Federal Reserve Bank of Minneapolis
Research Department Staff Report 365

December 2005

Tax Smoothing with Redistribution*

Iván Werning

Massachusetts Institute of Technology,
National Bureau of Economic Research,
and Federal Reserve Bank of Minneapolis

ABSTRACT

We study optimal labor and capital taxation in a dynamic economy subject to government expenditure and aggregate productivity shocks. We relax two assumptions from Ramsey models: that a representative agent exists and that taxation is proportional with no lump-sum tax. In contrast, we capture a redistributive motive for distortive taxation by allowing privately observed differences in relative skills across workers. We consider two scenarios for tax instruments: (i) taxation is linear with arbitrary intercept and slope; and (ii) taxation is non-linear and unrestricted as in Mirrleesian models. Our main result provides conditions for perfect tax smoothing: marginal taxes on labor income should remain constant over time and invariant to shocks. In addition, capital should not be taxed. We also discuss implications for optimal debt management. Finally, an extension highlights movements in the distribution of relative skills as a potential source for variations in optimal marginal tax rates.

* Email: iwerning@mit.edu. This paper is a revised version of the first part of my Ph.D. dissertation at the University of Chicago (Chapters 2 and 3 in Werning, 2002). I am grateful to my advisors Fernando Alvarez, Gary Becker, Pierre-Andre Chippori and Robert Lucas for their support. For useful discussions and comments I also thank Daron Acemoglu, Marios Angeletos, Paco Buera, Emmanuel Farhi, Mike Golosov and Pierre Yared. I am grateful for the hospitality during my visit at the Federal Reserve Bank of Minneapolis. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

How should a government set and adjust taxes on labor and capital over time in the face of shocks to government expenditure and aggregate productivity? Ramsey optimal tax theory provides two important insights into this question: taxes on labor income should be smoothed (Barro, 1979; Lucas and Stokey, 1983) while taxes on capital should be set to zero (Chamley, 1986; Judd, 1985; Zhu, 1992; Chari, Christiano, and Kehoe, 1994).

This paper addresses an important shortcoming in interpreting these cornerstone results. The Ramsey approach casts the optimal tax problem within a representative-agent paradigm; then, to avoid the first-best allocation, lump-sum taxes—or any combination of tax instruments that may replicate lump-sum taxes—are ruled out. The second-best problem chooses the right mix of distortive taxes to maximize the representative agent's utility subject to the government's intertemporal budget condition.

Societies may have their own good reasons for avoiding complete reliance on lump-sum taxation and resigning themselves to the use of distortionary taxes. Unfortunately, none of these reasons are explicitly captured by a representative-agent Ramsey framework. Although the first-best allocation is ruled out, an arbitrary second-best problem is set in its place. What confidence can we have that tax recommendations obtained this way accurately evaluate the trade-offs faced by society? If, for reasons unspecified in the model, lump-sum taxes are presumably undesirable for society yet still desirable within the model, how can we be sure that the tax prescriptions derived are not, for the same unspecified reasons, also socially undesirable?

Distributional concerns are a natural reason to resort to distortionary taxation (Mirrlees, 1971). If workers are heterogeneous with respect to their labor productivity, and if this trait is not observable, then society cannot attain almost any of the first-best allocations. In contrast, by taxing observable differences such as income, redistribution is possible, albeit at a loss in efficiency. Such a trade-off between redistribution and efficiency provides a solid microfoundation for the role of distortionary taxes. With this in mind, this paper reexamines optimal taxation in dynamic economies close to Ramsey models such as Chari, Christiano, and Kehoe (1994) and others, but modeling distributional concern explicitly.

In our model, workers are heterogenous with respect to the productivity of their work effort. These differences in relative skill are private information. The aggregate economy features neoclassical capital accumulation and is subject to fluctuations in government expenditures and technology. We consider two main scenarios regarding the set of tax instruments available to the government. In the first, labor income taxes are assumed linear, but we allow for an arbitrary lump-sum tax intercept in the schedule. In the second, we do not restrict tax instruments at all, so the government is limited then only by the inherent asymmetry of information, as in Mirrlees's (1971) nonlinear taxation model.

In the first case, the labor income tax schedule can be summarized at any moment by two variables: the intercept T_t, which we refer to as the lump-sum tax component, and the slope τ_t, which we refer to as the marginal tax rate. Thus, this simple linear tax case suffices for incorporating the essential missing tax instrument in Ramsey models: the lump-sum tax component. Indeed, if all workers had the same skill, then the lump-sum tax component can be used to attain the first-best allocation. However, with heterogeneity a positive marginal tax is generally preferable, since then more productive ("richer") workers bear a larger tax burden and alleviate that of the less productive ("poorer") workers. Since it is hard to justify restricting the lump-sum tax component, heterogeneity seems primordial for a well-motivated non-trivial tax problem (Mirrlees, 1971).[1]

Our main analytical result is that *perfect tax smoothing is optimal* for an interesting class of preferences. At the optimum, marginal income tax rates are constant over time and invariant to government expenditure and technology shocks. The government uses debt and the lump-sum tax component to smooth out these shocks. In addition, we find for the same class of preferences *the tax rate on capital is optimally set to zero*—a version of the Chamley-Judd result (Chamley, 1986; Judd, 1985; Zhu, 1992; Chari, Christiano, and Kehoe, 1994) for our heterogenous agent stochastic economy.

The intuition for the tax smoothing results is that, with heterogenous workers and a lump-sum tax component, it is *distributional concerns* that determine the desired level of distortive taxation. At any point in time, the current tax rate is a measure of redistribution across workers, while the distribution of relative skills determines the desired level of redistribution. In the model, this level is constant over time and invariant to government expenditure and technology shocks because these shocks do not directly affect the distribution of relative skills across workers.

To bring the distributional concerns to the forefront, we extend the model to allow for shocks to the distribution of relative skills. We show that the optimal tax rate does then respond to these shocks, but continues to remain unchanged in the face of aggregate shocks to the government and technology. In particular, tax rates rise when the dispersion of worker skills widens. This extension highlights a novel determinant in the dynamic evolution of optimal tax rates, one that cannot be addressed in a representative-agent Ramsey framework.

More generally, our normative model attributes a crucial role in optimal tax rates to the

[1] Indeed, most countries feature a negative intercept when one considers deductibles and transfer programs. This is the optimal outcome of our model as long as there is enough inequality in skills and a strong enough taste for equality.

distribution of relative skills. This relates to positive political economy models where the distribution of income has always played a prominent role—as in median voter models such as Meltzer and Richard (1981).

For the unrestricted Mirrleesian tax scenario, we find that tax rates should vary across workers, but should remain perfectly constant over time for any given worker. This tax smoothing result suggests a role for taxation based on current and past income averages, as a way of equating marginal tax rates over time while retaining the desired non-linearity across workers. Vickery (1947), for other reasons, was an early proponent of such income-averaging taxation schemes.

Our model also has some novel implications for public debt management. Ramsey models break Ricardian equivalence, which otherwise renders government debt indeterminate. In particular, in Barro (1979) and Lucas and Stokey (1983) public debt plays a crucial role in allowing the government to smooth tax rates over time. In contrast, in our model Ricardian equivalence reemerges, despite distortionary taxation, as long as a lump-sum tax component is available. In general, the government can smooth marginal tax rates with various mixes of debt and lump-sum tax financing. We briefly speculate on extensions of our model that may render debt management determinate.

