The Optimal Income Tax: Restatement and Extensions

By Stefan Homburg

School of Economics and Management, University of Hannover, Königsworther Platz 1, D-30167 Hannover, Germany. www.fiwi.uni-hannover.de.

ABSTRACT: The paper presents an axiomatic restatement of the standard approach to nonlinear income taxation. It assumes a finite number of taxpayers rather than an uncountable infinity. Analytical tools are developed which facilitate proving the existence of tax schedules that are continuous, differentiable almost everywhere and possess left- and right-derivatives at every point of non-differentiability. These tax schedules even become differentiable in the limit. Numerical examples are provided.

KEYWORDS: Optimal Taxation.

JEL-Classification: H2, D3

1. Introduction

What does an income tax schedule look like, which takes account of the trade-off between equity and efficiency? This question was first asked by Mirrlees (1971) who developed the standard model of the optimal nonlinear income tax. Since then innumerable papers have generalized, refined, or corrected his analysis. It has also been realized that the second-best approach to income taxation pertains to a wide variety of economic problems such as monopoly pricing or contract theory in general. In this respect Mirrlees' article has opened an important and fascinating strand of economic thought.

Yet, the theory of optimal income taxation is not in good shape. It contains a number of hidden assumptions and logical compromises, beginning with the problems that the Kuhn-Tucker conditions are neither sufficient nor necessary for global optima and that the usual condition for individual optima do not hold at corner solutions. Moreover, the analysis is so complex that many economists believe it's just not worth the time struggling mechanically with dozens of equations in order to obtain one of the common results, while others such as Diamond (1998, 86) have simply resigned to the quandaries resulting from "bunching" or from gaps in the income distribution. A good deal of these difficulties is due to a premise which was meant to simplify the theory and make it more elegant: The natural problem of taxing a finite number of persons was replaced by the problem of taxing an uncountable infinity. This continuum approach eventually yields the desired results, in a more or less satisfactory manner. But as Mirrlees (1971, p. 179) himself pointed out, the use of Pontriyagin's maximum principle fails to provide economic as opposed to mathematical insights. Thus, the logic behind some of the standard features of the optimal income tax remains obscure.

This paper is devoted to optimal nonlinear income taxation in an economy with finitely many taxpayers. Important first steps in this direction have been undertaken by Stiglitz (1982), who analyzed the situation of two taxpayers only, Guesnerie and Seade (1982), who introduced a special redistribution motive, Weymark (1986), who concentrated on quasi-linear utilities, and Brunner (1989). The basic model is stated in section 2. The assumptions are fairly general and close to those of the standard model except that there are finitely many persons.

Section 3 presents an axiomatic approach to characterizing second-best allocations. Such an approach has obvious advantages: Firstly, it is simpler in the mathematical sense. Secondly, it is easier to understand. Thirdly, it is self-contained without reference to other published or unpublished material or to additional calculations left to the reader. Two analytical tools are introduced. The first is the concept of a "utility curve" which was used by Matthews and Moore (1986) in a paper on monopoly pricing. The second is a revision of Seade's (1982) notion of "agent monotonicity". Equipped with utility curves and the new form of agent monotonicity it is possible to replace the usual graphical (revealed preferences) arguments by analytical proofs. As a side-product, new graphical illustrations emerge which can be used for classroom purposes. Every proof is designed in such a way as to make clear the economic meaning of the result.

It is only fair to say that the finite approach to optimal income taxation has not found many adherents; most recent papers still use the continuum model. The reason may be that the only tax schedules, which have been shown to support a second-best optimum, are step functions.

Due to their discontinuity these do not look appealing to tax theorists. Moreover, the optimum is characterized by strictly positive implicit marginal tax rates (defined in section 4) while the explicit tax rates are either zero or not defined in the case of step functions. Section 5 contains our main result which demonstrates that any second-best optimum can be supported by a tax schedule which is continuous, differentiable almost everywhere and for which left- and right-derivatives exist at every point of non-differentiability.

The relationship between the continuum approach and the finite approach has remained un-clear as yet since the former usually *assumes* tax schedules that are differentiable throughout whereas the latter *shows* that such schedules cannot exist. The central result of section 6 is a limit theorem which clarifies this tension. It demonstrates under which circumstances non-differentiable tax schedules for finite economies converge to a differentiable limit as the skill distribution becomes dense.

Another reason for the lacking popularity of the finite approach is presumably that not one single example of a tax schedule can be found in the literature. This is surprising because the simplicity to construct examples and to implement the model empirically is a particular strength of the finite model. Conversely, it is a weakness of the continuum model which, when it comes to applications, must be recast in finite form anyway since the integrals can only be solved numerically. Section 7 provides some illustrations for Cobb-Douglas and CES utility functions.

The methods developed below may perhaps also be of use in other branches of economics, such as in contract theory. They can certainly be applied to the numerous variants of the standard model, including consideration of public goods (Boadway and Keen 1993), nonlinear technologies (Stiglitz 1987), fixed labor supplies (Diamond 1980), luck as an additional income source (Strawczisnky 1999), or poverty as a public bad (Wane 2001).

2. The Basic Model

In the model economy there live finitely many persons who have different skills and thus earn different wage rates per hour, w^h, where $h = 0, 1...H$ and $0 = w^0 < w^1 < ... < w^H$. The premise that w^0 vanishes is not essential but shortens some arguments as there will always be at least one person with zero income. $f^h > 0$ denotes the probability mass of people with wage rate w^h. Each Person with wage rate w^h (person h, for short) chooses some *consumption* c^h and some *labor supply* ℓ^h such that the *commodity bundle* (c^h, ℓ^h) belongs to the uniform consumption space $\mathcal{C} = \mathcal{R}_{o+} \times [0; \ell^{max}]$ where $\ell^{max} > 0$. A person's *gross income* is denoted as $y^h = w^h \ell^h$.

Assumption A: The uniform utility function $u: \mathcal{C} \rightarrow \mathcal{R}$ is continuous, strictly monotonically increasing in c, strictly monotonically decreasing in ℓ and strictly concave on its entire domain. At least in the interior, it is twice continuously differentiable with partial derivatives $u_c(c, \ell) > 0$ and $u_\ell(c, \ell) < 0$ and a negative definite Hessian. Moreover, $-u_\ell(c, 0)/u_c(c, 0) < w^H$ for c small enough.

The assumptions are all standard, except the last one which deserves some comment. In an attempt to rule out corner solutions, several papers in the field suppose utility functions that are not continuous on their entire domain (Brito and Oakland 1977, p. 408) or aggregate consumption spaces that are not bounded (Guesnerie and Seade 1982, p. 159) or not closed

(Svensson and Weibull 1986, p. 167). In these cases the mere *existence* of an optimum becomes questionable and cannot be proven by standard arguments. The present approach explicitly allows corner solutions, but we need a minimum regularity to make the problem meaningful: If the marginal rate of substitution, evaluated $\ell^h = 0$, exceeded the highest wage rate at every consumption level, no one would choose to work even in the absence of taxes and it would be impossible to extract a strictly positive revenue by means of an income tax. An *allocation* is a vector $(c^h, \ell^h)_{h=0...H}$ in $\mathscr{C}^{H+1}$. The *social objective* reads

$$\underset{(c^h, \ell^h)_{h=0...H} \in \mathscr{C}^{H+1}}{max!} \quad \sum_{h=0}^{H} u(c^h, \ell^h) f^h \,. \tag{1}$$

Aggregate (per capita) consumption is $\Sigma\, c^h f^h$. Assuming a linear technology, aggregate (per capita) output is $\Sigma\, y^h f^h$. With an exogenous *tax revenue* $g > 0$, also defined in per capita terms, an allocation must satisfy the *resource constraint*

$$\sum_{h=0}^{H} (y^h - c^h) f^h \geq g \,. \tag{2}$$

As the differences $y^h - c^h$ are in fact tax payments, this inequality corresponds to the government's budget constraint. If the inequality holds strictly there is an *output surplus*. A *first-best optimum* maximizes the social objective (1) subject to (2).

The second-best approach to optimal income taxation holds that the government is unable to observe wage rates and working hours directly but can only observe consumption and income levels. Stipulating a tax schedule means confronting the persons with legal choices (c, y). Since wage rates are exogenous, the preference ordering over commodity bundles (c, ℓ) induces a preference ordering over *pairs* (c, y) of consumption and income which depends on the respective wage rate. Some person k accepts the pair (c^k, y^k) intended for him only if no other pair (c^h, y^h) exists which he prefers strictly. Otherwise person k will *mimic* person h by choosing c^h instead of c^k and y^h/w^k instead of ℓ^k. As a result person k has consumption c^h and gross income $w^k y^h/w^k = y^h$ and cannot be distinguished from person h. Therefore every *feasible allocation* must satisfy the resource constraint and the *self-selection constraints*

$$u(c^k, \ell^k) \geq u(c^h, y^h/w^k) \quad \text{for all k and h where} \quad y^h/w^k \leq \ell^{max}. \tag{3}$$

If $y^h/w^k > \ell^{max}$, person k cannot *reach* the pair (c^h, y^h) so that the respective self-selection constraint is not defined and irrelevant. A self-selection constraint is called *downward* if $k > h$ and *upward* if $k < h$; it is called *adjacent* if $k = h \pm 1$. A *second-best optimum* maximizes the social objective (1) subject to the resource constraint (2) and the self-selection constraints (3).

The social objective is often taken to presuppose interpersonal utility comparisons. However, it can be given a completely different interpretation which does not suffer from the philosophical and measurement problems raised by such comparisons. Consider a man who knows the entire model but does not yet know his own wage rate. The significance of this premise is that it induces an impartial choice. People with low wage rates are apt to vote for tax schedules with high transfers at the bottom of the income scale and vice versa: these are partial decisions. By contrast, a man deciding under a "veil of ignorance" will consider each feasible allocation to be a *lottery*, with f^h giving the probability to receive the commodity bundle

(c^h, ℓ^h) after the veil is lifted. If one defines a preference ordering over these lotteries which satisfies the familiar continuity and independence axioms, the social objective (1) turns out to be nothing but an expected utility representation of this preference ordering. Without any regard to interpersonal utility comparisons, the resulting optima are not optimal for the poor or the rich or somebody in between, but for a man who is completely uncertain about his future income. The man's optimal choice will reflect his risk aversion on the one hand and the deadweight losses from taxation on the other.

