Optimal Nonlinear Taxation of Income and Savings in a Two Class Economy*

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)

November 2005

*Craig Brett's research is generously supported by the Canada Research Chair Programme, the Social Sciences and Humanities Research Council of Canada, and the Canada Foundation for Innovation. We have benefitted from a conversation with Marcus Berliant about the subject of this article.

Abstract

"Optimal Nonlinear Taxation of Income and Savings in a Two Class Economy"

by

Craig Brett and John A. Weymark

Optimal nonlinear taxation of income and savings is considered in a two-period model with two individuals who only differ in their skill levels. When the government can commit to its second period policy, taxes on savings do not form part of the optimal tax mix. When commitment is not possible, the optimal tax scheme distorts private savings behavior. If the types are separated in period one, the savings of the low- (resp. high-) skilled individual are subsidized (resp. taxed) so as to relax an incentive compatibility constraint. If the types are pooled in period one, it is optimal for at least one type to have savings distorted, with the high-skilled individual facing a lower marginal tax rate on savings than the low-skilled individual.

Journal of Economic Literature classification numbers: D82, H21.

Keywords and phrases: asymmetric information, commitment, optimal income taxation, savings taxation, time consistency.

1. Introduction

Ever since the path breaking work of Mirrlees (1971), a government's lack of full information on the tax-relevant characteristics of its citizens has been viewed as a fundamental constraint on the design of nonlinear tax systems. In the context of redistributive income taxation, a taxation authority's egalitarian intentions may be hampered by its inability to identify the respective abilities to pay of different taxpayers. An astute taxation authority is aware, however, that at least some of the tax-relevant characteristics about which it would like more information help to shape the observable behavior of citizens. For example, a worker's unobservable skill level is a determinant of her potentially observable before-tax labor income. Thus, the observable choices that an individual makes convey information to the government about that individual's characteristics. Yet, this information is available only after the tax system has been designed and implemented. Moreover, this information is available at a cost in terms of distortions between consumer and producer prices. Indeed, much of the literature on optimal nonlinear income taxation is devoted to identifying, interpreting and, more rarely, quantifying the distortions in labor supply behavior associated with information-constrained optimal tax systems. Examples include Seade (1977), Guesnerie and Seade (1983), Stiglitz (1983), Tuomala (1990), and Guesnerie (1995).

The information revelation approach to taxation was originally developed for atemporal environments. A major impediment to extending the Mirrlees model to dynamic settings is that information revealed by taxpayers in one period can be used by the government in subsequent periods. Aware of this possibility, rational taxpayers may modify their behavior in early periods in an attempt to better conceal their characteristics. In particular, more able taxpayers might fear the Weitzman (1980) ratchet effect, whereby the government may use its knowledge of ability to pay to extract more taxes from them in the future. The ratchet effect would not arise if the government could commit to forgetting any information it learns at the beginning of each new tax year. However, such a commitment is not credible and, therefore, is not time consistent.

In this article, we investigate redistributive tax policy for a two class economy in which individuals of two productivity types work and consume in each of two periods. These individuals may also transfer resources forward in time through saving. A utilitarian government designs an optimal nonlinear tax system for these consumers in each time period. It can condition tax payments on individual savings. We assume that the preferences of the private individuals are additively separable both across time and between consumption and leisure. These assumptions on preferences guarantee that the optimal marginal tax rate on savings is zero for all individuals when the government can commit to ignore type information revealed in the first period.

We assume that the government is unable to commit to its second period tax policy in advance, and so any information about worker types revealed in the first period can be used when designing second period taxes. The dependence of second period taxes on information revealed in the first period is rationally anticipated by the taxpayers.

We consider the two possible classes of optimal tax systems. In one class, which is termed separating, type information is revealed in the first period. In the other class of systems, called pooling, type information remains hidden after the first period. Whether the optimal scheme is separating or pooling depends on a discrete comparison between the best separating scheme and the best pooling scheme. Such a comparison requires additional assumptions about the functional form of the utility functions. We do not explore the issue of identifying whether separating or pooling types is the global optimum.

A major focus of the literature on redistributive tax policy is the identification of tax instruments to supplement income taxation that, by relaxing an incentive compatibiity constraint, are welfare enhancing. See Boadway and Keen (2000, Section 4.4). When types are separated in the first period, we show that taxes on savings are such an instrument. In this case, it is optimal to subsidize the savings of the less productive individual and to tax the savings of the more productive individual. In other words, the former has her savings distorted upward and the latter has her savings distorted downward. As in atemporal models, in the first period, it is optimal for the high-skilled individual to face a zero marginal income tax rate, whereas the low-skilled individual faces a positive marginal income tax rate. Because there is complete revelation of types in the first period, personalized lump-sum taxes and transfers are optimal in period two.

When types are pooled in period one, it is optimal for the low-skilled individual to have the smaller intertemporal marginal rate of substitution (in absolute value). This pattern of savings distortions is consistent with either taxing or subsidizing the savings of both individuals. However, if only one person has savings subsidized, it must be the high-skilled individual. With pooling in the first period, the second period is a standard one-period optimal income tax problem, and so it is optimal for the high-skilled individual to face a zero marginal income tax rate and for this rate to be positive for the low-skilled individual. In the first period, both individuals are distorted in the labor market, with the low-skilled individual facing an implicit negative marginal income tax rate and the high-skilled individual facing a positive marginal income tax rate.

Much of the early literature on the time consistency of savings taxation, such as Fischer (1980), considered representative agent models with no asymmetric information, thereby excluding distributional concerns from the outset. Heterogeneity and asymmetric information can be introduced into representative agent models by subjecting individuals to person-specific shocks that are unobservable to the government. Redistributive taxation in such settings provides social insurance. For the most part, the contributions to this literature assume that the government can commit to its tax policies.[1] A notable exception is the work of Bisin and Rampini (2006), who assume that the government lacks such commitment. Bisin and Rampini consider a two-period model with both income and savings taxation in which there are shocks that affect incomes directly or indirectly through their effects on labor productivity (as in Mirrlees (1971)). Their main finding is that the power of the taxation authority to take advantage of information revealed in

[1]For references to and discussions of this literature, see Berliant and Ledyard (2005) and Bisin and Rampini (2006).

the first period can be somewhat mitigated if individuals have access to capital markets that prevent the government from observing their total savings. In other words, access to anonymous markets can serve as a welfare-improving constraint on a government that lacks commitment.

Berliant and Ledyard (2005) and Roberts (1984) have studied optimal nonlinear income taxation when the government cannot commit to ignore information gathered in earlier periods in models with and without savings. Berliant and Leydard have identified sufficient conditions for type information to be revealed in the first period of a two-period economy with a continuum of types, whereas Roberts has shown that types are never separated in an infinite horizon economy with a finite number of types provided that the government revenue requirement is not so large as to bankrupt any individual. Neither Berliant and Ledyard nor Roberts consider taxes on savings as a possible instrument.[2]

An important early contribution to redistributive tax policy in the presence of asymmetric information in dynamic settings when the government lacks the ability to commit to its tax policies is that of Ordover and Phelps (1979). They consider optimal nonlinear taxation of income and savings in an overlapping generations environment. In their model, individuals live for two periods, but work only when young. Because retirees make no labor-consumption tradeoffs, it is possible to treat taxes on second period consumption as being paid at the end of the first period. In this way, information revealed to the taxation authority by a young individual does not change the taxes that individual faces when old, and the ratchet effect does not operate. Ordover and Phelps show that it is optimal to tax the savings of most workers (but not the savings of the most skilled) whenever the marginal rate of substitution between consumption when young and consumption when old depends on labor supply. On the other hand, savings should remain untaxed at the margin whenever preferences are separable across time. Using a similar generational structure and informational assumptions, Pirttilä and Tuomala (2001) argue that distorting savings decisions can be optimal even when preferences are separable across time when future relative wages are sensitive to current savings via their effect on capital accumulation.

