

**THE IMPACT OF CHANGING SKILL LEVELS
ON OPTIMAL NONLINEAR INCOME TAXES**

by

Craig Brett and John A. Weymark



Working Paper No. 07-W08R

May 2007

Revised November 2007

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

The Impact of Changing Skill Levels on Optimal Nonlinear Income Taxes*

by

Craig Brett

*Department of Economics,
Mount Allison University,
144 Main Street, Sackville NB, E4L 1A7, Canada
(e-mail: cbrett@mta.ca)*

and

John A. Weymark

*Department of Economics,
Box 1819, Station B, Vanderbilt University,
Nashville, TN 37235, U.S.A.
(e-mail: john.weymark@vanderbilt.edu)*

May 2007

Revised, November 2007

Abstract. The impact of changing an individual's skill level on the solution to a finite population version of the Mirrlees optimal nonlinear income tax problem with quasilinear-in-leisure preferences is investigated. It is shown that it is possible to sign the directions of change in everyone's optimal consumptions and optimal marginal tax rates in response to such a change.

Journal of Economic Literature classification numbers: D82, H21.

Keywords and phrases: optimal income taxation, comparative statics.

*Craig Brett's research is generously supported by the Canada Research Chair Programme and by the Canada Foundation for Innovation. We are grateful to our referees and to Marcus Berliant for their comments.

1. Introduction

In the Mirrlees (1971) optimal nonlinear income tax problem, everyone has the same preferences for consumption and leisure, but they differ in their skills as measured by their labour productivities. Because labour markets are competitive, an individual's wage is equal to his labour productivity. While the distribution of these skills is publicly known, the skill of any particular individual is private information. The only role for the government is to design an income tax schedule in order to meet its redistributive goals given the constraints that it faces. Merely assuming that the common individual utility function is concave and increasing in both consumption and leisure yields few general qualitative properties of optimal income taxes. For this reason, more restrictive functional forms have been considered in order to obtain more clear-cut results. Assuming that preferences are quasilinear in leisure, as in Lollivier and Rochet (1983), or quasilinear in consumption, as in Diamond (1998), has been particularly fruitful.

One of the reasons that quasilinear versions of the Mirrlees model yield strong conclusions is that it is then possible to obtain a closed-form solution for the optimal allocation. This result was first shown for preferences that are quasilinear in leisure by Lollivier and Rochet (1983) for a continuum of individuals. Closed-form solutions when the population is finite have been obtained by Weymark (1986) for quasilinear-in-leisure preferences and by Simula (2007) for quasilinear-in-consumption preferences.

Weymark (1987) has used the closed-form solution for a finite population version of the Mirrlees optimal income tax problem with quasilinear-in-leisure preferences and a weighted utilitarian social welfare function to derive a number of comparative static properties of this solution. However, he does not provide any comparative statics for the skill levels. The purpose of this article is to extend Weymark's analysis by showing how the optimal marginal tax rates and optimal consumptions vary in response to a change in any person's skill level for the model that he considered. For this kind of parameter change, it does not seem possible to obtain unambiguous comparative statics for the optimal pretax incomes without further restrictions on the model. Indeed, for either kind of quasilinearity, there are few clear-cut comparative static results available for the choice variables that enter linearly in the utility function.

Skills are the key parameters of interest in the Mirrlees model because they are the only source of heterogeneity between individuals. It is therefore

important to know how the solution to the optimal income tax problem depends on the skills present in the economy. Aside from its intrinsic interest, knowing how optimal tax policy responds to changes in skills is of considerable practical significance because the skill distribution varies over time. Furthermore, the quasilinear optimal income tax model has many formal similarities to the models considered in the nonlinear pricing literature, such as that of Mussa and Rosen (1978). As a consequence, our results can be used to provide insight into how the solution to this kind of problem depends on the distribution of consumer tastes.

