Optimal dynamic nonlinear income taxes: facing an uncertain future with a sluggish government

Berliant, Marcus and Fujishima, Shota

Washington University in St. Louis, University of Tokyo

4 April 2014

Optimal Dynamic Nonlinear Income Taxes: Facing an Uncertain Future with a Sluggish Government*

Marcus Berliant† Shota Fujishima‡

April 4, 2014

Abstract

We consider the optimal nonlinear income taxation problem in a dynamic, stochastic environment when the government is sluggish in the sense that it cannot change the tax rule as uncertainty resolves. We show that the sluggish government cannot allow saving or borrowing regardless of the utility function. Moreover, we argue that the zero top marginal tax rate result in static models is of little practical importance because it is actually relevant only when the top earner in the initial period receives the highest shock in every period.

JEL classification: H21

Keywords: Optimal income taxation; New dynamic public finance

1 Introduction

Since the *New Dynamic Public Finance* was inaugurated, progress has been made in clarifying what the optimal dynamic nonlinear income tax looks like. This

*We are grateful to Eric Stephens, Yoshihiro Takamatsu, Jonathan Hamilton, Steven Slutsky, participants at the Southern Economic Association meeting in New Orleans and the Japanese Economic Association meeting at University of Toyama, and the seminar audience at Boston College for their helpful comments and suggestions. Shota Fujishima gratefully acknowledges travel support from the Center for Research in Economics and Strategy (CRES) in the Olin Business School, Washington University in St. Louis. Any remaining errors are our own.

†Department of Economics, Washington University in St. Louis, Campus Box 1208, One Brookings Drive, St. Louis, MO 63130-4899 USA. Phone: +1 314 935 8486, Fax: +1 314 935 4156, Email: berliant@artsci.wustl.edu

‡Center for Spatial Information Science, University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8568, Japan. Phone/Fax: +81 4 7136 4298, Email: sfujishima@csis.u-tokyo.ac.jp

agenda aims to extend the seminal work of Mirrlees (1971), who studies optimal income taxation in a static environment, to dynamic, stochastic environments.[1] Dynamic tax rules are in effect dynamic contracts because taxpayers have private information about their labor productivity, so the optimal dynamic income tax rule is generally complicated: it is nonstationary and depends on the entire history of income declared for any taxpayer. However, it is questionable whether governments can implement such complex tax rules because making tax rules time-dependent and tracking histories of income would entail large administrative and compliance costs. Indeed, neither of our governments (i.e., the US and Japanese governments) is tracking histories for income taxation.

In view of this observation, we contribute to the New Dynamic Public Finance literature by considering optimal dynamic income taxation when the government cannot change the tax rule over time. That is, the government can use only stationary tax rules. Our interpretation is that we must restrict our attention to a simple dynamic tax rule because our government is *sluggish*. Moreover, the stationarity of tax rule implies that the tax cannot depend on histories of income. Indeed, the sluggish government can look at only current incomes, just as it can only look at current incomes in the initial period. Naturally, we also assume that the sluggish government makes a full commitment to its tax system which is a collection of each period's tax rules. That is, the government cannot change the tax system once it is determined in the initial period. We are assuming that such commitment is not only possible, but perhaps unavoidable, due to political deadlock over the issue of tax policy, as in the US right now.[2] Thus, we may interpret our planner's problem

[1] See Kocherlakota (2010) for an overview of this literature.

[2] Indeed, the US government has not changed its income tax system in a major way since 1986. The Japanese government is more flexible, but it has not changed its income tax system in a major way since 2007. Therefore, once the tax systems are fixed, they persist for some time.

on a politician's short time-scale. Although our assumptions might be extreme, we believe that it is important and useful to have a sense about what the optimal dynamic income tax looks like when the set of tax systems is limited to ones that are feasible in practice.

We consider a finite horizon model in which the government would like to maximize the equal weight utilitarian social welfare function. Our economy is heterogeneous because we fix the type distribution in the initial period.[3] People receive idiosyncratic shocks in each period that are i.i.d. among people but otherwise, the stochastic structure is general.[4] Regarding intertemporal resource allocation, we assume that the government can save or borrow from an outside party so that it considers a single aggregate resource constraint. However, we *show* that the sluggish government cannot allow agents to save or borrow at all because allowing them to save or borrow requires the government to look at histories.

