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

17 May 2013

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

Marcus Berliant† Shota Fujishima‡

May 17, 2013

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 argue that the zero top marginal tax rate result in static models is of little practical importance because it actually holds 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 agenda aims to extend the seminal work of Mirrlees (1971), who studies optimal income taxation in a static environment, to dynamic, stochastic environments.[1]

*We are grateful to Eric Stephens, Jonathan Hamilton, Steven Slutsky, participants at the Southern Economic Association meeting in New Orleans, and the seminar audience at Boston College for their helpful comments and suggestions. Shota Fujishima gratefully acknowledges travel support from 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`

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

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 in the real world can implement such complex tax rules because changing the tax rule frequently and tracking histories of income would entail large administrative and political costs. Indeed, regarding the stationarity of the tax rule, 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 rules are fixed, they persist for some time.

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. Our interpretation is that we must restrict our attention to a simple dynamic tax rule because our government is *sluggish*. Naturally, we also assume that the sluggish government makes a full commitment to its tax rule. That is, the government cannot change the tax rule once it is determined in the initial period. Moreover, because the length of history is time-dependent, the sluggish government can look at only current incomes, just as it can only look at current incomes in the initial period.[2] 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 rules is limited to ones that are feasible in practice. *Our motivation is to develop a tractable, realistic model of dynamic income taxation.*

We consider a finite horizon model in which the government would like to maximize the equal weight utilitalian 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] We assume that the government confiscates all

[2]Technically speaking, this is explained as follows: If a tax function depends non-trivially on an individual's entire income history, its domain is time-dependent, because the length of income history is time-dependent. Therefore it is necessarily nonstationary.

[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 utilitalian social welfare function, the distinction is not essential for the optimal tax rule as Farhi and Werning (forthcoming) illustrate.

[4]In Section 2.2, we argue that focusing on idiosyncratic shocks is without loss of generality. We

incomes that are not consumed in each period, so agents cannot save nor borrow.[5] In fact, we show that a sluggish government cannot allow saving or borrowing as long as it would like to address the inherent incentive problems. Although the analytical characterizations and even numerical analysis of the optimal dynamic tax rule are difficult in general, we can analytically characterize the optimal tax rule because our problem can be reduced to a static one due to the sluggishness of the government. Specifically, this is because under a sluggish government, the tax rule depends on only the current income, so we can regard an agent living for T periods as distinct agents in each period and for each shock if we do not allow saving or borrowing. 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 tax rate is zero. That is, the top earner's marginal tax rate is zero. However, we cast doubt on its policy relevance as Diamond and Saez (2011) do.[6] 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. Therefore, the event that the marginal tax rate is zero at the ex post top actually has probability zero.

Our tax rule is fixed over time and therefore depends on only current income, so it would be a simple one in the literature. At the other extreme, Farhi and Werning (forthcoming) , 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 (forthcoming) and Battaglini and Coate (2008) consider Markov processes.[7] In the middle, Albanesi and Sleet (2006) study a nonstationary

could also consider aggregate shocks.

[5]This assumption is not uncommon in the literature. For example, it is also made by Battaglini and Coate (2008) who study optimal income taxation in a dynamic, stochastic environment and Brito *et al*. (1991) who study it in a dynamic, deterministic economy.

[6]The source of their claim is different from ours, though. In Diamond and Saez (2011), the source is a zero marginal social welfare weight on the utility of high income earners.

[7]On the other hand, whereas Kocherlakota (2005) does not consider the time-consistency of a tax rule, Battaglini and Coate (2008) provide conditions under which their rule is time-consistent. In a

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.[8]

Table 1: The position of this paper in the literature

	Shock	History	Stationary	Commitment
Farhi and Werning (forthcoming)	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.[9] At the beginning of period 1, an element of $z^T = \{z_t\}_{t=1}^T \in Z^T$ is drawn for each agent according to a density function f_z. 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 where $\phi_t(\cdot\,;\cdot)$ is continuously differentiable. For example, if we consider a linear technology, $\phi_t(w; z^T) = w + \sum_{s=1}^t z_s$. We assume

two-period deterministic environment, Berliant and Ledyard (forthcoming) study a time-consistent tax rule that is nonstationary and depends on history.

