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)

April 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.

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 an input into the production process. The presence of the public input implies, in contrast to most existing studies of the comparative statics of optimal nonlinear income tax problems, that wage rates are endogenous. The parameters for which comparative statics are obtained are the weights in a weighted utilitarian social welfare function, a taste parameter that measures the onerousness of working, and a technological parameter that determines the price of the publicly provided input.

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) describes the circumstances under which public provision of a private good is merited even in the presence of optimal nonlinear income taxes. A central feature in 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 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 able workers.[2]

[1] When 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.

[2] Similar justifications for production inefficiency in models of optimal nonlinear income taxation are

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.

Our comparative static analysis focuses on the effects of changes in the following variables: the weights in the social welfare function, a measure of the disutility of working, and the marginal cost of the publicly provided good. 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), Simula (2007), and Brett and Weymark (2008a,b).[3] In spite of the existence of the extra effects, we are able to obtain results on the sign of the comparative static responses to parameter changes for every individual's consumption allocation and for the sign of the responses of aggregate effective labor and publicly provided input usage to the welfare weights and the disutility parameter.

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, 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, it is possible to formulate a first-stage problem describing the choice of consumption and input allocations as arising out of a maximizing problem constrained only by the economy-wide resource constraint. We employ techniques borrowed from the theory of consumer demand to derive comparative static results for our reduced

provided, albeit in other contexts, by Naito (1999) and Blackorby and Brett (2004).

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

form.

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 $s_1 < s_2 < \cdots < s_N$. An individual's skill level measures the rate at which his labor time, l_i, is transformed into his effective labor supply, y_i. Specifically, $y_i = s_i l_i$.

The producer sector is described by a production function, f, that transforms a publicly provided input, R, and effective labor, y, into the output $f(R, y)$, where f is continuous, twice continuously differentiable, and strictly concave with $f(R, 0) = f(0, y) = 0$ for all nonnegative R and y. We also 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, c, or transformed into the public input according to a constant marginal rate of technical substitution process in which the opportunity cost of one unit of the public input is q units of consumption. Thus, the aggregate technology satisfies

$$c + qR \leq f(R, y). \tag{2.1}$$

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

$$w = \frac{\partial f(R, y)}{\partial y}. \tag{2.2}$$

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

$$z_i = w y_i = w s_i l_i. \tag{2.3}$$

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

$$V(c, l) = v(c) - \gamma l, \tag{2.4}$$

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 $v_c(c) \to 0$ as $c \to \infty$. The limiting

[4]This assumption is satisfied when the technology exhibits constant returns to scale.

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 preferences by the type-specific monotonic transformation of (2.4)

$$U^i(c, y) = s_i v(c) - \gamma y. \tag{2.5}$$

Equation (2.5) describes preferences over consumption and effective labor supply. The marginal rate of substitution between effective labor and consumption for an individual of type i is

$$\mathrm{MRS}^i(c_i, y_i) = \frac{\gamma}{s_i v'(c_i)}, \tag{2.6}$$

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 (2.5) 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

$$s_i v(c_i) - \gamma y_i \geq s_i v(c_j) - \gamma y_j, \quad \forall i, j = 1, \ldots, N. \tag{2.7}$$

It is well known that the self-selection conditions imply that the consumption allocations must satisfy the conditions

$$c_1 \leq c_2 \leq \cdots \leq c_n.^5 \tag{2.8}$$

The tax system consistent with an allocation satisfying (2.7) is typically nondifferentiable. Thus, marginal tax rates are only implicitly defined by the difference between producer and consumer prices at the an individual's consumption bundle. The implicit marginal tax rate (IMTR) for labor income is given by

$$\mathrm{IMTR}_i = 1 - \frac{\gamma}{w s_i v'(c_i)}. \tag{2.9}$$

An allocation is a vector $a = (y_1, \ldots, y_N; c_1, \ldots c_N; R)$ 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

$$\sum_{i=1}^{N} n_i c_i + qR \leq f(R, y), \tag{2.10}$$

[5] Effective labor supplies satisfy analogous monotonicity conditions, but these follow necessarily from (3.3) below.

where

$$y = \sum_{i=1}^{N} n_i y_i, \tag{2.11}$$

is the aggregate supply of effective labor.