Although the analytical characterizations and even numerical analysis of the optimal dynamic tax system are difficult in general, we can analytically characterize the optimal tax system because our problem can be reduced to a static one due to the sluggishness of the government.[5] Specifically, this is because under a sluggish government, the tax rule depends on only the current income and the individual saving or borrowing is not allowed, so we can regard an agent living for T periods as distinct agents in each period and for each shock. Therefore, we can directly apply the arguments for static models to our model.

A famous result in the static optimal income taxation is that the top marginal

[3]If we do not fix it, the model has identical agents facing uncertainty, which is like a macro model. However, as long as we consider the equal weight utilitarian social welfare function, the distinction is not essential for the optimal tax rule as Farhi and Werning (2013) illustrate.

[4]In particular, the initial type and subsequent shocks can be correlated for each agent.

[5]Naturally, gaining tractability in this way widens the analytical insights about optimal dynamic income taxation we could derive. For example, if we assumed the quasi-linear utility, we could conduct comparative static analysis as in Weymark (1987).

tax rate is zero. That is, the top earner's marginal tax rate is zero. However, we cast doubt on its policy relevance. In our dynamic stochastic economy, the support of types will move over time, and a direct application of the static arguments implies that the marginal tax rate is zero at the top of the *expanded* type space, or the union of supports over time. Thus, if the largest value of the shock is positive, the zero top marginal tax result would apply only when the top earner in the initial period receives the largest possible value of shock in every period. One can argue that the fraction of people to which the zero top marginal tax rate result applies (i.e., the top earners) is negligible in static models. However, our result is much stronger than this. Indeed, whereas someone certainly faces the zero marginal tax rate in most static models anyway, it is not true in our dynamic, stochastic model. In fact, no one faces the zero marginal tax rate almost surely.

Our tax rule is stationary and therefore depends on only current income, so it would be a simple one in the literature. At the other extreme, Farhi and Werning (2013) , Battaglini and Coate (2008), and Kocherlakota (2005) study the most general rule by considering nonstationary tax rules that depend on the history of income. Whereas the stochastic structure of the shock is general in Kocherlakota (2005), Farhi and Werning (2013) and Battaglini and Coate (2008) consider Markov processes.[6] In the middle, Albanesi and Sleet (2006) study a nonstationary tax rule that depends on only current income when the shock is i.i.d. See Table 1 for a comparison between our work and others' work.[7]

[6]On the other hand, whereas Kocherlakota (2005) does not consider the time-consistency of a tax system, Battaglini and Coate (2008) provide conditions under which their tax system is time-consistent. In a two-period deterministic environment, Berliant and Ledyard (2014) study a tax rule that is nonstationary and depends on history while addressing time-consistency.

[7]In a two-period deterministic environment, Gaube (2010) compares three types of nonstationary income taxation: a tax rule depends on history and the resulting tax system is time-consistent; a tax rule depends on history but the resulting tax system is not time-consistent; and a tax rule depends on only current income (but can change over time).

Table 1: The position of this paper in the literature

	Shock	History	Stationary	Commitment
Farhi and Werning (2013)	Markov	Yes	No	Yes
Battaglini and Coate (2008)	Markov	Yes	No	Yes/No
Kocherlakota (2005)	General	Yes	No	Yes
Albanesi and Sleet (2006)	i.i.d.	No	No	Yes
This paper	General	No	Yes	Yes

The rest of this paper proceeds as follows. In Section 2, we state the basic structure of the model, present our problem, and characterize the second-best tax rule. Section 3 contains our conclusions and discusses subjects for future research. Proofs omitted from the main text are provided in an Appendix.

2 The Model

We consider a finite horizon model with a unit mass of agents. The economy lasts for $T+1$ periods. In period 0, each agent is endowed with type $w \in W_0 \subseteq \mathbb{R}_{++}$ distributed with density function f_w. However, there are idiosyncratic shocks to the agents' types in the subsequent periods. At the beginning of period 1, an element of $z^T = \{z_t\}_{t=1}^T \in Z^T$ is drawn for agent with type w according to a density function $f(z^T \,|\, w)$. We assume that W_0 and $Z \subseteq \mathbb{R}$ are (non-degenerate) closed intervals. Note that, although shocks for all T periods are drawn in the initial period, the agent only learns them as time goes on, so that in period t the agent observes the history $(w, z_1, ..., z_t)$. If an agent is endowed with type w in the initial period, his type will change to $w_t \equiv \phi_t(w; z^T)$ in period t. For example, if we consider a linear technology, $\phi_t(w; z^T) = w + \sum_{s=1}^t z_s$. We assume that $\phi_t(\,\cdot\,;\,\cdot\,)$ is continuously differentiable and for any $w \in W_0$, $\phi_t(w; z^T) > 0$ for all $z^T \in Z^T$ in any period. Moreover, we assume

