

Gaube, Thomas

Conference Paper

Taxation of Annual Income as a Commitment Device

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session:
Dynamic Aspects in Optimal Taxation, No. F17-V3

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Gaube, Thomas (2010) : Taxation of Annual Income as a Commitment Device, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Dynamic Aspects in Optimal Taxation, No. F17-V3, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/37486>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Taxation of Annual Income as a Commitment Device

Thomas Gaube*

July 2010

Abstract

I study an intertemporal model of nonlinear, information constrained income taxation. It is shown that time-consistent taxation of annual income results in higher welfare than time-consistent taxation of the agents' income history if preferences are such that stationary allocations are efficient. If uncertainty is taken into account and if state-contingent fiscal policy is not feasible, time-consistent taxation of annual income can also welfare-dominate taxation of the agents' income history under full commitment. These findings may help in explaining why governments usually tax annual rather than lifetime earnings.

Keywords: nonlinear income taxation, tax base, time-consistency, commitment

JEL-Classification: H21, H24

*Department of Economics, University of Osnabrück, Rolandstrasse 8, D-49069 Osnabrück, Germany, Phone: +49-451-969-2747, Fax: +49-541-969-3528, Email: thomas.gaube@uos.de

1 Introduction

It is a matter of fact that income is mainly taxed on an annual basis and that governments do not commit to a particular income tax schedule for longer than one fiscal year. The aim of this paper is to discuss whether the decision to tax annual income rather than an agent's income history can be justified from the perspective of optimal nonlinear income taxation (Mirrlees, 1971).

To state the problem more precisely, consider an agent that lives for two periods with first-period income y_1 and second-period income y_2 . Since the tax burden in any period can only depend on what is observable in that period, the first-period tax can only depend on y_1 , whereas the second-period tax can either depend on current income y_2 , or the whole income history (y_1, y_2) . Hence, the question is whether one should tax the agent's annual income with tax functions $T_1(y_1)$ and $T_2(y_2)$ or their income history with tax functions $T_1(y_1)$ and $T_2(y_1, y_2)$.¹ Intuitively, the trade-off between these two regimes is clear: Taxation of annual income has the shortcoming that some allocations that can be implemented with function $T_2(y_1, y_2)$ may not be implementable with function $T_2(y_2)$. On the other hand, history-dependent taxation suffers from a commitment problem because the higher the agent's income in period one, the greater the temptation of the government to further increase the agent's tax burden in period two.

In the following, this trade-off will be investigated by comparing three regimes of intertemporal income taxation (see Table 1). The *full-commitment* regime FC relies on the assumption that the government can commit to tax functions $T_1(y_1)$ and $T_2(y_1, y_2)$ at the beginning of period one. In contrast, tax policy under *no commitment* NC must be time-consistent because function $T_2(y_1, y_2)$ can be re-designed at the beginning of period two. In the *partial-commitment* regime PC, time-consistency of function $T_2(y_2)$ is required as in regime NC, but with a guarantee that information about first-period income will be ignored. I take the view that time consistent taxation of annual income as in regime PC is closer to tax policy as observed in practice than regimes FC or NC.

The first question of the present paper is whether taxation of annual income in regime PC welfare-dominates taxation of the income history in regime NC. The answer to this question differs for different versions of the model outlined above, but the two main insights can be summarized as follows: Consider first a model

¹See Banks and Diamond (2010) for a comprehensive discussion of the choice of tax base in theory and practice.

| Regime | Decision in Period 1 | Decision in Period 2 |
|--------|--------------------------------|----------------------|
| FC | $T_1(y_1)$ and $T_2(y_1, y_2)$ | |
| NC | $T_1(y_1)$ | $T_2(y_1, y_2)$ |
| PC | $T_1(y_1)$ | $T_2(y_2)$ |

Table 1

without savings such that income means labor income as in the static Mirrlesian framework. In this context, it is shown that the *only* difference between regimes NC and PC is that allocations in regime PC have a tendency for being stationary while allocations in regime NC are inherently non-stationary. Therefore, the welfare comparison between these regimes just depends on whether stationary or non-stationary allocations are efficient under full commitment (Propositions 2-4). In the presence of a capital market, an additional commitment problem arises because first-period savings can also be observed in period two. In this framework, it is shown that time-consistent taxation of current earnings, current capital income, and wealth welfare-dominates time-consistent taxation of the agents' wealth and income history even when non-stationary allocations are efficient (Proposition 5).

The paper also discusses the trade-off between commitment in regime FC and flexibility in regime PC. For that purpose, I assume that preferences in period two are uncertain and that only state-invariant tax functions $T_2(y_1, y_2)$ can be implemented in period one. It is shown that regime FC can be welfare-inferior relative to regime PC where tax policy can flexibly react to a change in the agents' preferences (Proposition 7). This result again depends on whether stationary allocations are efficient. At least under this assumption, taxation of annual income is thus flexible, reliable, and sufficiently efficient at the same time.

Most contributions that deal with intertemporal extensions of the Mirrlesian model either assume full or no commitment. Regime FC was first investigated by Brito et al. (1991) and has recently gained renewed attention in the literature on 'new dynamic public finance'.² Regime NC was investigated by Berliant and Ledyard (2005) and by Brett and Weymark (2008).³ Kapicka (2006) and Gaube (2007) consider taxation of annual income, but they do not require time-

²See Golosov, Tsyvinski, and Werning (2007) and Kocherlakota (2010) for an overview.

³They consider models either without capital income (see Section 2) or with capital income (see Section 6). Krause (2009) investigates regime NC in a model with learning by doing.

consistency and concentrate on human capital accumulation or hidden savings. The present paper complements the literature by concentrating on time-consistent tax policy and by comparing regimes FC, NC, and PC. In showing that regime PC can be efficient, it also related to the concern that theories that build on the concept of history-dependent taxation might be of limited practical interest; see Judd (2007), Mankiw, Weinzierl, and Yagan (2009), Banks and Diamond (2010).

The idea that imperfect commitment can be better than no commitment is familiar in economics and has been discussed in the context of Mirrleesian income taxation by Roberts (1984) and Bisin and Rampini (2006). Roberts also compares regimes PC and NC. His argument in favor of regime PC, however, crucially relies on the assumption of an infinite time horizon with no discounting (see Section 4). The present paper asks whether regime PC welfare-dominates regime NC in a finite setting and irrespective of the discount factor. Bisin and Rampini consider a two-period model and show that the decision to ignore information about private savings can be welfare-improving. This issue will be discussed in Section 6.

In a prominent paper on dynamic principal-agent problems, Baron and Besanko (1984) have shown that a sequence of static contracts can replicate the optimal long-term contract if the latter is stationary or, equivalently, if randomization is inefficient. While the models of this literature differ from the income-tax model in several ways (see Berliant and Ledyard, 2005), Propositions 2 and 4 of the present paper rely on the same reasoning. Therefore, Hellwig's (2007a) assumption of decreasing risk-aversion, under which randomized income tax contracts are inefficient, plays an important role in the subsequent analysis.

The remainder of this paper is organized as follows: Section 2 presents the basic model and Section 3 briefly reviews randomized taxation. The welfare-comparison between regimes PC and NC is discussed in Section 4. Sections 5 and 6 extend this analysis by introducing random tax contracts and a capital market with observable savings. The trade-off between full commitment and flexibility is investigated in Section 7. Section 8 concludes. All proofs are relegated to the Appendix.

2 The model

I study an intertemporal extension of the two-type model of optimum income taxation (Stiglitz, 1982). There is a continuum of skilled agents with productivity

w^H and a continuum of unskilled agents with productivity $w^L < w^H$. Without loss of generality, the mass of skilled and unskilled agents is normalized to unity. It is assumed that the agents' type $i = L, H$ does not change over time $t = 1, 2$.