For Boadway et al. (1996), savings take the form of unobservable investments in education. There are two types of individuals who consume in both of two periods and supply labor at a common wage rate and invest in education in the first period. Labor supply is fixed in the second period, but the wage received depends on the returns to education, which are type specific. Because all individuals are obervationally equivalent in the first period, no private information is revealed until the second period when the government observes incomes (the returns to education), at which time nonlinear income taxation is used for redistributive purposes. As is the case here, Boadway, Marceau, and Marchand identify a policy instrument (in their case, mandating a minimum amount of time spent in publically-observable education) that can help to mitigate the distortions

[2] Dillén and Lundholm (1996) also analyse when it is optimal to separate or pool types in a two-period model with variable consumption and labor supply, but restrict attention to linear income taxation. They, too, do not consider savings taxation.

introduced because of the government's lack of commitment.[3]

Section 2 describes the economy. In Section 3, we characterize the optimal tax system under the assumption that the government can commit to a second period tax schedule before type information is revealed. Section 4 contains our results on optimal taxation when the government cannot commit. We offer concluding remarks in Section 5. Proofs are gathered in an Appendix.

2. The Model

The economy lasts for two time periods. There are two individuals, $i = 1, 2$, with person i suppling l_i^t units of labor and consuming c_i^t units of a single consumption good in period t, $t = 1, 2$.[4] Each consumer may transfer wealth between the two periods by saving the amount s_i of the consumption good. The individuals differ in labor productivity, with the skill level of person i given by the parameter w_i, with $w_1 < w_2$. In keeping with the literature on optimal nonlinear taxation, skill is interpreted as an enhancement to effective labor, so that person i's effective labor in period t is $y_i^t = w_i l_i^t$. The production technology exhibits constant returns to scale. In each period, one unit of effective labor is required to produce one unit of the consumption good. Each unit of the consumption good stored in the first period produces $1 + r$ units of the consumption good in the second period, where $r > 0$. As in Boadway et al. (1996), individuals may not borrow against future income. The labor market is perfectly competitive in each period, so that an individual's effective labor supply equals her income before taxes.

The government wishes to design a tax system that may redistribute income between the individuals. The taxation authority cannot observe labor supply or skill level, but it can observe before-tax income. Moreover, it has the ability to observe savings. Thus, after-tax income x_i^t of person i in period t, the difference between effective labor supply and tax payments, is contingent on before-tax income and savings. Individuals are free to divide their first period after-tax income between consumption and savings. Each unit of savings affords a consumer an additional $1 + r$ units of consumption in the second period over and above her second period after-tax income. Consumption in each period is, therefore, is given by

$$c_i^1 = x_i^1 - s_i, \quad c_i^2 = x_i^2 + (1 + r)s_i, \quad i = 1, 2. \tag{2.1}$$

The individuals have identical preferences over consumption and labor supply, additive in all goods and across time, and represented by the utility function

$$U(c_i^1, l_i^1, c_i^2, l_i^2) = u(c_i^1) - g(l_i^1) + v(c_i^2) - h(l_i^2), \quad i = 1, 2. \tag{2.2}$$

[3]The literature on the taxation of savings also considers a range of normative issues that are not discussed here. See Boadway and Wildasin (1994, Section VI) and Stiglitz (1987, Sections 12–15) for introductions to these topics.

[4]Subscripts index individuals, while superscripts index time periods.

The functions $u(\cdot)$ and $v(\cdot)$ are increasing, strictly concave, and twice continuously differentiable, while the functions $g(\cdot)$ and $h(\cdot)$ are increasing, strictly convex, and twice continuously differentiable. Preferences over the variables that the government can observe are given by

$$u(x_i^1 - s_i) - g\left(\frac{y_i^1}{w_i}\right) + v(x_i^2 + (1+r)s_i) - h\left(\frac{y_i^2}{w_i}\right), \quad i = 1, 2. \tag{2.3}$$

Person i's marginal rate of substitution between before-tax income and after-tax income in the first period is

$$\mathrm{MRS}_{y_i^1, x_i^1} = \frac{g'\left(\frac{y_i^1}{w_i}\right)}{w_i u'(c_i^1)}, \tag{2.4}$$

while her marginal rate of substitution between before-tax income and after-tax income in the second period is

$$\mathrm{MRS}_{y_i^2, x_i^2} = \frac{h'\left(\frac{y_i^2}{w_i}\right)}{w_i v'(c_i^2)}. \tag{2.5}$$

Holding incomes and consumption levels constant, the marginal rates of substitution between before-tax and after-tax income are decreasing in the skill level because the more highly-skilled worker must work fewer additional hours for each additional unit of before-tax income than does the lower-skilled worker. Thus, it takes a smaller increase in after-tax income to compensate the higher-skilled worker for increases in before-tax income than it does to compensate the lower-skilled worker.

Person i's marginal rate of substitution between after-tax income in period one and after-tax income in period two is

$$\mathrm{MRS}_{x_i^1, x_i^2} = -\frac{u'(c_i^1)}{v'(c_i^2)}. \tag{2.6}$$

This intertemporal marginal rate of substitution does not depend explicitly upon the skill level. Because of their common preferences over consumption and labor supply, the two consumers have the same willingness to trade consumption across time periods. The additive nature of preferences implies that the marginal rate of substitution between period one consumption and period two consumption does not depend on the amount of labor supplied in either period.

The government may also engage in saving by storing an amount s_G of the consumption good. The storage technology available to the government is exactly the same as the storage technology for the private sector. Thus, the materials balance constraints for the economy are

$$x_1^1 + x_2^1 + s_G \leq y_1^1 + y_2^1 \tag{2.7}$$

and

$$x_1^2 + x_2^2 \leq y_1^2 + y_2^2 + (1+r)s_G. \tag{2.8}$$

We assume that the government has a utilitarian objective function. Thus, the taxation authority evaluates outcomes using the social welfare function

$$\begin{aligned}\mathcal{W}(x_1^1, x_2^1, y_1^1, y_2^1, x_1^2, x_2^2, y_1^2, y_2^2, s_1, s_2) &= u(x_1^1 - s_1) - g\Big(\frac{y_1^1}{w_1}\Big) + v(x_1^2 + (1+r)s_1) \\ &- h\Big(\frac{y_1^2}{w_1}\Big) + u(x_2^1 - s_2) - g\Big(\frac{y_2^1}{w_2}\Big) + v(x_2^2 + (1+r)s_2) - h\Big(\frac{y_2^2}{w_2}\Big). \end{aligned} \tag{2.9}$$

Note that with this objective function, the government shares the intertemporal preference of the consumers. Browning and Burbidge (1990) have shown that when the government has a different rate of time preference than does the private sector, there is a case for distortionary taxation of savings.[5] In order to focus on the redistributional role of taxation, we do not consider differential time preferences as a rationale for savings taxation.