$W_{t-1} \cap W_t \neq \emptyset$ for any $t \geq 1$ where W_t is the range of $\phi_t(\,\cdot\,;\,\cdot\,)$.[8] As long as $\phi_t(w; \mathbf{0}) = w$, $0 \in Z$ is sufficient for this. Finally, we assume that the draws are i.i.d. among agents and the law of large numbers holds.[9] Thus, letting $f(z^T, w) = f(z^T \mid w) f_w(w)$, the joint density of z^T and w, denote the density of agents having type w in the initial period and getting shock z^T.

The agents supply labor and consume the good produced under constant returns to scale in each period. As is usual in optimal taxation models, they face a trade-off between consumption and leisure. The utility function is

$$U\left(\{c_t, \ell_t\}_{t=0}^{T}\right) = \sum_{t=0}^{T} \rho^t u(c_t, \ell_t) \tag{1}$$

where $\ell \in [0, 1]$ is labor, c is consumption, and $\rho > 0$ is the discount factor. We assume that $u(c, \ell)$ is twice continuously differentiable, strictly concave, increasing in c, and decreasing in ℓ. Moreover, we assume that leisure $1 - \ell$ is a noninferior good.[10] In our model, type represents the earning ability of agents. That is, if the labor supply of agent w is ℓ, his gross income is given by $y = w\ell$.

We suppose that there is a risk-free bond market with the interest rate $R > 0$ where b is the bond holding. Then, letting τ be a (lump-sum) component of an income tax, agents' budget constraint in period t is

$$c_t + b_{t+1} = y_t - \tau_t + (1 + R) b_t. \tag{2}$$

---

[8] Because the range of ϕ_t is in $\mathbb{R}$, ϕ_t is continuous, and $W_0 \times Z^T$ is connected and compact, W_t is a closed interval. We further assume that W_t is non-degenerate.

[9] Kocherlakota (2005), for example, also makes these assumptions. Regarding the law of large numbers, there are some technical issues for the case of *continuum* of i.i.d. random variables (Judd, 1985). However, Sun (2006) provides a solution to this issue by presenting a probability space in which the law of large numbers holds.

[10] Hellwig (2007) presents another assumption that is a cardinal property of u instead of the assumption that leisure is a noninferior good, which is an ordinal property.

The government would like to maximize social welfare. In this paper, we consider the following utilitarian social welfare function:

$$SW = \int_{W_0} \int_{Z^T} U\left(\{c_t, y_t/w_t\}_{t=0}^T\right) f(z^T, w) \mathrm{d}z^T \mathrm{d}w. \tag{3}$$

Since the one-period utility function is strictly concave and leisure is a noninferior good, it follows that redistribution is desirable under the utilitarian welfare function (Seade, 1982). The planner would like to carry out redistribution through income taxes, but he cannot observe the agents' types. Thus, the government needs to design a mechanism that makes the agents reveal their true types.

We consider a direct mechanism in which agents report their types and the government specifies the combination of consumption c, gross income y, and bond holding b for each report in each period. Specifically, we call $x_t(\cdot) \equiv (c_t(\cdot), y_t(\cdot), b_{t+1}(\cdot))$ an *allocation rule*. In general, the rule could be nonstationary and depend on histories of reports as in Battaglini and Coate (2008). However, because our planner is sluggish, he cannot enforce nonstationary tax rules. Moreover, as a consequence, he looks at only current reports, because the domain of his tax rule must be time-dependent if he looks at history. Therefore, we restrict our attention to the allocation rule that is time-invariant (i.e., $x_t(\cdot) = x(\cdot)$ for all t) and does not depend on history (or it depends on only the current report).[11] For example, if agent reports w_t in period t, his allocation in that period is $x(w_t) = (c(w_t), y(w_t), b(w_t))$.

