On the political economy of income taxation

Berliant, Marcus and Gouveia, Miguel

Washington University in St. Louis, Universidade Católica Portuguesa

12 December 2018

On the Political Economy of Income Taxation*

Marcus Berliant† and Miguel Gouveia‡

December 12, 2018

Abstract

The literatures dealing with voting, optimal income taxation, implementation, and pure public goods are integrated here to address the problem of voting over income taxes and public goods. In contrast with previous articles, general nonlinear income taxes that affect the labor-leisure decisions of consumers who work and vote are allowed. Uncertainty plays an important role in that the government does not know the true realizations of the abilities of consumers drawn from a known distribution, but must meet the realization-dependent budget. Even though the space of alternatives is infinite dimensional, conditions on primitives are found to assure existence of a majority rule equilibrium when agents vote over both a public good and income taxes to finance it. JEL numbers: D72, D82, H21, H41 Keywords: Voting; Income taxation; Public good

*Support provided by National Science Foundation grant SES-9007800 is gratefully acknowledged. The comments of Dilip Mukherjee (the Coeditor), two anonymous referees, Randy Calvert, Yan Chen, Roy Gardner, Hideo Konishi, John Ledyard, Wan-hsiang Pan, Michael Smart, Jim Snyder, Max Troshkin and John Weymark are appreciated, but the authors bear full responsibility for any errors. This work was originally a chapter of M. Gouveia's dissertation, completed at the University of Rochester.

†Department of Economics, Washington University, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899 USA. Phone: (1-314) 935-8486, Fax: (1-314) 935-4156, e-mail: berliant@wustl.edu

‡Universidade Católica Portuguesa, Lisbon, Portugal, e-mail: mig@ucp.pt

1 Introduction

1.1 Background

The theory of income taxation has been an important area of study in economics. Interest in a formal theory of income taxation dates back to at least J.S. Mill (1848), who advocated an equal sacrifice approach to the normative treatment of income taxes. In terms of the modern development, Musgrave (1959) argued that two basic approaches to taxation can be distinguished: the benefit approach, which puts taxation in a Pareto efficiency context; and the ability to pay approach, which puts taxation in an equity context. Some of the early literature, such as Lindahl (1919) and Samuelson (1954, 1955), made seminal contributions toward understanding the benefit approach to taxation and tax systems that lead to Pareto optimal allocations. Although the importance of the problems posed by incentives and preference revelation were recognized, scant attention was paid to solving them, perhaps due to their complexity and difficulty.

Since the influential work of Mirrlees (1971), economists have been quite concerned with incentives in the framework of income taxation. The model proposed there postulates a government that tries to collect a given amount of revenue from the economy. For example, the level of public good provision might be fixed. Consumers have identical utility functions defined over consumption and leisure, but differing abilities or wage rates. The government chooses an income tax schedule that maximizes some objective, such as a utilitarian social welfare function, subject to collecting the needed revenue, resource constraints, and incentive constraints based on the knowledge of only the overall distribution of wages or abilities. The incentive constraints derive from the notion that individuals' wage levels or characteristics (such as productivity) are unknown to the government. The optimal income tax schedule must separate individuals as well as maximize welfare and therefore is generally second best.[1] The necessary conditions for welfare optimization generally include a zero marginal tax rate for the highest wage individual. Intuitive and algebraic derivations of this result can be found in Seade (1977), where it is also shown that some of these necessary conditions hold for Pareto optima as well as utilitarian optima. Existence of an optimal tax schedule for a modified

[1] If the government knew the type of each agent, it could impose a differential head tax. As is common in the incentives literature, one must impose a tax that accomplishes a goal without the knowledge of the identity of each agent *ex ante*.

model was demonstrated in Kaneko (1981), and then for the classical model in Berliant and Page (2001, 2006). An alternative view of optimal income taxation is as follows. Head taxes or lump sum taxes are first best, since public goods are not explicit in the model and therefore Lindahl taxes cannot be used. Second best are commodity taxes, such as Ramsey taxes. Third best are income taxes, which are equivalent to a uniform marginal tax on all commodities (or expenditure). In our view, it is not unreasonable to examine these third best taxes, since from a pragmatic viewpoint, the first and second best taxes are infeasible.

1.2 A Positive Political Model

The main objective of this research is to derive testable hypotheses. How can we explain (or model) the income tax systems we observe in the real political world? We shall attempt to answer this question with a voting model, a positive political model, in combination with the standard income tax model described above. As noted in the introduction of Roberts (1977), one does not need to believe that choices are made through any particular voting mechanism; one need only be interested in whether choices mirror the outcomes of some voting process. Thus, what is described below is an attempt to construct a potentially predictive model with both political and economic content. It contains elements of the optimal income tax literature as well as positive political theory (an excellent survey of which can be found in Calvert (1986)).

Although much of the optimal income tax literature and most of the work cited above deals with the normative prescriptions of an optimal income tax, there is a relatively small literature on voting over income taxes. Most of this literature is either restricted to consideration of only linear taxes, or does not consider problems due to information (adverse selection and moral hazard), or both. Examples that might fit primarily into the linear tax category which also involve no labor disincentives on the part of agents are Foley (1967), Nakayama (1976) and Guesnerie and Oddou (1981). Aumann and Kurz (1977) use personalized lump sum taxes in a one commodity model. Hettich and Winer (1988) present an interesting politico-economic model in which candidates seek to maximize their political support by proposing nonlinear taxes. Work disincentives are not present in the model. Chen (2000) extends their work to the more standard optimal income tax model in the context of probabilistic voting. Romer (1975), Roberts (1977), Peck (1986), and Meltzer and Richard (1981, 1983) use linear taxes in voting models with work disincentives.

Roemer (1999) restricts to quadratic tax functions with no work disincentives but with political parties. Perhaps the model closest in spirit to the one we propose below is in Snyder and Kramer (1988), which uses a modification of the standard (nonlinear) income tax model with a linear utility function. The modification accounts for an untaxed sector, which actually is a focus of their paper. This interesting and stimulating paper considers fairness and progressivity issues, as well as the existence of a majority equilibrium when individual preferences are single peaked over the set of individually optimal tax schedules. (Sufficient conditions for single peakedness are found.) Röell (1996) considers the differences between individually optimal (or dictatorial) tax schemes and social welfare maximizing tax schemes when there are finitely many types of consumers. Of particular interest are the tax schedules that are individually optimal for the median voter type. This interesting work uses quasi-linear utility and restricts voting to tax schedules that are optimal for some type. Brett and Weymark (2017) push this further in a continuum of types model by characterizing individually optimal tax schedules. Then they show, under conditions including quasi-linear utility, that if the set of tax schedules is restricted to individually optimal ones, the individually optimal tax for the median voter is a Condorcet winner.

We propose in this paper to allow general nonlinear income taxes with work disincentives in a voting model. The main problem encountered in trying to find a majority equilibrium, as well as the reason that various sets of restrictive assumptions are used to obtain such a solution in the literature, is as follows. The set of tax schedules that are under consideration as feasible for the economy (under any natural voting rule) is large in both number and dimension. Thus, the voting literature such as Plott (1967) or Schofield (1978) tells us that it is highly unlikely that a majority rule winner will exist. Is there a *natural* reduction of the number of feasible alternatives in the context of income taxation?

1.3 The Role of Uncertainty and Feasibility

The answer appears to be yes. The (optimal) income tax model has a natural uncertainty structure that has yet to be exploited in the voting context. As in the classical optimal income tax model, all *worker/consumers* have the same well-behaved utility function, but there is a nonatomic distribution of wages or abilities. In standard models, such as the Mirrlees model or its modern descendants, the draw is known by all and the aggregate revenue requirement is fixed at a scalar; it is 0 in models of pure redistribution. (This applies whether the number of consumer/workers is finite or a continuum.) Suppose that a *finite* sample is drawn from this nonatomic distribution.[2] The finite sample will be the true economy, and the revenue requirement imposed by the government can depend on the draw. In fact this dependence is just a natural extension of the standard optimal income tax model. In that model, the amount of revenue to be raised (the revenue requirement in our terminology) is a fixed parameter, something that makes perfect sense since the population in the economy and the distribution of the characteristics of that population are both fixed, and thus we can take public expenditures also as fixed. But consider now the optimal tax problem for the cases when the characteristics of the population are unknown. That is exactly what happens when we consider that the true population is a draw from a given distribution. In such circumstances, it is not reasonable to fix the revenue requirement at some exogenously given target level, but instead the revenue requirement should be a function of the population characteristics. In our analysis below, *the revenue requirements for a particular draw will be derived from the Pareto efficient level of public good provision for that draw*, leading to intrinsic variation in revenue requirements across draws.

It seems natural for us to require that any proposed tax system must be feasible (in terms of the revenue it raises) for *any* draw, as no player (including the government) knows the realization of the draw before a tax is imposed. For example, an abstract government planner might not know precisely the top ability of individuals in the economy, and therefore might not be able to follow optimal income tax rules to give the top ability individual a marginal rate of zero. The key implication of using finite draws as the true economies is that requiring *ex ante* feasibility of any proposable tax system for any draw narrows down the set of alternatives, which we call the *feasible* set, to a manageable number (even a singleton in some cases).

To be clear, the assumption is that the government must commit to a tax system (as a function of income) before knowing the realization of the draw of abilities from the distribution of abilities, this tax system cannot depend on the draw, and it must raise sufficient draw-dependent revenue no matter the draw. If we allowed the tax system to depend on the draw of abilities, we would be back in the situation the rest of the literature has found unsolvable,

[2] This assumption is similar to the one used in Bierbrauer (2011), though the purpose of that work is entirely different from ours.

since in general any tax system can be defeated by a majority for a given draw. In other words, if the government doesn't have to commit and can propose a state or draw contingent tax, we have the same situation as if there is no uncertainty and a finite number of worker/consumers with given types, so there generally will be no Condorcet winner in any given state.

Our arguments apply to finite numbers of agents. The model has a discontinuity when one goes from a finite to an infinite number of agents. In this latter case there is no uncertainty about the composition of the draw, so we do not have a continuum of *ex ante* feasibility restrictions, one for each possible draw. Instead we have only that the revenue constraint needs to be satisfied for the known population. Thus, for our purposes, even a little uncertainty is sufficient, and it is possible to view perfect certainty about the draw as a knife-edge case.[3] Moreover, there are further conceptual issues pertaining to models with a pure public good and a continuum of consumers; see Berliant and Rothstein (2000).

What is key here is not only the set of assumptions on utility or preferences, but also assumptions concerning the revenue required from each draw. The revenue requirement function was proposed and examined to some extent in Berliant (1992), and is developed further in more generality in section 2 below. We do not claim that the particular games examined here are the "correct" ones in any sense. The point of this work is that there is a natural structure and set of arguments that can be exploited in voting games over income taxes to obtain existence and sometimes uniqueness and characterization results.[4]

In relation to the literature that deals with voting over *linear* taxes, our model of voting over nonlinear taxes will not yield a linear tax as a solution without very extreme assumptions. This will be explained in section 5 below. Moreover, our second order assumption for incentive compatibility will generally be much weaker than those used in the literature on linear taxes; compare our assumptions below with the Hierarchical Adherence assumption of Roberts (1977). As noted by L'Ollivier and Rochet (1983), these second order conditions are generally not addressed in the optimal income taxation literature, though they ought to be addressed there. In what follows, we employ the results contained in Berliant and Gouveia (2001) and more generally in Berliant and Page (1996) to be sure that the second order conditions for incentive compatibility hold in our model.

---

[3] We are indebted to Jim Snyder for some of these thoughts.

[4] The variation in revenue requirements can be seen as variation in fiscal pressure on the government; see Heathcote and Tsujiyama (2017) for discussion.

The structure of the paper is as follows. First, we introduce our framework and notation in section 2. In section 3 we provide a pair of motivating examples; more examples are provided as applications in section 4, where our main results on voting over both public goods and income taxes are stated. Section 5 contains a discussion of the techniques we use in the proofs. Finally, section 6 contains conclusions and suggestions for further research. The appendix contains proofs of most results.

2 The Model

2.1 Basic Notation and Definitions

We shall develop an initial model of an endowment economy as a tool. Although it might be of independent interest, our primary purpose is to apply this model and the results we obtain to the standard optimal income tax model in the succeeding sections.

There is a single consumption good c and consumers' preferences are identical and given by the utility function $v(c) = c$, with $c \in \mathbb{R}_+$. A consumer's endowment, which is also her type, is described by $w \in [\underline{w}, \overline{w}]$, where $[\underline{w}, \overline{w}] \subseteq \mathbb{R}_{++}$. In this section the endowment can also be seen as pre tax income or, following classical terminology in Public Finance, the *ability to pay* of each agent. References to measure are to Lebesgue measure on $[\underline{w}, \overline{w}]$.

The distribution of consumers' endowments has a measurable density $f(w)$, where $f(w) > 0$ *a.s.*[5].