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

[9]We can add a public or private aggregate shock without changing the main result. See Section 2.2.

that 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 $\phi_t(\cdot\,;\cdot)$.[10] 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 law of large number holds.[11] Thus, $f_z(z^T) f_w(w)$ denotes 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.[12] 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 $y = w\ell$. Temporarily, let there be no taxes. Then, assuming that the agents cannot save, their budget constraint in period t is $c_t = y_t$ since u is strictly increasing in c.

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

$$SW = \int_{W_0} \int_{Z^T} U\left(\{c_t, y_t/w_t\}_{t=0}^{T}\right) f_z(z^T) \mathrm{d}z^T f_w(w) \mathrm{d}w. \tag{2}$$

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.

[10]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.

[11]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 number holds.

[12]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.

We consider a direct mechanism in which agents report their types and the government specifies the pair of consumption c and gross income y for each report in each period. Specifically, we call $x_t(\cdot) \equiv (c_t(\cdot), y_t(\cdot))$ 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 complex rules that vary over time. 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).[13] For example, if agent reports w_t in period t, his allocation in that period is $x(w_t) = (c(w_t), y(w_t))$.

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, 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 $c_t = y_t - \tau(y_t)$.

Since the planner cannot observe the agents' types, he faces incentive compatibility (IC) constraints that require that the agents do not misreport their types. Recall that our allocation rule is independent of time (i.e., $x_t(\cdot) = x(\cdot)$ for all t). Thus, we omit the time subscripts and write our allocation rule as $x(\cdot) = (c(\cdot), y(\cdot))$. Let $u(x(w'), w) = u(c(w'), y(w')/w)$. 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. Let W_t be the range of $\phi_t(\cdot\,;\cdot)$ for $t \geq 1$. Then, the IC constraint in the last period is given by

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

On the other hand, assuming that the agents cannot save, the IC constraint in period $t \in \{0, 1, 2, ..., T-1\}$ is given by

$$\forall w \in W_t,\ u(x(w), w) + \sum_{s \geq t+1} \rho^s \int_{Z^T} \max_{\tilde{w} \in W_s} u(x(\tilde{w}), w_s) f_z(z^T) \mathrm{d}z^T$$

[13] 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.

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

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., $\sum_{s\geq t+1} \rho^s \int_{Z^T} \max_{\tilde{w}\in W_s} u(x(\tilde{w}),w_s) f_z(z^T)\mathrm{d}z^T$ does not depend on the report in period t). As a result, the IC constraint in period t reduces to

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

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 has the following resource constraint (RC):

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

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

$$V(x(\cdot),w) = \int_{Z^T} U\left(\{c(w_t), y(w_t)/w_t\}_{t=0}^T\right) f_z(z^T)\mathrm{d}z^T. \tag{3}$$

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

$$\begin{aligned} &\max_{x(\cdot)} \int_{W_0} V(x(\cdot),w) f_w(w)\mathrm{d}w \\ &\text{s.t.} \quad (\mathrm{RC}) \text{ and } (\mathrm{IC}_t) \text{ for all } t. \end{aligned} \tag{4}$$

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.[14]

Here is the key idea of our work. When we solve the problem (4), we exploit the fact that our mechanism is time-invariant and does not depend on history, and we consider 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 gives us the equivalence. Then, we take the standard approach for static optimal income taxation problems to solve (4). 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.[15]

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 1. *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 $w \in [\underline{w}, \overline{w})$.*

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

[15] 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, forthcoming).

Proof. See Appendix. □

Property (i) states that the allocation is efficient at the top of W and if income is strictly increasing at the bottom of W, the allocation is also efficient there. In addition, no allocation can be distorted upward from efficiency and in particular, if income is positive, the allocation is distorted downward from efficiency 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 1, 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 $\phi_t(w; z^T) = w + \sum_{s=0}^{t} z_s$. 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.