3. Optimal Taxation with Commitment

First-best taxation is infeasible in this economy because the government cannot distinguish *ex ante* between the two consumers. Thus, only anonymous tax systems are feasible. Because individuals are free to select their optimal work-consumption-savings combinations from the anonymous schedule on offer, the tax system must be incentive compatible; that is, each individual weakly prefers the bundle designed for her to the bundle designed for the other individual. In order to provide a benchmark for our analysis of the tax design problem without commitment, in this section, we assume that the government can commit to a tax system specifying after-tax income in both time periods as a function of labor supply in the two time periods and savings decisions. Specifically, the government is able to credibly commit not to use information about the respective skill levels of the individuals revealed in the first period to adjust taxes in the second period. Thus, incentive compatibility requires that an individual weakly prefers the entire allocation, over both time periods, designed for her to the bundle designed for the other individual; that is,

$$\begin{aligned}&u(x_2^1 - s_2) - g\Big(\frac{y_2^1}{w_2}\Big) + v(x_2^2 + (1+r)s_2) - h\Big(\frac{y_2^2}{w_2}\Big) \\ &\qquad \geq u(x_1^1 - s_1) - g\Big(\frac{y_1^1}{w_2}\Big) + v(x_1^2 + (1+r)s_1) - h\Big(\frac{y_1^2}{w_2}\Big)\end{aligned} \tag{3.1}$$

[5] Browning and Burbidge only consider linear taxation, but their point also applies when taxes are nonlinear.

and

$$u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_1}\right) + v(x_1^2 + (1+r)s_1) - h\left(\frac{y_1^2}{w_1}\right)$$
$$\geq u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_1}\right) + v(x_2^2 + (1+r)s_2) - h\left(\frac{y_2^2}{w_1}\right). \quad (3.2)$$

We assume that only the incentive compatibility condition (3.1) might potentially bind. Given its utilitarian objective, the government wishes to redistribute income from the more highly-skilled worker to the less highly-skilled worker. The natural limit on this redistribution is that, if taken too far, such redistribution might induce the higher-skilled worker to pretend to be the lower-skilled worker. Imposing (3.1) prevents this type of mimicking.[6] Thus, the problem faced by the taxation authority can be specified as follows:

The Second-Best Tax Design Problem with Commitment. The government chooses an allocation $(x_1^1, x_2^1, y_1^1, y_2^1, x_1^2, x_2^2, y_1^2, y_2^2, s_1, s_2, s_G)$ to maximize the social welfare function (2.9) subject to the materials balance constraints (2.7) and (2.8) and the two-period incentive compatibility constraint (3.1).[7]

The second-best tax design problem with commitment is a standard one-dimensional screening problem. Because there are five components to each individual's allocation, the taxation authority has more instruments than the minimum required to achieve separation.[8] Given the adverse selection problem faced by the government, some distortion to the behavior of at least one of the individuals is inevitable. Proposition 1 describes the pattern of distortions at a solution to the government's problem.

Proposition 1. *At a solution to the second-best tax design problem with commitment:*

(i) $\mathrm{MRS}_{y_2^1,x_2^1} = 1$, $\mathrm{MRS}_{y_2^2,x_2^2} = 1$ *and* $\mathrm{MRS}_{x_2^1,x_2^2} = -(1+r)$.

(ii) $\mathrm{MRS}_{y_1^1,x_1^1} < 1$, $\mathrm{MRS}_{y_1^2,x_1^2} < 1$ *and* $\mathrm{MRS}_{x_1^1,x_1^2} = -(1+r)$.

Part (i) of Proposition 1 is a familiar no distortion result for the high-skilled individual. In both periods, this individual faces a zero marginal income tax rate. Part (ii) of Proposition 1 describes the distortions caused by the asymmetric information. Because the first-best solution is not incentive compatible, constraint (3.1) must bind at a solution to the second-best problem with commitment. It follows from Brito et al. (1990, Proposition 5) that the marginal rate of substitution for person 1 is distorted only for those

[6]Indeed, at a solution to the first-best taxation problem for this economy, the government wishes to equalize the consumption of both individuals in each time period and to require the more skilled individual to work more. Thus, (3.1) is violated at the first-best allocation, while (3.2) is slack.

[7]In all of our tax design problems, we assume that the omitted nonnegativity constraints do not bind. We also assume that each of these problems has a solution.

[8]Separation is possible in two-good worlds when there is asymmetric information in one dimension.

pairs of goods for which persons 1 and 2 have a different marginal rate of substitution at person 1's allocation. Because the marginal rates of substitution between before-tax income and after-tax income vary by skill level, the effective labor-consumption margin is distorted in each period, and so person 1 faces a positive implicit marginal income tax rate in both periods.[9] On the other hand, the two individuals have the same intertemporal preferences. In particular, individual 1 and individual 2 considering the opportunity to mimic individual 1 are willing to trade consumption across time at the same implicit prices. Thus, there is no informational advantage to be had by changing the intertemporal relative price of consumption. Therefore, savings decisions are not distorted, and hence not taxed, at the margin.

4. Optimal Taxation without Commitment

The government's ability to commit in the first period to the second-period taxation scheme is not credible. The optimal two-period scheme with commitment offers different allocations to the two individuals in the first period. With full knowledge of the workings of the economy, this allows the taxation authority to infer the identities of the individuals at the end of the first period. The information asymmetry between the government and the private sector disappears, and there is no need to distort behavior in the second period. Because the optimal second-best scheme with commitment features a distortion in the period two labor supply of individual 1, it would not be chosen by a taxation authority with the ability to re-optimize after the first period. Furthermore, because savings decisions have already been fixed in the first period, the government will have an incentive to increase the implicit tax on savings of the high-skilled individual beyond what is optimal with commitment in order to further its redistributional goals. Hence, the optimal scheme with commitment is time inconsistent.

The private agents are aware that the government is able to use information gleaned in the first period when setting second-period taxes. In particular, the more highly-skilled individual understands that if her type is revealed in the first period, then the taxation authority will have a easier time redistributing income from her to the lower-skilled individual in the second period because it no longer needs to worry about incentive compatibility constraints. This can be accomplished by transfering more of the high-skilled individual's savings to the low-skilled individual and/or providing an incentive for the high-skilled individual to work more so that there is more of this person's income available to redistribute. Thus, there is an increased incentive for the more highly-skilled individual to conceal information in the first period.

For its part, the taxation authority is aware of the added incentive to hide information in the first period. It realizes that the full-commitment tax schedule may need to be modified in order to induce information revelation in the first period. As pointed out by

[9] Optimal income tax schedules may be nondifferentiable. Person i's implicit marginal income tax rate in period t is $1 - \mathrm{MRS}_{y_i^t, x_i^t}$. Because $\mathrm{MRS}_{y_i^t, x_i^t} > 0$, marginal income tax rates are bounded above by one.

Freixas et al. (1985) in a more general planning context, and by Dillén and Lundholm (1996) for linear income taxes, such modifications may be sufficiently costly to lead the government to prefer not to separate types in the first period. The taxation authority must compare the gains accruing from the use of first-best taxation in the second period to the costs incurred in the first period of extracting the information it needs to implement the second period first-best allocation.

Type information remains hidden only if both individuals choose the same before-tax income, after-tax income, and savings in the first period. Such an outcome is called a pooling outcome. If any component of the first-period allocation differs by type, information is revealed and the types are separated. The tax schedule on offer in the first period and the anticipated tax schedule for the second period shape the choices of the two individuals and implicitly determine whether there is pooling or separation in the first period. The first period revelation outcome is discrete; either there is pooling or there is separation. Deciding which of the two configurations is better requires a comparison between the maximized values of social welfare in the two cases. In general, such a comparison depends on the exact form of the utility function. Before making this comparison, the taxation authority must be able to determine the optimal separation mechanism and the optimal pooling mechanism.

4.1. Separation in the First Period

If the two individuals make different choices in the first period, then the government has sufficient information to carry out lump-sum taxation in the second period. The two private individuals and the government enter the second period with an endowment of the consumption good equal to the amount of their savings augmented by the factor $1+r$. The taxation authority may levy taxes on the savings of the private agents as well as on their labor incomes. In this way, it determines the net-of-all-taxes incomes of the two individuals in period two, x_1^2 and x_2^2. Second period social welfare is given by the sum of individual utilities, which, using (2.1) and (2.3), is

$$\mathcal{W}^2(x_1^2, x_2^2, y_1^2, y_2^2, s_1, s_2) = v(x_1^2 + (1+r)s_1) - h\left(\frac{y_1^2}{w_1}\right) + v(x_2^2 + (1+r)s_2) - h\left(\frac{y_2^2}{w_2}\right). \tag{4.1}$$

The problem faced by the taxation authority in the second period is:

The Second Period First-Best Problem. Given (s_1, s_2, s_G), the government chooses an allocation $(x_1^2, x_2^2, y_1^2, y_2^2)$ to maximize the second period social welfare function (4.1) subject to the materials balance constraint (2.8).

