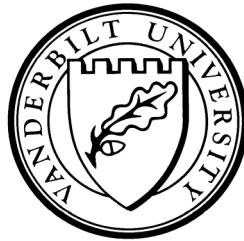


**COMPARATIVE STATICS OF OPTIMAL NONLINEAR INCOME TAXATION
IN THE PRESENCE OF A PUBLICLY PROVIDED INPUT**

by

Craig Brett and John A. Weymark



Working Paper No. 09-W10

August 2009

DEPARTMENT OF ECONOMICS
VANDERBILT UNIVERSITY
NASHVILLE, TN 37235

www.vanderbilt.edu/econ

Comparative Statics of Optimal Nonlinear Income Taxation in the Presence of a Publicly Provided Input*

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, Vanderbilt University,
VU Station B #35189, 2301 Vanderbilt Place,
Nashville, TN 37235-1819, U.S.A.
(e-mail: john.weymark@vanderbilt.edu)

August 2009

*Craig Brett's research is generously supported by the Canada Research Chair Programme. This project was supported by the Social Sciences and Humanities Research Council of Canada. We thank Robin Boadway, Katherine Cuff, Firhouz Gavhari, Richard Harris, Fernando Ruiz, and participants at the Public Economics conference in honor of Robin Boadway held at Queen's University, the 2009 Canadian Economics Association Meetings, and PET09 Galway for helpful comments.

Abstract

“Comparative Statics of Optimal Nonlinear Income Taxation
in the Presence of a Publicly Provided Input”

by

Craig Brett and John A. Weymark

Comparative static properties of the solution to an optimal nonlinear income tax problem are provided for a model in which the government both designs a redistributive income tax schedule and provides a public input into a nonlinear production process. These assumptions imply that wage rates are endogenous. The endogeneity of the wages necessitates taking account of general equilibrium effects of changes in the parameters of the model that are not present when the technology is linear.

Journal of Economic Literature classification numbers: D82, H21.

Keywords and phrases: asymmetric information, comparative statics, optimal income taxation, publicly provided inputs.

1. Introduction

The study of optimal nonlinear income taxation focuses on the tension between a government's assumed desire to set taxes according to an ability-to-pay criterion and the practical reality that the government cannot directly observe anyone's ability to pay. In order to focus attention on the tradeoffs required to reconcile this tension and the concomitant economic distortions, much of the literature on optimal nonlinear income taxation follows the lead of Mirrlees (1971) by assuming that the sole purpose of taxation is to redistribute income, typically from individuals with higher abilities-to-pay to individuals with lower abilities-to-pay. While redistribution is undoubtedly a significant component of what governments do, the provision of various kinds of goods and services features prominently on their agendas. These goods and services may be primarily of value as consumption goods, both public goods *per se* and publicly-provided private goods, or they may be publicly-provided inputs into production, such as infrastructure. In this article, we derive comparative static properties for an optimal nonlinear tax problem in which the government provides inputs into the production process in addition to redistributing income.

The literature on the interactions between optimal nonlinear taxation and governmental provision of consumption goods is well-developed. One of the key insights in this literature is that judicious deviations from first-best allocation rules can, in certain circumstances, be used to implicitly redistribute income, thereby providing a useful supplement to optimal distortionary income taxes. Christiansen (1981) and Boadway and Keen (1993) describe when deviations from the Samuelson (1954) Rule for the provision of public goods are justified on these grounds, while Boadway and Marchand (1995) describe the circumstances under which public provision of a private good is merited even in the presence of optimal nonlinear income taxes. A central feature of this class of arguments is the possibility that individuals of different abilities have different responses to public expenditures in their consumption-leisure choices. These diverse responses provide the government with additional information concerning abilities-to-pay, allowing it to carry out redistribution more effectively.¹

The study of interactions between distortionary income taxation and the provision of public inputs is perhaps less prominent in the literature. Gaube (2005) argues that the link between publicly provided inputs and redistributive income taxes, if one exists, must be more indirect because the provision of inputs has no direct influence on individual consumption or labor supply decisions. He shows that it is optimal to deviate from first-best public input decisions when the relative wages of different types of workers depend on the level of the publicly provided input. The resulting production inefficiency is justified by the implicit redistribution afforded by increasing the relative wages of less

¹When observable behavior is independent of public provision, as in, for example the Boadway–Keen model under the assumption of a common utility function that is weakly separable between consumption and labor supply, first-best provision rules remain optimal.

able workers.²

We develop a model of an economy with an arbitrary, finite number of individuals who only differ in labor productivities. There may be several individuals with the same labor productivity and the number of individuals may vary from skill class to skill class. All individuals have the same preferences over a single private consumption good and leisure. Unambiguous comparative static results can be obtained when these preferences are quasilinear. For concreteness, we assume that these preferences can be represented by a quasilinear-in-leisure utility function, as in Weymark (1987). Following Gaube (2005), our model features a strictly convex aggregate production technology, thereby abstracting from the issue of whether the first-best provision rule is marginal cost pricing or the Samuelson-like rules for the provision of a public input derived by Kaizuka (1965) and Sandmo (1972). The aggregate technology transforms total labor time in efficiency units and a publicly provided input into an output good. The output good can be either consumed or transformed into the publicly provided input at a constant marginal cost. The government simultaneously chooses a nonlinear income tax schedule and a level of the publicly provided input to maximize a weighted utilitarian social welfare function subject to incentive compatibility constraints and an economy-wide resource constraint.

As noted by Lollivier and Rochet (1983) for a model with a continuum of skill types and by Weymark (1987) with discrete types, it is possible to solve the optimal nonlinear income tax problem in two stages when preferences are quasilinear in leisure and the aggregate technology is linear. In the first stage, provided that it is optimal to separate individuals with different labor productivities by having them consume different amounts of the consumption good, a reduced-form unconstrained maximization problem is solved to determine the optimal allocation of consumption. The allocation of before-tax income (labor supply) is determined in a second stage. It is not possible to fully replicate the Lollivier–Rochet–Weymark argument when the technology is not linear. However, when it is optimal to separate different skill classes, it is possible to formulate a first-stage problem describing the choice of consumption and input allocations as arising out of a maximization problem constrained only by the economy-wide resource constraint. In this reduced-form problem, the individual utilities from consumption are summed using reduced-form welfare weights that depend on the original weights from the social welfare function, as well as some of the other parameters of our model.

We employ techniques borrowed from the theory of consumer demand, also used by Brett (2009), to derive comparative static results for our reduced-form problem. Our comparative static analysis focuses on the effects of changes in the following variables: the welfare weights in the reduced-form problem, a measure of the disutility of working, and the marginal cost of the publicly provided input. We also provide some limited comparative static results for the welfare weights in the social welfare function. We investigate how the optimal individual consumption levels, the aggregate effective labor supply, the provision of the publicly provided input, the shadow value of the resource

²Similar justifications for production inefficiency in models of optimal nonlinear income taxation are provided, albeit in other contexts, by Naito (1999) and Blackorby and Brett (2004).

constraint, and the implicit marginal tax rates respond to changes in these parameters.

The assumptions we make about the technology imply that relative wages do not vary with the level of the publicly provided input. Thus, there is production efficiency in our model. On the other hand, the wage paid per unit of effective labor does change as the model parameters vary. These wage effects lead to changes in optimal production and consumption plans that are not present in models of nonlinear income taxation with *linear* production possibilities frontiers, like the ones analyzed by Weymark (1987), Brett and Weymark (2008a,b), and Simula (2007).³ In spite of the existence of the extra effects, with two exceptions, we are able to sign the comparative static responses of the endogenous variables listed above to marginal changes in each of the parameters that appear in the reduced-form problem.

In Section 2, we present our model and describe the government's decision problem. We derive and characterize the solution to our reduced-form of the government's problem in Section 3. In Section 4, we conduct our comparative static exercises. We offer some concluding remarks in Section 5. Our proofs are gathered in an Appendix.

2. Model

The economy is populated by N types of individuals, where an individual of type i has skill level $s_i > 0$. The number of individuals of type i is $n_i > 0$. The types are numbered so that $s_1 < s_2 < \dots < s_N$. A type i individual's skill level measures the rate at which his labor time, l_i , is transformed into his effective labor supply, y_i . Specifically,

$$y_i = s_i l_i, \quad i = 1, \dots, N. \quad (2.1)$$

The private sector uses an aggregate production function, f , to transform R units of a publicly provided input and y units of effective labor into x units of output; that is,

$$x = f(R, y), \quad (2.2)$$

where f is twice continuously differentiable and strictly concave with $f(R, 0) = f(0, y) = 0$ for all nonnegative R and y . We assume that effective labor and the publicly provided input are complements in production in the sense that $f_{yR}(R, y) > 0$ for all input combinations.⁴ The output good may be used for consumption or transformed into the public input according to the constant marginal rate of technical substitution process given by

$$c + qR \leq x, \quad (2.3)$$

where c is the quantity of the consumption good and q is the opportunity cost of one unit of the public input using the consumption good as the numeraire. Thus, the aggregate production possibility set for this economy is characterized by

$$c + qR \leq f(R, y). \quad (2.4)$$

³Simula (2007) assumes that preferences are quasilinear in consumption.