The rest of this paper is organized as follows. Section 2 introduces the basic assumptions regarding preferences and technology. Section 3 considers the linear taxation scenario. The main tax smoothing and capital taxation results are obtained in Section 4 for the linear case, while Section 5 shows that these results extend to the Mirrleesian tax scenario. Section 6 contains our conclusions and some speculations regarding useful extensions for future work. An appendix collects some proofs and other derivations.

2 The Dynamic Economy

The main purpose of our model's assumptions is to extend Ramsey dynamic taxation frameworks in the direction of incorporating heterogeneity, in the spirit of Mirrlees's (1971) private information framework, in a simple and tractable way so that taxation can be microfounded by a desire for redistribution. To this end, our economy is populated by a continuum of infinitely lived workers with different relative skill levels. To focus on uncertainty at the aggregate level, we only consider fixed differences across worker types.[2] Initially we assume a fixed distribution of relative productivities. However, later we allow for shocks to this

[2] This distinguishes our approach here from an incipient and growing literature that attributes to taxation an important role in the *insurance of ongoing shocks* to workers' productivity (e.g. Golosov, Kocherlakota, and Tsyvinski, 2003; Albanesi and Sleet, 2004) by assuming the market cannot provide such insurance arrangements. This paper does not attempt to contribute to this interesting alternative line of work.

distribution of relative skills.

The population is divided into a finite number of types indexed by $i \in I$, with productivity level θ^i and relative size π^i. For now we assume that the distribution of productivities, summarized by $\{\theta^i, \pi^i\}_{i \in I}$, is fixed over time; we relax this assumption later in Section 4. Abler workers produce more efficiency units of labor for any given level of work effort. Importantly, individual productivity θ^i and work effort n^i are private information to the worker. Only the product of the two, the efficiency units of labor $L^i = \theta^i n^i$, is publicly observable. As a result, the government cannot levy discriminatory lump-sum taxes that are conditional on the worker's type $i \in I$, which would otherwise allow all first-best allocations of interest to be attained.

Following Mirrlees (1971), we assume productivity differences to be the only source of heterogeneity. Individuals have identical preferences represented by the additively separable utility function

$$\sum_{t=0}^{\infty} \beta^t \mathbb{E}\left[u(c_t) - v(n_t)\right] \tag{1}$$

In addition, we work with particular specification for the utility from consumption and the disutility from work effort functions and assume that both are of constant elasticity, so that $u(c) = c^{1-\sigma}/(1-\sigma)$ and $v(n) = \alpha n^{\gamma}/\gamma$, with $\sigma, \alpha > 0$ and $\gamma > 1$. As we will see below, this specification of preferences allows us to derive clean analytical results that are likely to provide a very useful benchmark for other cases.

There are two sources of aggregate uncertainty: government expenditures and technology shocks. Both can be captured rather abstractly by assuming a publicly observed state of the economy $s_t \in S$ in period t, where S is some finite set. Below, both government expenditure and the production function are functions of s_t. Let $s^t \equiv (s_0, s_1, s_2, ..., s_t) \in S^t$ denote the history of states; no restriction is placed on the probability distributions $\Pr(s^t)$ governing the evolution of the state. An allocation specifies consumption, labor and capital, in every period and history: $\{c^i(s^t),\, L^i(s^t),\, K_{t+1}(s^t)\}$.

Production combines labor with capital using a constant returns to scale technology. The resource constraints are

$$c(s^t) + K(s^t) + g_t(s_t) \leq F(L(s^t), K(s^{t-1}), s_t) + (1-\delta)K(s^{t-1}) \tag{2}$$

for all periods $t = 0, 1, \ldots$ and histories $s^t \in S^t$. Aggregates are denoted by $c(s^t) \equiv \sum_{i \in I} c^i(s^t)\pi^i$ and $L(s^t) \equiv \sum_{i \in I} L^i(s^t)\pi^i$.

Note the two roles that the state s_t is allowed to play. First, government expenditure may fluctuate over time or with the state of the economy according to the given $g_t(s_t)$ function.

Second, the production function may depend on the current state to allow for aggregate technology shocks.

3 Linear and Proportional Taxation

We begin by considering the case where in each period the tax schedule is a linear function of labor income: $\tau(s^t)y^i(s^t) + T(s^t)$. A virtue of this simple linear specification, compared to more complicated tax schemes such as the nonlinear case considered in Section 5, is that it allows us to focus on the evolution of the marginal tax rate on labor income, a one-dimensional policy variable. Thus, the notion of tax smoothing is straightforward. For completeness, we shall also consider the proportional tax case where the lump-sum tax component T_t is constrained to be zero.

In addition to taxing labor income, the government can levy a proportional tax, denoted $\kappa(s^t)$, on the net return to capital. In our model, as long as the taxation of initial wealth and the returns to capital are unrestricted, consumption taxes are superfluous and can be omitted from the analysis without loss in generality.

We allow for complete asset markets as in Lucas and Stokey (1983), Chari, Christiano, and Kehoe (1994), and many others. A literal interpretation of this envisions government debt to include a rich set of Arrow-Debreu state-contingent bonds; or equivalently, that the government issues a single bond but with well-tailored stochastic returns. An alternative, less literal, interpretation is provided by the fact that even with non-contingent debt there are other ways of replicating complete market outcomes. For example, Angeletos (2002) and Buera and Nicolini (2004) show how a portfolio composed of riskless bonds of different maturities might be used to this end.[3]

Agent Problem. With complete markets each individual $i \in I$ can be seen as facing a single intertemporal budget constraint:[4]

$$\sum_{t,s^t} p(s^t)\big(c^i(s^t) + K^i(s^t) - w(s^t)(1-\tau(s^t))L^i(s^t) - R(s^t)K^i(s^{t-1})\big) \leq -T, \tag{3}$$

where $T \equiv \sum_{t,s^t} p(s^t)T(s^t)$ is the present value of the lump-sum components of taxes. Here

[3] Nevertheless, in this paper we adopt the complete market assumption for its simplicity and as a mechanism for focusing attention on other issues, not for its realism. Our stated objective is to extend Ramsey models by incorporating heterogeneity and lump-sum taxation. Recent work featuring incomplete markets, but within representative-agent Ramsey economies, includes Aiyagari, Marcet, Sargent, and Seppälä (2002), Werning (2005) and Farhi (2005).

[4] It is standard to derive such a single intertemporal budget constraint from a sequence of temporary budget constraints, solving out for the bond holdings, see Lucas and Stokey (1983) or Chari, Christiano, and Kehoe (1994).

$p(s^t)$ represents the Arrow-Debreu price of the consumption good in period t and history s^t, and we normalize so that $p(s_0) = 1$; the real wage is $w(s^t)$; and $R(s^t) \equiv 1+(1-\kappa(s^t))(r(s^t)-\delta)$ is the after-tax gross rate of return on capital, where $r(s^t)$ denotes the rental rate of capital.

Firms. Each period firms solve a static maximization of profits, $F(K, L, s)-r(s^t)K-w(s^t)L$, leading to the usual marginal conditions

$$r(s^t) = F_K(L(s^t), K(s^{t-1}), s_t), \tag{4}$$

$$w(s^t) = F_L(L(s^t), K(s^{t-1}), s_t). \tag{5}$$

Profits are zero in equilibrium given constant returns to scale.

