On the Political Economy of Income Taxation

Berliant, Marcus and Gouveia, Miguel

Washington University in St. Louis, Universidade Católica Portuguesa

7 February 2018

On the Political Economy of Income Taxation*

Marcus Berliant[†] and Miguel Gouveia[‡]

February 7, 2018

Abstract

The literatures dealing with voting, optimal income taxation, and implementation are integrated here to address the problem of voting over income taxes. 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 tax requirements such that a majority rule equilibrium exists are found. Finally, conditions 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 Randy Calvert, Yan Chen, Roy Gardner, John Ledyard, Hideo Konishi, Wan-hsiang Pan, Michael Smart, Jim Snyder 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. It also seems reasonable to fix the revenue requirement, given that in many countries (such as the U.S.), the institutional national political structure separates decisions about taxes from decisions about expenditures. However, this will not be a requirement of our analysis below.

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 in 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. Suppose that a *finite* sample is drawn from this 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.

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).[3]

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

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

[3] Our arguments apply to finite numbers of agents. Jim Snyder has pointed out to us that 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. Instead we have only that the revenue constraint needs to be satisfied for the known population.

and set of arguments that can be exploited in voting games over income taxes to obtain existence and sometimes uniqueness and characterization results.

We then extend the analysis to an environment with incentives, along the lines of Mirrlees (1971). At this point the assumptions on utilities used in the optimal income tax literature alone are sufficient to obtain an analog of single peaked preferences over the feasible set of alternative tax functions.

The focus of this paper is on voting over income taxes without information transmission or opportunities for strategic behavior. We hope to address these issues in subsequent work.

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 (1999) to be sure that the second order conditions for incentive compatibility hold in our model.

The structure of the paper is as follows. First, we introduce our framework and derive some useful preliminary results in section 2. In section 3 we examine an optimal income tax economy based on the setting pioneered by Mirrlees (1971), where labor supply is endogenous. Of primary interest are the existence and characterization of a majority voting equilibrium. Section 4 studies voting over both public goods and taxes. Finally, section 5 contains conclusions and suggestions for further research.

2 The Model

2.1 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 $u(c) = c$, with $c \in \mathbb{R}_+$. A consumer's endowment, which is also her type, is described by $w \in W \equiv [\underline{w}, \overline{w}]$, where $[\underline{w}, \overline{w}] \subseteq \mathbb{R}_{++}$ is an interval contained in the positive real line. 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 $f(w) > 0$ $a.s.$[4].

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 $R_k : \mathcal{A}_k \to \mathbb{R}$. 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.

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.

Assumptions on R_k will be imposed and discussed below. One basic assumption that we will maintain throughout is that R_k is *attainable*. Formally, R_k is *attainable* if for every $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, $\sum_{i=1}^{k} w_i > R_k(w_1, w_2, ..., w_k)$. Strict inequality is used here in preparation for the models used in Sections 3 and 4.

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

[4]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.

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 major point about asymmetric information in this model is as follows. The government and the agents in the economy know the prior distribution f of types of agents in the economy[5] as well as the mapping R_k. In an endowment economy we assume that the timing of events is such that tax functions must be chosen before the composition of the draw is known to the voters.

Next we impose a topology on $\mathcal{A}_k$. For $\mathcal{A}_k$, we simply use the Euclidean norm $\|\cdot\|_k$ on the subspace $[\underline{w}, \overline{w}]^k$.

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 $\mathcal{A}_k$ or in $\mathcal{A}$. 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 (or of an unknown size) 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*[6] 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 k and R_k. Let

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

G_k is the set of all *individual* revenue requirements that collect enough revenue to satisfy R_k. $G_k \neq \emptyset$ if R_k is attainable. The restriction that the revenue requirement be satisfied for each draw restricts the feasible set significantly.

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

[6] 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.

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.

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 Assumptions and Preliminaries

2.3.1 Basic Assumptions

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

Next we state a natural assumption on R_k, which 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)).[7] We will also assume that R_k is smoothly monotonic:

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

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 draws with higher ability to pay owe more taxes. One could successfully use weaker assumptions with this framework, but at a cost of greatly complicating the proofs.[8]

Finally, we shall impose the condition that $\underline{w} > R_k(\overline{w}, \overline{w}, ..., \overline{w})/k$. Under smooth monotonicity, this assumption is sufficient for the conclusion $w > g(w) \forall w \in [\underline{w}, \overline{w}]$, so that after tax incomes are non-negative. This is important for implementation in the optimal income tax setting. Weaker assumptions could be used, again rendering proofs more complicated. Evidently, this assumption is stronger than attainability.

2.3.2 Single Crossing Individual Revenue Requirements

The next step in model development 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,[9] i.e. each pair of g's will cross only once.[10] These two sets of assumptions reflect how collective revenue requirements change with polarization of the draw, and are illustrated in Figures 1 and 2. Loosely speaking, they are opposites of one another. The first set of assumptions has collective revenue requirements increasing as a draw becomes more polarized, whereas the second set of assumptions has collective revenue requirements decreasing as a draw becomes more polarized.

We first present 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

[8] 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.

[9] 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.

[10] 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.

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.