Let k be a positive integer and let $\mathcal{A}_k \equiv [\underline{w}, \overline{w}]^k$, the collection of all possible draws of k individuals from the distribution with density f. Formally, a *draw* is an element $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$.

In order to be able to determine what any particular draw can consume, it is first necessary to determine what taxes are due from the draw. Hence, we first assume that there is a given net revenue requirement function $R_k : \mathcal{A}_k \to \mathbb{R}$.

[5]Note that $f(\cdot)$ plays almost no role in the development to follow, in contrast with its preeminent role in the standard optimal income tax model. It may be interpreted as a subjective distribution describing the planner beliefs about the characteristics of the agents in the economy, but that consideration is immaterial for the model presented here. We have implicitly assumed that the abilities are drawn independently, but since we never use this, correlation would also be permissible. In multistage voting in a representative democracy, the equilibria are likely to be a function of f, as is often the case in signaling games. We expect to study that problem in the future.

For each $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, $R_k(w_1, w_2, ..., w_k)$ represents the total taxes due from a draw. For example, if the revenues from the income tax are used to finance a good such as schooling, then $R_k(w_1, w_2, ..., w_k)$ can be seen as: the per capita revenue requirement for providing schooling to the draw $(w_1, w_2, ..., w_k)$ multiplied by k.[6]

Although we shall begin by taking revenue requirements as a primitive, in the end we will justify this postulate by deriving revenue requirements from the technology for producing a public good.

It is important to be clear about the interpretation of R_k. One easy interpretation is that the taxing authority provides a schedule giving the taxes owed by any draw. There are several reasons that revenue requirements might differ among draws, including differences in taste for a public good that is implicitly provided, a non-constant marginal cost for production of the public good, differences in the cost of revenue collection, and so forth.

The government and the agents in the economy know the prior distribution f of types of agents in the economy[7] as well as the mapping R_k. Before moving on to consider the game - theoretic structure of the problem, it is necessary to obtain some facts about the set of tax systems that are feasible for any draw in $\mathcal{A}_k$. These are the only tax systems that can be proposed, for otherwise the voters and social planner would know more about the draw than that it consists of k people drawn from the distribution with density f. Voters can use their private information (their endowment) when voting, but not in constructing the feasible set. For otherwise either each voter will vote over a different feasible set, or information will be transmitted just in the construction of the feasible set.

An *individual revenue requirement*[8] is a function $g : W \to \mathbb{R}$ that takes w to tax liability.

Clearly, there will generally be a range of individual revenue requirements consistent with any map R_k. Our next job is to describe this set formally. Fix

[6] Actually, regarding schooling, there is a separate literature on the political economy of public supplements for such goods. The formal structure is slightly different from what we consider in this paper; see Gouveia (1997).

[7] Actually, all they need to know is the support of that distribution.

[8] Even though this is simply a tax function on endowments, we will reserve the terminology "tax function" for an environment with incentives to simplify the exposition.

k and R_k. Let

$$\begin{aligned} G_k \equiv \Bigg\{ & g : [\underline{w}, \overline{w}] \to \mathbb{R} \mid g \text{ is measurable}, \\ & \sum_{i=1}^{k} g(w_i) \geq R_k(w_1, w_2, ..., w_k) \text{ a.s. } (w_1, w_2, ..., w_k) \in \mathcal{A}_k \Bigg\} \end{aligned}$$

G_k is the set of all *individual* revenue requirements that collect enough revenue to satisfy R_k. $G_k \neq \emptyset$ if almost surely for $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, $\sum_{i=1}^{k} w_i \geq R_k(w_1, w_2, ..., w_k)$. The constraint that the revenue requirement be satisfied for each draw restricts the feasible set significantly.

2.2 From Collective to Individual Revenue Requirements

In order to examine the set of feasible individual revenue requirements described above, more structure needs to be introduced. It is obvious that some feasible g's will raise strictly more taxes than necessary to meet $R_k(w_1, w_2, ..., w_k)$ for any $(w_1, w_2, ..., w_k)$. We now search for the minimal elements of the sets G_k. We call the set of such elements G_k^*. In other words, we search for individual revenue requirements $g \in G_k^* \subseteq G_k$ with the following property: there is no g' such that almost surely for $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, $R_k(w_1, w_2, ..., w_k) \leq \sum_{i=1}^{k} g'(w_i)$; almost surely for $w \in [\underline{w}, \overline{w}]$, $g'(w) \leq g(w)$; and there exists a set of positive Lebesgue measure in $[\underline{w}, \overline{w}]$ where $g'(w) < g(w)$.

To this end, define a binary relation $\succeq$ over G_k by $g \succeq g'$ if and only if $g(w) \geq g'(w)$ for almost all $w \in [\underline{w}, \overline{w}]$. Let

$$\mathcal{G}_k \equiv \{B \subseteq G_k | \ B \text{ is a maximal totally ordered subset of } G_k\}.$$

By Hausdorff's Maximality Theorem (see Rudin (1974, p. 430)), $\mathcal{G}_k \neq \emptyset$. Finally, define

$$G_k^* \equiv \{g : [\underline{w}, \overline{w}] \to \mathbb{R} | \ \exists B \in \mathcal{G}_k \text{ such that } g(w) = \inf_{g' \in B} g'(w) \text{ a.s.}\}.$$

G_k^* is nonempty.

If $g \in G_k \backslash G_k^*$ is proposed as an alternative to $g^* \in G_k^*$, $\exists g' \in G_k^*$ that is unanimously weakly preferred to g.

2.3 Notation for the Optimal Income Tax Model

Having dispensed with preliminaries, we now turn to the voting model with incentives based on Mirrlees (1971). The three goods in the model are a composite consumption good, whose quantity is denoted by c; labor, whose quantity is denoted by l; and a pure public good, whose quantity is denoted by x. Consumers have an endowment of 1 unit of *labor/leisure*, no consumption good, and no public good.[9] Let $u : \mathbb{R}_+ \times [0,1] \times \mathbb{R}_+ \times [\underline{w}, \overline{w}] \to \mathbb{R}$ be the utility functions of the agents, writing $u(c,l,x,w)$ as the utility function of type w, where u is twice continuously differentiable. Subscripts represent partial derivatives of u with respect to the appropriate arguments. The parameter w, an agent's type, is now to be interpreted as the wage rate or productivity of an agent. Thus w is the value of an agent of type w's endowment of labor. The gross income earned by an agent of type w is $y = w \cdot l$ and it equals consumption when there are no taxes.

A *tax system* is a function $\tau : \mathbb{R} \to \mathbb{R}$ that takes y to tax liability. A *net income function* $\gamma : \mathbb{R} \to \mathbb{R}$ corresponds to a given τ by the formula $\gamma(y) \equiv y - \tau(y)$.

First we discuss the typical consumer's problem under the premise that the consumer does not lie about its type, and later turn to incentive problems. A consumer of type $w \in [\underline{w}, \overline{w}]$ is confronted with the following maximization problem in this model:

$$\begin{aligned} \max_{c,l} u(c,l,x,w) \text{ subject to } w \cdot l - \tau(w \cdot l) &\geq c \text{ with } \tau, x \text{ given}, \\ \text{and subject to } c &\geq 0,\ l \geq 0,\ l \leq 1. \end{aligned}$$

For fixed τ, we call arguments that solve this optimization problem $c(w)$ and $l(w)$ (omitting τ and x) as is common in the literature. Define $y(w) \equiv w \cdot l(w)$.

The public good financed by the revenue raised through the income tax is usually excluded from models of optimal income taxation due to the complexity introduced, but here the cost of the public good will be used to derive the revenue requirements function. Let the cost function for the public good in terms of consumption good be $H(x)$, which is assumed to be C^2.

Let $F_k : \mathcal{A}_k \to T_k \times \mathbb{R}_+$ be a correspondence defined by:

$$F_k(w_1, w_2, ..., w_k) \equiv \left\{ (\tau, x) \in T_k \times \mathbb{R}_+ \mid \sum_{i=1}^{k} \tau(y(w_i)) \geq H(x) \right\}.$$

With this in hand, a straightforward definition of majority rule equilibrium follows: a *majority rule equilibrium* for draws of size k is a correspondence

[9] It would be easy to add an endowment of conumption good for consumers, but that would complicate notation.

M_k mapping $(w_1, w_2, ..., w_k)$ into $F_k(w_1, w_2, ..., w_k)$ such that for almost every $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, for every $(\tau, x) \in M_k(w_1, w_2, ..., w_k)$ (with associated $y(w)$), there is no subset D of $\{w_1, w_2, ..., w_k\}$ of cardinality greater than $k/2$ along with another pair $(\tau', x') \in F_k(w_1, w_2, ..., w_k)$ (with associated $y'(w)$) such that $u(y'(w) - \tau'(y'(w)), y'(w)/w, x', w) > u(y(w) - \tau(y(w)), y(w)/w, x, w)$ for all $w \in D$.

3 Motivating Examples

To motivate our analysis, we provide a pair of simple examples that can be solved.

Example 1: Take

$$\begin{aligned} u(c, l, x, w) &= c - \frac{l^2}{2} - w \cdot \frac{x^{-2}}{2} \\ H(x) &= \frac{x^2}{2} \end{aligned}$$

The marginal cost of the public good is x. The the marginal willingness to pay of type w for the public good is $w \cdot x^{-3}$, so the total marginal willingness to pay for the draw $(w_1, w_2, ..., w_k)$ is $x^{-3} \sum_{i=1}^{k} w_i$. Setting this equal to marginal cost to solve for the Pareto efficient level of public good provision (that will be unique), we obtain:

$$x(w_1, w_2, ..., w_k) = \left(\sum_{i=1}^{k} w_i \right)^{\frac{1}{4}}$$

A reason why the isoelastic case might be interesting comes from the fact that it is a suitable case for the purpose of carrying out empirical tests of the model, given that the correct way to aggregate abilities (or tastes) in this particular case is simply to sum them.

The aggregate revenue requirement function is:

$$R_k(w_1, w_2, ..., w_k) = H\left(x(w_1, w_2, ..., w_k)\right) = \frac{1}{2}\sqrt{\sum_{i=1}^{k} w_i}$$

Next, take $\underline{w} = 1$, $\overline{w} = 2$, and let $\widetilde{w}$ be the median type of a draw. Then as we will see below, if $k \geq 2$ and $1.5 \leq \widetilde{w} \leq 2$, the minimal individual revenue requirements are indexed by $\widetilde{w}$ and given by[10]

$$\begin{aligned} g(w;\widetilde{w}) &= \frac{1}{2}\left(k\widetilde{w}\right)^{\frac{1}{2}}/k + \frac{1}{4}\left(k\widetilde{w}\right)^{-\frac{1}{2}}\left(w-\widetilde{w}\right) \\ &= \frac{1}{4}\left[\sqrt{\frac{\widetilde{w}}{k}} + \sqrt{\frac{1}{k\widetilde{w}}}\cdot w\right] \end{aligned}$$

In an endowment economy, this is the tax on endowments most preferred by type $\widetilde{w}$ among those satisfying the aggregate revenue constraints. In this particular case, it is a linear tax. The next step is to implement it in an optimal income tax economy.

Applying the first order approach to incentive compatibility[11] given in the differential equation (3) and $\frac{d\tau}{dy} = 1 - \frac{d\gamma}{dy}$, the income tax function is given by the solution to:[12]

$$\frac{d\tau}{dy} = 1 - \frac{y}{w^2}$$

Inverting g and solving for w in terms of τ,

$$w = 4\sqrt{k\widetilde{w}}\tau - \widetilde{w}$$

so

$$\frac{d\tau}{dy} = 1 - \frac{y}{\left[4\sqrt{k\widetilde{w}}\tau - \widetilde{w}\right]^2}$$

This ordinary differential equation has a solution at through every point. To choose the best of these, take the one that has the marginal tax rate zero for the top type $\overline{w} = 2$. For the top type, it is the solution that goes through $(\overline{\tau},\overline{y}) = \left(\frac{1}{2\sqrt{k\widetilde{w}}} + \frac{1}{4}\sqrt{\frac{\widetilde{w}}{k}}, 4\right)$. This will be the Condorcet winner for any draw with median $\widetilde{w} \geq 1.5$.

Example 2: One point of this example is that although we will restrict to quasi-linear utility functions for the general theory, that might not be necessary. Take

$$\begin{aligned} u(c,l,x,w) &= \min\left(c, w\cdot[1-l]\right) - w\cdot\frac{x^{-2}}{2} \\ H(x) &= \frac{x^2}{2} \end{aligned}$$

[10] To keep calculations simple, we focus on draws where the median is at least 1.5.

[11] The second order condition for incentive compatibility will be satsified because $\frac{\partial g(w;\widetilde{w})}{\partial w} > 0$.

[12] Although we know that a solution exists and through any point it is unique, actually solving the ODE explicitly is another matter entirely.

