w) - w f(w)}{w^2}\right] h''\left(\frac{y(w)}{w}\right) + F(w) y(w) \frac{y(w)}{w^4} h'''\left(\frac{y(w)}{w}\right) \leq 0. \quad (21)$$

Under a maintained assumption that $h'''(l) > 0$, (21) can be satisfied only when the term in the square bracket on its left-hand side is negative. The numerator in this term tends to 2 as $w$ tends to infinity. Thus, if $\bar{w}$ is sufficiently large, it may well be the case that the second-order condition for an optimum of the relaxed version of the maxi-max problem will fail to be satisfied for values at the top end of the type distribution.

For reasons that are closely related to the possible failure of the second-order conditions in the maxi-max case, the function $y^M(\cdot)$ can be rather ill-behaved. A rote application of the implicit function theorem to the top line of (13) yields

$$\frac{dy^M(w)}{dw} = -\frac{\theta_w^M(w,y^M(w))}{\theta_y^M(w,y^M(w))}. \quad (22)$$

It is entirely possible that the left-hand side of (21) is negative for values of $w$ near $\bar{w}$. In fact, this must be the case if $h'''(l) = 0$. Thus, the denominator in (22) may change signs at least once on $[w,\bar{w}]$. This, in turn, implies that $y^M(\cdot)$ can have a vertical asymptote in the interior of the type distribution.

Unlike in the case of candidates distinct from $w$ and $\bar{w}$, the possible failure of the second-order conditions in the maxi-min and maxi-max cases are not the result of a downward discontinuity in the income schedule. Nevertheless, in order to obtain a complete solution to any candidate $k$’s problem for $k \in (w,\bar{w})$ as described in (10), we need to take into account not only the second-order conditions for his problem, but also those for the lowest- and highest-skilled types.
3.3. The Complete Solution

If either the maxi-min or maxi-max income schedule obtained using the first-order approach fails to satisfy the second-order incentive-compatibility condition (7) (i.e., the requirement that the schedule be non-decreasing), then it is necessary to bunch all types in a decreasing part of the schedule with some types who are in an increasing part, what is known as ironing. Any bunching region must be a closed interval. Its endpoints can be determined using the approach described by Guesnerie and Laffont (1984). Because ironing in this kind of situation is well understood and we do not need to know where the endpoints of these bunching regions are for our results, we shall simply suppose that these schedules have been ironed. We let \( y^{R*}(\cdot) \) and \( y^{M*}(\cdot) \) denote the optimal maxi-min and maxi-max income schedules, respectively.

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

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

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

(23)

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

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

(24)

Ironing does not affect the solution outside a bunching region, so no modifications to the integrands in (12) are needed for types that are not bunched. The intuition for these expressions when there is bunching is similar to that provided above for candidate \( k \)'s relaxed problem. Now, if an extra unit of consumption is given to type \( w \) individuals, it must be given to all individuals who are bunched with them, whose mass is \( \int_{w_-}^{w_+} f(t) dt \).

\[\text{[13] The first analysis of ironing in economics appears to have been by Arrow (1968). Arrow was concerned with devising an optimal capital policy with irreversible investment. The irreversibility of investment imposes a monotonicity constraint analogous to the one on incomes found here.}\]
For this reason, the $f(w)$ that appears in the first cases of both (23) and (24) is replaced by this integral in the second cases. When $w$ is bunched, in the maxi-max case, some of this extra consumption can be reclaimed from individuals of lower type than those bunched with $w$, whose mass is $F(w_-)$. The corresponding individuals in the maxi-min case are those individuals of higher type than those bunched with $w$, whose mass is $[1 - F(w_+)]$. In the second cases of (23) and (24), these expressions are used to replace the $F(w)$ and $[1 - F(w)]$ that appear in the first cases.

Using $y^R^*(\cdot)$ and $y^M^*(\cdot)$ instead of $y^R(\cdot)$ and $y^M(\cdot)$ in (18) for $k \neq w, \bar{w}$, we obtain the income schedule that must be ironed in order to determine the complete solution to candidate $k$’s problem. Because the only decreasing part of this schedule is the downward discontinuity at his type, only one bunching region needs to be introduced. In effect, we must build a bridge between the maxi-max and maxi-min parts of this schedule, as illustrated in Figure 3. All types with skill levels in the interval $[w_b, w_B]$ are bunched at a common allocation. The values of the bunching interval endpoints $w_b$ and $w_B$ are determined optimally so as to minimize the loss in candidate $k$’s utility that results from deviating from his relaxed solution.

The endpoints of the bunching interval can be determined using the control-theoretic approach of Guesnerie and Laffont (1984). However, we instead employ a much simpler procedure that was introduced by Vincent and Mason (1967, 1968) to smooth discontinuous control trajectories. Applied to our problem, in this approach, the optimal schedule is first selected for each fixed pair of values of the bridge endpoints $w_b$ and $w_B$. Then, among these schedules, the one that maximizes candidate $k$’s utility is selected. This is a two-dimensional unconstrained optimization problem. In other words, there is no need
to use optimal control theory to determine the bridge endpoints. Our choice of technique simplifies the comparative static exercises we perform in Section 4.

The derivatives of $G^{M^*}(w,y)$ and $G^{R^*}(w,y)$ with respect to income are denoted by $\theta^{M^*}(w,y)$ and $\theta^{R^*}(w,y)$, respectively. The before-tax income schedule that solves candidate $k$’s problem in (10) is described in Proposition 3.

**Proposition 3.** The optimal schedule of before-tax incomes $y^*(\cdot)$ for candidate $k$’s problem is given by

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

The optimal values of the bridge endpoints $w_b$ and $w_B$ are determined by the first-order condition

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

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

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

if $w_B < \bar{w}$.

As we have already noted, for types smaller than the lower endpoint of the bridge, candidate $k$’s optimal income schedule coincides with the maxi-max income schedule $y^{M^*}(\cdot)$, whereas for types larger than the upper endpoint of the bridge, it coincides with the maxi-min income schedule $y^{R^*}(\cdot)$. It is possible that $w_b = \underline{w}$, in which case the optimal income schedule starts with the bridge and then tracks the maxi-min solution. It is also possible that $w_B = \bar{w}$, in which case the optimal income schedule first tracks the maxi-max solution and then ends with the bridge.\textsuperscript{14} These two possibilities are illustrated in Figure 4. Provided that $w_B < \bar{w}$, the income for skill type $w_B$ is $y^R(w_B)$. Consequently, all individuals on the bridge receive this income. Analogously, if $w_b > \underline{w}$, then all individuals on the bridge receive $y^M(w_b)$. If both $w_b > \underline{w}$ and $w_B < \bar{w}$, then $y^M(w_b) = y^R(w_B)$. The first-order optimality conditions for the optimal placement of the bridge endpoints are given in (26) and (27). When both $w_b > \underline{w}$ and $w_B < \bar{w}$ hold, (26) and (27) are equivalent conditions. These two equations are similar to the standard ironing condition found in Guesnerie and Laffont (1984, eqn. (3.16), p. 347).