Such an expected utility interpretation accords well with Rawls' (1971) concept of "justice as fairness" but does not take recourse to his "difference principle". It places convenient restrictions upon the social objective which help sharpen the analysis: the Bernoulli utility functions must be identical across skill levels and transforming or weighting them is not allowed. Harsanyi (1955) has sponsored a similar interpretation of the utilitarian objective. However, he discussed an abstract "soul" that randomizes not only over skill levels but also over preferences, implying that interpersonal utility comparisons re-enter through the backdoor. The present interpretation was suggested first by Vickrey (1960).

3. Second-best Allocations

The above optimization problem does not appear particularly difficult at first sight. However, the set of feasible allocations is not convex. To see this, consider two feasible allocations with different (c^h, ℓ^h) for some h and the same (c^k, ℓ^k) for some k>h such that the self-selection constraint (3) holds with equality in both cases. As the preference ordering is strictly convex, every convex combination of these two allocations will violate the self-selection constraint and will thus not be feasible itself. It has sometimes been observed (Diamond 1998, p. 86) that if the marginal tax rate decreases, individual budget sets will be non-convex, too, but a true solution will allow for this fact. The deeper problem is the non-convexity of the feasible set which entails that the familiar Kuhn-Tucker conditions are not sufficient for an optimum. What is more, they are not even necessary unless one shows that some constraint qualification is satisfied; see Bertsekas (1999, proposition 3.3.12).

The following analysis makes no use of Kuhn-Tucker conditions or of the maximum principle. It proceeds in two steps. Firstly, the basic properties of optimal allocations are described without explicit reference to taxation. This task is performed in the present section and in section 4. Secondly, tax schedules which support the optimum are introduced in section 5. These schedules inherit certain features from the features of the underlying allocations. The theory is based on a few constructive proofs and otherwise on *reductiones ad absurdum*.

Proposition 1 (Existence and Basic Properties): a) For any tax revenue g>0 sufficiently small there exists a second-best optimum.

b) At any feasible allocation, $c^k>c^h$ is equivalent to $y^k>y^h$ and $c^k=c^h$ is equivalent to $y^k=y^h$.

c) At any feasible allocation, $w^k>w^h$ implies $u(c^k, \ell^k) \geq u(c^h, \ell^h)$.

Proof: a) The objective function is continuous and the set of feasible allocations is compact because (3) defines a closed subset of the compact set given by (2) and the requirements (c^h, ℓ^h)$\in \mathcal{C}$ for all h. Due to assumption A, at least person H chooses a strictly positive income in the absence of taxes so that, by continuity, some tax revenue can be raised without violat-

ing any self-selection constraint. Hence the set of feasible allocations is nonempty, and existence follows from the Weierstrass theorem.

b) If a pair (c^k, y^k) contained strictly more consumption than another pair (c^h, y^h) but the same or even a smaller amount of income (effort), nobody would accept the latter pair since utility is strictly increasing in c and strictly decreasing in ℓ. The converse conclusion and the second equivalence are also obvious.

c) From (3), $w^k > w^h$ and $u_\ell < 0$ we have $u(c^k, \ell^k) \geq u(c^h, y^h/w^k) \geq u(c^h, y^h/w^h) \equiv u(c^h, \ell^h)$ with the second inequality holding strictly if $y^h > 0$. ∎

A first-best tax schedule can raise any revenue as large as $\ell^{max} \Sigma w^h f^h$ where aggregate consumption and aggregate leisure vanish. This is not possible in the second-best context because such an allocation would violate the self-selection constraints. Hence, existence of an optimum can be shown only if the tax revenue requirement is sufficiently small. Of course it would be helpful to have an equally simple expression of maximum tax revenue in the second-best case, but there is none. We will return to this issue at the end of this section.

Following proposition 1b) consumption increases strictly monotonically in income which means that the marginal income tax rate (defined in (14) below) must be strictly smaller than one. This feature of optimal income tax schedules holds under the most general assumptions and allows writing $(c^k, y^k) > (c^h, y^h)$ without ambiguity: At any feasible allocation, the former pair must contain strictly more consumption *and* strictly more income. A further consequence of 1b) is that all persons with a strictly positive income enjoy a strictly positive consumption since $c^0 \geq 0$ and $y^0 = 0$. It is hard to derive further general characteristics of second-best optima (Brito et al. 1990) but this changes rapidly once the following mild assumption on individual preferences is made.

Assumption B: In the absence of tax distortions, consumption is a gross substitute for leisure; inequality (18) from the appendix implies $(u_\ell/u_c \, u_{c\ell} - u_{\ell\ell})\ell - u_\ell > 0$.

The assumption of gross substitutability is known from general equilibrium theory where it plays an important role in stability and uniqueness proofs. It requires that in the absence of tax distortions a person will choose more consumption and thus more income when the price of leisure, i. e. the wage rate, rises. Consequently the government can use income as a signaling device: The higher a person's income, the higher the wage rate will be. Because, as we shall see, the government wants to redistribute between persons with different wage rates rather than between persons with different incomes, this premise is fundamental to the very idea of income taxation. If it were violated there could be persons with different wage rates but identical incomes, making it impossible for the government to redistribute between these by means of an income tax. Note that assumption B does not preclude a backward bending labor supply and that superiority of consumption is sufficient for it to hold because then both the substitution and income effects of an increase in the wage rate are positive. Moreover the premise says nothing about a person's potential behavior in the presence of tax distortions. Assumption B corresponds to Mirrlees' (1971) assumption B but has been cast in terms of an empirically testable hypothesis.

Substituting $\ell = y/w$ in the utility function and selecting some $\bar{u}$ from the utility function's range, the equation $u(c, y/w) = \bar{u}$ defines a function $c(y)$ with derivative

$$c'(y) = -\frac{u_\ell(c(y), y/w)}{u_c(c(y), y/w)\,w} > 0 \,. \tag{4}$$

Assumption B implies that for all numbers $w > 0$ and all $y \in (0, w\ell^{max})$ the function $c(y)$ becomes flatter in y-c-space the higher the wage rate, for differentiating with respect to w yields

$$\frac{dc'(y)}{dw} = -\frac{u_{\ell\ell}(-y/w^2)u_c w - u_\ell(u_{c\ell}(-y/w^2)w + u_c)}{(u_c w)^2} = -\frac{(u_\ell/u_c\, u_{c\ell} - u_{\ell\ell})\ell - u_\ell}{u_c w^2} < 0\,. \tag{5}$$

A person with a higher wage rate gives up a smaller amount of leisure in order to earn an extra unit of income and therefore demands a smaller extra amount of consumption to remain on the same indifference curve. Seade (1982) introduced (5) as an assumption which he named *agent monotonicity*. From Figure 1 which illustrates the case of two persons indexed h−1 and h it is possible to derive properties of second-best optima using graphical arguments. This is the traditional technique.

Figure 1

The Problem in y-c-space.
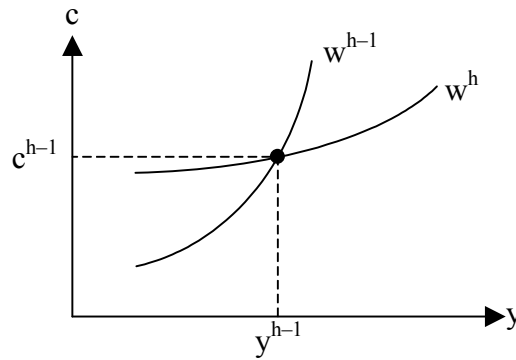
If one projects $u(c, y/w)$ into w-u-space rather than into y-c-space these graphical arguments can be replaced by analytical proofs. Such a projection is referred to as a *utility curve*. Points on a utility curve indicate the respective utility levels of persons with different wage rates who all have the same consumption and income. A utility curve is not an indirect utility function because consumption and income have not been chosen optimally; they are exogenous.

If $y = 0$ a utility curve is defined for all wage rates and runs parallel to the abscissa; every person has utility $u(c, 0)$ independent of the wage rate. If $y > 0$ a utility curve is defined for wage rates $w \geq y/\ell^{max}$ only because labor supplies are bounded from above. In this instance utilities are given by $u(c, y/w)$. The derivative $u_w(c, y/w) = -u_\ell(c, y/w)y/w^2 > 0$ shows that utility increases strictly in the wage rate. This is obvious because with consumption and income given, a rise in the wage rate simply means a reduction in labor supply. Differentiating anew shows that utility curves are strictly concave but this is immaterial and it will be more convenient to draw them as straight lines.

At any interior point with a given wage rate and a given utility level, utility curves become steeper the more consumption and income the pairs (c, y) contain. This has a pretty straightforward explanation: As stated above, the slope of the utility curve is positive because an increase in the wage rate reduces labor supply and thus increases leisure. This effect becomes

all the more pronounced, the more income a pair contains. To see this formally, differentiate the slope $u_w(c(y), y/w) = -u_\ell(c(y), y/w)\, y/w^2$ with respect to y, adjust c(y) in accordance with (4) and make use of assumption B to obtain

$$\frac{du_w(c(y), y/w)}{dy} = -\left(u_{\ell c}\left(-\frac{u_\ell}{u_c w}\right) + \frac{u_{\ell\ell}}{w}\right)\frac{y}{w^2} - \frac{u_\ell}{w^2} = \frac{(u_\ell/u_c\, u_{c\ell} - u_{\ell\ell})\ell - u_\ell}{w^2} > 0. \qquad (6)$$

This inequality allows developing an alternative notion of agent monotonicity which is very intuitive and which will be used in almost every proof below. Consider a person who is indifferent between two different pairs of consumption and income. Then *agent monotonicity*, as understood here, requires that all persons with a higher wage rate strictly prefer the pair with the higher income and vice versa. Under assumption B this is in fact true, as the following lemma shows. Figure 2 displays two utility curves associated with the pairs $(\bar{c}, \bar{y})$ and $(\underline{c}, \underline{y})$. The curves intersect at w^h, implying that person h is indifferent between the pairs. Obviously all persons with a higher wage rate strictly prefer $(\bar{c}, \bar{y})$ and all persons with a lower wage rate that can reach both pairs strictly prefer $(\underline{c}, \underline{y})$.