For the case in which an individual's skill level is increased, our findings can be summarized as follows. (i) It is optimal to increase the marginal tax rate of everyone with a lower skill and to decrease the marginal tax rate of everyone else except for the highest-skilled individual, whose marginal tax rate is always 0. (ii) It is optimal to decrease the consumption of everyone with a lower skill and to increase the consumption of everyone else except for the highest-skilled individual, whose consumption is only increased if it is his skill that has increased; otherwise his consumption is left unchanged. Some of these responses will reinforce one another when more than one skill is changed, so our analysis also provides some insight into the impact on the optimal solution of more complex changes in the skill distribution.

For quasilinear-in-leisure preferences, some comparative static results with respect to the skill levels have previously appeared in the literature, but only for the case in which there are two skill levels. Hamilton and Pestieau (2005) have investigated the effects on the individual utilities of changing the proportion of individuals with each skill when there are two fixed skill levels and the social welfare function is either maximin or maximax. Brett and Weymark (2008) have considered an extension of the Weymark (1987) model in which the government also chooses the level of provision of a public good and have shown how the optimal marginal tax rates and some components of the optimal allocation respond to a change in either of the skills.¹

For preferences that are quasilinear in consumption, Boadway and Pestieau (2006) have identified how the utility possibilities frontier responds to the kind of change in the skill distribution considered by Hamilton and Pestieau (2005). The most extensive comparative static results for this kind of quasilinearity have recently been obtained by Simula (2007) using the quasilinear-

¹Further comparative statics for this kind of quasilinearity may be found in Cebreiro-Gómez (2002) and Boone and Bovenberg (2007).

in-consumption preference version of the Weymark (1987) model. In particular, he has shown how the optimal marginal tax rates and optimal pretax incomes vary with a change in someone's skill. A striking feature of these comparative statics is that the only marginal tax rates and pretax incomes affected by a change in some person's skill are his own and those of the person with skill level just below his.

In Section 2, we introduce our model. We present our comparative static results in Section 3. We offer some concluding remarks in Section 4. All proofs are in the Appendix.

2. The Model

There are n individuals, indexed by $i = 1, \dots, n$. Individual i has skill level $w_i > 0$, which is also the wage he receives in a competitive labour market. We also refer to w_i as i 's type. To simplify the notation, we suppose that no two individuals have the same skill. Our comparative static results for a change in w_i are also valid if more than one person has this skill. Without loss of generality, we assume that

$$0 < w_1 < \dots < w_n. \quad (2.1)$$

Person i supplies l_i units of labour and therefore has a pretax income of

$$y_i = w_i l_i, \quad i = 1, \dots, n, \quad (2.2)$$

which is also i 's labour supply in efficiency units.

There is a constant-returns-to-scale technology that converts labour in efficiency units into a single consumption good. Units of this good are chosen so that one unit of consumption is produced using one unit of labour in efficiency units. The production sector is competitive, so the price of the consumption good is 1. Person i 's consumption is c_i . Because there is only one good and its price is 1, c_i is also i 's after-tax income.

A commodity bundle for person i is a vector $(l_i, c_i) \in \mathbb{R}_+^2$.² Everyone has the same quasilinear utility function $u: \mathbb{R}_+^2 \rightarrow \mathbb{R}$, where

$$u(l_i, c_i) = v(c_i) - \gamma l_i. \quad (2.3)$$

² \mathbb{R}_+ is the nonnegative subset of the real line \mathbb{R} .

The function v is assumed to be a twice continuously differentiable, increasing, strictly concave function with $v(0) = 0$, $v'(0) = \infty$, and $\lim_{r \rightarrow \infty} v'(r) = 0$. The marginal disutility of labour γ is assumed to be positive.

In terms of consumption and pretax income, i 's utility function is $u^i: \mathbb{R}_+^2 \rightarrow \mathbb{R}$, where

$$u^i(y_i, c_i) = v(c_i) - \gamma y_i/w_i, \quad i = 1, \dots, n. \quad (2.4)$$

Person i 's marginal rate of substitution between pretax income and consumption is

$$\text{MRS}^i(y_i, c_i) = \frac{\gamma}{w_i v'(c_i)}, \quad i = 1, \dots, n. \quad (2.5)$$

This marginal rate of substitution is independent of income, increasing in consumption, and decreasing in the skill level. Thus, preferences for income and consumption satisfy the standard single-crossing property.