⁴This assumption is satisfied when the technology exhibits constant returns to scale. Most of our results do not rely on the assumption that the inputs are complementary in production.

The aggregate production possibility set described by (2.4) bears some relation to the technology posited by Keen and Marchand (1997) in their study of the effects of fiscal competition on the mix of public spending.

The analysis that follows extends easily to the case in which the public input is produced at increasing marginal cost with, say, a strictly convex cost function $qC(R)$, where the parameter q is included as a way to capture shifts in marginal cost. This cost function can be derived from a *public sector* production technology that uses effective labor as its only input.⁵ The cost function approach abstracts from the allocation of effective labor between the public and private spheres. While this is a potentially important issue, it is beyond the scope of our analysis. Moreover, if the public sector technology uses effective labor as an input, then the government has no means by which to change relative wage rates, rendering the public-private division of labor of no redistributive consequence.

There is perfect competition in both input and output markets so that producer prices are equal to their respective marginal rates of transformation. In particular, the aggregate wage paid to effective labor, w , is

$$w = \frac{\partial f(R, y)}{\partial y}. \quad (2.5)$$

The before-tax income of an individual of type i is given by

$$z_i = wy_i = ws_i l_i, \quad i = 1, \dots, N. \quad (2.6)$$

All individuals have a common, cardinally significant utility function representing preferences that are quasilinear in leisure given by

$$V(c, l) = v(c) - \gamma l, \quad (2.7)$$

where $\gamma > 0$. The function v is assumed to be twice continuously differentiable at all $c \neq 0$, continuous and nondecreasing on \mathbb{R}_+ , strictly increasing on \mathbb{R}_{++} , and strictly concave on \mathbb{R}_{++} with $v(0) = 0$, $v_c(0) = \infty$, and $v_c(c) \rightarrow 0$ as $c \rightarrow \infty$. The limiting assumptions on v ensure that the optimal tax problem has a solution and that individuals of all types have positive consumption at this solution. The parameter γ measures the marginal disutility of labor.

Following Weymark (1986b, 1987), we conveniently represent the preferences of a type i individual by the type-specific monotonic transformation of (2.7) given by

$$U^i(c_i, y_i) = s_i v(c_i) - \gamma y_i, \quad i = 1, \dots, N, \quad (2.8)$$

where c_i is his consumption. Equation (2.8) describes preferences over consumption and effective labor supply. The marginal rate of substitution between effective labor and

⁵By concentrating on an aggregate technology, we also abstract from some thorny issues in the theory of public infrastructure. We are ruling out increasing returns to scale in the aggregate, which can be a characteristic of “factor-augmenting” public inputs. Manning, Markusen, and McMillan (1985), Feehan and Matsumoto (2000), and Gaube (2005) contain detailed treatments of these issues.

consumption for an individual of type i is

$$\text{MRS}^i(c_i, y_i) = \frac{\gamma}{s_i v'(c_i)}, \quad i = 1, \dots, N. \quad (2.9)$$

This marginal rate of substitution is decreasing in the skill level. Thus, preferences for income and consumption satisfy the standard single-crossing property. The representation of preferences given by (2.8) is linear in y and in the unobserved characteristic s . This linearity is heavily exploited in the analysis of Section 3.

As is common in models of nonlinear income taxation, for all i , the government can observe both c_i and z_i , but cannot observe l_i or s_i . It can observe the aggregate wage rate w , so that it can infer y_i at the individual level. Because l_i is unobserved, the government uses distortionary income taxes. The tax system specifies tax payments as a function of observed labor income. Equivalently, the government can be viewed as selecting consumption levels and effective labor time for each type of worker, subject to the standard incentive compatibility constraints

$$s_i v(c_i) - \gamma y_i \geq s_j v(c_j) - \gamma y_j, \quad i, j = 1, \dots, N. \quad (2.10)$$

It is well known that the self-selection conditions imply that the allocation of consumption and effective labor must satisfy

$$c_1 \leq c_2 \leq \dots \leq c_N \quad (2.11)$$

and

$$y_1 \leq y_2 \leq \dots \leq y_N. \quad (2.12)$$

The tax system consistent with an allocation satisfying (2.10) is typically nondifferentiable. Thus, marginal tax rates are only implicitly defined by the difference between producer and consumer prices at an individual's consumption bundle. The implicit marginal tax rate (IMTR) for labor income applicable to type i is given by

$$\text{IMTR}_i = 1 - \frac{\gamma}{w s_i v'(c_i)}, \quad i = 1, \dots, N. \quad (2.13)$$

An allocation is a vector $a = (y_1, \dots, y_N, c_1, \dots, c_N, R)$ consisting of the effective labor supply and consumption of each type of worker and a level of the publicly provided input. A production-feasible allocation satisfies

$$\sum_{i=1}^N n_i c_i + qR \leq f(R, y), \quad (2.14)$$

where

$$y = \sum_{i=1}^N n_i y_i \quad (2.15)$$

is the aggregate supply of effective labor. Inequality (2.14) is the materials balance constraint for this economy. In contrast to previous comparative static analyses of optimal nonlinear taxation, this constraint is nonlinear. It is this nonlinearity that prevents us from simply applying the methodology employed in previous studies without modification.

The government has the weighted utilitarian social welfare function $W: \mathbb{R}_+^{2N} \rightarrow \mathbb{R}$ given by

$$W(a) = \sum_{i=1}^N \mu_i n_i V(c_i, y_i/s_i) = \sum_{i=1}^N \lambda_i n_i [s_i v(c_i) - \gamma y_i] \quad (2.16)$$

for a collection of positive welfare weights $\mu = (\mu_1, \dots, \mu_n)$, where the skill-normalized welfare weights

$$\lambda_i = \mu_i/s_i, \quad i = 1, \dots, N, \quad (2.17)$$

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

$$0 < \lambda_N < \dots < \lambda_1. \quad (2.18)$$

This assumption is satisfied if the objective function is utilitarian, that is, if the weights μ_i are all equal. Because any welfare maximization problem is invariant to multiplying the social welfare function by an arbitrary constant, we assume that the normalized welfare weights sum to the total number of individuals in the economy; that is,

$$\sum_{i=1}^N n_i \lambda_i = \sum_{i=1}^N n_i. \quad (2.19)$$

The government's decision problem is defined formally as follows.

The Optimal Nonlinear Tax Problem. The government chooses an allocation a to maximize the social welfare function (2.16) subject to the self-selection constraints (2.10) and the materials balance constraint (2.14).

By using the economy-wide materials balance constraint as the government budget constraint, we are assuming that all pure profits are fully taxed. In stating the Optimal Nonlinear Tax Problem, we have not explicitly included nonnegativity constraints on the allocation vector a . Provided that $y_1 > 0$ at the solution to this problem, our assumptions ensure that all components of the optimal allocation are positive. *Henceforth, it is assumed that the optimal value of y_1 is positive.*

3. Preliminary Analysis

Monotonicity of the skill-normalized welfare weights implies that the government wishes to redistribute consumption toward and/or redistribute effective labor time away from lower-skilled individuals. The natural limit to this type of redistribution is a downward self-selection constraint. Lemma 1 demonstrates that, in fact, it is optimal for all of the downward adjacent self-selection constraints to bind.

Lemma 1. *At a solution a to the optimal nonlinear income tax problem:*

$$s_i v(c_i) - \gamma y_i = s_i v(c_{i-1}) - \gamma y_{i-1}, \quad i = 2, \dots, N. \quad (3.1)$$

The binding self-selection constraints (3.1) form a system of $N - 1$ linear equations in the N variables y_1, \dots, y_N . Given an aggregate supply of effective labor, y , (2.15) provides an N th linear equation in the y_i s. The solution to the resulting system of equations is given in Lemma 2.