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.[11]

Theorem 1: 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 $dg/dw > 0$, 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.

One class of functions for which limited complementarity and Edgeworth substitutability hold is given by the CES family:

$$R_k(w_1, ...w_k) = \alpha \cdot k^{1-1/p} \cdot \left(\sum_{i=1}^{k} w_i^p\right)^{1/p} \text{ where } p \geq 1 \text{ and } \alpha < \left(\frac{\overline{w} + \underline{w}}{2\overline{w}}\right)^{(p-1)/p}.$$

[11] This is the analog of condition (SC) of Gans and Smart (1996).

A particularly striking example of a case satisfying limited complementarity is given by the following collective revenue requirement function[12] and illustrated in Figure 1.

$$R_k(w_1, ...w_k) = \alpha \cdot \sum_{i=1}^{k} |w_i - w^M| + \beta \cdot \sum_{i=1}^{k} w_i,$$

where w^M is the median of the draw, $0 < \alpha < \beta < 1$, and $\alpha/(1-\beta) < \underline{w}/(\overline{w} - \underline{w})$.[13]

The Appendix contains a proof that limited complementarity holds for this case. The individual revenue requirement functions take the form:

$$\begin{aligned} g(w) &= \beta \cdot w + \alpha \cdot (\widetilde{w} - w) \text{ if } w < \widetilde{w} \text{ and} \\ g(w) &= \beta \cdot w + \alpha \cdot (w - \widetilde{w}) \text{ if } w \geq \widetilde{w} \end{aligned}$$

Figure 1 about here

In this case, the two branches of the function are linear and the marginal requirement is thus always higher for the taxpayers with higher ability to pay.

We now look at a second case that implies single crossing of revenue requirement functions in G_k^*.

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 $\frac{dQ}{d\sum w_i}$.

Theorem 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 $w \leq \widetilde{w} + (k-1) \cdot (\widetilde{w} - \underline{w})$.

[12] Although this example is not differentiable, it is simple and it has the basic properties leading to our result. It is monotonic in w_i and verifies a weaker version of Edgeworth substitutability: as other w's in the sample increase, the incremental requirement of a given w_i either stays the same or decreases (the latter when it goes from above to below the median).

[13] These constraints on parameters ensure that the individual revenue requirement function is strictly increasing and that it doesn't exceed the budget of a consumer.

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[14] 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.

Remark: Notice that the notion of single crossing used in Theorem 2 is weaker than the notion used in Theorem 1. Thus, we will use the notion of single crossing from Theorem 2 below. 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, as can be seen in Figure 2. 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. Figure 2 shows two possible individual revenue requirements corresponding to $\widetilde{w} = w_1$ and $\widetilde{w} = w_2$. 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. We can also relate this result to the intuition provided earlier for the limited complementarity case. In the argument-additivity case, concavity implies that per-capita revenue requirements decrease with the polarization of the sample.

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

Figure 2 about here

Such a result supports the idea 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 does not have in itself sufficient information to predict the shape of the income tax schedules chosen by majority rule. Revenue requirements R_k are also important.[15]

3 Voting Over Taxes in an Optimal Income Tax Economy

Having dispensed with preliminaries, we now turn to the voting model with incentives based on Mirrlees (1971). The objective here is to show that we can design a tax function on (endogenous) income implementing revenue requirements defined on the underlying (and unobservable) endowments. Given this, we can translate the results obtained for an endowment economy to the environment with incentives.

The two goods in the model are a composite consumption good, whose quantity is denoted by c, and labor, whose quantity is denoted by l. Consumers have an endowment of 1 unit of *labor/leisure* and perhaps an endowment of consumption good. $u(c, l)$ is the utility function of all consumers, 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_i is the value of agent i's endowment of labor. The gross income earned by agent i is $y_i = w_i \cdot \underline{l}_i$ and it equals consumption when there are no taxes.

The following assumptions are maintained throughout the remainder of this paper:

A1: Standard assumptions on preferences:
$u_1 > 0$, $u_2 < 0$, $u_{22} < 0$, $u_{11} < 0$.

A2: The utility function is strictly quasi-concave:
$u_{11} \cdot u_2^2 - 2u_{12} \cdot u_1 \cdot u_2 + u_{22} \cdot u_1^2 < 0$.

[15] 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.

A3: Consumption is normal:
$u_{21} \cdot u_2 - u_{22} \cdot u_1 > 0.$

A4: Leisure is normal:
$u_{11} \cdot u_2 - u_{12} \cdot u_1 > 0.$

A5: Boundary conditions:
$\lim_{c \to 0} u_1(c, l) = \infty$, $\lim_{l \to 1} u_2(c, l) = -\infty$, $\lim_{l \to 0} u_2(c, l) = 0$.