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

$$W(a) = \sum_{i=1}^{N} \mu_i n_i V(y_i, c_i) = \sum_{i=1}^{N} \lambda_i n_i [s_i v(c_i) - \gamma y_i] \tag{2.12}$$

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

$$\lambda_i = \mu_i / s_i, \quad i = 1, \dots, N, \tag{2.13}$$

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

$$0 < \lambda_N < \cdots < \lambda_1. \tag{2.14}$$

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 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,

$$\sum_{i=1}^{N} n_i \lambda_i = \sum_{i=1}^{N} n_i. \tag{2.15}$$

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 (2.12) subject to the self-selection constraints (2.7) and the materials balance constraint (2.10).

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

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

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

In the language of screening models, Lemma 1 states that optimality requires that all downward adjacent self-selection constraints bind. 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.

For a given consumption allocation, the binding self-selection constraints (3.1) form a system of $N-1$ linear equations in the N variables $y_1, \ldots, y_N$. Given an aggregate supply of effective labor, y, (2.11) provides an Nth linear equation in the y_is. The solution to the resulting system of equations is given in Lemma 2.

Lemma 2. *For a given* $(c_1, \ldots, c_n; y)$*, the system of equations* (2.11) *and* (3.1) *have a unique solution. Moreover, this solution can be written in the recursive form:*

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

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

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 (2.12). The resulting reduced-from objective function depends on consumption levels and aggregate effective labor. Maximizing this objective function subject to the production-feasibility constraint (2.10) yields optimal values $(c_1^*, \ldots, c_N^*; y^*, R^*)$. 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 associated with the Optimal Nonlinear Tax Problem can be found by solving*

$$\max_{c_1, \ldots, c_N; y, R} \sum_{i=1}^{N} \beta_i v(c_i) - \gamma y \quad \textit{subject to } (2.8) \textit{ and } (2.10), \tag{3.4}$$

where

$$\beta_i = n_i s_i + \left(\sum_{k=1}^{i} (n_k - n_k \lambda_k) \right) (s_{i+1} - s_i), \quad i = 1, \ldots, N \tag{3.5}$$

and s_{N+1} *is an arbitrary number.*[6]

Henceforth, we assume that the monotonicity constraints (2.8) *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.

[6]Note that the normalization (2.15) implies $\beta_N = n_N s_N$.

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

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), Simula (2007), and Brett and Weymark (2008b), even when it is assumed that the monotonicity constraints are not binding, (3.4) is not a fully unconstrained optimization problem. The nonlinearity of the production-feasibility constraint makes it inconvenient to substitute this constraint 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 (2.10), allows the first-order conditions for a solution to (3.4) to be written as

$$c_i\colon \beta_i v'(c_i) - \psi n_i = 0, \quad i = 1, \ldots, n; \tag{3.6}$$

$$y\colon -\gamma + \psi f_y = 0; \tag{3.7}$$

$$R\colon f_R - q = 0; \tag{3.8}$$

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

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).

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,*

$$\mathrm{IMTR}_N = 1 - \frac{\beta_N}{n_N s_N} = 0. \tag{3.10}$$

(iii) *The implicit marginal tax rate on the labor income of individuals of types $1, \ldots, N-1$ is positive; specifically,*

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

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 convey the standard pattern of labor market distortions arising in redistributive optimal nonlinear tax schemes: 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, aggregate effective labor supply, and provision of the publicly provided input respond to changes in some of the parameters of the economy. These are the endogneous variables in the first-stage optimization problem (3.4). As 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. In our model, the exogenous parameters are the technology parameter q, the taste parameter γ, the skill parameters $s_1, \ldots, s_N$, the welfare weights, $\lambda_1, \ldots, \lambda_n$, and the demographic parameters $n_1, \ldots, n_N$. The skills and welfare weights enter the problem (3.4) only through their influence on the reduced-form welfare weights $\beta_1, \ldots, \beta_N$. Thus, we will investigate how the optimal allocation depends on the reduced-form welfare weights. The parameter vector we vary is $\rho = (\beta_1, \ldots, \beta_N, q, \gamma)$.