\textsuperscript{14}It is conceivable that $w_b = \underline{w}$ and $w_B = \bar{w}$, in which case the bridge is the whole income schedule. This possibility is so unlikely that we do not consider it explicitly.
4. The Political Equilibrium

We now seek to determine the identity of the winning candidate. We consider a world in which there are no costs to candidate entry. If a Condorcet winner exists, this will be the chosen candidate. We show that preferences over candidates are single-peaked with respect to skill level. Therefore, a candidate of the median type is elected and the tax schedule that maximizes the utility of the median type is chosen. We now develop this argument.

Because we are now considering comparisons across tax schedules chosen by different candidates, we now index allocations by $k$. Let $(x(w,k), y(w,k))$ denote the optimal allocation assigned to an individual of type $w$ under candidate $k$’s optimal tax schedule and

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

The argument begins with a comparative static result with respect to a change in the candidate.

**Proposition 4.** $y(w,k') \geq y(w,k)$ whenever $k' \geq k$.

One intuition behind Proposition 4 is that there is a tendency to distort labor supply upwards for individuals below the candidate’s type. The higher is the the type of the candidate, the more individuals have their labor supply distorted upward.

Because all candidates face the same budget and incentive constraints,

$$V(w,w) \geq V(w,k), \quad \forall w, k.$$  

This observation is useful in proving the single-peakedness result.

**Proposition 5.** Preferences are single-peaked in the type of the candidate.
A direct consequence of Proposition 5 is that the tax schedule that maximizes the utility of the median type subject to the budget and incentive compatibility constraints is enacted. This tax schedule features marginal wage subsidies for workers of low types, a bunching region that includes the median type, and then positive marginal tax rates for the most highly skilled (with the possible exception of the very highest skilled). In particular, there must be a kink in the tax schedule.

Given the complexity of the policies under consideration, it is striking that a Condorcet winner exists. Indeed, if voting were over all possible (nonlinear) income tax schedules, there would be voting cycles. The main insight of Röell (2012) is that by restricting attention to the tax schedules that are optimal from some person’s perspective, there are no voting cycles. Proposition 5 confirms that this insight extends to an economy with a continuum of skill-types. In the early literature on voting over tax schedules, following Roberts (1977), cycles were avoided by restricting attention to affine schedules that are budget feasible.

5. Adding a Minimum-Utility Constraint

As described so far, the candidates for public office are extremely opportunistic. Electoral victory is viewed as a license to extract maximal rents from fellow citizens. The process of reaping these rents could leave the least skilled with an extremely low level of utility. It may be the case that there are social conventions forbidding extreme lows of utility, or candidates themselves might temper their self-interest with a desire to avoid extreme suffering by the least skilled. In light of these considerations, and in-keeping with the work of Röell (2012) and Bohn and Stuart (2013), we now consider a political environment in which each candidate’s choice of tax function must provide some minimal level of utility, say \( u_0 \), to all workers. Because incentive compatibility implies that utility is increasing in skill level, it is enough to impose the constraint that

\[
V(w) \geq u_0. \tag{30}
\]

To make matters non-trivial, we assume that \( u_0 \) is less than the utility that the least-skilled candidate would obtain under the allocation \( y^R(w) \).

As in Section 3, we begin by studying a relaxed version of the problem solved by an arbitrary candidate \( k \). Specifically, we consider the problem of maximizing (12) subject to \( V(w) \geq u_0 \). This takes account of all the constraints faced by the candidate, except for the second-order incentive conditions. When the first-order incentive and public budget constraints are accounted for, (A.6) in the Appendix can be used to reformulate the minimum-utility constraint as

\[
\int_{w_l}^{w_u} \left\{ \left[ y(w) - h \left( \frac{y(w)}{w} \right) \right] f(w) - h \left( \frac{y(w)}{w} \right) \right\} \left[ 1 - F(w) \right] dw \geq u_0 \tag{31}
\]

The integrand on the left-hand side of (31) is identical to both the second integrand in (12) and the top line on the right-hand side of (24) because they all measure utility of the very lowest type.
Figure 5: The optimal income schedule for candidate $k$’s relaxed problem with a minimum-utility constraint

Attaching a Lagrange multiplier $\lambda$ to the constraint (31), the first order conditions for the relaxed problem can be written

$$\begin{cases}
\theta^M(w, y) + \lambda \theta^R(w, y) = 0, & w \in [w, k], \\
\theta^R(w, y) = 0, & w \in (k, \bar{w}].
\end{cases}
$$

The solution to these first order conditions is illustrated in Figure 5, where $y^0(w)$ denotes the solution to the top line of (32). For types above the candidate’s, the solution is exactly the same as to the relaxed problem in the absence of a minimum utility constraint. This is because the addition of (30) effectively turns the candidate’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 constraint (30). The candidate still wishes to redistribute resources downward from types greater than his own, and redistribution from higher types is limited by downward incentive constraints. For types below his own, the candidate faces a problem of countervailing incentives. On the one hand, he wishes to redistribute upwards in the manner described in Section 3. On the other, he needs to move resources downward to the very lowest skill-type in order to satisfy the minimum-utility constraint. The before-tax income allocation that reconciles these motives, naturally, lies between the maxi-max and maxi-min allocations. The before-tax income schedule features a downward jump, so that the solution to the relaxed problem is not incentive compatible.

In contrast to the situation presented in Section 3, ill-behaved solutions to the relaxed problem can be ruled out by mild conditions on the primitives of the model. If $y^0(w)$ were
to asymptote, its deviation from the maxi-min allocation $y^R(w)$ would be unbounded. This, in turn, would lead to $y^0(w)$ yielding a utility for the lowest skill-type arbitrarily lower than the maxi-min utility, thereby causing a violation of the minimum-utility constraint. For this reason, we assume in this section that the second-order conditions for a strict maximum to the candidate’s optimum are always satisfied.\footnote{Formally, we assume that the Hessian of the associated Lagrange function is negative-definite in subspace orthogonal to the gradient of the function on the left-hand side of (31).}

As noted in Section 3, an unbounded $y^M(w)$ occurs most frequently for candidates of fairly high skill levels. Thus, it must be the case that $y^0(w)$ lies closer to $y^R(w)$ for more highly skilled candidates. This suggests that $\lambda$, the shadow value of the minimum utility constraint, increases in $k$. We show in Proposition 6 below that this is indeed the case in the complete (non-relaxed) version of the candidate’s problem.