Government Budget Constraint. With complete markets the government faces a single intertemporal budget constraint

$$\sum_{t,s^t} p(s^t)g(s^t) \leq T + \sum_{t,s^t} p(s^t)\big(\tau(s^t)w(s^t)L(s^t) + \kappa(s^t)(r(s^t) - \delta)K(s^{t-1})\big)$$

By a version of Walras law, the government budget constraint holds with equality whenever the resource constraints (2) and the worker budget constraints (3) all hold with equality.

Competitive Equilibria. A *competitive equilibrium* is a sequence of taxes $\{T(s^t),\ \tau(s^t),\ \kappa(s^t)\}$, prices $\{p(s^t),\ r(s^t),\ w(s^t)\}$, and quantities $\{c^i(s^t),\ L^i(s^t),\ K(s^t)\}$, such that: (i) workers maximize: consumption and labor choices $\{c^i(s^t), L^i(s^t)\}$ maximize utility taking prices and taxes as given for all individuals $i \in I$; (ii) firms maximize: capital and labor $K(s^{t-1})$ and $L(s^t)$ solve the static profit maximization taking the rental and real wage rate as given; (iii) the government's budget constraint holds; (iv) markets clear: the resource constraints (2) hold for all periods and histories.

Characterizing Equilibrium Allocations. Our first goal is to provide a useful characterization of the set of allocations that are sustainable by a competitive equilibrium for some taxes and prices. This later allows for a *primal* approach that formulates the taxation problem directly in terms of allocations.

The necessary and sufficient first-order conditions for the worker's maximization problem are

$$w(s^t)(1-\tau(s^t)) = \frac{1}{\theta^i}\frac{v'\left(\frac{L^i(s^t)}{\theta^i}\right)}{u'(c^i(s^t))}, \tag{6}$$

$$\frac{p(s^t)}{p(s_0)} = \beta^t \frac{u'\left(c^i(s^t)\right)}{u'\left(c^i(s_0)\right)} \Pr(s^t), \tag{7}$$

$$p(s^t) = \sum_{s_{t+1}\in S} p(s^t, s_{t+1}) R(s^t, s_{t+1}), \tag{8}$$

together with the budget constraint (3) holding with equality. Equation (6) is the intratemporal optimality conditions equating the marginal rate of substitution between consumption and labor with the real after-tax wage. Equation (7) is the standard intertemporal optimality condition. Equation (8) ensures no arbitrage for investment in capital.

The first-order conditions (6)–(7) and the assumption that marginal utility and disutility $u'(c)$ and $v'(n)$ are power functions imply that individual consumption and labor are proportional to their aggregates

$$c^i(s^t) = \omega_c^i\, c(s^t), \tag{9}$$

$$L^i(s^t) = \omega_L^i\, L(s^t), \tag{10}$$

with some fixed weights ω_c^i and ω_L^i with $\sum_{i\in I} \omega_j^i \pi^i = 1$, for $j = c, L$.

Substituting conditions (6)–(10) into the budget constraint (3), and using the fact that $u'(c)c = (1-\sigma)u(c)$ and $v'(n)n = \gamma v(n)$ yields

$$\sum_{t,s^t} \beta^t \left((1-\sigma)(\omega_c^i)^{1-\sigma} u(c(s^t)) - \gamma \left(\frac{\omega_L^i}{\theta^i}\right)^{\gamma} v(L(s^t)) \right) \Pr(s^t) = (\omega_c^i)^{-\sigma} c_0^{\sigma} (R_0 K_0^i - T), \tag{11}$$

which we call the *implementability constraint* for agent $i \in I$.

In Ramsey models the representative agent's implementability condition turns out to fully characterize, along with the resource constraints, the restrictions on competitive equilibria. In contrast, in our setup we need an additional restriction to capture the fact that all workers face the same marginal tax on labor. Using the proportionality conditions from (9)–(10) in equation (7) implies

$$w(s^t)(1-\tau(s^t)) \frac{u'(c(s^t))}{v'(L(s^t))} = \frac{1}{\theta^i \alpha} \frac{v'\left(\frac{\omega_L^i}{\theta^i}\right)}{u'(\omega_c^i)},$$

where we use the fact that with power functions: $u'(xy) = u'(x)u'(y)$ and $v'(xy) = v'(x)v'(y)/\alpha$. The right-hand side of this equation depends on i, but not on t nor s^t; in contrast, the left-hand side depends on t and s^t, but not on i. It follows that both sides must equal some constant independent of t, s^t and i. In particular,

$$\frac{1}{\theta^i \alpha} \frac{v'\left(\frac{\omega_L^i}{\theta^i}\right)}{u'(\omega_c^i)} = \phi \qquad i \in I \qquad \text{for some } \phi > 0 \tag{12}$$

is the additional restriction needed to characterize equilibrium allocations.[5]

We have shown that conditions (9)–(12) together with the resource constraint (2) are necessary for an equilibrium. It turns out that the converse is also true: these equations fully characterize allocations that can be supported as an equilibrium for some tax and price sequences. Indeed, these tax policies and prices are obtained from the first-order conditions and the definition of ϕ:

$$\tau(s^t) = 1 - \phi \frac{v'(L(s^t))}{u'(c(s^t))} \frac{1}{F_L(L(s^t), K(s^{t-1}), s_t)}, \tag{13}$$

$$\frac{p(s^t)}{p(s_0)} = \beta^t \frac{u'(c(s^t))}{u'(c(s_0))} \Pr(s^t), \tag{14}$$

Factor prices $r(s^t)$ and $w(s^t)$ are given by the marginal product conditions (4)–(5). The tax rates on capital income $\kappa(s^t)$ can be set in any way that satisfies condition (8) with prices (14).[6]

Proposition 1 *An allocation $\{c^i(s^t), L^i(s^t), K(s^{t-1})\}$ can be supported by a competitive equilibrium if and only if there are distributional weights (ω_L^i, ω_c^i) and aggregates $\{c(s^t), L(s^t)\}$ so that the resource constraint (2) and conditions (9)–(12) hold.*

4 Tax Smoothing and Zero Capital Taxation

Based on Proposition 1 we can write the optimal tax problem as maximizing a weighted sum of utilities using Pareto weights $\{\pi^i \lambda^i\}$:

$$\sum_{i \in I} \lambda^i \pi^i \sum_{t, s^t} \beta^t \Big((\omega_c^i)^{1-\sigma} u(c(s^t)) - \Big(\frac{\omega_L^i}{\theta^i}\Big)^\gamma v(L(s^t)) \Big) \Pr(s^t)$$

[5] This representation of an allocation in terms of $\{\omega_c^i, \omega_L^i\}$, $\{c_t, L_t\}$, ϕ and T is the most convenient for our present purposes. However, other equivalent representations are possible. Indeed, in Appendix B we adopt a slightly different one, which is more convenient for the purposes pursued there.

[6] Note that with a single agent equation (3) has $\omega_c = \omega_L = 1$ and this condition simply defines $\phi = 1$. Equation (13) is then the standard expression used in the Ramsey literature to back out tax rates.

subject to the resource constraint (2), the implementability condition (11) and the consistency condition (12).[7] The maximization is performed over the distributional weights (ω_L^i, ω_c^i), the aggregates $\{c(s^t),\ L(s^t)\}$, the constant ϕ and, if available, a lump-sum tax T. An interesting benchmark is the Utilitarian case where Pareto weights are simply group size, so that $\lambda^i = 1$.