The second period first-best problem has a strictly concave objective function and a single linear constraint, which can easily be shown to bind at the solution to this problem. Each of the four components of the solution to the second period first-best problem depends on the vector $\mathbf{s} = (s_1, s_2, s_G)$ of predetermined savings levels. Because the problem is so well-behaved, its comparative static properties with respect to each component of the savings vector can be derived using standard methods from consumer theory. The properties most pertinent to a characterization of the distortions arising in a two-period taxation system with separation in the first period are collected in the following lemma.

Lemma 1. *For a given savings vector* $\mathbf{s}$*, the second period first-best problem has a unique solution. Moreover, the solution functions* $x_1^2(\mathbf{s})$*,* $x_2^2(\mathbf{s})$*,* $y_1^2(\mathbf{s})$*, and* $y_2^2(\mathbf{s})$ *are continuously differentiable and satisfy the following conditions:*

(i) $v'(x_1^2(\mathbf{s}) + (1+r)s_1) = v'(x_2^2(\mathbf{s}) + (1+r)s_2) = \dfrac{1}{w_1} h'\Big(\dfrac{y_1^2(\mathbf{s})}{w_1}\Big) = \dfrac{1}{w_2} h'\Big(\dfrac{y_2^2(\mathbf{s})}{w_2}\Big).$

(ii) $\dfrac{\partial x_1^2(\mathbf{s})}{\partial s_i} + \dfrac{\partial x_2^2(\mathbf{s})}{\partial s_i} - \dfrac{\partial y_1^2(\mathbf{s})}{\partial s_i} - \dfrac{\partial y_2^2(\mathbf{s})}{\partial s_i} = 0, \quad i = 1, 2.$

With separation, in the second period, we have a full information planning problem in which the government has access to the interest-augmented savings from the first period to distribute as it wishes. Part (i) of Lemma 1 summarizes the marginal conditions for a first-best utilitarian optimum in the second period. The taxation authority wishes to equate the marginal utilities of consumption for the two individuals. Given identical additively separable preferences, equality of the marginal utilities of consumption implies equal consumption for the two individuals. Furthermore, for each worker, the marginal rate of substitution between labor and consumption equals the wage rate. Because person 2 has a higher wage rate, she also has a higher marginal disutility of labor at the first-best optimum. Given identical preferences with increasing marginal disutility of labor, person 2 must work more than does person 1. Because agreeing to work more than someone else for equal consumption is not incentive compatible, the taxation authority must make use of the skill information revealed in the first period in order to implement this scheme using person-specific lump sum taxes and transfers.

Part (ii) of Lemma 1 follows directly from the second period materials balance constraint. This does not mean that optimal second period before-tax and after-tax incomes are insensitive to individual wealth holdings. Indeed, it is feasible for the taxation authority to tax away all first-period savings. Instead, part (ii) simply says that changes in aggregate production are offset by changes in after-tax income. However, as is apparent from (2.1), the effect of a one unit increase in individual savings on aggregate second period consumption exceeds the effect of a one unit increase in individual savings on aggregate after-tax income by the factor $(1+r)$, the gross return on that unit of savings.

All decision makers in the economy, both private and public, recognize that the government is unable to commit to any second period taxation scheme apart from the one that is the second period optimum, given first period savings. The two private individuals take this into account when deciding on their first period courses of action, notably when making their savings decisions. Moreover, the taxation authority must provide sufficient incentive for individual 2 to reveal her type in the first period. Such an incentive is provided if the following condition is met:

$$u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) + v(x_2^2(\mathbf{s}) + (1+r)s_2) - h\left(\frac{y_2^2(\mathbf{s})}{w_2}\right)$$
$$\geq u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_2}\right) + v(x_2^2(\mathbf{s}) + (1+r)s_2) - h\left(\frac{y_2^2(\mathbf{s})}{w_2}\right). \quad (4.2)$$

The final two terms on each side of relation (4.2) are identical because, given that her true identity is revealed in period one, person 2 can not pretend to be person 1 in the second period. Additive separability of preferences across time, therefore, implies that the incentive compatibility condition (4.2) is equivalent to

$$u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) \geq u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_2}\right). \quad (4.3)$$

Condition (4.3) is formally equivalent to the incentive compatibility condition facing a taxation authority in a one-period economy.

The government designs its first period tax system fully aware of how it will respond in the second period to its own first period actions and to the savings decisions of the private individuals. Its first period objective function, which includes the social welfare accruing in the second period, is

$$\mathcal{W}^{sep}(x_1^1, x_2^1, y_1^1, y_2^1, s_1, s_2, s_G) = u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_1}\right) + v(x_1^2(\mathbf{s}) + (1+r)s_1)$$
$$- h\left(\frac{y_1^2(\mathbf{s})}{w_1}\right) + u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) + v(x_2^2(\mathbf{s}) + (1+r)s_2) - h\left(\frac{y_2^2(\mathbf{s})}{w_2}\right). \quad (4.4)$$

Because both the incentive compatibility condition (4.3) and the objective function (4.4) include the solution functions to the second period first-best problem, the materials balance constraint in period two is accounted for. However, the taxation authority must take account of the first period materials balance constraint. Thus, the government faces the following tax design problem in period one.

The First Period No-Commitment Tax Design Problem with Separation. The government chooses a first period allocation $(x_1^1, x_2^1, y_1^1, y_2^1, s_1, s_2, s_G)$ to maximize the objective function (4.4) subject to the first period materials balance constraint (2.7) and the incentive compatibility constraint (4.3).

The pattern of of distortions to labor supply and savings behavior arising at a solution to the first period no-commitment tax design problem with separation are given in the following proposition.

Proposition 2. *At a solution to the first period no-commitment tax design problem with separation:*

(i) $\mathrm{MRS}_{y_1^1,x_1^1} < 1$ *and* $\mathrm{MRS}_{y_2^1,x_2^1} = 1$.

(ii) $\mathrm{MRS}_{x_1^1,x_1^2} < -(1+r)$ *and* $\mathrm{MRS}_{x_2^1,x_2^2} > -(1+r)$.

Part (i) of Proposition 2 indicates that, at a solution to the no-commitment tax design problem with separation, person 1 faces a positive first period implicit marginal income tax rate, while person 2 faces a zero first period implicit marginal income tax rate. In this respect, the optimal policy mix is akin to an optimal nonlinear tax scheme for a one-period economy. Separability of preferences across time implies that the marginal rate of substitution between first period before-tax income and first period after-tax income is independent of the second period allocation. Therefore, the existence of a future period has no effect on the type of labor supply distortions needed to induce revelation. The magnitude of the implicit marginal tax rate on person 1's income may, however, differ from the corresponding implicit marginal tax rate in a one-period economy. Anticipated future events help to shape savings decisions, which directly affect first period consumption and person 1's marginal rate of substitution between consumption and labor supply in the first period.

Part (ii) of Proposition 2 shows that the government's lack of commitment to a second period tax scheme results in a distortion to each individual's savings decision, with the distortions running in opposite directions. Individual 1 has savings implicitly subsidized at the margin, while individual 2's savings are implicitly taxed.