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 efficient 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.

2.1 Saving

In the foregoing analysis, we assumed that people cannot save nor borrow because the government confiscates all incomes that are not consumed as in Battaglini and Coate (2008). In this section, we show that, under the linear technology of shock, the government actually *cannot* allow saving or borrowing under stationary tax rules that depend on only the current income.

To this end, suppose that there is a risk-free bond market with the interest rate $R > 0$ and let $b(w)$ be the endogenous bond holding of agent w. In the presence of bond holding, we consider a mechanism $x(\cdot) = (y(\cdot), c(\cdot), b(\cdot))$. Naturally, $b(\cdot)$ is time-invariant and depends on only current report because of the government's sluggishness. In this section, we focus on the linear technology of shock: $\phi_t(w; z^T) = w + \sum_{s=1}^{t} z_s$. Then, we obtain the following result:

Proposition 2. *Suppose $\phi_t(w; z^T) = w + \sum_{s=1}^{t} z_s$. Under Assumptions 1 and 2, if the government satisfies the IC constraints, $b(w) = 0$ for all $w \in W$.*

Proof. We argue by backward induction. At first, $b(w_T) = 0$ on W_T by terminal condition (i.e., in the last period, no one will save and borrowing is not permitted because people cannot repay). Now, suppose $b(w_t) = 0$ on W_t ($t \leq T$), and let $w_{t-1} \in W_{t-1}$. Then, there exists $w \in W_0$ and $z^{t-1} \in Z^{t-1}$ such that $w_{t-1} = w + \sum_{s=1}^{t-1} z_s$. Without loss of generality, suppose $\max Z > 0$.[16] Because $W_t \cap W_{t-1} \neq \emptyset$, $\min W_t \leq \max W_{t-1}$. If $w_{t-1} \geq \min W_t$, $b(w_{t-1}) = 0$ 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). Then, there exists $w' \in W_0$ and $z'^{t-1} \in Z^{t-1}$ such that $w'_{t-1} = w' + \sum_{s=1}^{t-1} z'_s$. Moreover, by the Intermediate Value Theorem, we can take $z_t, z'_t \in Z$ such that $w_t \equiv w_{t-1} + z_t = w'_{t-1} + z'_t \equiv w'_t$.[17] Then, if the IC constraints hold (i.e., people report truthfully), the budget constraints imply $\tau(w_t) = y(w_t) - c(w_t) - b(w_t) + (1+R)b(w_{t-1})$ 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 $b(w_{t-1}) = 0$ on W_{t-1}. □

Therefore, as long as the sluggish government would like to satisfy the IC constraints, it cannot allow saving or borrowing, *even if they are taxed*. Because our revenue constraint is integrated over time, the government can save and borrow for the agents. However, a sluggish government and IC constraints leave no room for saving and borrowing not because the government borrows and saves for the

[16] An analogous argument holds for $\max Z < 0$.

[17] 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.

consumers, but because it cannot actually address the intertemporal wedge.[18][19] Indeed, Farhi and Werning (forthcoming) also have a revenue constraint integrated over time, but according to their simulations, bond holdings are not zero.

We can see that the stochastic 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, this makes it impossible for the government to take care of these agents' IC constraints simultaneously due to its sluggishness.[20]

2.2 Aggregate shocks

We assumed that agents receive idiosyncratic shocks, but our argument remains intact even if we consider a private or public aggregate shock instead of or in addition to private idiosyncratic shocks.

