



Optimal Capital Taxation Revisited

V. V. Chari

University of Minnesota and Federal Reserve Bank of Minneapolis

Juan Pablo Nicolini

Federal Reserve Bank of Minneapolis and Universidad Di Tella

Pedro Teles

Banco de Portugal, Catolica Lisbon SBE, and CEPR

Staff Report 571

September 2018

DOI: <https://doi.org/10.21034/sr.571>

Keywords: Capital income tax; Time consistency; Production efficiency

JEL classification: E60, E61, E62

This is a revised version of Working Paper 752, which circulated with the same title. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, Banco de Portugal, or the European System of Central Banks.

Federal Reserve Bank of Minneapolis • 90 Hennepin Avenue • Minneapolis, MN 55480-0291

<https://www.minneapolisfed.org/research/>

Federal Reserve Bank of Minneapolis
Research Department Staff Report 571

First draft: September 2016. This version: September, 2018

Optimal Capital Taxation Revisited*

V. V. Chari

University of Minnesota
and Federal Reserve Bank of Minneapolis

Juan Pablo Nicolini

Federal Reserve Bank of Minneapolis and Universidad Di Tella

Pedro Teles

Banco de Portugal, Catolica Lisbon SBE, CEPR

ABSTRACT

We revisit the question of how capital should be taxed. We allow for a rich set of tax instruments that consists of taxes widely used in practice, including consumption, dividend, capital, and labor income taxes. We restrict policies to respect promises that the government has made in the previous period regarding the current value of wealth. We show that capital should not be taxed if households have preferences that are standard in the macroeconomics literature. We show that Ramsey outcomes that must respect such promises are time consistent. We show that the presumption in the literature that capital should be taxed for some length of time arises because the tax system is restricted.

Keywords: capital income tax; time consistency; production efficiency

JEL Codes: E60; E61; E62

*We thank Isabel Correia, João Guerreiro, Albert Marcet, Ellen McGrattan, Chris Phelan, Catarina Reis, and Ivan Werning for helpful discussions. Chari thanks the NSF for supporting the research in this paper and Teles is grateful for the support of FCT as well as the ADEMU project, “A Dynamic Economic and Monetary Union,” funded by the European Union’s Horizon 2020 Program under grant agreement 649396. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, Banco de Portugal, or the European System of Central Banks. E-mail addresses: varadarajanvchari@gmail.com, juanpa@minneapolisfed.org, pteles@ucp.pt.

1. Introduction

How should capital income be taxed? How should it be taxed in the long run and along the transition? An influential literature uses the Ramsey approach in the neoclassical growth model to answer these questions.¹ In this approach, the set of tax instruments is exogenously given. Some of this literature (see, for example, Chamley (1986) and Judd (1985)) considers tax systems in which only labor and capital income can be taxed and the tax rate on capital income is restricted to be below an upper bound that is less than or equal to 100%. This literature finds that capital income should be taxed at its maximum level for some length of time but should not be taxed in the steady state. More recently, Straub and Werning (2015) show that it may be optimal to tax capital income at its maximum level forever. This literature leads to the presumption that capital taxes should be high for some length of time.

In this paper, we take the view that the exogenously given set of tax instruments in the Ramsey approach should include taxes widely used in practice in most developed economies. In addition to taxes on capital and labor income, most economies tax dividends, consumption, and wealth. We refer to a tax system that potentially includes all of these taxes as a *rich tax system*. We also assume that the Ramsey planner cannot reduce the value of initial wealth in utility terms below an exogenously specified level and refer to this constraint on the planner as the *wealth constraint*.² The spirit of this wealth constraint is that agents in periods before period zero made decisions based on expectations of the value of their wealth in period zero and policies chosen in period zero should not violate those expectations.

As is well known, with a rich tax system, many tax policies can support the same allocations. This multiplicity issue has led the public finance literature to focus on wedges that the tax system induces in marginal conditions, rather than focusing on the taxes themselves. These considerations lead us to focus on whether the Ramsey policy yields intertemporal wedges, rather than focusing on the level of the capital income tax. We say that capital is not taxed if the Ramsey policy has no intertemporal wedges and that capital is taxed or subsidized depending on the sign of the intertemporal wedge.

¹See Judd (1985), Chamley (1986), Zhu (1992), Chari, Christiano, and Kehoe (1994), Chari and Kehoe (1999), Atkeson, Chari, and Kehoe (1999), Coleman (2000), Bassetto and Benhabib (2006), Werning (2007), and Straub and Werning (2015).

²See Armenter (2008) for a similar formulation.

We show that with a rich tax system, capital should not be taxed in the steady state of the neoclassical growth model. For general preferences, we show that along the transition, capital may be taxed or subsidized. We focus attention on a class of preferences that are standard in the macroeconomics literature. These preferences have constant intertemporal elasticity of substitution in consumption and a constant Frisch elasticity in labor. We show that with these preferences, capital income should never be taxed. These results hold for any value of the initial wealth that the government is exogenously required to deliver. We also consider environments with uncertainty and show that, with standard macroeconomic preferences, capital should not be taxed.