The motivation for subsidizing person 1's savings can be understood by considering the marginal effect on optimized social welfare of an infinitesimal increase in s_1, which we show in (A.34) in the Appendix to be equivalent to

$$-u'(c_1^1) + \psi u'(c_1^1) + (1+r)v'(c_1^2), \tag{4.5}$$

where ψ is the social value of a one unit relaxation of the incentive compatibility constraint (4.3). The first term in expression (4.5) is the reduction in person 1's utility due to the deferral of consumption. This deferral of first period consumption by person 1 makes the allocation on offer to person 1 less attractive to person 2, thereby relaxing her incentive constraint. The value of this relaxation is given by the second term in (4.5). The third term in (4.5) is the social benefit to using the returns to person 1's savings in period two. Each unit of savings produces $(1+r)$ units of resources to be distributed as the government wishes among increases in consumption and/or decreases in labor supply in period two. As noted in Lemma 1, each of the four competing uses of these additional resources are equally valued at the margin at a first best in the second period. Moreover, given the utilitarian objective, the marginal value of additional resources equals the common marginal utility of consumption of the two individuals. Thus, the marginal value of an increase in second period resources is $v'(c_1^2)$. At a solution to the no-commitment tax design problem with separation, the marginal value of an increase in s_1 in (4.5) must be zero, which implies that

$$\mathrm{MRS}_{x_1^1,x_1^2} = -\frac{u'(c_1^1)}{v'(c_1^2)} = -\frac{(1+r)}{(1-\psi)} < -(1+r). \tag{4.6}$$

The optimality of an implicit tax on the savings of individual 2 follows from a similar reasoning. Starting at an allocation for which $\mathrm{MRS}_{x_2^1,x_2^2} = -(1+r)$, a one unit decrease in s_2 financed by a $(1+r)$ decrease in the aggregate resources available in period two is a matter of direct welfare indifference. Yet, that decrease in s_2 should be encouraged, for the associated increase in c_1^2 slackens the incentive compatibility constraint.

In summary, it is the extra benefit of relaxing the incentive constraint (compared to the full information solution) that results when person 1's first period consumption is decreased and person 2's first period consumption is increased that accounts for the upward distortion of the savings of individual 1 and the downward distortion of the savings of person 2. In other words, compared to the savings required for the intertemporal substitution of consumption to be first-best optimal, it is socially beneficial to have person 1 to save more and to have person 2 to save less. We thus have another instance of the observation made by Boadway and Keen (2000, Section 4.4) that distortionary policy instruments that would not be used in the absence of asymmetric information are valuable when there is private information if these instruments can relax an incentive compatibility constaint.

The role played by the lack of commitment to a second period tax schedule can be illustrated by contrasting how a government like the one considered in Section 3 can improve upon the second-best optimum without commitment. Suppose that such a government is trying to find a way to improve upon an initial allocation in which $\mathrm{MRS}_{x_2^1,x_2^2} > -(1+r)$. It can decrease c_2^1 a small unit by, for example, increasing s_2, and simultaneously increasing c_2^2 by $|\mathrm{MRS}_{x_2^1,x_2^2}|$ while keeping the utility of individual 2 constant. Such a change has no effect on the two-period incentive compatibility constraint. Moreover, this change is resource saving, for the increase in c_2^2 is less than the second period return $(1+r)$ on the additional unit of savings. When the taxation authority lacks the ability to commit to the tax schedule in period two, the promise to compensate person 2 for a decrease in first period consumption is hollow. Instead, the extra resources will be used in the manner prescribed by the second period first-best optimum. Typically, any additional resources available in the second period are split among the two agents. Thus, the initial loss in utility due to a reduction in c_2^1 is only partially offset. The utility of person 1 actually increases because the taxation authority redistributes some of the additional second period resources to her. However, the welfare gain to person 1 is counteracted by a tightening of the incentive compatibility constraint due to the net redistribution from person 2 to person 1.

4.2. Pooling in the First Period

The only circumstance in which the taxation authority cannot infer the identities of the two individuals after the first period is when they make identical choices in that period. In particular, they choose a common level of savings s. The government can observe this level of savings and its own savings. Social welfare in the second period is affected by individual savings, and is given by

$$\mathcal{W}^{2,pool}(x_1^2, x_2^2, y_1^2, y_2^2, s) = v(x_1^2 + (1+r)s) - h\left(\frac{y_1^2}{w_1}\right) + v(x_2^2 + (1+r)s) - h\left(\frac{y_2^2}{w_2}\right). \tag{4.7}$$

Because the government enters the second period without full knowledge of the individuals' types, its tax design problem is constrained by the incentive compatibility requirement

$$v(x_2^2 + (1+r)s) - h\left(\frac{y_2^2}{w_2}\right) \geq v(x_1^2 + (1+r)s) - h\left(\frac{y_1^2}{w_2}\right). \tag{4.8}$$

The problem faced by the government in the second period is:

The Second Period Tax Design Problem with Pooling. Given $\tilde{\mathbf{s}} = (s, s_G)$, the government chooses a second period allocation $(x_1^2, x_2^2, y_1^2, y_2^2)$ to maximize the objective function (4.7) subject to the second period materials balance constraint (2.8) and the incentive compatibility constraint (4.8).

Apart from the dependence of the utility functions on the parameter s and the dependence of the resource constraint on s_G, the second period tax design problem with pooling is a standard optimal nonlinear taxation problem. Given the utilitarian nature of the objective function, the problem is strictly redistributive in the sense of Guesnerie (1995, p. 224). Hence, the optimal second period allocation features the usual pattern of distortions: individual 1 faces a positive implicit marginal income tax rate and individual two has a zero implicit marginal income tax rate. In other words, person 1's marginal rate of substitution between second period income and consumption is less than one, whereas this marginal rate of substitution is equal to one for person 2. Moreover, individual 1 has less of both second period consumption and income than does individual 2. This pattern of distortions is summarized in Lemma 2.

Lemma 2. $\mathrm{MRS}_{y_1^2, x_1^2} < 1$ *and* $\mathrm{MRS}_{y_2^2, x_2^2} = 1$ *at a solution to the second period no-commitment tax design problem with pooling.*

The taxation authority foresees the impact of second period decisions when designing the optimal first period scheme with pooling. Thus, given pooling in the first period, its objective function is

$$\begin{aligned}\mathcal{W}^{pool}(x^1, y^1, s, s_G) = {} & u(x^1 - s) - g\left(\frac{y^1}{w_1}\right) + v(x_1^2(\tilde{\mathbf{s}}) + (1+r)s) \\ & - h\left(\frac{y_1^2(\tilde{\mathbf{s}})}{w_1}\right) + u(x^1 - s) - g\left(\frac{y^1}{w_2}\right) + v(x_2^2(\tilde{\mathbf{s}}) + (1+r)s) - h\left(\frac{y_2^2(\tilde{\mathbf{s}})}{w_2}\right),\end{aligned} \tag{4.9}$$

where variables without subscripts indicate values that are identical for the two individuals. Because the objective function (4.9) incorporates the second period decisions of the taxation authority, it takes account of the period two materials balance condition (2.8) and the incentive compatibility condition (4.8). Because the two individuals have identical observable allocations in period one, overall incentive compatibility is equivalent to (4.8). Hence, the only constraint remaining on the first period tax design problem is (2.7), the materials balance constraint for period one. When the individuals are pooled, (2.7) reduces to

$$2x^1 + s_G \leq 2y^1. \tag{4.10}$$

The government's first period decision problem can, therefore, be described as follows:

The First Period No-Commitment Tax Design Problem with Pooling. The government chooses a first period allocation (x^1, y^1, s, s_G) to maximize the objective function (4.9) subject to the first period materials balance constraint (4.10).

The objective function of the first period no-commitment tax design problem with pooling depends on first period savings is two ways. There is a dependence due to the direct effects of private savings on consumption in each period. There are also indirect effects that depend on how the components of the optimal second period allocation depend on public and private savings. However, the exact comparative static responses of the optimal second period allocations to savings are difficult to determine. This is hardly surprising, for comparative static results for nonlinear income taxes have only been established when utility functions are quasilinear.[10] Without general comparative static results concerning the second period problem, it is impossible to fully characterize the pattern of distortions in the first period. However, as the following proposition demonstrates, some distortion of both first period labor supply and savings is required.

