

Optimal Nonlinear Taxation of Income and Savings without Commitment*

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)*

Current Version: January 2008

*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 benefited from a conversation with Marcus Berliant about the subject of this article and from correspondence with Francesco Cohen, Stéphane Gauthier, and Jonathan Hamilton. This article was previously circulated as "Optimal Nonlinear Taxation of Income and Savings in a Two Class Economy."

Abstract

“Optimal Nonlinear Taxation of Income and Savings without Commitment”

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 have additively separable preferences and 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, it is optimal to tax the savings of the high-skilled individual and to tax the savings of the low-skilled individual at a lower, possibly negative, rate. If the types are pooled in period one, it is optimal for the low-skilled (high-skilled) individual to face a marginal savings tax (subsidy). In both cases, the savings of the high-skilled individual are distorted because this individual rationally expects that some of his savings will be redistributed to the low-skilled individual in the second period. The savings of the low-skilled individual in the separating case are taxed at a lower rate so as to relax an incentive compatibility constraint.

Journal of Economic Literature classification numbers: D82, H21.

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

1. Introduction

Ever since the pathbreaking 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 schedules. 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.¹ It 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, an individual's unobservable skill level is a determinant of his 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 (1982), Stiglitz (1982), 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. For simplicity, we assume that there is only one person of each type. A utilitarian government designs an optimal nonlinear tax system for these individuals in each time period. It can condition tax payments on both labor income and 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

¹The only role of the government in this article is to redistribute income, so we speak interchangeably of the "government" and the "taxation authority."

in advance, and so any information about individual 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 kinds of optimal tax regimes. In a separating tax regime, type information is revealed in the first period, whereas in a pooling tax regime, type information remains hidden after the first period. Whether the optimal regime is separating or pooling depends on a discrete comparison between the best separating tax policy and the best pooling tax policy. 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 compatibility 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 tax the savings of the high-skilled individual. In addition, the less productive individual faces a lower marginal tax rate on savings than does the more productive individual, and may even receive a marginal savings subsidy. 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 face a positive marginal rate of savings taxation, while the savings of the high-skilled individual is subsidized. 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. In contrast, the recent literature on dynamic optimal taxation has supposed that individuals are heterogeneous and are subject 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.² 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. Their main finding is

²For overviews of this literature, see Kocherlakota (2006) and Golosov, Tsyvinski, and Werning (2007). For an illuminating perspective on this literature, see Diamond (2007).

that the power of the taxation authority to take advantage of information revealed in 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.

We do not consider shocks to the economy. In particular, we suppose that each individual has the same labor productivity in every period. Roberts (1984), Berliant and Ledyard (2005), and Apps and Rees (2006) have also employed a deterministic framework to analyze dynamic optimal nonlinear income taxation when the government cannot commit to ignore information gathered in earlier periods. None of these studies considers savings taxation as a possible instrument.³ 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.⁴

The work closest to our own is that of Apps and Rees (2006).⁵ As is the case here, Apps and Rees consider a two-type, two-period model. Their model differs from ours in two important respects. First, they have a continuum of individuals of each type. Second, they do not permit any intertemporal transfers of resources and, therefore, there are no savings to tax. By allowing for more than one person of each type, Apps and Rees are also able to consider a mixed case in which some, but not all, high-skilled individuals are pooled with the low-skilled individuals in period one. Working with a continuum of individuals simplifies their analysis when there is some separation in the first period because when a high-skilled individual mimics a low-skilled individual, his behavior does not change the distribution of types and, hence, does not change which regime applies (pooling, separating, or mixed). Because we only have one person of each type, with this kind of mimicking the regime switches from separating to pooling, and this must be taken into account when formulating this person's incentive compatibility constraint. In spite of these differences, the qualitative properties of the tax distortions on labor earnings that Apps and Rees identify are the same as those obtained here.

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

³There is also an extensive literature on dynamic optimal taxation with fixed types in which the government is assumed to be able to commit to its tax policy in period one. In an interesting contribution to this literature, Gaube (2007) argues that basing taxes only on current values of the variables in the tax base, in his case incomes, can serve as a partial commitment device.

⁴Dillén and Lundholm (1996) also analyze 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.

⁵The first version of our article was completed before we learned of their research.

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 (1979) show that it is optimal to tax the savings of most individuals (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, Marceau, and Marchand (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 observationally 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 publicly-observable education) that can help to mitigate the distortions introduced because of the government's lack of commitment.⁶

The rest of this article is organized as follows. Section 2 describes the economy. In order to provide a benchmark for our analysis of optimal taxation without commitment, in Section 3, we identify the qualitative properties of optimal taxes 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 supplying l_i^t units of labor and consuming c_i^t units of a single consumption good in period t , $t = 1, 2$.⁷ Individual i transfers wealth from period one to period two 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

⁶The 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.

⁷Subscripts index individuals, while superscripts index time periods.

the literature on optimal nonlinear income taxation, the skill level is interpreted as labor productivity, 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 Boddway, Marceau, and Marchand (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 his labor income.

The government designs a tax system to redistribute income between the individuals. It cannot observe an individual's labor supply or skill level, but it can observe his income from labor and knows that there is one person with each of the two possible skill levels. Moreover, it has the ability to observe savings. The total tax paid by consumer i in period t is T_i^t . The value of T_i^t can be made contingent on the amount saved and on the current and past values of labor income. In period t , the difference between labor income and the tax paid by consumer i is

$$x_i^t = y_i^t - T_i^t, \quad i, t = 1, 2. \quad (2.1)$$

Individuals are free to divide this net-of-tax labor income in the first period between consumption and savings. Each unit of savings provides an individual with an additional $1+r$ units of consumption in the second period over and above his second period net-of-tax labor income. Consumption in each period is, therefore, given by

$$c_i^1 = x_i^1 - s_i, \quad c_i^2 = x_i^2 + (1+r)s_i, \quad i = 1, 2. \quad (2.2)$$

Although strictly speaking, in period two, person i 's before-tax income is $y_i^2 + rs_i$ (the sum of labor and interest income) and his after-tax income is $y_i^2 + rs_i - T_i^2$, we shall abuse terminology somewhat and henceforth refer to y_i^t and x_i^t as, respectively, person i 's before-tax and after-tax incomes in period t .

