Rochester Center for

Economic Research

Incentive Compatible Income Taxation, Individual Revenue Requirements and Welfare

Berliant, Marcus and Miguel Gouveia

Working Paper No. 234

June 1990

University of

Rochester

Incentive Compatible Income Taxation,
Individual Revenue Requirements and Welfare

Marcus Berliant and Miguel Gouveia

Rochester Center for Economic Research
Working Paper No. 234

INCENTIVE COMPATIBLE INCOME TAXATION, INDIVIDUAL REVENUE REQUIREMENTS AND WELFARE*

Marcus Berliant
University of Rochester

and

Miguel Gouveia
University of Rochester and
Portuguese Catholic University

Working Paper No. 234

June 1990

---

*The authors would like to thank John H. Boyd III, Christophe Faugere and Theory Workshop participants at the University of Rochester for helpful comments. The financial support of the NSF (Grant SES-9007800) and Sloan Foundation are gratefully acknowledged. The authors retain responsibility for any errors. Address correspondence to the Department of Economics, University of Rochester, Rochester, NY 14627, USA.

Abstract

This paper deals with conditions that ensure the validity of the first order approach to the optimal income tax model by focusing on the incentive compatibility part of the problem and examining it in greater depth than is possible in the usual framework. The basic concept used is the specification of an individual revenue requirement function, 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 is provided. A welfare ranking of the solutions is used to select the Pareto optimal tax function that implements the individual revenue requirements. Finally, a number of implications for the general optimal income tax problem are drawn, including a clarification of the occurence and reasons for "bunching." Key Words: Optimal Income Taxation, Incentive Compatibility. *Journal of Economic Literature* Classification Number: 323.

1. Introduction

The purpose of this paper is to improve our knowledge of the properties of the optimal income tax model by relying on techniques developed in the incentive compatibility literature, so that a few steps more are taken to increase the usefulness of the model as a tool for policy analysis. To begin with, the usual approaches to deriving an optimal income tax are based on extensive manipulation of the first order conditions of the problem. Recently, counter-examples were found where, even with extremely simple preferences, the first order approach breaks down. Since it is well known that there are conditions which ensure the validity of the first order approach in models with moral hazard (see Rogerson (1985)), can we find similar conditions for the optimal income tax model, a self-selection problem as well as a moral hazard problem? This is the basic motivation for our paper.

There are a few aspects of the optimal tax problem that are specific to it. In the standard model of optimal income taxation 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. 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. The optimal income tax is found by maximizing a social welfare function subject to the constraints of technology and the agents' maximizing behavior, the incentive constraints. In contrast with other problems involving incentive compatibility, there is no individual rationality condition. [1]

In this model it is simply assumed that the government enforces whatever 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. This is the reason why we must allow for general preference structures and not just for the particular separable cases studied in the principal-agent literature.

---

[1] Extensions to an open economy, with the possibility of migration, might actually include a constraint of that type.

By preventing the possibility of a differentiated lump-sum tax, the information structure makes the problem interesting because it becomes conceptually closer to realistic situations. Unfortunately, this feature also makes the problem very complex. This is the reason why the traditional formulations of the problem have relied almost entirely on the information provided by the first order conditions of the problem to extract qualitative results.

Good examples of this approach are the papers by Seade ((1977) and (1982)). Seade (1982) explores the implications of the information contained in the first order conditions for the sign of the marginal tax rates and proves that they must be strictly positive when both consumption and leisure are normal goods. However, in terms of the agents' optimization behaviour the only assumption made is agent-monotonicity or, as it became known in other areas of economics, the single-crossing [2] condition. But this assumption alone does not guarantee the validity of the results. From this perspective, our paper is complementary to Seade's , because we investigate sufficient conditions for the approach he used to be valid. Not surprisingly, the results are also complementary in the sense that a look at sufficient conditions will give us more qualitative information about the structure of optimal income tax functions.

To our knowledge, the first explicit mention of problems with the first order approach is in Mirrlees (1986), which was actually written prior to 1980. Mirrlees draws on the principal-agent literature to provide a numerical example of the inadequacy of representing agents' behavior using the first order conditions. He makes some remarks about the mathematical structure of the problem but is short on implications for optimal income tax computation. Recently, however, L'Ollivier and Rochet (1983) provided an interesting example with quasi-linear preferences and a weighted utilitarian social welfare function where the first order approach breaks down. Hence, first order conditions might actually lead to errors in the computation of the optimal income tax function that are hard to detect. They stressed the need to check the second order conditions for incentive compatibility. Their paper estab-

[2] See, for example, Cooper (1984).

lishes that single crossing is only a necessary, but not a sufficient condition for the first order approach to work. They are able to reformulate the problem as a mathematical programming problem but did not provide any conditions for the validity of the first order approach nor any intuition as when to expect its breakdown.

This paper contributes to the solution of these problems by isolating the incentive compatibility part of the optimal income tax problem and examining it in greater depth than what is possible in the usual framework. The basic concept used is the specification of an individual revenue requirement function, a mapping from abilities to taxes. This function is implicit in the solution to any optimal income tax problem, but here it is taken as given and the discussion is centered on the derivation of a tax function on income such 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 the tax function 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 function is found. A discussion of the necessary and sufficient conditions for the validity of this technique, using intensively the particular structure of the problem, is presented. These conditions have a simple interpretation and are easy to check for any preference specification. We also discuss what we can expect to happen when the conditions are not met. An interpretation of the results in graphs is provided.