Lemma 2. *For a given (c_1, \dots, c_n, y) , the system of equations (2.15) and (3.1) have a unique solution. Moreover, this solution can be written in the recursive form:*

$$y_1(c_1, \dots, c_n, y) = \frac{1}{\sum_{i=1}^N n_i} \left(y - \frac{1}{\gamma} \sum_{j=2}^N \sum_{i=j}^N n_i s_j [v(c_j) - v(c_{j-1})] \right); \quad (3.2)$$

$$y_i(c_1, \dots, c_n, y) = y_1(c_1, \dots, c_n, y) + \frac{1}{\gamma} \sum_{j=2}^i s_j [v(c_j) - v(c_{j-1})], \quad i = 2, \dots, N. \quad (3.3)$$

Lemmas 1 and 2 imply that the optimal nonlinear tax problem can be solved in two steps. In the first step, (3.2) and (3.3) can be substituted into the social welfare function (2.16). The resulting reduced-form objective function depends on consumption levels and aggregate effective labor. Maximizing this objective function subject to the production-feasibility constraint (2.14) and the consumption monotonicity constraint (2.11) yields the optimal values $(c_1^*, \dots, c_N^*, y^*, R^*)$. In the second step, Lemma 2 is used to compute the optimal effective labor supplies for each type of individual.

Lemma 3. *The optimal consumption vector, optimal aggregate effective labor, and optimal level of the public input for the Optimal Nonlinear Tax Problem are unique and can be found by solving*

$$\max_{c_1, \dots, c_N, y, R} \sum_{i=1}^N \beta_i v(c_i) - \gamma y \quad \text{subject to (2.11) and (2.14)}, \quad (3.4)$$

where

$$\begin{aligned} \beta_i &= n_i s_i + \left(\sum_{k=1}^i (n_k - n_k \lambda_k) \right) (s_{i+1} - s_i) \\ &= n_i s_i + \left(\sum_{k=i+1}^N (n_k \lambda_k - n_k) \right) (s_{i+1} - s_i), \quad i = 1, \dots, N, \end{aligned} \quad (3.5)$$

and s_{N+1} is an arbitrary number.⁶ Furthermore, at the solution to (3.4), (2.14) binds.

⁶Note that the normalization (2.19) implies $\beta_N = n_N s_N$.

Henceforth, we assume that the monotonicity constraints on consumption (2.11) are all non-binding. That is, we rule out the possibility of bunching at the optimal solution.⁷ Alternatively, our comparative static results can be re-interpreted as applying to parameter changes that leave the pattern of bunching unchanged.

The problem (3.4) is considerably more tractable than the original statement of the Optimal Nonlinear Tax Problem. However, unlike the reduced forms obtained by Weymark (1986b), Brett and Weymark (2008b), and Simula (2007), even when it is assumed that the monotonicity constraints are not binding, (3.4) is not a fully unconstrained optimization problem. As is standard in models in which all pure profits are taxed, it is optimal to have aggregate production efficiency, so the production-feasibility constraint binds. However, the nonlinearity of this constraint makes it inconvenient to substitute it into the objective function. Characterizing the solution to and performing comparative static analysis concerning (3.4) is, nevertheless, fairly straightforward.

Introducing a multiplier ψ , the shadow value of the constraint (2.14), allows the first-order conditions for a solution to (3.4) to be written as

$$c_i: \beta_i v'(c_i) - \psi n_i = 0, \quad i = 1, \dots, N; \quad (3.6)$$

$$y: -\gamma + \psi f_y = 0; \quad (3.7)$$

$$R: f_R - q = 0; \quad (3.8)$$

$$\psi: f(R, y) - \sum_{i=1}^N n_i c_i - qR = 0.^8 \quad (3.9)$$

The first-order conditions have a recursive structure that greatly simplifies our analysis. Suppose that one can, perhaps by using information contained in all of equations (3.6)–(3.9), find the optimal value of the multiplier associated with the resource constraint, say $\tilde{\psi}$. Substituting $\tilde{\psi}$ into the first-order conditions (3.6), (3.7), and (3.8) renders each equation in (3.6) independent of the other of these first-order conditions. Thus, conditional on $\tilde{\psi}$, the optimal value of c_i can be found by solving the first-order condition associated with c_i . In addition, given $\tilde{\psi}$, the optimal values of y and R can be found by solving the two-equation system (3.7) and (3.8).

With the help of the first-order conditions (3.6) and (3.7), it is possible to interpret β_i as a reduced-form welfare weight applied to the utility from consumption of an individual of type i . This welfare weight takes account of both the effect of c_i on the utility of individuals of type i and, when $i \neq N$, the effects that an increase in c_i has on the self-selection constraints. By (2.4) and (2.5), a unit increase in c_i for each person of type i can be financed by increasing each of their effective labor supplies by $1/w$. When there are no self-selection constraints, this self-financing increase in consumption is worthwhile

⁷Conditions that guarantee that bunching does not occur at the optimum can be derived using the arguments found in Weymark (1986a) and Simula (2007).

⁸Our assumptions on f and v imply that the c_1, \dots, c_N, y , and R that solve (3.4) are all positive. Thus, because the production-feasibility constraint binds, the first-order conditions (3.6)–(3.9) are all equalities.

if and only if the utility gain $n_i s_i v'(c_i)$ from the extra consumption exceeds the utility loss $n_i \gamma / w$ due to the extra effective labor supplied, where utility has been cardinalized as in (2.8). That is, this change is worthwhile if and only if $n_i s_i v'(c_i) - \gamma n_i / w > 0$. Thus, in the absence of incentive effects, $n_i s_i$ is the social value of increasing the utility of consumption for every person of type i by one unit.

When $i \neq N$, the effects of an increase in c_i on the incentive compatibility conditions are captured by the final term on the right-hand side of (3.5).⁹ To see why this is the case, consider marginally increasing c_i , simultaneously increasing y_i so as to keep i 's utility unchanged. This change results in a violation of the downward adjacent self-selection constraint of type $i + 1$. So as to isolate the indirect effect of a self-financing marginal increase in c_i , it is necessary to restore feasibility without making any further change in the aggregate effective labor supply. This is accomplished in two steps. In the first step, the effective labor supply of each individual who is more productive than those of type i is decreased by $(s_{i+1} - s_i)v'(c_i)/\gamma$. Because preferences are quasilinear in effective labor, as a result of this adjustment, the self-selection constraints again bind. With the utility cardinalization (2.8), the marginal utility of effective labor for any individual is $-\gamma$. Thus, the social benefit of this change is $\sum_{k=i+1}^N n_k \lambda_k (s_{i+1} - s_i)v'(c_i)$. In the second step, everyone's effective labor is increased by $\frac{\sum_{k=i+1}^N n_k (s_{i+1} - s_i)v'(c_i)}{\gamma \sum_{k=1}^N n_k}$ so that in the aggregate there is no change in effective labor supply when both adjustments are implemented. The adjustments in the second step do not affect the self-selection constraints. The second-stage adjustment results in a social loss of $\frac{\sum_{k=1}^N n_k \lambda_k}{\sum_{k=1}^N n_k} \sum_{k=i+1}^N n_k (s_{i+1} - s_i)v'(c_i)$. Hence, using the normalization rule (2.19), the net indirect social benefit of a marginal increase in c_i is $\left[\left(\sum_{k=i+1}^N (n_k \lambda_k - n_k) \right) (s_{i+1} - s_i) \right] v'(c_i)$, which is negative. It then follows that when both the direct and indirect effects of increasing c_i are taken into account, the social value of increasing the utility from consumption of each person of type i by one unit is given by the right-hand side of (3.5).

Proposition 1 summarizes the qualitative features of the optimal allocations that follow directly from the first-order conditions.

Proposition 1. *The following statements hold at the solution a to the Optimal Nonlinear Tax Problem.*

- (i) *The marginal product of the publicly provided input equals its price.*
- (ii) *The labor supply of individuals of type N is not distorted; that is,*

$$\text{IMTR}_N = 1 - \frac{\beta_N}{n_N s_N} = 0. \quad (3.10)$$

- (iii) *The implicit marginal tax rate on the labor income of individuals of types $1, \dots, N - 1$*

⁹As previously noted, this term is zero when $i = N$.

is positive; specifically,

$$\begin{aligned} \text{IMTR}_i &= 1 - \frac{\beta_i}{n_i s_i} \\ &= \frac{1}{n_i s_i} \left(\sum_{k=1}^i (n_k \lambda_k - n_k) \right) (s_{i+1} - s_i) > 0, \quad i = 1, \dots, N-1. \end{aligned} \tag{3.11}$$

Part (i) of Proposition 1 states that there is no distortion in the provision of the publicly provided input. Gaube (2005) argues that distortions in publicly provided inputs are justified when relative wages vary with the level of the publicly provided input R . In that case, R provides a mechanism to carry out implicit redistribution. However, when relative wages are fixed, as they are here, changing R cannot enhance redistribution, so there is no reason to deviate from the first-best allocation rule for the provision of the public input. Parts (ii) and (iii) of Proposition 1 demonstrate that the standard pattern of labor market distortions arising in redistributive optimal nonlinear tax schemes apply here: no distortion at the top and positive marginal income tax rates for all other types of individuals.