Figure 2
Agent Monotonicity
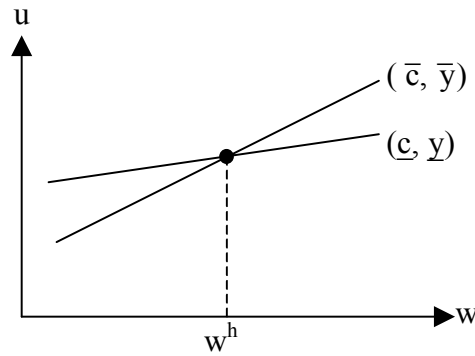
Lemma 1 (Agent Monotonicity): Assume $(\bar{c}, \bar{y}) > (\underline{c}, \underline{y})$ and let person h $(w^h \geq y/\ell^{max})$ be indifferent between these pairs. Then every person with wage rate $w > w^h$ strictly prefers $(\bar{c}, \bar{y})$ and every person with wage rate $w < w^h$ that can reach both pairs strictly prefers $(\underline{c}, \underline{y})$.

Proof: Consider the function $\Phi(w) = u(\bar{c}, \bar{y}/w) - u(\underline{c}, \underline{y}/w)$. By hypothesis $\Phi(w^h) = 0$. On the open interval $(\underline{y}, \bar{y})$ we have c(y) > 0, so that (6) holds and $u_w(c(y), y/w)$ increases strictly monotonically in income. Thus $\Phi'(w^h) = u_w(\bar{c}, \bar{y}/w^h) - u_w(\underline{c}, \underline{y}/w^h) > 0$, which implies $\Phi(w) > 0$ for $w > w^h$ and $\Phi(w) < 0$ for $w < w^h$ such that $w \geq y/\ell^{max}$. These two inequalities hold globally because the function Φ does not vanish at any other point. For assume there exists a smallest $w^* > w^h$ such that $\Phi(w^*) = 0$. By construction, $\Phi(w) > 0$ for all $w \in (w^h, w^*)$. If $\Phi(w)$ were strictly positive in a right-hand neighborhood of w^* then $\Phi'(w^*) = 0$, otherwise $\Phi'(w^*) \leq 0$, but both would contradict (6). An analogous argument for $w^* < w^h$ completes the proof. ∎

The statement takes account of the fact that the two pairs under consideration may be corner solutions, $\ell^h = 0$ or $\ell^h = \ell^{max}$, respectively. This causes no trouble because all pairs along the indifference curve between these are interior points, with intermediate values of labor and strictly positive consumption. As opposed to gross substitutability, which was assumed to hold only in the absence of distortionary taxes, agent monotonicity holds also in the presence

of such taxes. The following features of second-best optima, which have been established in the literature referred to in the introduction, give a first impression of how the agent monotonicity tool works.

Proposition 2 (Production Efficiency and Monotonicity): a) Any second-best optimum is production efficient in the sense that (3) holds with equality.

b) At any second-best optimum, $w^k > w^h$ implies $(c^k, y^k) \geq (c^h, y^h)$.

Proof: a) Assume there is an output surplus. If $\ell^H = \ell^{max}$ this can simply be used to increase c^H. No one else can reach y^H so that mimicking presents no problem, and person H becomes better off. If $\ell^H < \ell^{max}$, pairs $(\overline{c}, \overline{y}) > (c^H, y^H)$ exist such that $u(\overline{c}, \overline{y}/w^H) = u(c^H, \ell^H)$ and such that there is still an output surplus. Due to agent monotonicity, all persons $h < H$ that preferred their pairs weakly to (c^H, y^H) by hypothesis, prefer them strictly to $(\overline{c}, \overline{y})$ if they can reach this pair at all. Hence they will not mimic if the saved output surplus is used to make person H somewhat better off.

b) Assume $y^k < y^h$ which implies $c^k < c^h$ according to proposition 1b). Because k has a higher wage rate, both persons can reach both pairs. Everybody prefers (c^k, y^k) strictly to (c^k, y^h) but person h prefers (c^h, y^h) weakly to (c^k, y^k) by hypothesis. Hence, there exists $\overline{c} \in (c^k, c^h]$ such that h is indifferent between $(\overline{c}, y^h)$ and (c^k, y^k). Due to agent monotonicity, k must then prefer $(\overline{c}, y^h)$ strictly to (c^k, y^k) and must prefer (c^h, y^h) all the more, which contradicts (3). ∎

Production efficiency is an interesting feature in itself but will also play an important role in the following proofs. If, starting from some allocation, a reallocation is possible which satisfies all self-selection constraints, makes no one worse off and yields an output surplus, the original allocation cannot have been optimal. Figure 3 illustrates the monotonicity property. This states that persons with higher wage rates must be given pairs with higher consumption and income. Since the implication in proposition 2b) is a weak inequality, it may occur that persons with different wage rates obtain the same consumption and income. This is known as *bunching*.

Figure 3

Monotonicity and Bunching
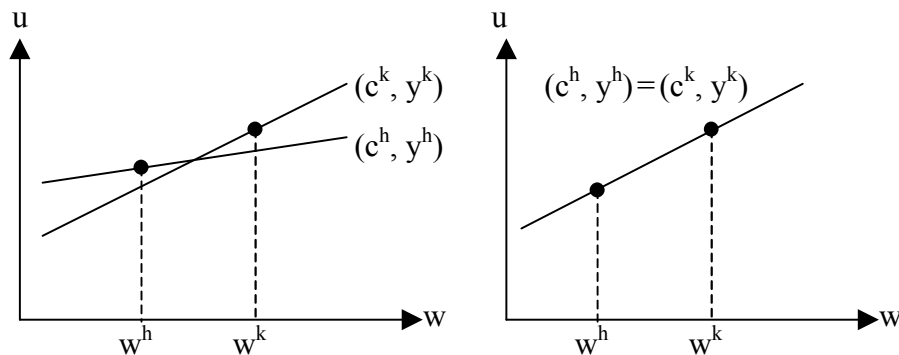
As yet we have assumed implicitly that two persons with identical wage rates will obtain identical pairs of consumption and income. Since Mirrlees (1971) this *equal treatment property* enters silently the formulation of the social objective and the resource constraint which would otherwise take a more complicated form. In the light of the theory of the core such a

presumption seems questionable because equal treatment is a necessary feature of core allocations only if all consumer types are of equal number (Hildenbrand and Kirman, 1988, p. 174). This requirement is not met in the present model which explicitly allows $f^k \neq f^h$ for $k \neq h$.
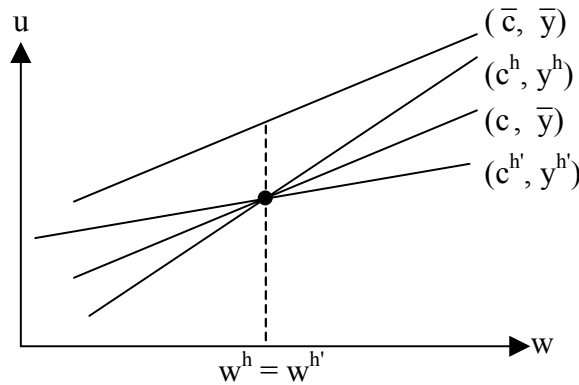
Proposition 3 (Equal Treatment Property): At any second-best optimum, $w^h = w^{h'}$ implies $(c^h, y^h) = (c^{h'}, y^{h'})$.

Proof: Assume $y^h > y^{h'}$ and $c^h > c^{h'}$ were second-best for some $w^h = w^{h'}$. The self-selection constraints imply $u(c^h, \ell^h) = u(c^{h'}, \ell^{h'})$. Set $\bar{c} = (c^h + c^{h'})/2$, $\bar{\ell} = (\ell^h + \ell^{h'})/2$ and $\bar{y} = (y^h + y^{h'})/2$. If persons h and h' are both given $(\bar{c}, \bar{\ell})$ they become better off as the preference ordering is strictly convex. Hence there exists $c < \bar{c}$ such that $u(c, \bar{\ell}) = u(c^h, \ell^h) = u(c^{h'}, \ell^{h'})$. If both persons are given $(c, \bar{\ell})$ their utilities remain unchanged and an output surplus emerges. Because person h is indifferent and $y^h > \bar{y}$, agent monotonicity implies that all persons with wage rates $w > w^h$ prefer their pairs strictly to $(c, \bar{y})$. The same is true for persons with wage rates $w < w^{h'}$ because person h' is indifferent and $y^{h'} < \bar{y}$. ∎

Figure 4 illustrates the proof. In the first step the two persons h and h' are made better off in that they are given convex combinations of their original bundles. But other persons, that possibly pay high taxes, may now be inclined to mimic h and h', making the new allocation infeasible. Therefore consumption is reduced in the second step so as to make h and h' as well off as before. Agent monotonicity implies that no one will mimic them. All utilities are unchanged but there is an output surplus which can be used profitably as described in proposition 2a). The existence of further persons with the same wage rates as h and h' does not matter because they must all have the same utilities as h and h', both before and after the change.

Figure 4

Equal Treatment Property

Because the equal treatment property did not enter the above proofs, proposition 3 renders this assumption valid and allows considering all persons with wage rate w^h as a *single* person h who has the weight f^h in the social objective and in the resource constraint. This yields semantic simplifications because we can refer to "consumption of person h", for instance, rather than to "consumption of every person with wage rate w^{h}".

The results derived so far hold under any Paretian ordering of social states. The optimum gets a more specific structure if one adds the assumption of non-inferiority of leisure whose relevance was first pointed out by Mirrlees (1974).

