

# A reexamination of the optimal nonlinear income tax

Udo Ebert\*

*Institute of Economics, University of Oldenburg, P.O. Box 2503, W-2900 Oldenburg, Germany*

Received July 1987, final version received May 1991

This paper investigates Mirrlees' model of optimal income taxation. It provides a concrete example of utility and density functions for which the solution to the usual (first-order) model is not implementable, i.e. an example where the first-order approach does not work. Adding second-order conditions leads to an extended model and to implementable solutions. If these conditions are binding one gets a kink in the optimal net-income schedule and bunching of individuals occurs. The properties of an optimal nonlinear income tax are reexamined within the extended model.

## 1. Introduction

This paper deals with the derivation and the properties of an optimal nonlinear income tax. These questions were first investigated by the important paper of Mirrlees published in 1971. In this and later papers [Mirrlees (1971, 1976, 1986)] he provides an economic model which takes into account some essential aspects of optimal income taxation, namely the choice between labour and leisure and the different abilities of workers. Furthermore, he proposes a corresponding formal model and succeeds in describing its solution.

In the formal model the first-order conditions resulting from utility maximization are taken into consideration. The question arises whether this procedure is sufficient to guarantee the implementability of the tax system. Mirrlees (1976) investigates this problem. He concludes that a tax system is implementable if and only if the first-order conditions are satisfied and if the resulting optimal gross income is a nondecreasing function of ability. The second part ensures that more able individuals also earn higher incomes.

*Correspondence to:* U. Ebert, Institute of Economics, University of Oldenburg, P.O. Box 2503, W-2900, Oldenburg, Germany.

\*I thank Martin Hellwig for initiating my interest in optimal nonlinear income taxation and for stimulating suggestions and comments. Furthermore, I am grateful to Heinz-Jürgen Büchner, Wolfgang Peters, Georg Tillmann, Oskar von dem Hagen, and Hans-Georg Zimmermann for most valuable discussions and to two unknown referees for helpful comments. This paper is based on results presented in Ebert (1986a, 1986b).

Mirrlees apparently assumes that this condition is satisfied automatically. He states [Mirrlees (1976, p. 335)]: 'It is this circumstance that makes the income-tax problem studied in Mirrlees (1971) manageable'. In his earlier paper [Mirrlees (1971, pp. 182–183)] he deals with the topic as well and writes '... This [equation (30)] does not yet prove global maximization of utility, but that also is true.' On the other hand Mirrlees was the first who pointed out that the first-order approach, i.e. using only the first-order conditions, can lead to the wrong results [cf. the unpublished paper Mirrlees (1975) and Mirrlees (1986)].

The present paper is concerned with this issue. It compares the first- and second-order approaches. The usual formal model used by Mirrlees and others is extended by the *additional* condition that the optimal gross income is a nondecreasing function of ability (which is equivalent to the second-order conditions for utility maximization). Furthermore, we present a class of problems which is completely and explicitly solved within the framework of the first-order model. Using this solution we demonstrate by a concrete example that the first-order approach may be inappropriate. The resulting optimal income tax function is *not* always nondecreasing. Thus it cannot be implemented. This implies that the second-order approach is necessary to cope with the problem of optimal income taxation. Within this extended model it is now possible to give a rigorous explanation for the problem of bunching. In this case there are some individuals who possess different abilities, but choose the same combination of gross income and consumption. The phenomenon is connected with a kink in the net-income schedule if it occurs at an interior income. It is shown that bunching is necessary if and only if the optimal tax obtained in the usual Mirrlees model is not implementable (i.e. if the first-order approach is insufficient).

An extension of the first-order model was derived in Brito and Oakland (1977), Mirrlees (1986) and independently in Ebert (1986a). But Brito and Oakland did not prove that the second-order conditions are *necessary*. Their contribution seems to have been overlooked by the profession. To the best knowledge of the author there is no paper in the literature using their model. On the other hand, the example provided in this paper proves that the implementability of the tax system must be taken into account explicitly. Therefore the final section of this paper reexamines the properties of an optimal nonlinear income tax within the *extended* model. In order to obtain definite results, again the necessary conditions for the solution to the planner's optimization problem are considered.

It turns out that most of the results known about the income tax are also true in the second-order model. Furthermore, some new results can be derived. It is possible that bunching occurs at the lower end of the income scale at a positive gross income  $z$ . In this case the marginal tax rate is strictly positive at  $z$ . This changes the result of Seade's theorem 2 [Seade (1977)].

(Seade admits bunching only at  $z=0$ .) Furthermore, it can be proved as well that bunching cannot occur at the highest income level. Finally, we get the result that the marginal rate of tax is *strictly* positive at all 'interior' incomes, i.e. at all income levels but possibly the smallest or highest level. This result supports Seade's interpretation [Seade (1977, 1982)] that a nonzero marginal tax is a means to transfer income.

The plan of this paper is as follows. Section 2 and section 3 describe the various models and their solutions. Section 4 presents the explicit solution and an example. In section 5 the properties of an optimal nonlinear income tax are discussed and section 6 concludes the paper. All proofs are relegated to the appendices.

## 2. The model

This section provides a description of Mirrlees' model. The presentation is based on Mirrlees (1971, 1986), Brito and Oakland (1977), Roberts (1979), and Ebert (1986a).

There are two commodities, consumption  $x$  and time  $y$  spent on work. All individuals have the same preference ordering which can be represented by an ordinal utility function. To impose certain distributional preferences the government chooses a special cardinalization  $u(x, y)$ .  $u$  is increasing in consumption  $x$ , decreasing in labour  $y$ , strictly concave, and twice differentiable. We assume that consumption is a normal good and leisure a noninferior good.

Every individual possesses a specific ability (marginal productivity)  $n$ . Working  $y$  units of time he or she is able to provide  $y \cdot n$  (efficiency) units of labour and to produce  $y \cdot n$  units of the consumption good. We choose consumption as numeraire and can interpret  $n$  as wage rate and  $z = y \cdot n$  as gross income. An individual of type  $n$  maximizes his or her utility function  $U(x, z, n) = u(x, z/n)$  subject to the budget constraint. Defining the marginal rate of substitution of gross income for consumption by

$$s(x, z, n) := - \frac{U_z}{U_x} > 0,$$

we have the property of

*Agent monotonicity (AM)*

$$s_n := \frac{\partial}{\partial n} s(x, z, n) < 0.$$

It is an (ordinal) property of the common preference ordering which is implied by the fact that consumption is a normal good.<sup>1</sup> As a consequence of (AM) the indifference curves in an  $(x, z)$  diagram are flatter the higher is an individual's wage rate. For a linear income tax (AM) yields that more able individuals earn higher incomes [cf. Hellwig (1986)]. The distribution of abilities is described by a continuous density function  $f(n)$  which is assumed to be strictly positive on an interval  $[\underline{n}, \bar{n}]$ ,  $\underline{n} > 0$ . Each individual is able to earn income.

