Incentive compatible income taxation, individual revenue requirements and welfare

Berliant, Marcus and Gouveia, Miguel

Washington University in St. Louis, Universidade Católica Portuguesa

1 October 2001

Incentive Compatible Income Taxation, Individual Revenue Requirements and Welfare[±]

Marcus Berliant* and Miguel Gouveia**

October 2001

---

*Department of Economics, Washington University, St. Louis, MO.

**Universidade Católica Portuguesa, Lisbon, Portugal.

± The authors would like to thank John H. Boyd III, Christophe Faugere, Tom Gresik, Patrick Legros, Wan-Hsiang Pan, Nirvikar Singh, Bill Vogt, John Weymark, and in particular two anonymous referees for corrections and comments. This research was funded by NSF grant SES-9007800. The authors retain responsibility for any errors. Address correspondence to the Department of Economics, Washington University, Campus Box 1208, 1 Brookings Drive, St. Louis, MO 63130-4899, U.S.A.

Abstract

This paper introduces the classical Public Finance concept of taxation according to ability to pay in the set-up of standard optimal income tax models. The fundamental concept used is the specification of an individual revenue requirement function, a mapping from abilities to taxes. The discussion is centered on the derivation of a tax function on income such that agents of a given ability pay exactly the amount specified by the revenue requirement function. The construction of the tax function is achieved by using the differentiable approach to the revelation principle. A basic differential equation is generated from which the tax function is found. A discussion of the necessary and sufficient conditions for the validity of this technique and an interpretation of the results in graphs are provided. A welfare ranking of the solutions is used to select the best tax function that implements the individual revenue requirements.

Journal of Economic Literature Classification Number: H21.

1. Introduction

Traditional discussions of income taxation are often based on the concept of "ability to pay", which has been one of the fundamental concepts used to study the horizontal and vertical equity characteristics of tax systems at both the theoretical and applied levels.[1] However, more recent developments in the normative theory of income taxation, namely the theory of optimal non-linear income taxes, have not used the concept at all. This may have misled some into thinking that ability to pay is an obsolete idea, incompatible with the emphasis that the modern theory of tax design places on incentives and efficiency.

This paper shows that is not the case. We provide an integration of the concept of ability to pay and the optimal income tax model and argue that the study of the politics and equity of income taxation can be pursued further with such an approach. In order to accomplish that desideratum we use basic techniques developed in the literature on incentives and mechanism design to obtain results in an optimal income tax framework.

The framework employed is essentially the one in Mirrlees (1971): there is a continuum of agents with the same preferences over consumption and labor but who differ by a parameter that can be interpreted as an index of ability, i.e. a wage rate. Given that the endowment of time is the same for all agents it is normalized to one without loss of generality. Thus the wage rate is also the market value of each agent's time endowment, a concept sometimes referred to as virtual income. The government tries to collect a given amount of revenue and knows the distribution of the ability parameter in the population but not each particular agent's identity. Income is defined to be the product of wage rate and labor supply. It is also assumed that the government is able to enforce whatever income tax function is chosen and agents can only adjust their labor supplies. However, the magnitudes of this adjustment are a crucial ingredient of the model. Thus, general preference structures should be allowed in order to accomodate general labor-leisure tradeoffs.

By preventing the possibility of a differentiated lump-sum tax, the information structure makes the optimal income tax problem interesting because it becomes conceptually closer to realistic situations. Unfortunately, this feature also makes the problem quite complex. This is a reason why the standard formulations of the problem, where the government

1 Cf. Musgrave (1959) and Pechman (1987).

maximizes a social welfare function, have relied almost entirely on the information provided by the first order conditions of the problem to extract qualitative results.

In contrast with the optimal income tax literature, the ability to pay approach does not necessarily coincide with the explicit maximization of a social welfare function. Instead it relies on the specification of an individual revenue requirement function, a mapping from abilities to taxes. Whereas in classical public finance "ability to pay" is defined as income (or wealth), in models with incentives where income is an endogenous variable, this role is played by the ability parameter.[2] Obviously, a revenue requirement function is implicit in the solution to any optimal income tax problem, but here it will be taken as a primitive concept. The discussion is then centered on the derivation of tax functions on income that implement a given revenue requirement. By this we mean that, in equilibrium, after all behavioral adjustments to the tax function have taken place, agents of a given ability pay exactly the amount specified by the revenue requirement function. The construction of such tax functions is achieved by using the differentiable approach to the revelation principle, as in Laffont and Maskin (1980): a basic differential equation is generated from which the tax functions are found. This paper provides a discussion of the necessary and sufficient conditions for the validity of this technique, using intensively the particular structure of the problem. These conditions have a simple interpretation and are easy to check for any preference specification. We provide a graphical interpretation of the results and discuss what can be expected to happen when such conditions are not met.

Since the solution is based on a differential equation, a welfare ranking of the many particular solutions is used to select the best tax function that implements the individual revenue requirements, a result inspired by Seade (1977).

The Mirrlees optimal tax model has two distinct components: 1) an economy with agents endowed with different productivities that are private information and 2) a government that maximizes a social welfare function. In this paper we retain the first component and replace the second component by the exogenous revenue requirement function. This approach gives us a good deal of flexibility in closing an income tax model: the determination of revenue requirement functions can be separated from incentive problems and considered

[2] The ability parameter, a wage rate, can also be interpreted as the value of the taxpayer's endowment of time.

as a problem with a life of its own, leaving room for politics or alternative ethical foundations. Obviously, social welfare functions are the standard way to close normative models and that is the way the literature has progressed. But there is certainly scope for other approaches. First, from a normative perspective, one can consider the case where revenue requirements are the ultimate consequence of adopting certain equity criteria to dictate tax policy.[3] This case includes applications of axiomatic equity theory to the income tax problem, where rather than focusing on optimizing the distribution of net income, tax design is defined as a cost-sharing problem. Our work here serves the function of translating tax problems with incentives to maps from abilities to tax liabilities. The latter can be compared directly using index numbers representing equity notions.

Second, we may want to take a positive rather than a normative perspective and use the Mirrlees framework to explain and predict economic behavior and institutions. It might seem more natural to examine policy choices at the revenue requirement level, given that it is usually at this level that the incidence and equity implications of tax design are discussed publicly. For example, recent debates on tax reform in the U.S. are explicitly stated in terms of how much people of each income class pay in taxes. Another example comes from the literature on the design of income maintenance programs, where non-utilitarian principles are seen as increasingly relevant.[4]

Since there is no reason to expect the political equilibrium to conform to the policies of a benevolent dictator, a more adequate formalization of a positive model for tax system formation may be to specify preferences and choices directly at the revenue requirement level. Indeed, a framework along these lines bears more resemblance to actual budgetary institutions in some countries, including the U.S., where taxing and spending decisions are made separately.

Formally, if we try to explain fiscal structures as the outcome of a political process, we find that revenue requirement functions are the natural objects where we should focus our attention. Specifying political economy models of income taxation as collective choices of revenue requirement functions simplifies the problem considerably in the sense that consideration of incentives becomes easy to handle, almost an afterthought. An immediate

[3] A detailed study of one such case is Berliant and Gouveia (1993).

[4] See, for example, Besley and Coate (1992).

example would be to use a model of direct democracy, where taxpayers vote over revenue requirement functions.[5] The case of a representative democracy can also be addressed by modeling a two stage process, where people choose among candidates and elected officials choose a revenue requirement function. In both these cases, income tax functions would then be the result of implementing the revenue requirement functions; the latter are chosen through a political process.

We wish to emphasize that, as suggested just above, one need not accept the "ability to pay" criterion (which might result in adverse welfare consequences) in order to use the approach advanced here. Instead, the approach proposed here can be viewed simply as a *tool* that can be used to separate the incentive compatibility aspects of an income tax question from its other considerations. This interpretation gives us another angle from which to analyze such questions. For example, it would be much more difficult to analyze positive questions concerning voting over income taxes without isolating the incentive compatibility part of the model. We employ the results presented here in other work, Berliant and Gouveia (1998), to accomplish this. Tax systems that are outcomes of such a voting scheme will be best for a given individual revenue requirement function, but not necessarily Pareto optimal for a given aggregate revenue requirement constraint. Similarly, the approach gives us another tool to analyze the validity of the first order approximation to the optimal income tax problem.