An allocation is a vector $a = (y, c)$, where $y = (y_1, \dots, y_n)$ is the vector of pretax incomes and $c = (c_1, \dots, c_n)$ is the vector of consumptions. The government has the weighted utilitarian social welfare function $W: \mathbb{R}_+^{2n} \rightarrow \mathbb{R}$ given by

$$W(a) = \sum_{i=1}^n \mu_i u^i(y_i, c_i). \quad (2.6)$$

The welfare weights μ_i are assumed to be positive. We also assume that the skill-normalized welfare weights

$$\lambda_i = \mu_i/w_i, \quad i = 1, \dots, n, \quad (2.7)$$

are decreasing in the skill level. Thus, the skill-normalized weights satisfy

$$0 < \lambda_n < \dots < \lambda_1. \quad (2.8)$$

This assumption is satisfied if the welfare weights in (2.6) are all equal, i.e., if the objective function is utilitarian.

The government uses tax revenue for redistributive purposes only. Hence, the materials balance constraint for this economy is

$$\sum_{i=1}^n c_i \leq \sum_{i=1}^n y_i. \quad (2.9)$$

Our comparative static results are unaffected if the government also has a fixed revenue requirement. For example, if individuals also benefit from

public goods but their quantities have been predetermined, then the revenue needed to fund their provision is also predetermined. By Walras' Law, (2.9) is equivalent to the government's budget constraint that the total tax revenue is nonnegative. In this formulation of the government budget constraint, we are regarding an income transfer as a negative tax.

The government knows the functional form of the utility function u and it knows that there is one person with each of the skill levels, w_i , $i = 1, \dots, n$. It can observe each person's pretax income and, hence, can infer what his consumption is. However, the government cannot observe the hours worked or the type of any particular individual. In order to pursue its redistributive goals as embodied in the social welfare function, it raises its tax revenue using an anonymous income tax schedule.

An allocation is incentive compatible if it can be achieved by individuals choosing their most-preferred combinations of pretax income and consumption given some anonymous tax schedule. It is well-known (see, e.g., Guesnerie and Seade, 1982; Guesnerie, 1995) that an allocation a is incentive compatible if and only if it satisfies the following self-selection constraints:

$$u^i(y_i, c_i) \geq u^i(y_j, c_j), \quad i, j = 1, \dots, n. \quad (2.10)$$

In view of this equivalence, we can think of the government as choosing the allocation directly subject to the self-selection constraints (2.10), rather than indirectly using an income tax schedule.

An implication of the self-selection constraints is that both consumption and pretax income must be nondecreasing in the skill level. Furthermore, two individuals have the same consumption if and only if they have the same pretax income. When this is the case, they are said to be bunched.

The government's optimal nonlinear income tax problem is to choose an allocation $a \in \mathbb{R}_+^{2n}$ to maximize the social welfare function W in (2.6) subject to the materials balance constraint (2.9) and the self-selection constraints (2.10).

3. Comparative Statics

At a solution to the optimal income tax problem, the materials balance constraint holds with equality. Furthermore, the monotonicity condition on the skill-normalized welfare weights implies that each of the adjacent downward self-selection constraints binds; i.e., for all $i \geq 2$, person i is indifferent between what he receives and what person $i - 1$ receives. An allocation with

this property is called a monotonic chain to the left by Guesnerie and Seade (1982). Our monotonicity condition on the skill-normalized weights also implies that our problem is strictly redistributive in the sense of Guesnerie (1995, p. 224); i.e., in the neighbourhood of the optimal allocation, a redistribution of consumption from a higher-skilled individual to a lower-skilled individual and/or a redistribution in the other direction of pretax income is socially desirable (ignoring the self-selection constraints).³