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 procedures is given in Proposition 2.

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

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

where

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

and

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

and where all expressions on the right-hand sides of (4.2) *and* (4.3) *are 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 the problem (3.4) to changes in the components of the parameter vector ρ. In the remainder of this section, 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 the parameters change.

Weymark (1987) bases his comparative static analysis in a model without a public input on an analysis of 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, and solving (3.6) yields,

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

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 (2.2) and (3.7), ψ varies inversely with the aggregate wage rate for fixed γ. Hence, the additional responses we analyze can be interpreted as 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 model parameters.

Proposition 3. *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.

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. Our results on consumption responses 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.*

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.

Conditional on the shadow value of resources, the optimal combination of aggregate effective labor and the public input is 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.

Proposition 5. *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.*

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. As the wage rate falls, the optimal amount of labor used increases. Because effective labor and the publicly provided goods are complements in production, it is optimal to use more R 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. Part (ii) of Proposition 5 states that the direct effect of an increase in γ on input usage is stronger than the general equilibrium effect on input usage operating through ψ.

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 R and its complement in production, y. There are also two effects 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. This second source of increase in the shadow value of the resource constraint may be sufficient to render the general equilibrium effects of a change in q stronger than the direct effect operating through input prices.

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 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 certain type to changes in the reduced-form welfare weight attached to other types of individuals.

It is possible to use our results to carry out other comparative static exercises. The underlying welfare weights, $\lambda_1, \ldots, \lambda_N$, and the skill levels of the various types of individuals, $s_1, \ldots, s_N$, 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. In a model without a public input, comparative static results for these parameters have been obtained by Weymark (1987) and by Brett and Weymark (2008a), respectively. The demographic structure of the economy, summarized by $n_1, \ldots, n_N$, enters both the reduced-form welfare weights and the economy's resource constraint. Thus, computing the effects of changes in the distribution of the population across skill types is more challenging, but not impossible. Hamilton and Pestieau (2005) and Boadway and Pestieau (2007) have analyzed 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.

A possible extension of our analysis would be to allow the relative wages of different types of workers to respond to the provision of the publicly provided input, as in Gaube (2005). Such an extension 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 identified by Gaube respond to changes in model parameters.

Appendix

Proof of Lemma 1. Let $a^* = (y_1^*, \ldots, y_N^*, c_1^*, \ldots, c_N^*, R^*)$ 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

$$s_j v(c_j) - \gamma y_j > s_j v(c_{j-1}) - \gamma y_{j-1}. \tag{A.1}$$

Then let

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

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

$$\varepsilon_1 \sum_{i=1}^{j-1} n_i = \varepsilon_2 \sum_{i=j}^{N} n_i.^8 \tag{A.3}$$

Because a^* does not violate any self-selection constraints, single-crossing and (A.1) implies that the allocation $\bar{a} = (\bar{y}_1, \ldots, \bar{y}_N, c_1^*, \ldots, c_N^*, R^*)$ does not violate any self-selection

[8] Our assumption that $y_i^* > 0$ for all i ensures that such ε_1 and ε_2 exist.

constraints for ε_1 (hence, ε_2) sufficiently small. Thus, the allocation $\bar{a}$ is feasible. Moreover,

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

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

by (2.14). Employing (2.14) again, along with (A.3) and(A.5), implies

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

contradicting the optimality of a^*. □

Proof of Lemma 2. The equation in (3.3) for type i follows straightforwardly from the equations in (3.1) for $j = 2, \ldots i$. Using (2.11), (3.3) implies

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

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

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

Equation (3.2) follows directly from (A.8). □

Proof of Lemma 3. Let V^i be the utility [as measured using (2.5)] of an individual of type i associated with an allocation that satisfies (3.1). By (3.1)