When the second-order incentive (monotonicity) constraint is added to problem, the solution features bridging. For values of $w$ near $\bar{w}$, the minimum-utility constraint is non-binding. For these candidates, the optimal schedule is identical to the one described in Section 3. Candidates with a skill level beyond some threshold choose schedules that are comprised of three segments. For the lower skill levels, $y^0(w)$; a bridge in a region containing the candidate’s type; and finally $y^R(w)$ for the highest skill types.\footnote{For distributions bounded on the right, this upper segment might not exist for candidates near $\bar{w}$. It is also possible that the bridge starts at $\underline{w}$, so that the lower segment might not exist.}

Because the minimum-utility constraint (31) contains an integral involving the entire before-tax income schedule, including its value along the bridge, it is impossible to use the two-stage technique of Section 3 to find the bridge points $w_b$ and $w_B$. Instead, the bridge points are determined jointly with the before-tax incomes and the Lagrange multiplier $\lambda$. Comparative static results on how these quantities change with the candidate’s type are contained in Proposition 6.

**Proposition 6.** Let $\lambda(k)$, $w_b(k)$, and $w_B(k)$ be the values the Lagrange multiplier associated with minimum-utility constraint, and the bridge points, respectively, when a candidate of type $k$ chooses the tax schedule. Then

1. $\lambda(k)$ increases in $k$ whenever $\lambda(k) > 0$, and;
2. both $w_b(k)$ and $w_B(k)$ increase in $k$ whenever $w_b(k) > \underline{w}$.
3. both $w_b(k)$ and $w_B(k)$ are locally constant if $w_b(k) = \underline{w}$.

One interpretation of part 1 of the proposition was given in our discussion of the relaxed problem. Further intuition can be gained by examining equation (14). As explained in Section 3, the final term in this equation measures resources that can be transferred upward from all types below the candidate’s. This term typically increases in $w$, so for large values of $k$ there will be some types just below it for which there are large gains to be had from this form of transfer. But the minimum-utility constraint acts to limit these upward transfers. A relaxation of the minimum-utility constraint, therefore, has a
greater value for candidates of a higher type. This is exactly the content of part 1 of the proposition.

Part 2 of Proposition 6 is geometrically intuitive, but the argument is somewhat subtle. When the minimum utility constraint does not bind, the analysis of Section 3 applies. When the minimum-utility constraint binds, part 1 of the proposition implies that an increase in $k$ induces a decrease in before-tax income for all skill-types to the left of the bridge. If $w_b$ did not increase, the entire before-tax income schedule would shift downward, causing an increase in left-hand side of (31). Intuitively, the entire schedule would be closer to the maxi-min schedule $y^R(w)$, so the utility of the least-skilled increases. But this slackens the minimum-utility constraint, thereby giving the candidate scope to increase his utility by increasing the before-tax incomes of all individuals of skills less than $w_b$. Thus, candidates of higher type optimally choose higher $w_b$. Figure 6 illustrates how the optimal before-tax income schedule changes when $k$ increases.

When $w_b(k) = w$, an increase in $k$ does increase the shadow value of the minimum-utility constraint, but there is no segment of $y^0(w)$ to shift down. Any change in the bridge portion of the before-tax income schedule would alter the left-hand side of (31). Decreasing this left-hand side is infeasible; increasing it would lead to a slack constraint, violating part (1) of Proposition 6.

If there happens to exist a set of candidates for which the lower bridge point is the minimum skill level, then this set must be an interval of types. Every candidate in that interval chooses the same allocation, and all voters are indifferent among these candidates. As long as such a set of candidates comprises less than half the population, their presence does not upset median-voter logic. If they do form a majority, then they
must contain the median voter.

We now turn our attention to how voters rank the candidates. Even in the presence of a minimum-utility constraint, a voter’s most preferred candidate is someone of his own skill-type. This is because the minimum-utility constraint is not type-specific. We have already described how voters rank alternatives for which (30) does not bind. We now consider the case when (30) does bind. If we let \( V^0(w, k) \) denote the utility of a worker of type \( w \) when a candidate of type \( k \) set the tax schedule, then (6) implies

\[
V^0(w, k) = u_0 + \int_w^w h'(y(t, k) \frac{y(t, k)}{t^2}) dt
\]

(33)

The partial derive of this function with respect to \( k \) is

\[
\frac{\partial V^0(w, k)}{\partial k} = \int_w^w \left[ h'' \left( y(t, k) \frac{y(t, k)}{t^2} \right) + h' \left( y(t, k) \frac{1}{t^2} \right) \right] \frac{\partial y(t, k)}{\partial k} dt.
\]

(34)

The term in the square bracket in (34) is always positive. The partial derivatives inside the integral sign are negative for low value of \( t \) (below \( w_b(k) \)) and can be positive for higher values. If \( w < k \), there is a preponderance of negative terms in the integral in (34), and \( V^0 \) is decreasing in \( k \). If \( w > k \) the positive terms enter into the integral via the effect of a rising bridge and, as we show in proof of Proposition 7 below, dominate the negative terms. These observations lead to the following proposition.

**Proposition 7.** In the presence of a minimum-utility constraint, preferences are single-peaked in the type of the candidate.

The qualitative features of the equilibrium tax schedule depend on whether the minimum-utility constraint binds for the median candidate or not. If it does not bind, then the equilibrium tax schedule is a described in Section 4. Otherwise, there is a positive marginal tax rate at the lowest skill level. This is because \( y^M(w) \) features no distortion at the bottom, but \( y^0(w) \) lies below \( y^M(w) \). There is also a bunching region that includes the median type, and positive marginal tax rates for types above the bunching range (except the very highest type, should one exist). The sign of the marginal tax rate is ambiguous below the bunching region, because it depends on the exact value of the multiplier \( \lambda \). If this multiplier is near zero for the median type, then \( y^0(w) \) lies near \( y^M(w) \), and the marginal tax rate becomes negative for some types below the median. In this case, there also exists an interior undistorted type. If \( \lambda \) is relatively large, \( y^0(w) \) lies near \( y^R(w) \), and marginal tax rates remain negative throughout \([w, w_b(k)]\). These are precisely the patterns of equilibrium tax rates found by Röell (2012) for a discrete population and by Bohn and Stuart (2013) for a continuous type space.

While imposing a minimum-utility constraint is a reasonable way to introduce concern for the least skilled, and thereby limit the extent to which they are disadvantaged through taxation, it is not the only way. One could imagine that candidates 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 a candidate has an objective function of the form

\[ V(k) + \lambda V(w) \]  

and no minimum-utility constraint is introduced. Indeed, the first-order conditions for the problem faced by such a candidate are exactly (32). However, because \( \lambda \) is fixed in this variant of the model, the second-order conditions for a candidate’s optimum might be violated. In that case, the analysis of Sections 3 and 4 are needed to characterize the voting equilibrium, even though the outcome is qualitatively similar to the voting equilibrium described in this section.