The aggregate revenue requirements function is unchanged from Example 1. Setting $c = w \cdot [1 - l]$,

$$y - \tau = w - y$$

Therefore,

$$\begin{aligned} \tau(y) &= 2y - w \\ &= 2y - 4\sqrt{k\widetilde{w}}\tau(y) + \widetilde{w} \end{aligned}$$

and thus

$$\tau(y) = \frac{2y + \widetilde{w}}{1 + 4\sqrt{k\widetilde{w}}}$$

Remarks: The single crossing of individual revenue requirements results from the combination of the assumptions on utility and the idea that the aggregate revenue requirements must be satisfied for any draw. We prove in Lemma 5 below that when we implement the individual revenue requirements and impose second best efficiency, the single crossing property is inherited by the income tax implementations. One common feature of our individual revenue requirement functions is that there is a *switch point*, indexed by $\widetilde{w}$ in our examples here, that represents the individual revenue requirement that minimizes that type's tax liability among all individual revenue requirements satisfying the aggregate revenue requirements for all draws. This is not actually necessary for our general results, and is not used in the proofs once we obtain single crossing of individual revenue requirements. However, as seen from Example 1, provided that g is strictly increasing, the optimization point for type w under the (optimal) income tax framework will correspond to tax liability $g(w; \widetilde{w})$. Therefore, using the standard diagrams from optimal tax theory,[13] the Condorcet winner will correspond to the best implementation (solution to the ordinary differential equation) of the revenue requirement function that minimizes the the tax liability of the median type of the draw, $g(w; \widetilde{w})$. Thus, the switch point is inherited by the optimal income tax implementation of the individual revenue requirements. The fact that we do not use the switch point once we have single crossing of individual revenue requirements allows room for expansion of our results.

[13] See Seade (1977).

4 Voting Over Income Taxes and a Public Good

4.1 Basic Assumptions

These basic assumptions will be maintained throughout the remainder of this paper.

We will now use ideas inspired by Bergstrom and Cornes (1983) to obtain a unique Pareto optimal level of public good for each draw, so the revenue requirement function is well-defined.

The major assumption that we make to obtain results, beyond requiring sufficient revenue to finance the public good for each draw, is that utility is quasi-linear and separable to a certain degree:[14]

$$u(c, l, x, w) = c + b(l, w) + r(x, w)$$

We assume throughout that $\partial b/\partial l < 0$*,* $\partial^2 b/\partial l^2 < 0$*,* $\partial r/\partial x > 0$*,* $\partial^2 r/\partial x^2 < 0$*;* $dH(x)/dx > 0$ *and* $d^2 H(x)/dx^2 \geq 0$*.*

From this, it follows that utility is strictly monotonic in consumption commodity (a good) and labor (a bad). There are several more remarks to be made. First, if we had more than 1 level of public good possible for given parameters, as is standard in public goods models without the Bergstrom-Cornes type of assumptions, then we would have another dimension to vote over, namely the level of the public good. Generally speaking, this would be death to our analysis. We would have the usual problems. Second, if we made utility more general, for example allowing the subutility function $r(x, w)$ to depend on consumption good c or labor l or both, then the public good level and hence the aggregate revenue requirement function would depend on the tax function, and that tax function would depend on the public good level and hence the aggregate revenue requirement function. Thus, the aggregate revenue requirement would not be exogenous and likely not uniquely defined. Probably it is a solution to a fixed point problem, possibly a contraction under some circumstances. Finally, when production of the public good is not constant returns to scale, there is a potential issue of profit distribution. However, when utility is quasi-linear, this isn't really an issue.

The bottom line is that something has to be done to shut down the feedback between tax liabilities and the optimal level of the public good. The Bergstrom

[14] In this case we are also using w as a taste parameter. That interpretation is quite common in both the optimal tax literature and the literature on self-selection.

and Cornes (1983) specification is a natural starting point and actually is more general than some of the separability assumptions used in the optimal nonlinear income tax literature.

We shall be using intensively the Lindahl-Samuelson condition for our specialized economy. It is given in (2) below.

Let $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, and let c_i and l_i denote the consumption and labor supply of the ith member of the draw respectively. Then production possibilities for this given draw are:

$$\sum_{i=1}^{k} w_i \cdot l_i - \sum_{i=1}^{k} c_i \geq H(x). \tag{1}$$

Fix $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$. We define an allocation to be *interior* if the associated level of public good x satisfies $x > 0$ and $H(x) < \sum_{i=1}^{k} w_i$.[15] Given our assumptions, a necessary and sufficient condition for an interior Pareto optimum is: $\sum_{i=1}^{k} \partial r(0, w_i)/\partial x > dH(0)/dx$ and there is $\overline{x}$ such that $\sum_{i=1}^{k} \partial r(\overline{x}, w_i)/\partial x < dH(\overline{x})/dx$ and $H(\overline{x}) < \sum_{i=1}^{k} w_i$. More usefully, we shall assume the following sufficient condition:

For all $w \in [\underline{w}, \overline{w}]$, $\partial r(0, w)/\partial x > dH(0)/dx$, and $k \cdot \partial r(H^{-1}(k\underline{w}), w)/\partial x < dH(H^{-1}(k\underline{w}))/dx$.

Lemma 1: Under the basic assumptions listed above, for any given draw $(w_1, w_2, ...w_k)$, there exists an interior Pareto optimal allocation; moreover, for all interior Pareto optimal allocations, the public good level is the same.

Proof: The Pareto optimal allocations are solutions to: $\max u(c_1, l_1, x, w_1)$ subject to $u(c_i, l_i, x, w_i) \geq \overline{u}_i$ for $i = 2, 3, ...k$ and subject to (1) where the maximum is taken over c_i, l_i, $(i = 1, ..., k)$ and x. Restricting attention to interior optima, we have the Lindahl-Samuelson condition for this problem:

$$\sum_{i=1}^{k} \partial r(x, w_i)/\partial x = dH(x)/dx. \tag{2}$$

Since this equation is independent of c_i and l_i for all i, the Pareto optimal level of public good provision is independent of the distribution of income and

[15] See Bergstrom and Cornes (1983) for an explanation of why we need to restrict the analysis to interior allocations.

consumption for the given draw. Given our assumptions on r and H, there exists a unique level of public good that solves (2).

For the class of utility functions defined above we can thus solve for x as an (implicit) function of $(w_1, w_2, ..., w_k)$, and obtain the revenue requirement function $R_k(w_1, w_2, ..., w_k) \equiv H(x(w_1, w_2, ..., w_k))$.

4.2 Main Results

To simplify notation, we shall abbreviate derivatives of functions of only one variable using primes, e.g. $H'(x) \equiv dH(x)/dx$.

Theorem 1: Let $k \geq 2$ and let $u(c, l, x, w) = c + b(l, w) + w \cdot s(\widehat{r}(x))$, $H(x) = m \cdot \widehat{r}(x)$, where $d\widehat{r}'(x) > 0$, $\widehat{r}''(x) \geq 0$, $s'(r) > 0$, $s''(r) < 0$, $2s''(r)^2 > s'''(r) \cdot s'(r)$, $m > 0$. Then for any draw in $\mathcal{A}_k$, the one stage voting game over interior (τ, x) has a majority rule equilibrium.

Proof: See the Appendix.

Examples covered by this theorem include the following:

A. $u(c, l, x, w) = c + b(l, w) + w \cdot s(x)$, $H(x) = m \cdot x$, where $\widehat{r}(x) = x$, $s'(x) > 0$, $s''(x) < 0$, $s'''(x) \leq 0$, and $m > 0$.

B. $u(c, l, x, w) = c + b(l, w) + \frac{w}{1-\alpha}x^{1-\alpha}$, $H(x) = \frac{m}{\beta} \cdot x^{\beta}$, with $\alpha > 1$, $\beta \geq 1$. In this case, $\widehat{r}(x) = \frac{1}{\beta}x^{\beta}$, $s(r) = \frac{\beta^{\frac{1-\alpha}{\beta}}}{1-\alpha} r^{\frac{1-\alpha}{\beta}}$.

Example 1 above is covered by B, with $\alpha = 3$, $\beta = 2$.

Theorem 2: Let $u(c, l, x, w) = c + b(l, w) + a \cdot \frac{\widehat{r}(x)}{\widehat{r}'(w)}$, where $a \geq 1$, $\widehat{r}'(x) > 0$, $\widehat{r}''(x) < 0$, $\lim_{x \to 0} \widehat{r}'(x) = \infty$, $2(\widehat{r}''(x))^2 > \widehat{r}'''(x) \cdot \widehat{r}'(x)$. Let $H(x) = m \cdot x$, where $m > 0$, $a \geq m$. Then for any draw in $\mathcal{A}_k$, the one stage voting game over interior (τ, x) has a majority rule equilibrium.

Proof: See the Appendix.

An example covered by this Theorem is the CES case: $u(c, l, x, w) = c + b(l, w) + a \cdot w^p \cdot \frac{x^{1-p}}{1-p}$, where $a \geq 1$, $p > 1$, and $H(x) = m \cdot x$, where $m > 0$, $a \geq m$. Here, $\widehat{r}(x) = \frac{x^{1-p}}{1-p}$.

It is important to note that Theorems 1 and 2 rely on different techniques of proof. That makes integration of the results tricky. A complicating factor is that revenue requirements derived here depend not only on the first derivative of the production or cost function, but on its level as well.

5 Discussion of the Underlying Techniques

5.1 Single Crossing Individual Revenue Requirements

With a view toward future extensions of Theorems 1 and 2, we state some natural assumptions on R_k that *will be satisfied by the revenue requirements derived in the the course of proving Theorems 1 and 2.* The first of these implies that position in the draw (first, second, etc.) does not matter. All that matters in determining the revenue to be extracted from a draw is which types are drawn from the distribution.

Definition: A revenue requirement function R_k is said to be *symmetric* if for each k and for each $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, for any permutation σ of $\{1, 2, ..., k\}$, $R_k(w_1, w_2, ..., w_k) = R_k(w_{\sigma(1)}, w_{\sigma(2)} ..., w_{\sigma(k)})$.

We will use the assumption that R_k is C^2. That is not a strong assumption. The reason is that the assumption that R_k is C^2 is generic in the appropriate topology; that is, C^2 R_k's will uniformly approximate any continuous R_k (Hirsch (1976, Theorem 2.2)).[16] We will also assume that R_k is smoothly monotonic:

Definition: A revenue requirement function R_k is said to be *smoothly monotonic* if for any $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, $\partial R_k(w_1, w_2, ..., w_k)/\partial w_i > 0$ for $i = 1, 2, ..., k$.

This assumption requires that increasing the ability or wage of any individual in a draw increases the total tax liability of the draw. One could successfully use weaker assumptions with this framework, but at a cost of greatly complicating the proofs.[17]

A major step in our analysis, that we have mainly relegated to other papers that are cited in the bibliography, Berliant and Gouveia (2001) and Berliant and Page (1996), is to implement the individual revenue requirement g using an income tax, an indirect mechanism. This is, in essence, what is known in the literature as the "Taxation Principle." A sufficient (and virtually necessary) condition is that g be increasing in type, w.[18] If g is anywhere decreasing

[16] This idea is also used to justify differentiability in the smooth economies literature.

[17] One particular case ruled out is the one of constant per capita revenues. In our model this situation implies constant individual revenue requirements, i.e. a head tax, clearly an uninteresting situation even though it is first-best. It can be handled as a limit of the cases considered here. This case includes the particular situation where the government wants to raise zero fiscal revenue.

[18] The case $g'(w) = 0$ for some types w could be handled, but it creates some technical problems because g is not necessarily invertible.

in type, the net income function can cut the indifference curve of an agent, creating a gap in the assignment of types to tax liability and ruining the implementation of g by an income tax. In order to use the first order approach to incentive compatibility, for example, we must make further assumptions, namely the second order conditions.[19] These second order conditions are equivalent to the property that g is increasing.

Turning next to aggregate revenue requirements R_k, we relate the property of increasing R_k ($\partial R_k(w_1, w_2, ..., w_k)/\partial w_i > 0$ for $i = 1, 2, ..., k$) to increasing $g \in G_k^*$. Suppose that there are $w, w' \in [\underline{w}, \overline{w}]$ with $w' > w$. Then, by definition of G_k^*, there is a draw $(w_1, w_2, ..., w_k)$ with $w = w_i$ for some i and $\sum_{j=1}^{k} g(w_j) = R_k(w_1, w_2, ..., w_k)$. Now replace w with w', namely set $w_i = w'$, leaving all other elements of the draw the same. Then $R_k(w_1, w_2, ..., w', ..., w_k) > R_k(w_1, w_2, ..., w_k)$. Since g is feasible, $g(w_1)+g(w_2)+\cdots+g(w')+\cdots+g(w_k) \geq R_k(w_1, w_2, ..., w', ..., w_k) > R_k(w_1, w_2, ..., w_k) = \sum_{j=1}^{k} g(w_j)$, so $g(w') > g(w)$.

The next step is to introduce two sets of assumptions where the elements of the set of feasible and minimal individual revenue requirements G_k^* are single crossing,[20] i.e. each pair of g's will cross only once.[21] The first set of assumptions will be implied by the postulates of Theorem 1, whereas the second set of assumptions will be implied by the postulates of Theorem 2. Thus, future generalizations of our main results will likely use the lemmas below.