The government knows the distribution of  $n$  and the utility function  $U(x, z, n)$ . It cannot observe an individual's ability  $n$  or time worked  $y$ , but only his or her gross income  $z = n \cdot y$ . Given a tax schedule  $T: \mathbb{R} \rightarrow \mathbb{R}$  which denotes the tax liability, the government can calculate an  $n$  individual's net income  $x(n)$  and gross income  $z(n)$ . Using this information the government wants to maximize social welfare (given by the utilitarian welfare function) by choosing  $T(z)$

$$\int_{\underline{n}}^{\bar{n}} U(x(n), z(n), n) f(n) dn \quad (\text{SWF})$$

subject to its budget constraint<sup>2</sup>

$$\int_{\underline{n}}^{\bar{n}} T(z(n)) f(n) dn = \int_{\underline{n}}^{\bar{n}} (z(n) - x(n)) f(n) dn = 0. \quad (\text{BC})$$

Deriving an optimal income tax the government has to take into account that each individual chooses a commodity bundle which is optimal *for him or her* given the tax schedule. Therefore the individual's reactions have to be taken into consideration a priori. Mirrlees (1976) proposes the following method. Let  $(x(n), z(n))$  denote the optimal plan of an  $n$  individual (as computed by the government). The  $n$  individual chooses this plan if it is better than all other possible plans

$$U(x(n), z(n), n) \geq U(x(m), z(m), n) \quad \text{for all } m \in [\underline{n}, \bar{n}], n \in [\underline{n}, \bar{n}]. \quad (\text{SSC})$$

If these conditions (self-selection constraints) are met, the tax system is implementable. [Remember that  $z(n) - x(n)$  is identical to an  $n$  individual's

<sup>1</sup> Actually (AM) is weaker than normality of consumption.

<sup>2</sup> In this case all taxes are redistributed. Of course it is also possible to raise a tax revenue  $\bar{T} \neq 0$ .

tax liability.] Mirrlees derives a minimization problem from (SSC). It yields the following necessary first-order conditions:<sup>3</sup>

$$\frac{du(n)}{dn} = U_n(x(n), z(n), n) \quad (\text{FOC})$$

or

$$U_x \cdot x'(n) + U_z \cdot z'(n) = 0, \quad [\text{FOC}]$$

i.e. the optimal level of utility  $u(n)$  of an  $n$  individual must coincide with the partial derivative  $U_n$ . Here and in what follows we make the assumption that the solutions  $x(n)$ ,  $z(n)$ , and  $u(n)$  are continuously differentiable almost everywhere. Moreover, the second-order conditions are equivalent to

$$\frac{dz(n)}{dn} \geq 0, \quad (\text{SOC})$$

which states that gross income is a nondecreasing function of ability  $n$ . More able individuals have to earn higher incomes.

Now we are in a position to describe Mirrlees' model precisely.

**Problem P.** The government chooses an income tax schedule to maximize social welfare (SWF) subject to the conditions of the individual utility maximization (SSC) and its budget constraint (BC).

This is a formulation of the underlying *economic* problem. In order to provide solutions we consider two further versions.

**Problem P\*.** Problem P\* is the same as P, but the self-selection constraints (SSC) are replaced by the first-order conditions (FOC) and with the additional constraint that the optimal commodity bundles are continuously differentiable almost everywhere.<sup>4</sup>

Problem P\* is called the first-order version of the problem. We are mainly interested in the following second-order version, its implications, and its relationship to P\*.

**Problem P\*\*.** Problem P\*\* is the same as P\* with the additional constraint that income is nondecreasing in ability (SOC).

<sup>3</sup>In Mirrlees (1971) eq. (FOC) is *directly* interpreted as the necessary first-order condition from utility maximization.

<sup>4</sup>In the light of the monotonicity of  $z(n)$  this assumption seems to be acceptable. Weibull (1989) shows that  $x(n)$  and  $z(n)$  can be chosen as continuous functions.

In the next section we determine the necessary conditions for a solution to  $P^*$  and  $P^{**}$ .

### 3. The solution to $P^*$ and $P^{**}$

Both optimization problems can be formulated as a maximum principle problem. The solution to  $P^*$  is well known in the literature. Here Problem  $P^{**}$  is investigated since it is an extension to  $P^*$ . We choose the derivative of gross income  $w(n) = dz(n)/dn$  as the control variable. Gross income  $z(n)$ , the optimal level of utility  $u(n)$ , and the tax payments of all individuals that have an arbitrary  $n' \leq n$ , namely

$$R(n) = \int_n^n (z(n') - x(n')) f(n') dn',$$

are defined as state variables. Because of the properties of the utility function  $U$  (monotonicity!) it is possible to compute  $x(n)$  for given  $u(n)$ ,  $z(n)$ , and  $n$ , i.e. there exists a function  $h$  such that

$$x = h(u, z, n),$$

where  $u$  is a level of utility. Therefore  $x(n)$  should not be taken into account explicitly.

In order to avoid dealing with singular solutions we transform the second-order condition (SOC) by means of an arbitrary strictly increasing twice differentiable nonlinear function  $g$  which has the properties  $g(0) = 0$  and  $g'(t) > 0$ . The solution does not depend on the choice of  $g$ . Then we obtain the following control problem. [Remember that  $x(n) = h(u(n), z(n), n)$ !]

Find  $w(n)$ ,  $z(n)$ ,  $u(n)$ , and  $R(n)$  s.t.

$$\int_n^{\bar{n}} u(n) f(n) dn \rightarrow \max,$$

$$\frac{dz}{dn} = w(n), \tag{1}$$

$$\frac{du}{dn} = U_n(x(n), z(n), n), \tag{2}$$

$$\frac{dR}{dn} = (z(n) - x(n)) f(n), \tag{3}$$

$$R(\underline{n}) = R(\bar{n}) = 0, \quad (4)$$

$$g\left(\frac{dz(n)}{dn}\right) = g(w(n)) \geq 0. \quad (5)$$

Obviously all restrictions of  $P^{**}$  are taken into account. The budget constraint (BC) is reflected by the differential equation (3) and the boundary conditions (4) for  $R(n)$ . The first-order conditions (FOC) determine  $u(n)$  by eq. (2). The second-order conditions (SOC) are incorporated directly by means of (1) and (5).

Applying the maximum principle technique we define the Hamiltonian

$$\begin{aligned} H = & u(n) \cdot f(n) + \lambda(n)(z(n) - x(n))f(n) \\ & + \mu(n)U_n(x(n), z(n), n) + v(n) \cdot w(n) + \kappa(n)g(w(n)), \end{aligned}$$

where  $\kappa$ ,  $\lambda$ ,  $\mu$  and  $v$  are adjoint variables, and we get<sup>5</sup> the following necessary (first-order) conditions (which are listed for completeness):

$$\frac{\partial H}{\partial w} = 0 = v + \kappa g'(w),$$

$$\frac{\partial H}{\partial z} = -v' = \lambda \frac{\partial}{\partial z} [(z(n) - x(n))f(n)] + \mu \frac{\partial}{\partial z} U_n(x(n), z(n), n),$$

$$\frac{\partial H}{\partial u} = -\mu' = f(n) + \lambda \frac{\partial}{\partial u} [(z(n) - x(n))f(n)] + \mu \cdot \frac{\partial}{\partial u} U_n(x(n), z(n), n),$$

$$\frac{\partial H}{\partial R} = -\lambda' = 0,$$

$$\left. \begin{aligned} \mu(\underline{n}) = \mu(\bar{n}) = 0 \\ v(\underline{n}) = v(\bar{n}) = 0 \end{aligned} \right\}, \quad \text{transversality conditions,}$$

$$\kappa \geq 0,$$

$$\kappa > 0 \Rightarrow \frac{dz}{dn} = 0,$$

$$\frac{dz}{dn} > 0 \Rightarrow \kappa = 0.$$

If we define  $\phi(n) := \kappa(n)g'(w(n))$  and observe that  $h_z = -U_z/U_x = s$ ,  $h_u = 1/U_x$  and that  $\lambda$  is constant, some tedious but elementary rearrangements and substitutions yield the following conditions, which are partly more familiar:

<sup>5</sup>Compare Bryson and Ho (1969).

$$-\mu(n) \cdot U_x s_n + \lambda(1-s(n)) \cdot \underline{f(n) - \phi'(n)} = 0, \quad (\text{P1}^*)$$

$$\mu'(n) + \mu(n) \cdot \frac{U_{nx}}{U_x} + \left(1 - \frac{\lambda}{U_x}\right) f(n) = 0, \quad (\text{P2}^*)$$

$$\mu(\underline{n}) = \mu(\bar{n}) = 0, \quad (\text{P3}^*)$$

$$\phi \geq 0, \quad (\text{P4}^{**})$$

$$\phi(n) > 0 \Rightarrow \frac{dz(n)}{dn} = 0, \quad (\text{P5}^{**})$$

$$\frac{dz(n)}{dn} > 0 \Rightarrow \phi(n) = 0, \quad (\text{P6}^{**})$$

$$\phi(\underline{n}) = \phi(\bar{n}) = 0. \quad (\text{P7}^{**})$$

These conditions can be directly interpreted. The important variable is  $\phi(n)$ . If  $\phi(n) = 0$ , the second-order conditions do not bind. Then the underlined term in eq. (P1\*) vanishes and (P4\*\*)–(P7\*\*) are irrelevant. Gross income  $z(n)$  is a strictly increasing function of the ability  $n$ . (P1\*)–(P3\*) are the well-known necessary conditions for a solution of problem P\*. In any case the expression  $(1-s)$  in eq. (P1\*) plays an important role since it equals the marginal tax rate. Therefore it allows us to describe the optimal tax schedule.

On the other hand, assume that  $\phi(n) > 0$  for  $n \in (\underline{n}, \bar{n})$  and  $\phi(\underline{n}) = \phi(\bar{n}) = 0$ . Then the second-order condition (SOC) is binding. (P5\*\*) implies that all individuals who possess an ability  $n \in [\underline{n}, \bar{n}]$  earn the same gross income  $z(n) = z(\underline{n})$  since  $dz(n)/dn = 0$  for  $n \in [\underline{n}, \bar{n}]$ . Moreover, these individuals also get the same net income  $x(n) = x(\underline{n})$ , since  $dz/dn = 0$  is equivalent to  $dx/dn = 0$  because of [FOC]. In this case a continuum of individuals who have different abilities choose the same bundle  $(x(\underline{n}), z(\underline{n}))$ . Then we observe ‘bunching of individuals’. As will be shown below, the tax schedule has a kink at  $z(\underline{n})$ . In this case the government is not able to discriminate the individuals perfectly. Compared with a tax, which is based on ability  $n$  and which allows a perfect discrimination of all individuals, the optimal *income* tax leads to a welfare loss.

#### 4. On the relevance of the second-order approach: An example

The necessary conditions (P1\*)–(P3\*) cannot be dealt with easily. Thus it is not surprising that one cannot find explicit solutions of the first-order



model  $P^*$  in the literature. Some authors try to obtain qualitative properties of an optimal income tax [e.g. Mirrlees (1971), Sadka (1976), Seade (1977, 1982)]. On the other hand, there are some articles that are concerned with numerical solutions or simulations of the model for special cases of the welfare function, the utility function  $u$ , and the density function  $f$  [e.g. Mirrlees (1971), Tuomala (1984)]. Nevertheless there is one example of an explicit solution to the optimal taxation problem, namely the paper of Lollivier and Rochet (1983).<sup>6</sup> They do not use Mirrlees' model, but propose a different method of solution. Moreover, they make the rather special assumption that ability  $n$  is uniformly distributed.

In what follows a concrete example will be presented which demonstrates that the first-order approach is not always sufficient to give a correct solution to the economic problem  $P$ . We assume that all individuals possess the quasilinear utility function

$$u^*(x, y) = v(x) - y,$$

where  $v(x)$  is a strictly concave, twice differentiable, positive function.  $u^*(x, y)$  is additively separable in consumption  $x$  and labour  $y$ . This property means that the substitution between consumption and labour does not depend on the amount of labour the individual provides. It implies that there is no income effect on consumption.

At first Problem  $P^*$  will be solved for *arbitrary* density functions  $f$ .<sup>7</sup> We need three more definitions:

$$F(n) := \int_{\underline{n}}^n f(m) dm,$$

$$G(n) := \int_{\underline{n}}^n \frac{f(m)}{m} dm,$$

and

$$\beta(n) = nf(n) + F(n) - G(n)/G(n),$$

where  $F$  is the cumulative distribution function of  $f$ ;  $G$  and  $\beta$  are useful abbreviations. Now the solution to  $P^*$  is given [as long as  $\beta(n) > 0$ ] by

$$x(n) = v'^{-1}(f(n)/\beta(n)),$$

<sup>6</sup>A similar (principal-agent) problem is dealt with in Guesnerie and Laffont (1984).

<sup>7</sup>Weymark investigates the problem of optimal income taxation for this class of utility functions in a *finite* economy [Weymark (1986a, 1986b, 1987)].

$$u(n) = \frac{1}{n} \left( K + \int_{\underline{n}}^n v(x(m)) \, dm \right),$$

$$z(n) = n(v(x(n)) - u(n)),$$

$$\mu(n) = n(G(\bar{n})F(n) - G(n)),$$

and

$$\lambda = G(\bar{n}),$$

where

$$K = \int_{\underline{n}}^{\bar{n}} \{ [mf(m) - 1 + F(m)]v(x(m)) - x(m)f(m) \} \, dm.$$

It is obvious that these functions can be determined uniquely and explicitly in terms of  $f$  and  $v$ . In appendix A it is proved that they satisfy the differential equations (2) and (3), the boundary condition (4), and the necessary conditions (P1\*)–(P3\*). Thus we have an explicit solution of Mirrlees' income tax problem, formulated as Problem P\*.

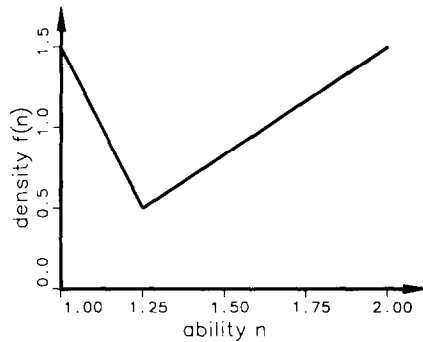


Fig. 1

Using this result we are in a position to solve P\* for arbitrary density functions  $f$ . We choose the interval  $[\underline{n}, \bar{n}] = [1, 2]$  and a particular density function  $f^*$  which is depicted in fig. 1.  $f^*$  possesses a minimum at  $n = 1.25$ . Now fig. 2 illustrates the solution of P\* for an arbitrary  $u^*$  and  $f^*$  qualitatively.<sup>8</sup> The diagram shows the relationship between gross income  $z$  and net income [consumption]  $x$ . To be more precise, all optimal combi-

<sup>8</sup>The solution of this example is derived in Appendix B.