The agents maximize a time-invariant utility function $U = u(c_1^i, l_1^i) + \rho u(c_2^i, l_2^i)$ with discount factor $\rho > 0$, where $u(c_t^i, l_t^i)$ is increasing in consumption c_t^i , decreasing in labor l_t^i , and strictly quasiconcave. In addition, the single-crossing property is assumed to hold. This property is related to the utility representation $v^i(c_t^i, y_t^i) \equiv u(c_t^i, y_t^i/w^i)$, where $y_t^i \equiv w^i l_t^i$ stands for labor income in period t . In order to simplify the notation, let $a_t^i \equiv (c_t^i, y_t^i)$ denote an allocation of type i in period t and $a_t \equiv (a_t^L, a_t^H)$ an allocation in period t .

Like most of the literature, the present analysis concentrates on the 'redistributive case' of optimum income taxation. Therefore, I assume that the government maximizes a weighted utilitarian welfare function

$$W(a_1, a_2) = \alpha[v^L(a_1^L) + \rho v^L(a_2^L)] + (1 - \alpha)[v^H(a_1^H) + \rho v^H(a_2^H)],$$

where parameter $\alpha \geq 1/2$ is chosen such that only the incentive constraint of the skilled agents is relevant under information constrained taxation.⁴

For the moment, it is assumed that neither a credit market nor a storage technology is available. The economy's budget constraints thus take the form

$$y_1^L + y_1^H - c_1^L - c_1^H \geq 0 \tag{1}$$

$$y_2^L + y_2^H - c_2^L - c_2^H \geq 0. \tag{2}$$

In the following, I will compare maximum welfare in three versions of the model that differ from each other with respect to the degree of government commitment at the beginning of period one.⁵ Assume first that the government can commit to a first-period tax function $T_1(y_1)$ and a second-period tax function $T_2(y_1, y_2)$ at the beginning of period one. This is equivalent to employing a direct

⁴Roëll (1985) and Hellwig (2007b) derive sufficiency conditions for optimality of the redistributive case under utilitarian welfare ($\alpha = 1/2$). Incentive compatibility implies that maximization of $W(a_1, a_2)$ with $\alpha = 1$ is equivalent to maximization of Rawlsian welfare. In this case, the incentive constraints of the skilled agents must be binding in the optimum.

⁵The definition of regimes FC and NC is the same as in Brito et al. (1991), Berliant and Ledyard (2005), and Brett and Weymark (2008). Regime PC is a time-consistent variant of the model in Gaube (2007).

mechanism where the government offers two allocations (a_1^H, a_2^H) and (a_1^L, a_2^L) . Therefore, the incentive compatibility constraint of the skilled agents

$$v^H(a_1^H) + \rho v^H(a_2^H) \geq v^H(a_1^L) + \rho v^H(a_2^L) \quad (3)$$

and the corresponding constraint of the unskilled agents must be taken into account. Since I assume that redistribution in favor of the unskilled agents is desirable, the latter constraint will be slack. The maximization problem under *full commitment* (FC) can thus be expressed as follows

$$\text{FC: } \max_{a_1, a_2} W(a_1, a_2) \quad \text{s.t.} \quad (1), (2), (3).$$

Consider next a situation where the government cannot commit to function $T_2(y_1, y_2)$ in period one. Under *no commitment* (NC), either a separating equilibrium or a pooling equilibrium can be optimal.⁶ In the *pooling regime* (NCP), one has $a_1^L = a_1^H$. The agents thus do not reveal their type in period one and the government's maximization problem in period two

$$\text{NCP2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (2), (3)$$

is equivalent to the static income tax problem. Accordingly, the first-period maximization problem can be written in the form

$$\text{NCP1: } \max_{a_1} W(a_1, a_2^{NCP}) \quad \text{s.t.} \quad (1), a_1^L = a_1^H,$$

where a_2^{NCP} is the solution to problem NCP2. Since utility is additively separable over time, this solution does not depend on a_1 .

In the *separating regime* (NCS), the agents reveal their type in the first period. This implies that welfare maximization in the second period is equivalent to the first-best problem

$$\text{NCS2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (2).$$

Since the agents anticipate that solution a_2^{NCS} to this problem will be implemented in period two, the first-period maximization problem can be written as

$$\text{NCS1: } \max_{a_1} W(a_1, a_2^{NCS}) \quad \text{s.t.} \quad (1), (3), a_1^L \neq a_1^H.$$

⁶For simplicity, I consider only equilibria with full separation and full pooling. In general, partial pooling (a fraction of agents of type H is pooled with a fraction of agents of type L) should also be taken into account. It will become clear from the subsequent analysis, however, that the results do not change if one allows for partial pooling as well.

Maximum welfare in the pooling and the separating regime will be denoted by $W(\text{NCP})$ and $W(\text{NCS})$, respectively. Maximum welfare under no commitment $W(\text{NC})$ is given by the maximum of these values.

Consider next tax policy under *partial commitment* (PC). In this case, the government designs a tax function $T_t(y_t)$ at the beginning of each period. This implies that incentive compatibility becomes more restrictive than under full commitment because a skilled agent can decide whether she wants to mimick an unskilled agent in both or in just one of the two periods. Therefore, constraint (3) must be replaced by two separate incentive constraints

$$v^H(a_1^H) \geq v^H(a_1^L) \quad (4)$$

$$v^H(a_2^H) \geq v^H(a_2^L). \quad (5)$$

Since the tax function in period two does not depend on labor income in period one, separation in the first period does not lead to type-specific taxation in the second period. Nevertheless, the second-period policy must be time-consistent. Similar to regime NC, we thus have to distinguish between welfare maximization at the beginning of period two

$$\text{PC2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (2), (5)$$

with solution a_2^{PC} and welfare maximization at the beginning of period one

$$\text{PC1: } \max_{a_1} W(a_1, a_2^{PC}) \quad \text{s.t.} \quad (1), (4).$$

Maximum welfare under partial commitment results from the solution to these problems and is denoted by $W(\text{PC})$.

3 Randomization in the static income tax model

The standard static income tax model is based on the assumption that only deterministic allocations a^i can be implemented. In this section, I will consider the generalized static income tax model, where the allocation $\tilde{a}^i = (\tilde{c}^i, \tilde{y}^i)$ is a discrete random variable with $K \geq 2$ realizations a_k^i that occur with probability π_k^i . This model has been investigated by Brito et al. (1995) who demonstrate that randomized taxation can Pareto-dominate deterministic taxation. Recently, Hellwig (2007a) worked out under which circumstances this puzzling phenomenon cannot occur.

Let $Ev^i(\tilde{a}^i) = \sum_{k=1}^K \pi_k^i v^i(a_k^i)$ denote expected utility, $E\tilde{c}^i$ expected consumption, and $E\tilde{y}^i$ expected income. The government maximizes expected welfare

$$EW(\tilde{a}^L, \tilde{a}^H) = \alpha Ev^L(\tilde{a}^L) + (1 - \alpha)Ev^H(\tilde{a}^H),$$

subject to the budget constraint and the incentive constraint

$$E\tilde{y}^L + E\tilde{y}^H - E\tilde{c}^L - E\tilde{c}^H \geq 0 \quad (6)$$

$$Ev^H(\tilde{a}^H) \geq Ev^H(\tilde{a}^L). \quad (7)$$

As before, it is assumed that redistribution in favor of the unskilled agents is desirable such that only the downward incentive constraint is relevant. The *generalized static* income tax problem (GS) is thus defined as follows

$$\text{GS: } \max_{a_k^i} EW(\tilde{a}^L, \tilde{a}^H) \quad \text{s.t.} \quad (6), (7).$$

An important question with respect to problem GS is whether randomization is actually desirable, that is, whether or not the solution to GS is deterministic such that $a_k^i = a^i$ for all k . Hellwig (2007a) argues that randomization must be inefficient if agents are risk averse and if it does not help in mitigating the incentive constraint (7) because the unskilled agents are more risk-averse than the skilled agents. He has formalized this idea by introducing the concept of a consumption-specific risk premium $z^i(\tilde{a})$ that is implicitly defined by the equation

$$Ev^i(\tilde{a}) = v^i(E\tilde{c} - z^i(\tilde{a}), E\tilde{y}).$$

An agent is risk averse if the risk premium is positive, and the unskilled agents are at least as risk-averse as the skilled agents if the premium is weakly decreasing in productivity w^i . This is the idea behind the assumption of *weakly decreasing consumption specific risk aversion (WDCRA)* which requires that $z^L(\tilde{a}) \geq z^H(\tilde{a}) > 0$ holds for any non-degenerate random allocation \tilde{a} .