We show that the presumption in the literature that capital income taxes should be high for some, possibly infinite, length of time arises from restrictions imposed on the tax system and that once we allow for a rich tax system, this presumption disappears. Given that our notion of a rich tax system contains taxes used in most countries, and given that macroeconomic models typically use the preferences we study, our analysis implies that conventional macroeconomic theory strongly suggests that tax systems that distort capital accumulation are inefficient.

One way of thinking about our wealth restriction is that the government in the period before the initial period made promises about the value of wealth in the initial period that the Ramsey planner is obliged to respect. Other than this restriction, the planner is free to choose current and future policies.

In this formulation, history matters only to the extent that promises made in the immediately previous period regarding the value of wealth must be respected. Suppose now that in all future periods, history matters only to this extent. That is, the government in each period must respect promises about the value of wealth made by the government in the previous period and can, in turn, make promises about the value of wealth in the next period in addition to choosing current policies. Other than this promise, the government in the current period has no ability to choose future policies. With this form of partial commitment, we then ask a natural question: What is the equilibrium outcome in an environment in which the government in any period can choose current policies as well as the value of wealth that the government in the following period must respect but has no other form of commitment? We

show that the Ramsey outcomes, which have commitment, are Markov equilibrium outcomes in the environment with partial commitment. In this sense, the Ramsey equilibrium with wealth constraints is time consistent. We view this time consistency as a justification for adding wealth constraints to Ramsey problems.

Suppose next that history does not matter at all or, alternatively, history matters only in that the government must respect one-period-ahead promises regarding current policies. Then, in models like ours with capital and debt, it is well known that Ramsey policies are time inconsistent. This time inconsistency problem raises concerns regarding the applicability of an analysis in which history does not matter to applied public policy.

We briefly analyze an economy with heterogeneous agents. This formulation allows for redistributive motives for taxation in addition to the need to raise taxes to finance government spending. For simplicity, we assume that the economy has two types of agents that differ on the level of wealth and possibly on preferences. We begin by considering tax systems that do not allow for type-specific taxes. We show that if both types of agents have identical standard macro preferences, it is optimal to never tax capital (see Werning (2007) for a similar result without wealth restrictions). If instead, preferences for each type of agent belong to the standard preference class but are different across the types of agents, it is optimal to distort capital accumulation. With type-specific tax rates, we show that even if preferences are different across agents, it is optimal to never tax capital.

Our result that it is not optimal to tax capital with standard preferences is related to results on uniform commodity taxation (Atkinson and Stiglitz, (1972)). Standard preferences are separable and homothetic in consumption and labor. With these preferences, the growth model can be recast as a model in which constant returns to scale technologies are used by competitive firms to produce one final composite consumption good and one composite labor input. The Ramsey planner in the recast economy faces a wealth constraint. We show that it is optimal to not distort the use of intermediate goods. These intermediate goods consist of consumption, labor, and capital at each date in the original economy. This result implies that in the original economy, capital income should never be taxed. This result is similar to the production efficiency result in Diamond and Mirrlees (1971) but is different in that they require full taxation of pure rents to obtain production efficiency, while with our wealth

constraint, we do not require such full taxation.

Our notion of zero taxation of capital is not equivalent to production efficiency. We show this result by demonstrating that with general preferences and a rich tax system, the Ramsey allocations are production efficient but do not have to have zero taxation of capital as we have defined it. The Ramsey allocation is production efficient because a rich tax system allows for taxes on all final consumption goods and on all types of labor, and allows pure rents to be taxed so as to meet the wealth constraint. The Ramsey allocation with general preferences typically does not have uniform taxation of consumption goods or labor types. Such uniform taxation is needed to achieve zero taxation of capital as we have defined it.

This recasting also allows us to develop a deeper understanding of our results in the heterogeneous agents economy. The recast heterogeneous agents economy with intermediate goods now has two final composite consumption goods and two final composite labor inputs. Each of these final goods corresponds to the consumption and labor input of each type of agent. If the preferences of the two agents are the same, the production technologies for the composite goods are identical, and it is optimal to tax the goods at the same rate. If instead, the preferences are different, in general the goods need to be taxed at different rates. If those tax rates are required to be the same, then production efficiency is not obtained in the recast economy. In the original economy, it may be optimal to tax or subsidize capital.

Our result that it is not optimal to tax or subsidize capital clearly conflicts with the general presumption in the literature that it is optimal to tax capital for some length of time. The difference in these results arises for two reasons. First, the literature restricts initial policies rather than the value of initial wealth. Second, the literature allows for restricted tax systems that tax only capital and labor income with a cap on capital tax rates, while we consider rich tax systems. It turns out that the restriction on initial policies rather than the value of initial wealth plays a relatively small role in the difference in results. We show this small role by considering an optimal taxation problem with a rich tax system and with restrictions on initial policies. We show that, with standard preferences and a rich tax system, capital is taxed for at most one period and is never taxed after the first period. This result is in stark contrast to the presumption in the literature.