Assumption C: Leisure is a non-inferior good; inequality (19) from the appendix implies $u_{\ell c} - u_\ell / u_c\, u_{cc} \leq 0$.

With any additively separable utility function, assumptions B and C are both satisfied because the cross-derivative vanishes. If leisure is inferior, the cross-derivative $u_{\ell c}$ is strictly positive and in the absence of taxes, persons with higher wage rates will choose to work longer because the income and the substitution effect go into the same direction. Therefore persons with higher wage rates may have higher marginal utilities of consumption even though they enjoy more consumption. If a first-best tax system became introduced now – aiming at equalizing marginal utilities of consumption – it would not necessarily redistribute from top to bottom but could possibly redistribute in the reverse direction. The following lemma, proven in the appendix, rests on assumption C and precludes this weird scenario where the rich receive cash payments from the government which are financed by taxes on the poor.

Lemma 2 (Redistribution): If $u(c^k, \ell^k) > u(c^h, \ell^h)$ for $w^k > w^h$, then it is possible to increase $u(c^k, \ell^k) + u(c^h, \ell^h)$ by redistributing some leisure and/or consumption from k to h.

The result carries over to a situation where there are many persons of each type: Assume that redistributing some consumption $\varepsilon > 0$ from one person with wage rate w^k to one person with wage rate w^h increased the sum of utilities. When there are two persons with wage rate w^k instead, the sum of utilities still rises if we take consumption $\varepsilon/2$ from each of them and give it to the person with wage rate w^h. More generally the social objective can always be increased under the assumptions of lemma 2 by redistributing some small enough quantity from each person with wage rate w^k to each person with wage rate w^h. One must only select a quantity small enough that every person of type k is still better off than every person of type h before the last redistributive move takes place.

Figure 5
Asymmetry Property
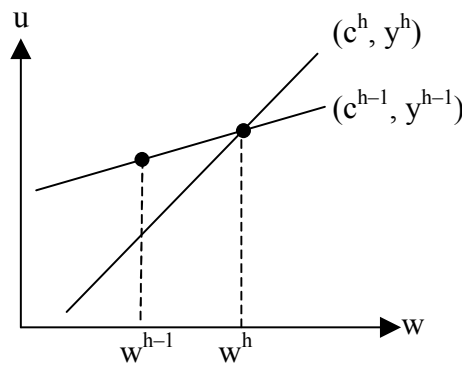
Before moving on we would like to point out two important features of feasible allocations. The first is referred to as the *asymmetry property*. It states that an upward self-selection constraint has slack if the corresponding downward self-selection constraint is binding and if the higher-skilled person has strictly more income. This is illustrated in figure 5 for an adjacent self-selection constraint. The figure makes clear that person h–1 will not mimic person h if the latter is made a bit better off. Asymmetry follows directly from agent monotonicity if one substitutes $(\bar{c}, \bar{y})$ for (c^h, y^h) and $(\underline{c}, \underline{y})$ for (c^{h-1}, y^{h-1}) in lemma 1. The second feature, referred to as the *transitivity property*, states that at any feasible allocation *all* downward self-

selection constraints are satisfied if only the *adjacent* downward constraints are satisfied. To see this, first consider a binding constraint $u(c^h, \ell^h) = u(c^{h-1}, y^{h-1}/w^h)$. Agent monotonicity implies that any person $k > h$ prefers (c^h, y^h) weakly to (c^{h-1}, y^{h-1}) because according to proposition 2b), the former pair contains at least as much income as the latter. As person k prefers his own pair to (c^h, y^h) by hypothesis, it follows that he also prefers it to (c^{h-1}, y^{h-1}), and even strictly so if the above constraint had some slack or if the two bundles were different. The significance of the asymmetry and the transitivity properties is that they allow neglecting all but the adjacent downward self-selection constraints. For any feasible allocation, the other constraints either have slack or hold tautologically if bunching prevails.

Proposition 4 (Chain Property): At any second-best optimum, all downward adjacent self-selection constraints are binding: $u(c^h, \ell^h) = u(c^{h-1}, y^{h-1}/w^h)$ for all $h > 0$.

Proof: Let k be the *largest* index such that $u(c^k, \ell^k) > u(c^{k-1}, y^{k-1}/w^k)$. By transitivity, this implies that persons with income y^k will not mimic persons with wage rates smaller than w^k if they are made a bit worse off. By asymmetry, they will also not mimic persons with incomes larger than y^k since all downward adjacent self-selection constraints above k are binding by construction.

Moreover, let h be the *smallest* index such that $u(c^{h+1}, \ell^{h+1}) > u(c^h, \ell^h/w^{h+1})$. By transitivity, this implies that persons with income y^h will not be mimicked by persons with wage rates higher than w^h if they are made a bit better off. By asymmetry, they will also not be mimicked by persons with incomes smaller than y^h since all downward adjacent self-selection constraints below h are binding by construction.

Each person with income y^k has a higher wage rate and a higher utility than each person with income y^h. Hence, the prerequisites of lemma 2 are fulfilled and the social objective can be increased by redistributing consumption ε or consumption and leisure of value ε from each person with income y^k to each person with income y^h, provided $\varepsilon > 0$ is small enough. It is insignificant if a person with income y^k mimics another with the same income after this redistribution because their commodity bundles all have identical values. The same applies for persons with income y^h. Because all other self-selection constraints and the resource constraint are satisfied, the original allocation could not have been optimal. Hence the assumed indices k and h do not exist. ∎

Figure 6

Chain Property
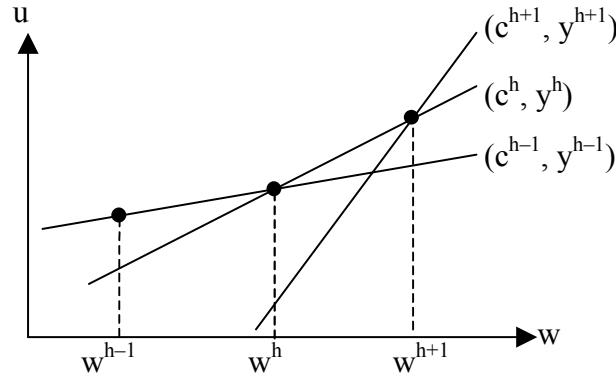
The proof takes account of two difficulties: Firstly, there may be bunching at income y^k meaning that several persons k ... $\bar{k}$ all have consumption c^k and income y^k at the outset. In this instance a number of upward adjacent self-selection constraints are binding and person k would immediately mimic person k+1 if he alone were made worse off. Therefore, all persons with income y^k must be made worse off and, conversely, all persons with income y^h must be made better off. Secondly, lemma 2 requires redistributing either consumption only or consumption and leisure together, depending on the sizes of the respective labor supplies. If there is bunching it may be necessary to practice both forms of redistribution simultaneously. In this instance persons with income y^k all having identical pairs at the outset receive different pairs after redistribution and one cannot exclude the possibility that, for instance, person k mimics person k+1. But this will not violate the resource constraint as all pairs have the same value after redistribution.

The chain property says that a second-best tax system redistributes from top to bottom until all downward adjacent self-selection constraints bind. This intuitive feature, proven under similar assumptions by Röell (1985), is illustrated in figure 6 and shows that each person is indifferent between his own pair and the pair of his left-hand neighbor. Combining with proposition 2b) we see that every second-best optimum satisfies the conditions

$$u(c^h, \ell^h) = u(c^{h-1}, y^{h-1}/w^h) \quad \text{and} \quad y^h \geq y^{h-1} \quad \text{for all } h > 0. \tag{7}$$

These necessary conditions also characterize solutions if the government's objectives differ from the objective considered here. With a *Rawlsian* objective, which requires maximizing the utility of person zero, the highest c^0 compatible with (7) will be selected. All downward adjacent self-selection constraints must bind, as otherwise the government could take some consumption from the person that prefers his pair strictly to that of his left-hand neighbor and give it to the persons with lower wage rates. Moreover, a *Leviathan* that aims at maximizing tax revenue will set c^0 to zero and will choose the other pairs in accordance with (7) because any slack in a downward adjacent self-selection constraint implied that a higher tax revenue could be extracted from the respective person. Therefore maximal tax revenue is characterized implicitly by $c^0 = 0$ and (7). Solving the related maximization problem yields the set of tax revenues compatible with existence of a second-best optimum under the present objective.

4. Implicit Marginal Tax rates

In the absence of distortionary taxes each person's marginal rate of substitution $-u_\ell/u_c$ will equal the wage rate at an interior solution. A positive marginal tax rate implies that the marginal rate of substitution falls short of the wage rate and vice versa. Therefore it seems natural to define *implicit marginal tax rates* for $h>0$ and $h<H$ as

$$L^h = 1 + \frac{u_\ell(c^h, \ell^h)}{u_c(c^h, \ell^h)w^h} \quad \text{and} \quad R^h = 1 + \frac{u_\ell(c^h, y^h/w^{h+1})}{u_c(c^h, y^h/w^{h+1})w^{h+1}}, \tag{8}$$

respectively. L^h denotes the implicit marginal tax rate of person h. R^h denotes the implicit marginal tax rate of person h+1 if the latter has the same income as the former. This occurs when person h+1 mimics person h or when there is bunching. Since (5) shows that the expression defined in (4) increases in the wage rate and since the latter becomes subtracted in the definitions of L^h and R^h, we obtain $R^h>L^h$ if y^h is strictly positive. Owing to the chain property the commodity bundles (c^h, ℓ^h) and $(c^{h-1}, y^{h-1}/w^h)$ must be on the same indifference curve of person h at an optimum. The former bundle contains weakly more consumption than the latter because of monotonicity. Hence the marginal rate of substitution is greater at (c^h, ℓ^h) and because it becomes subtracted we have $L^h \leq R^{h-1}$, holding as an equality only if $y^h=y^{h-1}$. To sum up, for all $h>0$ such that $y^h>0$ and $y^h>y^{h-1}$, we obtain

$$R^h>L^h \quad \text{and} \quad R^{h-1}>L^h, \tag{9}$$

respectively. The full meaning of the variables L^h and R^h will not become clear before the following section. The term "implicit tax rate" was coined by Weymark (1986) with respect to L^h whereas R^h has not been introduced explicitly as yet. Weymark has also derived results similar to the following for the special case of a quasi-linear utility function.