Proposition 1 (Hellwig 2007a) : *Assume WDCRA. Then randomization is not desirable in the generalized static income tax problem GS.*

Since Hellwig studies a more general model with a continuum of types, his proof is quite involved. For the present model, the result could easily be proved along the lines of the intuitive argument given above. Note that assumption WDCRA only requires that a risk-averse agent does not become more risk averse if she is allowed to work less without suffering from reduced consumption. Apart from Proposition 1, the assumption is also important from a technical point of view because it implies convexity of the deterministic income tax problem (see Hellwig, 2007a).

4 Comparison of regimes

I will now return to the model of Section 2 and ask whether regimes FC, PC, and NC can be ranked in terms of welfare. It is clear that full commitment cannot be worse than imperfect commitment, because any allocation that can be implemented in regimes PC or NC can also be implemented in regime FC. Therefore, we have

Lemma 1: $W(FC) \geq W(NC)$ and $W(FC) \geq W(PC)$.

By similar reasoning, it can easily be verified that any allocation that can be implemented in the pooling regime NCP can also be implemented under partial commitment. This implies

Lemma 2: $W(PC) \geq W(NCP)$.

The intuition behind Lemma 2 is straightforward. In regime NCP, the government does not want to use information about first-period behavior in period two, but this is only credible if a rather unattractive allocation is implemented in period one. Under partial commitment, the same effect is attained without tying the government's hands in the first period.

Lemma 2 implies $W(PC) \geq W(NC)$ if the threat of 'soaking the rich' in future periods is so high that a separating equilibrium does not exist in regime NC. This is the argument used by Roberts (1984) in a model with infinite time-horizon and no discounting. With a finite life-span, however, the threat of future exploitation can be rather small, in particular if no restriction on the discount factor ρ is imposed. In the following, we thus have to compare welfare between regimes PC and NCS as well.⁷

The main insight of this section is rather simple: Optimal allocations in regime NCS are inherently non-stationary, whereas optimal allocations in regime PC are inherently stationary. Therefore, the welfare comparison between these regimes just depends on whether stationary or non-stationary allocations are efficient under full commitment. This can be seen by restating the incentive constraint (3) of problem FC in the simplified form

$$IC_1 + IC_2 \geq 0, \tag{8}$$

⁷Note that a general welfare comparison between regimes NCS and NCP is not at hand. It can be shown, however, that NCS welfare-dominates NCP if the discount factor is sufficiently small (see Berliant and Ledyard, 2005).

where $IC_t \equiv \rho^{t-1}[v^H(a_t^H) - v^H(a_t^L)]$. Optimal taxation under partial commitment may differ from optimal taxation under full commitment because the somewhat stronger constraint

$$IC_1 \geq 0 \quad \text{and} \quad IC_2 \geq 0 \quad (9)$$

must hold in problem PC. Consider now regime NCS. In this regime, the government effectively commits to implement the static first-best allocation a^* in period two. Since it is assumed that redistribution in favor of the unskilled agents is desirable in second best, the first-best allocation violates the incentive constraint $IC_2 \geq 0$. Therefore, problem NCS is equivalent to welfare-maximization subject to the budget constraints and the additional constraint

$$IC_1 \geq -B \quad \text{and} \quad IC_2 = B, \quad (10)$$

where parameter $B \equiv \rho[v^H(a^{H*}) - v^H(a^{L*})] < 0$ is the value of function IC_2 when the static first-best allocation a^* is implemented.

Apart from the difference between (9) and (10), problems PC and NCS are equivalent. Therefore, the welfare comparison between these regimes only depends on whether the values of IC_1 and IC_2 in a solution to problem FC are close to zero as in (9) or close to $-B$ and B as in (10).

It has been shown for other multiperiod adverse-selection problems that the full-commitment optimum is stationary if randomization is not desirable in the static model.⁸ The following result points out that this property also holds in the present framework. Therefore, regime PC welfare-dominates regime NCS under assumption WDCRA.

Proposition 2: *Assume WDCRA. Then $W(FC) = W(PC) > W(NC)$.*

As noted above, it should not be surprising that assumption WDCRA is sufficient for $W(PC) > W(NC)$. Indeed, the most interesting implication of equations (8)-(10) is that counterexamples to this welfare-ranking can be constructed when assumption WDCRA is violated. This follows from the observation that $IC_1 \neq IC_2$ must hold in any non-stationary solution to problem FC. Hence, optimal non-stationary allocations violate the constraint $IC_t \geq 0$ in one period at the expense

⁸The idea goes back to Baron and Besanko (1984) who assume quasilinear utility and has been used by Brito et al. (1991) for investigating a model similar to that in Section 6 below. Note that non-desirability of randomization does not generally imply $W(FC) = W(PC)$, as has been shown by Gaube (2007) for a model with non-quasilinear utility and hidden savings.

of a tightened constraint in the other period. This is exactly the outcome in regime NCS. Therefore, this regime must welfare-dominate regime PC when the desire for randomization is sufficiently strong. In order to confirm this intuitive argument, consider the Cobb-Douglas example

$$U = c_1^i(e - l_1^i) + \rho c_2^i(e - l_2^i), \quad (11)$$

where e is the individuals' time endowment and $(e - l_t^i)$ is leisure time. This example is standard, except that the utility representation is not concave. Accordingly, assumption WDCRA is violated. Note that such an example must be investigated with care because the income tax problem will typically not be convex. Moreover, one has to verify that problems NCS and PC are indeed interesting in the sense that the incentive constraints (3)-(5) are binding in the optimum. At least for parameters

$$w^H = 1, \quad w^L = 1/4, \quad 2/3 \leq \alpha \leq 1, \quad 1/4 \leq \rho \leq 1/3, \quad (12)$$

it can be shown that redistribution from skilled to unskilled individuals is always desirable in both regimes.

Proposition 3: *Consider the Cobb-Douglas-example (11) with parameters according to (12). Then $W(FC) \geq W(NC) > W(PC)$.*

The intuition behind Proposition 3 is as follows: Under the assumptions (11) and (12), the skilled agents work full time with zero consumption in first best. Hence, their utility is zero in the second period of regime NCS. In the first period of that regime, they must thus be compensated with high consumption and plenty of leisure time. Since function (11) is convex, such a non-stationary utility profile leads to higher welfare than a stationary utility profile that occurs in regime PC. Hence, the non-stationary allocation in regime NCS welfare-dominates the stationary allocation in regime PC.

5 Randomization in the intertemporal model

The preceding analysis is based on the assumption that intertemporal taxation is deterministic even if randomized taxation were desirable. It is thus natural to ask whether the comparison of regimes will change if randomization is not ruled out in the intertemporal model from the outset.