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

A number of implications for the general optimal tax probem are drawn, including a clarification of the occurence and reasons for "bunching" of consumers. This is done by going back to the original optimal income tax model and examining what types of individual revenue requirement functions are implied by the solutions to the problem. Once this task is completed, we use our results to gain insight into the nature of the solutions to the optimal income tax problem.

This essay has direct implications for policy analysis. For one thing, policy might be designed with a revenue requirement function, rather than a social welfare function, in mind. Provided that certain conditions are obeyed, the two approaches are essentially equivalent, since we can always find a social welfare function that rationalizes a given revenue requirement function. But 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. It seems clear that any rationalization of the fiscal choices made through a social welfare function will not only be ex-post, but mainly ad-hoc. The reason for this assertion is that the mapping from social welfare functions to revenue requirement functions is not straightforward, as we will see in the second part of the paper. A more adequate formalization of a model for the tax system may be to specify preferences and choices directly at the revenue requirement level. Also, 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.

From a positive perspective, if we try to explain fiscal structures as the outcome of a political process, we find again that a revenue requirement function is the natural object where we should focus our attention. An example along those lines would be to use a model of direct democracy, where people vote over revenue requirement functions. 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, tax functions would then be the result of implementing the chosen revenue requirement functions.

Another implication of our results for policy is that they will be helpful when performing and evaluating numerical simulations of the optimal income tax model, thus improving on available literature such as Tuomala (1984). In order to perform simulations, the algorithms implicitly use the first order approach to solve the problem. Clear statements of sufficient conditions for the validity of the first order approach, together with extended knowledge about the bunching phenomenon, result in a significant improvement in the reliability and ease of numerical simulations. This may enhance the optimal income tax model's usefulness for applied policy purposes, such as revenue forecasting, incidence and efficiency analysis. Further developments in the model might enable it to address such issues as tax expenditures, tax evasion and even administration costs. In short, we believe that the optimal income tax model can have a destiny similar to general equilibrium models of tax incidence and make the transition from the theorist's ivory tower to the applied economist's toolbox.

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. The third section discusses the implications of the results in Section 2 for the optimal income tax problem and relates them to the literature, in particular Seade (1977) and (1982) and L'Ollivier and Rochet (1983). 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 $f(w)$, where $f(w) > 0$ a.s. 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 consumption c and labor l, represented by a C^2 utility function, $u(c, l)$. 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 $u(c, l)$. [3] Utility is employed in place of preferences purely as a matter of convenience.

We normalize the endowment of time to 1. 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 quasi-concave:

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

[3] See, for example, equation (3).

A3- Leisure is normal:

$$u_{11}u_2 - u_{21}u_1 > 0.$$

A4- Consumption is normal:

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

A5- Interior solutions:

$$\lim_{c \to 0} u_1(c, l) = \infty,$$

$$\lim_{l \to 1} u_2(c, l) = -\infty,$$

$$\lim_{l \to 0} u_2(c, l) = 0.$$

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 principal-agent 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 A5, 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 $y \equiv w \cdot l$ and $y = c$. Denote as Y the set of possible incomes, $Y = \{y | y = w \cdot l, \ \forall \ w \in W, \ \forall \ l \in [0, 1]\}$.

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

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

We assume 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 $\overline{y}(w) \equiv \text{argmax}_y u(y, y/w)$ and the indirect utility function as $v(w) \equiv u(\overline{y}(w), \overline{y}(w)/w)$.

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 $g : R_{++} \rightarrow R$, mapping from abilities to tax payments. Denote $T \subset R$ as the set of tax payments. Then we can be more precise and write $g : W \rightarrow T$.

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.

Notice that, in principle, we do not impose restrictions on the sign of $g(w)$. In particular, it can assume negative values with obvious meaning. However, we will need $g(w)$ to obey some conditions, summarized in the following assumption:

A6- Feasibility and regularity of $g(w)$:

$$g(w) \text{ is } C^2, \forall\, w \in W\ \exists\, \epsilon > 0 \text{ with } w - g(w) - \epsilon > 0.$$

For example, as in the optimal taxation literature we could specify a government budget constraint where the tax system must raise a non-negative aggregate revenue,

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

However, our results still hold even if g does not verify the inequality above, as in the case where the government disposes of profits realized by the public sector.

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. Define an income tax function $\tau : Y \to T$, which 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

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

and with solution $y(w)$.

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 adustments 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 $\tau(y(w)) \equiv g(w)$?

To find such a tax, we use the differentiable approach to the revelation principle [4]. 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 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, 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

[4] See Laffont and Maskin (1980).

strategy, as in the case of implementation in dominant strategies) is unaffected by other workers' strategies.

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.

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

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

A2 implies that ϕ is first best since the second order conditions are satisfied. A2 and the implicit function theorem imply that ϕ is well defined and C^1. In (w, c) space (defined in the positive orthant of R^2) the net income function defines two sets, where $u_1 + u_2/w$ is either negative or positive. Call the two areas ϕ^- and ϕ^+, according to the sign of the expression above.

(Fig 1 about here)

This last area is important in the proof of Proposition 1 that follows, as it corresponds to points where the second order conditions for incentive compatibiltity are verified. In particular, define $I(\overline{w}) = (0, \phi(\overline{w}))$. 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 A5 we must have $u_1 = \infty$. Hence $u_2 = -\infty$ so $l = 1$ and $w = g(w)$. Using A6, we have a contradiction.

We can now state our first result.

Proposition 1 : Assumptions A1-2, A4-6 and $\frac{dg}{dw} > 0$ are sufficient conditions to implement $g(w)$ by means of an income tax.