4. Comparative Statics

We now investigate how the optimal individual consumption levels, the aggregate effective labor supply, the provision of the publicly provided input, the shadow value of the resource constraint, and the implicit marginal tax rates respond to changes in some of the parameters of the economy. Except for the implicit marginal tax rates, these are the endogenous variables in the Lagrangian for the first-stage optimization problem (3.4).¹⁰ We provide complete comparative static results for all of these variables with respect to the taste parameter γ and the reduced-form welfare weights β_1, \dots, β_N . Except for the two inputs to the production process, we are also able to sign the effects of a change in the technology parameter q on these variables. We let $\rho = (\beta_1, \dots, \beta_N, q, \gamma)$ denote the vector of these parameters.

We also offer some partial results for the effects of changing the skill-normalized welfare weights, $\lambda_1, \dots, \lambda_N$.¹¹ These welfare weights only affect the solution to (3.4) through their influence on the reduced-form welfare weights β_1, \dots, β_N that appear in the objective function in (3.4). It is not possible to change only the welfare weight λ_i without violating the normalization constraint (2.19). However, it is possible to imagine a marginal increase in λ_i accompanied by a proportional reduction in λ_j for all $j \neq i$ sufficient to maintain (2.19). Henceforth, whenever we refer to an increase in λ_i , we

¹⁰As discussed by Weymark (1987) and Brett and Weymark (2008a), it is generally not possible to obtain unambiguous comparative static results for individual incomes when preferences are quasilinear in leisure, as is the case here.

¹¹For given values for the skills s_1, \dots, s_N , this is equivalent to performing comparative statics with respect to the original welfare weights μ_1, \dots, μ_N .

assume that all other λ_j are simultaneously rescaled in this way. Such a change affects all of the reduced-form welfare weights except for β_N , which is why few clear-cut comparative static results are available for changes in the welfare weights.¹² Note that, because the solution to the Optimal Nonlinear Tax Problem only depends on the relative welfare weights, our procedure is equivalent to increasing only the type i welfare weight in the absence of our normalization.

We begin our analysis by determining the effects of an increase in the welfare weight λ_i on the reduced-form welfare weights. Our findings are summarized in Lemma 4.

Lemma 4. *A marginal increase in λ_i (with a proportional reduction in λ_j , $j \neq i$) induces:*

- (i) *an increase in β_j for $j = 1, \dots, i - 1$.*
- (ii) *a decrease in β_j for $j = i, \dots, N - 1$.*
- (iii) *no change in β_N .*

The techniques we use to compute comparative static effects recognize the joint determination of all the endogenous variables in the system of first-order equations (3.6)–(3.9). The formal justification for our comparative statics procedure is given in Proposition 2.

Proposition 2. *The optimality conditions (3.6)–(3.9) define a continuously differentiable solution function $F: \mathbb{R}_+^{N+2} \rightarrow \mathbb{R}_{++}^{N+3}$ for the problem (3.4), where, for all $\rho \in \mathbb{R}_+^{N+2}$, $F(\rho) = (\tilde{c}_1(\rho), \dots, \tilde{c}_N(\rho), \tilde{y}(\rho), \tilde{R}(\rho), \tilde{\psi}(\rho))$. For all $\rho \in \mathbb{R}_+^{N+2}$, the derivative DF of F at ρ is given by*

$$DF(\rho) = (A^{-1}B)(\rho), \quad (4.1)$$

where

$$A(\rho) = \begin{bmatrix} \beta_1 v''(c_1) & 0 & \cdots & \cdots & 0 & 0 & 0 & -n_1 \\ 0 & \beta_2 v''(c_2) & 0 & \cdots & 0 & 0 & 0 & -n_2 \\ \vdots & 0 & \ddots & & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & 0 & 0 & 0 & \vdots \\ 0 & 0 & \cdots & 0 & \beta_N v''(c_N) & 0 & 0 & -n_N \\ 0 & 0 & \cdots & \cdots & 0 & \psi f_{yy} & \psi f_{yR} & f_y \\ 0 & 0 & \cdots & \cdots & 0 & f_{yR} & f_{RR} & 0 \\ -n_1 & -n_2 & \cdots & \cdots & -n_N & f_y & 0 & 0 \end{bmatrix} \quad (4.2)$$

¹²We do not consider changes in the the skill parameters s_1, \dots, s_N or the demographic parameters n_1, \dots, n_N . In general, each of them also affects more than one of the reduced-form welfare weights. In addition, they also enter into problem (3.4) through their effect on the production-feasibility constraint (2.14).

and

$$B(\rho) = \begin{bmatrix} -v'(c_1) & 0 & \cdots & \cdots & 0 & 0 & 0 \\ 0 & -v'(c_2) & 0 & \cdots & 0 & 0 & 0 \\ \vdots & 0 & \ddots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & -v'(c_N) & 0 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & R & 0 \end{bmatrix}, \quad (4.3)$$

with all expressions on the right-hand sides of (4.2) and (4.3) evaluated at the solution to (3.4).

The right-hand side of equation (4.1) contains the responses of each of the choice variables in problem (3.4) to changes in the components of the parameter vector ρ . We investigate the signs of the components of the right-hand side of (4.1) in order to deduce the respective directions of change in the choice variables when these parameters change.

Weymark (1987), in a model without a public input, obtains his comparative static results by analyzing the first-order conditions for the choice of the consumption levels. His first-order equation associated with c_i contains only c_i and model parameters, which allows him to obtain an explicit solution for c_i . The analogue of this equation in our model, equation (3.6), contains an additional endogenous variable, ψ , the shadow value of the economy's resource constraint. Thus, it is not possible to follow Weymark's strategy to compute the effects of parameter changes on the optimal consumption levels. However, recognizing the dependence of ψ on the model parameters, solving (3.6) yields

$$\tilde{c}_i(\rho) = v'^{-1} \left(\frac{n_i \tilde{\psi}(\rho)}{\beta_i} \right), \quad i = 1, \dots, N. \quad (4.4)$$

Thus, in addition to the comparative static effects described by Weymark, a parameter change induces consumption responses due to a change in the shadow value of the resource constraint. From (2.5) and (3.7), ψ varies inversely with the aggregate wage rate for fixed γ . Hence, the additional responses we analyze can be interpreted as being general equilibrium effects arising from the production side of the economy.

We begin our comparative static analysis by examining how ψ , the shadow value of the resource constraint, varies with the parameters in ρ .

Proposition 3. *At the solution to (3.4), a marginal increase in any of the components of ρ results in an increase in the shadow value of the resource constraint ψ .*

The intuition behind Proposition 3 is straightforward. In light of (3.4), an increase in any β_i increases the marginal value of consumption, and hence the social marginal value of the consumption good, ψ . When resources are optimally allocated, the social marginal value of output equals its social marginal cost. Thus, ψ increases when production

becomes more costly. An increase in either q or γ makes production more costly, either in physical terms or in utility terms. Thus, ψ increases with both q and γ .¹³

The responses of individual consumption levels to changes in the parameters can be deduced directly from (3.6) or (4.4) and (the proof of) Proposition 3. First, an increase in any component of ρ raises the shadow value of resources, thereby raising the social marginal cost of providing c_i . For changes in parameters that do not affect β_i , this results in the marginal cost of c_i exceeding its marginal benefit. It is, therefore, optimal for the government to adjust the value of c_i downward. When β_i increases, both the social marginal benefit and the social marginal cost of c_i increase at the initial optimal value. It turns out that the direct effect on the social marginal benefit via an increase in β_i itself is stronger than the indirect effect that operates through changes in ψ . From Lemma 4, when λ_1 increases, all of the β_j decrease except for β_N , which is unaffected. From the preceding discussion, we see that the effects of these changes in the β_j on c_N all reinforce one another and, hence, c_N increases. Similar reasoning shows that when λ_N increases, c_N decreases. Our comparative static results for consumption are collected in Proposition 4.

Proposition 4. *The consumption level for an individual of type i at the solution to (3.4):*

- (i) *increases when β_i increases marginally;*
- (ii) *decreases when β_j ($j \neq i$), q , or γ increases marginally;*
- (iii) *increases when λ_1 increases marginally for $i = N$;*
- (iv) *decreases when λ_N increases marginally for $i = N$.*

Weymark (1987) describes how consumption levels change in response to increases in reduced-form welfare weights β_i in his Proposition 5. Because the aggregate wage level is fixed in Weymark's model, his results capture only the direct effect of a change in β_i on c_i . As we have already noted, Part (i) of Proposition 4 states that the direct effect of a change in β_i outweighs its indirect effect. Thus, the sign of this comparative static result agrees with Weymark's findings. Part (ii) is at odds, however, with his results. In his model, c_i is unaffected by a change in the reduced-form welfare weights of the other types of individuals.