Proposition 5 (Implicit Marginal Tax Rates): At any second-best optimum

a) $L^h \geq 0$ for all h such that $y^h>y^{h-1}$,

b) $L^H=0$ if $\ell^H<\ell^{max}$ and

c) $L^h>0$ for all $h<H$ such that $y^h>y^{h-1}$.

Proof: a) The premise and proposition 1b) imply $y^h>0$, $\ell^h>0$ and $c^h>0$. Hence there exist numbers $\delta, \varepsilon>0$ such that $u(c^h-\varepsilon, \ell^h-\delta)=u(c^h, \ell^h)$. Because of agent monotonicity no person with wage rate $w>w^h$ will prefer the bundles constructed in this manner. Since the downward adjacent self-selection constraint holds with equality by the chain property and $y^h>y^{h-1}$, the upward self-selection constraints have slack so that no person with wage rate $w<w^h$ prefers the new bundles for sufficiently small δ. If L^h were strictly negative, person h's marginal rate of substitution would exceed the wage rate and it would be possible to set $\varepsilon>w^h\delta$. As this yields an output surplus we must have $L^h \geq 0$.

b) Because $\ell^H<\ell^{max}$, numbers $\varepsilon, \delta>0$ exist such that $u(c^H+\varepsilon, \ell^H+\delta)=u(c^H, \ell^H)$. Agent monotonicity implies that no person $h<H$ that can reach the new bundles will prefer them. If L^h were strictly positive, person H's marginal rate of substitution would fall short of the wage rate and it would be possible to set $\varepsilon<w^H\delta$. As this yields an output surplus we must have $L^H \leq 0$. Now suppose $y^H=y^{H-1}$. From (8) and (9) it follows that $L^H=R^{H-1}>L^{H-1} \geq 0$ but $L^H>0$

contradicts the result just found. Therefore we have $y^H > y^{H-1}$ so that result a) applies which gives $L^H \geq 0$. The two inequalities together imply $L^H = 0$.

c) Following the proof of a), commodity bundles $(c^h - \varepsilon, \ell^h - \varepsilon/w^h)$ exist that are preferred by no one else if $\varepsilon > 0$ is sufficiently small. The new bundles satisfy the resource constraint as the fall in aggregate consumption equals the fall in aggregate output. Now consider the utility $U(\varepsilon) = u(c^h - \varepsilon, \ell^h - \varepsilon/w^h)$ and the slack of the downward adjacent self-selection constraint $S(\varepsilon) = u(c^{h+1}, \ell^{h+1}) - u(c^h - \varepsilon, (w^h\ell^h - \varepsilon)/w^{h+1})$. Differentiating with respect to ε yields

$$U'(0) = -u_c(c^h, \ell^h) L^h \quad \text{and} \quad S'(0) = u_c(c^h, y^h/w^{h+1}) R^h. \tag{10}$$

If L^h equals zero an increase in ε leaves person h's utility unchanged up to first order but induces strictly positive slack in the self-selection constraint because $R^h > L^h$. This slack allows redistributing from persons with income y^{h+1} to persons with income y^h as described in the proof of the chain property, which increases the social objective according to lemma 2. Therefore $L^h = 0$ cannot be optimal, and combining with result a) it follows that $L^h > 0$. ∎
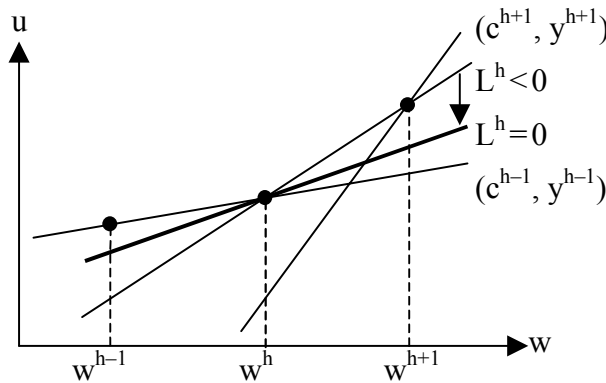
Figure 7

Non-negative Implicit Marginal Tax Rates

Figure 7 shows why a strictly negative implicit marginal tax rate can never be optimal. Starting from $L^h < 0$, person h's consumption and income are diminished; this corresponds to a clockwise rotation of the utility curve. Person h+1 that has been indifferent between (c^{h+1}, y^{h+1}) and (c^h, y^h) at the outset, now prefers his own pair strictly. Persons with wage rates below w^h originally preferred their own pairs strictly because of asymmetry and will stick to them if the rotation is small enough. Thus the new allocation satisfies all self-selection constraints. At the same time it yields an output surplus because a negative L^h implies that the marginal rate of substitution exceeds the wage rate which in turn allows a relatively strong reduction in consumption. Therefore a negative implicit marginal tax rate presents a useless waste of resources. It induces the familiar deadweight loss of taxation and has no incentive function because utility curves can always be rotated clockwise a bit without violating the incentive constraints.

Proposition 5b) utilizes the fact that the utility curve of the person with the highest wage rate can also be rotated *counter-clockwise*. Therefore a strictly positive tax rate at the top of the income distribution is also inefficient. Figure 6 makes clear why this argument does not apply to persons $h < H$: A counter-clockwise turn of any other utility curve induces mimicking of the respective right-hand neighbor. The famous "no distortion at the top" result first detected by

Phelps (1973) is not valid for a corner solution $\ell^H = \ell^{max}$, however. Corner solutions could be ruled out using an Inada condition $u_\ell(c, \ell)/u_c(c, \ell) \to -\infty$ for all $c > 0$ and $\ell \to \ell^{max}$ which implies that L^h approaches minus infinity as labor supply approaches its maximum. Because L^h must be non-negative according to proposition 5a) it follows that $\ell^h < \ell^{max}$ for all h. Substituting this into 5b) yields $L^H = 0$.

Proposition 5c) demonstrates that L^h will be strictly positive in general because it prevents the right-hand neighbor from mimicking. As person H has no such neighbor the argument is not valid in this case. Otherwise a strictly positive L^h facilitates additional redistribution which is a good thing according to lemma 2. The explanation of this fundamental result runs as follows: If one starts at $L^h = 0$, reduces person h's consumption by ε and increases his leisure by ε/w^h, the marginal rate of substitution declines which means that the implicit marginal rate becomes positive. The utility of person h remains unchanged up to first order. However, the utility of his right-hand neighbor declines in the case of mimicking because he also loses consumption ε, but only gains leisure ε/w^{h+1} which is smaller than ε/w^h. Hence the choice of person h is distorted because this distortion harms his right-hand neighbor more and can thus be used strategically to increase the degree of redistribution. This *manœuvre* works irrespective of whether persons h and h+1 have different or identical incomes at the outset. When $L^h = 0$ a possible bunching is broken up by introducing the tax distortion. But when $L^h > 0$ bunching cannot be excluded by a similar line of reasoning. Surely increasing L^h will separate the two persons but will also depress h's utility by a first order effect, and this loss cannot necessarily be counter-balanced by the redistributive gain.

Finally we should warn that proposition 5 says nothing about the sign of L^h at zero income. In the proofs new commodity bundles are constructed which contain less consumption and less income. At zero income such commodity bundles simply do not exist and the argument becomes void. This holds irrespective of whether the least skilled person has a zero wage rate, as assumed here, or has a strictly positive wage rate. Under the latter premise L^0 is defined and may have any sign if $y^0 = 0$. But if $y^0 > 0$, arguments analogous to those above imply $R^0 > L^0 > 0$. In this case, which seems less relevant empirically and has therefore been excluded via the assumption $w^0 = 0$, all numbers L^h must be strictly positive.

5. Second-best Tax Schedules

After these preparations it is easy to portray an attractive class of tax schedules supporting second-best optima. A *tax schedule* is a function T: $\mathcal{R}_{0+} \to \mathcal{R}$ which associates a positive or negative tax payment T(y) with any non-negative income. Negative tax payments are interpreted as welfare assistance. Subject to a given tax schedule each person solves

$$\max_{(c^h, \ell^h) \in \mathcal{C}}! \quad u(c^h, \ell^h) \quad \text{where} \quad c^h \leq y^h - T(y^h). \tag{11}$$

Every individual optimum satisfies $T(y^h) = y^h - c^h$ because of non-satiation. A tax schedule *supports* an allocation $(c^h, \ell^h)_{h=0...H}$ if each commodity bundle (c^h, ℓ^h) is chosen in accordance with (11). Because the schedule is defined for all non-negative incomes one must make sure that each person k not only prefers the pair (c^k, y^k) intended for him to all pairs (c^h, y^h) intended for somebody else but also prefers it to all (c, y) between these observable incomes.

Stiglitz (1987, p. 1003) argued that tax schedules for finite economies may be differentiable from the left, and Weibull (1989) showed that one can take them to be continuous. The following proposition provides a stronger characterization.

Proposition 6 (Tax Schedule): Any second-best optimum can be supported by a continuous tax schedule which has strictly positive derivatives from the left and the right at each income $y^h > 0$ ($h < H$) and is otherwise differentiable with $T'(y) > 0$ for $y < y^H$ and $T'(y^H) \geq 0$ for $y \geq y^H$.