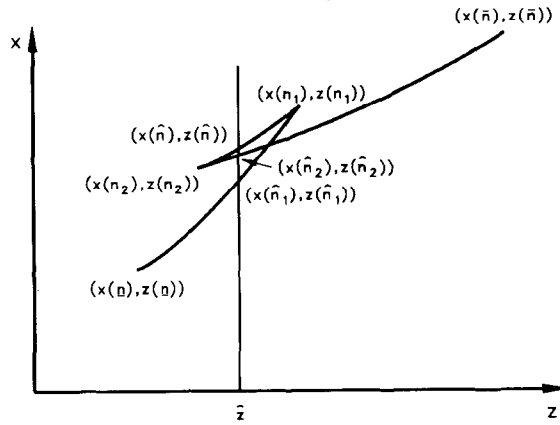


Fig. 2

nations  $(x(n), z(n))$  for  $n \in [\underline{n}, \bar{n}]$  are delineated. The figure demonstrates that there are levels of gross income  $z$  which are linked with three *different* levels of net income. Evidently there is something wrong with the optimal tax system: the second-order conditions are violated. Gross income  $z(n)$  [and net income  $x(n)$  as well] is not always a (weakly) increasing function of  $n$ . In the interval  $[n_1, n_2]$  gross and net incomes decrease, i.e. condition (SOC) is not fulfilled. In other words, we did not get a solution to Problem P\*\*.

The optimal income tax depicted in fig. 2 cannot be realized since the condition of implementability (SSC) is violated. Choose for instance the gross income  $\hat{z}$  which is earned by three different types of individuals, namely those having ability  $\hat{n}_1 < \hat{n} < \hat{n}_2$ . It is obvious that an individual possessing ability  $\hat{n}_2$  is better off if he or she receives  $x(\hat{n})$  units of the consumption good:

$$U(x(\hat{n}), z(\hat{n}), \hat{n}_2) > U(x(\hat{n}_2), z(\hat{n}_2), \hat{n}_2),$$

i.e.  $(x(\hat{n}_2), z(\hat{n}_2))$  is *not* a global maximum, but merely a *local* optimum for an  $\hat{n}_2$  individual. Furthermore, an  $\hat{n}_2$  individual can hide his or her true ability and can work with the productivity  $\hat{n} < \hat{n}_2$ . In that case the individual would be better off. This behaviour can only be avoided if the government is able to specify a budget set which depends on ability  $n$ . Since the government *cannot* observe true ability, this procedure is not possible. The government has to specify a budget set (an income tax) which is common to *all* individuals. Therefore the optimal tax system derived in this example cannot be implemented.

Thus the example demonstrates that the first-order approach is not the appropriate procedure to formulate Mirrlees' model. The second-order

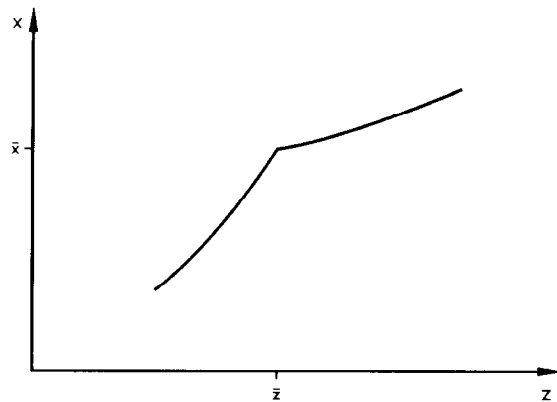


Fig. 3

conditions *are* relevant. Now it is interesting to compute a solution of  $P^{**}$  for the example above. Fig. 3 demonstrates this solution graphically.<sup>9</sup> Now gross income  $z(n)$  and net income  $x(n)$  are (weakly) *increasing* functions of ability  $n$ . There is one subinterval  $[\underline{n}, \tilde{n}]$  where both functions are constant. It corresponds to the kink in the net-income schedule. The optimal tax is implementable.

In this case we obtain bunching of individuals. There are individuals of different types who all choose the same pair  $(\bar{x}, \bar{z})$ . We have a kink in the net-income schedule at  $\bar{z}$ . Bunching occurs only within the framework of model  $P^{**}$ . The condition  $dz(n)/dn \geq 0$  must be binding. Then the solution implied by model  $P^*$  is not implementable.

## 5. Properties of the optimal income tax

The example demonstrates that the possibility of bunching *must* be taken into account when the optimal income tax is to be investigated, i.e. one must not examine the necessary conditions  $(P1^*)$ – $(P3^*)$ , but one has to look at the conditions  $(P1^*)$ – $(P7^{**})$ .

In what follows we reexamine the properties of the optimal income within this framework.

First we take stock of the literature. Here we find the following results:

(R1) The marginal tax rate is non-negative [Mirrlees (1971)].

<sup>9</sup>Compare this figure with fig. 1 in Rogerson (1985) who investigates the first-order approach in a different context. This figure is in general *not* identical to the lower part of the curve depicted in fig. 2.

- (R2) The marginal tax rate is less than 100 percent [Mirrlees (1971)].
- (R3) The marginal tax rate of the most able individual is zero [Seade (1977)].
- (R4) If there is no bunching at the lowest income, the marginal tax rate of the least able individual is zero [Seade (1977)].
- (R5) For every income  $z$  such that  $z(\underline{n}) < z < (\bar{n})$  the marginal tax rate is *strictly positive* [Seade (1982)].

All these properties are proven in a first-order model under essentially the same assumptions that are made in this paper. Of course it is necessary to *scrutinize* these results within the model  $P^{**}$ . We obtain:

*Theorem.* (R1)–(R5) hold in model  $P^{**}$  as well.

At first sight this theorem seems to be natural. Nevertheless it must be stressed that it is not easy to prove it in the extended framework, as can be seen from the proof in appendix C. Particularly the proof of (R5) is difficult. On the other hand, it should be emphasized again that the derivation of these properties in the second-order model is necessary since only in that case is bunching taken into account correctly.

Furthermore, some new properties of the optimal income tax can be derived.

- (R6) If bunching prevails at the lowest income, the marginal rate of tax is strictly positive at the end of the bunching interval.

Seade (1977, 1982) provides the following interpretation of a positive marginal rate of tax. Its main task is to redistribute income. A positive rate for any level of gross income *increases* the tax liability of all higher incomes and *decreases* the liability of all lower incomes. This fact explains why the marginal rate of tax is zero for the lowest and the highest incomes. In the first case there is no individual who benefits from a redistribution, and in the second case there are no individuals having higher incomes who would contribute to a potential redistribution.

If bunching occurs we know that the optimal tax schedule derived in model  $P^*$  cannot be implemented. We have to accept the corresponding welfare loss in order to guarantee the implementability of the tax schedule. Thus it is not surprising that in this case the redistribution of income is not optimal either: the marginal rate of tax is positive for the lowest gross

income; we are able to back up Seade's interpretation. Finally, we are in the position to state two results about bunching.

- (R7) If the marginal (dis)utility of labour is zero whenever an individual does *not* work ( $u_y(x,0)=0$  for all  $x>0$ ) and if the lowest income is zero ( $z(\underline{n})=0$ ), then no bunching can occur for the lowest income  $z(\underline{n})$ .

This case is discussed by Seade (1977). He needs the additional assumption that the marginal rate of tax for  $z(\underline{n})=0$  is less than 100 percent and he essentially argues by means of a diagram. Since the disutility of labour is zero if an individual does not work, the disutility will only slightly increase if the individual provides a small amount of labour. Therefore the government has no problems in designing a tax schedule which acts as an incentive to work for those individuals. Thus no bunching arises. At the top of the income scale bunching can be excluded.

- (R8) There is no bunching at the highest income  $z(\bar{n})$ .