One way of getting intuition for the presumption in the literature is to begin by

noting the well-known result that Ramsey planners without wealth constraints seek to tax away pure rents completely. In the growth model, these pure rents consist of the value of the wealth in utility terms. This value of wealth is the product of the initial marginal utility of consumption and the wealth in units of initial consumption goods. A Ramsey planner who cannot directly confiscate wealth in goods terms has a strong incentive to reduce the initial marginal utility of consumption, so as to indirectly confiscate the value of wealth. With a restricted tax system that allows taxation only of labor and capital, and with a bound on the capital tax rate, setting the capital income tax rate at its upper bound forever reduces the initial marginal utility of consumption by the greatest amount. The planner trades off the gain from this indirect confiscation with the losses from the induced intertemporal distortions. This trade-off determines the length of time that capital income taxes are set to the upper bound.

The central lesson of the public finance literature stemming from Ramsey (1927) and Diamond and Mirrlees (1971) is that tax systems that include taxes on all final consumption goods and taxes on all primary inputs (such as labor) and tax away all pure rents yield production efficiency. We have extended this result to environments in which the Ramsey planner faces a wealth constraint that limits the ability to fully tax away pure rents. Our notion of zero taxation of capital is stronger than production efficiency but follows from it for the kinds of preferences that are standard for the macroeconomics literature. These observations lead to our main result that standard macroeconomic models imply that capital taxation is inefficient if the planner has access to a rich tax system. These observations also lead us to conclude that systems that do not allow for a rich tax system may find capital taxation optimal but are of limited interest from an applied perspective, given that most countries already use the taxes that constitute a rich tax system.

2. A representative agent economy

Our benchmark framework is the deterministic neoclassical growth model with taxes. The representative household's preferences are defined over consumption c_t and labor n_t ,

$$(1) \quad U = \sum_{t=0}^{\infty} \beta^t u(c_t, n_t),$$

satisfying the usual properties. The production technology is described by

$$(2) \quad c_t + g_t + k_{t+1} - (1 - \delta) k_t \leq F(n_t, k_t),$$

where k_t is capital, g_t is exogenous government consumption, δ is the depreciation rate, and the production function F is constant returns to scale.

We now describe a competitive equilibrium with taxes. The government finances public consumption and initial debt, b_0 , with time-varying proportional taxes. We allow for a rich tax system that includes taxes on consumption τ_t^c , labor income τ_t^n , capital income τ_t^k , dividends τ_t^d , and a tax on initial wealth, l_0 .³

Capital accumulation is conducted by firms. Given that the technology is constant returns to scale, we assume without loss of generality that the economy has a representative firm. The household owns the firm and receives dividends.⁴ We now describe the household's and firm's problems and define a competitive equilibrium.

Households The representative household maximizes utility (1), subject to the present-value budget constraint

$$(3) \quad \sum_{t=0}^{\infty} q_t [(1 + \tau_t^c) c_t - (1 - \tau_t^n) w_t n_t] \leq (1 - l_0) [b_0 + \sum_{t=0}^{\infty} q_t (1 - \tau_t^d) d_t],$$

where q_t is the price of one unit of the good produced in period t in units of the good in period zero, so that $q_0 = 1$; w_t is the pretax wage rate; b_0 is the initial holdings of government debt; and d_t are the dividends paid by the firm.

Firms The representative firm maximizes the after tax present value of dividends

$$(4) \quad \sum_{t=0}^{\infty} q_t (1 - \tau_t^d) d_t,$$

³Note that we allow only for a tax on wealth in period zero. It turns out that allowing for taxes on wealth in future periods is equivalent to a consumption tax. Since we allow for consumption taxes, taxes on future wealth are redundant.

⁴In Appendix A, we describe an alternative, more widely used decentralization in which the households own the capital stock and firms rent capital from the households. The two decentralizations are equivalent, but it is easier to relate the taxes in the decentralization described here to the ones in existing tax systems.

where dividends, d_t , are given by

$$(5) \quad d_t = F(k_t, n_t) - w_t n_t - \tau_t^k [F(k_t, n_t) - w_t n_t - \delta k_t] - [k_{t+1} - (1 - \delta)k_t].$$

Note that the taxes on capital income, τ_t^k , are levied on income net of depreciation. Note also that the tax on dividends, τ_t^d , effectively allows firms to expense gross investment. This expensing turns out to imply that, as we show below, dividend taxes are similar to consumption taxes.

In this way of setting up the competitive equilibrium, dividends are net payments to claimants of the firm. These payments could be interpreted either as payments on debt or as payments to equity holders. To clarify this interpretation, consider an all-equity firm. In this case, our notion of dividends consists of cash dividends plus stock buybacks less issues of new equity. In particular, under this interpretation, taxes on capital gains associated with stock buybacks are assumed to be levied on accrual and at the same rate as cash dividends. Note also that dividends could be negative if returns to capital are smaller than investment. In this case, a positive tax on dividends would represent a subsidy to the firm.⁵

Remark: Note that the taxes paid by the firms are on accrued profits rather than on imputed profits. With taxation on accrued profits, if, in some period t , $\tau_t^k > 1$, the solution of the firm's problem would not be interior. This observation suggests that it may be reasonable to impose a restriction that $\tau_t^k \leq 1$. Similarly, it may be reasonable to impose restrictions on dividend taxes so that the present value of after-tax dividends must be nonnegative. If the taxes are on imputed profits, then it may be possible to have an interior solution without these restrictions. In what follows, we analyze equilibria with and without such restrictions.