Although we have included bond holding in the mechanism above, there is an important result regarding to what extent the sluggish government can control individual saving or borrowing. Note that, as far as a sluggish government is concerned, the tax rule, which is induced by an allocation rule via (2), also must

[11] We note that the government is aware that it is sluggish, so once it chooses its allocation rule, it knows the rule cannot be changed, and accounts for this when choosing the rule.

be stationary and history-independent. However, this will imply that the sluggish government cannot allow saving or borrowing:

Proposition 1. *Suppose that W_t is a non-degenerate closed interval and $W_{t+1} \cap W_t \neq \emptyset$ for any $t \geq 0$. Then, the sluggish government cannot allow agents to save or borrow. That is, $b(w) = 0$ for all $w \in W \equiv \bigcup_{t=0}^{T} W_t$.*

The proof can be found in the Appendix. The logic behind this result is pretty simple. Suppose that, in period $t-1$, there are agents having types w_{t-1} and w'_{t-1} respectively, and both of them become type w_t in period t. By the budget constraint, the period-t tax on the agent who changes from w_{t-1} to w_t is $\tau(w_t) = y(w_t) - c(w_t) - b(w_t) + (1+R)b(w_{t-1})$. Because the government is sluggish, the tax on the agent who changes from w'_{t-1} to w_t is also $\tau(w_t)$, but the budget constraint implies that $\tau(w_t) = y(w_t) - c(w_t) - b(w_t) + (1+R)b(w'_{t-1})$. Thus, we must have $b(w_{t-1}) = b(w'_{t-1})$. Then, because $b(w_T) = 0$ by the terminal condition, we obtain the result by induction. Note that this argument does not depend on the utility function. In particular, Proposition 1 holds regardless of people's risk attitude (i.e., whether they are risk-neutral or risk-averse).[12]

There are two remarks. First, because our revenue constraint is integrated over time as we will see, the government can save and borrow for the agents. However, the sluggishness of government leaves no room for saving and borrowing not because the government borrows and saves for the consumers, but because it cannot actually address the intertemporal wedge.[13] Indeed, Farhi and Werning (2013) also have a revenue constraint integrated over time, but according to their simulations, bond holdings are not zero. Second, we can see that the stochastic

[12]Note that, if we assume people are risk-neutral as in Battaglini and Coate (2008), the utility function is quasi-linear.

[13]The intertemporal wedge is related to Euler equation, or intertemporal substitution. See, for example, Kocherlakota (2004).

shocks are important for the argument above. Due to the shocks, each state can be reached by several agents who generally have different histories. When the government would like to address the intertemporal wedge, it is impossible for the government to take care of these agents' situations simultaneously due to its sluggishness.[14]

It might be more straightforward to consider an indirect mechanism in which the agents report their incomes and the government specifies income taxes for each report. However, now that the bond holding cannot be allowed, it readily follows that Hammond's (1979) result applies to our problem because, as we will see, our problem reduces to a static one. That is, characterizing the direct mechanism is equivalent to designing a tax rule $\tau(\cdot)$ and letting each agent choose his income y_t and consumption $c_t = y_t - \tau(y_t)$.

In view of Proposition 1, we henceforth drop the bond holding and let $x(\cdot) = (c(\cdot), y(\cdot))$. Since the planner cannot observe the agents' types, he faces incentive compatibility (IC) constraints that require that the agents do not misreport their types. Let $v(x(w'), w) = u(c(w'), y(w')/w)$. This is the one-period utility that agent w obtains when he reports w'. Since the agents report their types in each period, the IC constraints are imposed in each period. Recall that W_t is the range of $\phi_t(\cdot\,;\cdot)$ for $t \geq 1$. Then, the IC constraint in the last period is given by

$$\forall w \in W_T,\ v\,(x(w), w) \geq v\,(x(w'), w) \ \text{ for all } w' \in W_T. \tag{IC$_T$}$$

On the other hand, the IC constraint in period $t \in \{0, 1, 2, ..., T-1\}$ is given by

[14] Kapicka (2006) studies optimal income taxation in a dynamic, *deterministic* model where people can allocate their time to human capital investment. Focusing on steady states, the government can specify (constant) investment levels for each agent even though it is sluggish.