In contrast to our findings, in his Theorem 3, Weymark (1987) is able to sign the change in the consumption of every type of individual when any of the λ_i increase. Any increase in λ_i for $i \neq 1, N$, increases some reduced-form welfare weights and reduces other reduced-form welfare weights. Inspection of (3.5) reveals that the magnitudes of these changes depend upon numerous factors, including the differences between successive skill levels. Thus, in our model, it is impossible to make general statements about the effect of a change in any of $\lambda_2, \dots, \lambda_{N-1}$ on the shadow value of resources. Hence, the

¹³Only the comparative static result with respect to q depends on the sign of f_{yR} . As is apparent from equation (A.34) in the Appendix, this result is ambiguous when $f_{yR} < 0$.

comparative static effects of these parameters on consumption levels is indeterminate. While an increase in λ_1 or λ_N changes every reduced-form welfare weight in the same direction, except for β_N , which is unaffected, for any $i \neq N$, the effect on c_i of the induced change in β_i is of opposite sign to the effect on c_i of the induced change in the other β_j , rendering the overall effect indeterminate.¹⁴ The former effect vanishes when $i = N$ because there is no binding incentive constraint for this type of individual.

Despite some potential ambiguities concerning consumption responses, it is possible to determine the direction of change in implicit marginal tax rates to changes in every parameter that we consider. These results are collected in Proposition 5.¹⁵

Proposition 5. *The implicit marginal tax rate facing an individual of type i ($i \neq N$) at the solution to (3.4):*

- (i) *decreases when β_i increases marginally;*
- (ii) *does not change when β_j ($j \neq i$), q , or γ increases marginally;*
- (iii) *increases when λ_j increases marginally for $j = 1, \dots, i$;*
- (iv) *decreases when λ_j increases marginally for $j = i + 1, \dots, N$.*

Proposition 5 follows directly from equation (3.11) and Lemma 4. The sharpness of our findings is due to the invariance of the implicit marginal tax rates with respect to the aggregate wage level (and, also, with respect to the social marginal cost of resources). Because a change in the aggregate wage influences all workers proportionally, it has no screening value. Consequently, marginal distortions are not affected by parameters whose only influence is on the aggregate wage.

The two inputs in the production function (2.2) are the aggregate effective labor, y , and the public input, R . Conditional on the shadow value of resources, the optimal values for these inputs are determined by solving equations (3.7) and (3.8) simultaneously. Naturally, changes in the parameters appearing in these two equations affect the choice of inputs. So, too, do changes in the reduced-form welfare weights via their effects on ψ . The comparative static results for the two inputs in the production process are summarized in Proposition 6.

Proposition 6. *Both the amount of aggregate labor in efficiency units and the provision of the publicly provided input at the solution to (3.4):*

- (i) *increase when β_i increases marginally, for any type of individual i ;*
- (ii) *decrease when γ increases marginally;*
- (iii) *decrease when λ_1 increases marginally and increase when λ_N increases marginally.*

¹⁴See (A.41) in the Appendix.

¹⁵Recall that it is always optimal to set the implicit marginal tax rate for individuals of type N to zero.

It follows from Proposition 3 that the shadow value of resources increases when any reduced-form welfare weight increases. Thus, in light of (3.7), the aggregate wage rate decreases when any β_i increases. When this wage falls, the optimal amount of effective labor used increases. Because effective labor and the publicly provided goods are complements in production, it is optimal to use more of the public input as well. An increase in γ also produces an increase in ψ and, with it, a rationale for increasing input usage. However, an increase in γ also has a direct positive effect on the social marginal cost of effective labor. As the social marginal cost of labor increases, it is optimal to reduce the amount of aggregate effective labor and also to use less of the complementary publicly provided input. As Part (ii) of Proposition 6 demonstrates, the direct effect of an increase in γ on input usage is stronger than the general equilibrium effect on input usage operating through ψ .¹⁶

While it is possible to derive expressions for the marginal effect of an increase in the price of the publicly provided input on the optimal usage of the two inputs in the production process, it does not seem possible to sign these effects without further restrictions on the model. The reason for this ambiguity is that a change in q exerts three effects on governmental decisions. First, there is the direct effect on relative input prices, which tends to reduce the provision of the public input, R , and its complement in production, aggregate effective labor, y . There are also two effects that operate through their impact on ψ : the real wage effect described in the previous paragraph and a direct increase in ψ due to the increased cost of the initially optimal provision of R . As we have already seen, the former effect tends to increase the use of the two inputs. However, the second effect tends to reduce resource usage. The real wage effect may be sufficient for the general equilibrium effects of a change in q to dominate the direct effect operating through input prices.

Changes in any welfare weight λ_i influences input usage through its impact on the reduced-form welfare weights β_j , as described above. When λ_1 or λ_N changes, these effects all reinforce one another, decreasing the β_j ($j \neq N$) in the former case, and increasing them in the later. However, when any other welfare weight changes, as noted in our discussion of Proposition 4, the reduced-form welfare weights do not all move in the same direction, rendering the overall impact on input usage indeterminate.

5. Conclusion

Our results extend the literature on the comparative static properties of optimal nonlinear income taxation in several directions. Most obviously, we are able to describe how the optimal provision of a publicly provided input, a novel ingredient in our model, varies with changes in the underlying economy. In addition, we are able to extend the existing comparative static results on consumption allocations to an environment with a nonlinear resource constraint. When the resource constraint is nonlinear, parameter changes have

¹⁶Consistent with the discussion of this paragraph, the proof of Proposition 6 makes it clear that the sign of these comparative static results reverse when $f_{yR} < 0$.

general equilibrium effects that are absent from standard models with linear production functions. These general equilibrium effects are not strong enough to overturn existing results concerning the sign of the effect of a change in reduced-form welfare weights on own consumption. However, they do overturn existing results on the invariance of the consumption allocated to individuals of a given type to changes in the reduced-form welfare weight attached to other types of individuals.

The skill levels of the various types of individuals, s_1, \dots, s_N , enter into the reduced-form optimal nonlinear tax problem via the reduced-form welfare weights alone. Thus, it is possible to use our results to compute the marginal effects of changes in these parameters on the optimal allocations. However, as is the case here for our comparative statics with respect to the welfare weights in the social welfare function, we expect few clear-cut comparative statics for the skill levels. In a model without a public input, a number of comparative static results for these parameters have been obtained by Brett and Weymark (2008a), but their analysis makes extensive use of their assumption that wage rates are exogenous.

The demographic structure of the economy, summarized by n_1, \dots, n_N , affects both the reduced-form welfare weights and the economy's resource constraint. Thus, signing the effects of changes in the distribution of the population across skill types is challenging, but not necessarily impossible. Hamilton and Pestieau (2005) and Boadway and Pestieau (2007) have identified some of the effects of changes in the distribution of types when nonlinear income taxes are chosen optimally, but they assume that preferences are quasilinear in consumption, rather than quasilinear in labor (as we assume here), and they restrict attention to the case in which there are only two skill levels.

It is possible for the resource constraint to be nonlinear and, hence, for the aggregate wage rate to be endogenous, even if there is no public input. A natural extension of our analysis would be to investigate the comparative statics of such a model. If it is further supposed that different kinds of effective labor are not perfect substitutes, as in the two-type model of Stiglitz (1982), then it is necessary to also consider how the relative wages of different types of workers respond to changes in the model parameters. It would also be of interest to extend our model with a public input to allow relative wages to respond to the provision of this input, as in Gaube (2005). More general production technologies might be incorporated to study the comparative statics of public sector demand for labor. Such extensions would pose the technical challenge of analyzing the Weymark (1987) model without imposing a skill-normalization on the welfare weights. The reward for surmounting these challenges might include some results on how the production sector distortions respond to changes in the model parameters.