A less important but not irrelevant consideration is that the approach pursued here allows the computation of solutions for examples and easy computation of comparative statics. This is in contrast with the results obtained in the standard optimal income tax model employing a social welfare function, where simulations must be used in place of explicit closed form solutions and where comparative statics are complex.

The approach proposed here has an analogy in the optimal auction literature. The role of the individual revenue requirement function is quite similar to the role of the reduced form auction in this literature; see, for example, Border (1991) for a description as well as necessary and sufficient conditions on a reduced form auction for implementation. Here we inquire only about sufficient conditions on individual revenue requirements for implementation. Necessary conditions remain an open problem.

[5] For some work along these lines, see Berliant and Gouveia (1998).

The paper is organized as follows. The second section constructs the family of income tax functions that implement a smooth revenue requirement function and proves the existence of a best income tax function in that set. It also includes a new result on utilitarian and second best optimal tax functions. The third section contains an example with explicit solutions illustrating Propositions 1 and 2. A final section concludes.

2. Model and Results

2.1 Assumptions on Preferences and Technology

Consumers differ by an ability parameter, w, strictly positive, which can be interpreted as a wage rate or productivity. The support of w is W, an interval on the real line:

$$w \in W = [\underline{w}, \overline{w}] \subseteq R_{++}$$

where w has a population density function $f(w)$ and $f(w) > 0$ a.e. on W. The density function is common knowledge, but each agent's ability is private information. Thus, the only lump-sum taxes that can be used are necessarily uniform. Even in this case, such a tax must be bounded by the earning power of lowest ability individual in order to prevent bankruptcy.

Define C^p as the set of functions mapping from R^N to R that are p-times continuously differentiable. We assume that agents have preferences defined over non-negative values for consumption c and labor ℓ, represented by a C^2 utility function, $u(c, \ell)$. Without loss of generality we normalize the endowment of time to 1 and so we have that $\forall\ (c, \ell),\ (c, \ell) \in R_+ \times [0, 1]$.

In contrast with much of the optimal income tax literature, our utility function is purely ordinal. The arguments presented below are immune to continuous increasing transformations of $u(c, \ell)$. [6] Utility is employed in place of preferences purely as a matter of convenience.

[6] See, for example, equation (2).

Next, we present a list of assumptions on preferences often used in the optimal income tax model. Various *subsets* of these assumptions will be employed later on. Subscripts represent partial derivatives with respect to the appropriate arguments.

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:

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

A3- Consumption is normal:

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

A4- Interior solutions:

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

Restrictive as these assumptions may be, they allow for a considerably wider class of preferences than the quasi-linear or additively separable utility functions often used in the literature. The more restrictive preferences are also used in L'Ollivier and Rochet (1983) and papers following, such as Weymark (1987). We could weaken assumption A4, but at the cost of having more extensive proofs. The production technology side of the model is simple, in order to allow a sharper focus on the issues with which we are concerned. As is standard in the optimal income tax literature, we assume a production function with constant returns to scale in labor, with coefficient w for type w workers.

We define gross income as y and, when there are no taxes, we have that $y \equiv w \cdot \ell$ and $y = c$. Denote as Y the set of possible incomes, $Y = [0, \overline{w}]$.

Without taxes, the consumer's problem for type w is

$$\max_{y} u(y, y/w).$$

Given A4, we have interior solutions for this problem, satisfying the first order condition:

$$u_1(y, y/w) + u_2(y, y/w)/w = 0.$$

Denote the solution to the problem as $\overline{y}(w) \equiv \mathrm{argmax_y} u(y, y/w)$ and the indirect utility function as $v(w) \equiv u(\overline{y}(w), \overline{y}(w)/w)$. Notice that $\overline{y}(w)$ is well defined under A2, since the $\mathrm{argmax_y}$ is unique.

2.2 *The Revenue Requirement Function and Definition of the Problem*

In the basic model being discussed, the government wishes to impose a tax on individuals following the general principle of taxation according to ability to pay. This objective is formalized by the specification of a *revenue requirement function* g, mapping from abilities to tax payments. Denote $T \subset R$ as the set of possible tax payments. Then we can be more precise and write $g : W \to T$.

The concept of a revenue requirement function is a flexible way to formalize the tax payments implied by alternative equity concepts. It can also be interpreted as the object over which political debates on the distribution and redistribution of income are centered. Although it is cardinal, it has a natural scale in terms of numeraire.

We will require $g(w)$ to obey some feasibility and regularity conditions:

A5- Feasibility and regularity of $g(w)$:

$g(w)$ is C^2, $\forall\, w \in W$ it is assumed that $w > g(w)$, and

$$\int_{\underline{w}}^{\overline{w}} g(w)f(w)dw \geq 0.$$

These are not strong assumptions. Define G to be the space of all functions satisfying A5. Then the assumption that g is C^2 is generic in the appropriate topology on G; that is C^2 g's will uniformly approximate any continuous function (Hirsch (1976, Theorem 2.2)).[7] The specification of the budget constraint is standard in the optimal taxation literature.

[7] This justification is also used to explain differentiability assumptions in the smooth economies literature.

As is standard in that literature, we will not specify what is to be done with the revenue raised.

The government does not know each agent's identity, but we assume that it is able to monitor individual output, which can thus serve as a tax base. The *income tax function* $\tau : Y \to T$ specifies the tax paid by an individual with gross income y. The *net income function*, $\gamma(y)$, is defined as $\gamma(y) \equiv y - \tau(y)$.

Given that income will be taxed, the consumer's problem is now

$$\max_y u(\gamma(y), y/w),$$

with first order condition

$$u_1 \frac{d\gamma}{dy} + \frac{u_2}{w} = 0 \tag{1}$$

and with solution $y(w;\tau)$. [8]

We can now state the problem with which we are concerned: given that the government wants to impose taxes as described by the revenue requirement function $g(w)$, can we design an income tax function $\tau : Y \to T$ such that, in equilibrium, after all behavioral adjustments have taken place, agents of type w pay $g(w)$ in taxes? More formally, can we find a C^2 function τ that *implements* g in the sense that $\tau(y(w;\tau)) \equiv g(w)$?

To find such a tax, we use the differentiable approach to the revelation principle (see Laffont and Maskin (1980)). The direct mechanism is defined in the following way: agents are asked to report their type, the value of w. Given this reported value they are required to produce output $y(w) = \theta(w) + g(w)$ but can retain $\theta(w)$ for consumption. We find a family of C^2 functions $\theta(w)$ that satisfies incentive compatibility. In other words, these functions satisfy the first and second order conditions of the agent with respect to the type reported to the planner and furthermore do so for the case of truthful reporting.

We will define conditions under which the strategy chosen by the agents is unique, and consequently nothing is lost when we go from the indirect mechanism, taxation of income,

[8] In general $y(w;\tau)$ is a correspondence, but given our assumptions A1-A4 in equilibrium it will turn out to be a function.

to the direct mechanism just described. This is actually a trivial feature of the problem since the outcome for each worker's strategy (and not just his best strategy, as in the case of implementation in dominant strategies) is unaffected by other workers' strategies.

In contrast with the standard moral hazard problem, the optimal income tax problem is an adverse selection problem in the following sense. Each individual agent perceives no randomness in his individual decision problem since there is no randomness in the map from actions to consequences. The density function f represents the distribution of agents' types, while each agent knows her own type. This contrasts with moral-hazard problems where f might represent the conditional distribution of outcomes given an agent's action. A consequence of this is that the restrictions on density functions used in moral hazard environments to validate the first order approach are irrelevant to the adverse selection model studied here.

2.3 Results

It is helpful to consider briefly the first-best case, where perfect information allows the government to impose a differentiated lump-sum tax. It turns out that the solution to this simpler problem helps to understand the second order conditions for the second-best case and sheds light on a non-negativity condition we will require of $g(\underline{w})$.

The net income function for the first best case, ϕ, is defined implicitly by the first order condition of the agents :

$$u_1(\phi, \frac{\phi + g(w)}{w}) + u_2(\phi, \frac{\phi + g(w)}{w})/w = 0.$$

For the case with $g(w) = 0$, $\phi(w) = y(w)$. Given assumptions A1-A4, $y(w)$ exists and is unique. Assumptions A1-A5 also imply that the Jacobian of the equation above is not singular for strictly positive w. Thus, we can use the implicit function theorem to prove that $\phi(w)$ exists and is locally unique.