$$\forall w \in W_t,\ v(x(w), w) + \sum_{s \geq t+1} \rho^s \int_{Z^T} \max_{\tilde{w} \in W_s} v(x(\tilde{w}), w_s) f(z^T \mid w) \mathrm{d}z^T$$
$$\geq v(x(w'), w) + \sum_{s \geq t+1} \rho^s \int_{Z^T} \max_{\tilde{w} \in W_s} v(x(\tilde{w}), w_s) f(z^T \mid w) \mathrm{d}z^T \text{ for all } w' \in W_t \quad (\mathrm{IC}_t)$$

where $w_s = \phi_s(w; z^T)$. Since our mechanism does not depend on history, the report in the current period does not affect the expected continuation payoff (i.e., $\sum_{s \geq t+1} \rho^s \int_{Z^T} \max_{\tilde{w} \in W_s} v(x(\tilde{w}), w_s) f(z^T \mid w) \mathrm{d}z^T$ does not depend on the report in period t). As a result, the IC constraint in period t reduces to

$$\forall w \in W_t,\ v(x(w), w) \geq v(x(w'), w) \text{ for all } w' \in W_t.$$

In addition to the IC constraints, the government faces a resource constraint: it needs to finance G in units of consumption good through the tax. This revenue could be used for a public good that is fixed in quantity (and thus in cost) or the public good could enter utility as an additively separable term. We assume that the government can borrow or save at rate ρ. Then, the government faces the following resource constraint (RC):

$$G \leq \int_{W_0} (y(w) - c(w)) f_w(w) \mathrm{d}w + \sum_{t=1}^{T} \rho^t \int_{W_0} \int_{Z^T} (y(w_t) - c(w_t)) f(z^T, w) \mathrm{d}z^T \mathrm{d}w. \quad (\mathrm{RC})$$

Suppose that an agent is endowed with type w in the initial period. Let

$$V(x(\cdot), w) = \int_{Z^T} U\left(\{c(w_t), y(w_t)/w_t\}_{t=0}^{T} \right) f(z^T \mid w) \mathrm{d}z^T. \quad (4)$$

where $w_t = \phi_t(w; z^T)$. This is the expected lifetime utility that the agent obtains by reporting truthfully in each period. Then, the planner's problem is given by

$$\begin{aligned} &\max_{x(\cdot)} \int_{W_0} V(x(\cdot), w) f_w(w) \mathrm{d}w \\ &\text{s.t.} \quad \text{(RC) and (IC}_t\text{) for all } t. \end{aligned} \tag{5}$$

For reference, the *first-best allocation rule* $x^*(\cdot)$ maximizes the utilitarian welfare function subject to the resource constraint only, assuming that the government knows the type of each agent at each time.

Let $W = \bigcup_{t=0}^{T} W_t$. In what follows, we make the following assumption on the one-period utility function u, in addition to the regularity conditions stated before:

Assumption 1 (Spence-Mirrlees single crossing property: SCP)**.** *$\forall (c, y, w) \in \mathbb{R}_+^2 \times W$, $-wu_c(c, y/w)/u_\ell(c, y/w)$ is increasing in w.*[15]

Here is the key idea of our work. When we solve the problem (5), we exploit the fact that our mechanism is time-invariant and does not depend on history, and we consider a time-separable utility function and the utilitarian social welfare function. Therefore, the problem can be reduced to a static problem in which the total mass of agents is expanded to $\sum_{t=0}^{T} \rho^t$. That is, each person in each period is considered to be a different person in the static model. Utilitarianism with the time-separable utility gives us the equivalence. Then, we take the standard approach for static optimal income taxation problems to solve (5). That is, we consider a relaxed problem in which the IC constraints are replaced with weaker conditions and invoke the fact that a solution to the relaxed problem is also a solution to the original problem under Assumption 1.[16]

[15]This assumption is equivalent to assuming that the consumption good is a normal good. See p. 182 of Mirrlees (1971).

[16]This argument crucially depends on the fact that mechanism is static. Otherwise, general assumptions like the single crossing property that connect the relaxed problem to the original one are not known (Farhi and Werning, 2013).

Before stating our main result, let us summarize the regularity conditions we have imposed:

Assumption 2 (Regularity conditions)**.**

1. *$W_0 \subset \mathbb{R}_{++}$ and $Z \subset \mathbb{R}$ are non-degenerate closed intervals;*
2. *ϕ_t is continuously differentiable; for any $w \in W_0$, $\phi_t(w; z^T) > 0$ for all $z^T \in Z^T$ in any period; $W_{t-1} \cap W_t \neq \emptyset$ and W_t is non-degenerate for any $t \geq 1$;*
3. *$u(c, \ell)$ is twice continuously differentiable, strictly concave, increasing in c, and decreasing in ℓ; leisure $1 - \ell$ is a noninferior good;*

Let $\underline{w} = \min W$ and $\overline{w} = \max W$ (thus, $W = [\underline{w}, \overline{w}]$). Moreover, recall that $x^*(\cdot)$ is the first-best allocation rule that maximizes social welfare subject to the resource constraint. Then, the main properties of the planner's allocation rule are summarized in the following proposition.