Proof: The schedule is defined piecewise by $T(y) = y - c(y)$ for all $y \in [y^{h-1}, y^h]$ and all $y^h > y^{h-1}$ where the function $c(y)$ is defined on each such interval by

$$u(c(y), y/w^h) = u(c^h, \ell^h) = u(c^{h-1}, y^{h-1}/w^h). \tag{12}$$

The second equality follows from the chain property. Each person can reach the pair intended for him because $c(y^h) = c^h$ and $c(y^{h-1}) = c^{h-1}$ by construction. For every $y \in (y^{h-1}, y^h)$ any pair $(c(y), y)$ contains an income smaller than y^h and larger than y^{h-1}. Because person h is indifferent between these pairs and his own, agent monotonicity implies that every other person prefers his own pair strictly to all $(c(y), y)$ constructed this way. Hence the schedule in fact supports the optimum. On each open interval (y^{h-1}, y^h) the derivative $T'(y) = 1 - c'(y)$ exists since substituting $w = w^h$ into (4) yields

$$T'(y) = 1 + \frac{u_\ell(c(y), y/w^h)}{u_c(c(y), y/w^h)w^h}. \tag{13}$$

Definitions (8) imply $T'(y \downarrow y^{h-1}) = R^{h-1}$ and $T'(y \uparrow y^h) = L^h$. From (9) we have $R^{h-1} > L^h$ and L^h is non-negative according to proposition 5. Hence R^{h-1} and all values $T'(y)$ between L^h and R^{h-1} are strictly positive. Because $R^h > L^h$ the schedule is not differentiable at $y^h > 0$ ($h < H$). But with h denoting the smallest index such that $y^h > 0$ the derivative $T'(0) = R^{h-1} > L^h \geq 0$ exists. Setting $T(y) = (y - y^H)L^H + T(y^H)$ for all $y > y^H$ shows that $T'(y^H) = L^H \geq 0$ exists, too. ∎

The variables L^h and R^h turn out to be left-derivatives and right-derivatives, respectively. On the interval $[0, y^H]$ the above tax schedule is the point-wise minimum of all schedules supporting the optimum. Of course every schedule which prescribes the same tax payments at the observable incomes and higher payments otherwise supports the optimum, too, but neither has such appealing continuity and differentiability properties. In particular the present tax schedule is differentiable at $y = 0$. In proposition 5 we were unable to characterize what has now emerged as the left-derivative of the schedule at the bottom of the income distribution, but this left-derivative is meaningless because the schedule is only defined for non-negative incomes. According to (9) the right-derivative at $y = 0$ must be larger than the left-derivative at the smallest positive income, and the latter is non-negative according to proposition 5. Hence the marginal tax rate at the bottom must be strictly positive.

The proof of proposition 6 deliberately ignores persons with different wage rates but identical incomes and thus does not bring out what the tax schedule looks like in the case of bunching. Take h as the smallest and k as the largest index such that $y^h = y^k > 0$. The left-derivative and the right-derivative now correspond to L^h and R^k, respectively. At a bunching point, L^k and R^{k-1} are identical by definition. From (9) it follows that $R^k > L^k = R^{k-1} > ... > R^h > L^h$. So the non-differentiability at income y^h becomes more pronounced at a bunching point but it exists any-

way. Bunching at $y=0$ does not preclude differentiability because no left-derivative is defined at this point.
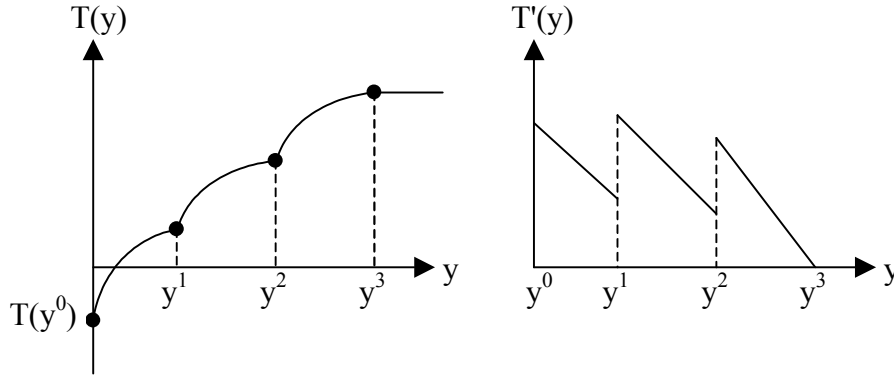
Figure 8

A Second-best Tax Schedule

Figure 8 shows a second-best tax schedule for an economy with four inhabitants. The marginal tax rate is strictly positive at the bottom of the income distribution and zero at the top. At the intermediate observable incomes y^1 and y^2 the schedule is not differentiable; the marginal tax rate jumps to a higher value. The marginal tax rates displayed on the right-hand side show a characteristic *saw-tooth pattern,* a direct consequence of (9). That second-best optima for finite economies cannot be supported by tax schedules which are differentiable on their entire domain, becomes evident from figure 1: If one adds an arbitrary straight line to the figure which runs through the point (c^{h-1}, y^{h-1}) there will always be some points on the line which are preferred by either person h or person h–1.

From integrating $T'(y)>0$ over an interval (y^{h-1}, y^h) it follows immediately that the *discrete marginal tax rate,* defined in the following equation, must be strictly positive for all $y^h>y^{h-1}$:

$$m^h = \frac{T(y^h) - T(y^{h-1})}{y^h - y^{h-1}}. \tag{14}$$

Thus in a finite economy there is a real zoo of marginal tax rates: L^h and R^h are implicit marginal tax rates which represent the left- and right-derivatives at the observable incomes. The $T'(y)$ are ordinary derivatives at the unobservable incomes. Finally, the m^h are discrete marginal tax rates indicating which additional tax payments are due when one moves up the income scale.

Following the traditional distinction between equity and efficiency objectives, the discrete marginal tax rate has the function of redistributing wealth whereas the local marginal tax rates have an incentive function. In a sense there are more tax instruments in a finite model than in the continuum case. For instance, if all left-derivatives vanished, the discrete marginal rates would nevertheless be strictly positive, though this would not be optimal. In the continuum model these different instruments collapse into one.

6. A Limit Theorem

The preceding section has shown that second-best tax schedules for finite economies cannot be differentiable throughout. This result is somewhat irritating for two reasons. Firstly, differentiability or at least differentiability from the left (Sadka 1976, p. 266) is habitually assumed in the continuum model. Secondly, most tax schedules used in reality are either differentiable or have only a small number of points of non-differentiability. One conjectures that the tax schedules presented above become smooth as the skill distribution grows dense.

To make this notion precise, assume a limit probability distribution function F with a strictly positive density on its support $[0, \overline{w}]$, where $\overline{w} > 0$. For $n = 1, 2, \ldots$ let $\mathcal{W}_n$ denote a partition of the support, i. e. a finite set of wage rates such that $0 = w^0 < \ldots < w^H = \overline{w}$. Subject to a chosen partition, the probability masses are given by $f^0 = F(0)$ and $f^h = F(w^h) - F(w^{h-1})$ for all $h > 0$. A finite economy is a quadruple $\mathcal{E}_n = (\mathcal{W}_n, F, u, g)$ consisting of a partition, the distribution function, the utility function and a specific aggregate tax revenue. A *sequence of increasingly fine economies* is a sequence $(\mathcal{E}_n)_{n=1, 2, \ldots}$ such that $w^h - w^{h-1} \to 0$ for all $h > 0$.

When the economies become increasingly fine, the number H grows without bound and the limit distribution function is approximated by a sequence of step functions. The associated tax schedules, however, are *not* step functions but possess the properties described in proposition 6. The length of the largest bunching interval is defined as

$$\beta(\mathcal{E}_n) = \max \{w^k - w^h : w^k, w^h \in \mathcal{W}_n \text{ and } y^k = y^h > 0\}. \tag{15}$$

This number vanishes if there is no bunching at any strictly positive income, otherwise it is strictly positive. Observe that bunching at $y = 0$ plays no role in the definition.

The *distance* of a particular economy from an economy with a differentiable tax schedule is denoted as $\Delta(\mathcal{E}_n)$. In the preceding section it has become clear that the tax schedules are differentiable almost everywhere and that, where this is not the case, finite left- and right-derivatives exist. Hence it is natural to define the said distance as the largest difference between the one-sided derivatives or, to put it more vividly, as the largest jump in the marginal tax rate schedule:

$$\Delta(\mathcal{E}_n) = \max \{R^k - L^h : w^{k+1}, w^h \in \mathcal{W}_n \text{ and } y^k = y^h > 0\}. \tag{16}$$

The distance is zero if $y = 0$ and $y^H > 0$ are the only observable incomes, because according to proposition 6 the tax schedule is always differentiable at the lowest and the highest income. Apart from this exception $\Delta(\mathcal{E}_n)$ must be strictly positive. Its limit behavior depends crucially on whether or not the length of the largest bunching interval converges to zero.

Proposition 7 (Limit Theorem): If a unique second-best optimum exists for each element of a sequence of increasingly fine economies, then $\Delta(\mathcal{E}_n) \to 0$ if and only if $\beta(\mathcal{E}_n) \to 0$.

Proof: The premise $\beta(\mathcal{E}_n) \to 0$ implies $w^k \to w^h$ for all $w^k, w^h \in \mathcal{W}_n$ where $y^k = y^h > 0$. As the economies become increasingly fine, $w^{k+1} \to w^k$ for all $w^{k+1}, w^k \in \mathcal{W}_n$, thus $w^{k+1} \to w^h$ for all $w^{k+1}, w^h \in \mathcal{W}_n$ where $y^k = y^h > 0$. Substituting into (8) shows that the differences $R^k - L^h$ become arbitrarily small, hence $\Delta(\mathcal{E}_n) \to 0$. Conversely, with bunching in the limit $R^k - L^h$ will be bounded away from zero for some k and h. ∎

This proposition renders the usual differentiability assumption of the continuum model valid in certain cases. If the tax schedules converge toward some limit and if there is no bunching in the limit, then the limit tax schedule will be differentiable on its entire domain. But in case of bunching even the limit tax schedule will be non-differentiable. From the proposition it follows that if there is exactly one bunching point in the limit, say at income y, then the limit schedule will be non-differentiable at y but will be differentiable at all other income levels. An example of non-differentiability at a single point was provided by Ebert (1992). The proof of the limit theorem does not require y^{k+1} to converge to y^h and hence holds irrespective of whether there are gaps in the limit income distribution.