Proposition 3. *At a solution to the first period no-commitment tax design problem with pooling:*

(i) $\mathrm{MRS}_{y_1^1, x_1^1} > 1 > \mathrm{MRS}_{y_2^1, x_2^1}$.[11]

(ii) $\mathrm{MRS}_{x_1^1, x_1^2} > \mathrm{MRS}_{x_2^1, x_2^2}$.[12]

Because the utility function is additively separable in labor and consumption, equal consumption in the first period implies equal marginal utility of consumption in that period. Equal incomes in period one imply that person 2 has a smaller marginal disutility

[10] In their comparative static analyses, Weymark (1987), Brett and Weymark (2004), and Hamilton and Pestieau (2005) consider preferences that are quasilinear in labor supply, while Boadway and Pestieau (2004) consider preferences that are quasilinear in consumption. Imposing quasilinearity is inappropriate in our model because it renders the second period first-best outcome under separation indeterminate.

[11] Using the first-order conditions for the first period no-commitment tax design problem with pooling, it is straightforward to show that $\mathrm{MRS}_{y_1^1, x_1^1} + \mathrm{MRS}_{y_2^1, x_2^1} = 2$.

[12] Recall that marginal rates of substitution between after-tax incomes in the two periods are negative, so in absolute value, person 1 has the smaller intertemporal marginal rate of substitution.

of labor in period one than does person 1. The monotonicity of second period consumption in type implies that individual 2 has a lower marginal utility of consumption in the second period than does individual 1. Part (ii) of Proposition 3 and the requirement that $\mathrm{MRS}_{y_1^1,x_1^1} > \mathrm{MRS}_{y_2^1,x_2^1}$ in part (i) then follow from the definitions of the marginal rates of substitution in (2.4) and (2.6).

If $\mathrm{MRS}_{y_2^1,x_2^1} \geq 1$ (and, hence, $\mathrm{MRS}_{y_1^1,x_1^1} > 1$), it is feasible to infinitesimally decrease their common first period consumption and before-tax income by the same amount holding savings fixed. Because savings are held constant, this change has no effect on the second period incentive compatibility constraint. But this change is a Pareto improvement, so it must be optimal to have $\mathrm{MRS}_{y_2^1,x_2^1} < 1$.[13] By reversing the direction of change in first period consumption and before-tax income, it follows that it is also optimal to have $\mathrm{MRS}_{y_1^1,x_1^1} > 1$.

In the absence of informational and commitment constraints, the government could engineer an increase in social welfare by facilitating an intertemporal trade of consumption between the two consumers. This can be accomplished by transferring a unit of consumption from person 1 to person 2 in period one and reversing the transfer in period two. Because the marginal utilities of consumption are equal in period one, there is no change in overall welfare as a result of the period one transfer. The overall sum of utilities increases when the second period transfers are taken into account because person 1's marginal utility of consumption exceeds that of person 2 in period two. However, if the individuals were to accept such a trade, the allocations of the two individuals would differ in period one and, hence, their identities would be revealed. The government would use this information to completely redesign the second period taxation schedule, thereby undermining its original intentions.

Unfortunately, it does not appear possible to ascertain the general direction of the distortion in savings behavior at a solution to the no-commitment tax design problem with pooling. This ambiguity arises because, without further assumptions, it is not possible to determine whether increases in savings tighten or slacken the second period incentive compatibility condition (4.8). At first glance, it appears that decreasing savings might slacken (4.8). Near the second period optimum, the marginal utility of person 1's consumption is higher than person 2's marginal utility of consumption. Thus, a decrease in second period consumption has a greater impact on person 2 as a mimicker of person 1 than it does on the utility of person 2 acting truthfully. However, this effect can be completely undone by appropriate changes in the second period after-tax incomes of the two individuals. Whether the government decides to carry out such adjustments to after-tax incomes or not depends on the functional form of the utility function.

Nevertheless, Proposition 3 tells us that in absolute value, person 1 has the smaller intertemporal marginal rate of substitution. This is the reverse of what was found when it is optimal to separate the types in the first period. Recall that in that case, subsidizing person 1's savings and taxing person 2's savings helped relax the incentive compatibility

[13]Except in the borderline case in which $\mathrm{MRS}_{y_2^1,x_2^1} = 1$, both people benefit from this change. In the borderline case, this change is a matter of indifference to person 2, but strictly benefits person 1.

constraint in period one. With pooling, this is no longer an issue. However, in order for there to be pooling in the first period and separation in the second, it is necessary for person 1 to have the smaller intertemporal marginal rate of substitution in absolute value. This pattern of savings distortions is consistent with either taxing or subsidizing the savings of both individuals. However, if only one person has savings subsidized, it must be person 2.

Because there is pooling, person 2 faces a higher implicit marginal income tax rate than person 1 in period one, which is the reverse of what we found when it is optimal to separate the types. Interestingly, person 1 faces a negative implicit marginal income tax rate (i.e., person 1's labor supply is subsidized) and person 2 faces a positive implicit marginal income tax rate.[14]

5. Conclusion

Our analysis suggests that problems of time inconsistency in the design of nonlinear income taxes provide an rationale for distortions in savings behavior. Extending our analysis to many-consumer economies is not straightforward. It is easy to construct models of static nonlinear income taxation that exhibit considerable bunching (see, for example, Weymark (1986)). Dynamic extensions of such models would invariably uncover cases of pooling, semi-pooling, and separation, each with its own distinct pattern of savings distortions. The fundamental insight of this paper—that time inconsistency renders some forms of savings distortions irreducible—is likely to carry over to economies with any number of consumers.

Despite the simplifications inherent in our model, it is interesting to compare the pattern of savings distortions to the pattern implicit in some forms of mandatory public pension schemes. The Canada Pension Plan, for example, is a mandatory defined contributions scheme that possibly acts as a form of forced savings for low-income workers. These workers, like the low-skilled workers in our model with separation, face an implicit marginal subsidy to savings. Higher income Canadians are more likely to access other forms of pensions. For these workers, the Canada Pension Plan does not constitute forced savings, and the marginal implicit marginal subsidy disappears. While not intended as a model of public pensions, our analysis does point to the potential role for the selective encouragement (or coercion) of savings as a way to alleviate informational constraints in dynamic settings.

[14]Dillén and Lundholm (1996) have found in their model of dynamic linear income taxation without commitment that pooling with a negative first period marginal income tax rate may be optimal.

Appendix

Proof of Proposition 1. The Lagrangian associated with the second-best tax design problem with commitment is

$$
\begin{aligned}
u(x_1^1 - s_1) - g\Big(\frac{y_1^1}{w_1}\Big) + v(x_1^2 + (1+r)s_1) - h\Big(\frac{y_1^2}{w_1}\Big) \\
+ u(x_2^1 - s_2) - g\Big(\frac{y_2^1}{w_2}\Big) + v(x_2^2 + (1+r)s_2) - h\Big(\frac{y_2^2}{w_2}\Big) \\
+ \lambda^1\Big[y_1^1 + y_2^1 - x_1^1 - x_2^1 - s_G\Big] + \lambda^2\Big[y_1^2 + y_2^2 + (1+r)s_G - x_1^2 - x_2^2\Big] \\
+ \mu\Big[u(x_2^1 - s_2) - g\Big(\frac{y_2^1}{w_2}\Big) + v(x_2^2 + (1+r)s_2) - h\Big(\frac{y_2^2}{w_2}\Big) \\
- u(x_1^1 - s_1) + g\Big(\frac{y_1^1}{w_2}\Big) - v(x_1^2 + (1+r)s_1) + h\Big(\frac{y_1^2}{w_2}\Big)\Big]. \quad \text{(A.1)}
\end{aligned}
$$