Proposition 2. *Under Assumptions 1 and 2, (i) $x(w) \leq x^*(w)$ for any $w \in W$ with equality at $w = \overline{w}$. If $y(w)$ is strictly increasing at $w = \underline{w}$, $x(\underline{w}) = x^*(\underline{w})$. Moreover, if $y(w) > 0$, then $x(w) \ll x^*(w)$ for any $w \in (\underline{w}, \overline{w})$; (ii) $\tau'(y(\overline{w})) = 0$ and if $y(w)$ is strictly increasing at $w = \underline{w}$, $\tau'(y(\underline{w})) = 0$. Moreover, if $y(w) > 0$, then $\tau'(y(w)) \in (0, 1)$ for any $w \in W$.*

The proof can be found in the Appendix. Property (i) states that the allocation is first-best at the top of W and if income is strictly increasing at the bottom of W, the allocation is also first-best there. In addition, no allocation can be distorted upward from the first-best allocation and in particular, if income is positive, the allocation is distorted downward from the first-best allocation in the interior of W. Property (ii) states that the marginal tax rate is zero at the top of W and if income is strictly increasing at the bottom of W, the marginal tax rate is also zero there. On the other hand, if income is positive, the marginal tax rate is more than 0 but less then 1 in the interior of W.

By Proposition 2, as long as everyone works so that $y(w) > 0$ for all $w \in W$, the allocation is generally first-best and the marginal tax rate is zero only at the top of the *expanded* type space W. For illustration, suppose $\phi_t(w; z^T) = w + \sum_{s=1}^{t} z_s$. Then, *if* $\max Z > 0$, *no one's allocation is generally first-best and no one's marginal tax rate is zero in the first T periods nor the last period except when the type of the top earner in the initial period reaches* $\overline{w} = \max W$. In practice, it is unlikely that the planner sets the marginal tax rate at the ex post top to zero because he does so only when the shock to the top earner in the initial period takes the largest value in every period. One can argue that the zero top marginal tax result is relevant to only a small fraction of people (i.e., the top earner) in static model. However, our result is much stronger than this. Indeed, whereas someone certainly faces the zero marginal tax rate in a static economy because the top earner necessarily exists, that is not true here. As in a static model, an agent faces the zero marginal tax rate when he attains the highest possible income, but this very top earner does not always exist ex post in our dynamic, stochastic economy.

Moreover, the results above are in sharp contrast with those of Battaglini and Coate (2008) in which the shock follows a Markov chain over two states (high and low). In their first-best tax rule, the allocation is distorted only when people's type is currently and has always been low. That is, the allocations of agents who are currently, or have at some point been high types are first-best. Therefore, the fraction of people whose allocations are distorted is decreasing over time. However, their results crucially depend on the following facts: the support of types is fixed over time, and the tax rule can depend on history. In our model, the support of types *moves* over time, and the tax rule can depend on only the current income. As a result, all people's allocations are almost surely distorted in any period.

3 Conclusion

We considered the optimal dynamic income taxation problem faced by a sluggish government that cannot change the tax rule over time. Because of the government's sluggishness, we could reduce our problem to a static one and analytically characterize the second-best tax rule. We argued that the zero top marginal tax result is of little importance in practice because it would apply only when the top earner in the initial period receives the largest value of shock in every period. This is a probability zero event, so ex post we ensure a positive tax rate for the top type.[17]

Regarding the sluggishness of the government, we have made an extreme assumption: the government cannot make its tax rule time-dependent and thus its tax rates cannot be history-dependent *at all*. It might be more realistic to consider the situation in which the government can make its tax rule time-dependent or look at past histories at some cost.

Moreover, because we considered i.i.d. idiosyncratic shocks, we could obtain the single resource constraint by invoking the law of large numbers. Besides the sluggishness of the government, this was also crucial for our results. In fact, if the agents face the common aggregate shocks, their types are correlated with each other, and the analytical approach of this paper will fail to apply. These should be subjects of future research.

Finally, although we characterized an optimal tax rule, we did not address its existence. This can probably be proved using the results of Berliant and Page (2001) for static optimal taxes.

[17]In this paper, we consider a finite-horizon model. Technically speaking, we use optimal control theory, so by replacing terminal conditions with transversality conditions, we would be able to extend both Propositions 1 and 2 to an infinite-horizon model.