Assumptions A1, A3 and A4 imply A2 but they are listed separately for convenience. Assumptions A1, A2, and A5 are standard. Assumption A3 is generally used in the optimal income tax literature to obtain the single crossing property for indifference curves. Assumption A4 is also common in the optimal income tax literature and is used to derive comparative statics there. The last part of A5 helps to ensure the existence of interior solutions.[16] Although these are strong assumptions, they seem necessary to obtain a tractable model. As mentioned in the introduction, they are weaker than assumptions used in the earlier literature in this area, such as quasi-linear preferences. Let $y \in \mathbb{R}$ be individual gross income. 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:

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

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

Next the production correspondence is defined formally. Let $I_k = [0, 1]^k$. For each $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, production possibilities are described by a set $Y_k(w_1, w_2, ..., w_k)$, where $Y_k(w_1, w_2, ..., w_k) \subseteq I_k \times \mathbb{R}$. For a given $(w_1, w_2, ..., w_k)$, $(\underline{l}, C) \in Y_k(w_1, w_2, ..., w_k)$ describes labor input $\underline{l}_i \equiv l(w_i)$ of person i, along with net output C in consumption good of the economy.

Notice that labor inputs are measured as positive numbers, and that labor inputs of those not in the draw $(w_1, w_2, ..., w_k)$ are zero. We assume throughout

[16] See Seade (1977,1982). A5 is used in our paper in the proof of Theorem 3.

that the endowment of consumption good of a draw as well as tax revenue due from a draw are independent of labor supply $\underline{l}$ and composite consumption good output C. As is almost universal in the optimal income tax literature, a constant returns to scale technology is postulated. Formally,

$$Y_k(w_1, w_2, ..., w_k) \equiv \left\{ (\underline{l}, C) \in I_k \times \mathbb{R} | \sum_{i=1}^{k} w_i \cdot \underline{l}_i \geq C + R_k(w_1, w_2, ..., w_k) \right\}.$$

This structure captures some important aspects of the optimal income tax model. First, labor is modeled as a differentiated product, so workers with different characteristics can have different wages (or productivities). Second, the production set can embody initial endowments of both labor and consumption good on the part of consumers, as well as revenue collections required by the government from any draw.

Here we assume that the revenue requirement function $R_k(w_1, w_2, ..., w_k) \geq 0$. This assumption is not strictly necessary but avoids the need to deal with a few technical problems. More discussion of this assumption can be found in Berliant and Gouveia (2001). One implication of this restriction is that the political process modeled here is best thought of as a cost sharing process rather than one of explicit redistribution. Foley (1967) used a simple endowment economy to prove that there is no majority rule equilibrium when voting over explicitly redistributional taxes. This nonexistence can be overcome if we place strong restrictions in the policy space, which typically means linear tax functions. This is precisely the kind of restriction we are trying to avoid. On the other hand, we should keep in mind that most government expenditures are exhaustive, that is they represent purchases of goods and services rather than transfers. Our model is geared to explaining the distribution of the tax burden financing these exhaustive expenditures.

For a given tax system, the labor income of each agent is observed by all, but the wage rate and hours worked of each agent are known only to the agent himself. This is an explicit statement of the information structure of the model.

Definition: Fix k and a revenue requirement function R_k. The set of *feasible* tax systems is defined to be

$$\begin{aligned} T_k \equiv \Bigg\{ & \tau : \mathbb{R} \to \mathbb{R} | \tau \text{ is measurable, and for all } (w_1, w_2, ..., w_k) \in \mathcal{A}_k, \\ & \sum_{i=1}^{k} \tau(y(w_i)) \geq R_k(w_1, w_2, ..., w_k) \Bigg\} \end{aligned}$$

Finally, the notion of a majority rule equilibrium can be defined. Fix a revenue requirements function R_k. A *majority rule equilibrium* for draws of size k is a correspondence $M_k : \mathcal{A}_k \to T_k$ such that for every $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, for every $\tau \in M_k((w_1, w_2, ..., w_k))$ (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 $\tau' \in T_k$ (with associated $y'(w)$) such that

$$u(y'(w) - \tau'(y'(w)), y'(w)/w) > u(y(w) - \tau(y(w)), y(w)/w) \text{ for all } w \in D. \quad (1)$$

We will show that we can restrict attention to continuous τ.

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 $\max_l u(\gamma(w \cdot l), l)$. The first order condition from this problem is $u_1 \cdot \frac{d\gamma}{dy} \cdot w + u_2 = 0$, where subscripts represent partial derivatives of u with respect to the appropriate arguments. Rearranging,

$$\frac{d\gamma}{dy} = -\frac{u_2(\gamma(w \cdot l), l)}{u_1(\gamma(w \cdot l), l)} \cdot \frac{1}{w}.$$

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,

$$\frac{d\gamma}{dy} = -\frac{u_2(\gamma, \frac{y}{g^{-1}(y-\gamma)})}{u_1(\gamma, \frac{y}{g^{-1}(y-\gamma)})} \cdot \frac{1}{g^{-1}(y-\gamma)} \equiv \Phi(\gamma, y). \quad (2)$$

As in Berliant (1992), a standard result from the theory of differential equations yields a family of solutions to this differential equation.[17] Berliant and Gouveia (2001) show that (2) has global solutions if $g \geq 0$.

[17] The method used above originates in the signaling model in Spence (1974), further developed by Riley (1979) and Mailath (1987). Equation (2) 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).