Proof/

We find a family of C^2 functions $\theta(w)$ that satisfy the first and second order conditions for incentive compatibility.

The problem solved by an agent of type w is

$$\max_{w'} u(\theta(w'), \frac{\theta(w') + g(w')}{w}) \tag{1}$$

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:

$$u_1(\theta(w), (\theta(w) + g(w))/w)\frac{d\theta}{dw} + u_2(\theta(w), (\theta(w) + g(w))/w)\frac{1}{w}(\frac{d\theta}{dw} + \frac{dg}{dw}) \equiv 0. \tag{2}$$

We now find $\theta(w)$.

It is immediate that we can rewrite (2) as:

$$\frac{d\theta}{dw} = -\frac{u_2(\theta, \frac{\theta+g(w)}{w})\frac{dg}{dw}/w}{u_1(\theta, \frac{\theta+g(w)}{w}) + u_2(\theta, \frac{\theta+g(w)}{w})/w} \tag{3}$$

Given our assumptions on $\frac{dg}{dw}$ and A1, the denominator is always different from zero for finite $\frac{d\theta(w)}{dw}$. This means (3) is well defined on both sides of ϕ because assumptions A1-A4 guarantee that this first order differential equation has a unique solution through any (w, θ) pair in ϕ^+ as well as ϕ^-.[5]

Recall that $I(\overline{w})$ is nonempty. Pick $\theta(\overline{w}) \in I(\overline{w})$, and let $\theta(\cdot)$ be the solution to (3) through $(\overline{w}, \theta(\overline{w}))$. Next we show that $\phi(w) > \theta(w) > 0,\ \forall\, w \in [\underline{w}, \overline{w}]$.

[5] See Spiegel (1967).

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

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 (3) $\lim_{n\to\infty} \frac{d\theta}{dw}|_{w_n} = \infty$. Now $\frac{d\phi}{dw}$ is bounded (at w^*), so $\exists\, w' > w^*$ with $\theta(w') \geq \phi(w')$, a contradiction, so $\theta(w) < \phi(w)$.

Using A1 and A5, $\theta(w) = 0$ implies by (3) that $\frac{d\theta}{dw} = 0$, so $\theta(w) = 0,\ \forall\, w \in W$. Since solutions to (3) cannot cross, any solution going through $I(\overline{w})$ has positive values for $\theta(w)$.[6]

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

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 (3) is C^1.

The second order condition for incentive compatibility is:

$$D^2 \equiv u_{11}(\frac{d\theta}{dw'})^2 + 2\frac{u_{21}}{w}\frac{d\theta}{dw'}(\frac{d\theta}{dw'} + \frac{dg}{dw'}) + u_{22}(\frac{d\theta}{dw'} + \frac{dg}{dw'})^2$$

$$+u_1\frac{d^2\theta}{dw'^2} + \frac{u_2}{w}(\frac{d^2\theta}{dw'^2} + \frac{d^2g}{dw'^2}) < 0. \tag{4}$$

If this condition holds everywhere, we have 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

$$D^2 - u_{21}\frac{(\theta+g)}{w^2}\frac{d\theta}{dw} - \frac{u_2}{w^2}(\frac{d\theta}{dw} + \frac{dg}{dw}) - u_{22}\frac{(\theta+g)}{w^3}(\frac{d\theta}{dw} + \frac{dg}{dw}) \equiv 0. \tag{5}$$

Combining the equations (2), (4) and (5), and using primes to denote derivatives in order to simplify the notation, we can write the second order condition for incentive compatibility as

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

We can use the first order condition (2) to rewrite the expression in brackets solely in terms of preferences:

[6] Since I(w) is open it is not crossed by the degenerate solution $\theta(w)$=0.

$$[\frac{u_{21}u_2 - u_{22}u_1}{u_2} + u_2].$$

Using assumptions A1 and A4, the expression above is always negative, so the second order condition is reduced to $\frac{d\theta(w)}{dw} + \frac{dg(w)}{dw} > 0$. Again using information in the first order conditions together with A1, this last inequality is equivalent to $\frac{d\theta}{dw} > 0$. Hence, under assumptions A1, A2, A4 and A5 an incentive compatible tax function exists if and only if $\theta'(w) > 0$. This amounts to requiring

$$u_1(\theta, \frac{\theta + g(w)}{w}) + u_2(\theta, \frac{\theta + g(w)}{w})/w > 0, \tag{6}$$

which follows since $\theta(w) \in \phi^+$, $\forall\, w \in [\underline{w}, \overline{w}]$.

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

$$\tau(y) = g(\eta(y)).// \tag{7}$$

Remarks: The monotonicity condition tells us that, in equilibrium, consumption must be an increasing function of ability. We note that this simple and intuitive result appears in most problems (see Laffont (1988)), but its appearance in this particular case is not trivial in the sense that we need to impose a substantial amount of structure on the preferences of the agents in order to obtain it.

The implications of the results derived above can be seen in a more intuitive manner if we ilustrate the problem by depicting two possible situations in (y, c) space, using graphs developed in Sadka (1976).

Normality of consumption (A2) 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 $\theta(y)$ in ascending ability order.

(Figure 3 about here)

In the case depicted in Figure 3, we have single crossing, but the consequence is a preverse 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 preversity of the situation is in the fact that taxes are 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 in a computer, and we had a program which solved the differential equation (3) 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 behaviour of the agents. Any $\tau(y)$ derived under these circumstances is meaningless.