Using the same notation as in Section 3, let $\tilde{a}_t = (\tilde{a}_t^L, \tilde{a}_t^H)$ denote the random allocation in period t . As before, the government maximizes expected welfare

$$EW(\tilde{a}_1, \tilde{a}_2) = \alpha[Ev^L(\tilde{a}_1^L) + \rho Ev^L(\tilde{a}_2^L)] + (1 - \alpha)[Ev^H(\tilde{a}_1^H) + \rho Ev^H(\tilde{a}_2^H)]$$

subject to the budget and incentive constraints

$$E\tilde{y}_t^L + E\tilde{y}_t^H - E\tilde{c}_t^L - E\tilde{c}_t^H \geq 0, \quad t = 1, 2 \quad (13)$$

$$Ev^H(\tilde{a}_1^H) + \rho Ev^H(\tilde{a}_2^H) \geq Ev^H(\tilde{a}_1^L) + \rho Ev^H(\tilde{a}_2^L) \quad (14)$$

$$Ev^H(\tilde{a}_t^H) \geq Ev^H(\tilde{a}_t^L), \quad t = 1, 2 \quad (15)$$

that replace constraints (1)-(5) of the deterministic model in Section 2. The definition of regimes FC, PC, NCS, and NCP is the same as before, except that expected welfare is maximized and that all constraints must hold in expectation. Therefore, I will only list the relevant constraints in each regime.

FC: constraints (13), (14)

PC: constraints (13), (15), time consistency

NCP: constraints (13), (14), $\tilde{a}_1^L = \tilde{a}_1^H$, time consistency

NCS: constraints (13), (14), $\tilde{a}_1^L \neq \tilde{a}_1^H$, time consistency

This list makes it clear that the constraints of regime FC are less restrictive than those of all other regimes and that the constraints of PC are less restrictive than those of NCP. Therefore, Lemmas 1 and 2 still hold and the crucial question is again whether regime PC welfare-dominates regime NCS.

Similar to Proposition 2, the following proposition complements earlier findings on multi-period contracting where it was shown for a quasilinear environment that the optimal long-term contract can be replicated by a sequence of randomized short-term contracts (see Laffont and Tirole, 1993, p. 103-104). Here it is pointed out that the same logic applies to the income tax model with utility $v^i(c, y)$.

Proposition 4: *If random tax contracts are available in the intertemporal income tax model, then $EW(FC) = EW(PC) \geq EW(NC)$.*

When random contracts are available, the potential advantage of regime NCS over regime PC vanishes because it is not necessary anymore to imitate a random static contract by a non-stationary intertemporal contract. In fact, the proof shows that any non-stationary allocation $(\tilde{a}_1^i, \tilde{a}_2^i)$ of regime FC can be replicated

by a stationary allocation where a convex combination between the lotteries \tilde{a}_1^i and \tilde{a}_2^i is implemented in each period. This stationary allocation is time-consistent and can thus be implemented in regime PC. Note that this argument does not extend to regime NC because optimal allocations in that regime are non-stationary even when randomization is not desirable. Therefore, $EW(PC) > EW(NC)$ will be the rule rather than the exception.

6 Introducing a capital market

In this section, I consider a small open economy where all agents and the government have access to a capital market that transforms one unit of first-period income into R units of second-period income – and vice versa. Hence, the agents' budget constraints now take the form $y_1^i - c_1^i - T_1^i - s^i \geq 0$ and $y_2^i - c_2^i - T_2^i + R s^i \geq 0$, and the economy's budget constraints can be written as

$$y_1^L + y_1^H - c_1^L - c_1^H - (s^L + s^H + s^G) \geq 0 \quad (16)$$

$$y_2^L + y_2^H - c_2^L - c_2^H + R(s^L + s^H + s^G) \geq 0, \quad (17)$$

where s^i and s^G stand for private and government savings, respectively.

Brito et al. (1991) have studied income taxation within this model under the assumption of full commitment and Brett and Weymark (2008) have done so under the assumption of no commitment.⁹ Both papers assume that the government can observe labor income and savings of each individual. This assumption means that fiscal policy is implemented by a tax function $T_1(y_1^i, s^i)$ for period one and a tax function $T_2(y_1^i, y_2^i, s^i)$ for period two.¹⁰ One can think of these functions as being designed at the beginning of period one, but function $T_2(y_1^i, y_2^i, s^i)$ must be time-consistent under no commitment.

As before, partial commitment means that second-period tax policy must be time-consistent, but without taking first-period labor income into account. This implies that the two tax functions take the form $T_1(y_1^i, s^i)$ and $T_2(y_2^i, s^i)$. Note that there is a one-to-one relationship between the savings of an agent, her wealth at

⁹Brett and Weymark (2008) consider a model with one individual of each type while Brito et al. (1991) and the present paper assume a continuum of individuals of each type. This difference, however, is not important for the subsequent analysis.

¹⁰Hence, income shifting between labor income and capital income is ruled out by assumption. See Christiansen and Tuomala (2008) for a critical discussion.

the end of period one, and her capital income in period two. Partial commitment can thus be interpreted as a regime where the tax burden in each period depends only on current labor income, current capital income, and end-of-period wealth.

The main difference between the present model and the model of Section 2 is that the government can now observe two actions of each agent – labor income and savings – at the end of period one. We thus have a double commitment problem in period two. Since capital income is taxed in practice, I assume that savings can affect the second-period tax burden in regime PC. While the savings tax creates a commitment problem in itself, it will be shown that the commitment to ignore first-period labor income leads to a welfare improvement relative to regime NC. The question whether the additional commitment to not tax savings will lead to a further welfare gain will be discussed at the end of this section.

Under full commitment, the government implements the tax functions $T_1(y_1^i, s^i)$ and $T_2(y_1^i, y_2^i, s^i)$ at the beginning of period one. As before, this allows each agent of type H to opt for the same allocation as an agent of type L . Therefore, the incentive constraint (3) remains valid, and the maximization problem differs from that in Section 2 only because the economy's budget constraints have changed and because the vector of savings $s \equiv (s^L, s^H, s^G)$ must also be taken into account.¹¹

$$\text{FC : } \max_{a_1, a_2, s} W(a_1, a_2) \quad \text{s.t.} \quad (3), (16), (17)$$

Under no commitment, we must again distinguish between equilibria with and without separation of types. Separation is impossible if skilled and unskilled agents do not differ with respect to labor income or savings in period one. This is equivalent to $(a_1^H, s^H) = (a_1^L, s^L)$. Therefore, the pooling regime NCP and the separating regime NCS can be defined as follows:

$$\text{NCP2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (3), (17)$$

$$\text{NCP1: } \max_{a_1, s} W(a_1, a_2^{NCP}(z)) \quad \text{s.t.} \quad (16), (a_1^L, s^L) = (a_1^H, s^H)$$

$$\text{NCS2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (17)$$

$$\text{NCS1: } \max_{a_1, s} W(a_1, a_2^{NCS}(z)) \quad \text{s.t.} \quad (3), (16), (a_1^L, s^L) \neq (a_1^H, s^H)$$

¹¹Note that s^i and s^G would not play a role in problem FC if the budget constraints (16) and (17) were replaced by the intertemporal budget constraint. However, the present exposition with two constraints is more suitable for defining the maximization problems NC and PC below.

Here, $a_2^{NCP}(z)$ and $a_2^{NCS}(z)$ denote the solutions to problems NCP2 and NCS2, respectively. Since the utility function is time-separable, these solutions only depend on aggregate savings $z \equiv s^L + s^H + s^G$, but not on the allocation a_1 or the composition of z .

Under partial commitment, the government ignores first-period labor income in period two. Nevertheless, agents with the same labor income in the second period may pay different taxes in that period because $T_2(y_2^i, s^i)$ is now a function of labor income and savings (capital income) alike. Hence, distinct levels of savings allow for type-specific tax contracts in period two. Therefore, we now have to distinguish between pooling and separation of types in regime PC. The pooling-regime PCP occurs if and only if $s^L = s^H$ and can be defined in the same way as problem PC in Section 2.

$$\text{PCP2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (5), (17)$$

$$\text{PCP1: } \max_{a_1, s} W(a_1, a_2^{PCP}(z)) \quad \text{s.t.} \quad (4), (16), s^L = s^H$$

The separating regime PCS is equivalent to regime NCS, except that constraint $(a_1^H, s^H) \neq (a_1^L, s^L)$ must be replaced by $s^H \neq s^L$.

$$\text{PCS2: } \max_{a_2} W(a_1, a_2) \quad \text{s.t.} \quad (17)$$

$$\text{PCS1: } \max_{a_1, s} W(a_1, a_2^{PCS}(z)) \quad \text{s.t.} \quad (3), (16), s^L \neq s^H$$

Similar to the definition of $W(NC)$, the welfare maximum under partial commitment $W(PC)$ is now defined as highest welfare that can be attained either in the pooling regime PCP or the separating regime PCS.