6. Conclusion

We have fully described the rather complex tax schedules that maximize the utility of an office holder who can freely design policy. The desire of types above the minimum to redistribute income upward produces a non-standard pattern of distortions reminiscent of those found in the literature on participation constraints and countervailing incentive in screening problems, as studied in detail by Jullien (2000). In addition, for types sufficiently above the minimum, complications arise that do not feature in the existing literature on countervailing incentives. These complications arise from potential violations of the second-order conditions associated with the tax-designer’s decision problem. Moreover, there is a fundamental tension between the desires to redistribute from the poor and from the rich that calls for special attention to be given to the second-order conditions associated with the workers’ decision problems. We have shown how an allocation schedule that reconciles all of these tensions is necessarily quite complicated.

The tools we use to analyze the candidates’ decisions are of potential interest in the study of other screening problems in which there might be a discontinuity in the principal’s objective function. For example, a policy design problem in which some group is explicitly excluded from the planner’s objective but is part of the set of contributors to the policy’s finances would give rise to a discontinuous objective function.

Despite the complexity of the tax schedule chosen by a particular candidate, the way in which schedules vary across candidates is relatively simple to characterize. This simplicity gives rise to single-peaked preferences over candidates (and their tax schedules). Given the relatively simple political process we posit in this paper, a median voter result follows. Our comparative static result on how candidates’ schedules change as one moves along the distribution can also be a building block in understanding the positions that different constituencies might hold and advocate in any model of politics.

While the tax schedule chosen by the median type appears to be at odds with casual observations about real world income tax schedules, it may not be a wildly unrealistic description of the entire redistributive apparatus of government. The tendency for democratic governments to redistribute from both the poor and the rich toward the middle class is termed Director’s Law. Recent empirical studies by Pamp and Mohl (2010) and
Scervini (2012) come to conflicting conclusions about the descriptive power of Director’s Law. Our conclusion that labor supply may be subject to marginal subsidy for a segment of low-skilled workers is consistent with the growing use of in-work benefit schemes, as surveyed in Brewer et al. (2009). A full treatment of the politics of income taxation for the low-skilled would also include issues related to labor supply at the extensive margin and interactions with policies targeted toward disabled and other disadvantaged workers. We leave these issues, and undoubtedly others, for future research.

Appendix

Proof of Proposition 1. By (6),

\[ V(w) = V(w) + \int_w^\infty \frac{y(t)}{t^2} h' \left( \frac{y(t)}{t} \right) dt. \quad (A.1) \]

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

\[ \int_w^\infty V(w)f(w)dw = \int_w^\infty V(w)f(w)dw + \int_w^\infty \int_w^w \frac{y(t)}{t^2} h' \left( \frac{y(t)}{t} \right) f(w)dt dw. \quad (A.2) \]

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

\[ \int_w^\infty V(w)f(w)dw = V(w) + \int_w^\infty \left( \int_t^w \frac{y(t)}{t^2} h' \left( \frac{y(t)}{t} \right) f(w)dt \right) dw \]

\[ = V(w) + \int_w^\infty \left( \frac{y(t)}{t^2} h' \left( \frac{y(t)}{t} \right) \right) [1 - F(t)] dt. \quad (A.3) \]

On the other hand, by (4),

\[ \int_w^\infty V(w)f(w)dw = \int_w^\infty x(w)f(w)dw - \int_w^\infty h \left( \frac{y(w)}{w} \right) f(w)dw. \quad (A.4) \]

Substituting (9) into (A.4) yields

\[ \int_w^\infty V(w)f(w)dw = \int_w^\infty y(w)f(w)dw - \int_w^\infty h \left( \frac{y(w)}{w} \right) f(w)dw. \quad (A.5) \]

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

\[ V(w) = \int_w^\infty y(w)f(w)dw - \int_w^\infty h \left( \frac{y(w)}{w} \right) f(w)dw \]

\[ - \int_w^\infty \frac{y(w)}{w^2} h' \left( \frac{y(w)}{w} \right) [1 - F(w)] dw. \quad (A.6) \]

The maximand in (12) is obtained by substituting (A.6) into (A.1) and setting \( w = k \). The preceding calculations have accounted for all the constraints in (11), and so the proof is complete. \( \blacksquare \)
Proof of Proposition 3. First, fix the bridge endpoints \( w_b \) and \( w_B \) and let \( y^*(w_b, w_B) \) denote the optimal income on the bridge \([w_b, w_B]\). A bridge cannot begin in the interior of a bunching interval of \( y^M(\cdot) \), nor can it end in the interior of a bunching interval of \( y^R(\cdot) \), so it is supposed in the rest of this proof that \( w_b \) and \( w_B \) satisfy these restrictions.

The rest of candidate \( k \)'s optimal income schedule is obtained by solving

\[
\max_{y(\cdot)} \left[ \int_{w_b}^{w_B} G^M(w, y(w))dw + \int_{w_B}^{w} G^R(w, y(w))dw \right]
\]

subject to \( y(w) = y^*(w_b, w_B), \forall w \in [w_b, w_B] \).

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

\[
\theta^M(w, y(w)) = 0, \quad \forall w \in [w, w_b],
\]

\[
\theta^R(w, y(w)) = 0, \quad \forall w \in [w_B, w].
\]

Thus,

\[
y^*(w) = y^M(w), \quad \forall w \in [w, w_b],
\]

and

\[
y^*(w) = y^R(w), \quad \forall w \in [w_B, \bar{w}].
\]

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

Case 1: \( w_B < \bar{w} \). Continuity of the solution implies that \( y^*(w_B) = y^R(\bar{w}) \). We also have \( y^*(w_B) = y^*(\bar{w}) \) because income is a constant on the bridge. If \( w_b > \bar{w} \), then by continuity, \( y^*(w_b) = y^M(w_b) \). Let

\[
\phi(w_B) = \begin{cases} y^{M-1}(y^R(w_B)), & \text{if } w_b > \bar{w}, \\ w_b, & \text{if } w_b = \bar{w}. \end{cases}
\]

Because \( y^M(\cdot) \) is continuous and piecewise continuously differentiable, so is \( \phi(\cdot) \). Candidate \( k \)'s choice of \( w_B \) is therefore the solution to

\[
\max_{w_B} \int_{w}^{\phi(w_B)} G^M(w, y^M(w))dw + \int_{\phi(w_B)}^{w_B} G^M(w, y^R(w_B))dw
\]

\[
+ \int_{w_B}^{\bar{w}} G^R(w, y^R(w_B))dw + \int_{\bar{w}}^{\bar{w} - w_B} G^R(w, y^R(w))dw,
\]

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

\[
G^M(w_b, y^M(w_b)) \frac{d\phi(w_B)}{dw_B} - G^M(w_b, y^M(w_b)) \frac{d\phi(w_B)}{dw_B}
\]

\[
+ \int_{w_b}^{k} \theta^M(w, y^R(w_B)) \frac{dy^R(w_B)}{dw_B}dw + \int_{w_B}^{w} \theta^R(w, y^R(w_B)) \frac{dy^R(w_B)}{dw_B}dw
\]

\[
+ G^R(y^R(w_b, y(w_B)) - G^R(y^R(w_b, y(w_B))) = 0.
\]
Simplifying (A.13) yields
\[
\frac{dy^R(w_B)}{dw_B} \left[ \int_{w_b}^{w} \theta^M(y^R(w_B)) dw + \int_{w}^{w_B} \theta^R(y^R(w_B)) dw \right] = 0, \tag{A.14}
\]
from which (27) follows because \(y^R(\cdot)\) is increasing at \(w_B\).