These two sets of assumptions reflect how collective revenue requirements change with polarization of the draw. Loosely speaking, they are opposites of one another. The first set of assumptions has collective revenue requirements decreasing as a draw becomes more polarized, whereas the second set of assumptions has collective revenue requirements increasing as a draw becomes more polarized.

Definition: A revenue requirement function $R_k(w_1, ...w_k)$ is *argument-additive* if $R_k(w_1, w_2, ..., w_k) \equiv Q(\sum_{i=1}^{k} w_i)$. Let Q' denote $\frac{dQ}{d\sum_{i=1}^{k} w_i}$.

[19] We note that much of the recent literature on optimal taxation verifies the second order conditions numerically, not analytically.

[20] In fact, under stronger assumptions, it is possible to show that the set of feasible and minimal individual revenue requirements is a singleton, rendering voting trivial. In that analysis, it's useful to have the size of the draw, k, unknown to the planner as well. We omit this analysis for the sake of brevity.

[21] A G_k^* with single crossing g's generates a trade-off where raising more taxes from one type of voter allows less revenue to be raised from another type, as in the conventional income tax model.

Lemma 2: Let $k \geq 2$ and let the revenue requirement function $R_k(w_1, w_2, ..., w_k)$ be argument-additive with $Q'' < 0$. Then, we have that $\forall g \in G_k^*$, g is as follows:

– For $\widetilde{w} \geq (\underline{w} + \overline{w})/2$, $g \in G_k^*$ implies:

A) $g(w; \widetilde{w}) = Q(k\widetilde{w})/k + Q'(k\widetilde{w}) \cdot (w - \widetilde{w})$ if $w \leq \widetilde{w} + (k-1) \cdot (\widetilde{w} - \underline{w})$.

B) $g(w; \widetilde{w}) = Q((k-1)\underline{w} + w) - ((k-1)/k) \cdot Q(k\widetilde{w}) + (k-1) \cdot Q'(k\widetilde{w}) \cdot (\widetilde{w} - \underline{w})$ if $w > \widetilde{w} + (k-1) \cdot (\widetilde{w} - \underline{w})$.

– For $\widetilde{w} < (\underline{w} + \overline{w})/2$, $g \in G_k^*$ implies:

C) $g(w; \widetilde{w}) = Q(k\widetilde{w})/k + Q'(k\widetilde{w}) \cdot (w - \widetilde{w})$ if $w \geq \widetilde{w} - (k-1) \cdot (\overline{w} - \widetilde{w})$

D) $g(w; \widetilde{w}) = Q((k-1)\overline{w} + w) - ((k-1)/k) \cdot Q(k\widetilde{w}) + (k-1) \cdot Q'(k\widetilde{w}) \cdot (\widetilde{w} - \overline{w})$ if $w < \widetilde{w} - (k-1) \cdot (\overline{w} - \widetilde{w})$

where $\widetilde{w} \in [\underline{w}, \overline{w}]$. Thus, $dg(w; \widetilde{w})/dw > 0$ except at a finite number of points. Furthermore, $\forall w \in [\underline{w}, \overline{w}]$, $g(w; \widetilde{w})$ is single caved[22] in $\widetilde{w}$ and attains a minimum at $\widetilde{w} = w$. Finally, any pair of g's in G_k^* will cross once: for any $g, g' \in G_k^*$, there exists a $\widetilde{w}, \widetilde{w}' \in [\underline{w}, \overline{w}]$, $\widetilde{w} < \widetilde{w}'$ such that $g(\underline{w}) > g'(\underline{w})$ implies $g(w) > g'(w)$ for all $w \in [\underline{w}, \widetilde{w})$, $g(w) = g'(w)$ for all $w \in [\widetilde{w}, \widetilde{w}']$ and $g(w) < g'(w)$ for all $w \in (\widetilde{w}', \overline{w}]$.

Proof: See the Appendix.

The implication of our feasibility approach in this case is that feasible tax functions turn out to be parameterized by $\widetilde{w}$. The intuition for this result is quite simple. Consider (for the moment) the case where the distribution of endowments is not bounded above or below. Since the revenue requirement Q is concave, so is the per capita revenue requirement Q/k. But then, only the tangents to Q/k can be tax functions, since any linear combination of taxes has to be greater than or equal to the per capita requirement. The $\widetilde{w}$'s correspond to the arguments of the per capita revenue functions at the tangency points. The statement of the theorem is slightly more complex because this intuition may not work near the limits $\underline{w}$ or $\overline{w}$.

Note that the marginal rates in branch B are lower than the rate in branches A and C (the tangent branches), that in turn is lower than those in branch D. In the argument-additivity case, concavity implies that per-capita revenue requirements decrease with the polarization of the sample.

[22] A function f is single-caved if $-f$ is single peaked.

We look next at a second case that implies single crossing of revenue requirement functions in G_k^*, using an assumption that we call *limited complementarity.* This assumption implies that revenue requirements are maximal for draws consisting of at most two types of consumers. Maximal revenue draws for type w are polarized draws, i.e. they consist of people of type w and people of the type most unlike w, either $\underline{w}$ or $\overline{w}$. This is a stark way to capture the idea that higher heterogeneity in an economy leads to higher fiscal revenue needs. As such, one may find it helpful to associate this property with some notion of convexity of the collective revenue requirement function on each individual endowment. The revenue requirements to be discussed thus have the properties that revenue collection must increase with the agents' endowments and with polarization of the distribution of endowments.

The basic idea detailed below is that there is just a little more freedom, namely one dimension of flexibility, in specifying the class of individual revenue requirements consistent with an aggregate revenue requirement in this second case. That flexibility is given by the choices of $g(\underline{w})$ and $g(\overline{w})$ subject to $R_k(\overline{w}_1, \overline{w}_2, ..., \overline{w}_{k/2}, \underline{w}_{k/2+1}, ..., \underline{w}_k) = \frac{k}{2}(g(\overline{w}) + g(\underline{w}))$. Once they are fixed, then we will have a class of single crossing individual revenue requirements functions (under some further assumptions).

Fix k. Four conditions on g evaluated at $\underline{w}$ and $\overline{w}$ are:

C1. $g(\underline{w}) \geq R_k(\underline{w}, \underline{w}, ..., \underline{w})/k$.
C2. $g(\overline{w}) \geq R_k(\overline{w}, \overline{w}, ..., \overline{w})/k$.
C3. For k even:

$$R_k(\overline{w}_1, \overline{w}_2, ..., \overline{w}_{k/2}, \underline{w}_{k/2+1}, ..., \underline{w}_k) = \frac{k}{2}(g(\overline{w}) + g(\underline{w})).$$

C4. For k odd:

$$R_k(\underline{w}_1, \underline{w}_2, ..., \underline{w}_{(k-1)/2}, \overline{w}_{(k+1)/2}, ..., \overline{w}_k) + R_k(\underline{w}_1, \underline{w}_2, ..., \underline{w}_{(k+1)/2}, \overline{w}_{(k+3)/2}, ..., \overline{w}_k) = k(g(\underline{w}) + g(\overline{w})).$$

C1 and *C2* are mere feasibility conditions. *C3* and *C4* mean that a draw consisting of both extreme types provides a "worst case scenario" against which feasibility of any upper and lower values of the individual revenue requirement function g must be assessed.

Definition: The *set of admissible extreme revenue requirements* is:

$$EG_k \equiv \left\{ (g(\underline{w}), g(\overline{w})) \in \mathbb{R}_+^2 \mid (g(\underline{w}), g(\overline{w})) \text{ verifies } C1\text{-}C4. \right\}.$$

Define the switching function $W : [\underline{w}, \overline{w}] \times [\underline{w}, \overline{w}] \to \{\underline{w}, \overline{w}\}$ by $W(w, w^*) \equiv \underline{w}$ if $w \geq w^*$, and $W(w, w^*) \equiv \overline{w}$ if $w < w^*$. This function defines which extreme type is the one most unlike a given type w, either $\underline{w}$ or $\overline{w}$, relative to w^*.

Definition: A revenue requirement function R_k is said to satisfy *limited complementarity* if for each $(g(\underline{w}), g(\overline{w})) \in EG_k$ there exists a switching point $w^* \in [\underline{w}, \overline{w}]$ such that for all $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$ the following holds:

– For k even:

$$R_k(w_1, w_2, ..., w_k) \leq \sum_{i=1}^{k} [R_k(w_1^i, w_2^i, ..., w_k^i) \cdot \frac{2}{k} - g(W(w_i, w^*))]$$

where $w_j^i = w_i$ for $j = 1, ..., k/2$ and $w_j^i = W(w_i, w^*)$ for $j = k/2 + 1, ..., k$.

– For k odd:

$$R_k(w_1, w_2, ..., w_k) \leq \sum_{i=1}^{k} [R_k(w_1^{ia}, w_2^{ia}, ..., w_k^{ia})/k + R_k(w_1^{ib}, w_2^{ib}, ..., w_k^{ib})/k - g(W(w_i, w^*))]$$

where $w_j^{ia} = w_i$ for $j = 1, ..., (k-1)/2$ and $w_j^{ia} = W(w_i, w^*)$ for $j = (k+1)/2, ..., k$ and $w_j^{ib} = w_i$ for $j = 1, ..., (k+1)/2$ and $w_j^{ib} = W(w_i, w^*)$ for $j = (k+3)/2, ..., k$.

Definition: A revenue requirement function R_k is said to satisfy *Edgeworth substitutability* if $\partial^2 R_k / \partial w_i \partial w_j < 0$ for $i \neq j$.

This assumption means that the individual marginal contributions for the revenue requirement out of a draw decline when the type of another individual in the draw increases.

The next result establishes that requirements in G_k^* cross exactly once.

Lemma 3: Let k be a positive integer. Suppose that R_k satisfies limited complementarity and Edgeworth substitutability. Then, $\forall g \in G_k^*$, g is continuous and $dg/dw > 0$ where it is differentiable, and for any $g, g' \in G_k^*$, there exists a $\widetilde{w} \in [\underline{w}, \overline{w}]$ such that $g(\underline{w}) > g'(\underline{w})$ implies $g(w) > g'(w)$ for all $w \in [\underline{w}, \widetilde{w})$ and $g(w) < g'(w)$ for all $w \in (\widetilde{w}, \overline{w}]$. Moreover, for any $g, g' \in G_k^*$ with switching points w^* and w'^*, respectively, $g(\underline{w}) > g'(\underline{w})$ implies $w^* > w'^*$. Finally, the $g \in G_k^*$ that minimizes $g(\widehat{w})$ has a switch point $w^* = \widehat{w}$.

Proof: See the Appendix.

Remark: Given the assumptions on R_k and the fact (proved in the Appendix) that

$$g(w_i) = R_k(w_1^i, w_2^i, ..., w_k^i) \cdot \frac{2}{k} - g(W(w^i, w^*))$$

for k even, we have that each $g(w)$ has at most one non-differentiable point, which is at the switch point w^*. A similar result holds for k odd.

Remark: Notice that the notion of single crossing used in Lemma 2 is weaker than the notion used in Lemma 3. Thus, we will use the notion of single crossing from Lemma 2 in what follows.

Lemmas 2 and 3 illustrate that existence of a political equilibrium determining the shape of tax schedules does not necessarily imply a given pattern of taxation. Notice also that the shape of the distribution of endowments f does not have in itself any relevant information to predict the shape of the income tax schedules chosen by majority rule, since we have not used it anywhere. Revenue requirements R_k is all that is needed.[23]

5.2 Single Crossing Optimal Tax Functions

Next, some results from the literature on optimal income taxation and implementation theory are used to construct the best income tax function that implements a given individual revenue requirement. The discussion will be informal, but made formal in the theorems and their proofs.

The problem confronting a *worker/consumer* of type w given net income schedule γ is $\max_l u(\gamma(w \cdot l), l, x, w)$. Using the particular form of utility that we have specified, the first order condition from this problem is $\frac{d\gamma}{dy} \cdot w + \partial b/\partial l = 0$. Rearranging,

$$\frac{d\gamma}{dy} = -\frac{\partial b(l, w)}{\partial l} \cdot \frac{1}{w}.$$

For this tax schedule, we want the consumer of type w to pay exactly the taxes due, which are $g(w)$ for some $g \in G_k^*$. If g is strictly increasing, g is invertible. If we assume (for the moment) that $g(w)$ is continuously differentiable, then g^{-1}, which maps tax liability to ability (or wage), is well-defined and continuously differentiable. Substituting into the last expression,

[23] With these preliminary results in hand, it would be possible to prove that a majority rule equilibrium exists for the endowment economy. Since this not our main aim, for the sake of brevity it is omitted.

$$\frac{d\gamma}{dy} = -\frac{\partial b(l,w)}{\partial l} \cdot \frac{1}{g^{-1}(y-\gamma)} \equiv \Phi(\gamma, y). \tag{3}$$

As in Berliant (1992), a standard result from the theory of differential equations yields a family of solutions to this differential equation.[24] Berliant and Gouveia (2001) show that (3) has global solutions if $g' > 0$, $g(\underline{w}) \geq 0$.