The arguments that follow only involve the optimality conditions with respect to aggregate variables, not those related to distributional weights or ϕ. The first-order conditions with respect to $c(s^t)$ for $t \geq 1$, and $L(s^t)$ for any $t \geq 0$, yield

$$u'(c(s^t)) \sum_{i \in I} (\omega_c^i)^{1-\sigma} (\lambda^i - (1-\sigma)\mu^i)\pi^i = \eta(s^t), \tag{15}$$

$$v'(L(s^t)) \sum_{I \in i} \left(\frac{\omega_L^i}{\theta^i}\right)^{\gamma} (\lambda^i - \gamma\mu^i)\pi^i = \eta(s^t) F_L(L(s^t), K(s^{t-1}), s_t), \tag{16}$$

where the multiplier on the resource constraint is $\beta^t \eta(s^t) \Pr(s^t)$ and that on the left-hand side of the implementability condition for agent $i \in I$ is $\mu^i \pi^i$.[8]

Our first result concerns the optimal taxation of capital. The first-order condition with respect to capital $K(s^t)$ is

$$\eta(s^t) = \beta \sum_{s_{t+1}} \eta(s^{t+1}) R^*(s^{t+1}) \Pr(s_{t+1}|s^t),$$

where $R^*(s^t) \equiv F_K(L(s^t), K(s^{t-1}), s_t) + 1 - \delta$ is the social marginal rate of return on capital. Condition (15) implies that $\eta(s^t)$ is proportional to $u'(c(s^t))$ for $t \geq 1$, so this gives

$$u'(c(s^t)) = \beta \sum_{s_{t+1}} u'(c_{t+1}(s^{t+1})) R^*(s^{t+1}) \Pr(s_{t+1}|s^t)$$

for $t \geq 1$. Comparing this condition with condition (8) with prices in (14) reveals that the tax on capital can be set to zero $\kappa(s^t) = 0$ for all $t \geq 2$.

We now show that a very strong form of marginal tax smoothing is optimal. Dividing equation (16) by (15) and rearranging, we obtain

$$\frac{v'(L(s^t))}{u'(c(s^t))} \frac{1}{F_L(L(s^t), K(s^{t-1}), s_t)} = \frac{\sum_{i \in I} (\omega_c^i)^{1-\sigma} (\lambda^i - (1-\sigma)\mu^i)\pi^i}{\sum_{i \in I} \left(\frac{\omega_L^i}{\theta^i}\right)^{\gamma} (\lambda^i - \gamma\mu^i)\pi^i}$$

[7] Note that for each individual $i \in I$ the left-hand side of the implementability condition is comparable to their contribution in the objective function except that more relative weight is placed on the disutility of work than on consumption, since $\gamma > 1$ and $1 - \sigma < 1$.

[8] The first-order condition for initial period consumption $c(s_0)$ is derived later. It is not crucial for any of our main results. As usual, it is slightly different due to the presence initial wealth.

for all $t \geq 1$. Combining this with equation (13) and using the definition of ϕ then gives

$$\tau(s^t) = \bar{\tau} \equiv 1 - \frac{\sum_{i \in I} u'(\omega_c^i)\omega_c^i(\lambda^i - (1-\sigma)\mu^i)\pi^i}{\sum_{i \in I} u'(\omega_c^i)\omega_L^i(\lambda^i - \gamma\mu^i)\pi^i}. \tag{17}$$

Hence, labor income tax rates are constant across time and states.

We summarize both results in the next proposition.

Proposition 2 *Perfect tax smoothing is optimal $\tau(s^t) = \bar{\tau}$ given by equation (17), for $t \geq 1$. The optimal tax on capital is zero $\kappa(s^t) = 0$ for all $t \geq 2$. Both results hold with or without a lump-sum tax component.*

Our interpretation for the zero tax on capital result is based on the well-known uniform taxation principles due to Diamond and Mirrlees (1971). The assumption that the utility function has constant elasticity of substitution implies a homotheticity in preferences over consumption time paths. This, in turn, implies that consumption at different dates should be taxed uniformly, which is only possible if capital income is untaxed. Basically, our model does not upset the main logic of the Chamley-Judd result.[9]

The intuition for the tax smoothing result is best conveyed by considering in turn the cases with and without the lump-sum tax component. With lump-sum taxes distortionary taxation is simply a redistribution mechanism. A positive marginal tax rate is the instrument by which the "rich" pay more taxes than the "poor", which is desirable whenever redistribution is. The optimal tax level balances concerns for redistribution and efficiency. Tax smoothing emerges as long as the determinants of inequality are invariant to government expenditure or aggregate technology shocks. The desired amount of redistribution is then constant over time, and a perfectly constant optimal tax rate results.

In representative-agent Ramsey models tax smoothing results are often informally explained by the following intuition: in order to minimize the total cost from distortions it is optimal to equate the marginal cost of distortions over time by equating tax rates over time. Our tax smoothing result refines this intuition. Consider first the more natural scenario where a lump-sum tax component is available. In this case, optimality dictates that the marginal cost of increased distortions be equated to the marginal benefit from increased redistribution. In our model the latter is invariant to government expenditure and aggregate technology shocks. Hence, the marginal cost from distortions should be equated over time and perfect tax smoothing is optimal.

Interestingly, the results regarding tax smoothing and capital taxation hold even when a lump-sum tax component is not available. However, there are some important differences.

[9] Indeed, in different ways, zero capital tax results have been derived allowing for heterogeneity and some forms of redistribution (e.g., see Chamley, 1986; Judd, 1985).

First, note that for this case the average *level* of taxation cannot be determined by distributional considerations. To take an extreme case, even in the absence of inequality—that is, if productivities were identical across workers—distortive taxation is still required. Moreover, for any given distribution of relative skills the distributional tastes are irrelevant. For example, even if equality is not valued—say, if the weights λ^i are higher for more productive workers—then positive distortive taxation will still be required to meet the government's budget constraint. In this sense, the level of taxation is determined by budgetary needs, not distributional concerns. This is, of course, in contrast to the case where a lump-sum tax component is available, where we have argued that distributional concerns are at center stage.

Turning to the timing of taxes, debt becomes critical when no lump-sum tax component is available. If, instead, a government were forced to balance the budget each period, tax smoothing would be simply infeasible. Debt allows the government to spread out the collection of taxes and meet a single present-value budget constraint. Intuitively, smoothing tax rates is then optimal because it minimizes total efficiency costs by equating the marginal efficiency costs from distortions over time.

Before moving on, note that our model nests the standard representative-agent Ramsey model. That is, if we restrict the government to no lump-sum tax and if the distribution of labor productivity is trivial, so that $\theta^i = 1$ for all i, our model is equivalent to the standard Ramsey framework. Restricted within this Ramsey case, our result on zero capital taxation echoes results in Chari, Christiano, and Kehoe (1994). As for labor income taxation, they report numerical results using various preference specifications that are different from the class required for our analytical results. Their simulations, however, turn out minuscule variations in labor income taxes, which they attribute intuitively to a strong tax smoothing motive.

These numerical results show the relevance of our analytical result away from our baseline class of preferences. Conversely, our analytical result, proving perfect tax smoothing for a baseline class of preferences, help explain numerical findings of near perfect tax smoothing away from our baseline class, such as those found by Chari, Christiano, and Kehoe (1994) and others. Thus, the results obtained in our model, which permits heterogeneity and lump-sum taxation, are also of interest when restricted to the Ramsey case with no heterogeneity nor lump-sum taxation.

Distributional Shocks. To bring out the importance of distributional concerns in determining the marginal tax rate, we now extend the model by and allow the distribution of relative skills to vary over time or with the state of the economy. The productivity of a worker of type $i \in I$ is now given by $\theta^i_t(s_t)$. The analysis of this extended model is contained in Appendix B. Here we discuss the main implications for our results.