The associated first-order conditions for an interior solution are:

$$x_1^1:\ u'(c_1^1) - \lambda^1 - \mu u'(c_1^1) = 0; \quad \text{(A.2)}$$
$$x_2^1:\ u'(c_2^1) - \lambda^1 + \mu u'(c_2^1) = 0; \quad \text{(A.3)}$$
$$y_1^1:\ -\frac{1}{w_1} g'\Big(\frac{y_1^1}{w_1}\Big) + \lambda^1 + \frac{\mu}{w_2} g'\Big(\frac{y_1^1}{w_2}\Big) = 0; \quad \text{(A.4)}$$
$$y_2^1:\ -\frac{1}{w_2} g'\Big(\frac{y_2^1}{w_2}\Big) + \lambda^1 - \frac{\mu}{w_2} g'\Big(\frac{y_2^1}{w_2}\Big) = 0; \quad \text{(A.5)}$$
$$x_1^2:\ v'(c_1^2) - \lambda^2 - \mu v'(c_1^2) = 0; \quad \text{(A.6)}$$
$$x_2^2:\ v'(c_2^2) - \lambda^2 + \mu v'(c_2^2) = 0; \quad \text{(A.7)}$$
$$y_1^2:\ -\frac{1}{w_1} h'\Big(\frac{y_1^2}{w_1}\Big) + \lambda^2 + \frac{\mu}{w_2} h'\Big(\frac{y_1^2}{w_2}\Big) = 0; \quad \text{(A.8)}$$
$$y_2^2:\ -\frac{1}{w_2} h'\Big(\frac{y_2^2}{w_2}\Big) + \lambda^2 - \frac{\mu}{w_2} h'\Big(\frac{y_2^2}{w_2}\Big) = 0; \quad \text{(A.9)}$$
$$s_1:\ -u'(c_1^1) + (1+r)v'(c_1^2) + \mu u'(c_1^1) - (1+r)\mu v'(c_1^2) = 0; \quad \text{(A.10)}$$
$$s_2:\ -u'(c_2^1) + (1+r)v'(c_2^2) - \mu u'(c_2^1) + (1+r)\mu v'(c_2^2) = 0; \quad \text{(A.11)}$$
$$s_G:\ -\lambda^1 + (1+r)\lambda^2 = 0. \quad \text{(A.12)}$$

The first equality of part (i) follows from solving each of (A.3) and (A.5) for λ^1 and rearranging the resulting equality. Similar algebra applied to (A.7) and (A.9) yields the second equality. From (A.11),

$$(1+\mu)u'(c_2^1) = (1+\mu)(1+r)v'(c_2^2), \quad \text{(A.13)}$$

from which the final equality of part (i) follows.

By (A.2) and (A.4),

$$(1-\mu)u'(c_1^1) = \frac{1}{w_1}g'\Big(\frac{y_1^1}{w_1}\Big) - \frac{\mu}{w_2}g'\Big(\frac{y_1^1}{w_2}\Big) = \lambda^1. \tag{A.14}$$

Because $w_1 < w_2$ and $g(\cdot)$ is strictly convex,

$$\frac{1}{w_1}g'\Big(\frac{y_1^1}{w_1}\Big) - \frac{\mu}{w_2}g'\Big(\frac{y_1^1}{w_2}\Big) > \frac{(1-\mu)}{w_1}g'\Big(\frac{y_1^1}{w_1}\Big). \tag{A.15}$$

Combining (A.14) and (A.15) yields

$$(1-\mu)u'(c_1^1) > \frac{(1-\mu)}{w_1}g'\Big(\frac{y_1^1}{w_1}\Big). \tag{A.16}$$

Because the multiplier on the resource constraint, λ^1, is positive, (A.14) implies that $(1-\mu)u'(c_1^1)$ is positive. Dividing both sides of (A.16) by $(1-\mu)u'(c_1^1)$ and rearranging yields the first inequality of part (ii). The second inequality follows from a similar argument applied to (A.6) and (A.8). From (A.10),

$$(1-\mu)u'(c_1^1) = (1-\mu)(1+r)v'(c_1^2), \tag{A.17}$$

from which the final equality of part (ii) follows. □

Proof of Lemma 1. The objective function of the second period first-best problem is strictly concave and the constraint set is convex. Hence, by Sundaram (1996, Theorem 7.14), the problem has a unique solution. The Lagrangian associated with the optimization problem is

$$v(x_1^2 + (1+r)s_1) - h\Big(\frac{y_1^2}{w_1}\Big) + v(x_2^2 + (1+r)s_2) - h\Big(\frac{y_2^2}{w_2}\Big) \\ + \lambda\Big[y_1^2 + y_2^2 + (1+r)s_G - x_1^2 - x_2^2\Big]. \tag{A.18}$$

The first-order conditions for an optimum are:

$$x_1^2\colon\ v'(c_1^2) - \lambda = 0; \tag{A.19}$$

$$x_2^2\colon\ v'(c_2^2) - \lambda = 0; \tag{A.20}$$

$$y_1^2\colon\ -\frac{1}{w_1}h'\Big(\frac{y_1^2}{w_1}\Big) + \lambda = 0; \tag{A.21}$$

$$y_2^2\colon\ -\frac{1}{w_2}h'\Big(\frac{y_2^2}{w_2}\Big) + \lambda = 0; \tag{A.22}$$

$$\lambda\colon\ y_1^2 + y_2^2 + (1+r)s_G - x_1^2 - x_2^2 = 0. \tag{A.23}$$

Part (i) of the lemma follows from solving each of (A.19)–(A.22) for λ.

The bordered Hessian matrix for this problem is

$$A = \begin{bmatrix} v''(c_1^2) & 0 & 0 & 0 & -1 \\ 0 & v''(c_2^2) & 0 & 0 & -1 \\ 0 & 0 & -\frac{h''(l_1^2)}{(w_1)^2} & 0 & 1 \\ 0 & 0 & 0 & -\frac{h''(l_2^2)}{(w_2)^2} & 1 \\ -1 & -1 & 1 & 1 & 0 \end{bmatrix}. \tag{A.24}$$

Its determinant is

$$|A| = v''(c_1^2)v''(c_2^2)\Big[\frac{h''(l_1^2)}{(w_1)^2} + \frac{h''(l_2^2)}{(w_2)^2}\Big] - \frac{h''(l_1^2)}{(w_1)^2}\frac{h''(l_2^2)}{(w_2)^2}\Big[v''(c_1^2) + v''(c_2^2)\Big]. \tag{A.25}$$

By the strict concavity of $v(\cdot)$, the first two factors of the first term on the right-hand side of (A.25) are negative. Strict convexity of $h(\cdot)$ implies that the term inside the first square bracket in (A.25) is positive. On the other hand, the first two factors in the second term are positive, while the term in square brackets is negative. Thus, (A.25) expresses $|A|$ as a positive quantity minus a negative quantity. Hence, $|A| > 0$ and A is invertible. It then follows from the Implicit Function Theorem (see Sundaram (1996, Theorem 1.77)) that the solution functions are continuously differentiable. Part (ii) of the lemma follows directly from differentiating both sides of the materials balance condition, which is also the first-order condition (A.23), with respect to each s_i, $i = 1, 2$, in turn. □

Proof of Proposition 2. The Lagrangian associated with the first-period no-commitment tax design problemwith separation is

$$\begin{aligned} &u(x_1^1 - s_1) - g\Big(\frac{y_1^1}{w_1}\Big) + v(x_1^2(\mathbf{s}) + (1+r)s_1) - h\Big(\frac{y_1^2(\mathbf{s})}{w_1}\Big) \\ &\qquad + u(x_2^1 - s_2) - g\Big(\frac{y_2^1}{w_2}\Big) + v(x_2^2(\mathbf{s}) + (1+r)s_2) - h\Big(\frac{y_2^2(\mathbf{s})}{w_2}\Big) \\ &+ \eta\Big[y_1^1 + y_2^1 - x_1^1 - x_2^1 - s_G\Big] + \psi\Big[u(x_2^1 - s_2) - g\Big(\frac{y_2^1}{w_2}\Big) - u(x_1^1 - s_1) + g\Big(\frac{y_1^1}{w_2}\Big)\Big]. \end{aligned} \tag{A.26}$$