Of course, as L'Ollivier and Rochet (1983) point out, the second order conditions must be checked to ensure that solutions to (3) do not involve bunching, which means that consumers do optimize in (3) at the tax liability given by g.[25] This was done in Berliant and Gouveia (2001), where the Revelation Principle[26] was used to construct strictly increasing post tax income functions $\theta(w) = y(w) - g(w)$ that implement $g(w)$, where $g'(w) > 0$. Since we then have that $y(w)$ is invertible, we immediately obtain $\gamma(y) = \theta(w(y))$ and $\tau(y) = g(w(y))$.

It is almost immediate from this development that the set of solutions to (3) for a given g is Pareto ranked. We focus on the best of these for each given

[24] The method used above originates in the signaling model in Spence (1974), further developed by Riley (1979) and Mailath (1987). Equation (3) is best seen as defining an indirect mechanism where gross income is the signal sent by each agent to the planner, much as in Spence's model education is the signal sent to the firm. However, finding the equilibria of this game is only part of the problem. The remaining part of the problem relates to implementation. By this we mean that the social planner's problem is to define reward/penalty functions that induce each type of agent to choose, in equilibrium, the behavior the planner desires of that type of agent. A reference closer to our work is Guesnerie and Laffont (1984). However, there is a difference between our results and the other literature on implementation using the differentiable approach to the revelation principle. The difference is that in the other literature the principal cares only about implementing the action profiles of the agents (labor supply schedules in our model). In contrast, we consider the implementation of explicit maps from types to tax liability. That is, the principal cares about agents' types, which are hidden knowledge. These maps from types to tax liability are not action profiles, and are motivated by the ability to pay approach in classical public finance. They play the same role here as reduced form auctions play in the auction literature.

[25] That is, we have a separating equilibrium.

[26] In (3) the planner first chooses a net income function $\gamma(y)$, the agents then take the chosen net income function as given and maximize utility by selecting a gross income level y (or the corresponding level of labor supply). This is the implementation approach described in Laffont (1988). The Revelation Principle allows us to write an equivalent mechanism where agents are simply asked to report their type w. It is easier to check second order conditions of the problem for this direct mechanism. They essentially say that both pre and post tax incomes should be increasing functions of w. In our case they are strictly increasing functions and there is no bunching.

g. Define

$$\begin{aligned} T_k^* \equiv & \{\tau | \gamma \text{ is a solution to (3) for some } g \in G_k^*, \tau(y) = y - \gamma(y), \\ & \text{and } \tau \text{ Pareto dominates all other solutions to (3) for the given } g\}. \end{aligned}$$

Any element of T_k^* has the property that the marginal tax rate for the top ability $\overline{w}$ consumer type is zero.

From a practical viewpoint, for instance in solving examples such as those presented here, the use of these techniques and in particular equation (3) makes sense. However, for the general theory, in our application we do not have the conditions required by Berliant and Gouveia (2001); for example, the standard boundary condition is not satisfied due to the quasi-linear form of utility. Thus, we use Berliant and Page (1996), which is more general than Berliant and Gouveia (2001), and even then we must modify the proof slightly.

Lemma 4: If G_k^* is a set of measurable functions that are non-decreasing, then for any k and any $\tau \in T_k$ there is a $\tau^* \in T_k^*$ such that the utility level of each agent under τ^* is at least as large as the utility level of each agent under τ and such that the marginal tax rate for the top ability $\overline{w}$ consumer type under τ^*, if it exists, is zero.[27]

Proof: We verify the assumptions of Berliant and Page (1996), Theorems 1 and 2, with a modification. Obviously, u is continuous and strictly decreasing in tax payment τ, where $c = y - \tau$. As is standard in the optimal income tax literature, single crossing is satisfied because c is a normal good. The modification we must make is that, instead of the boundary conditions (3) and (4) in that paper, we have quasi-linear utility. The boundary conditions are used in only one place in the proofs, namely in the first paragraph of the proof of Theorem 1 of that paper. So we provide a substitute for the argument in that paragraph; quasi-linear utility actually makes the proof simpler. As there, $Y = [0, \mathfrak{m}) \subseteq \mathbb{R}_+$, the set of all possible incomes. We can set $\mathfrak{m} = \overline{w}$. First, given income y_i assigned to person i in the draw, where people are ordered by income from lowest to highest, omit x since it is irrelevant to this argument. The proof is by induction downward, beginning with the highest type k. After the inductive argument, Berliant and Page (1996, p. 399) let $k \to \infty$. Let $y_k \in \left\{ y \in Y \mid u(y - g(w_k), \frac{y}{w_k}, w_k) \geq u(y' - g(w_k), \frac{y'}{w_k}, w_k) \forall y' \in Y \right\}$. Define $\tau(y) = g(w_k)$ for $y \in [y_k, \infty)$. We must show that there is $\widehat{y}_{i-1}$ such

[27] The result on the top marginal tax rate is extended to non-differentiable functions in Berliant and Page (1996), but is a little complicated and, in fact, irrelevant to our purpose here.

that $u(\widehat{y}_{i-1} - g(w_{i-1}), \frac{\widehat{y}_{i-1}}{w_{i-1}}, w_{i-1}) \leq u(y_i - g(w_i), \frac{y_i}{w_{i-1}}, w_{i-1})$. Taking $\widehat{y}_{i-1} = y_i - g(w_i) + g(w_{i-1}) \leq y_i$ and evaluating the utility difference for our specific form of utility function, $u(\widehat{y}_{i-1} - g(w_{i-1}), \frac{\widehat{y}_{i-1}}{w_{i-1}}, w_{i-1}) - u(y_i - g(w_i), \frac{y_i}{w_{i-1}}, w_{i-1}) = y_i - g(w_i) + g(w_{i-1}) - g(w_{i-1}) + b(\frac{\widehat{y}_{i-1}}{w_{i-1}}, w_{i-1}) - y_i + g(w_i) - b(\frac{y_i}{w_{i-1}}, w_{i-1}) = b(\frac{\widehat{y}_{i-1}}{w_{i-1}}, w_{i-1}) - b(\frac{y_i}{w_{i-1}}, w_{i-1}) \leq 0$. As in the proof, since $g(w_{i-1}) \leq g(w_i)$, $u(y_i - g(w_i), \frac{y_i}{w_{i-1}}, w_{i-1}) \leq u(y_i - g(w_{i-1}), \frac{y_i}{w_{i-1}}, w_{i-1})$, and hence $u(\widehat{y}_{i-1} - g(w_{i-1}), \frac{\widehat{y}_{i-1}}{w_{i-1}}, w_{i-1}) \leq u(y_i - g(w_i), \frac{y_i}{w_{i-1}}, w_{i-1}) \leq u(y_i - g(w_{i-1}), \frac{y_i}{w_{i-1}}, w_{i-1})$. So there exists $y_{i-1} \in [\widehat{y}_{i-1}, y_i]$ with $u(y_{i-1} - g(w_{i-1}), \frac{y_{i-1}}{w_{i-1}}, w_{i-1}) = u(y_i - g(w_i), \frac{y_i}{w_{i-1}}, w_{i-1})$. Then for $y \in (y_{i-1}, y_i)$ define $\tau(y) \equiv y + b(\frac{y}{w_{i-1}}, w_{i-1}) - y_{i-1} + g(w_{i-1}) - b(\frac{y_{i-1}}{w_{i-1}}, w_{i-1})$. For $i = 1$, for $y \in [0, y_1)$, define $\tau(y) \equiv y + b(\frac{y}{w_1}, w_1) - y_1 + g(w_1) - b(\frac{y_1}{w_1}, w_1)$. The remainder of the proofs proceed as in that paper.

Remark: The theorem says that any non-negative and feasible revenue requirement function can be implemented by a continuum of tax schedules. These tax schedules are Pareto ranked and furthermore a maximal tax schedule under the Pareto ranking exists.

The next step is to characterize a class of individual revenue requirements for which we will be able to obtain results. This class contains the cases discussed in Theorems 1 and 2 and may possibly include other sets of assumptions.

Definition: A collection $\mathbf{E}$ of functions mapping $[\underline{w}, \overline{w}]$ into $\mathbb{R}$ is called *strongly single crossing* if each $g \in \mathbf{E}$ is:

1. Continuous.

2. Twice continuously differentiable except possibly at a finite number of points.

3. $dg/dw > 0$ except possibly at a finite number of points.

4. Individual revenue requirements cross each other only once, i.e. for any pair $g, g' \in \mathbf{E}$, there exists a $\widetilde{w}, \widetilde{w}' \in [\underline{w}, \overline{w}]$, $\widetilde{w} < \widetilde{w}'$ such that $g(\underline{w}) > g'(\underline{w})$ implies $g(w) > g'(w)$ for all $w \in [\underline{w}, \widetilde{w})$, $g(w) = g'(w)$ for all $w \in [\widetilde{w}, \widetilde{w}']$ and $g(w) < g'(w)$ for all $w \in (\widetilde{w}', \overline{w}]$.

Lemma 5 proves that when individual revenue requirements are strongly single crossing, the income tax systems in T_k^* cross at most once. It will be used to prove Lemma 6.

Lemma 5: Let k be a positive integer. Suppose that R_k implies strongly single crossing minimal individual revenue requirements, G_k^*. Let $\tau, \tau' \in T_k^*$, and let $y(\cdot), y'(\cdot)$ be the gross income functions associated with τ and τ', respectively. For incomes $y_1, y_2, y_3 \in y([\underline{w}, \overline{w}]) \cap y'([\underline{w}, \overline{w}])$, $y_1 < y_2 < y_3$, $\tau(y_3) < \tau'(y_3)$ and $\tau(y_2) > \tau'(y_2)$ implies $\tau(y_1) \geq \tau'(y_1)$.[28]

Proof: See the Appendix.

The notion of single crossing of tax systems is the analog of condition (SC) of Gans and Smart (1996) in this specific context.

Lemmas 4 and 5 are used to prove Lemma 6:

Lemma 6: Let k be a positive integer. Suppose that R_k implies strongly single crossing minimal individual revenue requirements, G_k^*. Then for any draw in $\mathcal{A}_k$, the one stage voting game has a majority rule equilibrium, namely $\forall (w_1, w_2, ..., w_k) \in \mathcal{A}_k$, $M_k(w_1, w_2, ..., w_k) \neq \emptyset$.

Proof: See the Appendix.

Strongly single crossing is used intensively to prove this. It has the implication that induced preferences over tax systems appear to have properties shared by single peaked preferences over a one dimensional domain. The winners will be the tax systems most preferred by the median voter (in the draw) out of tax systems in T_k^*.

The proof consists of two parts. The first part shows that there is a tax schedule that is weakly preferred to all others by the median voter. The second part shows that this tax schedule is a majority rule winner. This second part could be replaced by Gans and Smart (1996, Theorem 1). But it would take as much space to verify the assumptions of that Corollary as it does to prove our more specialized result directly.

The proof of Lemma 6 characterizes the set of majority rule equilibria for each draw. It will be interesting to investigate the comparative statics properties of the equilibria. This will require the imposition of further conditions on the utility function so that we can solve explicitly not only for the individual revenue requirements as above, but also for the implementation of the individual revenue requirements in terms of an income tax function using equation (3).

[28] Outside of $y([\underline{w}, \overline{w}])$, for instance, τ can be extended in an arbitrary fashion subject to incentive compatibility, for example in a linear way.

6 Conclusions

Two different but related issues deserve some discussion at the outset. The first is whether information on the likelihood of each draw can be used. The second is how to deal with possible excess revenues. As for the opposite situation of insufficient revenues, the reader should note that imposing a penalty for not meeting the requirement simply results in a new revenue requirement function.

We first discuss the information issue. One obvious possibility would be to define as feasible all individual revenue requirement functions that generate an *expected* revenue equal to or larger than the collective revenue requirement for the expected draw. The problem with this notion is that single crossing conditions would likely fail to be satisfied for most cases including the ones studied in this paper. But one could consider weakening our feasibility restriction and still have enough "bite" to generate single crossing g's. Here is a suggestion:

One option is to use a class of revenue weighting functions and constrain the expectation of weighted revenues. Expected revenue would be one particular member of this class. The class could be chosen to generate a continuum of constraints, binding enough for the single crossing result to survive, and we would be back to our initial setup although with weaker feasibility conditions. This is similar to a model of government behavior using ambiguity aversion or Knightian uncertainty. Perhaps this could be justified as a way to aggregate risk averse voter preferences over budget deficits.

We now address the issue of excess revenue. Consider first the case of utility quasi-linear in consumption good that we have used throughout this paper. It is possible to return the *ex post* excess revenue in a lump-sum fashion, as there are no income effects. It is not true that one might want to reduce the amount of the public good produced to prevent the welfare loss caused by excessive revenue: in this case the structure of preferences is such that any decrease in public good provision will result in recontracting afterwards, so as to get the unique Pareto optimal level of the public goods. Reducing revenue requirements cannot possibly lead to better resource allocations *ex post*.