In (w, c) space (defined in the positive orthant of R^2) the net income function defines two sets:

$$\phi^- = \{(w, c) \mid u_1(c, \frac{c + g(w)}{w}) + u_2(c, \frac{c + g(w)}{w})/w < 0\} \quad \text{and}$$

$$\phi^+ = \{(w, c) \mid u_1(c, \frac{c + g(w)}{w}) + u_2(c, \frac{c + g(w)}{w})/w > 0\}.$$

(Fig 1 about here)

The last area, ϕ^+, is important in the proof of Proposition 1 that follows, as it corresponds to points where the second order conditions for incentive compatibility are verified. In particular, define the interval $I(\overline{w}) = (0, \phi(\overline{w}))$. We now show that $I(\overline{w})$ is not degenerate, or more generally that ϕ is always positive. Suppose that $\phi(\tilde{w}) = 0$ for some $\tilde{w} \in W$. But then by A4 we must have $u_1 = \infty$. Hence $u_2 = -\infty$ so $\ell = 1$ and $w = g(w)$. Using A5, we have a contradiction. We can now state our first result.

Proposition 1 : Assumptions A1-A5, $g(\underline{w}) \geq 0$ and $\frac{dg}{dw} > 0$ are sufficient conditions to implement $g(w)$ by means of an income tax.

The Appendix contains the proofs of all Propositions. However, it will be useful to mention here the basic steps involved. Using the first order condition of the agents, the government constructs net consumption functions $\theta(w)$ that induce truth telling from the agents. These functions are defined by the following differential equation:

$$\frac{d\theta}{dw} = -\frac{u_2(\theta, \frac{\theta+g(w)}{w})\frac{dg}{dw}/w}{u_1(\theta, \frac{\theta+g(w)}{w}) + u_2(\theta, \frac{\theta+g(w)}{w})/w} \tag{2}$$

A solution to (2) that satisfies monotonicity ($\theta'(w) > 0$) implements the revenue requirement $g(w)$.

Remark 1: The condition that $g(\underline{w}) \geq 0$ is needed only to ensure that the solution to the implementation problem implies non-negative labor supply. Given specific preferences and an arbitrary revenue requirement function assuming negative values for some range of W, we can solve (2) to obtain $\theta(w)$. As long as $\theta(w) + g(w) > 0$ for all $w \in W$ and A5 holds the remainder of our proof is still valid.[9]

[9] We are indebted to a referee for corrections and suggestions on this matter. Berliant and Page (1996) prove existence of solutions to the implementation problem for more general $g(w)$'s but their existence proofs are not constructive.

Remark 2: The monotonicity condition tells us that, in equilibrium, consumption and gross income must be an increasing function of ability. Proposition 1 is related to other results in the literature as can be seen in Myles (1995, Chap.5, sec. 2.4). However, one needs to keep in mind that Proposition 1 provides sufficient conditions on *exogenous* objects for implementation, while the remainder of the literature provides conditions on *endogenous* objects, like θ. There are some similarities with Brito and Oakland (1977) Proposition 4. However, the proof of the latter theorem relied on the optimality of the solution to a planner's maximization problem; our Proposition does not. The results in Proposition 1 are also related to the results in Theorem 2 of Guesnerie and Laffont (1984), where monotonicity is shown to be a necessary condition for implementability of "action profiles" (labor supply and consumption schedules). Papers dealing with monotonicity in the optimal income tax literature include L'Ollivier and Rochet(1983), Weymark (1986), and Ebert (1992). However, Guesnerie and Laffont are concerned with a principal-agent relationship where the welfare of the principal depends on the action profiles of the agents. In contrast, the problem we solve is in the tradition of classical Public Finance, where establishing a desirable mapping from abilities to taxes is a goal in itself. The principal or government cares not only about each agent's action, but also about her type, which is hidden knowledge. Our proposition also differs in that we discuss the entire set of implementable tax functions instead of just using them as tools (more precisely nonlinear price mechanisms) that induce patterns of behavior over which the principal optimizes. Moreover, the Guesnerie-Laffont result is implicitly restricted to additively separable utility functions since the objective (principal's utility) is assumed to be separable. Given the importance of the labor-leisure tradeoff this is a restrictive assumption.

The implications of the results derived above can be seen in a more intuitive manner if we illustrate the problem by depicting two possible situations in (y, c) space, using graphs developed in Sadka (1976). The reader is referred to that paper as well as Seade (1977) for a more detailed explanation of these graphs.

Normality of consumption (A3) yields the single crossing property for indifference curves of agents with different abilities. However, this alone is not enough to guarantee implementation. The second order conditions must be checked.

(Figure 2 about here)

In Figure 2 we have the case where the second order conditions for incentive compatibility are obeyed. Notice that when indifference curves of different agents cross, higher ability agents have smaller slopes. This is essentially what the single crossing condition says. In this situation, agents align themselves along $\gamma(y)$ in ascending ability order.

(Figure 3 about here)

In the case depicted in Figure 3, we have single crossing, but the consequence is a perverse situation where taxes decrease with ability. Agents line up in *descending* order of ability, i.e. high w agents choose to produce and consume less than low w agents. The perversity of the situation is its apparent "normality" because taxes are still an increasing function of income! This must be a serious concern to anyone using this type of framework. Suppose we were doing simulations of the model on a computer, and we had a program which solved the differential equation (2) and computed the corresponding $\tau(y)$. Just by looking at $\tau(y)$ we might never guess the solution was completely wrong.

In this last case the second order conditions for incentive compatibility do not hold. Consequently, critical points in the agents' Lagrangeans are local minima and hence useless for describing the actual behavior of the agents. Any $\tau(y)$ derived under these circumstances is meaningless.

Notice that $f(w)$ plays no role in Proposition 1.

For future reference, note that the inverse function theorem yields the following expression for the marginal tax rates:

$$\tau'(y) = \frac{g'(w)}{g'(w) + \theta'(w)}$$

So far we have proved that we can implement $g(w)$ by means of an income tax. This can be accomplished by solving the differential equation (2) and checking second order conditions. But we have not yet discussed the efficiency of the income tax thus found. Since Proposition 1 identifies a family of solutions, we can search for the best one. This is the content of our next proposition.

Proposition 2: Under assumptions A1-A5, if $g'(w) > 0$ and $g(\underline{w}) \geq 0$, then the C^2 income taxes implementing $g(w)$ are Pareto ranked, and there exists a best C^2 income tax under this ranking that implements g.

The results we have presented concern smooth tax functions. Do there exist other tax functions that may be superior on efficiency grounds and also implement g? The next proposition rules out such alternatives.

Proposition 3: Under the assumptions used in Proposition 2, the best C^2 income tax implementing g Pareto dominates *any* income tax implementing g.

One final remark is in order. Although we have found the best tax system implementing g, an important question remains open. Is this tax system (second best) Pareto optimal among tax systems satisfying the *aggregate* revenue constraint? To restate this question in a rigorous way, fix an aggregate revenue requirement R and consider the sets $G(R) = \{$ Continuously differentiable g with $\frac{dg}{dw} \geq 0 | \int_{\underline{w}}^{\overline{w}} g(w)f(w)dw \geq R\}$ and $\Theta(R) = \{\theta_g \mid g \in G(R)\}$, where θ_g is the net income function characterized in Proposition 2.

Is there any θ_g that is Pareto dominated in $\Theta(R)$? We conjecture that further conditions on both g and utility would be needed to answer the question in the negative. Here is some intuition for the conjecture. Suppose that $g, g' \in G(R)$ but that g' gives more tax liability to high ability agents than g. Since the incentive constraint is looser towards the top of the net income schedule induced by g as opposed to g', agents will generally work more under θ_g than under $\theta_{g'}$. It seems possible that although low ability agents pay more taxes and work harder under θ_g than under $\theta_{g'}$, they might also get more consumption good. Thus it seems to us that Pareto domination or lack thereof depends on the particular utility function employed. In our opinion, this method of attack might yield further characterizations of Pareto optimal income taxes (other than the classical zero top marginal rate), but that is a difficult subject outside the scope of the present paper.

However, we are able to offer one result that is stronger than the classical necessary condition for a second best tax and employs our individual revenue requirements framework.