7. Illustrations

A distinct advantage of the finite approach to optimal income taxation is that it allows calculating optima for arbitrary utility functions and skill distributions using standard optimization algorithms or a simple grid search. By contrast, the continuum approach has produced results for only a small number of utility functions (Tuomala 1986) because computations turn out to be hard; explicit solutions are available for quasi-linear utility functions only (Lollivier and Rochet, 1983).

Assume a Cobb-Douglas utility function $u(c, \ell)=[c(500-\ell)]^{0.4}$ (which increases strictly monotonically only in the interior of the consumption space, but this does not matter), an initial labor endowment ℓ^{max} of 500 working hours per month, and a per capita tax revenue of 100, implying that the sum of *positive* tax payments makes up roughly a fifth of national income in the following examples. The wage rates are distributed uniformly (the probability masses f^h all being equal to 20 per cent) unless stated otherwise, and their respective values are shown in the first columns of the following tables.

Table 1
The Falling Marginal Tax Rate

w	c	y	T	L	R	m
0	861	0	-861	--	57%	--
4	962	209	-753	46%	75%	52%
8	1548	1644	96	34%	64%	59%
12	2343	3122	779	19%	52%	46%
16	3381	4619	1239	0%	--	31%

The second column in Table 1 displays monthly consumption (net income) and the third gross income. The fourth column gives the difference between these, the tax payment. The last columns report the respective marginal tax rates. As one infers from the third and forth column, the average tax burden T/y increases strictly with income so that the tax schedule is progressive in this sense. But the left-derivatives L^h fall from 46% over 34% and 19% to zero. Slemrod et al. (1994) found a corresponding result for a two-bracket linear income tax. The

falling marginal tax rate is also present in Mirrlees' (1971) original and most subsequent simulations; counter-examples were constructed by Kanbur and Tuomala (1994) and Diamond (1998). An economic explanation of the falling marginal tax rate has never been offered, which is a pity since it contradicts the idea of a graduated income tax (though it agrees with the high implicit rates characterizing most contemporary transfer schemes).

In the several hundred simulations we have run, the falling marginal tax rate L^h has turned out to be a fairly robust feature. Proposition 5c) demonstrated that starting from an undistorted choice $L^h=0$ (h<H) and increasing the marginal tax rate a bit keeps the right-hand neighbor from mimicking and thus allows additional redistribution from person h+1 to person h. This story sufficed to prove the positivity of L^h but it has a natural continuation: Once person h+1 has been made worse off by the extra redistribution, additional slacks in the subsequent downward constraints will emerge, making it possible to also levy higher taxes on persons h+2, h+3, ... Hence the social value of distorting a person's choice becomes all the smaller the fewer persons with higher skills exist. Because the number of such persons decreases as one moves up the income scale, the left-derivative will decrease as a rule. But there are important exceptions to this rule. Since the optimization process weighs redistributive gains against deadweight losses, the marginal tax rate will be low if a person reacts particularly elastically or if the probability mass f^h is large.

The results reported in Table 2 rest on unchanged assumptions, except that the tax schedule was required to be linear. Owing to the complexity of the nonlinear income tax, linear income taxation has found some interest in the literature (Svensson and Weibull 1986). The required linearity has bold consequences. Comparing the tables one sees, for instance, that the consumption of the least skilled person falls from 861 to 553. More generally, low-skilled persons are worse off and high-skilled persons are better off under the linear income tax as compared with the nonlinear tax. This is due to the fact that with a linear tax schedule – as with *any* differentiable tax schedule – the downward adjacent self-selection constraints have slack, so that there is still room for further redistribution.

Table 2

An Optimal Linear Tax Schedule

w	c	y	T	L	m
0	553	0	-553	--	--
4	876	540	-336	40%	40%
8	1476	1539	63	40%	40%
12	2076	2539	463	40%	40%
16	2676	3539	863	40%	40%

Returning to nonlinear taxation, the setting behind Table 3 is the same as before except that f^1 has been increased to 36% and f^2 decreased to 4%, correspondingly. Bunching at gross income 388 emerges. The left-derivative at this income level is 34% and the right-derivative is 81% whereas the figures $R^1=L^2=71\%$ are meaningless implicit rates. The reason why bunching is optimal in this instance should be obvious: As there are many persons with wage rate 4, the aggregate deadweight loss from distortionary taxation is high. Hence a low tax rate seems

appropriate whereas keeping the persons with wage rate 8 away from this income level would require a high tax rate. The optimizing process accepts that the latter persons mimic the former, because the redistributive loss from doing so is insignificant. Thus bunching can be optimal in particular if adjacent probability masses differ substantially.

Table 3

The Social Value of Bunching

w	c	y	T	L	R	m
0	852	0	-852	--	57%	--
4	1057	388	-668	34%	71%	47%
8	1057	388	-668	71%	81%	71%
12	2170	3267	1097	21%	54%	61%
16	3205	4795	1591	0%	--	32%

Table 4 presents the last example with Cobb-Douglas utility. It was constructed by filling all integer values between skill levels zero and sixteen, leaving the rest of the assumptions behind Table 1 unchanged. Whereas "fighting unemployment" is usually considered an important business of social policy, the present case shows that it has a price: In order to induce the less skilled persons to work one must offer them attractive pairs containing more consumption and income. Then agent monotonicity implies that the higher skilled – who have to be prevented from mimicking – become better off. The present optimum accepts a certain degree of unemployment in order to make more redistribution possible. Somewhat paradoxically, insisting on full employment would reduce utilities at the bottom of the income scale and increase utilities at the top.

In accordance with the warning after proposition 5, there is a strictly negative left-derivative at zero income (-39% in Table 4). This figure is meaningless; the explicit marginal tax rate at income zero equals $R^2=54\%$. Starting at the first meaningful left-derivative ($L^3=53\%$) we encounter again the nice regularity of the falling marginal tax rate. Note also that the largest difference between left- and right-derivatives at strictly positive incomes was thirty points in Table 1 but is only twelve points in Table 4. This illustrates the limit theorem which states that the differences will become smaller and smaller as we add more and more skill levels to the original distribution.

Table 4

The Social Value of Unemployment

w	c	y	T	L	R	m
0	695	0	-695	--	-39%	--
1	695	0	-695	-39%	30%	--
2	695	0	-695	30%	54%	--
3	698	5	-693	53%	65%	53%
4	820	303	-517	52%	63%	59%
5	957	617	-340	49%	60%	56%
6	1109	943	-166	46%	57%	53%
7	1275	1277	2	43%	53%	50%
8	1457	1618	160	39%	49%	46%
9	1656	1963	307	35%	45%	43%
10	1871	2312	441	30%	41%	38%
11	2103	2664	561	26%	37%	34%
12	2353	3018	665	21%	32%	29%
13	2622	3375	754	16%	28%	25%
14	2909	3734	824	11%	23%	20%
15	3217	4094	877	6%	18%	15%
16	3545	4455	910	0%	--	9%

The last two tables were produced with a CES utility function $u(c, \ell)=[c^\rho+(500-\ell)^\rho]^{1/2\rho}$, all other premises the same as above. The results sustain the widely held expectation that optimal marginal tax rates would become smaller as the elasticity of substitution $\sigma=(\rho-1)/\rho$ increases. Stern (1976) showed this in the case of linear income taxation. With a CES utility function the continuum approach fails to provide examples of nonlinear income taxes because the differential equations cannot be solved. The finite approach avoids such roundabout procedures and allows optimizing directly with any conceivable utility function. An interesting route for future research would be to take some flexible utility function, fit it to the data and calculate optimal tax schedules for economies where individuals are grouped into finitely many classes.

Table 5

Elasticity of Substitution equals 2

w	c	y	T	L	m
0	1653	0	-1653	--	--
4	2237	1012	-1225	25%	42%
8	3309	3337	29	21%	54%
12	4363	5513	1151	14%	52%
16	5461	7659	2198	0%	49%

Table 6

Elasticity of Substitution equals 0.5

w	c	y	T	L	m
0	507	0	-507	--	--
4	544	128	-416	66%	71%
8	764	917	153	51%	72%
12	1086	1635	549	26%	55%
16	1456	2177	721	0%	32%

8. Conclusion

This paper presented an axiomatic restatement of the standard approach to nonlinear income taxation. The present approach differed from the standard model invented by Mirrlees (1971) in that it assumed a finite number of taxpayers rather than an uncountable infinity. All other assumptions were basically the same. The finite framework allowed proving the standard results in a simply and rigorous fashion which highlighted the economic forces at work. By contrast, the infinite approach is not rigorous since the usual optimality conditions are neither necessary nor sufficient for a true optimum.

The most important economic insight was that positive marginal tax rates have an important incentive function rather than merely a redistributive function: High marginal tax rates prevent more productive people from mimicking the less productive. As a result, marginal tax rates are likely to decline over the entire income range. This finding contradicts the traditional idea of income tax graduation.

Appendix

Firstly, the inequalities introduced with assumptions B and C are derived. Consider a person that maximizes utility $u(c, \ell)$ subject to the budget constraint $c \leq w\ell + e$, where e denotes a virtual income. Differentiating the Lagrangean $\mathcal{L} = u(c, \ell) + \lambda(w\ell + e - c)$ yields the usual first-order conditions for an interior individual optimum:

$$w\ell + e - c = 0,$$
$$u_c - \lambda = 0, \tag{17}$$
$$u_\ell + \lambda w = 0.$$

If consumption is a *gross substitute* for leisure, an increase in w will induce an increase in consumption. If leisure is *non-inferior,* an increase in e will induce an increase in leisure or, what amounts to the same, a reduction in labor supply. Both assumptions should hold in the entire interior of the consumption space but not necessarily on its boundary. The determinant of the Jacobian matrix J of (17) is strictly positive according to assumption A. From the implicit function theorem we obtain after substituting $\lambda = u_c$ and $w = -u_\ell/u_c$:

$$\frac{\partial c}{\partial w} = \frac{(u_\ell/u_c\, u_{c\ell} - u_{\ell\ell})\ell - u_\ell}{|J|} > 0 , \tag{18}$$

$$\frac{\partial \ell}{\partial e} = \frac{u_{\ell c} - u_\ell/u_c\, u_{cc}}{|J|} \le 0 . \tag{19}$$

These two inequalities are used in the text. In order to motivate the next proof observe that λ equals the marginal utility of consumption. Hence changes in the marginal utility of consumption following an increase in e or w, respectively, can again be calculated using the implicit function theorem. First we have $\partial u_c/\partial e = -|\text{Hess } u|/|J| < 0$ which says that the marginal utility of consumption is smaller for persons with a higher exogenous income. This follows directly from the strict concavity of the utility function and is pretty clear. However,

$$\frac{\partial u_c}{\partial w} = \frac{\partial u_c}{\partial e}\, \ell + \frac{\partial \ell}{\partial e}\, u_c . \tag{20}$$

Thus marginal utilities need not be smaller for persons with higher wage rates, see also Christiansen (1983, p. 367). But when leisure is a non-inferior good the second term will be non-positive and since the first term is strictly negative, the expected result is obtained.