Weymark (1986) has shown that it is possible to use these observations to develop explicit formulae for the optimal values of the consumptions and pretax incomes in our model. These formulae make use of the virtual wages β_i , $i = 1, \dots, n$, which can be written as

$$\beta_i = w_i + \left[\sum_{h=i+1}^n \frac{\mu_h}{w_h} - \frac{(n-i)}{n} \sum_{h=1}^n \frac{\mu_h}{w_h} \right] (w_{i+1} - w_i),$$

$$i = 1, \dots, n-1, \quad (3.1)$$

and

$$\beta_n = w_n. \quad (3.2)$$

Because the skill-normalized welfare weights are declining in type, $\beta_i < w_i$ for all $i \neq n$.

We assume that there is no bunching at the solution to the optimal tax problem and that the nonnegativity constraints on consumption and pretax income are not binding.⁵ When this is the case, person i 's optimal consumption is

$$c_i^* = v'^{-1}(\gamma/\beta_i), \quad i = 1, \dots, n, \quad (3.3)$$

and the optimal pretax incomes are given by

$$y_1^* = \frac{1}{n} \left\{ \sum_{j=1}^n c_j^* - \frac{1}{\gamma} \sum_{j=2}^n (n+1-j) w_j [v(c_j^*) - v(c_{j-1}^*)] \right\} \quad (3.4)$$

³Hellwig (2007) has recently introduced a weaker condition, “desirability of redistribution,” that also implies that the optimal allocation is a monotonic chain to the left.

⁴These expressions for the virtual wages are equivalent to those in equation (25) in Weymark (1986) once his normalization rule is adopted.

⁵Our comparative static results are unaffected if there is bunching provided that a marginal change in anyone’s skill does not affect the optimal pattern of bunching, which is generically the case.

and

$$y_i^* = y_1^* + \frac{1}{\gamma} \sum_{j=2}^n w_j [v(c_j^*) - v(c_{j-1}^*)], \quad i = 2, \dots, n. \quad (3.5)$$

Because the optimal allocation is a monotonic chain to the left, any tax schedule that generates this allocation is nondifferentiable at the pretax incomes chosen by all but the person with the highest skill level. For this reason, person i 's marginal tax rate is defined using his marginal rate of substitution by setting

$$\text{MTR}^i(y_i, c_i) = 1 - \text{MRS}^i(y_i, c_i) = 1 - \frac{\gamma}{w_i v'(c_i)}, \quad i = 1, \dots, n. \quad (3.6)$$

Because $\beta_n = w_n$ and $\beta_i < w_i$ for $i \neq n$, it follows from (3.3) and (3.6) that the optimal marginal tax rate is 0 for person n and it is strictly positive (but bounded above by 1) for everyone else.

Our findings are summarized in the following theorem.

Theorem. *If at the initial optimal allocation there is no bunching and the nonnegativity constraints on consumption and pretax income are not binding, then a marginal increase in w_i results in:*

- (i) *a decrease in c_j^* and an increase in $\text{MTR}^j(y_j^*, c_j^*)$, $j = 1, \dots, i - 1$;*
- (ii) *an increase in c_j^* and a decrease in $\text{MTR}^j(y_j^*, c_j^*)$, $j = i, \dots, n - 1$;*
- (iii) *no change in c_n^* if $i \neq n$, an increase in c_n^* if $i = n$, and no change in $\text{MTR}^n(y_n^*, c_n^*)$.⁷*

Because the marginal disutility of labour γ is fixed, from (3.3) we see that, for all i , the optimal consumption c_i^* only depends on the value of the virtual wage β_i . Hence, in order to provide some intuition for the comparative static results for the consumption variables, we first need to provide interpretations for the virtual wages and for the terms in the expressions for these wages in

⁶If $y_1^* > 0$, a necessary and sufficient condition for there to be no bunching and for all of the other nonnegativity constraints not to bind is $0 < \beta_1 < \dots < \beta_n$. If there is bunching, then the optimal consumption of individuals who are bunched together is given by (3.3) with β_i replaced by the average virtual wage of these individuals.