It can easily be verified that the constraints of problem FC are less restrictive than those of all other problems and that the constraints of problem NCP are more restrictive than those of problem PCP. Therefore, Lemmas 1 and 2 also hold in the present context. What is new here is that a similar argument can be made with respect to the separating regime NCS: Any allocation of that regime can now also be implemented under partial commitment. Taken together, these findings imply

Proposition 5: *Assume that a capital market is available and that savings can be taxed. Then $W(FC) \geq W(PC) \geq W(NC)$.*

The intuition behind this result is as follows: Under partial commitment, the government commits to neglect the information about first-period labor income,

but not about savings that can again be observed in period two. Since the individuals' savings decision also depends on tax policy in period one, separation of types becomes an option, whereas pooling does not depend on first-period labor supply. Therefore, the set of incentive compatible allocations is strictly larger under partial commitment than under no commitment.

Proposition 5 differs from Proposition 2 because regime PC now welfare dominates regime NC even though the optimum under partial commitment does typically not replicate the optimum under full commitment. In fact, these optima do only coincide if it is efficient to have identical savings of both types. The latter condition holds in the special case $\rho R = 1$.¹² If this assumption is made in addition to assumption WDCRA, one obtains the same welfare-ranking as in Proposition 2.

Proposition 6: *Assume that a capital market is available and that savings can be taxed. Assume also $\rho R = 1$ and WDCRA. Then $W(FC) = W(PC) \geq W(NC)$.*

Proposition 6 makes it clear that the full-commitment optimum can be implemented in regime PC, but only under special circumstances. This raises the question whether the additional commitment to not tax the individuals' savings will lead to a further welfare improvement relative to regime PC. This question is closely related to the analysis of Bisin and Rampini (2006) who also consider a two-period model with asymmetric information and show that private access to hidden capital markets can improve welfare. Their framework, however, differs from the present one in several ways. In particular, they assume that the agents do not work in period two.¹³ This difference is important because hidden capital markets do not only serve as a commitment device, but also distort incentive compatibility. The incentive problem occurs because access to a capital market reduces the skilled agents' cost of mimicking the unskilled agents in period two. In fact, Gaube (2007) has shown for a model with slightly stronger assumptions than those of Proposition 6 that the full-commitment optimum cannot be implemented with tax functions $T_1(y_1^i)$ and $T_2(y_2^i)$. Under the assumptions of Proposition 6, hidden savings thus lead to a welfare loss relative to regime PC. It is an open question for future research whether another regime with constrained

¹²Under the assumption $\rho R = 1$, problem FC becomes equivalent to problem GS of Section 3. Most findings of Brito et al. (1991) on problem FC rely on this assumption.

¹³This assumption is also made in the literature that extends the static Mirrlees model for investigating pension systems. See Tenhunen and Tuomala (2010) for a recent contribution.

taxation of savings can be found that welfare dominates tax policy under partial commitment.

7 Full commitment vs. flexibility

The preceding analysis relies on the assumption that tax policy must be time consistent. Under this assumption, the full-commitment regime just serves as a theoretical benchmark, and the relevant question is whether taxation of annual income in regime PC welfare dominates taxation of the agents' income history in regime NC. In this section, I will discuss another argument in favor of annual income taxation, namely that it is more flexible than history-dependent taxation (under commitment), where the tax schedule of all fiscal periods must be designed at the beginning of period one.

The general idea that a tax system should be flexible is well-established in public economics. It has been argued, for example, that a flexible tax system is required for reasonably rapid adjustments of fiscal policy to economic and political developments (see Meade, 1978, pp. 21-22). However, to my knowledge, such a preference for flexibility has not been formally investigated in the literature on information-constrained taxation.

Following previous approaches for modeling the trade-off between commitment and flexibility,¹⁴ I assume that flexibility is valuable because future preferences for public services are uncertain and because it is impossible to make tax law contingent on how this uncertainty will be resolved. Therefore, utility

$$\bar{v}^i(a_t, g_t, \beta_t) = v^i(a_t) + \beta_t f(g_t) \quad (18)$$

of an agent in period t now also depends on the public service g_t and a preference parameter β_t . It is assumed that $f(g_t)$ is strictly concave and that the government learns about β_t at the beginning of each period. Hence, at the beginning of period one, first-period preferences β_1 are known, but second-period preferences $\tilde{\beta}_2$ are uncertain. I assume that $\tilde{\beta}_2$ is a non-degenerate discrete random variable with realizations β_{2k} , probabilities π_k , and expectation $E\tilde{\beta}_2 = \beta_1$. This implies that expected utility is time-invariant as in the preceding analysis. Since state-contingent policies (a_{2k}, g_{2k}) are taken into account, the second-period allocation $(\tilde{a}_2, \tilde{g}_2)$ is also random from the perspective of period one.

¹⁴See Amador, Werning and Angeletos (2006) for a recent contribution. They consider the decision problem of an individual with preferences similar to those of the present analysis.

Using the notation $E\bar{v}^i(\tilde{a}_2^i, \tilde{g}_2, \tilde{\beta}_2) \equiv \sum_{k=1}^K \pi_k \bar{v}^i(a_{2k}^i, g_{2k}, \beta_{2k})$, expected lifetime utility of an agent and expected welfare can be written in the form

$$EV^i(a_1^i, g_1, \beta_1, \tilde{a}_2^i, \tilde{g}_2, \tilde{\beta}_2) = \bar{v}^i(a_1^i, g_1, \beta_1) + \rho E\bar{v}^i(\tilde{a}_2^i, \tilde{g}_2, \tilde{\beta}_2)$$

$$EW(a_1, g_1, \beta_1, \tilde{a}_2, \tilde{g}_2, \tilde{\beta}_2) = \alpha EV^L(\cdot) + (1 - \alpha)EV^H(\cdot).$$

Similar to the model of Section 2, I assume that neither a conventional capital market nor state contingent securities are available. Therefore, we get separate budget constraints for period one and each state of period two:

$$y_1^L + y_1^H - c_1^L - c_1^H - g_1 \geq 0 \quad (19)$$

$$y_{2k}^L + y_{2k}^H - c_{2k}^L - c_{2k}^H - g_{2k} \geq 0, \quad k = 1, \dots, K. \quad (20)$$

In the following, I will compare three regimes of intertemporal income taxation under uncertainty. Assume first that the government can commit to an allocation a_1^i and state-contingent allocations a_{2k}^i at the beginning of period one. In such a regime with *full commitment and complete contracting* (FCC), tax policy must take account of the intertemporal incentive constraint

$$\bar{v}^H(a_1^H, g_1, \beta_1) + \rho E\bar{v}^H(\tilde{a}_2^H, \tilde{g}_2, \tilde{\beta}_2) \geq \bar{v}^H(a_1^L, g_1, \beta_1) + \rho E\bar{v}^H(\tilde{a}_2^L, \tilde{g}_2, \tilde{\beta}_2) \quad (21)$$

and maximum welfare results from the solution to problem

$$\text{FCC : } \max_{a_1, g_1, \tilde{a}_2, \tilde{g}_2} EW(\cdot) \quad \text{s.t.} \quad (19), (20), (21).$$