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

On the other hand, by (2.5) and (2.11),

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

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

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

Now, for any allocation that satisfies (3.1),

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

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

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

The normalization rule (2.15) allows the simplification of (A.13) to

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

Collecting terms in (A.14) yields

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

Simplifying the term in braces in (A.15) gives

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

But the normalization rule (2.15) implies that the upper limit of first sum in the term in braces in (A.16) can be extended to N because the Nth term is zero. Thus, for any constant s_{N+1},

$$W = \left[\sum_i^N n_i s_i + \left\{ \sum_{i=1}^N \sum_{k=1}^i (n_k - n_k \lambda_k) \right\} (s_{i+1} - s_i) \right] v(c_i) - \gamma y, \tag{A.17}$$

which is exactly the objective function in (3.4). The constraint in (3.4) is the production-feasibility constraint, which has not been substituted into the objective function during the argument in this proof.

It remains to show that (3.1) and (2.8) imply (2.7). This implication follows from the analysis in Matthews and Moore (1987) because (2.8) and (3.3) imply $y_1 \leq \cdots \leq y_n$. Thus, the Matthews–Moore attribute ordering and ordering of marginal rates of substitution conditions are satisfied. Therefore, (2.7) is also satisfied. □

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

Solving (3.6) for $v'(c_i)$ and substituting the result into (2.9) yields

$$\mathrm{IMTR}_i = 1 - \frac{\gamma}{ws_i \frac{\psi n_i}{\beta_i}} = 1 - \frac{\gamma}{ws_i \frac{\gamma n_i}{w\beta_i}} = 1 - \frac{\beta_i}{n_i s_i}, \quad i = 1, \ldots, N, \tag{A.18}$$

where the second equality follows from (2.2) 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. Then

$$\sum_{k=1}^{i} n_k \lambda_k \leq \sum_{k=1}^{i} n_k. \tag{A.19}$$

Now, by (2.14)

$$\lambda_i \sum_{k=1}^{i} n_k < \sum_{k=1}^{i} n_k \lambda_k. \tag{A.20}$$

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

$$\lambda_k \sum_{k=1}^{i} n_k < \sum_{k=1}^{i} n_k. \tag{A.21}$$

which implies that $\lambda_i < 1$.

Next, note that (2.15) and (A.19) imply

$$\sum_{k=i+1}^{N} n_k \lambda_k \geq \sum_{k=i+1}^{N} n_k. \tag{A.22}$$

Now, by (2.14)

$$\lambda_{i+1} \sum_{k=i+1}^{N} n_k > \sum_{k=i+1}^{N} n_k \lambda_k. \tag{A.23}$$

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

$$\lambda_{i+1} \sum_{k=i+1}^{N} n_k > \sum_{k=i+1}^{N} n_k, \tag{A.24}$$

which implies that $\lambda_{i+1} > 1$. Therefore, (A.21) and (A.24) imply $\lambda_{i+1} > \lambda_i$, which violates (2.14). This contradiction proves the inequality in (3.11). □

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

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

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

$$A = \begin{bmatrix} H & Z \\ Z^T & 0 \end{bmatrix}, \tag{A.26}$$

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

$$Z^T = [-n_1, \ldots, -n_N, f_y, 0], \tag{A.27}$$

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

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

where

$$\theta = \frac{1}{Z^T H^{-1} Z} < 0. \tag{A.29}$$

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

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

$$\begin{bmatrix} \frac{\partial \tilde{\psi}}{\partial \beta_1} & \cdots & \frac{\partial \tilde{\psi}}{\partial \beta_N} & \frac{\partial \tilde{\psi}}{\partial q} & \frac{\partial \tilde{\psi}}{\partial \gamma} \end{bmatrix} = \begin{bmatrix} \theta Z^T H^{-1} & -\theta \end{bmatrix} B \tag{A.30}$$

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

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

where

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

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

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

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

$$\frac{\partial \tilde{\psi}}{\partial \gamma} = \frac{\theta f_y f_{RR}}{\Delta}. \tag{A.35}$$