The reason for this property is the following. We know that the marginal rate of tax for the most able individual must be zero (R3). Whenever bunching occurs the marginal rate of tax is a decreasing function of the ability  $n$  for all those individuals who choose this gross income. Thus the tax rate of an individual who has almost the highest ability must be negative if we have bunching at  $z(\bar{n})$ . This is excluded by (R3) and (R5).

## 6. Conclusion

The analysis of this paper draws attention to a problem in developing an optimal nonlinear income tax. The crucial point is the way in which the individual's reaction to the imposition of the tax is taken into account. It turns out that substituting the first-order conditions from utility maximization does not suffice to imply the implementability of the tax system. Also second-order conditions have to be taken into consideration. This leads to an extension of Mirrlees' model and a somewhat more complicated system of necessary conditions for an optimal tax. It is necessary to investigate the properties of an optimal nonlinear tax in the framework of the extended model.

The main properties of an optimal tax schedule which have been derived in Mirrlees' model can be proved in this extended model as well. Furthermore, some new results are presented in this paper. Particularly the problem of bunching can be investigated rigorously in this framework. Perhaps it will

be possible to describe the optimal nonlinear income tax even more precisely in the future.

#### Appendix A: Solution of Problem P\* for quasi-linear utility functions

Problem P\* has the following solution for  $u(x, y) = v(x) - y$  [as long as  $\beta(n) > 0$ ]:

$$x(n) = v'^{-1}(f(n)/\beta(n)),$$

$$u(n) = \frac{1}{n} \left( K + \int_{\underline{n}}^n v(x(m)) \, dm \right),$$

$$z(n) = n(v(x(n)) - u(n)),$$

$$\mu(n) = n(G(\bar{n})F(n) - G(n)),$$

and

$$\lambda = G(\bar{n}),$$

where

$$K = \int_{\underline{n}}^{\bar{n}} \{ [mf(m) - 1 + F(m)]v(x(m)) - x(m)f(m) \} \, dm,$$

$$F(n) = \int_{\underline{n}}^n f(m) \, dm,$$

$$G(n) = \int_{\underline{n}}^n \frac{f(m)}{m} \, dm,$$

and

$$\beta(n) = nf(n) + F(n) - G(n)/G(\bar{n}).$$

First we define some important derivatives which will be needed below.

We have  $U(x, z, n) = v(x) - z/n$  and therefore

$$U_x = v'(x), \quad U_z = -\frac{1}{n}, \quad U_n = \frac{z}{n^2}, \quad U_{nx} = 0,$$

$$s = -\frac{U_z}{U_x} = \frac{1}{nv'(x)}, \quad s_n = -\frac{1}{n^2 v'(x)}.$$

Now we will prove all conditions which must be satisfied by a solution.

$$(a) \quad \frac{du(n)}{dn} = U_n(x(n), z(n), n)'. \quad (2)$$

We get by insertion:

$$\begin{aligned} \frac{du(n)}{dn} &= \frac{d}{dn} \left[ \frac{1}{n} \left( K + \int_n^n v(x(m)) dm \right) \right] \\ &= \frac{v(x(n)) \cdot n - \left( K + \int_n^n v(x(m)) dm \right)}{n^2} \\ &= \frac{v(x(n))}{n} - \frac{1}{n} \left[ \frac{1}{n} \left( K + \int_n^n v(x(m)) dm \right) \right] \\ &= \frac{v(x(n))}{n} - \frac{u(n)}{n}. \end{aligned}$$

On the other hand:

$$U_n = \frac{z}{n^2} = \frac{n(v(x(n)) - u(n))}{n^2} = \frac{v(x(n))}{n} - \frac{u(n)}{n}.$$

$$(b) \quad \frac{dR}{dn} = (z(n) - x(n)) f'(n). \quad (3)$$

This equation holds by the definition of  $R$ .

$$(c) \quad R(\underline{n}) = R(\bar{n}) = 0. \quad (4)$$

We have

$$R(n) = \int_n^n (z(n') - x(n')) f'(n') dn'$$

and get immediately  $R(\underline{n}) = 0$ .

Now consider



$$\begin{aligned}
R(\bar{n}) &= \int_{\underline{n}}^{\bar{n}} (z(n) - x(n)) f(n) \, dn \\
&= \int_{\underline{n}}^{\bar{n}} \{n(v(x(n)) - u(n)) - x(n)\} f(n) \, dn \\
&= \int_{\underline{n}}^{\bar{n}} \left\{ nv(x(n)) \cdot f(n) - n \left[ \frac{1}{n} \left( K + \int_{\underline{n}}^n v(x(m)) \, dm \right) \right] f(n) - x(n) f(n) \right\} \, dn \\
&= \int_{\underline{n}}^{\bar{n}} \left\{ nv(x(n)) \cdot f(n) - \left[ \int_{\underline{n}}^n v(x(m)) \, dm \right] f(n) - x(n) f(n) \right\} \, dn - K.
\end{aligned}$$

Now by definition of  $K$  we get

$$\begin{aligned}
R(\bar{n}) &= \int_{\underline{n}}^{\bar{n}} \left\{ nv(x(n)) \cdot f(n) - \int_{\underline{n}}^n v(x(m)) \, dm \cdot f(n) - x(n) f(n) \right. \\
&\quad \left. - [(nf(n) - 1 + F(n))v(x(n)) - x(n)f(n)] \right\} \, dn \\
&= \int_{\underline{n}}^{\bar{n}} \left\{ - \int_{\underline{n}}^n v(x(m)) \, dm \cdot f(n) + v(x(n)) - F(n)v(x(n)) \right\} \, dn
\end{aligned}$$

since some terms disappear.

Now the order of integration can be reversed:

$$\begin{aligned}
\int_{\underline{n}}^{\bar{n}} \left[ \int_{\underline{n}}^n v(x(m)) \cdot dm \right] f(n) \, dn &= \int_{\underline{n}}^{\bar{n}} v(x(n)) \cdot \left[ \int_{\underline{n}}^{\bar{n}} f(m) \cdot dm \right] \cdot dn \\
&= \int_{\underline{n}}^{\bar{n}} v(x(n)) [1 - F(n)] \, dn.
\end{aligned}$$

Thus we obtain  $R(\bar{n}) = 0$ .

$$(d) \quad -\mu(n) \cdot U_x \cdot s_n + \lambda(1 - s(n))f(n) = 0. \quad (\text{P1}^*)$$

We get by insertion:

$$\begin{aligned}
& -\mu(n) \cdot U_x \cdot s_n + \lambda(1-s(n))f(n) \\
& = -n(G(\bar{n})F(n) - G(n)) \cdot v'(x(n)) \cdot \left( -\frac{1}{n^2 v'(x(n))} \right) \\
& \quad + G(\bar{n}) \left( 1 - \left( \frac{1}{nv'(x(n))} \right) \right) \cdot f(n) = (*).
\end{aligned}$$

Now observe that

$$\begin{aligned}
nv'(x(n)) &= nv'(v'^{-1}(f(n)/\beta(n))) \\
&= \frac{nf(n)}{\beta(n)} = \frac{nf(n)}{nf(n) + F(n) - G(n)/G(\bar{n})}.
\end{aligned}$$

Therefore we get

$$\begin{aligned}
(*) &= \frac{1}{n}(G(\bar{n}) \cdot F(n) - G(n)) \\
& \quad + G(\bar{n}) \frac{nf(n) - [nf(n) + F(n) - G(n)/G(\bar{n})]}{nf(n)} \cdot f(n) = 0.
\end{aligned}$$

$$(e) \quad \mu'(n) + \mu(n) \cdot \frac{U_{nx}}{U_x} + \left( 1 - \frac{\lambda}{U_x} \right) f(n) = 0. \quad (P2*)$$