In regime FCC, the government must commit to state-contingent tax functions $T_{2k}(y_1, y_{2k})$ at the beginning of period one. Such tax functions, however, are not observed in practice. One reason for this may be that it is impossible to verify the realization of $\tilde{\beta}_2$. Therefore, I will consider the more interesting case where the government can only commit to a state-invariant policy (a_2, g_2) in period one. This regime with *full commitment and incomplete contracting* (FCI) leads to the maximization problem

$$\text{FCI : } \max_{a_1, g_1, \tilde{a}_2, \tilde{g}_2} EW(\cdot) \quad \text{s.t.} \quad (19), (20), (21), (a_{2k}, g_{2k}) = (a_2, g_2) \forall k.$$

Consider now time-consistent taxation of annual income. Since uncertainty is resolved at the beginning of period two, time-consistency implies that state-contingent allocations (a_{2k}, g_{2k}) can be implemented, but that state-contingent

incentive constraints must also be taken into account. Incentive compatibility thus requires

$$\bar{v}^H(a_1^H, g_1, \beta_1) \geq \bar{v}^H(a_1^L, g_1, \beta_1) \quad (22)$$

$$\bar{v}^H(a_{2k}^H, g_{2k}, \beta_{2k}) \geq \bar{v}^H(a_{2k}^L, g_{2k}, \beta_{2k}), \quad k = 1, \dots, K. \quad (23)$$

In each state k of period two, the government maximizes welfare subject to the state-contingent budget and incentive constraint. This implies that second-period tax policy maximizes expected welfare in period two subject to these constraints. In analogy to the definition of regime PC in the preceding sections, we thus obtain the maximization problems

$$\text{PC2: } \max_{\bar{a}_2, \bar{g}_2} EW(\cdot) \quad \text{s.t.} \quad (20), (23)$$

$$\text{PC1: } \max_{a_1, g_1} EW(\cdot) \quad \text{s.t.} \quad (19), (22).$$

The trade-off between commitment and flexibility can be seen by comparing the constraints of regimes FCI and PC. Constraint $(a_{2k}, g_{2k}) = (a_2, g_2)$ in regime FCI means that this regime is inflexible because the planner must commit to an allocation for period two before uncertainty is resolved. Constraints (22)-(23) in regime PC reveal that this regime suffers from imperfect commitment relative to regime FCI where the weaker lifetime incentive constraint (21) is imposed. It turns out that the utility loss from imperfect commitment is less severe than the utility loss from inflexibility if the solution to problem FCI is stationary. Since the latter is implied by assumption WDCRA, one obtains

Proposition 7: *Assume WDCRA. Then $EW(FCC) \geq EW(PC) > EW(FCI)$.*

Under the assumption WDCRA, regime PC thus appears to be an attractive solution to the government's problem of finding a tax system that is flexible and reliable at the same time.

8 Conclusion

This paper shows that taxation of annual income can welfare-dominate a regime with history-dependent taxes if tax policy must be time-consistent, or if state-contingent fiscal policy is not feasible under full commitment. These findings imply that the practice of taxing income on an annual basis can be justified by

the theory of information-constrained income taxation. However, the paper also makes it clear that this conclusion rests on an important qualification, namely that stationary allocations are efficient under full commitment.

The analysis is confined to a framework with two types and two periods because the definition of regime NC becomes unduly complicated in a more general setting. However, the formal arguments make it clear that the reasoning extends to a model with a finite number of periods and types, provided that only downward incentive constraints are binding in the optimum.¹⁵ It is less clear, though, whether similar results can also be derived for related models with imperfect correlation of types over time or a non-benevolent government.¹⁶

Note that the commitment to ignore past earnings is not the only commitment device observed in practice. Up to now, governments have been able to commit, for example, to a comprehensive income tax, a linear tax on savings, or constitutional constraints that prevent excessive taxation of income and wealth. Apart from the commitment effect, such constraints also imply that real-world tax policies appear to be much simpler than the prescriptions that follow from the theory of information-constrained taxation (see Judd, 2007 and Mankiw, Weinzierl, and Yagan, 2009 for a critical assessment). So far, however, the profession has little to say about the advantages and disadvantages of the various restrictions one might impose on tax bases and tax schedules in an intertemporal setting.¹⁷

Appendix

Proof of Proposition 2

In the following, I will show that $a_1 = a_2$ must hold in a solution to problem FC. This implies that the same allocation can be implemented in regime PC.

¹⁵Assumption WDCRA makes randomization attractive rather than unattractive when upward incentive constraints are binding in the optimum.

¹⁶The full-commitment optimum can be renegotiation-proof when the agents' abilities are imperfectly correlated over time (see Battaglini and Coate, 2008), or time-consistent when self-interested politicians with infinite time-horizon are subject to electoral accountability (see Acemoglu, Golosov, and Tsyvinski, 2010).

¹⁷For example, Vickrey's (1939) proposal of a life-time income tax imposes such a constraint. While the concept makes use of history-dependent tax functions $T_1(y_1)$ and $T_2(y_1, y_2)$ due to tax-averaging, it also requires government commitment because the present value of T_1 and T_2 must be a function of the present value of y_1 and y_2 . To my knowledge, time-consistent nonlinear Vickrey-taxation has not been explored so far in the literature.

Therefore, $W(FC) = W(PC)$. Since $a_1 = a_2$ cannot be optimal in regime NC,¹⁸ we also have $W(FC) > W(NC)$.

Consider now problem FC. Dividing functions $v_t^i(c_t^i, y_t^i)$ by $(1 - \rho)$, one obtains the welfare function

$$\hat{W}(a_1, a_2) \equiv \frac{W(a_1, a_2)}{1 + \rho} = \sum_{t=1}^2 \pi_t [\alpha v^L(a_t^L) + (1 - \alpha) v^H(a_t^H)],$$

where $\pi_1 \equiv 1/(1 + \rho)$ and $\pi_2 \equiv \rho/(1 + \rho)$. Similarly, the incentive constraint (3) and the budget constraints (1)-(2) can be written in the form

$$\sum_{t=1}^2 \pi_t [v^H(a_t^H) - v^H(a_t^L)] \geq 0 \quad \text{and} \quad \pi_t [y_t^L + y_t^H - c_t^L - c_t^H] \geq 0. \quad (24)$$

Clearly, the problem of maximizing $\hat{W}(a_1, a_2)$ subject to (24) is equivalent to problem FC. At the same time, it is equivalent to the generalized static problem GS with parameters $K = 2$ and $\pi_k^i = \pi_t$, except that the budget constraints in (24) must hold separately rather than in expectation as in (6). However, the budget constraints in (6) and (24) are equivalent as long as $a_1 = a_2$. Since assumption WDCRA implies that $a_1 = a_2$ is optimal in problem GS, the same allocation must also be optimal in problem FC. ■

Proof of Proposition 3

The proof proceeds in three steps. Steps 1-2 deal with problem NCS and show $W(NCS) > e^2/4$. Step 3 deals with problem PC and shows $W(PC) < e^2/4$. Taken together, these findings prove the claim. Note that equation (12) is used throughout the proof without further mention.

It turns out that problems NCS and PC can more easily be analyzed in terms of consumption c_t^i and leisure $f_t^i = e - l_t^i$ than in terms of c_t^i and y_t^i . It is straightforward to verify that the utility functions $v^i(\cdot)$ can be written in the form $v^i(c_t^i, y_t^i) = c_t^i f_t^i$, $v_t^H(c_t^L, y_t^L) = c_t^L(e(1 - w^L) - f_t^L)$, and that the budget constraints (1)-(2) are equivalent to

$$e(1 + w^L) - f_t^H - w^L f_t^L - c_t^L - c_t^H \geq 0, \quad t \in \{1, 2\}. \quad (25)$$

Welfare in period t will be denoted by $W_t \equiv \alpha c_t^L f_t^L + (1 - \alpha) c_t^H f_t^H$ and total welfare by $W = W_1 + \rho W_2$.

¹⁸Problem NCP2 is equivalent to the static income tax problem. It is well-known that $a_2^L = a_2^H$ is not optimal in this problem (see Stiglitz, 1982). Therefore, we have $a_1 \neq a_2$ in regime NCP. It is clear from equation (10) that $a_1 \neq a_2$ must also hold in regime NCS.