Of course, as L'Ollivier and Rochet (1983) point out, the second order conditions must be checked to ensure that solutions to (2) do not involve bunching, which means that consumers do optimize in (2) at the tax liability given by g.[18] This was done in Berliant and Gouveia (2001), and more generally in Berliant and Page (1999), where the Revelation Principle[19] 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 (2) for a given g is Pareto ranked. We focus on the best of these for each given g. Define

$$\begin{aligned} T_k^* \equiv & \{\tau | \gamma \text{ is a solution to (2) for some } g \in G_k^*, \, \tau(y) = y - \gamma(y), \\ & \text{and } \tau \text{ Pareto dominates all other solutions to (2) for the given } g\} . \end{aligned}$$

Theorem 3: If G_k^* is a set of continuous, increasing functions that are twice continuously differentiable except at a finite number of points, 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^* is zero.

Proof: See the Appendix.

Remark: The theorem says that any non-negative and feasible revenue

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.

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

[19] In (2) 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.

requirement function can be implemented by a continuum of tax schedules. These tax schedules are Pareto ranked and furthermore a maximal tax schedule exists.

Remark: Note that when $g(w)$ is C^2 so is τ^* (see Berliant and Gouveia (2001)). Furthermore, when we have non-differentiability of g at w^*, τ^* is C^1: simple computations show that both the right-hand and left-hand derivatives of $\tau(y)$ at $y = y(w^*)$ are equal to $1 + u_2(y - \tau, y/w^*)/(u_1(y - \tau, y/w^*) \cdot w^*)$.

We begin by characterizing a class of individual revenue requirements for which we will be able to obtain results. This class corresponds to the cases discussed in Theorems 1 and 2 and may possibly include other sets of assumptions on collective revenue requirements.

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 1 proves that when individual revenue requirements are strongly single crossing, the income tax systems in T_k^* cross at most once.

Lemma 1: 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 $\tau(y_1) \geq \tau'(y_1)$.[20]

Proof: See the Appendix.

With this result in hand we now state the central theorem of this paper.

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

Theorem 4: 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.[21]

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^*.

Notice that the proof of Theorem 4 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. Also, notice that Theorem 4 is immediately applicable to the cases delineated in Theorems 1 and 2:

Theorem 1(IC): Let k be a positive integer. Suppose that R_k satisfies limited complementarity and Edgeworth substitutability. Then for any draw in $\mathcal{A}_k$, the one stage voting game has a majority rule equilibrium.

Theorem 2(IC): Let $k \geq 2$ and assume that $R_k(w_1, w_2, ..., w_k)$ is argument-additive with $Q'' < 0$. Then for any draw in $\mathcal{A}_k$, the one stage voting game has a majority rule equilibrium.

4 Simultaneous Voting Over a Public Good and Taxes

The public goods financed by the revenue raised through the income tax are usually excluded from models of optimal income taxation due to the complexity introduced. In the model considered here, voting over a public good is already captured to some degree on the revenue side, since R_k varies with the draw of

[21] 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, Corollary 2). But it would take as much space to verify the assumptions of that Corollary as it does to prove our more specialized result directly.

types and hence with the production of public goods for these types. Where it is not captured is in the utility functions of agents, where public goods should appear explicitly. Suppose that a public good, with some given cost function, is included in the model and incorporated in utility functions. Let $x \in \mathbb{R}_+$ be the quantity of the public good.[22] 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:

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

In this case, a straightforward extension of our definition of majority rule equilibrium is the following: a *majority rule equilibrium* for draws of size k is a correspondence M_k mapping $(w_1, w_2, ..., w_k)$ into $F_k((w_1, w_2, ..., w_k))$ such that for 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 $w \in D$.

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.[23] Let utility be given by a C^2 function $u(c, l, x, w) = a \cdot c + b(l, w) + r(x, w)$.[24] We assume that $\partial b/\partial l < 0$, $\partial^2 b/\partial l^2 < 0$, $\partial r/\partial x > 0$, $\partial^2 r/\partial x^2 < 0$. It is assumed that $dH(x)/dx > 0$ and $d^2 H(x)/dx^2 \geq 0$. Let $(w_1, w_2, ..., w_k) \in \mathcal{A}_k$, and let c_i and $\underline{l}_i$ denote the consumption and labor supply of the ith member of the draw respectively. Then production possibilities are given by:

$$\sum_{i=1}^{k} w_i \cdot \underline{l}_i - \sum_{i=1}^{k} c_i \geq H(x). \tag{3}$$

We define interior allocations to be vectors $(< c_i >_{i=1}^{k}, < l_i >_{i=1}^{k}, x) >> 0$. A pair (τ_j, x_j) is *interior* if the resulting allocation[25] is interior.

[22] We write $u(c, l, x, w)$ for the utility function.

[23] Revenue requirement correspondences are too difficult to handle at this stage of model development.

[24] 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.

[25] The allocation results when the agents in a draw each solve their consumer problem. See Bergstrom and Cornes (1983) for an explanation of why we need to restrict the analysis to interior allocations.

Lemma 2: Under the assumptions listed above, for any given draw $(w_1, w_2, ...w_k)$, for all interior $(\tau_j, x_j), (\tau_h, x_h) \in M_k$, $x_j = x_h$.