When we consider general preferences and technologies the problem becomes more difficult. Clearly, the excess revenue cannot be returned to taxpayers in a lump sum fashion, as it will affect their behavior in optimizing against the income tax. In the more general case, there will be a trade-off between decreasing the public good level for some draws and, at the same time, decreasing revenue requirements. However, once we deviate from quasi-linear utility, other issues would arise before we get to this point, most importantly the presence of multiple Pareto optimal levels of public good provision. From the point of view of applications, analysis of these more general models will be much more difficult.

The real question is whether the alternative models have more to offer. Is it better to restrict ourselves to fixed revenue and voting over a parameter of a prespecified functional form for taxes (as in the previous literature), which are also generally Pareto dominated, or is the model proposed here a useful complement? Differences of opinion are clearly possible.

We note here that unlike much of the earlier literature on voting over *linear* taxes, the majority equilibria are not likely to be linear taxes without strong assumptions on utility functions and on the structure of incentives. The reason is simple: in the optimal income tax model, Pareto optimality requires that the top ability individuals face a marginal tax rate of zero.[29] All majority rule equilibria derived in Section 3 of this paper are second best Pareto optimal (for a given individual revenue requirement), and hence satisfy this property. Hence, poll taxes are the only linear taxes that could possibly be equilibria. In our model, such taxes are not generally majority rule equilibria, since consumers at the lower ability end of the spectrum will object.

In that sense, the results obtained here are a step forward relative to Romer (1975) and Roberts (1977). In another sense, they also improve on Snyder and Kramer (1988) by using a standard optimal income tax model as the framework to obtain the results.

There are a few strategies that may be productive in pursuing research on voting over taxes. One strategy is to use probabilistic voting models such as in Ledyard (1984). Another is to take advantage of the structure built in this paper and, with our results in hand, look at multi-stage games in which players' actions at the earlier stages might transmit information about types. Of course, it might be necessary to look at refinements of the Nash equilibrium concept to narrow down the set of equilibria to those that are reasonable (at least imposing subgame perfection as a criterion).

A two-stage game of interest is one in which k is fixed and each player in a draw proposes a tax system in T_k^* (simultaneously). The second stage of the

[29] We know of only one case where an optimal tax is linear: Snyder and Kramer (1988). But this and other results derived in that paper are due to the use of a peculiar model that departs significantly from the other models used in the study of income taxation. There are no income nor substitution effects on effort induced by taxation up to the point where workers switch to the underground sector, and from that point on the same holds since, by definition, income realized in the underground sector is not taxed.

game proceeds as in the single stage game above, with voting restricted to only those tax systems in T_k^* that were proposed in the first stage.

A three stage game of interest is one in which k is again fixed and the players in a draw elect representatives and who then propose tax systems and proceed as in the two stage game (see Baron and Ferejohn (1989)).

Work remains to be done in obtaining comparative statics results. As seen from the examples, that can be a complex task. Finally, the predictive power of the models will be the subject of empirical research. That will certainly be the focus of future work.

7 Appendix

7.1 Proof of Theorem 1

For a draw $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, the Lindahl-Samuelson condition for this model is:

$$\sum_{i=1}^{k} w_i \cdot s'(\widehat{r}(x)) \cdot \widehat{r}'(x) = m \cdot \widehat{r}'(x)$$

Hence,

$$x = \widehat{r}^{-1}\left[s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)\right]$$

and thus

$$R_k(w_1, w_2, ..., w_k) = m \cdot s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)$$

Hence, R_k is argument additive. Computing the first derivative,

$$\frac{dR_k}{d\sum_{i=1}^{k} w_i} = -m^2 \cdot \frac{1}{s''\left(s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)\right) \cdot \left(\sum_{i=1}^{k} w_i\right)^2} > 0$$

Computing the second derivative,

$$
\begin{aligned}
&\frac{d^2 R_k}{d\left(\sum_{i=1}^{k} w_i\right)^2} \\
&= m^2 \cdot \frac{2s''\left(s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)\right) \cdot \left(\sum_{i=1}^{k} w_i\right) + s'''\left(s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)\right) \cdot \frac{-m}{s''\left(s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)\right)}}{\left[s''\left(s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)\right) \cdot \left(\sum_{i=1}^{k} w_i\right)^2\right]^2} \\
&= m^2 \cdot \frac{2s''(r) \cdot \left(\sum_{i=1}^{k} w_i\right) + s'''(r) \cdot \frac{-m}{s''(r)}}{\left[s''(r) \cdot \left(\sum_{i=1}^{k} w_i\right)^2\right]^2}
\end{aligned}
$$

where

$$
r = s'^{-1}\left(\frac{m}{\sum_{i=1}^{k} w_i}\right)
$$

Thus,

$$
\begin{aligned}
\frac{d^2 R_k}{d\left(\sum_{i=1}^{k} w_i\right)^2} &< 0 \text{ if and only if} \\
2s''(r) \cdot \left(\sum_{i=1}^{k} w_i\right) &< s'''(r) \cdot \frac{m}{s''(r)} \text{ or} \\
2s''(r)^2 &> s'''(r) \cdot s'(r)
\end{aligned}
$$

The last expression holds by assumption. Therefore, R_k is argument additive with negative second derivative. The result then follows from Lemmas 2 and 6.

7.2 Proof of Theorem 2

For a draw $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, the Lindahl-Samuelson condition for this model is:

$$a \cdot \sum_{i=1}^{k} \frac{\widehat{r}'(x)}{d\widehat{r}(w_i)/dw_i} = m$$

So applying the inverse function theorem,

$$a \cdot \sum_{i=1}^{k} \frac{\widehat{r}^{-1\prime}(\widehat{r}(w_i))}{\widehat{r}^{-1\prime}(\widehat{r}(x))} = m$$

Hence,

$$x = \left[\widehat{r}^{-1\prime}\left(\widehat{r}'\right)\right]^{-1} \left(\frac{a}{m} \sum_{i=1}^{k} \widehat{r}^{-1\prime}(\widehat{r}(w_i)) \right)$$

Fortunately, we can use the inverse function theorem to simplify this mess:

$$\widehat{r}^{-1\prime}(\widehat{r}(w)) = \frac{1}{\widehat{r}'(w)}$$

Inverting this function,

$$\left[\widehat{r}^{-1\prime}\left(\widehat{r}'\right)\right]^{-1}(\cdot) = \widehat{r}'^{-1}\left(\frac{1}{\cdot}\right)$$

and thus

$$R_k(w_1, w_2, ..., w_k) = m \cdot x = m \cdot \widehat{r}'^{-1}\left[\left(\frac{a}{m} \sum_{i=1}^{k} \frac{1}{\widehat{r}'(w_i)} \right)^{-1} \right] \tag{4}$$

Next we show that $\widehat{r}^{-1\prime}(\widehat{r}(w)) = \frac{1}{\widehat{r}'(w)}$ is an increasing, convex function. Evaluating derivatives,

$$\frac{d\widehat{r}^{-1\prime}(\widehat{r}(w))}{dw} = -\frac{\widehat{r}''(w)}{\left(\widehat{r}'(w)\right)^2} > 0$$

$$\frac{d^2\widehat{r}^{-1\prime}(\widehat{r}(w))}{dw^2} = \frac{-\widehat{r}'(w)^2\widehat{r}'''(w) + \widehat{r}''(w)2\widehat{r}'(w)\widehat{r}''(w)}{\widehat{r}'(w)^4}$$

This expression is non-negative if and only if

$$2(\widehat{r}''(w))^2 \geq \widehat{r}'''(w) \cdot r'(w)$$

that holds by assumption. Under the same condition, $\left[\widehat{r}^{-1\prime}\left(\widehat{r}'\right)\right]^{-1}(\cdot) = \widehat{r}'^{-1}(\frac{1}{\cdot})$ is concave. Even without this assumption, it is increasing. Notice also that according to the inverse function theorem,

$$\begin{aligned} \widehat{r}^{-1\prime}(\widehat{r}(0)) &= \frac{1}{\widehat{r}'(0)} = 0 \\ \left[\widehat{r}^{-1\prime}\left(\widehat{r}'\right)\right]^{-1}(0) &= \widehat{r}'^{-1}(\infty) = 0 \end{aligned}$$

Next we verify limited complementarity. We shall use the following facts. Let $S : \mathbb{R}_+ \to \mathbb{R}_+$ with $S(0) = 0$ and $S' > 0$. Let $w_1, w_2, ..., w_k > 0$, $\eta \geq 1$. If $S'' \geq 0$, $\sum_{i=1}^k S(w_i) \leq S(\sum_{i=1}^k w_i)$ and $\eta \cdot S(w_1) \leq S(\eta \cdot w_1)$. If $S'' \leq 0$, $\sum_{i=1}^k S(w_i) \geq S(\sum_{i=1}^k w_i)$ and $\eta \cdot S(w_1) \geq S(\eta \cdot w_1)$.

$$
\begin{aligned}
& R_k(w_1, w_2, ..., w_k) + \sum_{i=1}^k g\left(W(w_i, w^*)\right) \\
= \; & m \cdot \widehat{r}'^{-1}\left[\left(\frac{a}{m} \cdot \sum_{i=1}^k \frac{1}{\widehat{r}'(w_i)}\right)^{-1}\right] + \sum_{i=1}^k g\left(W(w_i, w^*)\right) \\
\leq \; & m \cdot \widehat{r}'^{-1}\left[\left(\frac{a}{m} \cdot \sum_{i=1}^k \frac{1}{\widehat{r}'(w_i)}\right)^{-1}\right] + \sum_{i=1}^k W(w_i, w^*)
\end{aligned}
$$

Since $\frac{1}{\widehat{r}'(w)}$ is an increasing, convex function

$$
\leq m \cdot \widehat{r}'^{-1}\left[\left(\frac{a}{m} \cdot \frac{1}{\widehat{r}'\left(\sum_{i=1}^k w_i\right)}\right)^{-1}\right] + \sum_{i=1}^k W(w_i, w^*)
$$

Since $\widehat{r}'^{-1}(\frac{1}{\cdot})$ is concave and increasing with $\widehat{r}'^{-1}(\frac{1}{0}) = 0$ and $a \geq m$,

$$
\begin{aligned}
\leq \; & a \cdot \widehat{r}'^{-1}\left[\widehat{r}'\left(\sum_{i=1}^k w_i\right)\right] + \sum_{i=1}^k W(w_i, w^*) \\
= \; & a \cdot \sum_{i=1}^k w_i + \sum_{i=1}^k W(w_i, w^*)
\end{aligned}
$$

Because $a \geq 1$,

$$
\leq a \cdot \sum_{i=1}^k w_i + a \cdot \sum_{i=1}^k W(w_i, w^*)
$$

$$
= m \cdot \frac{2}{k} \cdot \sum_{i=1}^k \widehat{r}'^{-1}\left[\widehat{r}'\left(\frac{ak}{2m} \cdot w_i + \frac{ak}{2m} \cdot W(w_i, w^*)\right)\right]
$$

$$
= m \cdot \frac{2}{k} \cdot \sum_{i=1}^k \widehat{r}'^{-1}\left[\left(\frac{1}{\widehat{r}'\left(\frac{ak}{2m} \cdot w_i + \frac{ak}{2m} \cdot W(w_i, w^*)\right)}\right)^{-1}\right]
$$

Since $\frac{1}{\widehat{r}'(\cdot)}$ is convex and increasing,

$$
\leq m \cdot \frac{2}{k} \cdot \sum_{i=1}^k \widehat{r}'^{-1}\left[\left(\left[\frac{1}{2} \cdot \frac{1}{\widehat{r}'(\frac{ak}{m} \cdot w_i)} + \frac{1}{2} \cdot \frac{1}{\widehat{r}'(\frac{ak}{m} \cdot W(w_i, w^*))}\right]\right)^{-1}\right]
$$

Since $a \geq m$ implies $a \cdot k \geq m$,

$$\leq \quad m \cdot \frac{2}{k} \cdot \sum_{i=1}^{k} \widehat{r}'^{-1} \left[\left(\left[\frac{ak}{2m} \cdot \frac{1}{\widehat{r}'(w_i)} + \frac{ak}{2m} \cdot \frac{1}{\widehat{r}'(W(w_i, w^*))} \right] \right)^{-1} \right]$$

$$= \quad m \cdot \frac{2}{k} \cdot \sum_{i=1}^{k} \widehat{r}'^{-1} \left[\left(\frac{a}{m} \left[\frac{k}{2} \cdot \frac{1}{\widehat{r}'(w_i)} + \frac{k}{2} \cdot \frac{1}{\widehat{r}'(W(w_i, w^*))} \right] \right)^{-1} \right]$$

$$= \quad \sum_{i=1}^{k} R_k(w_1^i, w_2^i, ..., w_k^i) \cdot \frac{2}{k}$$

Finally, we verify Edgeworth substitutability. For $1 \leq j, j' \leq k$, $j \neq j'$, tedious calculations yield:

$$\frac{\partial^2 R_k(w_1, w_2, ..., w_k)}{\partial w_j \partial w_{j'}} =$$

$$\frac{m}{a} \cdot \widehat{r}'(w_j)^{-2} \cdot \widehat{r}''(w_j) \cdot \widehat{r}'(w_{j'})^{-2} \cdot \widehat{r}''(w_{j'}) \cdot \left(\sum_{i=1}^{k} \frac{1}{\widehat{r}'(w_i)} \right)^{-3} \cdot \frac{1}{\widehat{r}'' \left[\widehat{r}'^{-1} \left[\left(\frac{a}{m} \cdot \sum_{i=1}^{k} \frac{1}{\widehat{r}'(w_i)} \right)^{-1} \right] \right]^3} \cdot$$

$$\left(-\widehat{r}''' \left[\widehat{r}'^{-1} \left[\left(\frac{a}{m} \cdot \sum_{i=1}^{k} \frac{1}{\widehat{r}'(w_i)} \right)^{-1} \right] \right] \cdot \widehat{r}' \left[\widehat{r}'^{-1} \left[\left(\frac{a}{m} \cdot \sum_{i=1}^{k} \frac{1}{\widehat{r}'(w_i)} \right)^{-1} \right] \right] \right.$$
$$\left. + 2\widehat{r}'' \left[\widehat{r}'^{-1} \left[\left(\frac{a}{m} \cdot \sum_{i=1}^{k} \frac{1}{\widehat{r}'(w_i)} \right)^{-1} \right] \right]^2 \right)$$

Under the condition:

$$2(\widehat{r}''(w))^2 > \widehat{r}'''(w) \cdot \widehat{r}'(w)$$

this last expression is negative.