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

$$U(c_i^1, l_i^1, c_i^2, l_i^2) = u(c_i^1) - g(l_i^1) + v(c_i^2) - h(l_i^2), \quad i = 1, 2. \quad (2.3)$$

The functional form of $U(\cdot)$ is known to the government. 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

$$u(x_i^1 - s_i) - g\left(\frac{y_i^1}{w_i}\right) + v(x_i^2 + (1+r)s_i) - h\left(\frac{y_i^2}{w_i}\right), \quad i = 1, 2. \quad (2.4)$$

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

$$\text{MRS}_{y_i^1, x_i^1} = \frac{g'\left(\frac{y_i^1}{w_i}\right)}{w_i u'(c_i^1)}, \quad (2.5)$$

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

$$\text{MRS}_{y_i^2, x_i^2} = \frac{h' \left(\frac{y_i^2}{w_i} \right)}{w_i v'(c_i^2)}. \quad (2.6)$$

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 individual must work fewer additional hours for each additional unit of before-tax income than does the lower-skilled individual. Thus, it takes a smaller increase in after-tax income to compensate the higher-skilled individual for increases in before-tax income than it does to compensate the lower-skilled individual.

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

$$\text{MRS}_{x_i^1, x_i^2} = -\frac{u'(c_i^1)}{v'(c_i^2)}. \quad (2.7)$$

This intertemporal marginal rate of substitution does not depend explicitly upon the skill level and, hence, is observable to the government. Because of their common preferences over consumption and labor supply, the two individuals 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

$$x_1^1 + x_2^1 + s_G \leq y_1^1 + y_2^1 \quad (2.8)$$

and

$$x_1^2 + x_2^2 \leq y_1^2 + y_2^2 + (1+r)s_G. \quad (2.9)$$

Using (2.2), these constraints can be equivalently written as

$$c_1^1 + c_2^1 + s_1 + s_2 + s_G \leq y_1^1 + y_2^1 \quad (2.10)$$

and

$$c_1^2 + c_2^2 \leq y_1^2 + y_2^2 + (1+r)(s_1 + s_2 + s_G). \quad (2.11)$$

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

$$\begin{aligned} \mathcal{W}(x_1^1, x_2^1, y_1^1, y_2^1, x_1^2, x_2^2, y_1^2, y_2^2, s_1, s_2) = & u(x_1^1 - s_1) - g \left(\frac{y_1^1}{w_1} \right) + v(x_1^2 + (1+r)s_1) \\ & - h \left(\frac{y_1^2}{w_1} \right) + u(x_2^1 - s_2) - g \left(\frac{y_2^1}{w_2} \right) + v(x_2^2 + (1+r)s_2) - h \left(\frac{y_2^2}{w_2} \right). \end{aligned} \quad (2.12)$$

Note that with this objective function, the government shares the intertemporal preference of the individuals. 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.⁸ In order to focus on the redistributive 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 individuals. Thus, only anonymous tax schedules are feasible.⁹ Because individuals are free to select their optimal work-consumption-savings combinations from the anonymous schedule that determines their tax payments in each period, the resulting allocation must be incentive compatible; that is, each individual must weakly prefer the allocation designed for him to the allocation 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 policy in period one. Specifically, the government is able to credibly commit not to use information about the skill levels of the individuals revealed in the first period to adjust taxes in the second period.

Given our information assumptions, we can equivalently think of the government as choosing the allocation for this economy directly subject to incentive compatibility and material balance constraints, rather than interacting indirectly with the individuals through a tax schedule. When there is full commitment, incentive compatibility requires that an individual weakly prefers the entire allocation, over both time periods, designed for him to the allocation designed for the other individual. Formally,

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

⁸Browning and Burbidge only consider linear taxation, but their point also applies when taxes are nonlinear.

⁹Because the economy has a finite number of individuals, the taxation authority can, by observing aggregate outcomes, tell when an individual misreports his true type. Piketty (1993) and Hamilton and Slutsky (2007) show how this knowledge can be used to implement first-best taxes in atemporal settings. To achieve the first-best, the taxation authority must condition an individual's tax payments on the behavior of other taxpayers. We follow the tradition of the optimal tax literature by requiring the government to set a tax schedule specifying an individual's taxes as a function of only the values of his own incomes and savings. This is a restriction, but a very natural one.

and

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

We assume that only the incentive compatibility condition (3.1) might potentially bind. Given its utilitarian objective, the government wants to redistribute income from the more highly-skilled individual to the less highly-skilled individual. The natural limit on this redistribution is that, if taken too far, such redistribution might induce the higher-skilled individual to pretend to be the lower-skilled individual. Imposing (3.1) prevents this type of mimicking.¹⁰ 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 (2.12) subject to the materials balance constraints (2.8) and (2.9) and the two-period incentive compatibility constraint (3.1).¹¹

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.¹² 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) $\text{MRS}_{y_2^1, x_2^1} = 1$, $\text{MRS}_{y_2^2, x_2^2} = 1$, and $\text{MRS}_{x_2^1, x_2^2} = -(1+r)$.
- (ii) $\text{MRS}_{y_1^1, x_1^1} < 1$, $\text{MRS}_{y_1^2, x_1^2} < 1$, and $\text{MRS}_{x_1^1, x_1^2} = -(1+r)$.