Definions: Let $M = \{\tau : Y \to T| \quad \tau \text{ is measurable}\}$. Given aggregate revenue requirement $R \in \Re$, a tax function $\tau \in$ M is called **Utilitarian Optimal** if $\int_{\underline{w}}^{\overline{w}} \tau(y(w;\tau))df(w) \geq R$ and if there is no other tax function $\bar{\tau}$ such that $\int_{\underline{w}}^{\overline{w}} \bar{\tau}(y(w;\bar{\tau}))df(w) \geq R$ and $\int_{\underline{w}}^{\overline{w}} u(y(w;\bar{\tau}) - \bar{\tau}(y(w;\bar{\tau}), y(w;\bar{\tau})/w)df(w) > \int_{\underline{w}}^{\overline{w}} u(y(w;\tau) - \tau(y(w;\tau), y(w;\tau)/w)df(w)$.

Given aggregate revenue $R \in \Re$, a tax function $\tau \in M$ is called **Second Best** if $\int_{\underline{w}}^{\overline{w}} \tau(y(w;\tau))df(w) \geq R$ and if there is no other tax function $\bar{\tau} \in M$ such that $\int_{\underline{w}}^{\overline{w}} \bar{\tau}(y(w;\bar{\tau}))df(w) \geq R$ and $u(y(w;\bar{\tau}) - \bar{\tau}(y(w;\bar{\tau}), y(w;\bar{\tau})/w) \geq u(y(w;\tau) - \tau(y(w;\tau), y(w;\tau)/w)$ a.s., with strict inequality holding for a set of positive measure (in f).

Existence of utilitarian optima and second best taxes is studies in Berliant and Page (1999).

Theorem: Suppose that τ^* is a tax function such that the first order conditions for incentive compatibility (1) hold, let $y^*(w)$ be the gross income function associated with τ^*, and suppose that $\tau^*(y^*(w)) \in G(R)$, where R is fixed and positive. Let $g \in G(R)$, let $\epsilon \in \Re$,[10] and let $\tau(y, \epsilon)$ be any implementation of $(1 - \epsilon)\tau^*(y^*(W)) + \epsilon g(w) \in$ G(R); let $Y(w, \epsilon)$ be the gross income function associated with $\tau(y, \epsilon)$. Suppose that u is C^1, that $\tau(y, \epsilon)$ is continuosly differentiable in ϵ at $\epsilon = 0$ over all $y \in [0, \overline{w}]$, and suppose that $Y(w, \epsilon)$ is continuously differentiable in ϵ at $\epsilon = 0$ over all $w \in [\underline{w}, \overline{w}]$. If τ^* is utilitarian optimal, then $\int_{\underline{w}}^{\overline{w}} u_1(y^*(w) - \tau^*(y^*(w)), y^*(w))/w) \cdot \partial\tau(y^*(w), \epsilon)/\partial\epsilon|_{\epsilon=0} \quad df(w) = 0$. If τ^* is second best and $u_1 > 0$, then either (A) there is a set of positive measure such that $\partial\tau(y^*(w), \epsilon)/\partial\epsilon|_{\epsilon=0} > 0$ and another set of positive measure such that $\partial\tau(y^*(w), \epsilon)/\partial\epsilon|_{\epsilon=0} < 0$, or (B) $\partial\tau(y^*(w), \epsilon)/\partial\epsilon|_{\epsilon=0} = 0$ a.s. in w.

Remarks: At a utilitarian optimal tax, a small movement toward any other individual

[10] Note that ϵ can be positive, negative, or zero

revenue requirement function satisfying the aggregate revenue constraint will balance gains and losses in marginal utilities across consumers. This is the analog of the idea that a first best utiliitarian tax equates the marginal utility of income across consumers, but it applies to second best utilitarian taxes.

At a second best tax, if a small move is made toward any other individual revenue requirement function satisfying the aggregate revenue constraint, then income taxes rise on a set of incomes associated with a set of consumers of positive measure at the second best tax, and taxes fall on another set of incomes associated with a set of consumers of positive measure. Of course, utilitarian optimal taxes are a special case of second best taxes.

Consider a utilitarian optimal tax for the population density f, and suppose that it is also utilitarian optimal for the population density $\bar{f}$. Then, for any $\lambda, \bar{\lambda} \in \Re$ with $\lambda f(w) + \bar{\lambda}\bar{f}(w) \geq 0$ a.s., $\int_{\underline{w}}^{\overline{w}} u_1(y^*(w) - \tau^*(y^*(w)), y^*(w))/w) \cdot \partial\tau(y^*(w), \epsilon)/\partial\epsilon|_{\epsilon=0} \quad d(\lambda f(w) + \bar{\lambda}\bar{f}(w)) = 0$. So if these first order conditions for a utilitarian optimal tax are also sufficient, [11] *then this tax is utilitarian optimal for any non-negative linear combination of the population densities* for the appropriate linear combination of aggregate revenue requirements. [12] An analogous statement applies for non-negative linear combinations of populations for second best taxes.

Finnaly, if u is quasi-linear so that u_1 is constant, then the necessary condition for utilitarian optimal taxes reduces to $\int_{\underline{w}}^{\overline{w}} \partial\tau(y^*(w), \epsilon)/\partial\epsilon|_{\epsilon=0} \quad df(w) = 0$.

3. An Explicit Example

This section contains an example illustrating Propositions 1 and 2. It provides a closed form interior solution to the problem of implementing a revenue requirement function. We assume agents have Cobb-Douglas preferences, $u(c, \ell) = log(c) + log(1 - \ell)$ and $W = [1, 2]$.

We wish to implement proportional taxes on the *endowments* of the agents, $g(w) = kw$, with $k \in (0, 1)$. The first-best ϕ is defined by

$$\max_{\phi} \quad log(\phi) + log(1 - \frac{\phi + g}{w}) \quad \text{with solution} \quad \phi(w) = \frac{w - g(w)}{2} = \frac{w(1-k)}{2}.$$

[11] In general, Frechet or Gateaux derivatives can be used to explore first and second order conditions in the context of optimal taxation. However, to reduce the mathematical complexity of this paper and of this result in particular, a parametrization is used. Second best conditions are certainly of interest, but are beyond the scope of this paper.

[12] A special case of interest is $\bar{\lambda}$=0,$\lambda \neq$1.

This equation defines the area over which the second-best net income functions must be constructed. The second-best net income function is obtained following the steps in the proof of Proposition 1. Agent w sends the message w' that

$$\max_{w'} \quad log(\theta(w')) + log(1 - \frac{\theta(w') + g(w')}{w}).$$

Truth telling requires

$$\frac{du}{dw'}|_{w'=w} = 0 \iff \frac{\frac{d\theta}{dw}}{\theta} - \frac{(\frac{d\theta}{dw} + \frac{dg}{dw})/w}{1 - \frac{\theta+g}{w}} = 0.$$

We obtain the differential equation for the net income functions that implement g:

$$\frac{d\theta}{dw} = \frac{\frac{dg}{dw}\theta}{w - g - 2\theta} = \frac{k\theta}{w(1-k) - 2\theta}.$$

This is an homogeneous differential equation. The solution is:

$$w = \frac{2\theta + b\theta^{k^{-1}-1}}{1-2k} \text{ for } k \neq .5 \quad \text{or} \quad w = 4\theta(b - log(\theta)) \text{ for } k = .5,$$

where b is the constant of integration. By picking points in $(0, \phi(2))$ we determine the value of b that defines a particular net income function implementing $g = kw$. A particularly simple result is obtained for $k < .5$. In this case if we set b to equal zero, we have $\theta(w) = (1-2k)w/2$ and $y(w) = (1-2k)w/2 + kw$ or $w = 2y$. This implies $\tau(y) = 2ky$, i.e. a proportional income tax implements proportional taxation of the endowment of time. Note that this simple result does not hold for $k \geq .5$, since the solutions going through $I(2) = (0, 1-k)$ require a non-zero constant of integration. We can now illustrate Proposition 2 and compute the best tax function that implements $g = kw$. The net income function corresponding to the best tax is the particular solution that goes through $(2, \phi(2)) = (2, 1-k)$. In order to compute an exact solution assume that $k = 1/3$.