We use the solution given above and obtain:

$$\begin{aligned}
& [n(G(\bar{n}) \cdot F(n) - G(n))]' + 0 + \left( 1 - \frac{G(\bar{n})}{v'(x(n))} \right) f(n) \\
& = \left[ G(\bar{n}) \cdot F(n) - G(n) + n \left( G(\bar{n}) f(n) - \frac{f(n)}{n} \right) \right] \\
& \quad + \left( 1 - \frac{G(\bar{n})}{v'(v'^{-1}(f(n)/\beta(n)))} \right) f(n)
\end{aligned}$$

$$\begin{aligned}
&= [G(\bar{n})F(n) - G(n) + nG(\bar{n})f(n) - f(n)] + \frac{f(n) - G(\bar{n}) \cdot \beta(n)}{f(n)} f(n) \\
&= [G(\bar{n})F(n) - G(n) + nG(\bar{n})f(n) - f(n)] \\
&\quad + [f(n) - G(\bar{n}) \cdot (nf(n) + F(n) - G(n)/G(\bar{n}))] \\
&= 0.
\end{aligned}$$

$$(f) \quad \mu(\underline{n}) = \mu(\bar{n}) = 0. \quad (P3^*)$$

Immediately by the definition of  $F$  and  $G$ :

$$F(\underline{n}) = G(\underline{n}) = 0 \quad \text{and} \quad F(\bar{n}) = 1.$$

## Appendix B: Solution of the example

We choose the utility function  $u(x, y) = v(x) - y$ , where  $v(x)$  is *any* strictly concave, twice differentiable, positive function, and the density function  $f^*$  which is depicted in fig. 1.  $f^*$  is defined on  $[\underline{n}, \bar{n}] = [1, 2]$  by

$$f^*(n) = \begin{cases} 5\frac{1}{2} - 4n & \text{for } 1 \leq n \leq 1\frac{1}{4}, \\ -1\frac{1}{6} + 1\frac{1}{3} \cdot n & \text{for } 1\frac{1}{4} < n \leq 2. \end{cases}$$

$f^*$  possesses a minimum at  $n = 1.25$ . The form of  $f^*$  allows us to derive the functions  $F$  and  $G$  as simple algebraic expressions:

$$F(n) = \begin{cases} 5\frac{1}{2}n - 2n^2 - 3\frac{1}{2} & \text{for } 1 \leq n \leq 1\frac{1}{4}, \\ -1\frac{1}{6}n + \frac{2}{3}n^2 + \frac{2}{3} & \text{for } 1\frac{1}{4} < n \leq 2, \end{cases}$$

and

$$G(n) = \begin{cases} 5\frac{1}{2} \ln n - 4n + 4 & \text{for } 1 \leq n \leq 1\frac{1}{4}, \\ -1\frac{1}{6} \ln n + 1\frac{1}{3}n - 1.179 & \text{for } 1\frac{1}{4} < n \leq 2. \end{cases}$$

Therefore it is not difficult to compute  $\beta(n)$  and to derive  $x(n)$ ,  $z(n)$  and  $u(n)$  after choosing any function  $v$  (cf. Appendix A). Figs. B1 and B2 illustrate the form of  $x(n)$  and  $z(n)$  diagrammatically. It turns out that  $x(n)$  and  $z(n)$  strictly decrease on a subinterval. This fact can be easily proved numerically as well: evaluate  $dx(n)/dn$  at  $\hat{n} = 1.245$ . Then one gets:

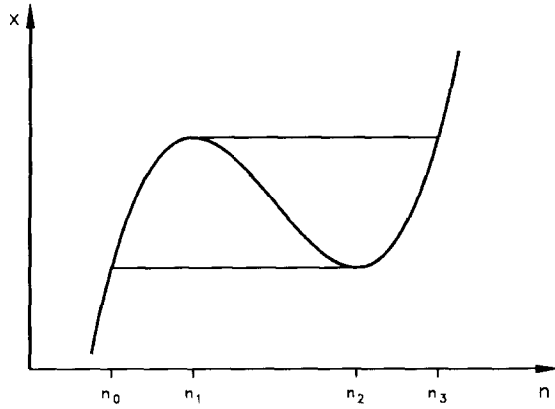


Fig. B1

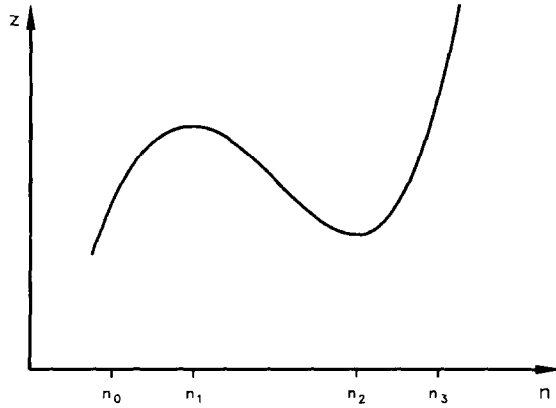


Fig. B2

$$\frac{dx(\hat{n})}{dn} = \frac{1}{v''(v'^{-1}(f(\hat{n})/\beta(\hat{n})))} \frac{f'(\hat{n})\beta(\hat{n}) - f(\hat{n})\beta'(\hat{n})}{\beta^2(\hat{n})} < 0.$$

In order to derive fig. 2 we choose  $n_i$ ,  $i=0,1,2,3$ , as indicated in fig. B1. Generally we have

$$\begin{aligned} nu(n) &= K + \int_{\underline{n}}^n v(x(m)) dm \\ &= n_i u(n_i) + \int_{n_i}^n v(x(m)) dm \quad \text{for } n \geq n_i \quad (\text{cf. Appendix A}), \end{aligned}$$

and for  $n = n_j > n_i$ :

$$n_j u(n_j) = n_i u(n_i) + \int_{n_i}^{n_j} v(x(m)) \, dm. \quad (\text{B1})$$

On the other hand:

$$nu(n) = nv(x(n)) - z(n).$$

Using this equation for  $n = n_j$  and  $n = n_i$  we can transform (B1) into

$$n_j v(x(n_j)) - z(n_j) = n_i v(x(n_i)) - z(n_i) + \int_{n_i}^{n_j} v(x(m)) \, dm. \quad (\text{B2})$$

Next we observe that

$$\int_{n_i}^{n_j} v(x_j) \, dm = n_j v(x_j) - n_i v(x_j)$$

and replace  $n_j v(x_j)$  in (B2) accordingly. Then we obtain by simple rearrangements for  $n_j > n_i$ :

$$z(n_j) = z(n_i) + n_i [v(x(n_j)) - v(x(n_i))] + \int_{n_i}^{n_j} [v(x(n_j)) - v(x(m))] \, dm.$$

Because of the monotonicity of  $v$  we get

$$z(n_1) > z(n_0), \quad z(n_2) < z(n_1) \quad \text{and} \quad z(n_3) > z(n_2)$$

and, moreover,

$$z(n_2) < z(n_0) \quad \text{and} \quad z(n_3) > z(n_1)$$

since

$$x(n_2) = x(n_0) \quad \text{and} \quad x(n_3) = x(n_1).$$

These inequalities imply the ordering

$$z(n_2) < z(n_0) < z(n_1) < z(n_3)$$

and the diagrams depicted in figs. B2 and B3 which form the basis of fig. 2.