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, Hamilton, Slutsky,

¹⁰Indeed, at a solution to the first-best taxation problem for this economy, the government wants 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.

¹¹In 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.

¹²Separation is possible in two-good worlds when there is asymmetric information in one dimension.

and Stiglitz (1990, Proposition 5) that the marginal rate of substitution for person 1 is distorted only for those pairs of goods for which the two individuals 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.¹³ On the other hand, the two individuals have the same intertemporal preferences. In particular, person 2 considering the opportunity to mimic person 1 is willing to trade consumption across time at the same implicit prices as is person 1. 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 tax schedule is not credible. The optimal two-period tax policy 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 has 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 redistributive goals. Hence, the optimal scheme with commitment is time inconsistent.

The two individuals 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 his type is revealed in the first period, then the taxation authority will have an easier time redistributing income from him to the lower-skilled individual in the second period because it no longer needs to worry about incentive compatibility constraints. This redistribution can be accomplished by transferring 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

¹³Optimal income tax schedules may be nondifferentiable. Person i 's implicit marginal income tax rate in period t is $1 - \text{MRS}_{y_i^t, x_i^t}$. Because $\text{MRS}_{y_i^t, x_i^t} > 0$, marginal income tax rates are bounded above by one.

modified in order to induce information revelation in the first period. As pointed out by Freixas, Guesnerie, and Tirole (1985) in a 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; i.e., there is pooling. If any component of the first period allocation differs by type, information is revealed and the types are separated. The tax schedule offered 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 the social welfare function 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 first determine the solution to the optimal tax problem conditional on there being (a) pooling and (b) separation in period one.

4.1. 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

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

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

$$v(x_2^2 + (1+r)s) - h\left(\frac{y_2^2}{w_2}\right) \geq v(x_1^2 + (1+r)s) - h\left(\frac{y_1^2}{w_1}\right). \quad (4.2)$$

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

The Second Period Tax Design Problem with Pooling. Given $\tilde{s} = (s, s_G)$, the government chooses a second period allocation $(x_1^2, x_2^2, y_1^2, y_2^2)$ to maximize the objective function (4.1) subject to the second period materials balance constraint (2.9) and the incentive compatibility constraint (4.2).

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 2 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 1.

Lemma 1. *$\text{MRS}_{y_1^2, x_1^2} < 1$ and $\text{MRS}_{y_2^2, x_2^2} = 1$ 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 solving for the optimal first period allocation with pooling. Thus, given pooling in the first period, its objective function is

$$\begin{aligned} \mathcal{W}^{pool}(x^1, y^1, s, s_G) = & u(x^1 - s) - g\left(\frac{y^1}{w_1}\right) + v(x_1^2(\tilde{\mathbf{s}}) + (1+r)s) \\ & - h\left(\frac{y_1^2(\tilde{\mathbf{s}})}{w_1}\right) + u(x^1 - s) - g\left(\frac{y^1}{w_2}\right) + v(x_2^2(\tilde{\mathbf{s}}) + (1+r)s) - h\left(\frac{y_2^2(\tilde{\mathbf{s}})}{w_2}\right), \end{aligned} \quad (4.3)$$

where variables without subscripts indicate values that are identical for the two individuals. Because the objective function (4.3) incorporates the second period decisions of the taxation authority, it takes account of the period two materials balance constraint (2.9) and the incentive compatibility constraint (4.2). Because the two individuals have identical observable allocations in period one, overall incentive compatibility is equivalent to (4.2). Hence, the only constraint remaining for the first period tax design problem is (2.8), the materials balance constraint for period one. When the individuals are pooled, (2.8) reduces to

$$2x^1 + s_G \leq 2y^1. \quad (4.4)$$

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 (4.3) subject to the first period materials balance constraint (4.4).

The objective function of the first period no-commitment tax design problem with pooling depends on first period savings in 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 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.¹⁴ The following proposition demonstrates that, even without general comparative static results concerning the second period problem, it is possible to characterize the pattern of distortions in the first period.

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

- (i) $\text{MRS}_{y_1^1, x_1^1} > 1 > \text{MRS}_{y_2^1, x_2^1}$.
- (ii) $\text{MRS}_{x_1^1, x_1^2} > -(1 + r) > \text{MRS}_{x_2^1, x_2^2}$.¹⁵

With pooling and no commitment, person 2 faces a higher implicit marginal income tax rate than person 1 in period one, which is the reverse of what occurs in the standard nonlinear income tax problem and in the full commitment problem. 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.¹⁶ Both individuals also have their savings decisions distorted, with person 1 (resp. person 2) having his savings taxed (resp. subsidized) at the margin.

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 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. The requirements that $\text{MRS}_{y_1^1, x_1^1} > \text{MRS}_{y_2^1, x_2^1}$ in part (i) of Proposition 2 and $\text{MRS}_{x_1^1, x_1^2} > \text{MRS}_{x_2^1, x_2^2}$ in part (ii) then follow from the definitions of these marginal rates of substitution in (2.5) and (2.7).

At an optimal allocation with first-period pooling, it is possible to infinitesimally decrease the common first period consumption and before-tax income by the same amount holding savings fixed without violating the materials balance constraints (2.8) and (2.9). Because savings are held constant, this change has no effect on the second period incentive compatibility constraint (4.2). If $\text{MRS}_{y_2^1, x_2^1} \geq 1$ and, hence, $\text{MRS}_{y_1^1, x_1^1} > 1$, this change is a Pareto improvement. Hence, it must be optimal to have $\text{MRS}_{y_2^1, x_2^1} < 1$.¹⁷ By reversing

¹⁴In their comparative static analyses, Weymark (1987), Hamilton and Pestieau (2005), and Brett and Weymark (2008) consider preferences that are quasilinear in labor supply, while Boadway and Pestieau (2006) and Simula (2007) 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.