The net income function is of the form:

$$\theta = -1/b + (9 + 3bw)^{.5}/3b.$$

Solve for b such that the function passes through $(2, 2/3)$ and has a positive derivative to obtain (for b= -1.5)

$$\theta = \frac{2 - 2(1 - w/2)^{.5}}{3} \text{ and } y = \frac{(2+w) - 2(1-w/2)^{.5}}{3}.$$

Inverting for w we find the best tax function:

$$\tau = y - 1 + (1 - 2y/3)^{.5}.$$

This tax function is concave (regressive) and has a zero marginal tax rate at the highest income level $y(2) = 4/3$. It differs quite substantially from the proportional tax function $\tau = 2y/3$ which generates exactly the same revenue from every agent but implies a higher level of deadweight loss. Figures 4 and 5 illustrate this example. In Figure 4, income is on the horizontal axis while tax is on the vertical axis. The proportional tax is the darker line while the optimal tax is the lighter line. Note that the optimal tax function lies below the proportional tax function everywhere. In Figure 5, income is on the horizontal axis while marginal tax rates are on the vertical axis. Figure 5 gives the derivatives of the tax functions in Figure 4. The solid line is the marginal tax function for the proportional tax, while the dashed line is the marginal tax function for the optimal tax. Notice that marginal tax rates decrease with income for the optimal income tax, so it is regressive.

(Figure 4 about here)

(Figure 5 about here)

4. Conclusions

Sufficient conditions for the implementation of well-behaved revenue requirement functions have been found. This allows us to ignore incentives when modeling the determination of the distribution of tax burdens. We can simply use revenue requirement functions to generate incentive compatible income tax functions.

An implication of this result is that modeling strategies of the politics of income taxation can emphasize revenue requirement constraints instead of the maximization of social welfare functions.

Future work should examine the Pareto optimality of the best tax function relative to an individual revenue requirement and its welfare properties relative to the aggregate revenue constraint. Solving this problem is equivalent to finding more necessary conditions for an optimal tax, a very difficult unsolved problem. Our techniques may turn out to be useful for dealing with such a problem. Other topics deserving additional work are 1) the extension of this approach to environments where agents differ not only in their productivity but possibly in other dimensions as well, and 2) an examination of the role of heterogeneity in determining the progressivity of the tax schedules implementing a given revenue requirement function, for instance in our example.

Appendix

Proof of Proposition 1:

The problem solved by an agent of type w is

$$\max_{w'} u(\theta(w'), \frac{\theta(w') + g(w')}{w}). \tag{3}$$

The first order condition for incentive compatibility is:

$$u_1 \frac{d\theta}{dw'} + \frac{u_2}{w}(\frac{d\theta}{dw'} + \frac{dg}{dw'}) = 0.$$

Truthful revelation requires $\theta(w)$ to be constructed in such way that the following identity in w holds:

$$u_1(\theta(w), (\theta(w) + g(w))/w)\frac{d\theta}{dw} + u_2(\theta(w), (\theta(w) + g(w))/w)\frac{1}{w}(\frac{d\theta}{dw} + \frac{dg}{dw}) \equiv 0. \tag{4}$$

We now find $\theta(w)$.[13]

It is immediate that we can rewrite (4) as:

$$\frac{d\theta}{dw} = -\frac{u_2(\theta, \frac{\theta+g(w)}{w})\frac{dg}{dw}/w}{u_1(\theta, \frac{\theta+g(w)}{w}) + u_2(\theta, \frac{\theta+g(w)}{w})/w} \tag{2}$$

Equation (2) is well defined on both sides of the function $\phi(w)$ because assumptions A1-A4 guarantee that this first order differential equation has a unique continuous solution through any (w, θ) pair in ϕ^+ as well as ϕ^-.[14]

Recall that $I(\overline{w})$ is a nonempty open interval. Pick $\theta(\overline{w}) \in I(\overline{w})$, and let $\theta(\cdot)$ be the solution to (2) through $(\overline{w}, \theta(\overline{w}))$. Next we show that $\phi(w) > \theta(w) > 0, \ \forall\, w \in [\underline{w}, \overline{w}]$.

Clearly $\phi(\overline{w}) > \theta(\overline{w}) > 0$.[15]

Let w^* be the largest w such that $\phi(w) \leq \theta(w)$. Take $\{w_n\}_{n=1}^{\infty}$ with $\lim_{n\to\infty} w_n = w^*$, $w_n > w^*\ \forall\, n$. Then from the definition of ϕ and (2), $\lim_{n\to\infty} \frac{d\theta}{dw}|_{w_n} = \infty$. Now from the definition of ϕ, $\frac{d\phi}{dw}$ is bounded (at w^*), so $\exists\ w' > w^*$ with $\theta(w') \geq \phi(w')$, a contradiction, so $\theta(w) < \phi(w)$.

[13] It is possible to solve the problem using an indirect mechanism as in Mirrlees (1971), where he uses a differential equation for utility instead of consumption. The approach used here seems to be easier to handle.

[14] See Brock and Malliaris (1989), Theorem 5.1.

[15] Note that the denominator of (2) is zero when $\theta(w)$=$\phi(w)$ and that u_1 is undefined when $\theta(w)$=0.

Using A1 and A4, $\theta(w)=0$ implies by (2) that $\frac{d\theta}{dw}=0$, so $\theta(w')\geq 0\ \forall\ w'\in W$. If $\exists w"$ such that $\theta(w")=0$ and $g(w")<w"$ then $\theta(w)=0$ for $w>w"$, contradicting $\theta(\overline{w})>0$. If $g(w")\geq w"$ we contradict A5.

Hence $\phi(w)>\theta(w)>0\ \ \forall\ w\in[\underline{w},\overline{w}]$.

Since $g(\underline{w})\geq 0$, this result is sufficient to obtain the positivity of income, $\theta(w)+g(w)$.

Since θ never intersects ϕ, $\theta(w)\in\phi^+,\ \forall\ w\in[\underline{w},\overline{w}]$.

Our assumptions on preferences also imply that $\theta(w)$ is C^2 since the right hand side of (2) is C^1.

The second order condition for incentive compatibility is:

$$D\equiv u_{11}(\frac{d\theta}{dw'})^2+2\frac{u_{21}}{w}\frac{d\theta}{dw'}(\frac{d\theta}{dw'}+\frac{dg}{dw'})+\frac{u_{22}}{w^2}(\frac{d\theta}{dw'}+\frac{dg}{dw'})^2$$
$$+u_1\frac{d^2\theta}{dw'^2}+\frac{u_2}{w}(\frac{d^2\theta}{dw'^2}+\frac{d^2g}{dw'^2})<0. \tag{5}$$

If this condition holds everywhere, we have strict concavity of the agents' utility over w', guaranteeing uniqueness of a solution to each agent's maximization problem.

We can differentiate (4) with respect to w and obtain

$$D-u_{21}\frac{(\theta+g)}{w^2}\frac{d\theta}{dw}-\frac{u_2}{w^2}(\frac{d\theta}{dw}+\frac{dg}{dw})-u_{22}\frac{(\theta+g)}{w^3}(\frac{d\theta}{dw}+\frac{dg}{dw})\equiv 0. \tag{6}$$

Combining the equations (4), (5) and (6), and using primes to denote derivatives in order to simplify the notation, we write the second order condition for incentive compatibility as

$$\frac{\theta'(w)+g'(w)}{w^2}[\frac{\theta(w)+g(w)}{w}(u_{22}-\frac{u_{12}u_2}{u_1})+u_2]<0.$$

Using assumptions A1 and A3, the expression in brackets above is always negative [16], so the second order condition is reduced to $\theta'(w)+g'(w)>0$.

Since $\theta(w)\in\phi^+,\ \forall\ w\in[\underline{w},\overline{w}]$, we have

$$u_1(\theta,\frac{\theta+g(w)}{w})+u_2(\theta,\frac{\theta+g(w)}{w})/w>0. \tag{7}$$

With $g'(w)>0$ and A1, we have that (2) implies $\theta'(w)>0$. Since $g'(w)>0$, the second order condition holds.[17]

[16] The negativity of the expression in brackets coincides with Assumption B in Mirrlees (1971).

[17] We could also study the implementation of revenue requirements with $g'(w)<0$ in ϕ^- even though they are not very appealing as ability to pay criteria. However, they pose an additional problem, since the second order condition would require $\theta'(w)>-g'(w)$, and it is difficult to find *apriori* conditions ensuring that outcome.