Appendix

Proof of Lemma 1. Let $a^* = (y_1^*, \dots, y_N^*, c_1^*, \dots, c_N^*, R^*)$ be a candidate solution to the optimal nonlinear income tax problem with the property that, contrary to the statement

of the lemma, there exists a type of individual j such that

$$s_j v(c_j^*) - \gamma y_j^* > s_j v(c_{j-1}^*) - \gamma y_{j-1}^*. \quad (\text{A.1})$$

Then let

$$\bar{y}_i = \begin{cases} y_i^* - \varepsilon_1, & i = 1, \dots, j-1; \\ y_i^* + \varepsilon_2, & i = j, \dots, N, \end{cases} \quad (\text{A.2})$$

for positive ε_1 and ε_2 chosen so that $\bar{y}_i \geq 0$ for all i and so as to preserve the amount of total effective labor supply in the economy; that is, so that

$$\varepsilon_1 \sum_{i=1}^{j-1} n_i = \varepsilon_2 \sum_{i=j}^N n_i. \quad (\text{A.3})$$

Because a^* does not violate any self-selection constraints, single-crossing and (A.1) imply that the allocation $\bar{a} = (\bar{y}_1, \dots, \bar{y}_N, c_1^*, \dots, c_N^*, R^*)$ does not violate any self-selection constraints for ε_1 (and, hence, ε_2) sufficiently small. Thus, the allocation \bar{a} is feasible. Moreover,

$$W(\bar{a}) - W(a^*) = \gamma \left[\varepsilon_1 \sum_{i=1}^{j-1} n_i \lambda_i - \varepsilon_2 \sum_{i=j}^N n_i \lambda_i \right] \quad (\text{A.4})$$

$$\geq \gamma \left[\varepsilon_1 \lambda_{j-1} \sum_{i=1}^{j-1} n_i - \varepsilon_2 \lambda_j \sum_{i=j}^N n_i \right], \quad (\text{A.5})$$

by (2.18). Employing (2.18) again, along with (A.3) and (A.5), implies

$$W(\bar{a}) - W(a^*) \geq \gamma \left(\varepsilon_1 \sum_{i=1}^{j-1} n_i \right) [\lambda_{j-1} - \lambda_j] > 0, \quad (\text{A.6})$$

contradicting the optimality of a^* . \square

Proof of Lemma 2. The equation in (3.3) for type i follows straightforwardly from the equations in (3.1) for $j = 2, \dots, i$. Using (2.15), (3.3) implies

$$y = \sum_{i=1}^N n_i y_i = \sum_{i=1}^N n_i y_1 + \frac{1}{\gamma} \left(\sum_{i=2}^N n_i \sum_{j=2}^i s_j [v(c_j) - v(c_{j-1})] \right). \quad (\text{A.7})$$

Reversing the order of the double summation in (A.7) yields

$$y = y_1 \sum_{i=1}^N n_i + \frac{1}{\gamma} \sum_{j=2}^N \sum_{i=j}^N n_i s_j [v(c_j) - v(c_{j-1})]. \quad (\text{A.8})$$

Equation (3.2) follows directly from (A.8). \square

¹⁷Because the self-selection constraints imply that effective labor is nondecreasing in type, our assumption that $y_1^* > 0$ ensures that such ε_1 and ε_2 exist.

Proof of Lemma 3. Let U^i be the utility [as measured using (2.8)] of an individual of type i associated with an allocation that satisfies (3.1). By (3.1),

$$\begin{aligned} \sum_{i=1}^N n_i U^i &= \sum_{i=1}^N n_i U^1 + \sum_{i=2}^N n_i \sum_{j=1}^{i-1} (s_{j+1} - s_j) v(c_j) \\ &= \sum_{i=1}^N n_i U^1 + \sum_{i=1}^{N-1} \left(\sum_{j=i+1}^N n_j \right) [(s_{i+1} - s_i) v(c_i)]. \end{aligned} \quad (\text{A.9})$$

On the other hand, by (2.8) and (2.15),

$$\sum_{i=1}^N n_i U^i = \sum_{i=1}^N n_i s_i v(c_i) - \gamma \sum_{i=1}^N n_i y_i = \sum_{i=1}^N n_i s_i v(c_i) - \gamma y. \quad (\text{A.10})$$

Combining (A.9) and (A.10) yields

$$U^1 = \frac{1}{\sum_{i=1}^N n_i} \left(\sum_{i=1}^N n_i s_i v(c_i) - \gamma y - \sum_{i=1}^{N-1} \left(\sum_{j=i+1}^N n_j \right) [(s_{i+1} - s_i) v(c_i)] \right). \quad (\text{A.11})$$

Now, for any allocation that satisfies (3.1),

$$\begin{aligned} W &= \left(\sum_{i=1}^N n_i \lambda_i \right) U^1 + \sum_{i=2}^N n_i \lambda_i \left[\sum_{j=1}^{i-1} (s_{j+1} - s_j) v(c_j) \right] \\ &= \left(\sum_{i=1}^N n_i \lambda_i \right) U^1 + \sum_{i=1}^{N-1} \left[\left(\sum_{j=i+1}^N n_j \lambda_j \right) (s_{i+1} - s_i) v(c_i) \right]. \end{aligned} \quad (\text{A.12})$$

Substituting (A.11) into (A.12) yields

$$\begin{aligned} W &= \frac{\sum_{i=1}^N n_i \lambda_i}{\sum_{i=1}^N n_i} \left[\sum_{i=1}^N n_i s_i v(c_i) - \gamma y - \sum_{i=1}^{N-1} \left(\sum_{j=i+1}^N n_j \right) (s_{i+1} - s_i) v(c_i) \right] \\ &\quad + \sum_{i=1}^{N-1} \left[\left(\sum_{j=i+1}^N n_j \lambda_j \right) (s_{i+1} - s_i) v(c_i) \right]. \end{aligned} \quad (\text{A.13})$$

The normalization rule (2.19) allows the simplification of (A.13) to

$$\begin{aligned} W &= \sum_{i=1}^N n_i s_i v(c_i) - \sum_{i=1}^{N-1} \left(\sum_{j=i+1}^N n_j \right) (s_{i+1} - s_i) v(c_i) \\ &\quad + \sum_{i=1}^{N-1} \left[\left(\sum_{k=1}^N n_k - \sum_{k=1}^i n_i \lambda_k \right) (s_{i+1} - s_i) v(c_i) \right] - \gamma y. \end{aligned} \quad (\text{A.14})$$

Collecting terms in (A.14) yields

$$W = \sum_{i=1}^N n_i s_i v(c_i) + \sum_{i=1}^{N-1} \left[\left(\sum_{k=1}^i (n_k - n_i \lambda_k) \right) (s_{i+1} - s_i) v(c_i) \right] - \gamma y. \quad (\text{A.15})$$

The normalization rule (2.19) implies that the upper limit of the outer summation in the second term in (A.15) can be extended to N because the term in the inner summation is zero when $i = N$. Thus, for any constant s_{N+1} ,

$$W = \sum_i^N \left[n_i s_i + \left(\sum_{k=1}^i (n_k - n_k \lambda_k) \right) (s_{i+1} - s_i) \right] v(c_i) - \gamma y, \quad (\text{A.16})$$

which is exactly the objective function in (3.4). The constraints in (3.4) are the consumption monotonicity constraint (2.11) and the production-feasibility constraint (2.14), neither of which has been substituted into the objective function during the preceding argument. The strict concavity of f and v implies that any solution to (3.4) is unique.

To complete the proof that the values of c_1, \dots, c_N , y , and R that solve (3.4) are optimal for the Optimal Nonlinear Tax Problem, it remains to show that the self-selection constraints (2.10) are satisfied. By Lemma 2, we know that (3.3) holds at a solution to the optimal nonlinear income tax problem. It follows from (2.11) and (3.3) that $y_1 \leq \dots \leq y_N$. Thus, the attribute ordering and ordering of marginal rates of substitution conditions of Matthews and Moore (1987) are satisfied. Therefore, (2.10) is also satisfied.

Let $(c_1^*, \dots, c_N^*, y^*, R^*)$ solve (3.4). Contrary to the statement of the lemma, suppose that

$$\sum_{i=1}^N n_i c_i^* + q R^* < f(R^*, y^*). \quad (\text{A.17})$$

The assumption that $f(R, 0) = 0$ implies that $y^* > 0$. Thus, it is possible to marginally decrease y^* by some amount $\delta > 0$ without violating the production-feasibility constraint (2.14). But such a change increases the value of the objective function in (3.4), contradicting the optimality of $(c_1^*, \dots, c_N^*, y^*, R^*)$. \square

Proof of Proposition 1. Part (i) follows directly from (3.8).

Solving (3.6) for $v'(c_i)$ and substituting the result into (2.13) yields

$$\text{IMTR}_i = 1 - \frac{\gamma}{w s_i \frac{\psi n_i}{\beta_i}} = 1 - \frac{\gamma}{w s_i \frac{\gamma n_i}{w \beta_i}} = 1 - \frac{\beta_i}{n_i s_i}, \quad i = 1, \dots, N, \quad (\text{A.18})$$

where the second equality follows from (2.5) and (3.7). Part (ii) follows directly from (A.18) because $\beta_N = n_N s_N$.