For the extended model, the results in Proposition 2 on zero capital taxation are unchanged. The appendix shows that the optimal tax rate on labor income is given by

$$\tau(s^t) = \bar{\tau}_t(s_t) \equiv 1 - \frac{\sum_{i\in I} u'(\omega_c^i)\omega_c^i \left(\lambda^i - (1-\sigma)\,\mu^i\right)\pi^i}{\sum_{i\in I} u'(\omega_c^i)\omega_{L,t}^i\,(s_t)\left(\lambda^i - \gamma\mu^i\right)\pi^i}, \tag{18}$$

which generalizes equation (13).[10] The only difference is that now an individual's share of labor income $\omega_{L,t}^i(s_t)$ potentially varies over time and with the current state, due to potential underlying variations in the skill distribution. Indeed, the share $\omega_{L,t}^i(s_t)$, and hence the tax $\bar{\tau}_t(s_t)$, is solely a function of the current distribution $\{\theta_t^i(s_t)\}$. Thus, tax smoothing continues to be optimal in that tax rates remain unresponsive to shocks that affect government expenditure or aggregate technology. The only source of variations in tax rates are changes in the distribution of relative skills.

With a redistributive motive, optimal tax rates are higher during times of higher dispersion of relative skills. The reason is that the redistribution from "rich" to "poor" workers that can be engineered by labor income taxation is more powerful then. To see this clearly, consider the Utilitarian case where $\lambda^i = 1$ and suppose a lump-sum tax is available. Then formula (18) becomes

$$\bar{\tau}_t(s_t) = 1 - \frac{\mathbb{E}[u'(\omega_c^i)] + \mathrm{Cov}(\omega_c^i, u'(\omega_c^i)(1-(1-\sigma)\mu^i))}{\mathbb{E}[u'(\omega_c^i)] + \mathrm{Cov}\left(\omega_L^i(s_t), u'(\omega_c^i)(1-\gamma\mu^i)\right)} \tag{19}$$

where the expectations and covariances use the population fractions $\{\pi^i\}$ as probabilities. This formula uses the first-order condition for the nondiscriminatory lump-sum tax T which implies that $\sum_{i\in I} u'(\omega_c^i)\mu^i\pi^i = 0$. Although only nondiscriminatory lump-sum taxation is allowed in our model, here $u'(\omega_c^i)\mu^i$ represents the value of a small fictitious discriminatory lump-sum tax to individual i. Thus, redistribution is desirable whenever $u'(\omega_c^i)\mu^i$ increases with skill θ^i, making the covariance term in the denominator positive. A period of high skill dispersion then increases dispersion of the labor income share $\omega_L^i(s_t)$, which increases the covariance term in the denominator. As a result, the optimal tax rate is then higher.

Movements in the distribution of relative skills turn out to be the only source for tax rate fluctuations in our model. This underscores the point made earlier, that a crucial determinant for tax rates is distributional concerns. Indeed, as discussed above, when a lump-sum tax component is available distributional concerns are the main determinant of the overall level of tax rates. Fluctuations in the distribution of skills then lead to optimal

[10] Once again, this formula holds when T is free (the case with a lump-sum tax component), as well as when it is restricted to being zero (the case without a lump-sum tax component).

fluctuations in tax rates over time.

Recall the intuition that, with a lump-sum tax component, the marginal cost from distortions should equal the marginal benefit from increased redistribution in each period. Then, as long as the skill distribution does not vary, the marginal benefit from redistribution is unchanging. Thus, the marginal cost from distortions should be equated over time, which in turn implies that tax rates should be constant. However, when the distribution of skills does shift, the marginal benefit from redistribution shift with it, so the marginal cost from distortions should not be equated over time. As a result, the optimal tax rate responds to such shifts.

Interestingly, changes in the distribution of skills affect tax rates with or without the availability of a lump-sum tax component. As argued previously, if no lump-sum tax component is available, then distributional concerns simply cannot shape the overall level of tax rates. They can, however, affect their timing: during times of high skill dispersion a greater fraction of taxes are paid by those who have the most; hence, it is optimal to concentrate taxation then.

Debt Management. Since Barro (1979), second-best tax problems have been used to avoid the neutrality results implied by Ricardian equivalence. In Ramsey models the optimal timing of taxes implies an optimal management of debt. Barro was the first to argue that distortionary tax rates should be smoothed: by analogy with permanent income theory, tax rates should be set with an eye towards permanent government spending, as opposed to current spending. As a result, government debt should be used to buffer any resulting deficits and surpluses. Lucas and Stokey (1983) extended this argument by allowing state-contingent debt: then taxes should also be smoothed across states of the world, as well as across time.

Both models share the essential feature that the solution to the tax problem determines a debt management policy. This is the case because, with proportional taxation, average and marginal taxes coincide.

However, in our model, with a lump-sum tax component available, this link is broken and marginal tax rates alone do not determine revenue. Ricardian equivalence is then recovered, rendering the debt level indeterminate in our model. Indeed, government debt is simply irrelevant. Nothing is lost if for some reason the government is required to balance its budget each period—the lump-sum component can do all the work.[11]

Things are quite different if we rule out the lump-sum tax component in our model. Government debt is then key to smoothing tax rates over time, just as in representative-

[11] However, note that even though the government may not need to issue bonds, in our model the asset market may still be important to allow the heterogenous agents to trade with each other.

agent Ramsey models. The optimal tax policy then uniquely determines an optimal debt management policy.

Initial Period Taxation. We now turn to the determinants of optimal taxation in the very first periods. There are two distortive taxes to consider. First, there is the initial tax rate on labor income τ_0. Second, there is the capital income tax $\kappa(s^1)$ which distorts investment in the initial period. Finally, there is also the initial time-zero capital levy κ_0 which is not distortive.

The first-order condition for initial aggregate consumption c_0 contains a few extra terms relative to equation (15):

$$u'(c_0)\sum_{i\in I} u'(\omega_c^i)\big(\omega_c^i(\lambda^i-\mu^i)+(R_0K_0^i+T)\mu^i\big)\pi^i=\eta_0.$$

Of course, if no lump-sum tax is available the same condition holds but with $T=0$. When a lump-sum tax is available, the first-order condition with respect to T is $\sum_{i\in I} u'(\omega_c^i)\mu^i\pi^i=0$. In both cases, it follows that the term involving T vanishes:

$$u'(c_0)\sum_{i\in I} u'(\omega_c^i)\big(\omega_c^i(\lambda^i-\mu^i)+R_0K_0^i\mu^i\big)\pi^i=\eta_0.$$

Now, unless the term involving initial capital income also drops out, the initial tax rate τ_0, determined by equation (13), will be affected and differ from the constant tax rate $\bar{\tau}$ we found for all other periods. Likewise, the tax on capital income $\kappa(s^1)$ will not be zero. However, the term involving capital does drop out in two important cases.

First, whenever the time-zero capital levy κ_0 is unrestricted, its first-order condition yields $\sum_{i\in I}(\omega_c^i)^{-\sigma}\mu^i K_0^i\pi^i=0$. As a result, the first-order condition for c_0 is then identical to that of any other periods. In this case the conclusions in Proposition 2 extend and $\tau(s^t)=\bar{\tau}$ for all $t\geq 0$, and $\kappa(s^t)=0$ for all $t\geq 1$.