¹⁵Recall 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.

¹⁶Dillén 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.

¹⁷Except in the borderline case in which $\text{MRS}_{y_2^1, x_2^1} = 1$, 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.

the direction of change in first period consumption and before-tax income, it follows that it is also optimal to have $\text{MRS}_{y_1^1, x_1^1} > 1$.

As we have seen, at an optimal allocation with first-period pooling, it is necessary for person 1 to have the smaller intertemporal marginal rate of substitution in absolute value. Suppose that $\text{MRS}_{x_1^1, x_1^2} \leq -(1+r)$ and, hence, $\text{MRS}_{x_2^1, x_2^2} < -(1+r)$. Consider modifying the optimal allocation by having each individual transfer a common infinitesimally small amount from savings into first period consumption and then decreasing second period consumption by $-\text{MRS}_{x_1^1, x_1^2}$ for person 1 and by $-\text{MRS}_{x_2^1, x_2^2}$ for person 2. This composite change has no effect on the variables that appear in the first period materials balance constraint (2.8), but it relaxes the second-period materials balance constraint (2.9); i.e., it is resource saving. Furthermore, this reallocation is a matter of indifference for each individual. From the definition of the intertemporal marginal rate of substitution (2.7), we see that second period utility has decreased by $u'(c_2^1)$ for person 2 and by $u'(c_1^1)$ for person 1. Note that $u'(c_1^1)$ is also the change in second period utility for person 2 if he mimics person 1. Because there is pooling in the first period, $c_1^1 = c_2^1$. Hence, this composite change does not violate the incentive compatibility constraint (4.2). The resource savings can now be used to increase each person's second period consumption without violating any of the constraints, contradicting the optimality of the initial allocation. Thus, it must be optimal to have $\text{MRS}_{x_1^1, x_1^2} > -(1+r)$. A similar argument can be used to show that it is also optimal to have $\text{MRS}_{x_2^1, x_2^2} < -(1+r)$, for otherwise it would be possible to obtain a Pareto improvement by transferring consumption from the first period to the second period.

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 individuals. This trade 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.

4.2. 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 after-tax 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 (2.2) and (2.4), is

$$\mathcal{W}^2(x_1^2, x_2^2, y_1^2, y_2^2, s_1, s_2) = v(x_1^2 + (1+r)s_1) - h\left(\frac{y_1^2}{w_1}\right) + v(x_2^2 + (1+r)s_2) - h\left(\frac{y_2^2}{w_2}\right). \quad (4.5)$$

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

The Second Period First-Best Problem. Given $\mathbf{s} = (s_1, s_2, s_G)$, 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 (4.5) subject to the materials balance constraint (2.9).

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. Its value function, $\mathcal{V}^{2,sep}(\mathbf{s})$, is also well-behaved. The properties of the solution to the second period first-best problem most pertinent to a characterization of the distortions arising in our two-period optimal tax problem when there is separation in the first period are collected in the following lemma.

Lemma 2. *For a given savings vector $\mathbf{s} = (s_1, s_2, s_G)$, 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) $v'(x_1^2(\mathbf{s}) + (1+r)s_1) = v'(x_2^2(\mathbf{s}) + (1+r)s_2) = \frac{1}{w_1}h'\left(\frac{y_1^2(\mathbf{s})}{w_1}\right) = \frac{1}{w_2}h'\left(\frac{y_2^2(\mathbf{s})}{w_2}\right).$
- (ii) $\frac{\partial \mathcal{V}^{2,sep}(\mathbf{s})}{\partial s_i} = (1+r)(v'(x_i^2(\mathbf{s}) + (1+r)s_i)), \quad i = 1, 2.$
- (iii) $\frac{\partial x_1^2(\mathbf{s})}{\partial s_i} + \frac{\partial x_2^2(\mathbf{s})}{\partial s_i} - \frac{\partial y_1^2(\mathbf{s})}{\partial s_i} - \frac{\partial y_2^2(\mathbf{s})}{\partial s_i} = 0, \quad i = 1, 2.$
- (iv) $\frac{\partial y_2^2(\mathbf{s})}{\partial s_2} - \frac{\partial x_2^2(\mathbf{s})}{\partial s_2} > 0.$

With separation, we have a full information planning problem in the second period in which the government has access to the interest-augmented savings from the first period to distribute as it wishes. Part (i) of Lemma 2 summarizes the marginal conditions for a first-best utilitarian optimum in the second period. The taxation authority wants 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. This common marginal utility of consumption is the rate at which social welfare increases in the second period if there is a marginal increase in either individual's consumption in this period. Therefore, as stated in part (ii), the marginal value of savings, both private and public, is $1 + r$ times the marginal utility of consumption in the second period. Furthermore, for each individual, the marginal rate of substitution between labor and consumption equals his wage rate (i.e., his skill level). Because person 2 has a higher wage rate, he 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 (iii) of Lemma 2 follows directly from the second period materials balance constraint. This result does not imply 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 (iii) simply says that changes in aggregate production are offset by changes in after-tax income. However, as is apparent from (2.2), 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.

Part (iv) of Lemma 2 states that the total tax payment of person 2 is increasing in his savings. Thus, time inconsistency is a real concern in this economy, rather than a mere possibility. If person 2 does reveal his type and does save for the future, then the taxation authority will take some of those resources away through taxation in the second period. Given the purely redistributive nature of taxation, these resources are then given to person 1.