Government

The government budget constraint is

$$\sum_{t=0}^{\infty} q_t (\tau_t^c c_t + \tau_t^n w_t n_t + \tau_t^d d_t - g_t) + l_0 [b_0 + \sum_{t=0}^{\infty} q_t (1 - \tau_t^d) d_t] = 0.$$

⁵In a steady state of the competitive equilibrium, it is possible to show that dividends are always positive.

Competitive and Ramsey equilibrium

A *competitive equilibrium* is a set of allocations $\{c_t, n_t, k_{t+1}, d_t\}$, prices $\{q_t, w_t\}$, and policies $\{\tau_t^c, \tau_t^n, \tau_t^d, \tau_t^k, l_0\}$, given $\{k_0, b_0\}$ such that the households maximize utility subject to their constraints, firms maximize the present value of dividends, the government budget constraint is satisfied, and markets clear in that resource constraints (2) are satisfied. We refer to a subset of the allocations $\{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$ as *implementable allocations* if they are part of a competitive equilibrium.

A *Ramsey equilibrium* is the competitive equilibrium that yields the highest utility for the representative household. The *Ramsey allocation* is the associated implementable allocation.

In order to characterize the Ramsey equilibrium, we begin by deriving the conditions that any competitive equilibrium must satisfy. The first-order conditions of the households' problem include

$$(6) \quad -\frac{u_{c,t}}{u_{n,t}} = \frac{(1 + \tau_t^c)}{(1 - \tau_t^n) w_t}, \quad t \geq 0,$$

$$(7) \quad \frac{u_{c,t}}{1 + \tau_t^c} = \frac{q_t}{q_{t+1}} \frac{\beta u_{c,t+1}}{(1 + \tau_{t+1}^c)}, \quad t \geq 0,$$

where $u_{c,t}$ and $u_{n,t}$ denote the marginal utilities of consumption and labor in period t .

The first-order conditions of the firm's problem include

$$(8) \quad w_t = F_{n,t} \quad \text{and}$$

$$(9) \quad \frac{q_t}{q_{t+1}} = \frac{(1 - \tau_{t+1}^d) [1 + (1 - \tau_{t+1}^k) (F_{k,t+1} - \delta)]}{1 - \tau_t^d},$$

where $F_{n,t}$ and $F_{k,t}$ denote the marginal products of capital and labor in period t .

Substituting for d_t from (5) and using (8) and (9), it is possible to show that the present discounted value of dividends is given by

$$(10) \quad \sum_{t=0}^{\infty} q_t (1 - \tau_t^d) d_t = (1 - \tau_0^d) [1 + (1 - \tau_0^k) (F_{k,0} - \delta)] k_0.$$

The budget constraint (3) can then be written as

$$(11) \quad \sum_{t=0}^{\infty} q_t [(1 + \tau_t^c) c_t - (1 - \tau_t^n) w_t n_t] = W_0,$$

where the initial wealth of the households is given by

$$(12) \quad W_0 \equiv (1 - l_0) [b_0 + (1 - \tau_0^d) [1 + (1 - \tau_0^k) (F_{k,0} - \delta)] k_0].$$

The full set of equilibrium conditions can then be summarized by the household's first-order marginal conditions (6) and (7), the firm's conditions first-order (8) and (9), the budget constraint (11) with (12), together with the expression for dividends (5) and the market clearing condition (2). The government's budget constraint is implied by the household budget constraint and market clearing.

Implementability These equilibrium conditions can be used to provide a compact characterization of the set of implementable allocations. Substituting prices and taxes from the first-order conditions for the households into the households' budget constraint (11), we obtain that any competitive equilibrium must satisfy the following *implementability constraint*:

$$(13) \quad \sum_{t=0}^{\infty} \beta^t (u_{c,t} c_t + u_{n,t} n_t) = \mathcal{W}_0, \text{ where}$$

$$(14) \quad \mathcal{W}_0 = \frac{u_{c,0}}{1 + \tau_0^c} W_0.$$

Clearly, any competitive equilibrium must also satisfy the resource constraints (2). Thus, we have shown that any competitive equilibrium must satisfy the implementability constraint, (13), and the resource constraints (2).