The conditions usually used in optimal tax models to avoid bunching of agents, such as single crossing, are also part of the sufficient conditions of our problem. [7] However, in most cases such conditions are not imposed, raising the question of the validity of the results in such environments. Even when they are used, there usually is no assumption equivalent to $\frac{dg}{dw} > 0$. We postpone a more detailed discussion of this problem to the next part of the paper.

[7] See Seade (1977).

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:

$$\tau'(y) = g'(\cdot)\phi'(\cdot) = \frac{g'(w)}{g'(w) + \theta'(w)} \tag{8}$$

So far we have proved that we can implement $g(w)$ by means of an income tax. This can be done by solving the differential equation (3) 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,A2,A4-A6, if $g'(w) > 0$, 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.

Proof/

Under the assumptions, every income tax implementing g has an associated net income function satisfying (3). Since all are solutions to the same differential equation, they do not intersect. Hence they are Pareto ranked,[8] and so are the income taxes implementing g.

To find the best of these, let $\{\theta_n(\cdot)\}_{n=1}^{\infty}$ be a sequence of admissible solutions to (3) 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

[8] For each w and each net income level c, less labor is required to obtain the same consumption c using a schedule ranked higher.

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

$\forall\ (w^*, \theta^*) \in Z,\ \ \forall\ w'$ 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 (3) 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 (3) passing through an increasing sequence of points in $I(\overline{w})$. In the end, $\theta(\overline{w}) = \phi(\overline{w})$, so equation (3) 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, and the associated income tax is incentive compatible. Moreover, since $\theta(w) \geq \theta'(w)$ for all admissible solutions [9] $\theta'(w)$, $\theta(w)$ 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 (3). 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 the denominator of (3) vanishes (and $\frac{d\theta(w)}{dw}$ approaches infinity), which also implies that the marginal tax rate is zero at that point. [10]

Appendix 1 contains an example with explicit solutions illustrating Propositions 1 and 2.

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.

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

[10] See the Appendix of Seade (1977) for an equivalent result.

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

<u>Proof/</u>

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. 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. The last fact follows since u is C^2 and quasi-concave. Then, without loss of generality (and without 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

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

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

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

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

One final remark is in order. Although we have found the best tax system implementing g, one important question remains open. Is this tax system (second best) Pareto optimal among tax systems satisfying the *aggregate* revenue constraint? We conjecture that further conditions on both g and utility would be needed to answer the question in the affirmative.

3. Implications for the Optimal Income Tax Problem

Using the notation above we can define formally the optimal income tax problem as:

$$\max_{\gamma()} \int_{\underline{w}}^{\bar{w}} u(\gamma(y(w)), y(w)/w) f(w) dw \tag{9}$$

$$\text{subject to} \qquad R(s) \leq \int_{\underline{w}}^{\overline{w}} [y(w) - \gamma(y(w))] f(w) dw \tag{10}$$

$$\text{and} \qquad y(w) = \{y \in R \mid y \quad \text{maximizes} \quad u(\gamma(y), y/w)\}. \tag{11}$$

Equation (11) is the incentive compatibility constraint for an agent of type w. It can be written as a continuum of inequalities,

$$y \in y(w) \quad \text{if} \quad u(\gamma(y), y/w) \geq u(\gamma(y'), y'/w),\ \forall\, y' \in Y.$$

Since we have a continuum of agents, it is evidently of the utmost importance to replace (11) by something more manageable. Ideally, one would like to be able to use a first order condition instead. This is what is actually done in the optimal income tax literature. What are the implications of the analysis presented in the second section of this paper for the optimal income tax problem? We can start by looking at the $g(w)$ function implicit in any derived solution to the optimal income tax problem defined by $g(w) \equiv \tau(y(w))$ and check whether the assumptions used in Proposition 1 are satisfied. If that is the case, we know a solution to an optimal income tax problem satisfies incentive compatibility. Alternatively, we can check whether the $\theta(w)$ and $y(w)$ implicit in the solution have positive derivatives. These ex-post verifications will not be useful in computing the solution to (9), (10) and (11), but they may help rule out candidates for solutions that fail to meet the requirements. A simple application of this idea would be to include in the programs that solve numerically for $\tau(y)$ a final routine computing $g(w)$. Equivalently, (6) can be checked at each stage of the computation. Simple as this suggestion may be, it will certainly help to avoid gross mistakes.

Another route that might be explored is to transform the problem defined by (9), (10) and (11). In this case we include an additional constraint in the problem, that $g'(w)$ be positive. This approach is in the same line as L'Olliver and Rochet (1983).

The most satisfactory way to solve the problem would be to find *ex ante* conditions on preferences and social welfare functions such that the g resulting from the solution to an optimal income tax problem would be well-behaved (i.e. satisfy A6 with $g' > 0$). A starting point will then be the consideration of specific forms for the social welfare function, namely of the utilitarian type. Unfortunately, even in this particular case we need to make strong additional assumptions in order to obtain positive results. For this particular case, we can prove the following:

Proposition 4: If the Social Welfare Function is utilitarian, the utility function is additively separable and C^3, Assumptions A1 and A5 hold, the marginal utility of consumption is concave and the marginal utility of leisure is convex, then the first order approach is valid.

Proof/

See Appendix 2.//

The proof is based on Section 2 and Seade (1977). Mirrlees (1986) contains material somewhat related. The proof consists of the computation of the first and second order conditions of the variational problem associated with (9) and (10), with the first order conditions for the agents in place of (11). These conditions yield an implicit g with a positive derivative.