Proof: The Pareto optimal allocations are solutions to: $\max u(c, l, x, w_1)$ subject to $u(c, l, x, w_i) \geq \overline{u}_i$ for $i = 2, 3, ...k$ and subject to (3) 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:

$$\sum_{i=1}^{k} \frac{1}{a} \cdot \partial r(x, w_i)/\partial x = dH(x)/dx. \tag{4}$$

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 consumption for the given draw. Given our assumptions on r and H, there is a unique level of public good that solves (4).

Q.E.D.

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 $R_k(w_1, w_2, ..., w_k) \equiv H(x(w_1, w_2, ..., w_k))$.

A simple example where equation (4) can be solved explicitly for x is given by the cost function $H(x) = m \cdot x$ and preferences over the public good $R_k(x, w_i) = w_i \cdot (x - x^2)$, which (for interior solutions) generate the revenue requirement

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

This case constitutes an example where the revenue requirements satisfy the conditions of Theorem 2(IC), namely concavity and argument-additivity. A majority rule equilibrium exists for any given sample size.

In general, it is difficult to trace the properties of the revenue requirement function back to the structure of both the cost function $H(x)$ and the subutility function $r(x, w)$. However, we now show that the case where both primitive functions are isoelastic has a simple solution that, given reasonable values for the parameters, verifies the conditions for existence of a majority rule equilibrium.

Theorem 5: Let $u(c,l,x,w) = a \cdot c - b(l,w) + \frac{r \cdot w}{1-\alpha} x^{1-\alpha}$, and let $H(x) \equiv m \cdot x^{\beta}$, with $\alpha > 1$, $\beta \geq 1$. 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.

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 Appendix contains an example of interest.

To compute the efficient level for the public good in a "Bergstrom-Cornes" economy, one need only maximize the sum of utilities over the feasible set. The revenue requirement function is the cost of providing this level of the public good. Even though a condition such as limited complementarity are not always verified, for the cases where they hold we obtain existence of a majority rule equilibrium when voting occurs simultaneously over income tax functions and levels of public good provision. This result is considerably stronger than existing results where the tax functional form is taken as given and voting occurs over the value of one parameter of the tax function.

5 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 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. Even though our example illustrating Theorem 1 generates exactly the revenues needed, that will not generally happen and excess revenue will be raised in most cases. Consider first how this point might apply to the economies included in Theorem 5. There, 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 the type of example used in Theorem 5, 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.[26]

[26] We know of only one case where an optimal tax is linear: Snyder and Kramer (1988).

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 (1976) 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. 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 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.

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 and 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.

6 Appendix

6.1 Proof of Theorem 1

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 $g(w^*) = \frac{2}{k} \cdot R_k(w^*, w^*, ..., \underline{w}) - g(\underline{w}) = \frac{2}{k} \cdot R_k(w^*, w^*, ..., \overline{w}) - g(\overline{w})$.[27] 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, $g(\underline{w}) < \widehat{g}(\underline{w}) \Longrightarrow w^* <$

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

$\widehat{w}^* \implies g(\widehat{w}^*) - \widehat{g}(\widehat{w}^*) = \widehat{g}(\underline{w}) - g(\underline{w}) > 0$, proving the last statement in the Theorem.

Q.E.D.

6.2 An example satisfying limited complementarity

For the sake of brevity, we consider the case of k even. Define w^M as a median of $(w_1, w_2, ...w_k)$, namely if $w_1, w_2, ...w_k$ is ordered in a non-decreasing fashion, $w^M \in [w_{k/2}, w_{k/2+1}]$. We have that:

$$R_k(w_1, w_2, ...w_k) = \alpha \cdot \sum_{i=1}^{k} |w_i - w^M| + \beta \cdot \sum_{i=1}^{k} w_i,$$

and $0 < \alpha < \beta < 1$ as well as $\alpha/(1-\beta) < \underline{w}/(\overline{w} - \underline{w})$.

The adding-up restriction ($C3$ in this case) is as follows:

$$\frac{2}{k} \cdot R_k(\underline{w}, .., \underline{w}, \overline{w}, ..., \overline{w}) = \alpha \cdot (\overline{w} - \underline{w}) + \beta \cdot (\overline{w} + \underline{w}) = g(\underline{w}) + g(\overline{w}).$$

From the definition of limited complementarity, since the worst case scenario for type w is when $w^M = W(w, w^*)$, the minimal feasible individual revenue requirements with switch point w^* are:

$$g(w; w^*) = \beta \cdot (w + W(w, w^*)) + \alpha \cdot |w - W(w, w^*)| - g(W(w, w^*)). \qquad (5)$$

In particular we have that

$$g(w^*; w^*) = \beta \cdot (w^* + \underline{w}) + \alpha \cdot |w^* - \underline{w}| - g(\underline{w}) = \beta \cdot (w^* + \overline{w}) + \alpha \cdot |w^* - \overline{w}| - g(\overline{w}).$$

Thus,

$$\begin{aligned} g(\overline{w}) &= \beta \cdot (w^* + \overline{w}) + \alpha \cdot (\overline{w} - w^*) - \beta \cdot (w^* + \underline{w}) - \alpha \cdot (w^* - \underline{w}) + g(\underline{w}) \\ &= \beta \cdot (\overline{w} - \underline{w}) + \alpha \cdot (\overline{w} + \underline{w}) - 2\alpha w^* + g(\underline{w}) \end{aligned}$$

Setting $w = \underline{w}$ and substituting for $g(\overline{w})$ in (5), we obtain $g(\underline{w}; w^*) = \beta \cdot \underline{w} + \alpha \cdot (w^* - \underline{w})$. Using the adding-up restriction, we obtain $g(\overline{w}; w^*) = \beta \cdot \overline{w} + \alpha \cdot (\overline{w} - w^*)$. Hence $g(W(w_i, w^*); w^*) = \beta \cdot W(w_i, w^*) + \alpha \cdot |w^* - W(w_i, w^*)|$.