Case 2: \(w_b > w\). Continuity of the solution now implies that \(y^*(w_b) = y^M(w_b)\). The proof that (26) holds parallels that of Case 1 where now the optimization is with respect to \(w_b\).

\[\square\]

Proof of Proposition 4. Given the nature of the candidate’s optimum, it suffices to show that the bridge points \(w_b\) and \(w_B\) move to the right as \(k\) increases.

The objective function in (A.12) depends only on \(w_B\) and \(k\). Let this function be denoted by \(B(w_B, k)\). It follows directly from differentiating (A.14) along with (23) and (24) that \(B(w_B, k)\) is supermodular in its arguments. The calculations are as follows:

\[
\frac{\partial^2 B}{\partial w_B \partial k} = \frac{dy^R(w_B)}{dw_B} \left[ \theta^M(y^R(w_B), k) - \theta^R(y^R(w_B), k) \right] = \frac{dy^R(w_B)}{dw_B} \left[ h'' \left( \frac{y^R(w_B)}{k} \right) y^R(w_B) k^3 + h' \left( \frac{y^R(w_B)}{k} \right) \frac{1}{k^2} \right] > 0. \tag{A.15}
\]

It follows that \(w_B\) increases with \(k\). A similar argument can be used to establish that \(w_b\) also increases in \(k\).

\[\square\]

Proof of Proposition 5. Take three types \(w, k_1\) and \(k_2\) such that \(w < k_1 < k_2\). We need to show that
\[
V(w, w, k_1) \geq V(w, k_1) \geq V(w, k_2). \tag{A.16}
\]

The first inequality in (A.16) follows from (29).

Now, by (6),
\[
V(w, k_1) = V(k_1, k_1) - \int_{w}^{k} h' \left( \frac{y(t, k_1)}{t} \right) \frac{y(t, k_1)}{t^2} dt. \tag{A.17}
\]

Therefore, using (29),
\[
V(w, k_1) \geq V(k_1, k_2) - \int_{w}^{k} h' \left( \frac{y(t, k_1)}{t} \right) \frac{y(t, k_1)}{t^2} dt. \tag{A.18}
\]

Also by (29),
\[
V(w, k_2) = V(k_1, k_2) - \int_{w}^{k} h' \left( \frac{y(t, k_2)}{t} \right) \frac{y(t, k_2)}{t^2} dt. \tag{A.19}
\]

Subtracting (A.19) from (A.18) yields
\[
V(w, k_1) - V(w, k_2) = \int_{w}^{k} \left[ h' \left( \frac{y(t, k_2)}{t} \right) \frac{y(t, k_2)}{t^2} - h' \left( \frac{y(t, k_1)}{t} \right) \frac{y(t, k_1)}{t^2} \right] dt. \tag{A.20}
\]
By Proposition 4, \( y(t, k_2) \geq y(t, k_1) \) for all \( t \). Therefore, by the convexity of \( h \), the right hand side of (A.20) is non-negative. This establishes the second inequality in (A.16). The proposition follows.

**Proof of Proposition 6.** As in Section 3, assume that \( \theta^R_y(w, y) < 0 \) for all \((w, y)\). Moreover, the strict second-order conditions associated with \( y^0(w) \) can be written

\[
\theta^M_y(w, y) + \lambda \theta^R_y(w, y) < 0, \quad \forall (w, y),
\]

where \( \lambda \) is the endogenous Lagrange multiplier. These assumptions are sufficient to ensure that \( y(w) \) is increasing everywhere except on the single bridging interval. They also ensure that \( \theta^R(w, y) \) is negative everywhere above the graph of \( y^R(w) \) and that \( \theta^M(w, y) + \lambda \theta^R(w, y) \) is positive below the graph of \( y^0(w) \).

It is convenient to solve the candidate’s problem in a number of steps. In the first step, the value of the multiplier associated with the minimum-utility constraint is taken as given, as are the bridge points \( w_b \) and \( w_B \). Outside of the bridge, the before-tax income schedule satisfies (32). Let \( \tilde{y}(w, \lambda) \) denote the solution to these first-conditions for a fixed value of \( \lambda \). Using (A.21) and the Implicit Function Theorem,

\[
\frac{\partial \tilde{y}(w, \lambda)}{\partial \lambda} = \frac{-\theta^R_y(w, y)}{\theta^M_y(w, y) + \lambda \theta^R_y(w, y)} < 0, \quad w < w_b;
\]

and,

\[
\frac{\partial \tilde{y}(w, \lambda)}{\partial \lambda} = \frac{\partial y^R(w)}{\partial \lambda} = 0, \quad w > w_B.
\]

The inequality in (A.22) follows from the observations immediately below (A.21).

The results from the first step of the candidate’s optimization problem are then substituted into the objective function and the minimum-utility constraint, and the optimal values of \( w_b, w_B, \) and \( \lambda \) are found.

We denote the integrand on the left-hand side of (31) by \( G^R(y, w) \).

Three cases arise.

**Case 1:** \( w < w_b < w_B < \bar{w} \).

In this case, the choice of \( w_b, w_B, \) and \( \lambda \) is determined by following three equations.

\[
\tilde{y}(w_b, \lambda) - y^R(w_B) = 0; \quad (A.24)
\]

\[
\int_{w_b}^{w_B} G^R(\tilde{y}(w, \lambda), w)dw + \int_{w_b}^{w_B} G^R(y^R(w_B), w)dw + \int_{w_B}^{\bar{w}} G^R(y^R(w), w)dw = u_0; \quad (A.25)
\]

\[
\int_{w_b}^{k} \theta^M_y(y^R(w_B), w)dw + \int_{w_b}^{k} \theta^R_y(y^R(w_B), w)dw + (1 + \lambda) \int_{w_b}^{w_B} \theta^R_y(y^R(w_B), w)dw = 0.
\]

(A.26)
The first of these equations, (A.24), is the requirement that the before-tax income function be continuous at the two ends of the bridge. The second, (A.25), is the minimum-utility constraint with the value of the before-tax income function substituted therein. We have used (24) in this derivation. The third, (A.26), is the first-order condition for the placement of the bridge points. We note that the left-hand sides of (A.24)–(A.26) are functions of $w_b, w_B, \lambda,$ 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.