As pointed out in the introduction, we do not consider the difficulty of establishing a mapping from social welfare functions to revenue requirement functions to be important, since we believe modeling at the revenue requirement level is more profitable. But we certainly want to rule out extreme cases where we can be sure that there does not exist a reasonable social welfare function which rationalizes a given $g(w)$. From this perspective, it is interesting to extend the analysis to cases where g may not have all the required properties. To begin with, consider the case of a strictly decreasing $g(w)$. It is straightforward to construct analogs of Propositions 1, 2 and 3 for this case, but this can hardly be considered of any practical relevance. We are not aware of any instances where income taxes are negatively related to ability to pay. Nevertheless it is a perfectly admissible structure as far as incentive compatibility is concerned. However, a negative $\frac{dg}{dw}$ can never be a solution to the optimal income tax model. This has been proved for the utilitarian case (assuming the first order approach is valid) provided that A1-A5 hold and that the utility functions are concave. [11]

Where $g(w)$ is (locally) flat, we can distinguish two cases. In the first, agents with different w's have different consumption and gross incomes, much as if there were no taxes (or only head taxes) as far as these agents are concerned. This situation does not arise in the usual models of optimal income taxation. In the second case, agents with different w's have the same consumption and gross income (but not the same welfare since higher w's work less to earn the same gross income). "Bunching" occurs here. Bunching is also known as pooling in the economics of information (e.g. Rothschild and Stiglitz (1976)). Bunching in optimal income tax models was extensively discussed in Mirrlees (1971), Seade (1977) and L'Ollivier and Rochet (1983). However, the situation referred to as "bunching" in Mirrlees

[11] See Seade (1982).

and Seade does not seem to be the same as the one mentioned by L'Ollivier and Rochet. In the situation we call MS-bunching, agents are grouped at a zero work point because the non-negativity constraint for labor supply becomes binding. This situation is eliminated in models with assumption A5, first made by Seade (1977). The other type of bunching, that we call OR-bunching, originates in a failure of the second order conditions for incentive compatibility. In this case we may have that the first order condition is verified, but in fact at a local minimum or inflexion point of an agent's Lagrangean. The global maximum for the agent is then at a corner of the budget set. The two types of bunching are clearly distinct: MS-bunching corresponds to corner solutions in (convex) individual opportunity sets, while OR-bunching corresponds to a failure of these sets to be convex. If these differences are to be recognized, then the results of numerical calculations that take the first type of bunching into account still have to be checked for possible OR-bunching. For example, two circumstances that require extra care are:

– The use of double log utility functions (which do not have concave marginal utility of consumption, as in the conditions for Proposition 4) to perform numerical calculations.

–The S shape form of the tax schedule, a result claimed by Seade (1982). This interesting result says that since marginal tax rates are zero at the top and bottom of the ability distribution and strictly positive in the interior (given regularity assumptions on the density function $f(w)$) the optimal income tax function has an S shape. The potential problem lies in the fact that such a shape may well generate a singularity in the implicit net income function. The upper part of the tax function may actually generate a budget constraint with a non convexity, as depicted in Fig.4.

(Figure 4 about here)

If a situation like the one illustrated in Figure 3 or Figure 5 is a real possibility, we can only emphasize that a computation of $\theta(w)$ or $g(w)$ is necessary, in order to be sure the solutions computed are indeed incentive compatible.

4. Conclusions

The results in the paper allow us to draw some conclusions.

First, sufficient conditions for the implementation of well-behaved revenue requirement functions have been found. This allows us to ignore incentives when modelling the determination of the distribution of tax burdens. Simply use revenue requirement functions to generate incentive compatible income tax functions.

Second, modeling strategies in income taxation should emphasize revenue requirement constraints instead of the maximization of social welfare functions. The results are transparent, in terms of the implications for the shape of the tax function, making it easier to interpret the result as the outcome of an equity criterion or of a political process. In optimal tax models the mapping from social welfare functions to tax payments by types of agents can be quite unpredictable.

Third, optimal income taxes have a complex structure. This complexity makes the fulfillment of conditions for incentive compatibility a somewhat less than trivial requirement. Problems arise calculating numerical solutions to the model and might even affect some of the general results in the literature. We suggest a simple way to verify *ex post* whether there might be problems with a particular computation.

Appendix 1

This appendix contains an example illustrating Propositions 1 and 2. Even though it does not meet assumption A5, it provides a closed form interior solution to the problem of implementing a revenue requirement function. We assume agents have Cobb-Douglas preferences, $u(c,l) = log(c) + log(1-l)$ 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

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

The solution is

$$\phi(w) = \frac{w - g(w)}{2} = \frac{w(1-k)}{2}.$$

This equation defines the area over which the second-best net income functions must be constructed.

The second-best best net income function is obtained following the steps of Proposition 1. The agent sends the message w' that

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

Truthtelling requires

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

So, we obtain the differential equation that generates 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 solvable by standard methods. The solution is:

$$w = \frac{2\theta + c\theta^{k^{-1}-1}}{1-2k} \text{ for } k \neq .5$$

$$w = 4\theta(c - log(\theta)) \text{ for } k = .5,$$

where c is the constant of integration. By picking points in $(0, \phi(2))$ we determine the value of c 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 c 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. 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:

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

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

Appendix 2

This appendix contains a discussion of the optimal income tax problem, with statements of the first and second order conditions for a solution. The proof of Proposition 4 is an immediate consequence of the results.

Proposition 4: If the Social Welfare Function is utilitarian, the utility function is additively separable and C^3, Assumptions A1 and A5 hold, the marginal utility of consumption is concave and the marginal utility of leisure is convex, then the first order approach is valid.