We now check feasibility, which says:

$$R_k(w_1, w_2, ...w_k) = \alpha \cdot \sum_{i=1}^{k} |w_i - w^M| + \beta \cdot \sum_{i=1}^{k} w_i$$

$$\leq \frac{2\alpha}{k} \cdot \sum_{i=1}^{k} \frac{k}{2} \cdot |w_i - W(w_i, w^*)| + \beta \cdot \sum_{i=1}^{k} (w_i + W(w_i, w^*)) - \sum_{i=1}^{k} g(W(w_i, w^*)).$$

This can be simplified to:

$$\sum_{i=1}^{k}|w_i - w^M| \leq \sum_{i=1}^{k}|w_i - W(w_i, w^*)| - \sum_{i=1}^{k}|w^* - W(w_i, w^*)|.$$

Now use the definition of $W(w, w^*)$ to rewrite the expression above as:

$$\sum_{i=1}^{k}|w_i - w^M| \leq \sum_{w_i \geq w^*}[(w_i - \underline{w}) - (w^* - \underline{w})] + \sum_{w_i < w^*}[(\overline{w} - w_i) - (\overline{w} - w^*)],$$

resulting in

$$\sum_{i=1}^{k}|w_i - w^M| \leq \sum_{i=1}^{k}|w_i - w^*|.$$

Since the median w^M is the parameter relative to which the sum of the absolute deviations is minimized we have that the inequality above necessarily holds. Furthermore we have that it holds as an equality when $w^* = w^M$. Since that will be a majority rule outcome we have that, in this particular instance, the sum of the individual tax payments matches exactly the collective revenue requirement.

6.3 Proof of Theorem 2

It is straightforward to prove by direct calculation that $\forall g(w; \widetilde{w}) \in G_k^*$ as given in the Theorem, $g(w; \widetilde{w})$ 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 Theorem. 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 Theorem 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:

$$g(w_1; \widetilde{w}) = Q((k-1) \cdot \widehat{w} + w_1) - (k-1)/k \cdot Q(k \cdot \widetilde{w}) - (k-1) \cdot Q'(k \cdot \widetilde{w}) \cdot (\widehat{w} - \widetilde{w}) \quad (6)$$

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 (6) 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 Theorem. 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 Theorem. 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 Theorem, 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: $\widetilde{w}' = \widetilde{w}$, $\widetilde{w}' < \widetilde{w}$, $\widetilde{w}' > \widetilde{w}$.

Q.E.D.

6.4 Proof of Theorem 3

Fix $\tau \in T_k$. If $y(w)$ is the gross income function associated with τ, then $\widetilde{g}(w) \equiv \tau(y(w)) \in B$ for some $B \in \mathcal{G}_k$. Pick $g(w) \in B \cap G_k^*$. If $g(w)$ is twice continuously differentiable, the remainder of the proof follows from Berliant and Gouveia (2001), Propositions 1 – 3.

Now suppose there exists one w^* such that $dg/dw\mid_{w^*}$ or $d^2g/dw^2\mid_{w^*}$ does not exist or is not continuous. Define two segments of g, $g^1(w)$ and $g^2(w)$, on the intervals $W^1 \equiv [\underline{w}, w^*]$ and $W^2 \equiv [w^*, \overline{w}]$, respectively. τ^* and θ^* over W^2 are again given by results in Berliant and Gouveia (2001), Propositions 1–3. Using the results in the proof of Proposition 1 in Berliant and Gouveia (2001), there is an extension of θ^* (and consequently of $\tau^*(y)$) through $(w^*, \theta^*(w^*))$ implementing g over W^1. By construction, incentive compatibility holds within both segments and since w^* is common to both intervals any solution to (2) that Pareto dominates τ^* over W^1 must necessarily violate global incentive compatibility.

The general problem with a finite number of non-differentiable points is solved by using repeatedly the technique above.

Q.E.D.

6.5 Proof of Lemma 1[28]

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 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 $y'(w^a) > y(w^a)$, $g'(w^a) > g(w^a)$.[29] 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.

Q.E.D.

[28] 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).

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

6.6 Proof of Theorem 4

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 (see the proof of Theorem 3). 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 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 Theorem 3, $\tau^* \in T_k^*$.

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 Theorem 3, we can take τ to be in T_k^* without loss of generality. Using Lemma 1, $\tau^*$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.

Q.E.D.