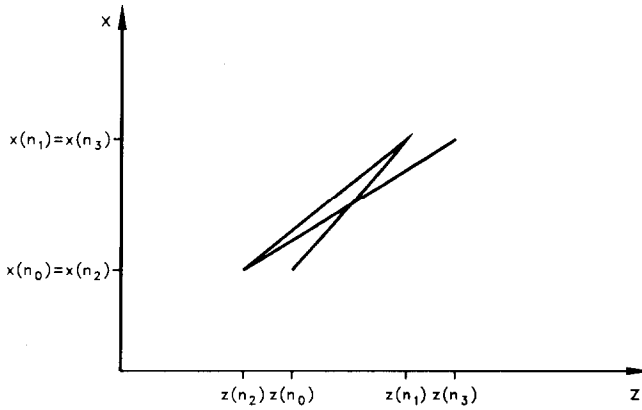


Fig. B3

### Appendix C

First some useful relationships are stated or derived. The first-order conditions (P1\*) and (P2\*) are equivalent to (C1) and (C2), respectively:

$$\mu' U_z + \mu U_{nz} + (U_z + \lambda) f - \phi' = 0, \quad (C1)$$

$$\mu' U_x + \mu U_{nx} + (U_x - \lambda) f = 0. \quad (C2)$$

(C1) follows immediately by observing that  $U_x s_n = -(U_{nx} s + U_{nz})$  and using (P2\*).

Solving the differential equations (C1) and (C2) we get directly:

$$\mu(n) = \int_n^{\bar{n}} \left[ \left( 1 + \frac{\lambda}{U_z} \right) f(n') - \frac{\phi'(n')}{U_z} \right] \exp \left( \int_n^{n'} \frac{U_{nz}}{U_z} dn'' \right) dn', \quad (C3)$$

$$\mu(n) = \int_n^{\bar{n}} \left( 1 - \frac{\lambda}{U_x} \right) f(n') \exp \left( \int_n^{n'} \frac{U_{nx}}{U_x} dn'' \right) dn'. \quad (C4)$$

Here the initial conditions (P3\*) are taken into account.

Moreover, from (C4) we obtain the inequality

$$\lambda > 0 \quad (C5)$$

(observe that the integrand must change its sign at least once;  $U_x$  is positive and cannot be constant).

Finally, the sign of the marginal tax rate  $t'(n)$  is mostly equivalent to the sign of the adjoint variable  $\mu$

$$\text{sign}(t'(n)) = -\text{sign}(\mu(n)) \text{ if there is no bunching at } z(n). \quad (\text{C6})$$

(C6) is implied by the (P1\*), agent monotonicity (AM), (C5), and the fact that  $t'(n) = (1-s)$ .

Now we turn to the proof of (R1)–(R8). (R1) and (R5) will be considered at the end of the appendix because the proof is long.

(R2) has already been proved by Mirrlees (1971).

(R3) is implied by (P1\*). We have

$$(1-s(\bar{n})) = \frac{\mu(\bar{n}) \cdot U_x s_n + \phi'(\bar{n})}{\lambda f(\bar{n})} = \frac{\phi'(\bar{n})}{\lambda f(\bar{n})}$$

since  $\mu(\bar{n}) = 0$  by (P3\*). Now  $\lambda$  is positive by (C5) and  $\phi'$  is nonpositive by (P4\*\*) and (P7\*\*). Thus  $(1-s(\bar{n})) \leq 0$ . The rest follows from (R5).

(R4) is due to Seade (1977).

Next we look at (R6). For  $n = \underline{n}$ , (P1\*), (P3\*), and (P7\*\*) imply

$$\lambda(1-s(\underline{n}))f(\underline{n}) = \mu(\underline{n})U_x s_n + \phi'(\underline{n}) = 0.$$

If bunching obtains,  $(1-s(n))$  increases with  $n$ :

$$\begin{aligned} \frac{d}{dn}(1-s(n)) &= \frac{d}{dn}(1-s(x(n), z(n), n)) \\ &= -s_x \cdot x'(n) - s_z z'(n) - s_n = -s_n > 0 \end{aligned}$$

because of  $x'(n) = z'(n) = 0$  and (AM). This proves (R6).

We now consider (R7). Suppose that there is bunching in the interval  $[\underline{n}, \bar{n}]$  and that  $z(n) = 0$  for  $n \in [\underline{n}, \bar{n}]$ . Because of the form of the utility function we have  $U_z(n) = u_y(x(n), 0) = 0$  and therefore  $U_{nz}(n) = 0$  for all  $n \in [\underline{n}, \bar{n}]$ . (C1) implies in particular  $\phi'(n) = \lambda f(n) > 0$  for all  $n \in [\underline{n}, \bar{n}]$ . On the other hand, we have  $\phi(\underline{n}) = \phi(\bar{n}) = 0$  and  $\phi(n) > 0$  for  $n \in (\underline{n}, \bar{n})$ . This is a contradiction.

(R8) follows from (R3). The marginal tax rate is zero. If we have bunching at  $z(\bar{n})$  the marginal rate must be negative for some  $n < \bar{n}$  since  $(1-s(n))$  increases [cf. the proof of (R6)]. This fact contradicts (R1), which will be proved now.

The proof of (R1) is rather lengthy. If no bunching prevails, the result is proved by Seade (1982). Therefore we have to discuss the implications of

bunching. Suppose that bunching obtains in a (maximal) interval  $[\underline{n}, \tilde{n}]$ . If  $\underline{n} = \underline{n}$ , property (R6), which was proved above, demonstrates that the marginal tax rate is positive at  $n = \tilde{n}$ . Thus we assume  $\underline{n} < \tilde{n}$ . If the marginal tax rate is non-negative for  $n = \underline{n}$ , it is positive as well for  $n = \tilde{n}$  since the rate increases in  $[\underline{n}, \tilde{n}]$  because of agent monotonicity (AM). So we must consider the situation where  $t'(\underline{n}) = (1 - s(\underline{n}))$  is less than zero. We have to distinguish between two cases. In case A the tax rate is still nonpositive at  $n = \tilde{n}$ , in case B it is strictly positive for  $n = \tilde{n}$ .

*Case A.* Here we can apply Seade's method of proof. Since  $(1 - s(\tilde{n})) \leq 0$ ,  $\mu(\tilde{n})$  cannot be negative [cf. (P1\*)]. There must exist  $n_* \leq \underline{n}$  and  $n^* \geq \tilde{n}$  such that

$$\mu(n_*) = \mu(n^*) = 0 \quad \text{and} \quad \mu(n) > 0 \quad \text{for } n \in (n_*, n^*)$$

We obtain  $\mu'(n_*) \geq 0$  and  $\mu'(n^*) \leq 0$ .

Condition (C2) implies  $U_x(n_*) \leq \lambda$  and  $U_x(n^*) \geq \lambda$ . Since  $(1 - s(n_*)) = 0$  we get

$$U_x(n_*) = -U_z(n_*)$$

by the definition of  $s$ .

Application of (C1) implies, because of  $\phi'(n^*) \leq 0$ ,

$$\lambda \leq -U_z(n^*).$$

All this yields (remember that  $U_x = u_x$  and  $U_z = u_y/n$ ):

$$u_x(n_*) \leq u_x(n^*), \tag{C7}$$

$$-u_y(n_*) < -u_y(n^*) \frac{n_*}{n^*} \leq -u_y(n^*), \tag{C8}$$

and of course we have

$$u(n_*) < u(n^*) \tag{C9}$$

(since  $du/dn > 0$ ).