Appendix

Proof of Proposition 1. We argue by backward induction. First, $b(W_T) = \{0\}$ by the terminal condition (i.e., in the last period, no one will save and borrowing is not permitted because people cannot repay). Now, suppose $b(W_t) = \{0\}$ where $t \leq T$, and let $w_{t-1} \in W_{t-1}$. Without loss of generality, suppose $\min W_t \leq \max W_{t-1}$. Because $W_t \cap W_{t-1} \neq \emptyset$, and W_t and W_{t-1} are closed intervals, $b(w_{t-1}) = 0$ if $w_{t-1} \geq \min W_t$ by assumption. Thus, we assume $w_{t-1} < \min W_t$. Take $w'_{t-1} > w_{t-1}$ in a neighborhood of w_{t-1} with diameter less than the length of Z. By the Intermediate Value Theorem, we can take $z_t, z'_t \in Z$ such that $w_t \equiv w_{t-1} + z_t = w'_{t-1} + z'_t \equiv w'_t$.[18] Then, because the sluggish government's tax can depend on only the current state, $\tau(w_t) = y(w_t) - c(w_t) - b(w_t) + (1 + R)b(w_{t-1})$ and $\tau(w_t) = y(w_t) - c(w_t) - b(w_t) + (1 + R)b(w'_{t-1})$, and thus $b(w_{t-1}) = b(w'_{t-1})$. Therefore, $b(w_{t-1})$ is constant in (the upper half of) its neighborhood. Because w_{t-1} ($< \min W_t$) is arbitrary and $b(\hat{w}_{t-1}) = 0$ for $\hat{w}_{t-1} \geq \min W_t$, it follows that $b(W_{t-1}) = \{0\}$. □

Proof of Proposition 2. We show that due to the sluggishness of the government, our problem can be reduced to a static problem and then invoke the results of Hellwig (2007) who analyzes a static optimal taxation problem under the utilitarian welfare function. As in Hellwig (2007), we consider a relaxed problem by replacing the IC constraint with a weaker condition that is called the *downward IC constraint*:

$$\forall w \in W_t,\ v\left(x(w), w\right) \geq v\left(x(w'), w\right) \text{ for all } w' \in \{\tilde{w} \in W_t : \tilde{w} \leq w\}. \qquad (\text{IC}'_t)$$

for each t. Thus, the downward IC constraint takes care of only downward devia-

[18] Because the diameter of the neighborhood is smaller than the length of Z, $w'_{t-1} + \min Z < w_{t-1} + \max Z$. Thus, for $w_t \in [w'_{t-1} + \min Z, w_{t-1} + \max Z]$, we can take $z_t, z'_t \in Z$ such that $w_t = w_{t-1} + z_t$ and $w_t = w'_{t-1} + z'_t$ respectively.

tions. By Lemma 6.2 of Hellwig (2007), $x(\cdot)$ with nondecreasing $c(\cdot)$ satisfies (IC_t') if and only if $\frac{\mathrm{d}v(x(w),w)}{\mathrm{d}w} \geq v_w(x(w), w)$ for all $w \in W_t$. Thus, when we solve the problem, we impose the constraints that $c(w)$ is nondecreasing and $\frac{\mathrm{d}v(x(w),w)}{\mathrm{d}w} \geq v_w(x(w), w)$ on $W = \bigcup_{t=0}^{T} W_t$ instead of the downward IC constraints.

Next, we rewrite the welfare function as

$$
\begin{aligned}
&\int_{W_0} V(x(\cdot), w) f_w(w) \mathrm{d}w \\
&= \int_{W_0} \left[v(x(w), w) + \sum_{t=1}^{T} \rho^t \int_{Z^T} v(x(w_t), w_t) f(z^T \mid w) \mathrm{d}z^T \right] f_w(w) \mathrm{d}w \\
&= \int_{W_0} v(x(w), w) f_w(w) \mathrm{d}w + \sum_{t=1}^{T} \rho^t \int_{W_t} v(x(w_t), w_t) f_t(w_t) \mathrm{d}w_t
\end{aligned}
$$

where $f_t(w) = \int_{Z^T} f(z^T, \phi_t^{-1}(w; z^T)) \frac{\mathrm{d}\phi_t^{-1}(w;z^T)}{\mathrm{d}w} \mathrm{d}z^T$. Let $\overline{f_w}$ be an extension of f_w to W (i.e., $\overline{f_w}(w) = f_w(w)$ on W_0 and $\overline{f_w}(w) = 0$ on $W \setminus W_0$). Similarly, let $\overline{f_t}$ be an extension of f_t to W. Then, the above expression reduces to