Substituting (3.5) into (A.18) and simplifying yields the final equation in (3.11). It remains to show that the inequality in (3.11) is satisfied. To that end, suppose, by way of contradiction, that the inequality is not satisfied for some type i . Then

$$\sum_{k=1}^i n_k \lambda_k \leq \sum_{k=1}^i n_k. \quad (\text{A.19})$$

Now, by (2.18)

$$\lambda_i \sum_{k=1}^i n_k < \sum_{k=1}^i n_k \lambda_k. \quad (\text{A.20})$$

Hence, by (A.19) and (A.20),

$$\lambda_i \sum_{k=1}^i n_k < \sum_{k=1}^i n_k, \quad (\text{A.21})$$

which implies that $\lambda_i < 1$.

Next, note that (2.19) and (A.19) imply

$$\sum_{k=i+1}^N n_k \lambda_k \geq \sum_{k=i+1}^N n_k. \quad (\text{A.22})$$

Now, by (2.18)

$$\lambda_{i+1} \sum_{k=i+1}^N n_k > \sum_{k=i+1}^N n_k \lambda_k. \quad (\text{A.23})$$

Hence, by (A.22) and (A.23),

$$\lambda_{i+1} \sum_{k=i+1}^N n_k > \sum_{k=i+1}^N n_k, \quad (\text{A.24})$$

which implies that $\lambda_{i+1} > 1$. Therefore, (A.21) and (A.24) imply $\lambda_{i+1} > \lambda_i$, which violates (2.18). This contradiction proves the inequality in (3.11). \square

Proof of Lemma 4. The proportional reduction in λ_k for every $k < i$ increases $\sum_{k=1}^j (n_k - n_k \lambda_k)$ for every $j < i$. Therefore, by (3.5), β_j increases for every $j < i$. The increase in λ_i , however, induces an decrease in $\sum_{k=1}^j (n_k - n_k \lambda_k)$ for every $k = i, \dots, N - 1$. Consequently, β_j decreases for every $j = i, \dots, N - 1$. Finally, $\beta_N = s_N w_N$, which does not depend on λ_i . \square

Proof of Proposition 2. Totally differentiating the optimality conditions (3.6)–(3.9) with respect to the endogenous variables and the components of ρ (and suppressing the dependence of $A(\rho)$ and $B(\rho)$ on ρ) yields

$$A \begin{bmatrix} dc_1 \\ \vdots \\ dc_N \\ dy \\ dR \\ d\psi \end{bmatrix} = B \begin{bmatrix} d\beta_1 \\ \vdots \\ d\beta_N \\ dq \\ d\gamma \end{bmatrix}, \quad (\text{A.25})$$

where use has been made of (3.8). Proposition 2 follows from the Implicit Function Theorem if the matrix A is invertible. In order to establish invertibility of A , rewrite A in the form

$$A = \begin{bmatrix} H & Z \\ Z^T & 0 \end{bmatrix}, \quad (\text{A.26})$$

where H is the $(N + 2) \times (N + 2)$ upper-left block of A ,

$$Z^T = [-n_1, \dots, -n_N, f_y, 0], \quad (\text{A.27})$$

and the zero in (A.26) is scalar. Because v and f are both strictly concave, H is negative definite. Hence, H is invertible. It is straightforward to check that

$$A^{-1} = \begin{bmatrix} H^{-1} - \theta H^{-1} Z Z^T H^{-1} & \theta H^{-1} Z \\ \theta Z^T H^{-1} & -\theta \end{bmatrix}, \quad (\text{A.28})$$

where

$$\theta = \frac{1}{Z^T H^{-1} Z} < 0. \quad (\text{A.29})$$

The inequality in (A.29) holds because H^{-1} is negative definite. \square

Proof of Proposition 3. The partial derivatives of $\tilde{\psi}(\rho)$ are found in the bottom row of (4.1). It follows from (A.28) that

$$\left[\frac{\partial \tilde{\psi}}{\partial \beta_1} \quad \dots \quad \frac{\partial \tilde{\psi}}{\partial \beta_N} \quad \frac{\partial \tilde{\psi}}{\partial q} \quad \frac{\partial \tilde{\psi}}{\partial \gamma} \right] = \left[\theta Z^T H^{-1} \quad -\theta \right] B. \quad (\text{A.30})$$

The matrix H is block diagonal. It contains an upper-left block of size $N \times N$ which is, itself, diagonal, along with a 2×2 lower-right block. Thus, it is clear that

$$H^{-1} = \begin{bmatrix} \frac{1}{\beta_1 v''(c_1)} & 0 & \dots & \dots & 0 & 0 & 0 \\ 0 & \frac{1}{\beta_2 v''(c_2)} & 0 & \dots & 0 & 0 & 0 \\ \vdots & 0 & \ddots & & \vdots & \vdots & \vdots \\ \vdots & \vdots & & \ddots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & \frac{1}{\beta_N v''(c_N)} & 0 & 0 \\ 0 & 0 & \dots & \dots & 0 & \frac{f_{RR}}{\Delta} & -\frac{\psi f_{yR}}{\Delta} \\ 0 & 0 & \dots & \dots & 0 & -\frac{f_{yR}}{\Delta} & \frac{\psi f_{yy}}{\Delta} \end{bmatrix}, \quad (\text{A.31})$$

where

$$\Delta = \psi \left[f_{RR} f_{yy} - (f_{yR})^2 \right] > 0. \quad (\text{A.32})$$

The inequality in (A.32) holds because f is strictly concave and, by (3.7), $\psi > 0$. Substituting (4.3), (A.27), and (A.31) into the right-hand side of (A.30) and performing the resulting matrix multiplications yields

$$\frac{\partial \tilde{\psi}}{\partial \beta_i} = \frac{\theta n_i v'(c_i)}{\beta_i v''(c_i)}, \quad i = 1, \dots, N; \quad (\text{A.33})$$

$$\frac{\partial \tilde{\psi}}{\partial q} = \frac{-\theta \psi f_y f_{yR}}{\Delta} - \theta R; \quad (\text{A.34})$$

$$\frac{\partial \tilde{\psi}}{\partial \gamma} = \frac{\theta f_y f_{RR}}{\Delta}. \quad (\text{A.35})$$

The right-hand side of (A.33) is positive because $\theta < 0$, $v'(c_i) > 0$, and $v''(c_i) < 0$. The first term on the right-hand side of (A.34) is positive because $f_y(R, y) > 0$, $f_{yR}(R, y) > 0$, $\Delta > 0$, $\psi > 0$, and $\theta < 0$. Because $\theta < 0$, it then follows that the right-hand side of (A.34) is positive. Finally, the right-hand side of (A.35) is positive because $f_y(R, y) > 0$, $f_{RR}(R, y) < 0$, $\Delta > 0$, and $\theta < 0$. \square

Proof of Proposition 4. Let μ denote the argument of the function v'^{-1} . Differentiating (4.4) yields

$$\frac{\partial \tilde{c}_i}{\partial \zeta} = \frac{\partial v'^{-1}}{\partial \mu} \frac{n_i}{\beta_i} \frac{\partial \tilde{\psi}}{\partial \zeta}, \quad \zeta = q, \gamma, \beta_j \ (j \neq i). \quad (\text{A.36})$$

By the concavity of v , v' is decreasing. Hence, v'^{-1} is also decreasing. Thus, by Proposition 3, the right-hand side of (A.36) is negative. Part (ii) of Proposition 4 follows from these observations.

Differentiating (4.4) with respect to β_i yields

$$\frac{\partial \tilde{c}_i}{\partial \beta_i} = \frac{\partial v'^{-1}}{\partial \mu} \left[\frac{n_i}{\beta_i} \frac{\partial \tilde{\psi}}{\partial \beta_i} - \frac{n_i}{\beta_i^2} \tilde{\psi} \right]. \quad (\text{A.37})$$

Using (3.6) and (A.33) to substitute for $\tilde{\psi}$ and its partial derivative, respectively, in (A.37) yields

$$\frac{\partial \tilde{c}_i}{\partial \beta_i} = \frac{\partial v'^{-1}}{\partial \mu} \frac{v'(c_i)}{\beta_i} \left[\frac{\theta n_i^2}{\beta_i v''(c_i)} - 1 \right]. \quad (\text{A.38})$$

Now, using (A.27), (A.29), and (A.31),

$$\frac{1}{\theta} = \sum_{j=1}^N \frac{n_j^2}{\beta_j v''(c_j)} + \frac{f_y^2 f_{RR}}{\Delta} < \frac{n_i^2}{\beta_i v''(c_i)}. \quad (\text{A.39})$$

The inequality in (A.39) holds because $\Delta > 0$ and the strict concavity of v and f imply that every one of the $N + 1$ terms in the middle expression in (A.39) is negative. Because $\theta < 0$, (A.39) implies

$$1 > \frac{\theta n_i^2}{\beta_i v''(c_i)}. \quad (\text{A.40})$$

Thus, the term in square brackets on the right-hand side of (A.38) is negative. Because v'^{-1} is decreasing, the entire right-hand side of (A.38) is positive. Part (i) of Proposition 4 then follows.