The result follows from Lemmas 3 and 6.

7.3 Proof of Lemma 2

It is straightforward to prove by direct calculation that $\forall g(w; \widetilde{w}) \in G_k^*$ as given in the statement of the Lemma, $g(w; \widetilde{w})$ is continuously differentiable in each of w and $\widetilde{w}$ and is strictly increasing in w. Since R_k is argument-additive, $R_k(w_1, w_2, ..., w_k) = Q(\sum_{i=1}^{k} w_i) = Q(k \cdot w^A)$, where w^A is the average ability in the draw.

Next focus on branches A and C of the statement of the Lemma. Since R_k is concave, on these branches,

$$g(w; \widetilde{w}) = Q(k \cdot \widetilde{w})/k + Q'(k \cdot \widetilde{w})(w^A - \widetilde{w}) \geq Q(\sum_{i=1}^{k} w_i)/k.$$

This shows that the branches A and C in the statement of the Lemma are feasible. We now prove that they are minimal. Consider branch A. Clearly, if a draw consists of k individuals of type $\widetilde{w}$, $g(\widetilde{w};\widetilde{w})$ is minimal. To show that $g(w;\widetilde{w})$ is minimal, suppose the opposite. Take $h(w)$ to be minimal, with $h(\widetilde{w}) = Q(k\cdot\widetilde{w})/k$ and $h(w) \leq g(w;\widetilde{w})$ with strict inequality for some $w_1 \in [\underline{w}, \widetilde{w} + (k-1)\cdot(\widetilde{w}-\underline{w})]$. It is feasible to have a draw $(w_1, w_2, ...w_k)$ with mean $\widetilde{w}$ and $w_i \in [\underline{w}, \widetilde{w} + (k-1)\cdot(\widetilde{w}-\underline{w})]$ for $i = 1, 2, ...k$. Then, $R_k(w_1, w_2, ...w_k) = Q(k\cdot\widetilde{w}) = \sum_{i=1}^{k} g(w_i;\widetilde{w})$. But $\sum_{i=1}^{k} h(w_i) < \sum_{i=1}^{k} g(w_i;\widetilde{w})$, so $h(w)$ is not feasible. Similar reasoning holds for branch C.

Now consider branch B and $w_1 \in (\widetilde{w} + (k-1)(\widetilde{w}-\underline{w}), \overline{w}]$. The logic used for branches A and C does not hold in this case: it is not possible to find $k-1$ ability levels in order to construct a draw with mean $\widetilde{w}$. Consider a draw with $w_j \in [\underline{w}, \widetilde{w} + (k-1)\cdot(\widetilde{w}-\underline{w})]$ for $j = 2, 3, ...k$. Due to argument-additivity, for any fixed draw mean w^A, we can take all w_j's ($j = 2, 3, ..., k$) to be equal to $\widehat{w} = (k\cdot w^A - w_1)/(k-1)$, without loss of generality. Feasibility requires

$$g(w_1;\widetilde{w}) + (k-1)\cdot g(\widehat{w};\widetilde{w}) \geq Q((k-1)\cdot\widehat{w} + w_1).$$

Take this as an equality and replace $g(\widehat{w};\widetilde{w})$ by

$$Q(k\cdot\widetilde{w})/k + Q'(k\cdot\widetilde{w})\cdot(\widehat{w}-\widetilde{w})$$

to obtain:

$$g(w_1;\widetilde{w}) = Q((k-1)\cdot\widehat{w}+w_1) - (k-1)/k\cdot Q(k\cdot\widetilde{w}) - (k-1)\cdot Q'(k\cdot\widetilde{w})\cdot(\widehat{w}-\widetilde{w}) \quad (5)$$

By construction, this revenue requirement is minimal (particularly at $\widehat{w} = \widetilde{w}$).

Next, notice that

$$\begin{aligned}(k-1)\cdot\widehat{w} + w_1 &> (k-1)\cdot\widehat{w} + \widetilde{w} + (k-1)\cdot(\widetilde{w}-\underline{w}) \\ &\geq (k-1)\cdot\underline{w} + \widetilde{w} + (k-1)\cdot(\widetilde{w}-\underline{w}) = k\cdot\widetilde{w}\end{aligned}$$

Since Q is concave,

$$(k-1)\cdot Q'((k-1)\cdot\widehat{w} + w_1) < (k-1)\cdot Q'(k\cdot\widetilde{w})$$

Hence, expression (5) is maximized over $\widehat{w} \in [\underline{w}, \overline{w}]$ at $\widehat{w} = \underline{w}$, so feasibility requires

$$g(w_1;\widetilde{w}) = Q((k-1)\cdot\underline{w} + w_1) - (k-1)/k\cdot Q(k\cdot\widetilde{w}) + (k-1)\cdot Q'(k\cdot\widetilde{w})\cdot(\widetilde{w}-\underline{w}).$$

It is easy to prove that allowing for draws with different compositions, namely more than one ability in the interval $[\widetilde{w}+(k-1)\cdot(\widetilde{w}-\underline{w}),\overline{w}]$, does not violate feasibility. We thus obtain branch B in the statement of the Lemma. Branch D is obtained following similar reasoning.

Next, suppose there is $h \in G_k^*$ that is not of the form given in the statement of the Lemma. Then there is some $w \in [\underline{w},\overline{w}]$ with $h(w) < g(w)$. Then $k \cdot h(w) < k \cdot g(w;w) = Q(k \cdot w)$, implying that h is not feasible.

To prove single cavedness in $\widetilde{w}$, one need only differentiate $g(w;\widetilde{w})$ with respect to the parameter $\widetilde{w}$. For branches A and C we obtain:

$$\frac{\partial g(w;\widetilde{w})}{\partial \widetilde{w}} = Q''(k \cdot \widetilde{w}) \cdot k \cdot (w - \widetilde{w}).$$

The derivative above is positive if $w < \widetilde{w}$ and negative for $w > \widetilde{w}$.

For branch B we have:

$$\frac{\partial g(w;\widetilde{w})}{\partial \widetilde{w}} = k \cdot (k-1) \cdot Q''(k \cdot \widetilde{w}) \cdot (\widetilde{w} - \underline{w}) < 0.$$

which applies only for $w > \widetilde{w}$.

Finally, for branch D we get:

$$\frac{\partial g(w;\widetilde{w})}{\partial \widetilde{w}} = k \cdot (k-1) \cdot Q''(k \cdot \widetilde{w}) \cdot (\widetilde{w} - \overline{w}) > 0.$$

which applies only for $w < \widetilde{w}$.

These results imply that $\arg\min_{\widetilde{w}} g(w;\widetilde{w}) = w$. Furthermore, we claim that these g's are single crossing. To see this, first note that from the definition of $g(w;\widetilde{w})$ in the statement of the Lemma, direct calculation yields that $\frac{\partial g(w;\widetilde{w})}{\partial w}$ is weakly decreasing in $\widetilde{w}$ for each w. Therefore, if $g(w;\widetilde{w})$ and $g(w;\widetilde{w}')$ cross twice, there exist $w, w', w'' \in [\underline{w},\overline{w}]$, $w < w' < w''$ such that $g(w;\widetilde{w}) = g(w;\widetilde{w}')$, $g(w';\widetilde{w}) \neq g(w';\widetilde{w}')$, $g(w'';\widetilde{w}) = g(w'';\widetilde{w}')$. But this cannot happen in each case: $\widetilde{w}' = \widetilde{w}$, $\widetilde{w}' < \widetilde{w}$, $\widetilde{w}' > \widetilde{w}$.

7.4 Proof of Lemma 3

We present the proof for k even. Adaptation of the proof for the case when k is odd is straightforward. Here we use the notation introduced in the definition of limited complementarity.

Fix $g \in G_k^*$. By assumption, $R_k(w_1, w_2, ..., w_k) \leq \sum_{i=1}^{k}[\frac{2}{k} \cdot R_k(w_1^i, w_2^i, ..., w_k^i) - g(W(w^i, w^*))]$. Since g is feasible, $g(w_i) \geq \frac{2}{k} \cdot R_k(w_1^i, w_2^i, ..., w_k^i) - g(W(w^i, w^*))$ for each i. If $g(w_i) > \frac{2}{k} \cdot R_k(w_1^i, w_2^i, ..., w_k^i) - g(W(w^i, w^*))$ for some i, then

g is not minimal in the sense that $g \notin G_k^*$, which is a contradiction. Hence $g(w_i) = \frac{2}{k} \cdot R_k(w_1^i, w_2^i, ..., w_k^i) - g(W(w^i, w^*))$ and in particular $g(w^*) = \frac{2}{k} \cdot R_k(w^*, w^*, ..., \underline{w}) - g(\underline{w}) = \frac{2}{k} \cdot R_k(w^*, w^*, ..., \overline{w}) - g(\overline{w})$.[30] Hence g is continuous. Moreover, since R_k is smoothly monotonic, $dg/dw > 0$.

Let $g, g' \in G_k^*$, with switching points w^* and w'^*. Suppose without loss of generality that $g(\underline{w}) > g'(\underline{w})$. Since g and g' belong to EG_k, we have that $g(\overline{w}) < g'(\overline{w})$. Since $g - g'$ is a continuous function defined over a connected, domain the intermediate value theorem says that it must have at least one zero. Take $\widetilde{w}$ as one such case. Assume that $\widetilde{w} \geq w^*$, $\widetilde{w} \geq w'^*$. Then $0 = g(\widetilde{w}) - g'(\widetilde{w}) = g'(\underline{w}) - g(\underline{w}) < 0$, a contradiction. Now assume that $\widetilde{w} \leq w^*$, $\widetilde{w} \leq w'^*$. Then $0 = g(\widetilde{w}) - g'(\widetilde{w}) = g'(\overline{w}) - g(\overline{w}) > 0$, another contradiction. Hence, either $w^* > \widetilde{w} > w'^*$ or the reverse must hold. Assume the former. Over the open interval (w'^*, w^*) we have:

$$\frac{d(g' - g)}{dw} = \frac{2}{k} \cdot [\frac{\partial R_k(w_i, ..., w_i, \underline{w}, ..., \underline{w})}{\partial w_i} - \frac{\partial R_k(w_i, ..., w_i, \overline{w}, ..., \overline{w})}{\partial w_i} > 0$$

by Edgeworth substitutability. Since the difference is increasing we have that there is a single zero, i.e. the revenue requirements g and g' cross only once.

Now assume $w'^* > \widetilde{w} > w^*$. Then, over (w^*, w'^*), $\frac{d(g'-g)}{dw}$ is negative (again by Edgeworth substitutability), contradicting continuity since we started with $g(\underline{w}) > g'(\underline{w})$ and $g(\overline{w}) < g'(\overline{w})$.

Notice that this proof of single crossing of the individual revenue requirements also proves that $g(\underline{w}) > g'(\underline{w}) \Longrightarrow w^* > w'^*$.