Proof of lemma 2: Let $c(z) = zc^k + (1-z)c^h$ and $\ell(z) = z\ell^k + (1-z)\ell^h$. The strict concavity of the utility function implies $u(c(z), \ell(z)) > zu(c^k, \ell^k) + (1-z)u(c^h, \ell^h)$ for all $z \in (0; 1)$. An analogous inequality holds if one substitutes $1-z$ for z. Adding the two inequalities yields

$$u(c(z), \ell(z)) + u(c(1-z), \ell(1-z)) > u(c^k, \ell^k) + u(c^h, \ell^h) . \tag{21}$$

Assume $\ell^k < \ell^h$. Because proposition 2b) implies $c^k \ge c^h$, consumption of k decreases and consumption of h increases; aggregate consumption remains unchanged. As person k has a higher wage rate aggregate output rises. Reducing person k's work to $\ell^k + w^h/w^k (1-z)(\ell^h - \ell^k) < \ell(z)$ increases overall utility anew and leaves output unchanged. Thus consumption and leisure of value $\varepsilon = (1-z)(c^k - c^h) + (1-z)w^h(\ell^h - \ell^k) > 0$ are distributed from k to h.

Now assume $\ell^k \ge \ell^h$. Let $\Delta c = c^k - c^h$ and $\Delta\ell = \ell^k - \ell^h$. The functions $c(z)$ and $\ell(z)$ defined above possess derivatives $c'(z) = \Delta c$ and $\ell'(z) = \Delta\ell$. The premise $u(c^k, \ell^k) > u(c^h, \ell^h)$ and strict convexity of the preference ordering imply $u(c(z), \ell(z)) > u(c^h, \ell^h)$ for all $z \in (0; 1]$ so that the derivative of the utility function with respect to z is strictly positive almost everywhere:

$$\frac{du(c(z), \ell(z))}{dz} = u_c\Delta c + u_\ell\Delta\ell > 0 \quad \Rightarrow \quad \Delta c > -u_\ell/u_c\Delta\ell . \tag{22}$$

All derivatives are evaluated at $(c(z), \ell(z))$. Differentiating the marginal utility of consumption with respect to z yields

$$\frac{du_c(c(z), \ell(z))}{dz} = u_{cc}\Delta c + u_{c\ell}\Delta\ell . \tag{23}$$

Since $u_{cc} < 0$ the right-hand inequality in (22) is equivalent to $u_{cc}\Delta c < -u_{cc}\, u_\ell/u_c\Delta\ell$. Hence

$$\frac{du_c(c(z), \ell(z))}{dz} < -u_{cc}\, u_\ell/u_c\, \Delta\ell + u_{c\ell}\Delta\ell = (u_{\ell c} - u_\ell/u_c\, u_{cc})\Delta\ell \le 0 . \tag{24}$$

The last inequality follows from $\Delta\ell \ge 0$ and (19). Thus du_c/dz is strictly negative so that

$$u_c(c^k, \ell^k) - u_c(c^h, \ell^h) = \int_0^1 \frac{du_c(c(z), \ell(z))}{dz} dz < 0. \tag{25}$$

This implies $u(c^k - \varepsilon, \ell^k) + u(c^h + \varepsilon, \ell^h) > u(c^k, \ell^k) + u(c^h, \ell^h)$ for $\varepsilon > 0$ small enough. As person k has higher utility and $\ell^k \geq \ell^h$ we must have $c^k > c^h \geq 0$ implying that such reallocations, which leave aggregate consumption unchanged, are indeed feasible. ∎

The proof shows that one must follow different redistribution strategies depending on which labor supply is greater. If $\ell^k < \ell^h$, which may well occur at a second-best optimum, the sum $u(c^k, \ell^k) + u(c^h, \ell^h)$ can be increased by distributing consumption *and* leisure from k to h. But if $\ell^k \geq \ell^h$ the same goal is reached if one distributes consumption *only* from k to h. Dixit and Seade (1979) pointed out that when one moves up the utility scale, at least one marginal utility must decline if both goods are non-inferior. The premise of non-inferiority of consumption is not needed here. Dixit and Seade's conclusion is too weak for our purposes because at corner solutions $\ell^k = \ell^{max}$ or $\ell^h = 0$ it becomes impossible to redistribute leisure from person k to person h. Both corner solutions imply $\ell^k \geq \ell^h$, however, and the second part of the above proof shows that redistribution of consumption only suffices in this case.

REFERENCES

Bertsekas, Dimitri P. (1999), Nonlinear Programming, Belmont.

Boadway, R. and M. Keen (1993), Public Goods, Self-selection and Optimal Income Taxation, International Economic Review 34, 463–478.

Brito, D. L. and W. H. Oakland (1977), Some Properties of the Optimal Income Tax, International Economic Review 18, 407–423.

Brito, D. L., J. H. Hamilton, S. M. Slutzky and J. E. Stiglitz (1990), Pareto Efficient Tax Structures, Oxford Economic Papers 42, 61–77.

Brunner, J. K. (1989), Theory of Equitable Taxation, Berlin.

Christiansen, V. (1983), Some Important Properties of the Social Marginal Utility of Income, Scandinavian Journal of Economics 85, 359–371.

Diamond, P. A. (1980), Income Taxation with Fixed Hours of Work, Journal of Public Economics 13, 101–110.

Diamond, P. A. (1998), Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Rates, American Economic Review 88, 83–95.

Dixit, A. and J. Seade (1979), Utilitarian versus Egalitarian Redistribution, Economics Letters 4, 121–124.

Ebert, U. (1992), A Reexamination of the Optimal Nonlinear Income Tax, Journal of Public Economics 49, 47–73.

Guesnerie, R. and J. Seade (1982), Nonlinear Pricing in a Finite Economy, Journal of Public Economics 17, 157–179.

Harsanyi, J. C. (1955), Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility, Journal of Political Economy 63, 309–321.

Hildenbrand, W. and A. P. Kirman (1988): Equilibrium Analysis, Amsterdam.

Kanbur, R. and M. Tuomala (1994), Inherent Inequality and the Optimal Graduation of Marginal Tax Rates, Scandinavian Journal of Economics 96, 275–282.

Lollovier, S. and J.-C. Rochet (1983), Bunching and Second-Order Conditions: A Note on Optimal Tax Theory, Journal of Economic Theory 31, 392–400.

Matthews, S. and J. Moore (1987), Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening, Econometrica 55, 441–467.

Mirrlees, J. A. (1971), An Exploration in the Theory of Optimum Income Taxation, Review of Economic Studies 38, 175–208.

Mirrlees, J. A. (1974), Notes on Welfare Economics, Information and Uncertainty, in: M. Balch, D. McFadden and S. Wu (eds.) Essays on Economic Behaviour under Uncertainty, Amsterdam, 243–261.

Phelps, E. S. (1973), The Taxation of Wage Income for Economic Justice, Quarterly Journal of Economics 87, 331–354.

Rawls, J. (1971), A Theory of Justice, Cambridge.

Röell, A. (1985), A Note on the Marginal Tax Rate in a Finite Economy, Journal of Public Economics 28, 267–272.

Sadka, E. (1976), On Income Distribution, Incentive Effects and Optimal Income Taxation, Review of Economic Studies 43, 261–267.

Seade, J. K. (1982), On the Sign of the Optimum Marginal Income Tax, Review of Economic Studies 49, 637–643.

Slemrod, J., S. Yitzhaki, J. Mayshar and M. Lundholm (1994), The Optimal Two-Bracket Linear Income Tax, Journal of Public Economics 53, 269–290.

Stern, N. (1976), On the Specification of Models of Optimum Income Taxation, Journal of Public Economics 6, 123–162.

Stiglitz, J. E. (1982), Self-selection and Pareto Efficient Taxation, Journal of Public Economics 17, 213–240.

Stiglitz, J. E. (1987), Pareto Efficient and Optimal Taxation and the New Welfare Economics, in: Auerbach and M. S. Feldstein (eds.) Handbook of Public Economics Vol. II, North-Holland, 991–1042.

Strawczynski, M. (1998), Social Insurance and the Optimum Piecewise Linear Income Tax, Journal of Public Economics 69, 371–388.

Svensson, L.-G. and J. W. Weibull (1986), An Upper Bound on Optimal Income Taxes, Journal of Public Economics 30, 165–181.

Tuomala, M. (1984), On the Optimal Income Taxation, Journal of Public Economics 23, 351–366.

Vickrey, W. (1960), Utility, Strategy, and Social Decision Rules, Quarterly Journal of Economics 74, 507–535.

Wane, W. (2001), The Optimal Income Tax when Poverty is a Public ´Bad´, Journal of Public Economics 82, 271–299.

Weibull, J. W. (1989), A Note on the Continuity of Incentive Schedules, Journal of Public Economics 39, 239–243.

Weymark, J. A. (1986), A Reduced-Form Optimal Nonlinear Income Tax Problem, Journal of Public Economics 30, 199–217.

Stefan Homburg
Chair for Public Finance
University of Hannover
Königsworther Platz 1
30167 Hannover
Germany
homburg@fiwi.uni-hannover.de