Next we show that given any arbitrary allocation and period zero policies that satisfy (13) and (2), it is possible to construct prices and policies so that these outcomes constitute a competitive equilibrium. Consider one such implementation. Pin down the wage rates w_t from (8). Set the consumption tax rate to zero, $\tau_t^c = 0$ for all $t \geq 1$. Pin down the tax rate

on labor τ_t^n from (6). Set the intertemporal prices q_t for $t \geq 1$ from (7). Set $\tau_t^k = 0$ for $t \geq 1$. Given τ_0^d , pin down the time path of dividend taxes, τ_t^d , $t \geq 1$, from (9). Obtain dividends d_t from (5). It is immediate that these allocations satisfy all the equilibrium conditions for households and firms. Thus, the so-constructed allocation, prices, and policies are a competitive equilibrium. We then have the following

Proposition 1: (Characterization of the implementable allocations) Any implementable allocation satisfies the implementability constraint (13) and the resource constraints (2). Furthermore, if a sequence $\{c_t, n_t, k_{t+1}\}$, initial conditions k_0, b_0 , and period zero policies $(\tau_0^c, \tau_0^d, \tau_0^k, l_0)$, satisfy (13) and (2), it is implementable.

Wedges and multiple implementations In proving this proposition, we used one particular implementation of policies. We emphasize that any equilibrium allocation can be implemented with numerous other policies. To see this result, note that any competitive equilibrium pins down wedges together with initial wealth. The wedges are implicitly given by an *intra-temporal wedge*,

$$(15) \quad -\frac{u_{c,t}}{u_{n,t}} = \frac{(1 + \tau_t^c)}{(1 - \tau_t^n) F_{n,t}},$$

a *consumption intertemporal wedge*,

$$(16) \quad \frac{u_{c,t}}{\beta u_{c,t+1}} = \frac{(1 - \tau_{t+1}^d)(1 + \tau_t^c)}{(1 - \tau_t^d)(1 + \tau_{t+1}^c)} [1 + (1 - \tau_{t+1}^k)(F_{k,t+1} - \delta)],$$

and a *labor intertemporal wedge*,

$$(17) \quad \frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{(1 - \tau_{t+1}^d)(1 - \tau_t^n)}{(1 - \tau_t^d)(1 - \tau_{t+1}^n)} \frac{F_{n,t}}{F_{n,t+1}} [1 + (1 - \tau_{t+1}^k)(F_{k,t+1} - \delta)].$$

Note that the labor intertemporal wedge condition, (17), is implied by (15) and (16). We include it here to analyze when it is optimal to not distort the labor intertemporal margin.

Remark: Notice that a constant dividend tax does not distort any of the marginal conditions. Such a tax of course raises revenues by reducing the value of the firm at the beginning of period zero, which in turn reduces the household's wealth, as can be seen from

(12). In this sense, a constant dividend tax is equivalent to a levy on the initial capital stock. This dividend tax resembles the tax proposed by Abel (2007) as a way of collecting lump-sum revenue from the taxation of the initial capital stock. Note also that a constant consumption tax does not distort intertemporal conditions but does reduce the value of initial wealth, as can be seen from (14). Notice also that a tax on capital income distorts intertemporal decisions in the same way as do time-varying taxes on consumption, dividends, and labor income. We will use these properties in implementing the Ramsey equilibrium.

To see how a competitive equilibrium can be implemented in multiple ways, consider alternative implementations of some arbitrary competitive equilibrium.

Consider first an alternative implementation that uses a system that levies taxes only on consumption, labor, and initial wealth. We refer to such a system as the *Diamond-Mirrlees system* because it is in the spirit of their tax system that allows taxes only on final goods, primary inputs, and pure rents. Clearly, τ_t^c and τ_t^n can be chosen to satisfy (15)-(17), and l_0 can be chosen to yield the same initial wealth as in the arbitrary equilibrium.

Consider next a version of the implementation used in the proof of Proposition 1 that levies taxes only on labor income and dividends. We refer to this system as the *Abel system* because it resembles the proposal in Abel (2007). Here, τ_t^d and τ_t^n can be chosen to satisfy (15)-(17), and τ_0^d can be chosen to yield the same initial wealth as in the arbitrary equilibrium. This implementation may require τ_0^d to be greater than 100%. So if we impose the restriction that $\tau_t^d \leq 1$, it may not be possible to implement some competitive equilibria.

Finally, consider an alternative implementation that uses taxes only on labor and capital income referred to as the *Chamley-Judd system*. Again, clearly, τ_t^n and τ_{t+1}^k can be chosen to satisfy (15)-(17), and τ_0^k in the alternative implementation can be chosen to yield the same initial wealth as in the arbitrary equilibrium. Analogously to the Abel implementation, note that in this implementation, the tax rates on capital income may need to be greater than one. So if we impose the restriction that $\tau_t^k \leq 1$, it may not be possible to implement some competitive equilibria.

Intertemporal distortions We turn now to the question of whether capital should be taxed. Given our results on multiple implementations, it is clear that setting capital

income taxes to zero does not mean that the economy has no intertemporal wedges. We will say that capital income is not taxed if the Ramsey allocation has no intertemporal wedges. Formally, a competitive equilibrium has *no intertemporal distortions in consumption* from period s onward if there is no wedge in (16) in that

$$(18) \quad \frac{u_{c,t}}{\beta u_{c,t+1}} = 1 + F_{k,t+1} - \delta, \text{ for all } t \geq s.$$

Similarly, a competitive equilibrium has *no intertemporal distortions in labor* from period s onward if there is no wedge in (17) if

$$(19) \quad \frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{F_{n,t}}{F_{n,t+1}} (1 + F_{k,t+1} - \delta) \text{ for all } t \geq s.$$

Finally, a competitive equilibrium has *no taxation of capital* from period s onward if (18) and (19) hold.