Finally, recall that we have $y(w) \equiv \theta(w) + g(w)$. By our previous results we have that $y'(w) > 0$. Thus, we can invert it to obtain $w = \eta(y)$. The desired tax function is then

$$\tau(y) = g(\eta(y)).// \tag{8}$$

<u>Proof of Proposition 2:</u> Under the assumptions, every income tax implementing g has an associated net income function satisfying (2). Since all are solutions to the same differential equation, they do not intersect. Hence, in ϕ^+, they are Pareto ranked as for each w and each net income level c, less labor is required to obtain the same consumption c using a schedule ranked higher. [18] and so are the income taxes implementing g. The first-best net income function ϕ is an upper bound on admissible solutions to (2).

Let $\{\theta_n(\cdot)\}_{n=1}^{\infty}$ be a sequence of admissible solutions to (2) that are increasing to a maximal function in the Pareto ranking as $n \to \infty$. Hence $\forall\ n$, $\theta_{n+1}(w) \geq \theta_n(w)$, $\forall\ w \in W$.

Fix $w \in [\underline{w}, \overline{w})$. Since $\theta_n(w) \leq \phi(w)\ \forall\ n$, there is a pointwise limit $\theta(w) \equiv \lim_{n\to\infty} \theta_n(w)$. Hence $\theta(w) \leq \phi(w)$. If $\theta(w) = \phi(w)$, then

$$u_1(\theta, \frac{\theta + g(w)}{w}) + u_2(\theta, \frac{\theta + g(w)}{w})/w = 0,$$

and there is a neighborhood Z of $(w, \theta(w))$ in $R^2 \cap (\phi^-)^c$ and $\epsilon > 0$ such that

$$-\frac{u_2(\theta^*, \frac{\theta^*+g(w^*)}{w^*})\frac{dg(w^*)}{dw}/w*}{u_1(\theta^*, \frac{\theta^*+g(w^*)}{w^*}) + u_2(\theta^*, \frac{\theta^*+g(w^*)}{w^*})/w^*} > \frac{d\phi}{dw}|_{w'} + \epsilon,$$

$\forall\ (w^*, \theta^*) \in Z,\ \forall\ w'$ with $(w', \phi(w')) \in Z$. Choosing n large, it follows that there exists $\hat{w}$ with $\theta_n(\hat{w}) \geq \phi(\hat{w})$, a contradiction. So $\theta < \phi\ \forall\ w \in [\underline{w}, \overline{w})$. Since (2) holds for each θ_n on $[\underline{w}, \overline{w})$, it also holds for θ on $[\underline{w}, \overline{w})$. So θ is smooth on $[\underline{w}, \overline{w})$.

Implicitly, we have taken a sequence θ_n of solutions to (2) passing through an increasing sequence of points in $I(\overline{w})$. In the end, $\theta(\overline{w}) = \phi(\overline{w})$, so equation (2) is undefined at $\overline{w}$. However, θ can be defined at $\overline{w}$ by $\theta(\overline{w}) \equiv \lim_{w\to\overline{w}} \theta(w)$. Using the same arguments as in the proof of Proposition 1, $\theta(w)$ satisfies the second order conditions for incentive compatibility,

[18] Riley (1979) studies a different problem where a set of solutions to a fundamental differential equation is also Pareto ranked.

and the associated income tax $\tau(y)$ is incentive compatible. Moreover, since $\theta(w) \geq \hat{\theta}(w)$ for all admissible solutions [19] $\hat{\theta}(w)$, $\theta(w)$ is Pareto optimal among those C^2 net income functions implementing g.//

Remark : The initial restriction of the arguments in the proof to $[\underline{w}, \overline{w})$ is necessary because the upper limit of the Pareto optimal net income schedule is a singularity point of the differential equation (2). Intuitively, we push θ as far up as we can, until the top person's utility no longer increases. This happens at a point where the denominator of (2) vanishes (and $\frac{d\theta(w)}{dw}$ approaches infinity), which also implies that the marginal tax rate is zero at that point. [20]

Proof of Proposition 3: Let τ be any income tax implementing g, and let $\gamma(y)$ be its associated net income function as function of income. We will prove first that, without loss of generality, γ can be taken to be C^1.

It is easy to see that γ must be continuous on the relevant incomes $Y = \{y \mid \exists \; w \text{ with } y = \arg\max_{y'} u(\gamma(y'), y'/w)\}$, for otherwise incentive compatibility is violated. Fix y and define L and R to be the sets of left and right derivatives of γ at y:

$$L \equiv \{q \mid q = \lim_{y_n \to y} \frac{\gamma(y) - \gamma(y_n)}{y - y_n} \quad \text{where} \quad y_n \leq y \; \forall \, n\}$$

$$R \equiv \{r \mid r = \lim_{y_n \to y} \frac{\gamma(y) - \gamma(y_n)}{y - y_n} \quad \text{where} \quad y_n \geq y \; \forall \, n\}$$

If $\exists \; q \in L$, $r \in R$ with $q > r$, then there is a non-degenerate interval in W where all produce the same gross income y. Thus, $y - \gamma(y)$ is the same for all w in this interval, which contradicts the assumption that γ implements g, since $g'(w) > 0$. Hence for all $q \in L, r \in R, r \geq q$. If $\exists \; q \in L, r \in R$ with $r > q$, then γ is convex in a neighborhood around y. In addition, there is an open set A contained in this neighborhood, $y \in A$, such that no consumer produces gross income in A. In other words, A is a gap. The last fact follows since u is C^2 and quasi-concave. Then, without loss of generality (and without

[19] Those solutions satisfying the first and second order conditions.

[20] See the Appendix of Seade (1977) for a proof of this last statement in another context.

changing the consumption and labor supply of any consumer), γ can be made C^2 on A such that it is unchanged outside of A.

Thus, without loss of generality, γ is differentiable. Notice that the first order condition for incentive compatibility of γ implies

$$y = \arg\max_{y'} \quad u(\gamma(y'), y'/w)\} \quad \text{or}$$

$$u_1\gamma' + u_2/w = 0 \quad \text{or}$$

$$\gamma'(y) \equiv -\frac{u_2}{u_1 w}.$$

Hence γ is C^1 and, in fact, is C^2 since the last equation is a functional identity. The result follows from Proposition 2.//

Proof of Theorem:

Fix $w \in [\underline{w}, \overline{w}]$.

Now $\partial u(Y(w,\epsilon) - \tau(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0}$

$= \{u_1(Y(w,\epsilon) - \tau(w,\epsilon),\epsilon), Y(w,\epsilon)/w)[\partial Y(w,\epsilon)/\partial\epsilon - ((\partial\tau(Y(w,\epsilon),\epsilon)/\partial Y)(\partial Y(w,\epsilon)/\partial\epsilon) - \partial\tau(Y(w,\epsilon),\epsilon)/\partial\epsilon] + u_2(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)[\partial Y(w,\epsilon)/\partial\epsilon]/w\}|_{\epsilon=0}$.

Using equation (1) and $\gamma(y) = y - \tau(y)$,

$= \{u_1(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)[(\partial Y(w,\epsilon)/\partial\epsilon)(\partial\tau(Y(w,\epsilon),\epsilon)/\partial Y) - \partial\tau(Y(w,\epsilon),\epsilon)/\partial Y)(\partial Y(w,\epsilon)/\partial\epsilon) - \partial\tau(Y(w,\epsilon),\epsilon)/\partial\epsilon]\}|_{\epsilon=0}$.

$= [-u_1(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)\partial\tau(Y(w,\epsilon),\epsilon)\partial\epsilon]|_{\epsilon=0}$.

Now suppose that τ^* is utilitarian optimal. Then $\partial[\int_{\underline{w}}^{\overline{w}} u(y(w;\tau) - \tau(y(w;\tau), y(w;\tau)/w)df(w)]/\partial\epsilon|_{\epsilon=0}$. By Goffman (1965, Theorem 19), we can interchange the integral and derivative operators, so $[\int_{\underline{w}}^{\overline{w}} \partial u(y(w;\tau) - \tau(y(w;\tau), y(w;\tau)/w)/\partial\epsilon|_{\epsilon=0} \quad df(w)] = 0$. Plugging in the equalities of the last paragraph, $\int_{\underline{w}}^{\overline{w}} u_1(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)\partial\tau(Y(w,\epsilon),\epsilon)\partial\epsilon]|_{\epsilon=0} \quad df(w) = 0$, or $\int_{\underline{w}}^{\overline{w}} u_1(y^*(w) - \tau^*(y^*(w), y^*(w)/w)\partial\tau^*(y^*(w))\partial\epsilon]|_{\epsilon=0} \quad df(w) = 0$.