Proof/

Using the notation defined in the body of the paper, we state the optimal income tax problem as

$$\max_{\theta(),y()} \int_{\underline{w}}^{\overline{w}} u(\theta(w), y(w)/w) f(w) dw \qquad A2.1$$

$$\text{subject to} \qquad R(s) \leq \int_{\underline{w}}^{\overline{w}} [y(w) - \theta(w)] f(w) dw \qquad A2.2$$

$$\text{and} \qquad u_1 \theta'(w) + \frac{u_2}{w} y'(w) = 0. \qquad A2.3$$

This is a calculus of variations problem, linear in the derivatives. It is easy to show that it is equivalent to other statements of the problem such as Mirrlees (1971), which is not easy to interpret directly in our revelation framework.

Define the function F as:

$$F = [u(\theta, y/w) + \lambda(y - \theta)] f(w) + \mu(w)(u_1 \theta' + u_2/wy').$$

The Euler equations for θ and y are:

$$\frac{\partial F}{\partial \theta} = \frac{d(\frac{\partial F}{\partial \theta'})}{dw} \Leftrightarrow (u_1 - \lambda) f(w) = \mu'(w) u_1 - \mu(w) u_{12} \frac{y}{w^2}$$

and

$$\frac{\partial F}{\partial y} = \frac{d(\frac{\partial F}{\partial y'})}{dw} \Leftrightarrow (u_2/w + \lambda) f(w) = \mu'(w) u_2/w - \mu(w)(u_{22} \frac{y}{w^3} + \frac{u_2}{w^2}).$$

Solving the first equation for μ' and substituting in the second equation, we obtain

$$(1+\frac{u_2}{u_1 w}) = \frac{\mu(w)}{\lambda f(w)}[(\frac{u_{12}u_2 - u_{22}u_1}{u_1 w})\frac{y}{w^2} - \frac{u_2}{w^2}]. \qquad A2.4$$

The Euler equations are necessary conditions for a solution of the optimal income tax problem. To prove sufficiency, and hence validity of the first order approach, we need to check the concavity of F. Since F is linear in θ' and y' the Legendre and Weierstrass conditions are trivially satisfied. We need only investigate the concavity of F in the state variables. This amounts to requiring the matrix of the second partial derivatives with respect to θ and y to be negative semi-definite.

$$D=\begin{pmatrix} u_{11}f(w)+\mu(w)[u_{111}\theta'+u_{211}/wy'] & u_{21}f(w)/w+\mu(w)[u_{211}/w\theta'+u_{122}/w^2y'] \\ u_{21}f(w)/w+\mu(w)[u_{211}/w\theta'+u_{122}/w^2y'] & u_{22}f(w)/w^2+\mu(w)[u_{122}/w^2\theta'+u_{222}/w^3y'] \end{pmatrix}$$

By A1 the utility function is concave. Consequently we can decompose D as a sum of two matrices where one of these matrices is negative definite, since it is just the Hessian of the utility function (multiplied by the density function). Thus sufficient conditions for the utilitarian optimal income tax are verified if the other matrix, derived from the incentive compatibility constraint is concave in θ and y i.e. we require the matrix

$$\Delta = \begin{pmatrix} u_{111}\theta'+u_{211}/wy' & u_{211}/w\theta'+u_{122}/w^2y' \\ u_{211}/w\theta'+u_{122}/w^2y' & u_{122}/w^2\theta'+u_{222}/w^3y' \end{pmatrix}$$

to be negative semi-definite.

If we introduce the assumption that the utility function is additively separable the matrix D becomes

$$D' = \begin{pmatrix} u_{11}f(w)+\mu(w)u_{111}\theta' & 0 \\ 0 & u_{22}f(w)/w^2+\mu(w)u_{222}/w^3y' \end{pmatrix}$$

and Δ', obtained from D' in a similar way is:

$$\Delta' = \begin{pmatrix} u_{111}\theta' & 0 \\ 0 & u_{222}/w^3y' \end{pmatrix}.$$

Sufficient conditions for a solution to the optimal income tax problem are verified if Δ' is negative semi-definite. That amounts to $u_{111}\theta' \leq 0$ and $u_{222}y' \leq 0$. If both third derivatives are negative, i.e. the marginal utility of consumption is concave and the marginal utility of leisure is convex, then an admissible solution to the optimal income tax problem requires positivity of θ' and y'. Thus, by the results in the proof of Proposition 1, the solution to the optimal income tax using the first order condition of the agents instead of (11) is incentive compatible.//

Remarks – This is the only simple case we have come across where the individual incentive compatibility second order conditions are explicitly obeyed when we solve the planner's problem. But even in this case we have not ruled out that other solutions to the Euler equations might fail to satisfy the second order conditions of the planner's problem.

– Equation A2.1 is the basic equation defining the tax schedule. The left hand side is the marginal tax rate. By the results in Seade (1982) and A3-A4 [12] it is positive. Numerical computations of optimal taxes are usually based on this equation, after a suitable change of variables. No explicit computations of the functions $\theta(w)$ and $y(w)$ are usually performed. Thus, we cannot be sure the functions implicit in the numerical calculations are both increasing.

The result discussed in Proposition 4 is related to the results of Proposition 1, as shown by the following lemma.

Lemma: The assumptions listed in Proposition 4 imply that the revenue requirement function implicit in the solution to the optimal income tax problem has a positive derivative.