Note that it follows from (16) and (17) that no taxation of capital implies constant intratemporal distortions in (15).

A. Ramsey equilibrium

Given Proposition 1, it follows that the Ramsey allocation, together with period zero policies, maximizes utility subject to (13) and (2). We assume that the Ramsey planner faces a wealth constraint in the sense that households must be allowed to keep an exogenous value of initial wealth $\bar{\mathcal{W}}$, measured in units of utility. Specifically, we impose the following restriction on the Ramsey problem:

$$(20) \quad \mathcal{W}_0 \geq \bar{\mathcal{W}},$$

which we refer to as the *wealth restriction in utility terms*.

With this restriction, policies, including initial policies, can be chosen arbitrarily but the households must receive a value of initial wealth in utility terms of $\bar{\mathcal{W}}$ (see Armenter (2008) for an analysis with such a restriction). The spirit of this restriction is that agents in periods before period zero made decisions based on expectations of the value of their wealth

in period zero, and policies chosen in period zero should not violate those expectations. We make this idea precise in the section on partial commitment below.

We now characterize the first-order necessary conditions for an interior solution to the Ramsey problem. These are

$$(21) \quad -\frac{u_{c,t}}{u_{n,t}} = \frac{1 + \varphi(1 + \sigma_t^n - \sigma_t^{nc})}{1 + \varphi(1 - \sigma_t - \sigma_t^{cn})} \frac{1}{F_{nt}}, \quad t \geq 0,$$

$$(22) \quad \frac{u_{c,t}}{\beta u_{c,t+1}} = \frac{1 + \varphi(1 - \sigma_{t+1} - \sigma_{t+1}^{cn})}{1 + \varphi(1 - \sigma_t - \sigma_t^{cn})} (1 + F_{k,t+1} - \delta), \quad t \geq 0,$$

$$(23) \quad \frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{1 + \varphi(1 + \sigma_{t+1}^n - \sigma_{t+1}^{nc})}{1 + \varphi(1 + \sigma_t^n - \sigma_t^{nc})} \frac{F_{n,t}(1 + F_{k,t+1} - \delta)}{F_{n,t+1}}, \quad t \geq 0,$$

together with the implementability and resource constraints. Here,

$$\sigma_t = -\frac{u_{cc,t}C_t}{u_{c,t}}, \quad \sigma_t^n = \frac{u_{nn,t}N_t}{u_{n,t}}, \quad \sigma_t^{nc} = -\frac{u_{nc,t}C_t}{u_{n,t}}, \quad \sigma_t^{cn} = -\frac{u_{cn,t}N_t}{u_{c,t}},$$

and φ is the multiplier of the implementability condition.

These conditions make it clear that the optimal wedges depend on their own and cross elasticities of consumption and labor. If those elasticities are constant, it is optimal to not have intertemporal distortions. In this case, intratemporal wedges are constant and in general positive. If the elasticities are not constant over time, it is optimal to have intertemporal distortions, but whether it is optimal to effectively tax or subsidize capital accumulation depends on whether elasticities are increasing or decreasing over time.

Note that if consumption and labor are constant over time, then the relevant elasticities are also constant, so that it is optimal to have no intertemporal distortions. This observation leads to the following well-known proposition.

Proposition 2: (No intertemporal distortions in the steady state) If the Ramsey equilibrium converges to a steady state, it is optimal to have no intertemporal distortions asymptotically.

Consider now preferences that are standard in the macroeconomics literature. These

preferences take the form

$$(24) \quad U = \sum_{t=0}^{\infty} \beta^t \left(\frac{c_t^{1-\sigma} - 1}{1-\sigma} - \eta n_t^\psi \right).$$

In this case, the elasticities are constant, so that we have the following proposition.

Proposition 3: (No intertemporal distortions ever) Suppose that preferences are given by (24) and the wealth restriction (20) must be satisfied. Then, the Ramsey solution has no intertemporal distortions for all $t \geq 0$.

Proposition 3 extends in a straightforward manner to environments with uncertainty. In Appendix B, we extend our deterministic model to fluctuations in government spending and technology. There we show that the analog of Proposition 3 holds.

Remark: For general preferences, it is difficult to prove that the economy converges to a steady state. For standard preferences, it is straightforward to prove that it does so.

Note that the preferences above are separable and homothetic in both consumption and labor. (In Appendix C, we show that they are the only time-separable preferences with those properties.) We use these properties in Section 4 to relate our results to those on uniform commodity taxation and production efficiency.