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 the 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 his type in the first period. In order to do this, the total utility experienced by person 2 across the two periods when there is separation in the first period must exceed the total utility experienced by person 2 when person 2 pretends to be person 1 in the first period. The potential gains to mimicking in the first period are twofold. First, person 2 could directly benefit from a favorable first period allocation offered to person 1. Second, mimicking in the first period is a way for person 2 to keep his true type hidden until the second period. If person 2 does mimic person 1 in the first period, then there is pooling. As a consequence, person 2 rationally anticipates that the government will solve the second period tax design problem with pooling with s_1 and s_G as the levels of private and public savings, respectively. In order

to provide person 2 with an incentive to report his true type, it must therefore be the case that the following incentive compatibility constraint is satisfied:

$$\begin{aligned} u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) + v(x_2^2(\mathbf{s}) + (1+r)s_2) - h\left(\frac{y_2^2(\mathbf{s})}{w_2}\right) \\ \geq u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_2}\right) + v(\hat{x}_2^2(s_1, s_G) + (1+r)s_1) - h\left(\frac{\hat{y}_2^2(s_1, s_G)}{w_2}\right), \end{aligned} \quad (4.6)$$

where $(\hat{x}_2^2(\cdot), \hat{y}_2^2(\cdot))$ is the consumption-income pair received by person 2 in the second period when there is pooling in the first period.

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

$$\begin{aligned} \mathcal{W}^{sep}(x_1^1, x_2^1, y_1^1, y_2^1, s_1, s_2, s_G) \\ = u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_1}\right) + u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) + \mathcal{V}^{2,sep}(s_1, s_2, s_G). \end{aligned} \quad (4.7)$$

Because both the incentive compatibility condition (4.6) and the objective function (4.7) 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 also 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 (4.7) subject to the first period materials balance constraint (2.8) and the incentive compatibility constraint (4.6).

The pattern 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 3. *At a solution to the first period no-commitment tax design problem with separation:*

- (i) $\text{MRS}_{y_1^1, x_1^1} < 1$ and $\text{MRS}_{y_2^1, x_2^1} = 1$.
- (ii) $\text{MRS}_{x_1^1, x_2^1} < \text{MRS}_{x_2^1, x_2^2}$ and $\text{MRS}_{x_2^1, x_2^2} > -(1+r)$.

Part (i) of Proposition 3 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 tax policy shares the qualitative properties of an optimal nonlinear income tax schedule 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. Hence, person 1 has less income before and after tax than person 2. 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 3 shows that, in contrast with many atemporal optimal income tax schedules, the government's lack of commitment to a second period tax scheme results in a distortion at the top. Indeed, unlike the general atemporal optimal nonlinear tax problem described in Brito, Hamilton, Slutsky, and Stiglitz (1990, Proposition 3), Proposition 3 shows that every person faces some marginal distortion in the optimal two-period allocation when the type information is revealed in the first period.

The optimality of an implicit tax on the savings of individual 2 is a result of the time-consistency problem faced by the government. We know from part (iv) of Lemma 2 that if person 2 increases his savings, then he will face a higher tax payment in the second period. He anticipates this outcome and so, in the absence of any countervailing incentive, he rationally distorts his savings downward when he reveals his type. Even though the tax authority wants to encourage person 2 to save so as to use the proceeds to redistribute income in the second period, it accepts the downward distortion of savings as a cost of extracting type information in the first period.

Because the second period allocation is first best, it is optimal to equalize $v'(c_1^2)$ and $v'(c_2^2)$. Thus, person 1 is provided with a smaller marginal tax on savings than person 2 if (and only if) $u'(c_1^1) > u'(c_2^1)$. If this were not the case, then a marginal transfer of consumption in period one from person 1 to person 2 would increase (or at least not decrease) the sum of utilities without violating the materials balance constraint. Furthermore, this transfer would relax the incentive compatibility constraint because person one's allocation is no longer as attractive to person 2 as it was. Because this incentive constraint no longer binds, it is possible to make additional adjustments to the individual allocations so as to further increase social welfare without violating any of the constraints. Thus, providing person 1 with the larger savings subsidy is an effective way for the taxation authority to screen the two individuals. We therefore 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 constraint.

5. Conclusion

It is sometimes argued that conditioning taxes on both current and past values of the variables in the tax base, as is done here, has no real-world counterpart. As Kocherlakota (2006, p. 269) notes, citing features of the U.S. tax code, this is not in fact the case. More importantly, he argues that (see p. 296) such criticisms miss the point of a normative analysis whose objective is to identify better tax policies than have been employed in the past.

Our analysis suggests that the lack of government commitment in the design of an intertemporal tax policy provides a rationale for distortions in savings behavior. Extending our analysis to many-person 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. Nevertheless, the fundamental insight of this article—that when a government cannot commit to its tax policy, some form of savings distortions are necessary to counteract the incentive the government has to exploit in subsequent periods any type information revealed to it—is likely to carry over to economies with any number of individuals.