The implications of (C7), (C8), and (C9) are ruled out by a theorem proved in Dixit and Seade (1979) because  $u$  has to be twice continuously differentiable, monotone, and strictly concave, and because consumption and leisure are normal and noninferior, respectively [cf. Seade (1982)].



*Case B.* Here we have  $\mu(\tilde{n}) < 0$ . We choose  $n_*$  as before and get  $U_x(n_*) = -U_z(n_*)$ . Define  $n^*$  by the condition that  $n^* \in [\tilde{n}, \tilde{n}]$  and  $(1 - s(n^*)) = 0$ . The sign of  $\mu(n^*)$  is unknown. We consider three different cases:  $\mu(n^*) < 0$ ,  $\mu(n^*) = 0$ , and  $\mu(n^*) > 0$ .

(1) Suppose  $\mu(n^*) < 0$ . Then there exists  $\hat{n} \in [n_*, n^*]$  such that  $(1 - s(\hat{n})) < 0$ ,  $\mu(\hat{n}) = 0$ , and  $\mu'(\hat{n}) \leq 0$ . Condition (C2) implies  $U_x(\hat{n}) \geq \lambda$  and therefore we have  $U_x(\hat{n}) \geq U_x(n_*)$ .  $(1 - s(\hat{n})) < 0$  yields  $\lambda \leq U_x(\hat{n}) < -U_z(\hat{n})$ . Thus we can argue as Seade again.

(2) If  $\mu(n^*) = 0$  we can use the same arguments. (Choose  $\hat{n} = n^*$ .)

(3) Suppose  $\mu(n^*) > 0$ . If  $U_x(n^*) = -U_z(n^*) \geq U_x(n_*) = -U_z(n_*)$ . We use the same kind of proof as above. Otherwise we must have

$$U_x(n^*) = -U_z(n^*) < U_x(n_*) = -U_z(n_*) \leq \lambda.$$

Now we consider in the bunching interval

$$\frac{d}{dn}(\mu \cdot U_z) = \mu' \cdot U_z + \mu \cdot \frac{dU_z}{dn} = \mu' U_z + \mu \cdot U_{nz}$$

( $x'(n) = z'(n) = 0$ !). Therefore we know by (C1) that

$$\frac{d}{dn}(\mu \cdot U_z) = -(U_z + \lambda) f + \phi'.$$

Integrating both identities we get

$$\begin{aligned} \int_{n^*}^{\tilde{n}} \frac{d}{dn}(\mu U_z) &= \mu(\tilde{n}) U_z(\tilde{n}) - \mu(n^*) U_z(n^*) =: (*) \\ &= - \int_{n^*}^{\tilde{n}} (U_z + \lambda) f(n) dn + \phi(\tilde{n}) - \phi(n^*) =: (**). \end{aligned}$$

Since  $U_z < 0$ ,  $\mu(\tilde{n}) < 0$ , and  $\mu(n^*) > 0$  expression (\*) is positive. On the other hand,  $U_z(n^*) + \lambda$  is positive. Furthermore, it increases on  $[n^*, \tilde{n}]$  since

$$\frac{d}{dn} U_z = \frac{-u_{yy} \frac{z(n^*)}{n^2} - u_y}{n^2}$$

is positive. Thus (\*\*) is negative ( $\phi(\tilde{n}) = 0$  and  $\phi(n^*) > 0$ ) and we have a contradiction.

*Proof of (R5).* Assume that the marginal rate of tax is zero for  $z(n^*)$ ,  $z(\underline{n}) < z(n^*) < z(\bar{n})$ . This implies  $\mu(n^*) = \mu'(n^*) = 0$ . By (C2) we obtain  $U_x(n^*) = \lambda$ . It follows from (R1) that there can be no bunching in a small interval  $(\hat{n}, n^*)$  since otherwise the marginal tax rate would be negative, therefore  $-U_z(n^*) = \lambda$  by (C1).

We choose  $n_* \in (\hat{n}, n^*)$  such that  $\mu'(n_*) > 0$ . Since  $U_{nz}$  is positive, eq. (C1) implies that  $U_z(n_*) + \lambda > 0$ , i.e.  $-U_z(n_*) < \lambda$ . This inequality implies that  $1 + (\lambda/U_z) < 0$  in a small neighbourhood of  $n$ . ( $\mu(n)$  is increasing!), i.e. the integrand in (C3) is negative, since  $\phi'(n) = 0$ . The same must be true for the integrand in (C4). Therefore we have

$$1 - \frac{\lambda}{U_x} < 0$$

and hence  $U_x(n_*) < \lambda$ .

Thus we can apply Seade's proof again. The case considered cannot occur if consumption and leisure are noninferior. Q.E.D.

## References

- Brito, D.L. and W.H. Oakland, 1977, Some properties of the optimal income tax, *International Economic Review* 18, 407–423.
- Bryson, A.E. and Y.-C. Ho, 1969, *Applied optimal control* (Ginn and Company, Waltham).
- Dixit, A.K. and J.K. Seade, 1979, Utilitarianism vs. egalitarian redistributions, *Economics Letters* 4, 121–124.
- Ebert, U., 1986a, On the characterization of the optimal nonlinear income tax, Discussion Paper A-65, SFB 303, Bonn.
- Ebert, U., 1986b, Properties of the optimal nonlinear income tax, Discussion Paper A-66, SFB 303, Bonn.
- Guesnerie, R. and J.-J. Laffont, 1984, A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm, *Journal of Public Economics* 25, 329–369.
- Hellwig, M., 1986, The optimal linear income tax revisited, *Journal of Public Economics* 31, 163–179.
- 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.
- Mirrlees, J., 1971, An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, 175–208.
- Mirrlees, J., 1975, The theory of moral hazard and unobservable behavior – Part 1, Mimeo. (Nuffield College).
- Mirrlees, J., 1976, Optimal tax theory, a synthesis, *Journal of Public Economics* 7, 327–358.
- Mirrlees, J., 1986, The theory of optimal taxation, in: K.J. Arrow and M.D. Intriligator, eds., *Handbook of mathematical economics*, Vol. III (North-Holland, Amsterdam) ch. 24, pp. 1197–1249.
- Roberts, K.W.S., 1979, Welfare considerations of nonlinear pricing, *Economic Journal* 89, 66–83.
- Rogerson, W., 1985, The first-order approach to principal-agent problems, *Econometrica* 53, 1357–1367.
- Sadka, E., 1976, On income distribution, incentive effects and optimal income taxation, *Review of Economic Studies* 43, 261–267.
- Seade, J., 1977, On the shape of optimal tax schedules, *Journal of Public Economics* 7, 203–235.

- Seade, J., 1982, On the sign of the optimum marginal income tax, *Review of Economic Studies* 49, 637–643.
- Tuomala, M., 1984, On the optimal income taxation: Some further numerical results, *Journal of Public Economics* 23, 351–366.
- Weibull, J.W., 1989, A note on the continuity of incentive schedules, *Journal of Public Economics* 39, 239–243.
- Weymark, J.A., 1986a, Bunching properties of optimal nonlinear income taxes, *Social Choice and Welfare* 3, 213–232.
- Weymark, J.A., 1986b, A reduced-form optimal nonlinear income tax problem, *Journal of Public Economics* 30, 199–217.
- Weymark, J.A., 1987, Comparative static properties of optimal nonlinear income taxes, *Econometrica* 55, 1165–1185.