The Ramsey outcomes characterized in Proposition 3 can be implemented with a variety of systems. Each of these systems is a restricted version of our rich tax system. Some of these restricted tax systems allow for implementation of our Ramsey equilibrium for any initial conditions, while others allow for implementations for only some set of initial conditions. The Diamond-Mirrlees system, which allows for taxes on consumption, labor, and initial wealth, can implement the Ramsey equilibrium for any initial conditions. For example, one implementation has constant tax rates on consumption, sets the initial wealth tax to satisfy the wealth constraint, and sets the labor tax to zero, while another implementation has constant tax rates on labor, sets the initial wealth tax appropriately, and sets consumption taxes to zero.

Next we consider systems that allow for implementations only for a subset of initial conditions. Consider first a Diamond-Mirrlees system with a zero wealth tax. If W_0 and \bar{W} are of the same sign, then a consumption tax acts in exactly the same fashion as a wealth

tax, so that such a system can implement the Ramsey outcome. For example, a constant consumption tax could be set so as to have the same effect as the wealth tax, and the labor tax could be set appropriately to satisfy the intratemporal condition. If W_0 and \bar{W} are of different signs, then a wealth tax greater than one can implement the Ramsey equilibrium, while the consumption tax cannot. Thus, this system implements the Ramsey outcome only for a subset of initial conditions.

Consider next an Abel system. Again, if W_0 and \bar{W} are of the same sign, an Abel system with an unrestricted dividend tax can implement the Ramsey outcome because the dividend tax is similar to the wealth tax. As with the Diamond and Mirrlees implementation, a constant dividend tax could be set so as to have the same effect as the wealth tax, and the labor tax could be set to satisfy the intratemporal condition. If dividend taxes are restricted to be less than 100%, then the Abel system may not be able to implement the Ramsey allocation. To see this, let W_0 and τ_0^c denote the initial wealth in units of goods and the initial tax rate on consumption in the Diamond-Mirrlees implementation without a wealth tax. Then the Abel system implements the Ramsey outcome if and only if the following condition is met:

$$(25) \quad \frac{W_0}{1 + \tau_0^c} \geq b_0.$$

Note that $W_0 > b_0$, so that the Abel system implements the Ramsey outcome if the effective tax on initial wealth arising from the tax on consumption τ_0^c is not too large.

As with the Abel system, the Chamley-Judd system can implement the Ramsey outcome if the capital income tax is unrestricted. Here, the labor tax implements the intratemporal wedge, and if $F_{k,0} - \delta > 0$, the initial capital income tax can be used to satisfy the initial wealth constraint. If capital income taxes are restricted to be less than 100%, then this system can implement the Ramsey outcome if and only if the following condition is met:

$$(26) \quad \frac{W_0}{1 + \tau_0^c} \geq b_0 + k_0.$$

Notice that condition (25) is stronger than condition (26). The reason for this difference is that the dividend taxes are levied both on the period zero net capital income and the

value of the capital, while the capital income tax is levied only on the period zero net capital income.

Debt of multiple maturities Next we analyze Ramsey equilibria when the inherited debt has longer maturities than the single maturity debt that we have assumed so far. Let b_t^0 be the amount of debt inherited in period 0 that matures in period t . Then initial wealth is given by

$$W_0 = (1 - l_0) \left[\sum_{t=0}^{\infty} q_t b_t^0 + (1 - \tau_0^d) [k_0 + (1 - \tau_0^k) (F_{k,0} - \delta) k_0] \right].$$

The implementability condition is the same as (13), with

$$\mathcal{W}_0 = (1 - l_0) \left[\sum_{t=0}^{\infty} \frac{\beta^t u_{c,t}}{(1 + \tau_t^c)} b_t^0 + \frac{u_{c,0} (1 - \tau_0^d)}{(1 + \tau_0^c)} [1 + (1 - \tau_0^k) (F_{k,0} - \delta)] k_0 \right].$$

Clearly, the solution of the Ramsey problem is the same as in the model with one-period debt. Note that the same reasoning applies if the government is committed to lump-sum transfers in future periods.

B. Ramsey equilibria with restrictions on taxes

Here we relate our results to an extensive and influential literature. This literature differs from our analysis in two ways. First, the literature typically imposes restrictions on initial policies, as opposed to our wealth restriction. Second, it considers tax systems that are more restricted than our rich tax system. For example, Chamley (1986), Judd (1985), and Straub and Werning (2015) consider systems in which the only taxes allowed are taxes on capital and labor income and in which the tax rate on capital is restricted to be below an upper bound, typically 100% in both the initial period and subsequent periods. This literature finds that the optimal tax rate on capital income is at its upper bound for some length of time. Straub and Werning (2015) show that the tax rate on capital can be at its upper bound forever.⁶ While both types of restrictions play a role in the results in the

⁶Bassetto and Benhabib (2006) obtain a similar result in a political economy model.

literature, it turns out that restricted tax systems play a much more important role than do restrictions on initial policies.