⁷Note that if $i = 1$, then part (i) does not apply, whereas if $i = n$, then part (ii) does not apply.

(3.1) and (3.2). We also need to determine how the value of each β_i is affected by an increase in someone's skill level.

As shown in Weymark (1986, p. 209), increasing c_i is socially desirable if and only if $\beta_i v'(c_i) - \gamma > 0$.⁸ Thus, β_i is the social value of increasing person i 's utility from consumption by one unit.

A unit increase in c_i can be financed by increasing i 's labour supply by $1/w_i$. If the self-selection constraints are ignored, this self-financing increase in consumption is worthwhile if and only if the utility gain $v'(c_i)$ from the extra consumption exceeds the utility loss γ/w_i due to the extra labour supplied. Equivalently, this change is worthwhile if and only if $w_i v'(c_i) - \gamma > 0$. Thus, in the absence of incentive effects, w_i is the social value of increasing person i 's utility from consumption by one unit. Increasing w_i therefore has a direct positive impact on β_i . For person n , this is the only effect that increasing his skill has on his virtual wage.

If $i \neq n$, in order to obtain the true social value β_i , we need to add the second term on the right-hand-side of (3.1), which is negative, so as to take account of the social cost that increasing i 's consumption imposes due to the tightening of the incentive constraints of the higher-skilled individuals. To see why this is the case, consider marginally increasing c_i , simultaneously increasing y_i so as to keep i 's utility unchanged. When $(i + 1)$'s adjacent downward incentive constraint binds, (2.4) implies that

$$w_{i+1}v(c_{i+1}) - \gamma y_{i+1} = w_i v(c_i) - \gamma y_i + (w_{i+1} - w_i)v(c_i),$$

$$i = 1, \dots, n - 1. \quad (3.7)$$

Thus, holding all other consumption levels fixed, a new monotonic chain to the left can be obtained by reducing the pretax income of each individual who is more productive than person i by $(w_{i+1} - w_i)v'(c_i)/\gamma$. For any individual h , the marginal utility of income is $-\gamma/w_h$. Thus, the social benefit of this change is $\sum_{h=i+1}^n \frac{\mu_h}{w_h} (w_{i+1} - w_i)v'(c_i)$. In this adjustment, the aggregate reduction in income is $(n - i)(w_{i+1} - w_i)v'(c_i)/\gamma$. Materials balance can be restored by increasing everyone's pretax income by $\frac{(n-i)}{n}(w_{i+1} - w_i)v'(c_i)/\gamma$ so as to make up for this loss in aggregate income. This adjustment creates a social loss of $\sum_{h=1}^n \frac{\mu_h}{w_h} \frac{(n-i)}{n} (w_{i+1} - w_i)v'(c_i)$. Hence, the net indirect social benefit of a marginal increase in c_i is $\left[\sum_{h=i+1}^n \frac{\mu_h}{w_h} - \frac{(n-i)}{n} \sum_{h=1}^n \frac{\mu_h}{w_h} \right] (w_{i+1} -$

⁸Therefore, the optimal c_i^* must satisfy $\beta_i v'(c_i^*) - \gamma = 0$, from which (3.3) follows.

$w_i)v'(c_i)$, which is negative. Therefore, the sum of the direct and indirect social benefits of increasing i 's utility from consumption by one unit is β_i .⁹

Increasing w_i flattens i 's indifference curves. Thus, for a given increase in c_i , a greater increase in y_i is required to maintain the indifference of individual i with his initial allocation. As a consequence, when w_i increases, the pretax incomes of the higher-skilled individuals do not need to be reduced by as much as before the increase in w_i in order to restore incentive compatibility when there is a utility compensated increase in c_i , nor do the pretax incomes need to be increased by as much in order to restore materials balance. Therefore, for $i \neq n$, an increase in w_i increases β_i by reducing the cost in the second term on the right-hand-side of (3.1), thereby reinforcing the direct effect on β_i of an increase in w_i .