Proof/

The incentive compatibility conditions θ' and y' positive are satisfied. To prove

[12] A1 and additive separability imply A2-A4.

g' is positive, use A2.3 to derive

$$g' \equiv y' - \theta' = (1 + \frac{u_2}{u_1 w})y'.$$

By A2.4 and Seade (1982) the result follows due to the positivity of the marginal tax rate.//

To assess the implications of these results, consider the case of the utility function $u(c, y/w) = log(c) + log(1 - \frac{y}{w})$. This utility function has been used by Mirrlees (1971) and by Tuomala (1984), among others, to solve numerically for the optimal income tax. A potential problem with this utility function is that the marginal utility of consumption is convex. The first element of the second partials matrix is then

$$u_{11} f(w) + \mu(w) u_{111} \theta' = -c^{-2} f(w) + \mu(w) 2c^{-3} \theta'.$$

For small enough values of c the expression above may be positive!

References

[1] **Berliant, M. (1989),** "On Income Taxation and the Core", *University of Rochester ,* mimeo .

[2] **Cooper, R. (1984),** "On Allocation Distortions in Problems of Self-Selection", *Rand Journal of Economics,* **15**, 568-577 .

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

[4] **Lollivier, S. and J-C Rochet (1983),** "Bunching and Second Order Conditions: A Note on Optimal Tax Theory", *Journal of Economic Theory,* **31**, 392-400 .

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

[6] **Mirrlees, J. (1986),** "The Theory of Optimal Taxation", in Arrow and Intriligator (eds.), *Handbook of Mathematical Economics, Volume III* (Amsterdam : North-Holland) .

[7] **Rogerson, W. (1985),** "The First Order Approach to Principal Agent Problems", *Econometrica,* **53**, 1357-1367 .

[8] **Rothschild, M. and J. E. Stiglitz (1976),** "Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Imperfect Information", *Quaterly Journal of Economics,* **90**, 629-650. .

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

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

[11] **Seade, J. (1982)** , "On the Sign of the Optimuum Marginal Income Tax ", *Review of Economic Studies,* **49**, 637-643 .

[12] **Spiegel, M.R. (1967),** *Applied Differential Equations* (Englewood Cliffs, N.J. : Prentice-Hall) .

[13] **Stiglitz, J.E. (1982),** "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics,* **17**, 213 - 240 .

[14] **Tuomala, M. (1984),** "On Optimal Taxation: Some Further Numerical Results", *Journal of Public Economics,* **23**, 351-366 .

