On the Political Economy of Income Taxation

Berliant, Marcus and Miguel Gouveia

Working Paper No. 288
August 1991

On the Political Economy of Income Taxation

Marcus Berliant and Miguel Gouveia

Rochester Center for Economic Research
Working Paper No. 288

August 1991

On the Political Economy of Income Taxation±

Marcus Berliant*
and
Miguel Gouveia**

and

The Rochester Center for Economic Research
Working Paper No. 288

August 1991

---

*Department of Economics, University of Rochester, Rochester, NY 14627.

**Department of Economics, University of Pennsylvania. Support provided by a Sloan Foundation Fellowship is gratefully acknowledged.

± Support provided by National Science Foundation grant SES-9007800 is gratefully acknowledged. The comments of Randy Calvert, Roy Gardner, Hideo Konishi and Jim Snyder are appreciated, but the authors bear full responsibility for any errors.

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. 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.

1. Introduction

The theory of income taxation has been an important area of study in economics. Interest in a formal theory of income taxation goes 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 important 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 important 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, so that the optimal income tax schedule must separate individuals as well as maximize welfare and therefore is generally second best[1]. The necessary conditions for 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 model) was demonstrated in Kaneko (1981). 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

[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*.

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 proposed below.

The main objective of this research is actually to derive some 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 both 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. Romer (1975), Roberts (1977), and Peck (1986) use linear taxes in voting models with work disincentives. 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 (1984) 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. Unfortunately, this interesting work does not contain an explicit game-theoretic voting model; the individually optimal income tax might not be a majority rule equilibrium or in the voting core.

We propose in this paper to allow general nonlinear income taxes with work disincentives (adverse selection and moral hazard) in a voting model. We are indebted to J. Snyder for emphasizing to us that 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, 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?

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. 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 population is unknown or, more important, when the characteristics of the population are variable. 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 size and its 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 general incentive compatibility requirement and its implications are developed in detail in section III below. 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). We then have that 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.

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 requirements function was proposed and examined to some extent in Berliant (forthcoming), and is developed further in more generality in section II below.

Once the feasible set is explored, we will examine several games to see if equilibria exist, are unique, and can be characterized (the latter step having predictive value). Of primary interest at the start is the existence and characterization of a majority equilibrium studied in section III. Voting over both a public good and taxes is studied in section IV. Section V contains conclusions and suggestions for further research.

We do not claim that the particular games examined here are the "correct" ones in any sense. The point of this work is that there is a natural structure and set of arguments that can be exploited in voting games over income taxes to obtain existence and sometimes uniqueness and characterization results.

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 V 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 assumption 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 the Berliant and Gouveia (1991) to be sure that the second order conditions for incentive compatibility hold in our model.

2. The Model and Notation

A classical model of optimal income taxation can be developed formally as follows. Consumers' characteristics or wages are described by a single variable $w \in [\underline{w}, \overline{w}]$, where $[\underline{w}, \overline{w}]$ is an interval contained in the positive real line. References to measure are to Lebesgue measure on $[\underline{w}, \overline{w}]$. The distribution of consumer characteristics or wage levels has a measurable density $f(w)$, where $f(w) > 0$ *a.s* [2]. 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. $u(c, \ell)$ 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 following assumptions are maintained throughout this paper:

A1: Standard assumptions on preferences:

$u_1 > 0,\ u_2 < 0,\ u_{22} < 0,\ u_{11} < 0.$

A2: The utility function is quasi-concave:

$u_{11}u_2^2 - 2u_{12}u_1u_2 + u_{22}u_1^2 < 0.$

A3: Consumption is normal:

$u_{21}u_2 - u_{22}u_1 > 0.$

A4: Leisure is normal:

$u_{11}u_2 - u_{12}u_1 > 0.$

A5: Boundary conditions:

$\lim_{c\to 0} u_1(c, \ell) = \infty, \quad \lim_{\ell\to 1} u_2(c, \ell) = -\infty, \quad \lim_{\ell\to 0} u_2(c, \ell) = 0.$

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

2 Note that $f(w)$ plays almost no role in the developments 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. In the multistage games of voting in a representative democracy that we expect to study in the future, the equilibria are likely to be a function of f, as is often the case in signaling games.

tax literature to obtain the single crossing property for indifference curves. Assumption A5 is also common in the optimal income tax literature and is used to derive comparative statics there. 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.

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

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:

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

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

The production side of the economy is given as follows. Let $\mathcal{A}_k \equiv \{S | S = (w_1, ..., w_k)$ where $w_1, ..., w_k \in [\underline{w}, \overline{w}]\}$, the collection of all draws of k individuals from the distribution with density f. Define $\mathcal{A} \equiv \cup_{k=1}^{\infty} \mathcal{A}_k$. $\mathcal{A}$ is the collection of all draws of all sizes from the distribution f. A *draw* is an element, say S, of $\mathcal{A}$. In order to be able to determine what any particular draw can produce or 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 $R : \mathcal{A} \to \mathbf{R}$. For each $S \in \mathcal{A}$, $R(S)$ represents the taxes due from a draw less endowment of consumption good. For example, if the revenues from the income tax are used to finance a good such as schooling, then $R(S)$ can be seen as: the per capita revenue requirement for providing schooling to the draw S multiplied by the population in S.

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 will be imposed and discussed below. One basic assumption that we will maintain throughout is that R is *attainable* in the sense that there is some labor supply that will generate enough tax revenue to satisfy R. Formally, R is *attainable* if for every k and every $(w_1, ..., w_k) \in \mathcal{A}_k$, $R((w_1, ..., w_k)) < \sum_{i=1}^{k} w_i$.

Next the production correspondence is defined formally. Let $I_k = [0,1]^k$. For each

$S \in \mathcal{A}_k$, production possibilities are described by a set $Y_k(S)$, where $Y_k(S) \subset I_k \times \mathbf{R}$. For a given S, $(\underline{\ell}, C) \in Y_k(S)$ describes labor input $\underline{\ell}_i \equiv \ell(w_i)$ of person i for $w_i \in S$, 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 S are zero. We assume throughout that the endowment of consumption good of a draw as well as tax revenue due from a draw are independent of labor supply $\underline{\ell}$ and composite good consumption C. As is almost universal in the optimal income tax literature, a constant returns to scale technology is postulated. Formally,

$$Y_k(S) \equiv \{(\underline{\ell}, C) \in I_k \times \ \mathbf{R} | \sum_{i=1}^{k} w_i \underline{\ell}_i \geq C + R(S)\}.$$

It is important to be clear about the interpretation of R. One easy interpretation is that the taxing authority provides a schedule giving the taxes owed by any draw. There are several reasons that revenue requirements might differ among draws, including differences in taste for a public good that is implicitly provided, a non-constant marginal cost for production of the public good, differences in the cost of revenue collection, and (perhaps most importantly) a government attempt to achieve income redistribution in the realized draw.