Second, even if the initial capital levy is restricted, if there is no initial inequality in asset wealth so that $R_0K_0^i=R_0K_0^j$, then the term involving capital also drops out using the first-order condition for T.[12] Once again, the conclusions from Proposition 2 extend to all periods. Moreover, in this case the time-zero capital levy κ_0 can be set to zero and any ad hoc restriction on initial wealth taxation is nonbinding.[13]

Time-Zero Capital Tax Levy. In Ramsey models a striking contrast emerges between

[12] This is an interesting benchmark as it corresponds to the canonical optimal taxation situation where heterogeneity is due solely to productivity differences (Mirrlees, 1971).

[13] However, it seems difficult to justify ad hoc restrictions on the taxation of initial wealth. Moreover, even if explicit taxes on wealth are limited, consumption taxes could perfectly replicate their effects.

long-run and short-run capital tax prescriptions: eventually capital should go untaxed, but initially it should be taxed heavily. Time-zero capital levies provide revenues without distortions, mimicking the desired missing lump-sum tax.[14] The tension between long-run and short-run tax prescriptions has been viewed as an important source for time inconsistency of government policy.

In contrast, as discussed above in our framework, time-zero capital levies have the potential of being completely irrelevant. Indeed, the reason for their irrelevance is precisely what makes them so desirable in Ramsey models: capital levies that imitate them bring nothing new to the table when a lump-sum tax is already available. This is a noteworthy difference of our model with the more standard Ramsey framework. Thus, if positive time-zero capital levies are ever desired in our model, it must be for different reasons than in the standard Ramsey models.

In our model, capital levies cease to be neutral if we assume unequal initial asset holdings. For example, consider an extreme case where more productive workers are also wealthier, so that $\theta^i > \theta^j$ implies $K^i > K^j$. A proportional tax on initial wealth then acts as an ideal redistributive device, taking more from the rich, as income taxation does, but without introducing distortions. In such a case, as long as equality is valued, an initial wealth tax is desirable. Indeed, Pareto improvements may be possible if the tax on assets is coupled with a reduction in the distortionary tax rate on labor income.

In a nutshell, the Ramsey model is about the need to "redistribute" from the private to the public sector, in order to finance the government. Then any initial wealth in the hands of the private sector is best expropriated. In contrast, in our model the government may also need resources from the private sector, but the central tension is not getting these without distortions—which it could always do by raising the lump-sum tax. Rather, it is the distributional concern about *who* it is extracting resources from. Instead of redistribution from private to public sector, it is redistribution *within* the private sector that is at center stage.

Thus, a desire for initial wealth taxation can also be generated in our model. Moreover, this desire may also provide a source for time-inconsistent policy as long as more productive workers tend to accumulate more assets as time passes.[15] However, the mechanism is entirely different and suggests new issues. In particular, the distribution of assets within the private sector is brought to the forefront, something which cannot be addressed in a representative-agent Ramsey model.

[14] Hence, to avoid the first-best, most analyses proceed by imposing ad-hoc upper bounds on the amount of such levies.

[15] Note that a similar time-inconsistency issue arises if taxation based on past income were possible.

Replicating Completing Markets with Taxes. For Ramsey models, Chari, Christiano, and Kehoe (1994) show how capital taxation can help implement complete-market outcomes even when markets are incomplete. When markets are complete the tax rate on capital can be set in advance, so that at date t it is conditioned only on s^{t-1}; the relevant tax rate is then known at the moment of investment. However, if one conditions further on s^t, so that in addition to s^{t-1} the tax now depends on s_t, one can replicate any state-contingent profile of revenue, without introducing additional distortions to investment. As a result, state-contingent debt becomes inessential when the government can tax capital flexibly enough.

In our model, however, such a scheme will generally not work. The reason is related to our previous discussion on the role that capital levies then play: redistribution. The lump-sum tax can already provide a non-distortive source of state-contingent revenue. However, when markets are incomplete, it is not simply a source of non-distortive state-contingent revenue that is missing. Indeed, replicating complete markets requires replicating the insurance arrangements that heterogenous worker's were able to provide for each other. It is easy to see that, in general, a proportional tax on capital will not do the trick.

However, we speculate that there are possibly some interesting tax schemes that do work and relieve the role played by asset markets. We postpone exploring this issue further for future work.

5 Mirrleesian Taxation: Constrained Efficiency

We now consider the Mirrleesian scenario, where no restrictions are placed on tax instruments. That is, we study the constrained efficient allocations for our economy. The tax schemes required to implement these allocations are necessarily more complicated than the linear schemes we have considered so far. Our main goal is to characterize the shadow marginal tax rates for constrained efficient allocations. In particular, we examine whether some form of tax smoothing emerges and whether capital should not be distorted. Towards the end of this section we also discuss a relatively simple tax scheme with features suggested from the analysis.

Pareto Problem. Invoking the revelation principle, we set up a direct truth-telling mechanism. Workers submit reports regarding their type and receive an allocation as a function of this report. Incentive compatibility constraints then ensure that individuals report truthfully.

To simplify the exposition and notation, we treat the case with two types $\theta_H > \theta_L$ of equal size. The analysis trivially extends to more general cases. Thus, we maximize the utility of the high type

$$\sum_{t,s^t} \beta^t \Big(u(c^h(s^t)) - v\Big(\frac{L^h(s^t)}{\theta^h}\Big)\Big) \Pr(s^t)$$

subject to delivering some lower bound utility for the low type

$$\sum_{t,s^t} \beta^t \left(u(c^l(s^t)) - v\Big(\frac{L^l(s^t)}{\theta^l}\Big)\right) \Pr(s^t) \geq U^L,$$

the incentive constraints

$$\sum_{t,s^t} \beta^t \Big(u(c^h(s^t)) - v\Big(\frac{L^h(s^t)}{\theta^h}\Big)\Big) \Pr(s^t) \geq \sum_{t,s^t} \beta^t \left(u(c^l(s^t)) - v\Big(\frac{L^l(s^t)}{\theta^h}\Big)\right) \Pr(s^t),$$
$$\sum_{t,s^t} \beta^t \left(u(c^l(s^t)) - v\Big(\frac{L^l(s^t)}{\theta^l}\Big)\right) \Pr(s^t) \geq \sum_{t,s^t} \beta^t \Big(u(c^h(s^t)) - v\Big(\frac{L^h(s^t)}{\theta^l}\Big)\Big) \Pr(s^t),$$

and the resource constraints (2).

In the discussion that follows we consider the case of greatest interest, where only the first incentive compatibility constraint binds. This amounts to the region where redistribution takes place from high to low types, for high enough U^L. Indeed, a Utilitarian planner always favors this region of the Pareto frontier.

Implicit Marginal Taxes. Letting the multipliers be γ for the participation constraint, μ for the first incentive constraint and $\beta^t \eta(s^t)\pi(s^t)$ for the resource constraint, the first-order conditions are

$$\frac{1}{\theta^h} v'\Big(\frac{L^h(s^t)}{\theta^h}\Big)(1+\mu) = \eta(s^t) W^*(s^t), \tag{20}$$
$$u'(c^h(s^t))(1+\mu) = \eta(s^t), \tag{21}$$

$$\gamma \frac{1}{\theta^l} v'\Big(\frac{L^l(s^t)}{\theta^l}\Big) - \frac{\mu}{\theta^h} v'\Big(\frac{L^l(s^t)}{\theta^h}\Big) = \eta(s^t) W^*(s^t), \tag{22}$$
$$u'(c^l(s^t))(\gamma - \mu) = \eta(s^t), \tag{23}$$

$$\eta(s^t) = \beta \sum_{s_{t+1}} \eta_{t+1}(s^{t+1}) R^*(s^{t+1}) \Pr(s_{t+1}|s^t), \tag{24}$$

where $R^*(s^t) = F_K(L(s^t), K(s^{t-1}), s_t) + 1 - \delta$ denotes the social gross rate of return to capital and where $W^*(s^t) \equiv F_L(L(s^t), K(s^{t-1}), s_t)$ is the marginal product of labor.