Now suppose that τ^* is second best. Since $u_1 > 0$, the string of equalities above implies that $\partial u(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0}$ if and only if

$\partial\tau(Y(w,\epsilon),\epsilon)/\partial\epsilon|_{\epsilon=0} \equiv d\tau(y^*(w),\epsilon)/\partial\epsilon|_{\epsilon=0} > 0$;

$\partial u(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0} \leq 0$ if and only if $\partial\tau(Y(w,\epsilon),\epsilon)/\partial\epsilon)|_{\epsilon=0} \equiv d\tau(y^*(w),\epsilon)/\partial\epsilon|_{\epsilon=0} \geq 0$.

Thus, if τ^* is second best it is *not* the case that

$\partial u(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0} \leq 0$ a.s. with

$\partial u(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0} > 0$ for a set of positive measure. In other words, there is a set of positive measure such that

$\partial u(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0} < 0$, or

$\partial u(Y(w,\epsilon) - \tau(Y(w,\epsilon),\epsilon), Y(w,\epsilon)/w)/\partial\epsilon|_{\epsilon=0} \leq 0$, a.s. in w. Thus, there is a set of positive measure such that $d\tau(y^*(w),\epsilon)/d\epsilon|_{\epsilon=0} > 0$, or $d\tau(y^*(w),\epsilon)/d\epsilon|_{\epsilon=0} \geq 0$, a.s. in w.

Since ϵ can also be negative (and it is assumed that the derivative at $\epsilon = 0$ is well-defined), there is also a set of positive measure such that $d\tau^*(y),\epsilon)/\epsilon)/d\epsilon|_{\epsilon=0} < 0$, or $d\tau(y^*(w),\epsilon)/d\epsilon|_{\epsilon=0} \leq 0$, a.s. in w, then $d\tau(y^*(w),\epsilon)/d\epsilon|_{\epsilon=0} = 0$, a.s. in w.

Q.E.D.

REFERENCES

[1] **Berliant, M. (1992),** "On Income Taxation and the Core", *Journal of Economic Theory,* **56**, 121-141.

[2] **Berliant, M. and M. Gouveia (1993),** "Equal Sacrifice and Incentive Compatible Income Taxation", *Journal of Public Economics,* **51**, 219-240.

[3] **Berliant, M. and M. Gouveia (1998),** "On the Political Economy of Income Taxation", *mimeo,* Washington University.

[4] **Berliant, M. and F. Page (1999),** "Income Taxes and the Provision of Public Goods: Optima and Balanced Government Budgets", *mimeo,* Washington University.

[5] **Berliant, M. and F. Page (1996),** "Incentives and Income Taxation: The Implementation of Individual Revenue Requirement Functions", *Ricerche Economiche,* **50**, 389-400.

[6] **Besley, T. and S. Coate (1992),** "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs", *American Economic Review,* **82**, 249-261.

[7] **Border, K. (1991),** "Implementation of Reduced Form Auctions: A Geometric Approach", *Econometrica,* **59**, 1175-1187.

[8] **Brito, D. and W. Oakland (1977),** "Some Properties of the Optimal Income Tax", *International Economic Review,* **18**, 407-423.

[9] **Brock, W. A. and A.G. Malliaris (1989),** *Differential Equations, Stability and Chaos in Dynamic Economics* (Amsterdam : North-Holland).

[10] **Ebert, U. (1992),** "A Reexamination of the Optimal Nonlinear Income Tax", *Journal of Public Economics,* **49**, 47-73.

[11] **Goffman, C. (1965),** *Calculus of Several Variables* (New York: Harper and Row).

[12] **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.

[13] **Hirsch, M. W. (1976),** *Differential Topology* (New York: Springer-Verlag).

[14] **Laffont, J.J. and E. Maskin (1980),** "A Differential Approach to Dominant Strategy Mechanisms", *Econometrica,* **48**, 1507-1520.

[15] **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.

[16] **Mirrlees, J. (1971),** "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies,* **38**, 175-208.

[17] **Musgrave, R. (1959),** *The Theory of Public Finance* (New York : McGraw Hill).

[18] **Myles, G. (1995),** *Public Economics* (Cambridge, UK: University Press).

[19] **Pechman, J. (1987),** *Federal Tax Policy* (Washington D.C. : The Brookings Institution).

[20] **Riley, J. (1979),** "Informational Equilibrium", *Econometrica,* **47**, 331-359.

[21] **Sadka, E. (1976),** "On Income Distribution Incentive Effects and Optimal Income Taxation", *Review of Economic Studies ,* **43**, 261-268.

[22] **Seade, J. (1977) ,** "On the Shape of Optimal Tax Schedules ", *Journal of Public Economics,* **7**, 203-236.

[23] **Weymark, J. A. (1986),** "Bunching Properties of Optimal Nonlinear Income Taxes", *Social Choice and Welfare,* **3**, 213-232.

[24] **Weymark, J.P. (1987),** "Comparative Statics Properties of Optimal Nonlinear Income Taxes", *Econometrica,* **32**, 203-214.