This structure captures some important aspects of the optimal income tax model. First, labor is modelled 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.

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[3] as well as the mapping R. 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.

Next we impose a topology on $\mathcal{A}_k$. A topology will be induced on $\mathcal{A}$ as a consequence, but will not be useful to us since draws of different sizes will never be "close". In fact, we could start by defining a topology on $\mathcal{A}$, but this would obfuscate rather than clarify the development. For $\mathcal{A}_k$, we use the Euclidean norm $\|\cdot\|_k$ on the subspace $[\underline{w}, \overline{w}]^k$.

---

3 Actually, all they need to know is the support of that distribution.

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 ability or wage) 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.

Definition: Fix k and a revenue requirement function R. The set of *feasible* tax systems is defined to be $T_k \equiv \{\tau : \mathbf{R} \rightarrow \mathbf{R} | \ \tau$ is measurable, for all $(w_1, ..., w_k) \in \mathcal{A}_k, \sum_{i=1}^{k} \tau(y(w_i)) \geq R((w_1, ..., w_k))\}$. Define the set of tax systems that are feasible for draws of all sizes to be $T \equiv \cap_{k=1}^{\infty} T_k$.

Finally, the notion of a majority rule equilibrium can be defined. Fix a revenue requirements function R. In general, one cannot expect that the same tax system is unblocked by a majority (or even Pareto optimal) for *every* draw. Hence, it is reasonable to associate a set of majority rule winners with *each* draw, and allow this set to vary with the draw. To this end, a *majority rule equilibrium* for draws of size k is a correspondence $M_k : \mathcal{A}_k \rightarrow T_k$ such that for every $(w_1, ..., w_k) \in \mathcal{A}_k$, for every $\tau \in M_k((w_1, ..., w_k))$ (with associated $y(w)$), there is no subset D of $\{w_1, ..., w_k\}$ of cardinality greater than $k/2$ along with another $\tau' \in T$ (with associated $y'(w)$) such that $u(y'(w) - \tau'(y'(w)), y'(w)/w) > u(y(w) - \tau(y(w)), y(w)/w)$ for all $w \in D$. We will show that we can restrict attention to continuous τ. An obvious extension of the definition to $\mathcal{A}$ is possible: $M : \mathcal{A} \rightarrow T$ is a correspondence such that for all k, for all $(w_1, ..., w_k) \in \mathcal{A}_k$, for all $\tau \in M((w_1, ..., w_k))$ (with associated $y(w)$), there is no subset D of $\{w_1, ..., w_k\}$ of cardinality greater than $k/2$ along with another $\tau' \in T$ (with associated $y'(w)$) such that $u(y'(w) - \tau'(y'(w)), y'(w)/w) > u(y(w) - \tau(y(w)), y(w)/w)$ for all $w \in D$.

3. Voting Over Taxes in an Optimal Income Tax Economy

3.1 Overview

This section starts by studying the derivation of *individual* revenue requirements consistent with *overall* revenue requirements that possess certain desirable properties. Next, some results from the literature on optimal income taxation are used to construct the best income tax function that implements a given individual revenue requirement. Then, we demonstrate the power of our approach in two particularly simple cases where the feasibility conditions are so strict that voting need not occur in order to select an income tax function. Finally we study two cases where the feasible sets are non-trivial. We prove the existence of a voting equilibrium for each case and provide a partial characterization of the equilibria.

3.2 From Revenue Requirements to Tax Functions

In order to examine the set of feasible tax systems described above, more structure needs to be introduced. In particular, since revenue requirements described so far are in the form of a set function, *individual* revenue requirements remain unspecified.[4] Clearly, there will generally be a range of individual revenue requirements consistent with any map R. Our next job is to describe this set formally. Fix k and R. Let $G_k \equiv \{g : [\underline{w}, \overline{w}] \to \mathbf{R} |\ g$ is measurable, $\forall\ (w_1, ..., w_k) \in \mathcal{A}_k, \sum_{i=1}^{k} g(w_i) \geq R((w_1, ..., w_k)), g(w_i) < w_i\ \forall\ i\}$. Define $G \equiv \cap_{k=1}^{\infty} G_k$. G_k is the set of all *individual* revenue requirements that collect enough revenue to satisfy R, and G is the set of all *individual* revenue requirements that satisfy R for all k. $G_k \neq \emptyset$ if R is attainable. We now search for the minimal elements of these sets. 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 $\mathcal{G}_k \equiv \{B \subseteq G_k | B$ 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 $G_k^* \equiv \{g : [\underline{w}, \overline{w}] \to \mathbf{R} |\ \exists B \in \mathcal{G}_k$ such that $\forall\ w \in [\underline{w}, \overline{w}] \quad g(w) = \inf_{g' \in B} g'(w)$ a.s.$\}$. G_k^* is nonempty.

We now consider draws of any size. Define a binary relation $\succeq$ over G by $g \succeq g'$ if and only if $g(w) \geq g'(w)$ for almost all $w \in [\underline{w}, \overline{w}]$. Let $\mathcal{G} \equiv \{B \subseteq G | B$ is a maximal totally ordered subset of $G\}$. Finally, define $G^* \equiv \{g : [\underline{w}, \overline{w}] \to \mathbf{R} |\ \exists B \in \mathcal{G}$ such that $\forall\ w \in [\underline{w}, \overline{w}] \quad g(w) = \inf_{g' \in B} g'(w)$ a.s.$\}$. It is possible that $G^* = \emptyset$.

Next we need to impose some further conditions on R.

---

4 An individual revenue requirement gives the amount of tax an individual of a given *ability* should pay. Notice that the ability level of each individual is a characteristic of that individual not observable to others. The individual revenue requirement differs from the tax function: the tax function gives the amount of tax that an individual with a certain *income* should pay. Income is observable by the government and an individual of a certain ability may choose to earn different amounts of income, depending on the incentives provided by different tax structures. The relation between individual revenue requirements and tax functions will be examined later in this section of the paper.

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

This is a natural assumption on R, 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 coalition is which types are drawn from the distribution.

Definition: A revenue requirement function R is said to be *monotonic* if for any $(w_1, ..., w_k)$,$(w'_1, ..., w'_k) \in \mathcal{A}_k$ with $w_i \geq w'_i$ for all i and strict inequality holding for some j, then $R((w_1, ..., w_k)) > R((w'_1, ..., w'_k))$.