Step 1: Consider the first-best problem NCS2 which is equivalent to

$$\max_{c_2^i, f_2^i} W_2 = \alpha c_2^L f_2^L + (1 - \alpha) c_2^H f_2^H \quad \text{s.t.} \quad (25). \quad (26)$$

In addition to (25), the non-negativity constraints $c_2^i \geq 0$ and $0 \leq f_2^i \leq e$ have to be taken into account. It is clear that only allocations with $c_2^i f_2^i > 0$ or $c_2^i = f_2^i = 0$ can be optimal. Therefore, we can distinguish between the three cases $c_2^L f_2^L > 0, c_2^H = f_2^H = 0$ (Case 1), $c_2^L = f_2^L = 0, c_2^H f_2^H > 0$ (Case 2), and $c_2^L f_2^L > 0, c_2^H f_2^H > 0$ (Case 3). In Case 1, problem (26) is well behaved, except that the non-negativity constraint $f_2^L \leq e$ is binding. This implies $f_2^L = c_2^L = e$ and $W_2 = \alpha e^2$. In Case 2, problem (26) is also well-behaved, but maximum welfare is lower than in Case 1. In Case 3, problem (26) is degenerate because first-order conditions lead to a minimum and the constraint set is not closed. However, it can easily be verified that welfare is lower than in Case 1 for any feasible allocation of Case 3. Therefore, the unique solution to problem NCS2 is $f_2^H = c_2^H = 0$ and $f_2^L = c_2^L = e$ with second-period welfare $W_2 = \alpha e^2$.

Step 2: The solution to problem NCS2 implies $v_2^H(c_2^H, y_2^H) = 0$ and $v_2^H(c_2^L, y_2^L) = e^2$. The incentive constraint (3) of problem NCS1 can thus be written in the form

$$c_1^H f_1^H - c_1^L (e(1 - w^L) - f_1^L) \geq \rho e^2. \quad (27)$$

Hence, any allocation of period one that satisfies (27) and the budget constraint (25) is feasible. Note that the allocation $c_1^H = f_1^H = (e/2)(1 + w^L), c_1^L = f_1^L = 0$ is feasible. Hence, first-period welfare W_1 cannot fall below $(1 - \alpha)(e/2)^2(1 + w^L)^2 > (1 - \alpha)e^2\rho$. Because of $W_2 = \alpha e^2$, we thus have $W(NCS) > \rho e^2 \geq e^2/4$.

Step 3: Consider now problem PC. First, note that an allocation (a_1, a_2) with welfare $W_1 \neq W_2$ cannot be optimal because total welfare $W = W_1 + \rho W_2$ could be increased by implementing the allocation with higher W_t in both periods. Therefore, we can restrict attention to stationary allocations $a_1 = a_2$ without loss of generality. Next, note that $v^H(c_t^L, y_t^L) \geq v^L(c_t^L, y_t^L)$ must hold, which implies that any allocation that satisfies the incentive constraints (4)-(5) also satisfies the constraint $v^H(c_t^H, y_t^H) \geq v^L(c_t^L, y_t^L)$. Therefore, welfare in regime PC cannot be higher than in the solution to the relaxed problem

$$\max_{c_t^i, f_t^i} W = (1 + \rho)[\alpha c_t^L f_t^L + (1 - \alpha) c_t^H f_t^H] \quad \text{s.t.} \quad (25), \quad c_t^H f_t^H \geq c_t^L f_t^L.$$

This problem can be solved in the same way as problem NCS2 in Step 1. Due to the relaxed incentive constraint, however, the problem is well-behaved in all

three cases. It can easily be verified that maximum welfare is attained in Case 3 with solution $c_t^H = f_t^H = e(5/12)$, $c_t^L = e(5/24)$, $f_t^L = e(5/6)$ and welfare $W = (1 + \rho)e^2(5/12)^2 < e^2/4$. Therefore, $W(PC) < e^2/4$. ■

Proof of Proposition 4

Like in the deterministic case, any allocation that can be implemented in regimes PC or NC can also be implemented in regime FC. Therefore, $EW(FC) \geq EW(PC)$ and $EW(FC) \geq EW(NC)$. In the following, I will show $EW(PC) \geq EW(FC)$. Taken together, these findings prove the claim.

Let $\tilde{a} = (\tilde{a}_1, \tilde{a}_2)$ denote a solution to problem FC. I will first point out that this allocation can be implemented in problem PC as long as $\tilde{a}_1 = \tilde{a}_2$. Then I will show that we can restrict attention to allocations with $\tilde{a}_1 = \tilde{a}_2$ in problem FC without loss of generality. These two findings imply $EW(PC) \geq EW(FC)$.

Assume first $\tilde{a}_1 = \tilde{a}_2$. Then feasibility of \tilde{a} in problem FC implies that the constraints (13) and (15) of problem PC must hold. Moreover, since \tilde{a}_t maximizes $EW(\tilde{a}_1, \tilde{a}_2) = (1 + \rho)[\alpha Ev^L(\tilde{a}_t^L) + (1 - \alpha)Ev^H(\tilde{a}_t^H)]$, the allocation \tilde{a}_t is also time-consistent. Therefore, \tilde{a} can be implemented in problem PC.

Now assume $\tilde{a}_1 \neq \tilde{a}_2$ and consider another random allocation \bar{a} with $\bar{a}_1 = \bar{a}_2$. The elements \bar{a}_t^i of this allocation can be interpreted as the outcome of a two-stage lottery, where a random draw in the first stage determines whether the agent is assigned to lottery \tilde{a}_1^i or \tilde{a}_2^i in the second stage. The probabilities of the first-stage lottery are $\pi_1 = 1/(1 + \rho)$ and $\pi_2 = \rho/(1 + \rho)$. Since \bar{a}_t^i is a convex combination of \tilde{a}_1^i and \tilde{a}_2^i , we have

$$E\bar{a}_t^i = \pi_1 E\tilde{a}_1^i + \pi_2 E\tilde{a}_2^i \quad \text{and}$$

$$(1 + \rho)Ev^i(\bar{a}_t^i) = (1 + \rho)[\pi_1 Ev^i(\tilde{a}_1^i) + \pi_2 Ev^i(\tilde{a}_2^i)] = Ev^i(\bar{a}_1^i) + \rho Ev^i(\bar{a}_2^i)$$

These equalities and feasibility of \tilde{a} imply that \bar{a} is also feasible in problem FC. Moreover, both allocation lead to the same expected welfare $EW(\cdot)$. Therefore, we can restrict attention to random allocations with $\bar{a}_1^i = \bar{a}_2^i$ in problem FC without loss of generality. ■

Proof of Proposition 5

In the following, I will show $W(PC) \geq W(NCS)$. Because of Lemmas 1 and 2, this proves the claim. Consider a vector $(\hat{a}_1, \hat{a}_2, \hat{s})$ that can be implemented

in regime NCS. This implies that at least one of the inequalities $\hat{a}_1^H \neq \hat{a}_1^L$ and $\hat{s}^H \neq \hat{s}^L$ must hold. Assume first $\hat{s}^H \neq \hat{s}^L$. Then it is obvious from the definition of problems NCS and PCS that $(\hat{a}_1, \hat{a}_2, \hat{s})$ can also be implemented in regime PCS.