Holding the marginal disutility of labour γ constant, from (3.3), it is optimal to increase c_i^* if and only if β_i is increased. Thus, person's i 's optimal consumption is increased when he becomes more productive.¹⁰ Furthermore, for $i \neq n$, because he has become more productive, we want to distort his choices less, and therefore his marginal tax rate is decreased.

For $j \neq i$, an increase in w_i has no direct effect on j 's virtual wage β_j , but, except for $j = n$, β_j is indirectly affected by the adjustments in pretax incomes described above that are used to ensure that the allocation is a monotonic chain to the left and that the materials balance constraint binds. This is most clearly seen from (3.1). In this expression, it is obvious that an increase in w_i has an impact on everyone's virtual wage except for the most-productive individual. In the Appendix, we show that when w_i increases, then β_j decreases for $j < i$ and it increases for $j > i$ when $j \neq n$. Thus, c_j^* decreases (resp. increases) for $j < i$ (resp. $j > i$ when $j \neq n$). Furthermore, because person j 's skill level has not changed and preferences are quasilinear in leisure, his marginal tax rate increases if and only if his consumption decreases when $j \neq n$. Thus, if person i 's skill increases, it is optimal to further distort the choices of the lower-skilled individuals by increasing their marginal tax rates and, except for the highest-skilled individual, who faces no marginal distortion, it is optimal to reduce the distortions on the higher-skilled individuals by decreasing their marginal tax rates.

Because it is optimal to change the consumption of everyone except for

⁹For further interpretation of the virtual wages, see Weymark (1986, pp. 210–211).

¹⁰For the highest-skilled individual, this conclusion follows immediately from the observation that because raising w_n flattens his indifference curves, it is necessary to increase his consumption in order to restore the value of his marginal rate of substitution to 1.

possibly person n when w_i is increased, it is not possible to obtain unambiguous comparative statics for the pretax incomes without making further assumptions on the model. The problem is that the expressions for the optimal incomes in (3.4) and (3.5) depend on $v(c_j^*) - v(c_{j-1}^*)$ for some values of j . When both c_{j-1}^* and c_j^* move in the same direction, the sign of the change in this expression cannot be determined.

For the quasilinear-in-consumption version of this model, Simula (2007) has shown that an increase in w_i leads to an increase in y_i^* , a decrease in y_{i-1}^* , and no change in anybody else's optimal pretax income. In addition, if $i \neq n$, then the only marginal tax rates that change are those of individuals i and $i - 1$, with the former decreasing and the latter increasing. If $i = n$, the only marginal tax rate that changes is that of person $n - 1$, and it increases. Thus, an increase in w_i has no effect on the optimal pretax incomes and marginal tax rates of anyone but individuals i and $i - 1$. In contrast, with quasilinear-in-leisure preferences, we have shown that increasing w_i affects the optimal consumptions and marginal tax rates of everyone except for possibly the highest-skilled individual. We conclude this section with some remarks that are aimed at clarifying the reasons for this asymmetry between the two models.

As we have seen, increasing a single skill level has a widespread impact on the optimal consumptions and marginal tax rates in our model because, for all $i \neq n$, the social value β_i of increasing person i 's utility from consumption by one unit depends on everyone's skills. This dependence arises because the social benefit of a compensated increase in i 's consumption is equal to the social value of the adjustments to everyone's pretax incomes required to restore both binding adjacent downward incentive constraints and a binding materials balance constraint. As a consequence, each individual's marginal utility of income, which depends on his skill, is used to help compute this social value.

With quasilinear-in-consumption preferences, the relevant social value is the social value of increasing an individual's utility from pretax income by one unit. The social benefit of a compensated increase in this individual's pretax income is equal to the social value of the adjustments to everyone's consumption required to restore the binding constraints. Because the marginal utility of consumption is a constant that is independent of type, these adjustments do not introduce a dependence on the skill levels into the social benefit calculation. It is nevertheless the case that an increase in w_i affects these social values for both individuals i and $i - 1$. As a consequence, the pretax in-

comes (and possibly the consumptions) of both i and $i - 1$ are modified in order to restore i 's binding adjacent downward incentive constraint when i 's indifference curves become less steep.