Appendix

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

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

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

$$x_1^1: u'(c_1^1) - \lambda^1 - \mu u'(c_1^1) = 0; \quad (\text{A.2})$$

$$x_2^1: u'(c_2^1) - \lambda^1 + \mu u'(c_2^1) = 0; \quad (\text{A.3})$$

$$y_1^1: -\frac{1}{w_1} g'\left(\frac{y_1^1}{w_1}\right) + \lambda^1 + \frac{\mu}{w_2} g'\left(\frac{y_1^1}{w_2}\right) = 0; \quad (\text{A.4})$$

$$y_2^1: -\frac{1}{w_2}g'\left(\frac{y_2^1}{w_2}\right) + \lambda^1 - \frac{\mu}{w_2}g'\left(\frac{y_2^1}{w_2}\right) = 0; \quad (\text{A.5})$$

$$x_1^2: v'(c_1^2) - \lambda^2 - \mu v'(c_1^2) = 0; \quad (\text{A.6})$$

$$x_2^2: v'(c_2^2) - \lambda^2 + \mu v'(c_2^2) = 0; \quad (\text{A.7})$$

$$y_1^2: -\frac{1}{w_1}h'\left(\frac{y_1^2}{w_1}\right) + \lambda^2 + \frac{\mu}{w_2}h'\left(\frac{y_1^2}{w_2}\right) = 0; \quad (\text{A.8})$$

$$y_2^2: -\frac{1}{w_2}h'\left(\frac{y_2^2}{w_2}\right) + \lambda^2 - \frac{\mu}{w_2}h'\left(\frac{y_2^2}{w_2}\right) = 0; \quad (\text{A.9})$$

$$s_1: -u'(c_1^1) + (1+r)v'(c_1^2) + \mu u'(c_1^1) - (1+r)\mu v'(c_1^2) = 0; \quad (\text{A.10})$$

$$s_2: -u'(c_2^1) + (1+r)v'(c_2^2) - \mu u'(c_2^1) + (1+r)\mu v'(c_2^2) = 0; \quad (\text{A.11})$$

$$s_G: -\lambda^1 + (1+r)\lambda^2 = 0. \quad (\text{A.12})$$

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),

$$(1-\mu)u'(c_1^1) = \frac{1}{w_1}g'\left(\frac{y_1^1}{w_1}\right) - \frac{\mu}{w_2}g'\left(\frac{y_1^1}{w_2}\right) = \lambda^1. \quad (\text{A.14})$$

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

$$\frac{1}{w_1}g'\left(\frac{y_1^1}{w_1}\right) - \frac{\mu}{w_2}g'\left(\frac{y_1^1}{w_2}\right) > \frac{(1-\mu)}{w_1}g'\left(\frac{y_1^1}{w_1}\right). \quad (\text{A.15})$$

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

$$(1-\mu)u'(c_1^1) > \frac{(1-\mu)}{w_1}g'\left(\frac{y_1^1}{w_1}\right). \quad (\text{A.16})$$

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),

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

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

Proof of Proposition 2. The Lagrangian associated with the second period no-commitment tax design problem with pooling is

$$\begin{aligned} & v(x_1^2 + (1+r)s) - h\left(\frac{y_1^2}{w_1}\right) + v(x_2^2 + (1+r)s) - h\left(\frac{y_2^2}{w_2}\right) + \zeta[y_1^2 + y_2^2 - x_1^2 - x_2^2] \\ & + \phi\left[v(x_2^2 + (1+r)s) - h\left(\frac{y_2^2}{w_2}\right) - v(x_1^2 + (1+r)s) + h\left(\frac{y_1^2}{w_1}\right)\right]. \end{aligned} \quad (\text{A.18})$$

Let $\mathcal{V}^{2,pool}(s, s_G)$ be the value function for the second period problem with pooling. The associated first-order conditions include:

$$x_1^2: v'(c_1^2) - \zeta - \phi v'(c_1^2) = 0; \quad (\text{A.19})$$

$$x_2^2: v'(c_2^2) - \zeta + \phi v'(c_2^2) = 0. \quad (\text{A.20})$$

Rearranging (A.19) and (A.20) yields:

$$(1 - \phi)v'(c_1^2) = (1 + \phi)v'(c_2^2) = \zeta. \quad (\text{A.21})$$

By the envelope theorem,

$$\frac{\partial \mathcal{V}^{2,pool}}{\partial s} = (1+r)v'(c_1^2) + (1+r)v'(c_2^2) + \phi(1+r)v'(c_2^2) - \phi(1+r)v'(c_1^2). \quad (\text{A.22})$$

Equations (A.21) and (A.22) imply

$$\frac{\partial \mathcal{V}^{2,pool}}{\partial s} = 2(1+r)\zeta \quad (\text{A.23})$$

By (4.3) and (4.4), the Lagrangian associated with first period non-commitment problem with pooling can be written as

$$u(x^1 - s) - g\left(\frac{y^1}{w_1}\right) + u(x^1 - s) - g\left(\frac{y^1}{w_2}\right) + \mathcal{V}^{2,pool}(s, s_g) + \xi[2y^1 - 2x^1 - s_G]. \quad (\text{A.24})$$

The first-order condition associated with s is

$$-2u'(c^1) + \frac{\partial \mathcal{V}^{2,pool}}{\partial s} = 0. \quad (\text{A.25})$$

By (A.23) and (A.25),

$$u'(c^1) = (1+r)\zeta. \quad (\text{A.26})$$

Equations (A.21) and (A.26) imply:

$$-\frac{u'(c^1)}{v'(c_1^2)} = -(1-\phi)(1+r) > -(1+r), \quad (\text{A.27})$$

and

$$-\frac{u'(c^1)}{v'(c_2^2)} = -(1 + \phi)(1 + r) < -(1 + r), \quad (\text{A.28})$$

which establishes part (ii).