$$
\int_W v(x(w), w) g(w) \mathrm{d}w \tag{6}
$$

where $g(w) \equiv \overline{f_w}(w) + \sum_{t=1}^{T} \rho^t \overline{f_t}(w)$. Likewise, the resource constraint is reduced to

$$
G \leq \int_W \tau(w) g(w) \mathrm{d}w. \tag{7}
$$

Therefore, our relaxed problem is given by

$$
\begin{aligned}
&\max_{x(\cdot)} \int_W v(x(w), w) g(w) \mathrm{d}w \\
&\text{s.t.} \quad G \leq \int_W \tau(w) g(w) \mathrm{d}w, \\
&\qquad c(w) \text{ is nondecreasing and } \frac{\mathrm{d}v(x(w),w)}{\mathrm{d}w} \geq v_w(x(w), w) \text{ on } W.
\end{aligned} \tag{8}
$$

On the other hand, Hellwig (2007) considers a standard static optimal taxation problem. Specifically, in our notations, his problem is written as

$$\begin{aligned} &\max_{x(\cdot)} \int_{W_0} v(x(w), w) f_w(w) \mathrm{d}w \\ &\text{s.t.} \quad G \leq \int_{W_0} \tau(w) f_w(w) \mathrm{d}w, \\ &\qquad c(w) \text{ is nondecreasing and } \tfrac{\mathrm{d}v(x(w),w)}{\mathrm{d}w} \geq v_w(x(w), w) \text{ on } W_0. \end{aligned} \tag{9}$$

Hence, we can see that our problem can be viewed as a static problem in which the total mass of agents is $\sum_{t=0}^{T} \rho^t$, the support of type distribution is W, and the welfare weight for type w is $g(w)$, and therefore, the arguments of Hellwig (2007) directly apply. In particular, the property (i) follows from Theorem 6.1 and the property (ii) from Theorems 4.1 and 6.1 of Hellwig (2007). □

References

Albanesi, S. and C. Sleet. (2006), "Dynamic optimal taxation with private information," *Review of Economic Studies*, **73**, 1-30.

Battaglini, M. and S. Coate. (2008), "Pareto efficient income taxation with stochastic abilities," *Journal of Public Economics*, **92**, 844-868.

Berliant, M. and J. Ledyard. (2014), "Optimal dynamic nonlinear income taxes with no commitment," *Journal of Public Economic Theory*, **16**, 196-221.

Berliant, M. and F. H. Page. (2001), "Income taxes and the provision of public goods: existence of an optimum," *Econometrica*, **69**, 771-784.

Farhi, E. and I. Werning. (2013), "Insurance and taxation over the life cycle," *Review of Economic Studies*, **80**, 596-635.

Gaube, T. (2010), "Taxation of annual income as a commitment device," Unpublished manuscript.

Hammond, P. (1979), "Straightforward individual incentive compatibility in large economies," *Review of Economic Studies*, **46**, 263-282.

Hellwig, M. F. (2007), "A contribution to the theory of optimal utilitarian income taxation," *Journal of Public Economics*, **91**, 1449-1477.

Judd, K. L. (1985), "The law of large numbers with a continuum of i.i.d. random variables," *Journal of Economic Theory*, **35**, 19-25.

Kapicka, M. (2006), "Optimal income taxation with human capital accumulation and limited record keeping," *Review of Economic Dynamics*, **9**, 612-639.

Kocherlakota, N. (2010), *The New Dynamic Public Finance*, Princeton University Press.

Kocherlakota, N. (2005), "Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation," *Econometrica*, **73**, 1587-1621.

Kocherlakota, N. (2004), "Wedges and taxes," *American Economic Review*, **94**, 109-113.

Mirrlees, J. A. (1971), "An exploration in the theory of optimal income taxation," *Review of Economic Studies*, **38**, 175-208.

Seade, J. K. (1982), "On the sign of the optimum marginal income tax," *Review of Economic Studies*, **49**, 637-643.

Sun, Y. (2006), "The exact law of large numbers via Fubini extension and characterization of insurable risks," *Journal of Economic Theory*, **126**, 31-69.

Weymark, J. A. (1987), "Comparative static properties of optimal nonlinear income taxes," *Econometrica*, **55**, 1165-1185.