From (A.36) and (A.37), it follows that

$$\frac{\partial \tilde{c}_i}{\partial \lambda_j} = \frac{\partial v'^{-1}}{\partial \mu} \frac{n_i}{\beta_i} \sum_{k \neq i} \frac{\partial \tilde{\psi}}{\partial \beta_k} \frac{\partial \beta_k}{\partial \lambda_j} + \frac{\partial v'^{-1}}{\partial \mu} \left[\frac{n_i}{\beta_i} \frac{\partial \tilde{\psi}}{\partial \beta_i} - \frac{n_i}{\beta_i^2} \tilde{\psi} \right] \frac{\partial \beta_i}{\partial \lambda_j}, \quad i, j = 1, \dots, N. \quad (\text{A.41})$$

By Lemma 4, an increase in λ_1 (resp. λ_N) induces a decrease (resp. increase) in each of $\beta_1, \dots, \beta_{N-1}$, but no change in β_N . In both of these cases, when $i = N$, the second term

on the right-hand side of (A.41) vanishes. Hence, if λ_1 (resp. λ_N) increases, by Part (ii), the first term is positive (resp. negative), which establishes Part (iii) (resp. Part (iv)) of Proposition 4. \square

Proof of Proposition 5. Parts (i) and (ii) of Proposition 5 follow directly from (3.11). By Part (ii) of Lemma 4, an increase in λ_j induces a decrease in β_i for $i \geq j$ ($i \neq N$). Part (iii) of Proposition 5 then follows from Part (i). Similarly, by Part (i) of Lemma 4, an increase in λ_j induces an increase in β_i for $i < j$. As above, Part (iv) of Proposition 5 then follows from Part (i). \square

Proof of Proposition 6. We present heuristic calculations that are justified by the Implicit Function Theorem. The same results can be obtained by carrying out the matrix calculations in (4.1).

In light of (4.2), rearranging rows $N + 1$ and $N + 2$ of (A.25) yields

$$\begin{bmatrix} \psi f_{yy} & \psi f_{yR} \\ f_{yR} & f_{RR} \end{bmatrix} \begin{bmatrix} dy \\ dR \end{bmatrix} = \begin{bmatrix} d\gamma - f_y d\psi \\ dq \end{bmatrix}. \quad (\text{A.42})$$

The solution to the matrix equation (A.42) is

$$\begin{bmatrix} dy \\ dR \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} f_{RR} & -\psi f_{yR} \\ -f_{yR} & \psi f_{yy} \end{bmatrix} \begin{bmatrix} d\gamma - f_y d\psi \\ dq \end{bmatrix}. \quad (\text{A.43})$$

It follows from (A.43) that

$$\frac{\partial y}{\partial \beta_i} = -\frac{f_{RR} f_y}{\Delta} \frac{\partial \psi}{\partial \beta_i} > 0, \quad i = 1, \dots, N. \quad (\text{A.44})$$

Because $\Delta > 0$, the inequality in (A.44) follows from the strict concavity of f and Proposition 3. Also from (A.43),

$$\frac{\partial R}{\partial \beta_i} = \frac{f_{yR} f_y}{\Delta} \frac{\partial \psi}{\partial \beta_i} > 0, \quad i = 1, \dots, N. \quad (\text{A.45})$$

The inequality in (A.45) follows from the positivity of Δ , the strict concavity of f , the complementarity of y and R in production, and Proposition 3. Equations (A.44) and (A.45) establish Part (i) of Proposition 6.

Employing (A.43) once more yields

$$\frac{\partial y}{\partial \gamma} = \frac{f_{RR}}{\Delta} - \frac{f_{RR} f_y}{\Delta} \frac{\partial \psi}{\partial \gamma}. \quad (\text{A.46})$$

Substituting (A.35) into (A.46) and rearranging gives

$$\frac{\partial y}{\partial \gamma} = \frac{f_{RR}}{\Delta} \left[1 - \frac{\theta f_y^2 f_{RR}}{\Delta} \right]. \quad (\text{A.47})$$

Now, by an argument analogous to the one used to justify (A.39),

$$\frac{1}{\theta} < \frac{f_y^2 f_{RR}}{\Delta}, \quad (\text{A.48})$$

and, because $\theta < 0$,

$$1 > \frac{\theta f_y^2 f_{RR}}{\Delta}. \quad (\text{A.49})$$

Hence, the term in square brackets on the right-hand side of (A.47) is positive. Thus, $\Delta > 0$ and the strict concavity of f imply that $\frac{\partial y}{\partial \gamma} < 0$.

Using (A.43) yet again yields

$$\frac{\partial R}{\partial \gamma} = -\frac{f_{yR}}{\Delta} + \frac{f_{yR} f_y}{\Delta} \frac{\partial \psi}{\partial \gamma}. \quad (\text{A.50})$$

Substituting (A.35) into (A.50) and rearranging gives

$$\frac{\partial R}{\partial \gamma} = -\frac{f_{yR}}{\Delta} \left[1 - \frac{\theta f_y^2 f_{RR}}{\Delta} \right]. \quad (\text{A.51})$$

We have already established that the term in square brackets on the right-hand side of (A.51) is positive. Because $f_{yR} > 0$ and $\Delta > 0$, the entire right-hand side is negative, thereby establishing Part (ii) of Proposition 6.

Part (iii) of the Proposition 6 follows directly from Part (i) and Lemma 4. \square

References

- Blackorby, C., Brett, C., 2004. Production efficiency and the direct-indirect tax mix. *Journal of Public Economic Theory* 6, 165–180.
- Boadway, R., Keen, M., 1993. Public goods, self-selection and optimal income taxation. *International Economic Review* 34, 463–477.
- Boadway, R., Marchand, M., 1995. The use of public expenditures for redistributive purposes. *Oxford Economic Papers* 47, 45–59.
- Boadway, R., Pestieau, P., 2007. Tagging and redistributive taxation. *Annales d'Économie et de Statistique* 83–84, 123–147.
- Brett, C., 2009. The effects of population aging on optimal redistributive taxes in an overlapping generations model, Downloadable from the Munich Personal RePEc Archive, <http://mpra.ub.uni-muenchen.de/8585/>.
- Brett, C., Weymark, J. A., 2008a. The impact of changing skill levels on optimal nonlinear income taxes. *Journal of Public Economics* 92, 1765–1771.
- Brett, C., Weymark, J. A., 2008b. Public good provision and the comparative statics of optimal nonlinear income taxation. *International Economic Review* 49, 255–290.

- Christiansen, V., 1981. Evaluation of public projects under optimum taxation. *Review of Economic Studies* 48, 447–457.
- Feehan, J. P., Matsumoto, M., 2000. Productivity-enhancing public investment and benefit taxation: The case of factor-augmenting public inputs. *Canadian Journal of Economics* 33, 114–121.
- Gaube, T., 2005. Public investment and income taxation: Redistribution vs. productive performance. *Journal of Economics* 86, 1–18.
- Hamilton, J., Pestieau, P., 2005. Optimal income taxation and the ability distribution: Implications for migration equilibria. *International Tax and Public Finance*, 29–45.
- Kaizuka, K., 1965. Public goods and decentralization of production. *Review of Economics and Statistics* 47, 118–120.
- Keen, M., Marchand, M., 1997. Fiscal competition and the pattern of public spending. *Journal of Public Economics* 66, 33–53.
- Lollivier, S., Rochet, J.-C., 1983. Bunching and second-order conditions: A note on optimal tax theory. *Journal of Economic Theory* 31, 392–400.
- Manning, R., Markusen, J. R., McMillan, J., 1985. Paying for public inputs. *American Economic Review* 75, 235–238.
- Matthews, S., Moore, J., 1987. Monopoly provision of quality and warranties: An exploration in the theory of multi-dimensional 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.
- Naito, H., 1999. Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics* 71, 165–188.
- Samuelson, P. A., 1954. The pure theory of public expenditure. *Review of Economics and Statistics* 36, 387–389.
- Sandmo, A., 1972. Optimality rules for the provision of collective factors of production. *Journal of Public Economics* 1, 149–157.
- Simula, L., 2007. Optimality conditions and comparative static properties of non-linear income taxes revisited. Working Paper No. 2007-15, Paris School of Economics.
- Stiglitz, J. E., 1982. Self-selection and Pareto efficient taxation. *Journal of Public Economics* 17, 213–240.
- 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.