We first address the issue of capital taxation. Equations (21), (23) and (24) imply that

$$u'(c^i(s^t)) = \beta \sum_{s_{t+1}} u'(c^i_{t+1}(s^{t+1}))R^*(s^{t+1})\Pr(s_{t+1}|s^t), \tag{25}$$

which is the standard undistorted intertemporal Euler equation for individuals. Hence, capital income should go untaxed.[16]

Define the shadow marginal tax rate on labor as the solution to

$$\frac{v'\left(\frac{L^i(s^t)}{\theta^i}\right)}{u'(c^i(s^t))} = (1-\tau^i(s^t))\theta^i W^*(s^t)$$

It follows from equations (20) and (21) that $\tau^h(s^t) = 0$, so that the highest type is not distorted—a standard result in optimal tax theory. As for the low type, combining equations (22) and (23) gives

$$\tau^l(s^t) = \mu \frac{1-(\theta^l/\theta^h)\,\varepsilon(L^l(s^t))}{\gamma-(\theta^l/\theta^h)\mu\,\varepsilon(L^l(s^t))},$$

where $\varepsilon(L) \equiv v'(L/\theta^h)/v'(L/\theta^l)$. Thus, the tax rate is constant if $\varepsilon(L)$ is independent of L, which is true if $v(n) = \alpha n^\gamma/\gamma$.

Proposition 3 *At the constrained efficient allocation: (a) capital accumulation is not distorted, that is, the standard Euler equation (25) holds; (b) if $v(n) = \alpha n^\gamma/\gamma$, each type's marginal income tax is constant: $\tau^i(s^t) = \bar{\tau}^i$ for all t and s^t.*

This result provides an interesting benchmark for zero capital taxation and constant marginal tax rates on labor income. Tax smoothing is optimal in that tax rates for a given worker are constant across time and states, as in the case of linear taxation. However, here the optimal tax schedule is nonlinear in the sense that marginal tax rates vary across individuals.

Income Tax Averaging. The analysis suggests tax schemes that equate marginal tax rates over time but still allow them to vary across individuals. One arrangement with such a feature is taxation based on income averages, as opposed to only current income. Such rules were advocated by Vickery (1947) for different reasons.[17]

[16] Distorting the standard Euler condition is optimal if there are ensuing privately observed productivity shocks at the individual level (Diamond and Mirrlees, 1977; Rogerson, 1985; Werning, 2002; Golosov, Kocherlakota, and Tsyvinski, 2003; Farhi and Werning, 2005).

[17] Vickery (1947) argued based of horizontal equity: if the tax schedule is convex then individuals with highly fluctuating earnings would otherwise pay more taxes on average than individuals with steadier earnings. In our case, they serve to implement constant marginal income tax rates over time, while retaining the non-linearities across individuals.

We briefly discuss one such simple scheme that works for a deterministic dynamic economy, and which may be suggestive for other cases. The government does not tax capital income and instead sets a non-linear income tax payment as a function of the present value of lifetime earnings. In the first period workers must pay

$$\Psi\left(\sum_{t=0}^{\infty} p_t W_t^* L_t\right),$$

for some typically non-linear tax function Ψ.[18] Worker $i \in I$ maximizes $\sum_t \beta^t (u(c_t) - v(L_t/\theta))$ subject to $\sum_t p_t c_t \geq \sum_t p_t W_t^* L_t - p_0 \Psi\left(\sum_t p_t W_t^* L_t\right)$. The first-order conditions yield

$$\frac{v'\left(\frac{L_{t+1}^i}{\theta^i}\right)}{u'(c_t^i)} = \left(1 - \Psi'\left(\sum_{t=0}^{\infty} p_t L_t^i\right)\right) W_t^*.$$

so that the derivative $\Psi'(\cdot)$ plays the role of the constant marginal tax rate in all periods. Thus, taxation based on income averages automatically ensures that individuals' marginal tax rates are constant over time, as we found for our implicit tax rates in our constrained efficient allocation, while allowing these to vary across individuals. Indeed, it is easy to see that one can always find a smooth Ψ function that implements the allocation.

6 Conclusions

This paper provides a tractable framework to address issues of optimal taxation in dynamic economies. Unlike representative-agent Ramsey models, distortive taxation is microfounded by a concern for redistribution. Our framework is tractable and can handle rich specifications of the dynamic economy, such as those used in representative-agent Ramsey analyses. Indeed, our model nests the standard representative-agent Ramsey model as a special case, and our analytical results are also of interest restricted within this context.

Our results provide interesting benchmarks for perfect tax smoothing of labor income taxes and for zero taxation of capital income. Although the mechanisms and insights are quite different, it is comforting that a microfounded model of taxation does not disturb these cornerstone results in Ramsey tax theory. Our model does suggest a novel source for variations in optimal tax rates—for deviating from perfect tax smoothing. In particular, movements in the relative skill distribution induce changes in the optimal amount of redistribution, and thus, in the optimal tax rate.

[18] It is easy to change things slightly so that the worker does not make a single period in the first period, but instead pays each period as a function of past labor incomes.

Unlike Ramsey models, our model recovers a form of Ricardian equivalence that renders debt management indeterminate. We speculate that extensions of our model that overcome perfect Ricardian neutrality are likely to provide a determinate theory of debt management. One interesting possibility is to model some individuals as having limited participation in asset markets.[19] A simple, but extreme, example might be to suppose that one group of workers is hand-to-mouth, with no initial assets and no access to asset markets whatsoever. The desire to smooth consumption, and hence income net of taxes, for such nonparticipants could then pin down the lump-sum tax component, and public debt with it.

Appendix

A Proof of Proposition 1

Condition (2) directly implies condition (iv) for an equilibrium. Factor prices given by (4)–(5) ensure that the firm maximizes, so that condition (ii) for an equilibrium is met. Conditions (9)–(12) ensure that consumers are maximizing given taxes and prices given by (13)–(14), so that condition (i) for a competitive equilibrium is met. Finally, condition (iii) is automatically met given that the resource constraints and budget constraints of all individuals hold with equality.