4. Concluding Remarks

Quasilinearity, in either of its forms, is a restrictive assumption to make about preferences. Ideally, one would like to obtain comparative static results for more general kinds of preferences. While we can make no claims to the robustness of our results to large departures from quasilinearity, we believe that our conclusions are quite robust to small departures from our form of quasilinearity. With concave utility functions, the optimal allocations and marginal tax rates are continuous functions of the skill levels, as is the case here. Except for possibly the highest-skilled individual, in our model, every optimal consumption and marginal tax rate changes in response to an increase in anyone's skill level. Continuity considerations suggest that the signs of these responses will be the same if preferences are approximately quasilinear in leisure. The same can not be said for the quasilinear-in-consumption model. In this case, except for at most two individuals, there is no adjustment to the optimal pretax incomes and marginal tax rates in response to an increase in someone's skill level. Continuity considerations do not rule out the possibility that an individual whose optimal pretax income and marginal tax rate are unaffected by a change in some skill level will see these variables adjusted if there is a small departure from quasilinearity-in-consumption preferences.

Appendix

Proof of the Theorem. We normalize the weights in (2.7) by setting

$$\sum_{i=1}^n \lambda_i = n. \tag{A.1}$$

Thus, the average skill-normalized welfare weight is 1. This normalization has no effect on the optimal allocation because the social ranking of allocations generated by the social welfare function in (2.6) only depends on the relative welfare weights. With this normalization, the expressions for the virtual

wages in (3.1) and (3.2) can be rewritten as

$$\beta_i = w_i + \left(i - \sum_{h=1}^i \lambda_h \right) (w_{i+1} - w_i), \quad i = 1, \dots, n, \quad (\text{A.2})$$

where w_{n+1} is an arbitrary number. See Weymark (1986, p. 209). Note that $i - \sum_{h=1}^i \lambda_h < 0$ for all $i \neq n$.

It is not possible to change a single skill level without violating the normalization rule in (A.1). However, because the solution to the optimal income tax problem only depends on the relative welfare weights in (2.6), the optimal response to changing someone's skill in the absence of the normalization rule is the same as the optimal response to first changing this skill and then scaling all the nonnormalized welfare weights μ_i proportionally until (A.1) is again satisfied. In the subsequent argument, we rescale the original welfare weights in this way so as to be able make use of (A.2).

From (A.2) and (3.3), we see that the skill levels only affect the value of c_i^* through their effects on β_i for all $i = 1, \dots, n$. Furthermore, because v is strictly concave, a marginal increase (resp. decrease) in β_i results in an increase (resp. decrease) in c_i^* . Therefore, in order to determine the impact of a change in a skill level on the optimal consumptions, we need to determine how the values of the virtual wages vary in response to a change in someone's skill.

Consider increasing w_i marginally and then proportionally increasing all of the welfare weights μ_h , $h = 1, \dots, n$, so that (A.1) holds. Because the sum of the skill-normalized welfare weights is fixed, as a result of these changes, λ_i decreases and λ_j increases for $j \neq i$. Furthermore, $\sum_{h=1}^j \lambda_h$ increases for $j < i$, $\sum_{h=1}^j \lambda_h$ decreases for $j \geq i$ if $j \neq n$, and $\sum_{h=1}^n \lambda_h$ is unchanged. We consider the cases in which $i \neq n$ and $i = n$ separately.

Case 1. $i \neq n$.

From (A.2), for $j \leq i - 2$, the change in w_i only affects β_j through its impact on $\sum_{h=1}^j \lambda_h$. Because $\sum_{h=1}^j \lambda_h$ increases, β_j decreases and, hence, c_j^* decreases. Similarly, if $j \neq n$ and $j \geq i + 1$, both β_j and c_j^* increase because $\sum_{h=1}^j \lambda_h$ decreases.

For $j = n$, $\beta_n = w_n$, which has not changed. Hence, there is no change in c_n^* .