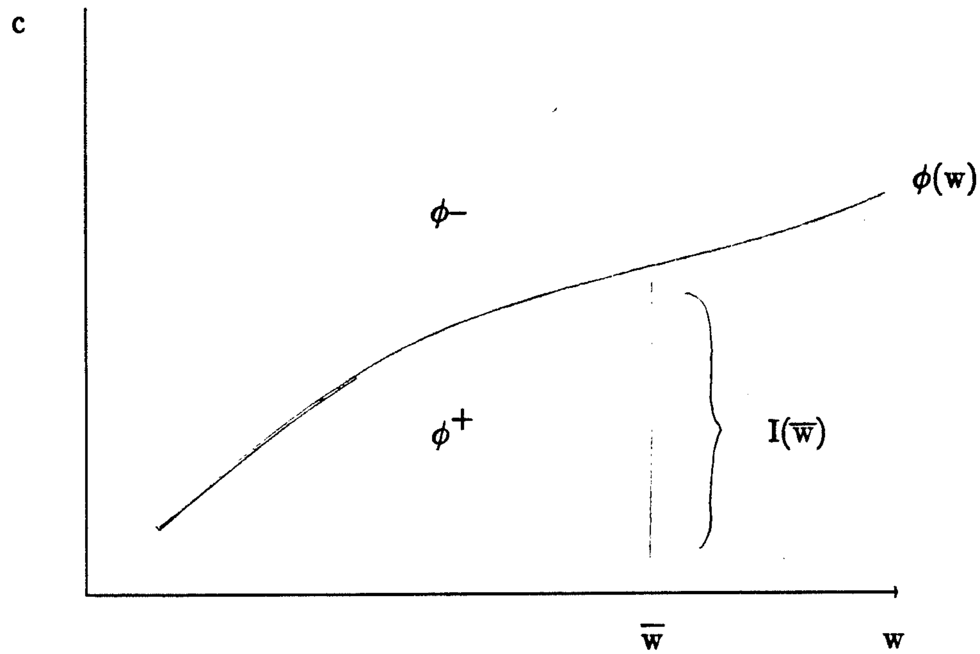

Figure 1 – The First–Best Net Income Function

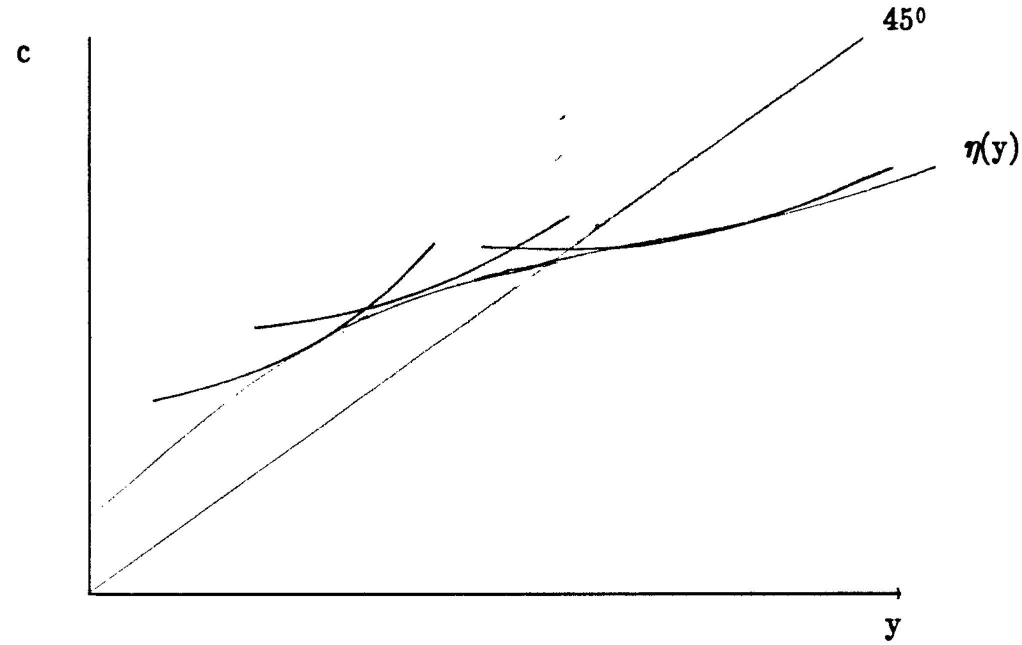

Figure 2 – Single Crossing with Monotonicity

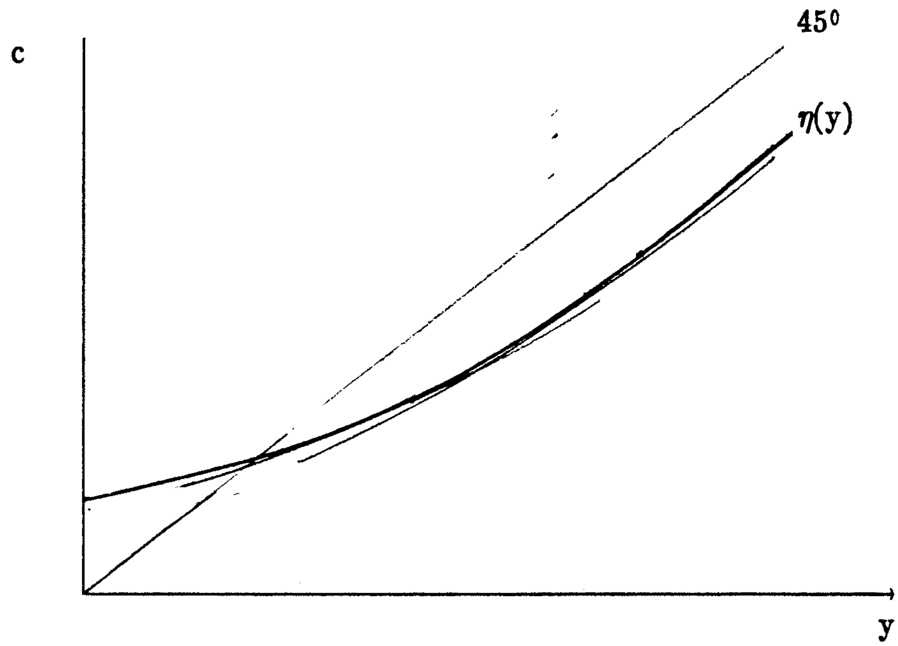

Figure 3 – Single Crossing Without Monotonicity

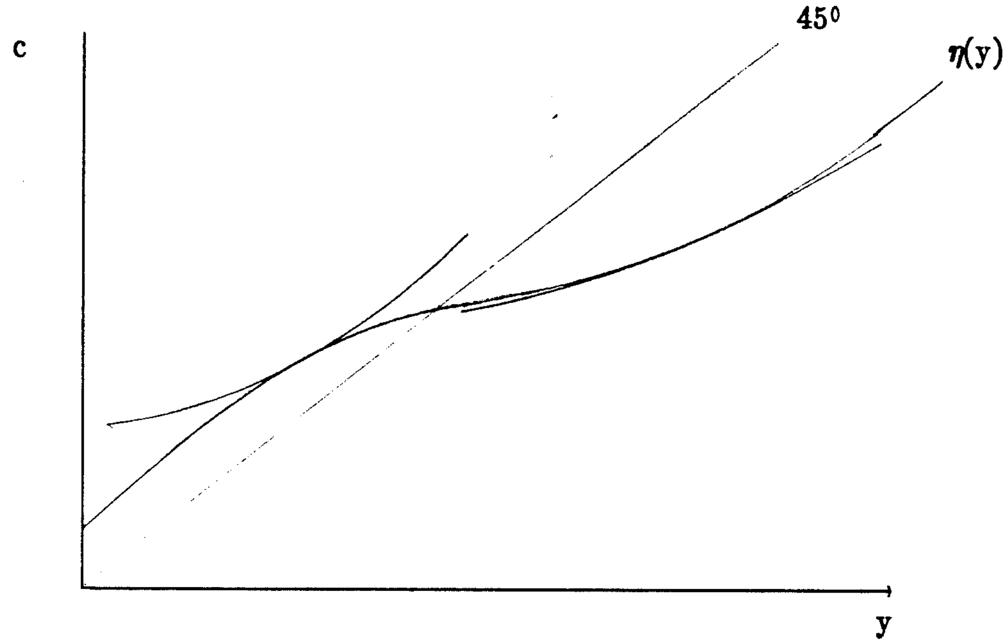

Figure 4 – Seade's Result and a Violation of Monotonicity