Consider first the case of only an aggregate shock but no idiosyncratic shocks. Suppose that draws from $f_z(z^T)$ are aggregate shocks, so all agents face the same shock. At period 1, people's expected utility is $\int_{Z^T}\left(\sum_{t=1}^{T} u(c_t, y_t/w_t)\right) f_z(z^T)\mathrm{d}z^T$. Thus, the social welfare is

$$\begin{aligned}&\int_{W_0}\left\{\int_{Z^T}\left(\sum_{t=1}^{T} u(c_t, y_t/w_t)\right) f_z(z^T)\mathrm{d}z^T + u(c_0, y_0/w)\right\} f_w(w)\mathrm{d}w\\ &= \int_{W_0}\int_{Z^T} U\left(\{c_t, y_t/w_t\}_{t=0}^{T}\right) f_z(z^T)\mathrm{d}z^T f_w(w)\mathrm{d}w = SW.\end{aligned}$$

Therefore, the social welfare function is unchanged. Moreover, if the aggregate shock is private, the IC constraints are identical to those in the idiosyncratic case, as long as the idiosyncratic shocks are i.i.d. across people. Thus, our argument is unchanged. However, if the aggregate shock is public and the government can also observe the shock, we need some modifications because the IC constraints are imposed for each realization of the shock, as opposed to the idiosyncratic case.

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

[19]It is worth noting that Proposition 2 holds regardless of people's risk attitude (i.e., whether they are risk-neutral or risk-averse).

[20]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.

Suppose that the aggregate shock is public and let W_{t,z^T} be the range of $\phi_t(\cdot\,; z^T)$. Because of the government's sluggishness, the IC constraint in period t when the shock is z^T reduces to

$$\forall w \in W_{t,z^T},\ u(x(w), w) \geq u(x(w'), w) \text{ for all } w' \in W_{t,z^T}.$$

Then, letting $W = \bigcup_{t=0}^{T} \bigcup_{z^T \in Z^T} W_{t,z^T}$, the same argument as used for the idiosyncratic case would follow.[21]

Finally, by combining our earlier analysis with the analysis in this subsection, we can readily see that having both aggregate and idiosyncratic shocks makes it necessary to expand the type space, but does not change the main result.

3 Conclusion

We considered the optimal dynamic income taxation problem faced by a sluggish government that cannot change its 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.[22]

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. This should be a subject 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.

[21] Note that if the government could look at history, it would suffice to let agents report only in the initial period. Indeed, as long as they report truthfully in the initial period, the government can see through their misreports in the subsequent periods because it can observe the shock.

[22] 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 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*:

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

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

Next, we rewrite the welfare function as

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

where $f_t(w) = \int_{Z^T} f_z(z^T) f_w(\phi_t^{-1}(w; z^T)) \frac{\mathrm{d}\phi_t^{-1}(w;z^T)}{\mathrm{d}w}\mathrm{d}z^T$. 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

$$\int_{W} u(x(w), w) g(w)\mathrm{d}w \tag{5}$$

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

$$G \leq \int_{W} \tau(w) g(w)\mathrm{d}w. \tag{6}$$

Therefore, our relaxed problem is given by

$$
\begin{aligned}
&\max_{x(\cdot)} \int_W u(x(w),w)g(w)\mathrm{d}w \\
&\text{s.t.} \quad G \leq \int_W \tau(w)g(w)\mathrm{d}w, \\
&\qquad c(w) \text{ is nondecreasing and } \tfrac{\mathrm{d}u(x(w),w)}{\mathrm{d}w} \geq u_w(x(w),w) \text{ on } W.
\end{aligned}
\tag{7}
$$

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

$$
\begin{aligned}
&\max_{x(\cdot)} \int_{W_0} u(x(w),w)f_w(w)\mathrm{d}w \\
&\text{s.t.} \quad G \leq \int_{W_0} \tau(w)f_w(w)\mathrm{d}w, \\
&\qquad c(w) \text{ is nondecreasing and } \tfrac{\mathrm{d}u(x(w),w)}{\mathrm{d}w} \geq u_w(x(w),w) \text{ on } W_0.
\end{aligned}
\tag{8}
$$

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., "Optimal dynamic nonlinear income taxes with no commitment," forthcoming, *Journal of Public Economic Theory*.

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

Brito, L. D., 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.

Diamond, P. and E. Saez. (2011), "The case for a progressive tax: From basic research to policy recommendation," *Journal of Economic Perspectives*, **25**, 165-190.

Farhi, E. and I. Werning., "Insurance and taxation over the life cycle," forthcoming, *Review of Economic Studies*.

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.