This assumption requires that higher ability draws owe more taxes. One could successfully use weaker assumptions with this framework, such as $w_i \geq w'_i$ for all i implies $R((w_1, ..., w_k)) \geq R((w'_1, ..., w'_k))$, but at a cost of greatly complicating the proofs. We refer to Berliant (forthcoming, pp. 23-26) for the appropriate techniques.

There are two conceptual steps in the process of going from revenue requirement functions defined over draws to income tax functions defined over individual incomes. The first step involves the construction of G^*_k. Results for this step are presented in Theorems 2–4, where we show that the individual revenue requirement functions are increasing and continuously differentiable except at a finite number of points in several important cases.

The second step is to construct incentive compatible tax systems out of these individual revenue requirements. This is done in Theorem 1, using standard methods from optimal income taxation. In the paragraphs following, we give intuition for this implementation result.

The problem confronting a *worker/consumer* of type w given net income schedule γ is $\max_\ell u(\gamma(w \cdot \ell), \ell)$. 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,

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

For this new 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 each such g is 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,

$$\frac{d\gamma}{dy} = -\frac{u_2(\gamma, \frac{y}{g^{-1}(y-\gamma)})}{u_1(\gamma, \frac{y}{g^{-1}(y-\gamma)})} \cdot \frac{1}{g^{-1}(y-\gamma)} \equiv F(\gamma, y) \tag{1}$$

As in Berliant (forthcoming, page 25), a standard result from the theory of differential equations yields a family of solutions to this differential equation. Of course, as L'Ollivier and Rochet (1983) point out, the second order conditions must be checked to ensure that solutions to (1) do not involve bunching which means that consumers do optimize in (1) at the tax liability given by g. This was done in Berliant and Gouveia (1991), where the Revelation Principle (see, for example, Laffont (1988, chapter 6)) was used to construct strictly increasing net income functions $\theta(w)$ that implement $g(w)$. Since we then have that $y(w) = g(w) + \theta(w)$ is invertible, we immediately obtain $\gamma(y) = \theta(w^{-1}(y))$ and $\tau(y) = g(w^{-1}(y))$.

It is almost immediate from this development that the set of solutions to (1) for a given g is Pareto ranked. We focus on the best of these for each given g. Define $T_k^* \equiv \{\tau |\ \gamma$ is a solution to (1) for some $g \in G_k^*$, $\tau(y) = y - \gamma(y)$, and τ Pareto dominates all other solutions to (1) for the given $g\}$.

Theorem 1: 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: Note that when $g(w)$ is C^2 so is τ^* (see Berliant and Gouveia (1991)). 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 $1 + u_2(y - \tau, y/w^*)/(u_1(y - \tau, y/w*)w^*)$.

If $\tau \in T_k \backslash T_k^*$ is proposed as an alternative to $\tau^* \in T_k^*$, $\exists\ \tau' \in T_k^*$ that is unanimously weakly preferred to τ. In multistage games, it will be interesting to allow players to propose incredible threats, such as those in $T_k \backslash T_k^*$, which might then be ruled out in equilibrium by using refinements. For now, however, we will focus on a one stage game and consequently on the set T_k^*. This is what the natural information structure implied by the optimal tax model buys us. The restriction that the revenue requirement be satisfied for each draw restricts the feasible set significantly.

Define $T^* \equiv \{\tau \in T|$ there does not exist $\tau' \in T$ such that $\max_y u(y - \tau'(y), y/w) \geq \max_y u(y - \tau(y), y/w)$ for almost all w, with strict inequality holding for a set of positive measure$\}$. There is no guarantee that $T^* \neq \emptyset$ without further assumptions, but T^* represents the set of best tax systems feasible for any finite population and any draw.

3.3 Singleton Feasible Sets

Next we illustrate some special cases in which the feasible set is quite small.

Definition : A revenue requirement function R is said to satisfy *individual increasing returns to scale* if for each k and each $(w_1, ..., w_k) \in \mathcal{A}_k$, $R((w_1, ..., w_k)) \leq \sum_{i=1}^k R((w_i))$.

This assumption says that larger economies need less revenue per capita than 1-consumer economies[5]. A class of examples of satisfying individual increasing returns is the collection of additively separable functions, $R((w_1, ..., w_k)) = \sum_{i=1}^k s(w_i)$, where s is smooth and increasing.

Theorem 2: Suppose that R is twice continuously differentiable, symmetric and monotonic on $\mathcal{A}_k$ for all k. Suppose further that R satisfies individual increasing returns to scale. Then T^* consists of exactly one tax system, which is the best one obtained as the solution to (1) where $g(w)$ is replaced by $R((w))$ and where the marginal tax rate for the top ability $\overline{w}$ consumer type is zero.

Proof: Let $g^*(w) \equiv R((w))$. Then g^* is C^2. Applying the results in Berliant and Gouveia (1991), Propositions 1-3, there exists a C^2 tax system τ^* implementing g^* that Pareto dominates any other tax system implementing g^* and is such that the marginal tax rate for the top ability consumer type is zero. Now let $\tau \in T, \tau \neq \tau^*$. Define $\hat{y}(w) \equiv \arg\max_y u(y - \tau(y), y/w)$ and $g(w) \equiv \tau(\hat{y}(w))$. By definition of τ^*, for $k = 1$ and any draw $(w_1) \in \mathcal{A}_1$, $g(w_1) \geq R((w_1))$. If $g(w_1) > R((w_1))$ for some w_1, then τ is dominated by τ^*. If $\tau(\hat{y}(w_1)) = R((w_1))$ for each w_1, then $\tau = \tau^*$ or τ is Pareto dominated by τ^*. Hence $\forall\, \tau \in T$, either

[5] The more restrictive assumption $R((w_1,...,w_k)) \leq R((w_1,...,w_j)) + R((w_{j+1},...,w_k))$ for each j and k is not needed.

$\tau = \tau^*$ or τ is weakly Pareto dominated by τ^*. So $\tau^* \in T^*$ and τ^* is the only element of T^*.

Q.E.D.

Definition : A revenue requirement function is said to satisfy *revenue complementarity* if for each k and for each $(w_1, ..., w_k) \in \mathcal{A}_k$, $R((w_1, ..., w_k)) \leq \sum_{i=1}^k R((w_i, ..., w_i))/k$.

This assumption says that the presence of different types in a draw reduces the per capita revenue needs. This could be due to implicit complementarities either in the production of public good or in the decision process used to arrive at the level of public good provision. The class of additively separable functions mentioned earlier also verifies revenue complementarity. Another example, studied in Appendix III, is given by functions of the form $R((w_i, ...w_k)) = z(\sum_{i=1}^k w_i)$, with z convex. However, in general revenue complementarity need not be related to standard concavity or convexity assumptions.