B Shocks to Skill Distributions

The first-order conditions from the worker's maximization problem are

$$\begin{aligned} \alpha \left(L_t^i\left(s^t\right)\right)^{\gamma-1} &= \nu^i \left(\theta_t^i\left(s_t\right)\right)^{\gamma} w(s^t)p(s^t)(1-\tau\left(s^t\right)), \\ \left(c_t^i\left(s^t\right)\right)^{-\sigma} &= \nu^i p(s^t), \end{aligned}$$

where ν^i is the multiplier individual $i \in I$ has on its own budget constraint (3). Solving

$$L_t^i\left(s^t\right) = \left(\frac{\nu^i}{\alpha}\left(\theta_t^i\left(s_t\right)\right)^{\gamma}\right)^{\frac{1}{\gamma-1}} \left(w(s^t)p(s^t)(1-\tau\left(s^t\right))\right)^{\frac{1}{1-\gamma}},$$

[19] Another interesting direction may be an overlapping generations framework.

it follows that $L_t^i(s^t) = \omega_L^i(s_t)\, L(s^t)$ with

$$\omega_L^i(s_t) \equiv \frac{\left(\frac{\nu^i}{\alpha}\right)^{\frac{1}{\gamma-1}} \left(\theta_t^i(s_t)\right)^{\frac{\gamma}{\gamma-1}}}{\sum_{i\in I}\left(\frac{\nu^i}{\alpha}\left(\theta_t^i(s_t)\right)^{\gamma}\right)^{\frac{1}{\gamma-1}}\pi^i}.$$

For future reference, note that

$$\left(\frac{\omega_L^i(s_t)}{\theta^i(s_t)}\right)^{\gamma} = \frac{\left(\frac{\nu^i}{\alpha}\right)^{\frac{\gamma}{\gamma-1}-1+1}\left(\theta_t^i(s_t)\right)^{\frac{\gamma}{\gamma-1}}}{\left(\sum_{i\in I}\left(\frac{\nu^i}{\alpha}\left(\theta_t^i(s_t)\right)^{\gamma}\right)^{\frac{1}{\gamma-1}}\pi^i\right)^{\gamma-1+1}} = \frac{\frac{\nu^i}{\alpha}\omega_L^i(s_t)}{\left(\sum_{i\in I}\left(\frac{\nu^i}{\alpha}\left(\theta_t^i(s_t)\right)^{\gamma}\right)^{\frac{1}{\gamma-1}}\pi^i\right)^{\gamma-1}}. \tag{26}$$

Dividing the worker's two first-order conditions by each other (i.e., the intratemporal optimality condition):

$$\theta^i(s_t)^{-\gamma}\frac{\left(\omega_L^i(s_t)\right)^{\gamma-1}}{u'(\omega_c^i)} = \frac{u'(c(s^t))}{v'(L(s^t))} w_t\left(s^t\right)\left(1-\tau\left(s^t\right)\right)$$

Using equation (26) and rearranging it follows that

$$\frac{\nu^i}{\alpha}\frac{1}{u'(\omega_c^i)} = \left(\sum_{i\in I}\left(\frac{\nu^i}{\alpha}\left(\theta_t^i(s_t)\right)^{\gamma}\right)^{\frac{1}{\gamma-1}}\pi^i\right)^{\gamma-1}\frac{u'(c(s^t))}{v'(L(s^t))} w_t\left(s^t\right)\left(1-\tau\left(s^t\right)\right)$$

must be equal to some constant ϕ because one side depends on i but not on s^t, and the other depends on s^t but not on i.

The rest of the analysis proceeds along the same lines as in Section 4. In particular the first-order conditions with respect to the aggregates $c(s^t)$ and $L(s^t)$ are identical to equations (15)–(16), which can be solved for $w_t(s^t)u'(c(s^t))/v'(L(s^t))$. Using equation (13) we then arrive at

$$\begin{aligned}
\tau(s_t) &= 1-\phi\left(\sum_{i\in I}\left(\frac{\nu^i}{\alpha}\left(\theta_t^i(s_t)\right)^{\gamma}\right)^{\frac{1}{\gamma-1}}\pi^i\right)^{1-\gamma}\frac{\sum_{i\in I} u'\left(\omega_c^i\right)\omega_c^i\left(\lambda^i-(1-\sigma)\mu^i\right)}{\sum_{i\in I}\left(\frac{\omega_L^i(s_t)}{\theta^i(s_t)}\right)^{\gamma}\left(\lambda^i-\gamma\mu^i\right)} \\
&= 1-\phi\frac{\sum_{i\in I} u'\left(\omega_c^i\right)\omega_c^i\left(\lambda^i-(1-\sigma)\mu^i\right)}{\sum_{i\in I}\frac{\nu^i}{\alpha}\omega_L^i(s_t)\left(\lambda^i-\gamma\mu^i\right)} \\
&= 1-\frac{\sum_{i\in I} u'(\omega_c^i)\omega_c^i\left(\lambda^i-(1-\sigma)\mu^i\right)}{\sum_{i\in I} u'(\omega_c^i)\omega_L^i(s_t)\left(\lambda^i-\gamma\mu^i\right)},
\end{aligned}$$

where the last equality follows from the definition of ϕ.

References

Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppälä, 2002. "Optimal Taxation without State-Contingent Debt." *Journal of Political Economy*. 110 (6). 5

Albanesi, Stefania, and Christopher Sleet, 2004. "Dynamic Optimal Taxation with Private Information." Discussion Paper 140, Federal Reserve Bank of Minneapolis. 3

Angeletos, George-Marios, 2002. "Fiscal Policy with Non-contingent Debt and the Optimal Maturity Structure." *Quarterly Journal of Economics*. 117 (2): 1105–1131. 5

Barro, Robert, 1979. "On the Determination of the Public Debt." *Journal of Political Economy*. 87 (5): 940–71. 1, 3, 13

Buera, Francisco J., and Juan Pablo Nicolini, 2004. "Optimal Maturity of Government Debt without State Contingent Bonds." *Journal of Monetary Economics*. 4 (3): 531–554. 5

Chamley, Christophe, 1986. "Optimal Taxation of Capital in General Equilibrium." *Econometrica*. 54: 607–22. 1, 2, 10

Chari, V.V., Lawrence J. Christiano, and Patrick J. Kehoe, 1994. "Optimal Fiscal Policy in a Business Cycle Model." *Journal of Political Economy*. 102 (4): 617–52. 1, 2, 5, 11, 16

Diamond, Peter A., and James A. Mirrlees, 1971. "Optimal Taxation and Public Production II: Tax Rules." *American Economic Review*. 61 (3): 261–78. 10

Diamond, Peter A., and James A. Mirrlees, 1977. "A Model of Social Insurance with Variable Retirement." Working papers 210, Massachusetts Institute of Technology, Department of Economics. 18

Farhi, Emmanuel, 2005. "Capital Taxation and Ownership When Markets are Incomplete." Mimeo. 5

Farhi, Emmanuel, and Iván Werning, 2005. "Capital Taxation: Quantitative Explorations of the Inverse Euler Equation." Mimeo. 18

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski, 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*. 70 (3): 569–587. 3, 18

Judd, Kenneth L., 1985. "Redistributive Taxation in a Perfect Foresight Model." *Journal of Public Economics*. 28: 59–83. 1, 2, 10

Lucas, Robert E., Jr., and Nancy L. Stokey, 1983. "Optimal Fiscal and Monetary Policy in an Economy without Capital." *Journal of Monetary Economics*. 12: 55–93. 1, 3, 5, 13

Meltzer, Allan H., and Scott F. Richard, 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy*. 89 (5): 914–27. 3

Mirrlees, James A., 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*. 38 (2): 175–208. 1, 2, 3, 4, 14

Rogerson, William P., 1985. "Repeated Moral Hazard." *Econometrica*. 53 (1): 69–76. 18

Vickery, William S., 1947. *Agenda for Progressive Taxation*, New York: Ronald. 3, 18

Werning, Iván, 2002. "Optimal Dynamic Taxation." PhD Dissertation, University of Chicago. 0, 18

Werning, Iván, 2005. "Tax Smoothing with Incomplete Markets." Mimeo. 5

Zhu, Xiaodong, 1992. "Optimal Fiscal Policy in a Stochastic Growth Model." *Journal of Economic Theory*. 58: 250–89. 1, 2