Next, I will show that any allocation (\hat{a}_1, \hat{a}_2) that can be implemented in regime NCS with savings $\hat{s}^H = \hat{s}^L$ can also be implemented in NCS with savings $\bar{s}^H \neq \bar{s}^L$. Because of the argument in the preceding paragraph, this implies that any allocation that can be implemented in regime NCS can also be implemented in regime PCS: The allocation (\hat{a}_1, \hat{a}_2) specifies income \hat{y}_1^i and consumption $\hat{c}_1^i = \hat{y}_1^i - \hat{s}^i - \hat{T}_1(\hat{y}_1^i, \hat{s}^i)$ of the two types in period one. Assume now that the planner offers two other tax contracts where the new variables $\bar{y}_1^L, \bar{s}^L, \bar{T}_1^L$ and \bar{y}_1^H are the same as before, but savings and taxation of the high skilled type are replaced by $\bar{s}^H = \hat{s}^H + \epsilon$ and $\bar{T}_1^H = \hat{T}_1^H - \epsilon$. Then we have $\bar{a}_1 = \hat{a}_1$ and aggregate savings $\hat{z} = \hat{s}^L + \hat{s}^H + \hat{s}^G$ remain constant. Therefore, the time-consistent policy $\hat{a}_2(\hat{z})$ will also not be affected. Each agent thus has the same options and the same incentive constraints as before. Accordingly, the initial allocation can also be implemented with savings $\bar{s}^H \neq \bar{s}^L$. ■

Proof of Proposition 6

Under the assumption $\rho R = 1$, problem FC is equivalent to the generalized static income tax problem GS for the special case $K = 2$ and $\pi_k^L = \pi_k^H$. (The proof of this claim is identical to the argument in the proof of Proposition 2 and can thus be omitted.) Since it is assumed that randomization is not desirable in problem GS, we thus have $a_1 = a_2$ in a solution to problem FC. Any allocation with $a_1 = a_2$ can be implemented with savings $s^L = s^H = s^G = 0$. Therefore, the solution to problem FC can also be implemented in regime PCP. Because of Lemma 1, this implies $W(FC) = W(PC)$. ■

Proof of Proposition 7

It is clear that the constraints of regime PC are more restrictive than the constraints of regime FCC. This implies $EW(FCC) \geq EW(PC)$.

Because of $E\tilde{\beta}_2 = \beta_1$ and (18), we have $E\bar{v}^i(a_2^i, g_2, \tilde{\beta}_2) = \bar{v}^i(a_2^i, g_2, \beta_1)$. This implies that problem FCI is equivalent to problem FC in Section 2, except that utility now also depends on government expenditures g_t . Strict concavity of $f(g_t)$ means that the agents are risk-averse with respect to a change in g_t and additive

separability between $v(a_t^i)$ and $f(g_t)$ implies that these expenditures do not affect the incentive compatibility constraints (21)-(23). By the same argument as in the proof of Proposition 2, assumption WDCRA thus implies that the solution to problem FCI is stationary such that $(a_1, g_1) = (a_2, g_2)$. In this case, constraint (21) and constraints (22)-(23) are equivalent. Hence, the solution to problem FCI can also be implemented in regime PC. This implies $EW(PC) \geq EW(FCI)$.

Note that $EW(PC) = EW(FCI)$ can only hold if a state-invariant allocation $(a_{2k}, g_{2k}) = (a_2, g_2)$ is optimal in problem PC. Note also that the maximization problem PC2 in state k is equivalent to the static income tax problem with optimal public-good provision g_{2k} and preference-parameter β_{2k} . Since utility is separable between (c_t^i, y_t^i) and g_t , the Samuelson-rule

$$\alpha \frac{\beta_{2k} f'(g_{2k})}{v_c^L(c_{2k}^L, y_{2k}^L)} + (1 - \alpha) \frac{\beta_{2k} f'(g_{2k})}{v_c^H(c_{2k}^H, y_{2k}^H)} = 1, \quad k = 1, \dots, K$$

holds in a solution to this problem (see Boadway and Keen, 1993). This condition implies that different allocations (a_{2k}, g_{2k}) are optimal for each realization β_{2k} . Therefore, we have $EW(PC) > EW(FCI)$. ■

References

- Acemoglu, D., Golosov, M., and A. Tsyvinski (2010) Dynamic Mirrlees taxation under political economy constraints, *Review of Economic Studies* 77, 841-881.
- Amador, M., Werning, I., and G.-M. Angeletos (2006) Commitment vs. flexibility, *Econometrica* 74, 365-396.
- Banks, J. and P. Diamond (2010) The Base for Direct Taxation, in: Mirrlees, J. et al. (eds.) *Dimensions of Tax Design: The Mirrlees Review*. Oxford: Oxford University Press, 548-648.
- Baron, D., and D. Besanko (1984) Regulation and information in a continuing relationship, *Information Economics and Policy* 1, 267-302.
- Battaglini, M., and S. Coate (2008) Pareto efficient income taxation with stochastic abilities, *Journal of Public Economics* 92, 844-868.
- Berliant, M., and J. O. Ledyard (2005) Optimal dynamic nonlinear income taxes with no commitment, Economics Working Paper Archive.
- Bisin, A., and A. A. Rampini (2006) Markets as beneficial constraints on the government, *Journal of Public Economics* 90, 601-629.

- Boadway, R. and M. Keen (1993) Public goods, self-selection, and optimal income taxation, *International Economic Review* 34, 463-478.
- Brett, C., and J. A. Weymark (2008) Optimal nonlinear taxation of income and savings without commitment, Working Paper No. 08-W05, Department of Economics, Vanderbilt University.
- Brito, D. L., J. H. Hamilton, S. M. Slutsky, and J. E. Stiglitz (1991) Dynamic optimal income taxation with government commitment, *Journal of Public Economics* 44, 15-35.
- Brito, D. L., J. H. Hamilton, S. M. Slutsky, and J. E. Stiglitz (1995) Randomization in optimal income tax schedules, *Journal of Public Economics* 56, 189-223.
- Christiansen, V. and M. Tuomala (2008) On taxing capital income with income shifting, *International Tax and Public Finance* 15, 527-545.
- Gaube, T. (2007) Optimum taxation of each year's income, *Journal of Public Economic Theory* 9, 127-150.
- Golosov, M., Tsyvinski, A. and I. Werning (2007) New dynamic public finance: A user's guide, In: *NBER Macroeconomics Annual 2006*. Cambridge, MA: MIT-Press, 317-363.
- Hellwig, M. (2007a) The undesirability of randomized income taxation under decreasing risk aversion, *Journal of Public Economics* 91, 791-816.
- Hellwig, M. (2007b) A contribution to the theory of optimal utilitarian income taxation, *Journal of Public Economics* 91, 1449-1477.
- Judd, K. (2007) Comment, In: *NBER Macroeconomics Annual 2006*. Cambridge, MA: MIT-Press, 381-384.
- Kapicka, M. (2006) Optimal income taxation with human capital accumulation and limited record keeping, *Review of Economic Dynamics* 9, 621-639.
- Kocherlakota, N. (2010) *The New Dynamic Public Finance*. Princeton: Princeton-University Press.
- Krause, A. (2009) Optimal nonlinear income taxation with learning-by-doing, *Journal of Public Economics* 93, 1098-1110.
- Laffont, J.-J., and J. Tirole (1993) *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.
- Mankiw, N.G., Weinzierl, M., and D. Yagan (2009) Optimal taxation in theory and practice, *Journal of Economic Perspectives* 23 (4), 147-174.
- Meade, J. E. (1978) *The Structure and Reform of Direct Taxation. Report of a Committee Chaired by Professor J.E. Meade*. London: George Allen & Unwin..

- Mirrlees, J. A. (1971) An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, 175-208.
- Roberts, K. (1984) The theoretical limits to redistribution, *Review of Economic Studies* 51, 177-195.
- Roëll, A. (1985) A note on the marginal tax in a finite economy, *Journal of Public Economics* 28, 267-272.
- Stiglitz, J. (1982) Self-selection and Pareto efficient taxation, *Journal of Public Economics* 17, 213-240.
- Tenhunen, S. and M. Tuomala (2010) On optimal lifetime redistribution policy, *Journal of Public Economic Theory* 12, 171-198.
- Vickrey, W. (1939) Averaging of income for income-tax purposes, *Journal of Political Economy* 47, 379-397.