The first-order conditions of (A.24) with respect to x^1 and y^1 are

$$2u'(c^1) - 2\xi = 0 \quad (\text{A.29})$$

and

$$-\frac{1}{w_1}g'\left(\frac{y^1}{w_1}\right) - \frac{1}{w_2}g'\left(\frac{y^1}{w_2}\right) + 2\xi = 0. \quad (\text{A.30})$$

It follows from (2.5), (A.29), and (A.30) that

$$\text{MRS}_{y_1^1, x_1^1} + \text{MRS}_{y_2^1, x_2^1} = 2. \quad (\text{A.31})$$

Because both of these marginal rates of substitution are positive and that of person 1 is larger than that of person 2, it must be the case that $\text{MRS}_{y_1^1, x_1^1} > 1$ and $\text{MRS}_{y_2^1, x_2^1} < 1$, which establishes part (i). \square

Proof of Lemma 2. 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 this optimization problem is

$$\begin{aligned} v(x_1^2 + (1 + r)s_1) - h\left(\frac{y_1^2}{w_1}\right) + v(x_2^2 + (1 + r)s_2) - h\left(\frac{y_2^2}{w_2}\right) \\ + \lambda [y_1^2 + y_2^2 + (1 + r)s_G - x_1^2 - x_2^2]. \end{aligned} \quad (\text{A.32})$$

The first-order conditions for an optimum are:

$$x_1^2: v'(c_1^2) - \lambda = 0; \quad (\text{A.33})$$

$$x_2^2: v'(c_2^2) - \lambda = 0; \quad (\text{A.34})$$

$$y_1^2: -\frac{1}{w_1}h'\left(\frac{y_1^2}{w_1}\right) + \lambda = 0; \quad (\text{A.35})$$

$$y_2^2: -\frac{1}{w_2}h'\left(\frac{y_2^2}{w_2}\right) + \lambda = 0; \quad (\text{A.36})$$

$$\lambda: y_1^2 + y_2^2 + (1 + r)s_G - x_1^2 - x_2^2 = 0. \quad (\text{A.37})$$

Part (i) of the lemma follows from solving each of (A.33)–(A.36) for λ .

The bordered Hessian matrix for this problem is

$$A = \begin{bmatrix} v''(c_1^2) & 0 & 0 & 0 & -1 \\ 0 & v''(c_2^2) & 0 & 0 & -1 \\ 0 & 0 & -\frac{h''(l_1^2)}{(w_1)^2} & 0 & 1 \\ 0 & 0 & 0 & -\frac{h''(l_2^2)}{(w_2)^2} & 1 \\ -1 & -1 & 1 & 1 & 0 \end{bmatrix}. \quad (\text{A.38})$$

Its determinant is

$$|A| = v''(c_1^2)v''(c_2^2) \left[\frac{h''(l_1^2)}{(w_1)^2} + \frac{h''(l_2^2)}{(w_2)^2} \right] - \frac{h''(l_1^2)}{(w_1)^2} \frac{h''(l_2^2)}{(w_2)^2} [v''(c_1^2) + v''(c_2^2)]. \quad (\text{A.39})$$

By the strict concavity of $v(\cdot)$, the first two factors of the first term on the right-hand side of (A.39) are negative. Strict convexity of $h(\cdot)$ implies that the term inside the first square bracket in (A.39) 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.39) 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 (iii) of the lemma follows directly from differentiating both sides of the materials balance condition, which is also the first-order condition (A.37), with respect to each s_i , $i = 1, 2$, in turn. Part (ii) of the lemma follows from the envelope theorem.

After totally differentiating the first-order conditions, a routine application of Cramer's rule to the resulting linear system yields the following results:

$$|A| \frac{\partial x_2^2}{\partial s_2} = -(1+r)v''(c_2^2) \left\{ v''(c_1^2) \left[\frac{h''(l_1^2)}{(w_1)^2} + \frac{h''(l_2^2)}{(w_2)^2} \right] - \frac{h''(l_1^2)h''(l_2^2)}{(w_1)^2(w_2)^2} \right\}; \quad (\text{A.40})$$

$$|A| \frac{\partial y_2^2}{\partial s_2} = -(1+r)v''(c_2^2)v''(c_1^2) \frac{h''(l_1^2)}{(w_1)^2}. \quad (\text{A.41})$$

Subtracting (A.40) from (A.41) and rearranging yields

$$|A| \left[\frac{\partial y_2^2}{\partial s_2} - \frac{\partial x_2^2}{\partial s_2} \right] = (1+r)v''(c_2^2) \frac{h''(l_2^2)}{(w_2)^2} \left[v''(c_1^2) - \frac{h''(l_1^2)}{(w_1)^2} \right]. \quad (\text{A.42})$$

Because $v(\cdot)$ is strictly concave and $h(\cdot)$ is strictly convex, the second and fourth factors on the right-hand side of (A.42) are negative, while the first and third factors are positive. Thus, the entire right-hand side of (A.42) is positive. Because $|A| > 0$, part (iv) of the lemma follows. \square

Proof of Proposition 3. The Lagrangian associated with the first period no-commitment tax design problem with separation is

$$\begin{aligned} & u(x_1^1 - s_1) - g\left(\frac{y_1^1}{w_1}\right) + u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) + \mathcal{V}^{2,sep}(s_1, s_2, s_G) \\ & + \eta [y_1^1 + y_2^1 - x_1^1 - x_2^1 - s_G] + \psi \left[u(x_2^1 - s_2) - g\left(\frac{y_2^1}{w_2}\right) + v(x_2^2(\mathbf{s}) + (1+r)s_2) \right. \\ & \left. - h\left(\frac{y_2^2(\mathbf{s})}{w_2}\right) - u(x_1^1 - s_1) + g\left(\frac{y_1^1}{w_2}\right) - v(\hat{x}_2^2(s_1, s_G) + (1+r)s_1) + h\left(\frac{\hat{y}_2^2(s_1, s_G)}{w_2}\right) \right]. \end{aligned} \quad (\text{A.43})$$