Theorem 3: Fix k and suppose that R is continuously differentiable, symmetric and monotonic on $\mathcal{A}_k$. Suppose further that R satisfies revenue complementarity. Then T_k^* consists of exactly one tax system, which is obtained as the best solution to (1) where $g(w)$ is replaced by $R((w, ..., w))/k$ and where the marginal tax rate for the top ability consumer $\overline{w}$ is zero.

Proof: Let $g(w) \equiv R((w, ..., w))/k$. Then g is C^2. Using revenue complementarity, for any draw $(w_1, ..., w_k) \in \mathcal{A}_k$, $R((w_1, ..., w_k)) \leq \sum_{i=1}^k R((w_i, ..., w_i))/k = \sum_{i=1}^k g(w_i)$, so $g \in G_k$. Since $R((w, ..., w)) = k \cdot g(w)$, $g \in G_k^*$. Hence applying Theorem 1, the tax system τ^* defined in the Theorem is in T_k^*.

Now let $\tau \in T_k^*$, $\tau \neq \tau^*$. Define $\hat{y}(w) \equiv \arg\max_y u(y - \tau(y), y/w)$ and $h(w) \equiv \tau(\hat{y}(w))$, where $h \in G_k^*$. By definition of T_k^*, $h(w) \geq R((w, ..., w))/k$. If $h(w) > R((w, .., w))/k$, $h \notin G_k^*$, a contradiction. If $\tau(\hat{y}(w)) = R((w, ..., w))/k$ for each w, then $\tau = \tau^*$. So $\tau^* \in T_k^*$ and τ^* is the only element of T_k^*.

Q.E.D.

Of course, in each of these two instances, there is no need to vote over alternatives, since the set over which voting occurs (credible and feasible tax systems) is a singleton in each case. That is, given the assumptions of Theorem 2, $M((w_1,...,w_k)) \equiv T^*$. Given the assumptions of Theorem 3, for any given k, $M_k((w_1,...,w_k)) \equiv T^*_k$.

3.4 Voting Over Non-trivial Sets

Having defined the feasible set of tax systems T^*_k when the population of the economy is known and T^* when the population of the economy is unknown, we shall next turn to voting games over the feasible sets when feasible sets are not quite as trivial. There are many voting games that might be natural and interesting.

We now present a set of definitions that will play an important role in the results to follow.

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

$C1$. $g(\underline{w}) \geq R((\underline{w},\underline{w},...,\underline{w}))/k$.

$C2$. $g(\overline{w}) \geq R((\overline{w},\overline{w},...,\overline{w}))/k$.

$C3$. For k even:

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

$C4$. For k odd:

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

Definition: The *set of admissible extreme revenue requirements* is $EG_k \equiv \{(g(\underline{w}),g(\overline{w})) \mid (g(\underline{w}),g(\overline{w}))$ verifies C1-C4. $\}$.

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 $w < w^*$.

Definition: A revenue requirement function R 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 $(w_1,...,w_k) \in \mathcal{A}_k$ the following holds:

– For k even:

$$R((w_1,...,w_k)) \leq \sum_{i=1}^{k}[R((w^i_1,...,w^i_k))2/k - g(W(w_i,w^*))],$$

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

–For k odd:

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

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 $j = (k+3)/2,...,k$.

This assumption means that revenue requirements are maximal for draws consisting of at most two types of players. Maximal revenue draws for type w consist of people of type w and people of the type most unlike w, either $\underline{w}$ or $\overline{w}$.

We will provide an example satisfying this assumption below.

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

Consider first one-stage voting over the feasible set, say T_k for some k. Here, all agents in a draw simply submit their votes over every possible pair of tax systems, given that they can vote for only one out of each pair. An arbitrator would pick the majority rule winner (if there is one). There is no information transmission in voter behavior in this type of game (until the game is over). Preferences of a *voter/worker/consumer* are induced over tax systems by their utility levels after a tax system is imposed. The next result establishes the individual revenue requirements in G_k^* cross exactly once.

Lemma 1: Let k be a positive integer. Suppose that R is twice continuously differentiable, symmetric, and monotonic. Finally, suppose that R satisfies limited complementarity and Edgeworth substitutability. Then, $\forall\, g \in G_k^*$, g is strictly increasing and for any $g, g' \in G_k^*$, there exists a $\tilde{w} \in [\underline{w}, \overline{w}]$ such that, $g(\underline{w}) \geq g'(\underline{w})$ implies $g(w) \geq g'(w)$ for all $w \in [\underline{w}, \tilde{w}]$ and $g(w) \leq g'(w)$ for all $w \in [\tilde{w}, \overline{w}]$.

Moreover, for any $g, g' \in G_k^*$, with switching points w^* and w'^*, $g(\underline{w}) > g'(\underline{w})$ implies $w^* > w'^*$.

Finally, the $g \in G_k^*$ that minimizes $g(\hat{w})$ has a switching point $w^* = \hat{w}$.

Proof: See the Appendix.

Remark: Given the assumptions on R and the fact (proved in the appendix) that $g(w_i) = [R((w_k^i, ..., w_k^i))2/k - g(W(w^i, w^*))]$ 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.

Besides generating single crossing individual revenue requirements, the cases where limited complementarity holds have an interesting characterization: they constitute a perfect illustration of the principle known as Director's Law of Income Redistribution. This principle, first formalized in Stigler (1970), says that middle income or ability classes minimize their tax burdens by pressuring for heavier burdens on the poor and/or the rich.

To see why that is the case here, consider the incidence of the individual revenue requirements function chosen by the voter with the median wage rate w^M in any given draw, taking the wage rate of voter i as the measure of his ability to pay. The first thing to notice is that the $g(w)$ chosen is the one that minimizes $g(w^M)$, and raises $g(\underline{w})$ and $g(\overline{w})$ relative to feasible alternatives preferred by either the top or the bottom ability levels. In fact, while Stigler describes circumstances where the middle income or ability class forms coalitions with either the poor or the rich in order to always win minimum tax burdens, here we have that the median voter shifts the tax burden simultaneously to both extrema of the ability distribution. A second and related characteristic is the progressivity of the marginal rates in the resulting revenue requirements. The proof of Lemma 1 shows that for low abilities (i.e. $w < \tilde{w}$) we have

$$\frac{dg}{dw} = \frac{k}{2} \frac{\partial R((w, ..., w, \overline{w}, ..., \overline{w})}{\partial w}$$

and that for high abilities ($w > \tilde{w}$) we obtain

$$\frac{dg}{dw} = \frac{k}{2} \frac{\partial R((w, ..., w, \underline{w}, ..., \underline{w})}{\partial w}.$$

Using the property of Edgeworth substitutability we can prove that in a neighborhood of w^M the marginal revenue requirement rates will be higher for higher abilities. In general Edgeworth substitutability will push the marginal $g's$ down for low abilities and the converse for high abilities.