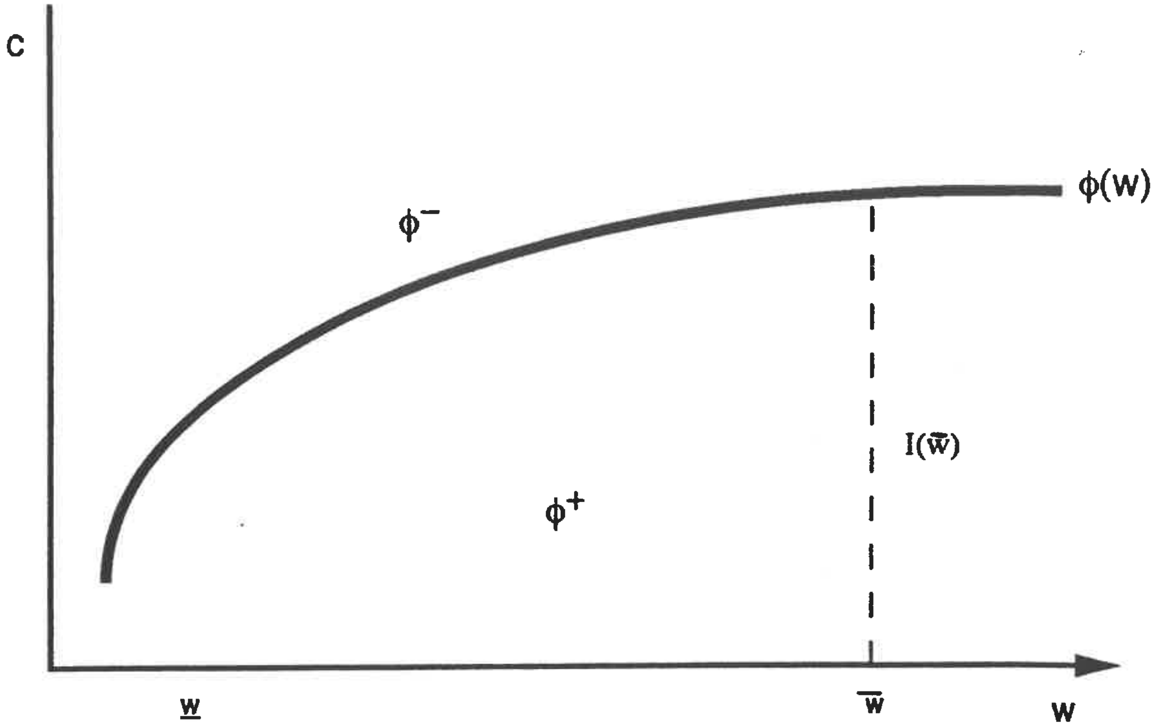

Figure 1: The First-Best Net Income Function

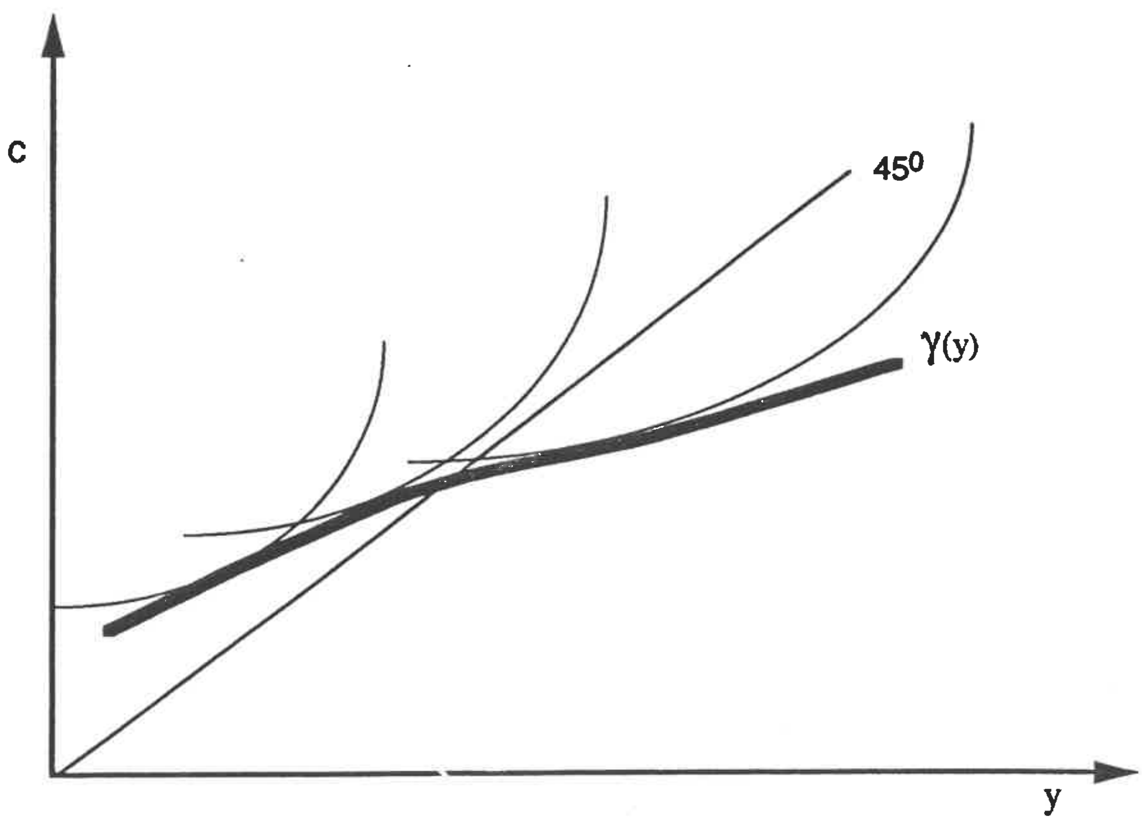

Figure 2: Single Crossing with Monotonicity

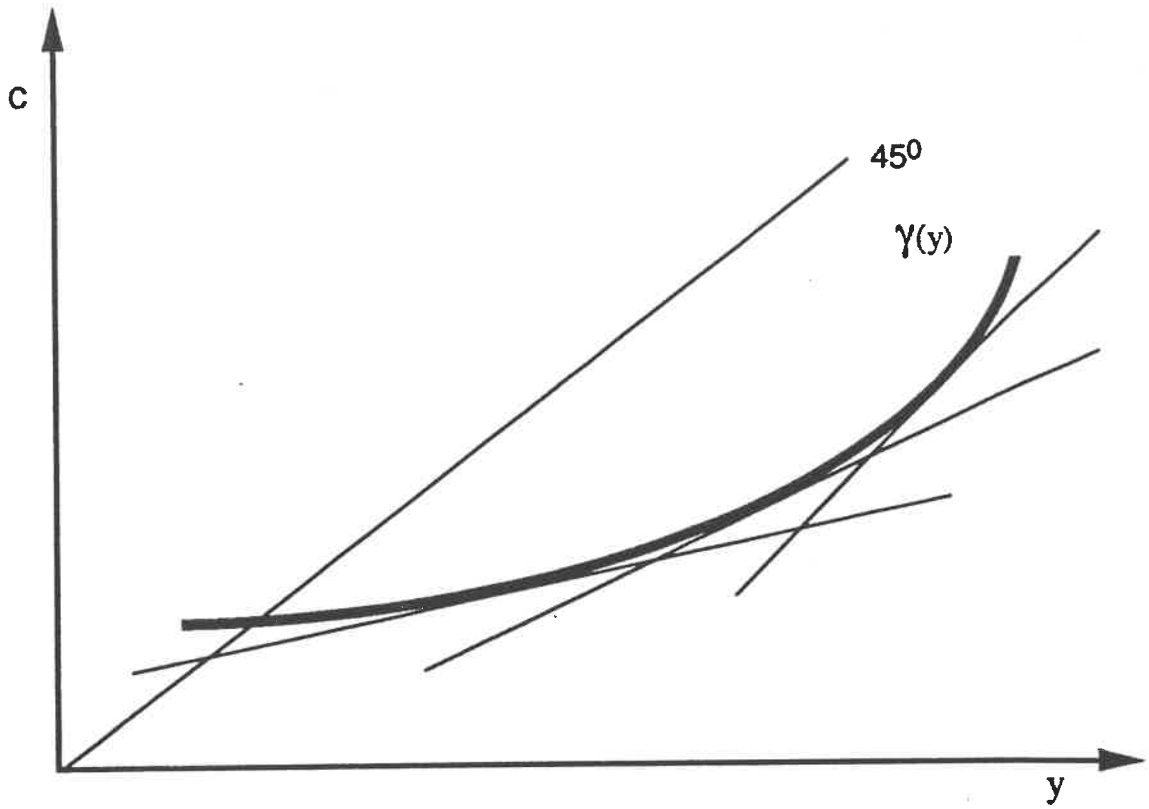

Figure 3: Single Crossing Without Monotonicity

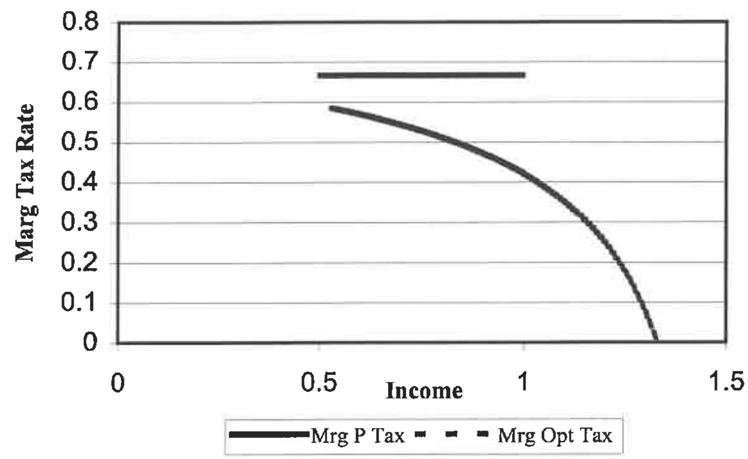

Figure 5: Marginal Tax Rates

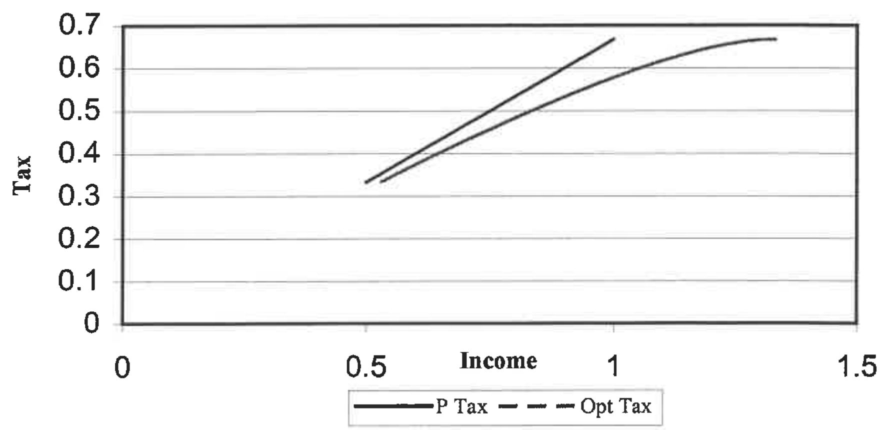

Figure 4: Tax Functions