For $j = i - 1$, $(i - 1) - \sum_{h=1}^{i-1} \lambda_h$ decreases; i.e., it increases in absolute value. Because $w_i - w_{i-1}$ is positive and increasing in w_i , the second term on

the right-hand-side of the formula for β_{i-1} decreases. Hence, both β_{i-1} and c_{i-1}^* decrease.

For $j = i$, $i - \sum_{h=1}^i \lambda_h$ increases; i.e., it decreases in absolute value. Furthermore, $w_{i+1} - w_i$ is positive and decreasing in w_i . Thus, the second term on the right-hand-side of (A.2) increases. Because the first term is simply w_i , it follows that β_i and, hence, c_i^* increases.

Turning now to the marginal tax rates, for all $j \neq i$, from (3.6), because v is strictly concave, j 's marginal tax rate increases if and only if c_j^* decreases. Thus, $\text{MTR}^j(y_j^*, c_j^*)$ increases for $j < i$, decreases for $j > i$ if $j \neq n$, and does not change if $j = n$.

From (A.2) and (3.3), we have

$$\begin{aligned} w_i v'(c_i^*) &= \frac{w_i \gamma}{\beta_i} \\ &= \frac{\gamma}{1 + \left[i - \sum_{h=1}^i \lambda_h \right] \left[\frac{w_{i+1}}{w_i} - 1 \right]}. \end{aligned} \tag{A.3}$$

As we have seen, when w_i is marginally increased and all the welfare weights μ_h are adjusted so as to satisfy (A.1), the first term in square brackets in the denominator of the last expression in (A.3) increases; i.e., it decreases in absolute value. The second term in square brackets is positive and decreasing in w_i . Thus, $w_i v'(c_i^*)$ decreases. It then follows from (2.5) and (3.6) that $\text{MRS}^i(y_i^*, c_i^*)$ increases and $\text{MTR}^i(y_i^*, c_i^*)$ decreases.

Case 2. $i = n$.

For all $j \neq n$, the proof is identical to the argument in Case 1. Because $\beta_n = w_n$, the increase in w_n results in an increase in c_n^* . Furthermore, because $\text{MTR}^n(y_n^*, c_n^*)$ is always 0, there is no change in n 's marginal tax rate. \square

References

- Boadway, R., Pestieau, P., 2006. Tagging and redistributive taxation. *Annales d'Économie et de Statistique*, forthcoming.
- Boone, J., Bovenberg, L., 2007. The simple economics of bunching: Optimal taxation with quasi-linear preferences. *Journal of Public Economic Theory* 9, 89–105.
- Brett, C., Weymark, J. A., 2008. Public good provision and the comparative statics of optimal nonlinear income taxation. *International Economic Review*, forthcoming.

- Cebreiro-Gómez, A., 2002. Optimal income taxation, production efficiency and mobility of labour, unpublished manuscript, Department of Economics, University of Essex.
- Diamond, P. A., 1998. Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review* 88, 83–95.
- Guesnerie, R., 1995. *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, Cambridge.
- Guesnerie, R., Seade, J., 1982. Nonlinear pricing in a finite economy. *Journal of Public Economics* 17, 157–179.
- Hamilton, J., Pestieau, P., 2005. Optimal income taxation and the ability distribution: Implications for migration equilibria. *International Tax and Public Finance*, 29–45.
- Hellwig, M., 2007. A contribution to the theory of optimal utilitarian income taxation. *Journal of Public Economics* 91, 1449–1477.
- Lollivier, S., Rochet, J.-C., 1983. Bunching and second-order conditions: A note on optimal tax theory. *Journal of Economic Theory* 31, 392–400.
- Mirrlees, J. A., 1971. An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38, 175–208.
- Mussa, M., Rosen, S., 1978. Monopoly and product quality. *Journal of Economic Theory* 18, 301–317.
- Simula, L., 2007. Optimality conditions and comparative static properties of non-linear income taxes revisited. Working Paper No. 2007-15, Paris School of Economics.
- Weymark, J. A., 1986. 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.