We now employ the Implicit Function Theorem to compute how the endogenous variables in the system (A.24)–(A.26) respond to a change in $k$. Implicit differentiation of (A.24)–(A.26) yields

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

(A.27)

We now compute the entries in (A.27). Using (A.24),

$$
\Upsilon_{w_b} = \frac{\partial \tilde{y}(w_b, \lambda)}{\partial w_b} > 0;
$$

(A.28)

$$
\Upsilon_{w_B} = \frac{\partial y^R(w_B)}{\partial w_B} > 0;
$$

(A.29)

$$
\Upsilon_{\lambda} = \frac{\partial \tilde{y}(w_b, \lambda)}{\partial \lambda} < 0;
$$

(A.30)

$$
\Upsilon_k = 0.
$$

(A.31)

The inequality in (A.30) follows from implicitly differentiating (32).

Using (A.25),

$$
\Psi_{w_b} = G^R(\tilde{y}(w_b, \lambda), w_b) - G^R(y^R(w_B), w_b) = 0, \text{ by (A.24)};
$$

(A.32)

$$
\Psi_{w_B} = G^R(y^R(w_B), w_B) + \int_{w_b}^{w_B} \theta^R(y^R(w_B), w) dw - G^R(y^R(w_B), w_B)
$$

(A.33)

$$
= \int_{w_b}^{w_B} \theta^R(y^R(w_B), w) dw < 0;
$$

$$
\Psi_{\lambda} = \int_{w_b}^{w_B} \theta^R(\tilde{y}(w, \lambda), w) \frac{\partial \tilde{y}(w, \lambda)}{\partial \lambda} dw > 0;
$$

(A.34)

$$
\Psi_k = 0.
$$

(A.35)

The inequalities in (A.33) and (A.34) follow from the fact that $\theta^R(y, w) < 0$ for all $(y, w)$ above the $y^R(w)$ locus. This, in turn, follows from $\theta^R_y(y, w) < 0$ everywhere.

Using (A.26)

$$
\Phi_{w_b} = -\theta^M(y^R(w_B), w_b) - \lambda \theta^R(y^R(w_B), w_b) = 0, \text{ by (32) and (A.24)};
$$

(A.36)
\[
\Phi_{w_B} = \left( \int_{w_b}^k \left[ \theta^M(y^R(w_B), w) + \lambda \theta^R(y^R(w_B), w) \right] dw \right. \\
\left. + (1 + \lambda) \int_k^{w_B} \theta^R(y^R(w_B), w) \frac{\partial y^R(w_B)}{w_B} \right) < 0; (A.37)
\]
\[
\Phi_\lambda = \int_{w_b}^{w_B} \theta^R(y^R(w_B), w) dw = \Psi_{w_B} < 0; (A.38)
\]
\[
\Phi_k = \theta^M(y^R(w_B), k) - \theta^R(y^R(w_B), k) > 0, \quad \text{as in (A.15).} (A.39)
\]

The inequality in (A.37) follows from the second-order conditions.

Let \( A \) be the matrix on the left-hand side of (A.27). The determinant \(|A|\) is given by
\[
|A| = \Upsilon_{w_B} \left[ \Psi_{w_B} \Phi_\lambda - \Psi_\lambda \Phi_{w_B} \right] > 0. (A.40)
\]

The inequality in (A.40) follows from (A.28), (A.33), (A.38), (A.34), and (A.37).

Now, by Cramer’s Rule, and making use of (A.28)–(A.40),
\[
\frac{\partial w_B}{\partial k} = \frac{\begin{vmatrix} 0 & \Upsilon_{w_B} & \Upsilon_\lambda \\ 0 & \Psi_{w_B} & \Psi_\lambda \\ 0 & -\Phi_k & \Phi_\lambda \end{vmatrix}}{|A|} = \frac{-\Phi_k \Upsilon_{w_B} \Psi_\lambda - \Upsilon_\lambda \Psi_{w_B}}{|A|} = -\frac{(+)\left[(+)\left[(-)(+)-(--)\right]\right]}{(+)} > 0; (A.41)
\]
\[
\frac{\partial w_B}{\partial k} = \frac{\begin{vmatrix} \Upsilon_{w_B} & 0 & \Upsilon_\lambda \\ 0 & 0 & \Psi_\lambda \\ 0 & -\Phi_k & \Phi_\lambda \end{vmatrix}}{|A|} = \frac{\Phi_k \Upsilon_{w_B} \Psi_\lambda}{|A|} = \frac{(+)\left[(+)\left[(+)\right]\right]}{(+)} > 0; (A.42)
\]
\[
\frac{\partial \lambda}{\partial k} = \frac{\begin{vmatrix} \Upsilon_{w_B} & \Upsilon_{w_B} & 0 \\ 0 & 0 & \Psi_{w_B} \\ 0 & \Phi_{w_B} & -\Phi_k \end{vmatrix}}{|A|} = \frac{-\Phi_k \Upsilon_{w_B} \Psi_{w_B}}{|A|} = \frac{-(+)\left[(+)\left[(-)ight]\right]}{(+)} > 0. (A.43)
\]

This ends the proof for this case.

**Case 2: \( w < w_b < w_B = \bar{w} \).**

In this case, the value of before-tax income for the types on the bridge is not \( y^R(w_B) \), as in Case 1. Instead, the types in \([w_b, \bar{w}]\) receive \( \bar{y}(w_b, \lambda) \). In addition, equation (A.24) no longer holds. The two variables \( w_b \) and \( \lambda \) are determined by the following modified versions of (A.25) and (A.26):
\[
\int_{w}^{w_b} G^R(\bar{y}(w_b, \lambda), w) dw + \int_{w_b}^{\bar{w}} G^R(\bar{y}(w_b, \lambda), w) dw = u_0; (A.44)
\]

29
\[
\int_{w_b}^{k} \theta^M(\tilde{y}(w_b, \lambda), w)dw + \lambda \int_{w_b}^{k} \theta^R(\tilde{y}(w_b, \lambda), w)dw + (1 + \lambda) \int_{k}^{w} \theta^R(\tilde{y}(w_b, \lambda), w)dw = 0.
\] (A.45)