Such a result cannot be obtained formally in models restricted to linear tax functions, such as Meltzer and Richard (1981,1983). Increasing marginal tax rates imply that taxes decrease little when going from the middle to the bottom of the ability distribution but increase faster when you proceed in the opposite direction. Overall, increasing marginal revenue requirement rates tend to decrease the fiscal burden imposed on the middle income classes[6]. This result agrees entirely with the comments in Foley (1967) about the role of progressivity and with the results obtained by Snyder and Kramer (1988), despite the fact that they use a substantially different model.

A particularly striking example of a case satisfying limited complementarity is given by the following collective revenue requirement function [7]

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

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

This function lends itself to a plausible interpretation: the term $\beta \sum_{i=1}^{k} w_i$ can be seen as the direct cost of public expenditure demanded by the types in the sample and the term $\alpha \sum_{i=1}^{k} |w_i - w^M|$ is related to the costs of arriving at a collective decision, since they increase as the divergence or disagreement among types in a draw increases (measuring divergence by the sum of deviations to the median type).

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

$$g(w) = \beta w + \alpha(\tilde{w} - w) \quad \text{if } w < \tilde{w} \quad \text{and}$$

$$g(w) = \beta w + \alpha(w - \tilde{w}) \quad \text{if } w \geq \tilde{w}.$$

In this case the two branches of the individual revenue requirements function are linear and the marginal requirement is thus always higher for the higher ability taxpayers.

We now look at a different case that implies single crossing of individual revenue requirements in G_k^*.

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

---

[6] The reader will notice that this statement applies to revenue requirements and not directly to tax functions. That may not be important because tax functions will inherit the basic characteristics of the underlying revenue requirements. Also, the fact remains that in an optimal income tax economy the proper measure of ability to pay is the taxpayer's ability level or wage rate, so characterizing the distributional incidence of a tax structure in terms of abilities and the revenue requirements on these abilities is perfectly adequate.

[7] 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 (when it goes from above to below the median). Finding smooth closed form examples has proved to be a difficult task, but there are smooth functions uniformly close to this example satisfying all conditions.

Lemma 2: Let $k \geq 2$ and let the revenue requirement function $R((w_1, ..., w_k))$ be symmetric, monotonic and twice continuously differentiable. Assume also that $R((w_1, ..., w_k))$ is argument-additive and concave.

Then, we have that $\forall\, g \in G_k^*$, g is as follows:

– For $\tilde{w} \geq (\underline{w} + \overline{w})/2$, $g \in G_k^* =>$

$$\text{A)} \quad g(w; \tilde{w}) = Q(k\tilde{w})/k + Q'(k\tilde{w})(w - \tilde{w}) \qquad \text{if } w \leq \tilde{w} + (k-1)(\tilde{w} - \underline{w})$$

$$\text{B)} \quad g(w; \tilde{w}) = Q((k-1)\underline{w} + w) - ((k-1)/k)Q(k\tilde{w}) + (k-1)Q'(k\tilde{w})(\tilde{w} - \underline{w})$$

$$\text{if } w > \tilde{w} + (k-1)(\tilde{w} - \underline{w}).$$

– For $\tilde{w} < (\underline{w} + \overline{w})/2$, $g \in G_k^* =>$

$$\text{C)} \quad g(w; \tilde{w}) = Q(k\tilde{w})/k + Q'(k\tilde{w})(w - \tilde{w}) \qquad \text{if } w \geq \tilde{w} - (k-1)(\overline{w} - \tilde{w})$$

$$\text{D)} \quad g(w; \tilde{w}) = Q((k-1)\overline{w} + w) - (k-1)/kQ(k\tilde{w}) + (k-1)Q'(k\tilde{w})(\overline{w} - \tilde{w})$$

$$\text{if } w < \tilde{w} - (k-1)(\overline{w} - \tilde{w}).$$

for $\tilde{w} \in [\underline{w}, \overline{w}]$. Furthermore, $\forall\, w \in [\underline{w}, \overline{w}]$, $g(w)$ is single caved[8] in $\tilde{w}$ and attains a minimum at $\tilde{w} = w$.

Proof: See the Appendix.

Remark: The intuition for this result is quite simple. Consider (for the moment) a case where the distribution of abilities is not bounded above or below. Since the revenue requirement R is concave, so is the per capita revenue requirement R/k. But then, the only functions that can be individual revenue requirements are the tangents to R/k, since any linear combination of individual revenue requirements has to be greater than or equal to the per capita requirement. The statement of the Lemma is slightly more complex because this intuition may not work near the limits $\underline{w}$ or $\overline{w}$.

Note that the marginal revenue requirement rates in branch B are lower than the rate in branches A and C (the tangent branches), which is lower than those in branch D. In a

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

certain sense the individual revenue requirements resulting from argument-additivity and concavity are the opposite of those obeying Director's Law: marginal revenue requirement rates are constant for middle ability levels and, for some draws with extreme values, we may see lower marginal requirement rates for high abilities or higher marginal requirement rates for low abilities. 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. In that sense, Director's Law is no more of a logical necessity than its converse. Notice also that the shape of the distribution of abilities (and consequently the distribution of incomes) does not have in itself sufficient information to predict the shape of the income tax schedules chosen by majority rule.

Lemmata 1 and 2 established that, in the two cases considered, the minimal individual revenue requirements consistent with the aggregate revenue requirement for finite economies possess some important properties. Most likely there are other other classes of examples that also generate individual requirements with those properties.

Definition: Individual revenue requirements are *strongly single crossing* if they are :

i) Continuously differentiable except at a finite number of points.

ii) Strictly increasing.

iii) Any two individual revenue requirements cross each other only once.[9]

Lemma 3 proves that when individual revenue requirements are strongly single crossing, the income tax systems in T_k^* cross at most once.

Lemma 3: Let k be a positive integer. Suppose that R implies strongly single crossing individual revenue requirements.

Then for any $\tau, \tau' \in T_k^*$ and incomes $y_1 > y_2 > y_3$, $\tau(y_1) < \tau'(y_1)$ and $\tau(y_2) > \tau'(y_2)$ implies $\tau(y_3) > \tau'(y_3)$.

Proof: See the Appendix.

Theorem 4: Let k be a positive integer. Suppose that R implies strongly single crossing individual revenue requirements. Then for any draw in $\mathcal{A}_k$, the one stage voting game has a majority rule winner.

[9] See the precise statement in the first paragraph of Lemma 1.

Strongly single crossing is used in a strong way 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^*.

Definition: Let C_k^1 be the space of continuously differentiable functions (with domain $[\underline{w}, \overline{w}]^k$ and range $\mathbf{R}$) endowed with the uniform topology. We consider T_k and T_k^* as subsets of this space.