The associated first-order conditions include:

$$x_1^1: u'(c_1^1) - \eta - \psi u'(c_1^1) = 0; \quad (\text{A.44})$$

$$x_2^1: u'(c_2^1) - \eta + \psi u'(c_2^1) = 0; \quad (\text{A.45})$$

$$y_1^1: -\frac{1}{w_1} g'\left(\frac{y_1^1}{w_1}\right) + \eta + \frac{\psi}{w_2} g'\left(\frac{y_1^1}{w_2}\right) = 0; \quad (\text{A.46})$$

$$y_2^1: -\frac{1}{w_2} g'\left(\frac{y_2^1}{w_2}\right) + \eta - \frac{\psi}{w_2} g'\left(\frac{y_1^1}{w_2}\right) = 0; \quad (\text{A.47})$$

$$s_2: -u'(c_2^1) + \frac{\partial \mathcal{V}^{2,sep}}{\partial s_2} + \psi \left\{ -u'(c_2^1) + (1+r)v'(c_2^2) + v'(c_2^2) \frac{\partial x_2^2}{\partial s_2} - \frac{1}{w_2} h'\left(\frac{y_2^2}{w_2}\right) \frac{\partial y_2^2}{\partial s_2} \right\} = 0. \quad (\text{A.48})$$

Equations (A.44)–(A.47) 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 parts (i) and (ii) of Lemma 2,

(A.48) is equivalent to

$$-u'(c_2^1) + (1+r)v'(c_2^2) - \psi u'(c_2^1) + \psi(1+r)v'(c_2^2) + \psi v'(c_2^2) \frac{\partial x_2^2}{\partial s_2} - \psi v'(c_2^2) \frac{\partial y_2^2}{\partial s_2} = 0. \quad (\text{A.49})$$

Rearranging (A.49) yields

$$\frac{u'(c_2^1)}{v'(c_2^2)} = (1+r) - \frac{\psi}{1+\psi} \left[\frac{\partial y_2^2}{\partial s_2} - \frac{\partial x_2^2}{\partial s_2} \right]. \quad (\text{A.50})$$

By part (iv) of Lemma 2, the term in square brackets in (A.50) is positive, so that

$$\frac{u'(c_2^1)}{v'(c_2^2)} < (1+r), \quad (\text{A.51})$$

and $\text{MRS}_{x_2^1, x_2^2} > -(1+r)$.

The first-order conditions (A.44) and (A.45) for x_1^1 and x_2^1 imply that

$$u'(c_1^1)(1-\psi) = u'(c_2^1)(1+\psi) = \eta. \quad (\text{A.52})$$

Because, $\psi > 0$, (A.52) implies that $u'(c_1^1) > u'(c_2^1)$ and, hence, that $c_1^1 < c_2^1$ because $u(\cdot)$ is strictly concave. By part (i) of Lemma 2, we have $c_1^2 = c_2^2$. Therefore,

$$\frac{u'(c_1^1)}{v'(c_1^2)} > \frac{u'(c_2^1)}{v'(c_1^2)} = \frac{u'(c_2^1)}{v'(c_2^2)}, \quad (\text{A.53})$$

which establishes part (ii) of the proposition. \square

References

- Apps, P., Rees, R., 2006. Repeated optimal nonlinear income taxation, unpublished manuscript, University of Sydney and University of Munich.
- 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/wpawuwpe/0403004.htm>.
- Bisin, A., Rampini, A. A., 2006. Markets as beneficial constraints on government. *Journal of Public Economics* 90, 601–629.
- 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., 2006. Tagging and redistributive taxation. *Annales d'Économie et de Statistique*, forthcoming.
- Boadway, R., Wildasin, D., 1994. Taxation and savings: A survey. *Fiscal Studies* 15, 19–63.
- Brett, C., Weymark, J. A., 2008. Public good provision and the comparative statics of optimal nonlinear income taxation. *International Economic Review* 49, 255–290.
- 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.
- Diamond, P., 2007. Comment on Golosov, Tsyvinski, and Werning (2007). In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), *NBER Macroeconomics Annual 2006*. Vol. 21. MIT Press, Cambridge, MA, pp. 365–379.
- 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 disassembling 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.
- Gaube, T., 2007. Optimum taxation of each year's income. *Journal of Public Economic Theory* 9, 127–150.
- Golosov, M., Tsyvinski, A., Werning, I., 2007. New dynamic public finance: A user's guide. In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), *NBER Macroeconomics Annual 2006*. Vol. 21. MIT Press, Cambridge, MA, pp. 317–363.
- Guesnerie, R., 1995. *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, Cambridge.
- Guesnerie, R., Seade, J., 1982. Nonlinear pricing in a finite economy. *Journal of Public Economics* 17, 157–179.

- Hamilton, J., Pestieau, P., 2005. Optimal income taxation and the ability distribution: Implications for migration equilibria. *International Tax and Public Finance* 12, 29–45.
- Hamilton, J., Slutsky, S., 2007. Optimal nonlinear income taxation with a finite population. *Journal of Economic Theory* 132, 548–556.
- Kocherlakota, N. R., 2006. Advances in dynamic optimal taxation. In: Blundell, R., Newey, W. K., Persson, T. (Eds.), *Advances In Economics and Econometrics. Theory and Applications, Ninth World Congress. Vol. I*. Cambridge University Press, Cambridge, pp. 269–297.
- Mirrlees, J. A., 1971. An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38, 175–208.
- Ordovery, 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.
- Piketty, T., 1993. Implementation of first-best allocations via generalized tax schedules. *Journal of Economic Theory* 61, 23–41.
- 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.
- Simula, L., 2007. Optimality conditions and comparative static properties of non-linear income taxes revisited. Working Paper No. 2007-15, Paris School of Economics.
- Stiglitz, J. E., 1982. Self-selection and Pareto efficient taxation. *Journal of Public Economics* 17, 213–240.
- 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.