The associated first-order equations include:

$$x_1^1\colon\ u'(c_1^1) - \eta - \psi u'(c_1^1) = 0; \tag{A.27}$$

$$x_2^1\colon\ u'(c_2^1) - \eta + \psi u'(c_2^1) = 0; \tag{A.28}$$

$$y_1^1\colon\ -\frac{1}{w_1}g'\Big(\frac{y_1^1}{w_1}\Big) + \eta + \frac{\psi}{w_2}g'\Big(\frac{y_1^1}{w_2}\Big) = 0; \tag{A.29}$$

$$y_2^1\colon\ -\frac{1}{w_2}g'\Big(\frac{y_2^1}{w_2}\Big) + \eta - \frac{\psi}{w_2}g'\Big(\frac{y_1^2}{w_2}\Big) = 0; \tag{A.30}$$

$$\begin{aligned} s_1\colon\ &-u'(c_1^1) + v'(c_1^2)\Big[\frac{\partial x_1^2}{\partial s_1} + (1+r)\Big] + v'(c_2^2)\frac{\partial x_2^2}{\partial s_1} - \frac{1}{w_1}h'\Big(\frac{y_1^2}{w_1}\Big)\frac{\partial y_1^2}{\partial s_1} \\ &- \frac{1}{w_2}h'\Big(\frac{y_2^2}{w_2}\Big)\frac{\partial y_2^2}{\partial s_1} + \psi u'(c_1^1) = 0; \end{aligned} \tag{A.31}$$

$$
\begin{aligned}
s_2\colon\quad & -u'(c_2^1) + v'(c_1^2)\frac{\partial x_1^2}{\partial s_2} + v'(c_2^2)\Big[\frac{\partial x_2^2}{\partial s_2} + (1+r)\Big] - \frac{1}{w_1}h'\Big(\frac{y_1^2}{w_1}\Big)\frac{\partial y_1^2}{\partial s_2} \\
& -\frac{1}{w_2}h'\Big(\frac{y_2^2}{w_2}\Big)\frac{\partial y_2^2}{\partial s_2} - \psi u'(c_2^1) = 0.
\end{aligned} \tag{A.32}
$$

Equations (A.27)–(A.30) are identical to equations (A.2)–(A.5), except that λ^1 is replaced by η and μ is replaced by ψ. Thus, the arguments used in the proof of Proposition 1 may be repeated to prove part (i) of the proposition.

By part (i) of Lemma 1, (A.31) is equivalent to

$$
-u'(c_1^1) + (1+r)v'(c_1^2) + \psi u'(c_1^1) + v'(c_1^2)\Big[\frac{\partial x_1^2}{\partial s_1} + \frac{\partial x_2^2}{\partial s_1} - \frac{\partial y_1^2}{\partial s_1} - \frac{\partial y_2^2}{\partial s_1}\Big] = 0. \tag{A.33}
$$

By part (ii) of Lemma 1, the term in square brackets on the left-hand side of (A.33) is zero, so that

$$
-u'(c_1^1) + (1+r)v'(c_1^2) + \psi u'(c_1^1) = 0, \tag{A.34}
$$

which is equivalent to

$$
(1-\psi)u'(c_1^1) = (1+r)v'(c_1^2). \tag{A.35}
$$

Because both $u(\cdot)$ and $v(\cdot)$ are increasing, it follows from (A.35) that $(1-\psi) > 0$. Rearranging (A.35) yields

$$
\frac{u'(c_1^1)}{v'(c_1^2)} = \frac{(1+r)}{(1-\psi)} > (1+r), \tag{A.36}
$$

from which the first inequality of part (ii) of the proposition follows.

A similar argument may be used to show that (A.32) is equivalent to

$$
-u'(c_2^1) + (1+r)v'(c_2^2) - \psi u'(c_2^1) = 0. \tag{A.37}
$$

Rearranging (A.37) yields

$$
\frac{u'(c_2^1)}{v'(c_2^2)} = \frac{(1+r)}{(1+\psi)} < (1+r), \tag{A.38}
$$

from which the second inequality of part (ii) of the proposition follows. □

References

Berliant, M., Ledyard, J. O., 2005. Optimal dynamic nonlinear income taxes with no commitment, downloadable from the Working Paper Archive at Washington University at St. Louis, `http://econpapers.repec.org/paper/wpawuwppe/0403004.htm`.

Bisin, A., Rampini, A. A., 2006. Markets as beneficial constraints on government. Journal of Public Economics, in press.

Boadway, R., Keen, M., 2000. Redistribution. In: Atkinson, A. B., Bourguignon, F. (Eds.), Handbook of Income Distribution. Vol. 1. North-Holland, Amsterdam, pp. 677–789.

Boadway, R., Marceau, N., Marchand, M., 1996. Investment in education and the time inconsistency of redistributive tax policy. Economica 63, 171–189.

Boadway, R., Pestieau, P., 2004. Tagging and redistributive taxation, unpublished manuscript, Department of Economics, Queens University and Centre de Recherche en Economie Publique et de la Population, Université de Liège.

Boadway, R., Wildasin, D., 1994. Taxation and savings: A survey. Fiscal Studies 15, 19–63.

Brett, C., Weymark, J. A., 2004. Public good provision and the comparative statics of optimal nonlinear income taxation. Working Paper No. 04-w15, Vanderbilt University.

Brito, D. L., Hamilton, J. A., Slutsky, S. M., Stiglitz, J. E., 1990. Pareto efficient tax structures. Oxford Economic Papers 42, 61–77.

Browning, M., Burbidge, J., 1990. Consumption and income taxation. Oxford Economic Papers 42, 281–292.

Dillén, M., Lundholm, M., 1996. Dynamic income taxation, redistribution, and the ratchet effect. Journal of Public Economics 59, 69–93.

Fischer, S., 1980. Dynamic inconsistency, cooperation and the benevolent dissembling government. Journal of Economic Dynamics and Control 2, 93–107.

Freixas, X., Guesnerie, R., Tirole, J., 1985. Planning under incomplete information and the ratchet effect. Review of Economic Studies 52, 173–191.

Guesnerie, R., 1995. A Contribution to the Pure Theory of Taxation. Cambridge University Press, Cambridge.

Guesnerie, R., Seade, J., 1983. Nonlinear pricing in a finite economy. Journal of Public Economics 17, 157–179.

Hamilton, J., Pestieau, P., 2005. Optimal income taxation and the ability distribution: Implications for migration equilibria. International Tax and Public Finance 12, 29–45.

Mirrlees, J. A., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175–208.

Ordover, J. A., Phelps, E. S., 1979. The concept of optimal taxation in the overlapping–generations model of capital and wealth. Journal of Public Economics 12, 1–26.

Pirttilä, J., Tuomala, M., 2001. On optimal non-linear taxation and public good provision in an overlapping generations economy. Journal of Public Economics 79, 485–501.

Roberts, K., 1984. The theoretical limits to redistribution. Review of Economic Studies 51, 177–195.

Seade, J., 1977. On the shape of optimal tax schedules. Journal of Public Economics 7, 203–236.

Stiglitz, J. E., 1983. Self-selection and Pareto efficient taxation. Journal of Public Economics 17, 213–240.

Stiglitz, J. E., 1987. Pareto efficient and optimal taxation in the new new welfare economics. In: Auerbach, A. J., Feldstein, M. (Eds.), Handbook of Public Economics. Vol. 2. North-Holland, Amsterdam, pp. 991–1042.

Sundaram, R. K., 1996. A First Course in Optimization Theory. Cambridge University Press, Cambridge.

Tuomala, M., 1990. Optimal Income Tax and Redistribution. Clarendon Press, Oxford.

Weitzman, M. L., 1980. The ratchet principle and performance incentives. Bell Journal of Economics 11, 302–308.

Weymark, J. A., 1986. Bunching properties of optimal nonlinear income taxes. Social Choice and Welfare 3, 213–232.

Weymark, J. A., 1987. Comparative static properties of optimal nonlinear income taxes. Econometrica 55, 1165–1185.