Mimicking the proof of Case 1, let \( \overline{\Psi}(w_b, \lambda, k) \) and \( \overline{\Phi}(w_b, \lambda, k) \) denote the expressions on the left-hand sides of (A.44) and (A.45), respectively. The calculations needed to apply the Implicit Function Theorem are as follows:

\[
\begin{align*}
\overline{\Psi}_{w_b} &= G^R(\tilde{y}(w_b, \lambda), w_b) - G^R(\tilde{y}(w_b, \lambda), w_b) + \int_{w_b}^{w} \theta^R(\tilde{y}(w_b, \lambda), w) \frac{\partial \tilde{y}(w_b, \lambda)}{\partial w_b} < 0; \quad (A.46) \\
\overline{\Psi}_{\lambda} &= \int_{w_b}^{w} \theta^R(\tilde{y}(w_b, \lambda), w) \frac{\partial \tilde{y}(w_b, \lambda)}{\partial w_b} + \int_{w_b}^{w} \theta^R(\tilde{y}(w_b, \lambda), w) \frac{\partial \tilde{y}(w_b, \lambda)}{\partial \lambda} > 0; \quad (A.47) \\
\overline{\Psi}_k &= 0; \quad (A.48) \\
\overline{\Phi}_{w_b} &= -\theta^M(\tilde{y}(w_b, \lambda), w_b) - \lambda \theta^R(\tilde{y}(w_b, \lambda), w_b) \\
&\quad + \left( \int_{w_b}^{k} \left[ \theta^M_y(\tilde{y}(w_b, \lambda), w) + \lambda \theta^R_y(\tilde{y}(w_b, \lambda), w) \right] dw \\
&\quad + (1 + \lambda) \int_{k}^{w} \theta^R_y(\tilde{y}(w_b, \lambda), w)dw \right) \frac{\partial \tilde{y}(w_b, \lambda)}{w_b} < 0; \quad (A.49) \\
\overline{\Phi}_{\lambda} &= \int_{w_b}^{w} \theta^R(\tilde{y}(w_b, \lambda), w) < 0; \quad (A.50) \\
\overline{\Phi}_{w_b} &= \theta^M(\tilde{y}(w_b, \lambda), k) + \lambda \theta^R(\tilde{y}(w_b, \lambda), k) - (1 + \lambda)\theta^R(\tilde{y}(w_b, \lambda), k) \\
&= \theta^M(\tilde{y}(w_b, \lambda), k) - \theta^R(\tilde{y}(w_b, \lambda), k) > 0. \quad (A.51)
\end{align*}
\]

The arguments establishing the signs of these derivatives are virtually identical to those for the previous case. One notable additional argument is that the top line on the right-hand side of (A.49) vanishes by (32).

Using a notation and style of argument similar to the proof of Case 1, and (A.46)–(A.51), we consider the matrix equation:

\[
\begin{bmatrix}
\overline{\Psi}_{w_b} & \overline{\Psi}_{\lambda} \\
\overline{\Phi}_{w_b} & \overline{\Phi}_{\lambda}
\end{bmatrix}
\begin{bmatrix}
dw_b \\
d\lambda
\end{bmatrix} =
\begin{bmatrix}
0 \\
-\overline{\Phi}_k
\end{bmatrix}
dk,
\] (A.52)

and conclude that

\[
|\overline{A}| = \begin{vmatrix}
\overline{\Psi}_{w_b} & \overline{\Psi}_{\lambda} \\
\overline{\Phi}_{w_b} & \overline{\Phi}_{\lambda}
\end{vmatrix} = (-)(-) - (+)(-) > 0; \quad (A.53)
\]

\[
\frac{\partial w_b}{\partial k} = \frac{0}{|A|} = \frac{\overline{\Psi}_{\lambda} \overline{\Phi}_k}{|A|} = \frac{(+)(+)}{(+) > 0}; \quad (A.54)
\]

\[
\frac{\partial \lambda}{\partial k} = \frac{\overline{\Psi}_{w_b} 0}{|A|} = \frac{-\overline{\Psi}_{w_b} \overline{\Phi}_k}{|A|} = \frac{-(+) > 0. \quad (A.55)}{(+) > 0}.
\]
The result now follows for this case.

**Case 3:** \( w = w_b < w_B < \bar{w} \).

This case is similar to Case 1, except that (A.24) is not used, and the functions \( \Phi \) and \( \Psi \) are evaluated with \( w_b = w \). Notably, this results in \( \lambda \) not appearing in the \( \Psi \) function. The term \( \Upsilon_{w_b} \) disappears from the analogs to (A.40), (A.42), and (A.43), but all other calculations remain essentially unchanged from Case 1. Specifically,

\[ |A| = \Psi_{w_B} \Phi_\lambda > 0; \quad \text{(A.56)} \]
\[ \frac{\partial w_B}{\partial k} = \frac{\Phi_k \Psi_\lambda}{|A|} = \frac{(+) (0)}{(+) - (-)} = 0; \quad \text{(A.57)} \]
\[ \frac{\partial \lambda}{\partial k} = \frac{-\Phi_k \Psi_{w_B}}{|A|} = \frac{(-)(+)}{(+) - (-)} > 0. \quad \text{(A.58)} \]

**Proof of Proposition 7.** We present the proof for the case of \( w_b(k) > w \) for all \( k \). We have discussed in the text how to account for the possibility that \( w_b(k) = w \) for some \( k \).

We first show that the function \( V^0(w, k) \) is single-peaked in \( k \). That is, we show that \( V^0(w, k) \) is decreasing in \( k \) for \( k > w \) and increasing in \( k \) for \( k < w \). In order to do so, we evaluate the right-hand side of (34). The following preliminaries are needed:

\[ \frac{\partial y(w, k)}{\partial k} = \frac{\partial y(w, \lambda(k))}{\partial \lambda} \frac{\partial \lambda(k)}{\partial k} < 0, \quad w < w_b(k); \quad \text{(A.59)} \]
\[ \frac{\partial y(w, k)}{\partial k} = \frac{\partial y^R(w_B(k))}{\partial w_B} \frac{\partial w_B(k)}{\partial k} > 0, \quad w_b(k) < w < w_B(k); \quad \text{(A.60)} \]
\[ \frac{\partial y(w, k)}{\partial k} = \frac{\partial y^R(w)}{\partial k} = 0, \quad w > w_B(k); \quad \text{(A.61)} \]

The inequalities in (A.59) and (A.60) follow from Proposition 6. We note that the expression in the middle of (A.60) is constant in \((w_b, w_B)\). Intuitively, it is change in the height of the bridge due to a change in \( k \).