Proof: Fix k and let $(w_1, ..., w_k) \in \mathcal{A}_k$. 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. Using Ascoli's theorem (see Munkres (1975, *p.* 290)), $\bar{T}_k^*$ (the closure of T_k^* in C_k^1) is compact. Let τ^* be a maximal element of $\bar{T}_k^*$ using $v(\cdot, w^M)$ as the objective, where w^M is the median ability level in $(w_1, ..., w_k)$ 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 1, $\tau^* \in T_k^*$.

Now suppose there exists $\tau \in T_k$ such that there is a subset D of $\{w_1, ..., w_k\}$ 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 1, we can take τ to be in T_k^* without loss of generality. Using Lemma 3, 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 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.

Notice that the proof of Theorem 4 characterizes the set of majority rule winners 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.

4. Voting Over Taxes in an Endowment Economy

One can also interpret the results obtained for the optimal income tax economy as being necessarily valid for an endowment economy with a single good (income), as in Foley (1967).

it corresponds to the cost of providing the Pareto optimum level of public good for each draw. In the next section we examine an application of a similar idea to an optimal income tax economy.

As above, the theorems below can be simplified to results concerning lump-sum tax functions (individual revenue requirement functions) when considering an endowment economy. We shall discuss this further at the end of next section.

5. 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 varies with the draw of 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 $x \in \mathbf{R}_+$ be the quantity of the public good. [10] Let the cost function for the public good in terms of consumption good be $H(x)$, which is assumed to be C^2.

Let $F_k : \mathcal{A}_k \to T_k \times \mathbf{R}_+$ be a correspondence defined by $F_k((w_1, ..., w_k)) \equiv \{(\tau, x) \in T_k \times \mathbf{R}_+ \mid \sum_{i+1}^{k} \tau(y(w_i)) \geq H(x)\}$.

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

As we previously mentioned, we will 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.[11] Let utility be given by a C^2 function $u(c, \ell, x, w) =$

---

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

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

In this case, the type of each agent is defined to be the income of that agent, so in terms of the notation used in this paper w is an income (or wealth) level. A tax is simply an amount to be paid by each type, since there is no labor supply. Hence, a tax function is simply an individual revenue requirements function $g \in G_k$, a lump-sum tax schedule.

We need to allow the government to have the ability to monitor and tax individual income. Thus, in this case, the interpretation of the informational asymmetry assumption must differ slightly. In the optimal income tax economy the government is able to observe income but not the abilities of the agents. In an endowment economy we assume that the timing of events is such that the lump-sum income tax function g must be defined before the composition of the draw is known. Then the lump-sum incomes $(w_1, ..., w_k)$ are revealed and taxed according to schedule g. Assume that utility is monotone increasing in the commodity. Given this structure all the results established in Theorems 1–4 concerning the individual revenue requirement (which plays the role of a lump-sum tax function here), are directly applicable to the income tax function in an endowment economy with one commodity. In particular, properties of the individual revenue requirement function g do not have to be translated to an income tax τ.

Thus, under individual increasing returns with an unknown population, there is a unique feasible lump-sum tax function g. Under revenue complementarity and a known population k , the same result obtains. Under known population k and either limited complementarity and Edgeworth substitutability or argument-additivity and concavity, there exists a majority rule equilibrium lump-sum tax function.

The endowment economy is also a good starting point to inquire about the nature of the collective revenue requirement function. Even though there may be several rationales for $R(S)$, such as income redistribution or public provision of private goods, the case involving provision of a pure public good stands as a benchmark: it is a clear-cut case and it is often invoked as the primary reason for existence of a public sector. In this case, $R(S)$ is the expenditure on the public good provision to the draw S.

A natural option is to model the decision on the provision of the public good also as a collective choice by majority voting. Having simultaneous voting over tax structures and public goods provision is a difficult problem that we will not address here in its full generality. However, for endowment economies, there is an interesting case studied by Bergstrom and Cornes (1983), where the Pareto optimal levels of the public good are unique for each draw. In that case, the revenue requirement function is well-defined and

$ac + b(\ell, w) + r(x, w)$.[12] We assume that $\partial b/\partial \ell < 0$, $\partial^2 b/\partial \ell^2 < 0$, $\partial r/\partial x > 0$, $\partial^2 r/\partial x^2 < 0$. It is assumed that $dH(x)/dx > 0$ and $d^2 H(x)/dx^2 \geq 0$. Let $(w_1, ..., w_k) \in \mathcal{A}_k$, and let c_i and ℓ_i denote the consumption and labor supply of the i^{th} member of the draw respectively. Then production possibilities are given by:

$$\sum_{i=1}^{k} w_i \cdot \ell_i - \sum_{i=1}^{k} c_i \geq H(x). \qquad (2)$$

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 [13] is interior.

Lemma 4: Under the assumptions listed above, for any given draw $(w_1, ...w_k)$, 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 $u(c, \ell, x, w_1)$ subject to $u(c, \ell, x, w_i) \geq \overline{u}_i$ for $i = 2, ...k$ and subject to (2) 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:

$$\sum_{i=1}^{k} (1/a) \cdot \partial r(x, w_i)/\partial x = dH(x)/dx. \qquad (3)$$

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 (3).

Q.E.D.

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

When the marginal cost of production of the public good is constant, equation (3) can be solved explicitly for x, and the revenue requirement function can be found. A simple example is given by the cost function $H(x) = mx$ and preferences over the public good

---

12 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.

13 The allocation results when the agents in a draw each solve their consumer problem. See Bergstrom and Cornes (1982) for an explanation of why we need to restrict the analysis to interior allocations.

$r(x, w_i) = w_i(x - x^2)$, which (for interior solutions) generate the revenue requirement

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

This case constitutes an example satisfying the conditions of Lemma 2, namely concavity and argument-additivity, so we know that 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 which, given reasonable values for the parameters, verifies the conditions for existence of a majority rule equilibrium.

Theorem 5: Let $u(c, \ell, x, w) = ac - b(\ell, w) + \frac{rw}{1-\alpha} x^{1-\alpha}$, and let $H(x) \equiv mx^{\beta}$, with $\alpha, \beta > 0, \alpha + \beta > 1$. Then for any draw in $\mathcal{A}_k$, the one stage voting game over interior (τ, x) has a majority rule winner.