Finally, suppose we have g and $\widehat{g}$ in G_k^*, with switching points w^* and $\widehat{w}^*$ respectively. By the previously mentioned result, $g(\underline{w}) > \widehat{g}(\underline{w}) \Longrightarrow w^* > \widehat{w}^* \Longrightarrow g(\widehat{w}^*) - \widehat{g}(\widehat{w}^*) = \widehat{g}(\overline{w}) - g(\overline{w}) > 0$. Similarly, $g(\underline{w}) < \widehat{g}(\underline{w}) \Longrightarrow w^* < \widehat{w}^* \Longrightarrow g(\widehat{w}^*) - \widehat{g}(\widehat{w}^*) = \widehat{g}(\underline{w}) - g(\underline{w}) > 0$, proving the last statement in the Lemma.

7.5 Proof of Lemma 5[31]

Let g and g' be the elements of G_k^* associated with τ and τ', respectively. The proof is by contradiction. Suppose that there exist incomes $y_1 < y_2 < y_3$ with $\tau(y_1) < \tau'(y_1)$, $\tau(y_2) > \tau'(y_2)$ and $\tau(y_3) < \tau'(y_3)$. Then by the intermediate value theorem applied to utility differences as a function of w, there exists w^a

[30] Otherwise either g is not minimal or g is not feasible.

[31] To see how this critical proof works, it is useful to draw the graphs from optimal taxation, net income as a function of gross income, that are standard in the literature; see Seade (1977).

such that $u(y(w^a) - \tau(y(w^a)), y(w^a)/w^a) = u(y'(w^a) - \tau'(y'(w^a)), y'(w^a)/w^a)$, $y'(w^a) > y(w^a)$, $\tau'(y'(w^a)) < \tau(y'(w^a))$ and $\tau(y(w^a)) < \tau'(y(w^a))$. Moreover, $g(w^a) = \tau(y(w^a)) < \tau'(y(w^a))$ and since $y'(w^a) > y(w^a)$, $g'(w^a) > g(w^a)$.[32] There also exists $w^b > w^a$ with $u(y(w^b) - \tau(y(w^b)), y(w^b)/w^b) = u(y'(w^b) - \tau'(y'(w^b)), y'(w^b)/w^b)$, $y(w^b) > y'(w^b)$, $\tau(y'(w^b)) > \tau'(y'(w^b))$ and $\tau'(y(w^b)) > \tau(y(w^b))$. Hence $\tau(y'(w^b)) > \tau'(y'(w^b)) = g'(w^b)$ and since $y(w^b) > y'(w^b)$, $g(w^b) > g'(w^b)$.

Using strongly single crossing, $g(\overline{w}) > g'(\overline{w})$.

By construction of T_k^*, $\tau(y(\overline{w})) > \tau'(y'(\overline{w}))$. Note that since the marginal tax rate at $y(\overline{w})$ and $y'(\overline{w})$ is zero, what we have are essentially lump sum taxes at the top ability level. Hence, $u(y'(\overline{w}) - \tau'(y'(\overline{w})), y'(\overline{w})/\overline{w}) > u(y(\overline{w}) - \tau(y(\overline{w})), y(\overline{w})/\overline{w})$. Normality of leisure implies $y(\overline{w}) > y'(\overline{w})$. Moreover, continuity of τ implies $\tau(y'(\overline{w})) > \tau'(y'(\overline{w}))$ Since $\tau'(y(w^b)) > \tau(y(w^b))$, there exists $y'(\overline{w}) > y^* > y(w^b)$ with $\tau(y^*) = \tau'(y^*)$, so there exists w^c with $u(y(w^c) - \tau(y(w^c)), y(w^c)/w^c) = u(y'(w^c) - \tau'(y'(w^c)),\ y'(w^c)/w^c)$, $y'(w^c) > y(w^c)$, $\tau'(y'(w^c)) < \tau(y'(w^c))$ and $\tau(y(w^c)) < \tau'(y(w^c))$. As above, $g(w^c) < \tau'(y(w^c))$ and since $y'(w^c) > y(w^c)$, $g'(w^c) > g(w^c)$.

This contradicts strongly single crossing. So the hypothesis is false, and the lemma is established.

7.6 Proof of Lemma 6

Definition: Let C^1 be the space of continuously differentiable functions (with domain $[\underline{w}, \overline{w}]$ and range $\mathbb{R}$) endowed with the uniform topology. We consider T_k^* to be a subset of this space by extending any $\tau \in T_k^*$ to the whole domain, if necessary, in a C^1 and linear fashion.

Fix $\tau \in T_k^*$. First we claim that $0 \leq d\tau/dy \leq 1$. The first inequality holds because $d\tau/dy = dg/dw \cdot dw/dy$, and $dg/dw > 0$ (except possibly at a finite number of points) by assumption whereas $dw/dy \geq 0$ is demonstrated in the course of proving the implementation result, Proposition 1, in Berliant and Gouveia (2001), so it holds except possibly at a finite number of points. Since, in spite of the exceptions at finitely many points, τ will be C^1, $0 \leq d\tau/dy$. The second inequality can be written $d\gamma/dy \geq 0$, which reduces to $d\gamma/dw \cdot dw/dy \geq 0$. As before, $dw/dy \geq 0$, and $d\gamma/dw \geq 0$ is demonstrated in the same place as $dw/dy \geq 0$. (Note that $d\gamma/dw > 0$ is the second order condition for incentive

[32] $d\tau/dy = dg/dw \cdot dw/dy > 0$ holds because $dg/dw > 0$ and $dw/dy > 0$ is proved in Proposition 1 of Berliant and Gouveia (2001).

compatibility in this model.) So every $\tau \in T_k^*$ is Lipschitz in income with constant 1, and thus T_k^* is equicontinuous. Since $k \cdot g(w) \geq R_k(w, w, ..., w) \geq 0$, T_k^* is also norm bounded by $\overline{w}$. Using Ascoli's theorem (see Munkres (1975, p. 290)), $\overline{T}_k^*$ (the closure of T_k^* in C^1) is compact.

Fix k and let $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$. For any $\tau \in T_k$, let $v(\tau, w) = \max_y u(y - \tau(y), y/w)$, the utility induced by the tax system τ for type w. It is easy to verify that for each w, $v(\tau, w)$ is continuous in its first argument.

Let τ^* be a maximal element of $\overline{T}_k^*$ using $v(\cdot, w^M)$ as the objective, where w^M is the median ability level in $(w_1, w_2, ..., w_k)$ if k is odd, and $w^M \in [w_{k/2}, w_{k/2+1}]$ (where the wage rates are ordered in an increasing fashion) if k is even. Using Lemma 4, $\tau^* \in T_k^*$.

Now suppose there exists $\tau \in T_k$ such that there is a subset D of $\{w_1, w_2, ..., w_k\}$ with $v(\tau, w) > v(\tau^*, w)$ for all $w \in D$ and where the cardinality of D is greater than $k/2$. Then using Lemma 4, we can take τ to be in T_k^* without loss of generality. Using Lemma 5, $\tau^*$and τ are single crossing, or alternatively, their after tax income functions are single crossing. Thus, there exist intervals $W, W' \subseteq [\underline{w}, \overline{w}]$ such that W and W' partition $[\underline{w}, \overline{w}]$ and $D \subseteq W$. Let W be the smallest interval (in the sense of set inclusion) such that W and its complement are both intervals, W and W' partition $[\underline{w}, \overline{w}]$, and $D \subseteq W$.

Then by definition of τ^*, $w^M \notin W$. Hence D cannot contain a majority of the draw, a contradiction. Hence the hypothesis is false and τ^* cannot be defeated by any other feasible tax system.

References

[1] Aumann, R.J. and M. Kurz, 1977, Power and taxes, Econometrica 45, 1137-1161.

[2] Baron, D. and J. Ferejohn, 1989, Bargaining in legislatures, American Political Science Review 83, 1181-1206.

[3] Bergstrom,T. and R. Cornes, 1983, Independence of allocative efficiency from distribution in the theory of public goods, Econometrica 51, 1753-1765.

[4] Berliant, M., 1992, On income taxation and the core, Journal of Economic Theory 56, 121-141.

[5] Berliant, M. and M. Gouveia, 2001, Incentive compatible income taxation, individual revenue requirements and welfare. https://mpra.ub.uni-muenchen.de/84374/1/MPRA_paper_84374.pdf

[6] Berliant, M. and F.H. Page, Jr., 1996, Incentives and income taxation: The implementation of individual revenue requirement functions, Ricerche Economiche 50, 389-400.

[7] Berliant, M. and F.H. Page, Jr., 2001, Income taxes and the provision of public goods: Existence of an optimum, Econometrica 69, 771-784.

[8] Berliant, M. and F.H. Page, Jr., 2006, Budget balancedness and optimal income taxation, Journal of Public Economic Theory 8, 409-451.

[9] Berliant, M. and P. Rothstein, 2000, On models with an uncongestible public good and a continuum of consumers, Journal of Urban Economics 48, 388-396.

[10] Bierbrauer, F.J., 2011, On the optimality of optimal income taxation, Journal of Economic Theory 146, 2105-2116.

[11] Brett, C. and J.A. Weymark, 2017, Voting over selfishly optimal nonlinear income tax schedules, Games and Economic Behavior 101, 172-188.

[12] Calvert, R., 1986, Models of Imperfect Information in Politics (Harwood Academic Publishers, Chur).

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

[14] Foley, D., 1967, Resource Allocation and the Public Sector, Yale Economic Essays 7, 45-98.

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

[16] Gouveia, M., 1997. Majority rule and the public provision of a private good, Public Choice 93, 221-244.

[17] Guesnerie, R. and J.-J. Laffont, 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.

[18] Guesnerie, R. and C. Oddou, 1981, Second best taxation as a game, Journal of Economic Theory 25, 67-91.

[19] Heathcote, J. and H. Tsujiyama, 2017, Optimal income taxation: Mirrlees meets Ramsey. Working paper.

[20] Hettich, W. and S.L. Winer, 1988, Economic and political foundations of tax structure, American Economic Review 78, 701-713.

[21] Hirsch, M., 1976, Differential Topology (Spinger-Verlag, New York, NY).

[22] Kaneko, M., 1981, On the existence of an optimal income tax schedule, Review of Economic Studies 48, 633-642.

[23] Laffont, J.-J., 1988, Fundamentals of Public Economics (MIT Press, Cambridge, MA).

[24] Ledyard, J., 1984, The pure theory of large two-candidate elections, Public Choice 44, 1-43.

[25] Lindahl, E., 1919, Positive lösung, die gerechtigkeit der besteuerung, translated as Just taxation — a positive solution, in: R.A. Musgrave and A.T. Peacock, eds., Classics in the Theory of Public Finance (Macmillan, London).

[26] L'Ollivier, S. and J.-C. Rochet, 1983, Bunching and second order conditions: A note on optimal tax theory, Journal of Economic Theory 31, 392-400.

[27] Mailath, G., 1987, Incentive compatibility in signaling games with a continuum of types, Econometrica 55, 1349-1365.

[28] Meltzer, A. and S. Richard, 1981, A rational theory of the size of government, Journal of Political Economy 89, 914-27.

[29] Meltzer,A. and S. Richard, 1983, Tests of a rational theory of the size of government, Public Choice 41, 403-418.

[30] Mill, J.S., 1848, Principles of Political Economy (Longmans Green, London).

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

[32] Munkres, J.R., 1975, Topology: A First Course (Prentice-Hall, Englewood Cliffs, NJ).

[33] Musgrave, R.A., 1959, The Theory of Public Finance (McGraw Hill, New York).

[34] Nakayama, M., 1976, A note on the core of an income redistribution game, Journal of Economic Theory 12, 347-349.

[35] Peck, R.M., 1986, Power and linear income taxes: An example, Econometrica 54, 87-94.

[36] Plott, C.R., 1967, A notion of equilibrium and its possibility under majority rule, American Economic Review 67, 787-806.

[37] Riley, J., 1979, Informational equilibrium, Econometrica 47, 331-359.

[38] Roberts, K.W.S., 1977, Voting over income tax schedules, Journal of Public Economics 8, 329-340.

[39] Röell, A., 1996, Voting over nonlinear income tax schedules, mimeo.

[40] Roemer, J.E., 1999, The democratic political economy of progressive income taxation, Econometrica 67, 1-19.

[41] Romer, T., 1975, Individual welfare, majority voting, and the properties of a linear income tax, Journal of Public Economics 4, 163-186.

[42] Rudin, W., 1974, Real and Complex Analysis (McGraw-Hill, New York, NY).

[43] Samuelson, P.A., 1954, The pure theory of public expenditure, Review of Economics and Statistics 36, 387-389.

[44] Samuelson, P.A., 1955, Diagrammatic exposition of a theory of public expenditure, Review of Economics and Statistics 37, 350-356.

[45] Schofield, N., 1978, Instability of simple dynamic games, Review of Economic Studies 45, 575-594.

[46] Seade, J.K., 1977, On the shape of optimal tax schedules, Journal of Public Economics 7, 203-235.

[47] Snyder, J.M. and G.H. Kramer, 1988, Fairness, self-interest, and the politics of the progressive income tax, Journal of Public Economics 36, 197-230.

[48] Spence, M., 1974, Market Signaling (Harvard, Cambridge, MA).