6.7 Proof of Theorem 5:

Using Lemma 2 and (4), the unique interior Pareto optimal level of x is given by

$$x_{PO} = \left[\frac{r}{am\beta} \sum_{i=1}^{k} w_i \right]^{\frac{1}{\alpha+\beta-1}},$$

which implies

$$R_k(w_1, w_2, ...w_k) = m \cdot \left[\frac{r}{am\beta} \sum_{i=1}^{k} w_i \right]^{\frac{\beta}{\alpha+\beta-1}}. \tag{7}$$

Since $\alpha + \beta > 1$, $\alpha > 1$, $R_k(w_1, w_2, ..., w_k)$ is concave and argument-additive. The remainder of the proof follows from Theorem 2(IC).

Q.E.D.

6.8 An Example of Theorem 5

Set $\alpha = 3$, $\beta = 2$, $r = 1$, so $u(c, l, x, w) = a \cdot c - l^2 - w \cdot x^{-2}/2$, and let $H(x) \equiv mx^2$. Then, from (7), we have

$$R_k(w_1, w_2, ..., w_k) = m \cdot \left[\frac{1}{2am} \cdot \sum_{i=1}^{k} w_i \right]^{\frac{1}{2}}.$$

Concavity and argument-additivity hold so, if we rule out bankruptcy problems, there is a majority rule equilibrium. Next we apply Theorem 2.

Since we can actually index all admissible g's by their $\widetilde{w}$'s, to find the choice of the median voter w^M we need only solve the problem: $\min_{\widetilde{w}} g(w^M; \widetilde{w})$. The solution to this problem is obtained when $w^M = \widetilde{w}$. Take $w \in [1, 2]$. Suppose we have a draw where $k \geq 2$ and the median voter is the type $w = 1.5$. Hence, by Theorem 2, we have that

$$g(w; \widetilde{w}) = \left(\frac{m}{2a} \right)^{\frac{1}{2}} \cdot \left[(\frac{\widetilde{w}}{k})^{\frac{1}{2}} + \frac{1}{2} (k\widetilde{w})^{-\frac{1}{2}} (w - \widetilde{w}) \right].$$

For notational simplicity define $\mu = \left(\frac{m}{2ak}\right)^{\frac{1}{2}}$. The majority equilibrium tax function implements the individual revenue requirement function

$$g(w) = \mu \cdot \left[(1.5)^{\frac{1}{2}} + \frac{1}{2} \cdot (1.5)^{-\frac{1}{2}} \cdot (w - 1.5) \right].$$

Applying equation (2) and $\frac{d\tau}{dy} = 1 - \frac{d\gamma}{dy}$, the income tax function is given by the solution to:

$$\frac{d\tau}{dy} = 1 - \frac{2y}{aw^2} = 1 - \frac{2y}{a\left[\frac{2\tau}{\mu} \cdot (1.5)^{\frac{1}{2}} - 1.5\right]^2},$$

with upper boundary at $(\overline{\tau}, \overline{y}) = \left(\mu \cdot \frac{7}{3^{1/2} \cdot 2^{3/2}}, 2a\right)$.

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

[4] 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

[5] Berliant, M. and F.H. Page, Jr., 1999, Incentives and income taxation: The implementation of individual revenue requirement functions," Ricerche Economiche 50, 389-400.

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

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

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

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

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

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

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

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

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

[15] 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.

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

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

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

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

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

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

[22] 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).

[23] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[43] Seade, J.K., 1982, On the sign of the optimum marginal income tax, Review of Economic Studies 49, 637-643.

[44] 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.

[45] Spence, M., 1974, Market Signaling (Harvard, Cambridge, Mass.).