Proof: Using Lemma 4, the unique interior Pareto optimal level of x is given by

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

which implies

$$R((w_1, ..w_k)) = m[\frac{r}{am\beta} \sum_{i=1}^{k} w_i]^{\frac{\beta}{\alpha+\beta-1}}.$$

There are two basic cases to consider. If $\alpha \leq 1$, $R((w_1, ..., w_k))$ is convex. In Appendix III we show that convex argument-additive functions $R(\sum w_i)$ satisfy revenue complementarity. The result then follows from Theorem 3. If $\alpha > 1$, $R((w_1, ..., w_k))$ is concave and argument-additive. In this case the remainder of the proof follows from Lemma 2 and Theorem 4.

Q.E.D.

Another 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 in this particular case is simply to sum them.

Next we focus on two examples of interest.

Example 1. Let $u(c,\ell,x,w) = ac - \ell^2 + wln(x)$, and let $H(x) \equiv mx^2$. Then (3) becomes $[\sum_{i=1}^k w_i]/(ax) = 2mx$ or $x = [\sum_{i=1}^k w_i/(2ma)]^{1/2}$ and $R((w_1,...,w_k)) = [\sum_{i=1}^k w_i]/2a$. In this case, the revenue requirement function satisfies both individual increasing returns and revenue complementarity. By Theorem 2 (if draws of all sizes are considered) or Theorem 3 (if only draws of size k are considered), $g(w) = w/2a$ is the one feasible individual revenue requirement function. Using the results in Berliant and Gouveia (1991), the unique feasible tax function $\tau(y)$ is the solution to the differential equation $d\tau/dy = 1 - [y/(2a^3\tau^2)]$ that goes through the point $(\bar{y},\bar{\tau}) = (a\bar{w}/2, \bar{w}/2a)$.

Example 2. Let $u(c,\ell,x,w) = ac - \ell^2 - wx^{-2}/2$, and let $H(x) \equiv mx^2$. Then, by Lemma 4, we have $R((w_1,...,w_k)) = m[\frac{1}{2am}\sum_{i=1}^k w_i]^{1/2}$. Concavity and argument-additivity hold so, if we rule out bankruptcy problems, there is a majority rule equilibrium.

By Lemma 2, we have that $g(w;\tilde{w}) = (\frac{2m}{a})^{1/2}[(k\tilde{w})^{1/2} + 1/2(k\tilde{w})^{-1/2}(w - \tilde{w})]$.

Take $w \in [1,2]$. For notational simplicity define $\mu = (m/2a)^{1/2}$. Since we can actually index all admissible g's by their $\tilde{w}$'s, to find the choice of the median voter w^M we need only solve the problem $\min_{\tilde{w}} g(w^M;\tilde{w})$. The solution to this problem is obtained when $w^M = \tilde{w}$.

Suppose we have a draw where the median voter is the type $w = 1.5$. The majority winner tax function implements the individual revenue requirement function $g(w) = \mu[(1.5/k)^{1/2} + (1/2)(1.5k)^{1/2}(w - 1.5)]$.

Applying Theorem 1, the income tax function is given by the solution to:

$$\frac{d\tau}{dy} = 1 - \frac{2y}{w^2} = 1 - \frac{2y}{[2\tau/\mu - (1.5k)^{-1/2} - 2/k + (1.5)]^2},$$

with upper boundary at $(\bar{\tau},\bar{y}) = (\mu[(1.5)^{1/2}(k^{-1/2} + (1/4)k^{1/2}, a)$.

As mentioned in the last section, the results can be simplified to deal with the case of lump-sum taxation in a "Bergstrom-Cornes" economy, with a private good and a public good. In this case we do not need to restrict ourselves to utility functions that are quasi-linear in the public good. Instead Bergstrom and Cornes (1983) show that quasi-concave utility functions with the general form

$$u_i(c,x) = A(x)c_i + B(x,w_i) \text{ with } i = 1,...k,$$

have the desired property that the interior Pareto optimal levels of the public good are independent of the distribution of income[14].

[14] They also need to assume that the set of feasible allocations is convex.

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 given by (2). The revenue requirement function is the cost of providing this level of the public good. Even though conditions such as individual increasing returns, revenue complementarity or 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.

6. Conclusions

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[15]. All majority rule equilibria derived in this paper are 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. However, they are generally infeasible in this model. When they are feasible, they are first-best.

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, instead of a model expressly designed to study the problem of voting over taxes. However, their most interesting results still hold under our assumptions of limited complementarity and Edgeworth substitutability; Director's Law, which says that middle income classes use the political system to shift some of the tax burden to the tails of the income distribution, holds. On the other hand, an alternative set of assumptions including concavity and argument-additivity results in the existence of equilibria with characteristics opposite to Director's Law, namely non-increasing marginal tax rates.

[15] We know of only one case where an optimal tax is linear: Snyder and Kramer (1988). But as explained in the text this and other results are due to the use of a peculiar model that departs significantly from the 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.

Theorems 2 and 3, where no voting need actually occur, can also have a different interpretation, pointed out to us by Roy Gardner: the assumptions provide an axiomatic characterization of a unique tax system. Under this interpretation, the overall information structure of the model (the uncertainty about the composition of the economy in particular) has a role akin to the veil of ignorance often invoked when discussing the creation of rules at the constitutional stage (see Rawls (1971)).

With these results in hand, it will be interesting to 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)).

It would be interesting to see what the feasible set would look like if we were to require less than full feasibility for each draw. One would have to define the consequences of failing to meet the revenue requirement. Our guess is that this would simply result in a modified revenue requirement function.

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.

Appendix

1. Proof of Theorem 1.

Fix $\tau \in T_k$. If $y(w)$ is the gross income function associated with τ, then $\tilde{g}(w) \equiv \tau(y(w)) \in B$ 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 (1991), Propositions 1 – 3.

Now suppose there exists one w^* such that $dg/dw\,|_{w^*}$ or $d^2g/dw^2\,|_{w^*}$ 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 (1991), Propositions 1–3. Using the results in the proof of Proposition 1 in Berliant and Gouveia (1991), there is an extension of $\theta^*(w)$ (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 (1) 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.

2. An example with revenue complementarity.

Let $R((w_1, ..., w_k)) = z(\sum_{i=1}^{k} w_i)$. Revenue complementarity holds if and only if

$$z(\sum_{i=1}^{k} w_i) \leq \sum_{i=1}^{k} z(kw_i)/k.$$

But this is the same as

$$z(Q) \leq \sum_{i=1}^{k} z(q_i)/k \text{ with } Q = \sum_{i=1}^{k} \frac{1}{k} q_i \quad \text{and } q_i = kw_i,$$

which is just the statement that z is convex.

3. Proof of Lemma 1.

We present the proof for k even. Adaptation of the proof for the case when k is odd is straightforward.