The right hand side of (A.33) is positive because $\theta < 0$, $v'(c_i) > 0$, and $v''(c_i) < 0$. Both terms on the right hand side of (A.34) are positive. The first is positive because $f_y(R, y) > 0$, $f_{yR}(R, y) > 0$, $\Delta > 0$, and $\theta < 0$. The second is positive because $\theta < 0$. Finally, the right hand side of (A.35) is positive because $f_y(R, y) > 0$, $f_{RR}(R, y) < 0$, $\Delta > 0$, and $\theta < 0$. □

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

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

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

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

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

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

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

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

The inequality in (A.39) holds because $\Delta > 0$ and the strict concavity of v and f imply that every term in the sum appearing in the middle term of (A.39) is negative. Because $\theta < 0$, (A.39) implies

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

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. □

Proof of Proposition 5. 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

$$\begin{bmatrix} \psi f_{yy} & \psi f_{yR} \\ f_{yR} & f_{RR} \end{bmatrix} \begin{bmatrix} dy \\ dR \end{bmatrix} = \begin{bmatrix} d\gamma - f_y d\psi \\ dq \end{bmatrix}. \tag{A.41}$$

The solution to the matrix equation (A.41) is

$$\begin{bmatrix} dy \\ dR \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} f_{RR} & -\psi f_{yR} \\ -f_{yR} & \psi f_{yy} \end{bmatrix} \begin{bmatrix} d\gamma - f_y d\psi \\ dq \end{bmatrix}. \tag{A.42}$$

It follows from (A.42) that

$$\frac{\partial y}{\partial \beta_i} = -\frac{f_{RR} f_y}{\Delta} \frac{\partial \psi}{\partial \beta_i} > 0, \quad i = 1, \ldots, N. \tag{A.43}$$

Because $\Delta > 0$, the inequality in (A.43) follows from the strict concavity of f and Proposition 3. Also from (A.42),

$$\frac{\partial R}{\partial \beta_i} = \frac{f_{yR} f_y}{\Delta} \frac{\partial \psi}{\partial \beta_i} > 0, \quad i = 1, \ldots, N. \tag{A.44}$$

The inequality in (A.44) follows from the positivity of Δ, the strict concavity of f, the complementarity of y and R in production, and Proposition 3. Equations (A.43) and (A.44) establish Part (i) of Proposition 5.

Employing (A.42) once more yields

$$\frac{\partial y}{\partial \gamma} = \frac{f_{RR}}{\Delta} - \frac{f_{RR} f_y}{\Delta} \frac{\partial \psi}{\partial \gamma}. \tag{A.45}$$

Substituting (A.35) into (A.45) and rearranging gives

$$\frac{\partial y}{\partial \gamma} = \frac{f_{RR}}{\Delta} \left[1 - \frac{\theta f_y^2 f_{RR}}{\Delta} \right]. \tag{A.46}$$

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

$$\frac{1}{\theta} < \frac{f_y^2 f_{RR}}{\Delta}, \tag{A.47}$$

and, because $\theta < 0$,

$$1 > \frac{\theta f_y^2 f_{RR}}{\Delta}. \tag{A.48}$$

Hence, the term in square brackets on the right hand side of (A.46) is positive. Thus, $\Delta > 0$ and the strict concavity of f imply that $\dfrac{\partial y}{\partial \gamma} < 0$.

Using (A.42) yet again yields

$$\frac{\partial R}{\partial \gamma} = -\frac{f_{yR}}{\Delta} + \frac{f_{yR} f_y}{\Delta} \frac{\partial \psi}{\partial \gamma}. \tag{A.49}$$

Substituting (A.35) into (A.49) and rearranging gives

$$\frac{\partial R}{\partial \gamma} = -\frac{f_{yR}}{\Delta} \left[1 - \frac{\theta f_y^2 f_{RR}}{\Delta} \right]. \tag{A.50}$$

We have already established that the term in square brackets on the right hand side of (A.50) is positive. Because $f_{yR} > 0$ and $\Delta > 0$, the entire right hand side is negative, thereby establishing Part (ii) of Proposition 5. □

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., 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.

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.

Lollivier, S., Rochet, J.-C., 1983. Bunching and second-order conditions: A note on optimal tax theory. Journal of Economic Theory 31, 392–400.

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.

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.