Now, suppose that \( k > w \). By (34),

\[ \frac{\partial V^0(w, k)}{\partial k} = \int_w^{w_b(k)} \left[ \Theta^M(t, y(t, k)) - \Theta^R(t, y(t, k)) \right] \frac{\partial y(t, k)}{\partial k} \, dt + \int_{w_b(k)}^{w} \left[ \Theta^M(t, y(t, k)) - \Theta^R(t, y(t, k)) \right] \frac{\partial y(t, k)}{\partial k} \, dt. \quad \text{(A.62)} \]

If \( w \leq w_b(k) \), (A.62) reduces to

\[ \frac{\partial V^0(w, k)}{\partial k} = \int_w^{w} \left[ \Theta^M(t, y(t, k)) - \Theta^R(t, y(t, k)) \right] \frac{\partial y(t, k)}{\partial k} \, dt. \quad \text{(A.63)} \]
Because $\theta^M(w, y) > \theta^R(w, y)$ for all $(w, y)$, (A.59) implies that the integrand in (A.63) is everywhere negative. Thus, $\frac{\partial V^0(w, k)}{\partial k}$ is negative, as claimed. On the other hand, if $w_b(k) < w < k$, substituting (32) into (A.62) yields

\[
\frac{\partial V^0(w, k)}{\partial k} = -(1 + \lambda(k)) \int_w^{w_b(k)} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt + \int_{w_b(k)}^{k} \theta^M(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt
\]

\[
- \int_{w_b(k)}^{k} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt - \int_w^{w_b(k)} [\theta^M(t, y(t, k)) - \theta^R(t, y(t, k))] \frac{\partial y(t, k)}{\partial k} dt. \tag{A.64}
\]

Using the first-order condition for the placement of the bridge points (A.26) and the observation that the expression in the middle of (A.60) is constant in $(w_b, w_B)$ to re-express the second term on the right-hand side of (A.64), we obtain

\[
\frac{\partial V^0(w, k)}{\partial k} = -(1 + \lambda(k)) \int_w^{w_b(k)} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt
\]

\[
- \lambda(k) \int_{w_b(k)}^{k} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt - (1 + \lambda(k)) \int_{w_b(k)}^{w_B(k)} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt
\]

\[
- \int_{w_b(k)}^{k} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt - \int_w^{w_b(k)} [\theta^M(t, y(t, k)) - \theta^R(t, y(t, k))] \frac{\partial y(t, k)}{\partial k} dt. \tag{A.65}
\]

which reduces to

\[
\frac{\partial V^0(w, k)}{\partial k} = -(1 + \lambda(k)) \int_w^{w_B(k)} \theta^R(t, y(t, k)) \frac{\partial y(t, k)}{\partial k} dt
\]

\[
- \int_{w}^{w_B(k)} [\theta^M(t, y(t, k)) - \theta^R(t, y(t, k))] \frac{\partial y(t, k)}{\partial k} dt. \tag{A.66}
\]

Differentiating the minimum utility-constraint (A.25) with respect to $k$ and using (A.61) reveals that the first term on the right-hand side of (A.66) vanishes. By (A.60), the integrand in the second term of the right-hand side of (A.66) is always positive. Hence $\frac{\partial V^0(w, k)}{\partial k}$ is negative, as claimed.

We now show that $k < w$ implies that $V(w, k)$ is increasing in $k$. By (34),

\[
\frac{\partial V^0(w, k)}{\partial k} = \int_w^{w_b(k)} [\theta^M(t, y(t, k)) - \theta^R(t, y(t, k))] \frac{\partial y(t, k)}{\partial k} dt
\]

\[
+ \int_{w_b(k)}^{w_B(k)} [\theta^M(t, y(t, k)) - \theta^R(t, y(t, k))] \frac{\partial y(t, k)}{\partial k} dt
\]

\[
+ \int_{w_B(k)}^{w} [\theta^M(t, y(t, k)) - \theta^R(t, y(t, k))] \frac{\partial y(t, k)}{\partial k} dt. \tag{A.67}
\]
Using (32) in the first term on the right-hand side of (A.67) and rearranging yields

$$
\frac{\partial V^0(w,k)}{\partial k} = -(1 + \lambda(k)) \int_w^{w_B(k)} \theta^R(t, y(t,k)) \frac{\partial y(t,k)}{\partial k} dt + \int_k^{w_B(k)} \theta^M(t, y(t,k)) \frac{\partial y(t,k)}{\partial k} dt 
- \int_0^w \theta^R(t, y(t,k)) \frac{\partial y(t,k)}{\partial k} dt + \int_k^{w_B(k)} \left[ \theta^M(t, y(t,k)) - \theta^R(t, y(t,k)) \right] \frac{\partial y(t,k)}{\partial k} dt 
+ \int_{w_B(k)}^w \left[ \theta^M(t, y(t,k)) - \theta^R(t, y(t,k)) \right] \frac{\partial y(t,k)}{\partial k} dt. \quad (A.68)
$$

Using (A.26) in the second term on the right-hand side of (A.68) and the observation that the expression in the middle of (A.60) is constant in \((w_b, w_B)\) yields, after some re-grouping of terms,

$$
\frac{\partial V^0(w,k)}{\partial k} = -(1 + \lambda(k)) \int_w^{w_B(k)} \theta^R(t, y(t,k)) \frac{\partial y(t,k)}{\partial k} dt 
+ \int_k^{w_B(k)} \left[ \theta^M(t, y(t,k)) - \theta^R(t, y(t,k)) \right] \frac{\partial y(t,k)}{\partial k} dt. \quad (A.69)
$$

As in the case of \(k < w\), the first term on the right-hand side of (A.69) vanishes, so that

$$
\frac{\partial V^0(w,k)}{\partial k} = \int_k^{w_B(k)} \left[ \theta^M(t, y(t,k)) - \theta^R(t, y(t,k)) \right] \frac{\partial y(t,k)}{\partial k} dt. \quad (A.70)
$$

Now, either \(w < w_B(k)\) or \(w > w_B(k)\). In the first case, the integrand on the right-hand side of (A.70) is always positive, by (A.60). In the second case, we can rewrite (A.70) as

$$
\frac{\partial V^0(w,k)}{\partial k} = \int_k^{w_B(k)} \left[ \theta^M(t, y(t,k)) - \theta^R(t, y(t,k)) \right] \frac{\partial y(t,k)}{\partial k} dt 
+ \int_{w_B(k)}^w \left[ \theta^M(t, y(t,k)) - \theta^R(t, y(t,k)) \right] \frac{\partial y(t,k)}{\partial k} dt. \quad (A.71)
$$

The first-term on the right-hand side of (A.71) is positive by (A.60) and the second term is zero by (A.61). Thus, the right-hand side of (A.70) is always positive, and so is \(\frac{\partial V^0(w,k)}{\partial k}\).

This completes the proof that \(V^0(w,k)\) is single-peaked in \(k\). If the minimum-utility constraint were to bind for all candidates, the proof would be complete. When, as we assume, the minimum-utility constraint is slack for a candidate of type \(w\), we need to combine single-peakedness of \(V^0(w,k)\) with Proposition 5, which describes each voter’s preferences over candidates whose optimal schedules leave the minimum-utility constraint slack. To do this, we need continuity of the Lagrange multiplier \(\lambda\), and hence, the allocation function, as the solution passes from a region where (30) does not bind to where it does bind. This continuity is assured by Fiacco (1983, Theorem 2.4.5, p. 43), as long as the second-order conditions are satisfied with strict inequalities.
References


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


Osborne, M. J., Slivinski, A., 1996. A model of political competition with citizen-