Fix $g \in G_k^*$. By assumption, $R((w_1, ..., w_k)) \leq \sum_{i=1}^{k} [\ R((w_1^i, ..., w_k^i))\ 2/k - g(W(w^i, w^*))\ ]$. Since g is feasible, $g(w_i) \geq R((w_1^i, ..., w_k^i))2/k - g(W(w^i, w^*))$ for each i. If $g(w_i) > R((w_1^i, ..., w_k^i))2/k - g(W(w^i, w^*))$ for some i, then g is not minimal in the sense that $g \notin G_k^*$, which is a contradiction. Hence $g(w_i) = R((w_1^i, ..., w_k^i))2/k - g(W(w^i, w^*))$ and in particular $g(w^*) = R((w^*, ..., \underline{w}))2/k - g(\underline{w}) = R((w^*, ..., \overline{w}))2/k - g(\overline{w})$.[16] Hence g is continuous.

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 $\tilde{w}$ as one such case. Assume that $\tilde{w} \geq w^*$, $\tilde{w} \geq w'^*$. Then $g(\tilde{w}) - g'(\tilde{w}) = g'(\underline{w}) - g(\underline{w}) < 0$, a contradiction. Now assume that $\tilde{w} \leq w^*$, $\tilde{w} \leq w'^*$. Then $g(\tilde{w}) - g'(\tilde{w}) = g'(\overline{w}) - g(\overline{w}) > 0$, another contradiction. Hence, either $w^* > \tilde{w} > w'^*$ or the reverse must hold. Assume the former. Over (w'^*, w^*) we have:

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

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 $w'^* > \tilde{w} > w^*$. 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 $g(\underline{w}) > g'(\underline{w}) \Rightarrow w^* > w'^*$.

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

Q.E.D.

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

4. An example satisfying limited complementarity.

Define w^M as the median of $(w_1, ...w_k)$. We have that:

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

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

The adding-up restriction is as follows:

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

Using the definition of limited complementarity, the minimal individual revenue requirements are:

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

In particular we have that

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

Substituting for $g(\overline{w})$ we obtain $g(\underline{w}; w^*) = \beta\underline{w} + \alpha(w^* - \underline{w})$. Using the adding-up restriction we obtain $g(\overline{w}; w^*) = \beta\overline{w} + \alpha(\overline{w} - w^*)$.

We now check feasibility, which says

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

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

This can be simplified to

$$\sum_{i=1}^{k} |w_i - w^M| \leq \sum_{i=1}^{k} |w_1 - W(w^*, W_i)| - \sum_{i=1}^{k} |w^* - W(w^*, w_i)|.$$

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

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

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.

5. Proof of Lemma 2.

It is straightforward to prove that $\forall\ g(w;\tilde{w}) \in G_k^*$, $g(w;\tilde{w})$ is continuously differentiable in both w and $\tilde{w}$ and strictly increasing in w. [17] Since R is argument-additive $R((w_1,...,w_k)) = Q(\sum_{i=1}^k w_i) = Q(kw^A)$, where w^A is the average ability in the draw.

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

Now consider branch B and $w_1 \in [\tilde{w} + (k-1)(\tilde{w} - \underline{w}), \overline{w}]$. 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 $\tilde{w}$. Consider a draw with $w_j \in [\underline{w}, \tilde{w} + (k-1)(\tilde{w} - \underline{w})]$ for $j = 2,...k$. Due to argument-additivity, for any fixed draw mean w^A, we can take all w_j's to be equal to $\hat{w} = (kw^A - w_1)/(k-1)$, without loss of generality. Feasibility requires $g(w;\tilde{w}) + (k-1)g(\hat{w};\tilde{w}) \geq Q((k-1)\hat{w} + w)$. Take this as an equality and replace $g(\hat{w};\tilde{w})$ by $Q(k\tilde{w})(k-1)/k + (k-1)Q'(k\tilde{w})(\tilde{w} - \hat{w})$. By construction this revenue requirement is minimal. It is maximized over $\hat{w} \in [\underline{w}, \overline{w}]$ for $\hat{w} = \underline{w}$, so feasibility requires $g(w;\tilde{w}) = Q((k-1)\underline{w} + w) - (k-1)/kQ(k\tilde{w}) + (k-1)Q'(k\tilde{w})(\tilde{w} - \underline{w})$. It is easy to prove that allowing for draws with different compositions, namely more than one ability in the interval $[\tilde{w}+(k-1)(\tilde{w}-\underline{w}), \overline{w}]$, does not violate feasibility. We thus obtain branch B in the statement of the Lemma. Branch D is obtained following a similar reasoning.

[17] Proofs available from the authors on request.

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

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

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

For branch B we have:

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

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

Finally, for branch D we get:

$$\frac{\partial g(w;\tilde{w})}{\partial \tilde{w}} = -k(k-1)Q''(k\tilde{w})(\underline{w}-\tilde{w}) > 0,$$

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

These results also imply that $\text{argmin}_{\tilde{w}} g(w;\tilde{w}) = w$.

Q.E.D.

6. Proof of Lemma 3.

Let $y(\cdot)$ and $y'(\cdot)$ be the gross income functions associated with τ and τ', respectively. 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 $\tau(y_1) < \tau'(y_1)$ and $\tau(y_2) > \tau'(y_2)$ and $\tau(y_3) < \tau'(y_3)$. Then 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 $\tau(y(w^a)) < \tau'(y(w^a))$. 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 $\tau(y'(w^b)) > g'(w^b)$ and since $y(w^b) > y'(w^b)$, $g(w^b) > g'(w^b)$. Similarly, $g(w^a) < \tau'(y(w^a))$ and since $y'(w^a) > y(w^a)$, $g'(w^a) > g(w^a)$.

Using strongly single crossing, $g(\overline{w}) \geq g'(\overline{w})$. If $g(\overline{w}) = g'(\overline{w})$, then using $g'(w^a) > g(w^a)$ and the proofs of Lemma 1, g is infeasible. Hence $g(\overline{w}) > g'(\overline{w})$.

By normality of leisure and construction of T_k^*, $\tau(y(\overline{w})) > \tau'(y'(\overline{w}))$ and $y(\overline{w}) > y'(\overline{w})$. Since $\tau'(y(w^b)) > \tau(y(w^b))$, there exists $y^* > y(w^b)$ 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.

REFERENCES

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

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

Berliant, M., forthcoming, On income taxation and the core, Journal of Economic Theory.

Berliant, M. and M. Gouveia, 1991, Incentive compatible income taxation, individual revenue requirements and welfare, mimeo.

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

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

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

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

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

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).

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.

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

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

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

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

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

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

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

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

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

Rawls, J., A Theory of Justice (Harvard, Cambridge, Mass.).

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

Röell, A., 1984, Individually optimal income taxation, mimeo.

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

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

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

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

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

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

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.

Stigler, G. (1970), Director's law of income redistribution, Journal of Law and Economics 13, 1-10.