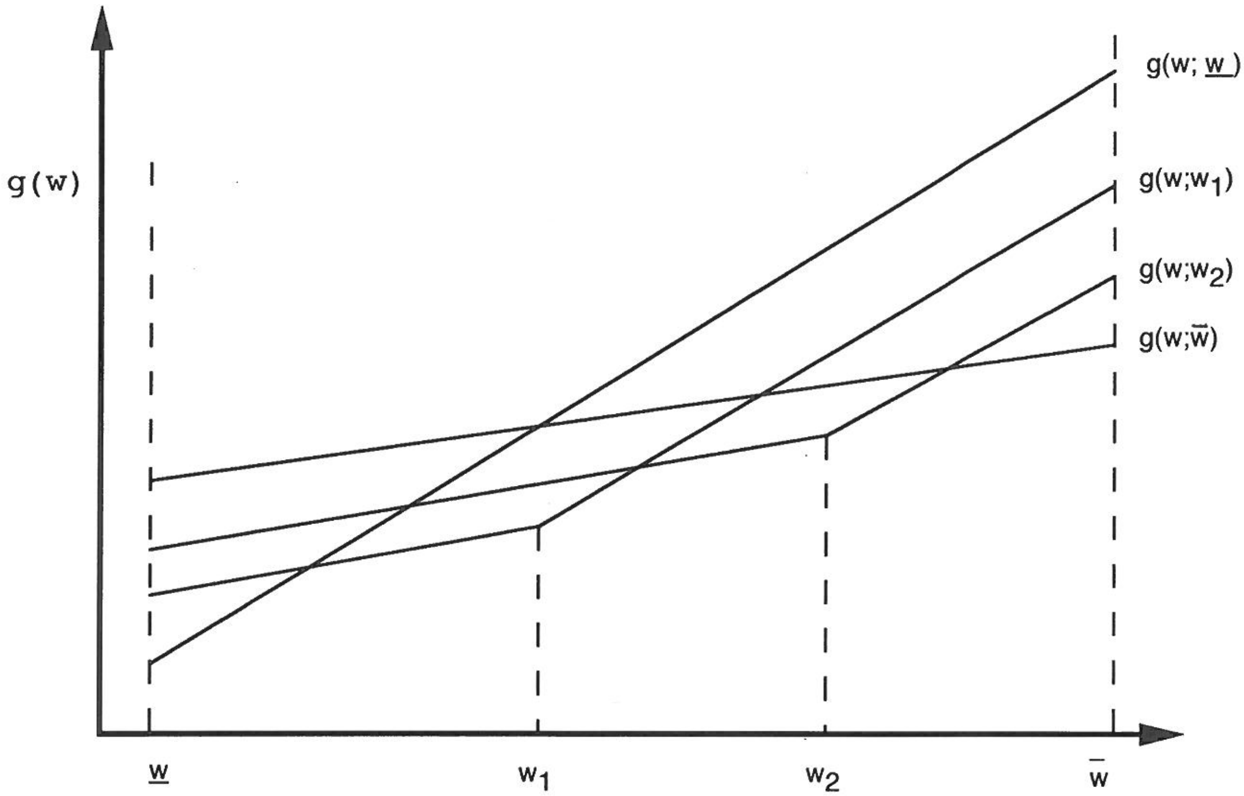

FIGURE 1: LIMITED COMPLEMETARITY

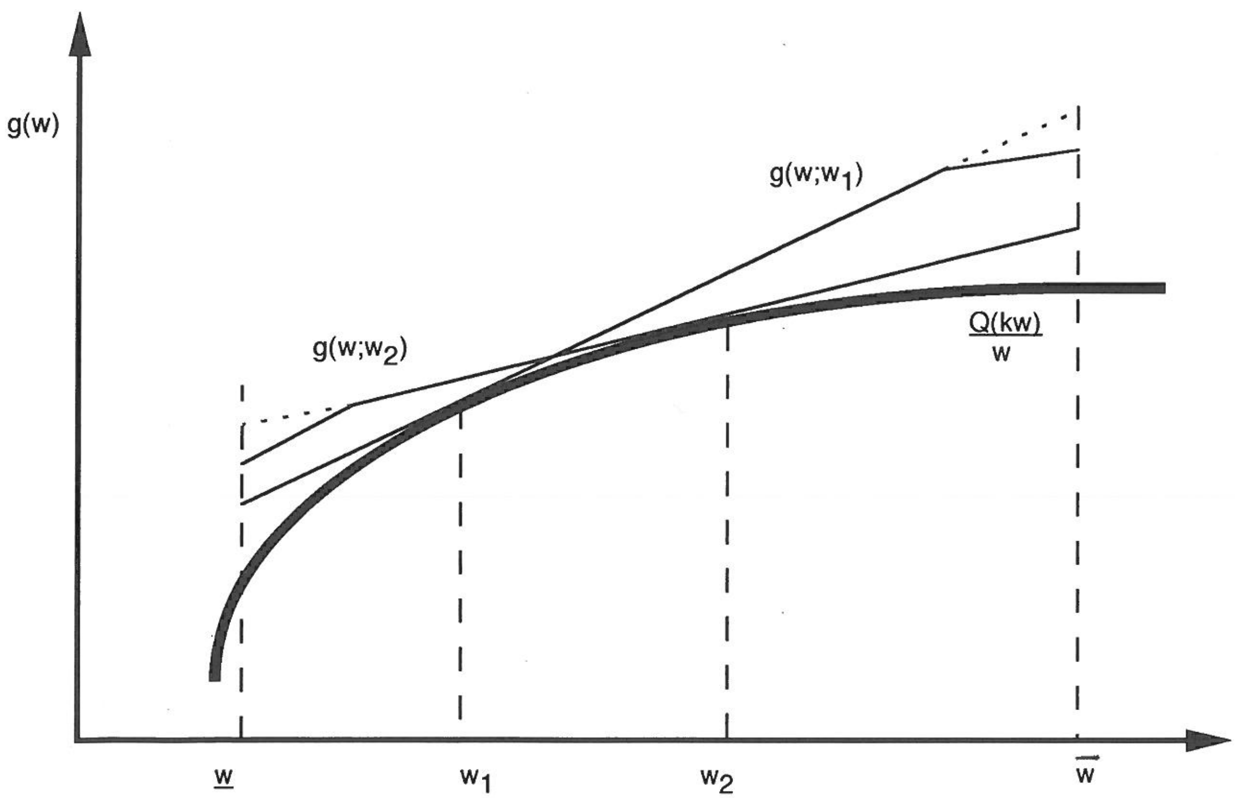

FIGURE 2: CONCAVE ARGUMENT-ADDITIVITY