Consider first a rich tax system with restrictions on initial policies. Specifically, we assume that l_0 , τ_0^d , τ_0^k , and τ_0^c are exogenously given.⁷ The implementability constraint becomes

$$(27) \quad \sum_{t=0}^{\infty} \beta^t (u_{c,t} c_t + u_{n,t} n_t) = \frac{u_{c,0} (1 - l_0)}{1 + \tau_0^c} [b_0 + (1 - \tau_0^d) [1 + (1 - \tau_0^k) (F_{k,0} - \delta)] k_0].$$

The Ramsey problem is to maximize utility subject to (2) and (27). The first-order conditions of the Ramsey problem are the same as before in (21), (22), and (23), for all $t \geq 1$. The other first-order conditions for period zero are different from those in our benchmark problem. The intertemporal condition for consumption between periods zero and one is now

$$(28) \quad \frac{u_{c,0}}{\beta u_{c,1}} = \frac{1 + \varphi (1 - \sigma_1 + \sigma_1^{cn})}{1 + \varphi \left(1 - \sigma_0 + \sigma_0^{cn} + \frac{\sigma_0 \mathcal{V}}{c_0} \right)} (1 + F_{k,1} - \delta),$$

for $\mathcal{V} = ((1 - l_0) / (1 + \tau_0^c)) [b_0 + (1 - \tau_0^d) [1 + (1 - \tau_0^k) (F_{k,0} - \delta)] k_0]$. We omit the intratemporal condition for period zero since it is not used in deriving our main results.

With standard macro preferences, since elasticities are constant over time, it is optimal to have no intertemporal distortions from period one onward. Consider now intertemporal distortions in period zero. With standard macro preferences, $\sigma_1 = \sigma_0$ and cross elasticities are zero, so that if $\mathcal{V} > 0$, (28) implies that

$$(29) \quad \frac{u_{c,0}}{\beta u_{c,1}} < 1 + F_{k,1} - \delta.$$

Comparing (16) with (29), we see that the effective implied tax rate on capital income in period one is strictly positive. One intuition for this result is as follows. The Ramsey planner finds it optimal to reduce the right side of the implementability constraint (27) or, equivalently, the value of the household wealth in utility terms. This value can be reduced

⁷Alternatively, we could have assumed that each of these tax rates has an upper bound, in which case the solution would be to trivially set them equal to their upper bounds.

by decreasing the marginal utility of period zero consumption. This decrease is achieved by inducing households to increase their period zero consumption relative to consumption in all future periods. We summarize this discussion in the following proposition.

Proposition 4: (No intertemporal distortions after one period) Suppose preferences satisfy (24) and initial policies are exogenously specified, with no wealth restriction. Then the Ramsey solution has no intertemporal distortions for all $t \geq 1$. If $\mathcal{V} > 0$, it is optimal to effectively tax capital accumulation from period zero to period one.

As usual, the Ramsey outcome can be implemented in a variety of ways. One implementation uses dividend and labor income taxes alone, together with the exogenously given initial policies. In this implementation, labor income taxes are set to satisfy the intratemporal wedge condition (15). Note that, for $t \geq 1$, the labor income tax rate is constant. Dividend taxes are set to satisfy the intertemporal wedge condition (16). From this condition, it is clear that the dividend tax is below one in period one and is zero thereafter.

An alternative implementation uses consumption and labor income taxes alone. Consumption taxes are set to satisfy the intertemporal wedge condition (16), and they are constant starting in period one. Given these consumption taxes, labor income taxes are set to satisfy (15). Inspecting (16) and (29), we clearly see that the consumption tax rate in period one, τ_1^c , must be greater than the tax rate in period 0, τ_0^c . Indeed, it is possible that τ_1^c is so large that the associated labor income tax rate needed to satisfy the intratemporal condition (15) might be negative. This possibility may be a disadvantage for a consumption tax implementation. Note that the dividend tax implementation does not have this disadvantage.

The third implementation uses labor and capital income taxes alone, together with the exogenously given initial policies. This implementation is the one widely used in the literature. As in the other implementations, the labor tax is used to satisfy (15). The capital income tax rate is used to satisfy (16). Note that, for $t \geq 2$, the capital income tax rate is zero. Inspecting (16) and (29), we see that it is possible that the capital income tax rate in period one, τ_1^k , may be greater than 100%. The reason why the dividend tax is bounded below one with the dividend tax implementation, while in this implementation the capital income tax may be greater than one, is that the dividend tax is a tax on the gross return on capital, while the capital income tax is a tax on the net return.

These three implementations show that, with a rich tax system, intertemporal decisions are distorted for one period at most. This finding implies that the results in the literature arise not just from restrictions on initial policies but also from departures from a rich tax system.

To understand the role of restrictions on the tax system, consider a tax system that is restricted in that only capital and labor income can be taxed and that the tax rate on capital income is restricted to be below an exogenously specified level $\bar{\tau} \leq 1$. Rearranging (16), it is immediate that this additional restriction imposes additional constraints on the Ramsey problem given by

$$(30) \quad \frac{u_{c,t}/\beta u_{c,t+1} - 1}{F_{k,t+1} - \delta} \geq 1 - \bar{\tau}, \text{ for all } t.$$

In this case, the Ramsey problem is to choose allocations $\{c_t, n_t, k_{t+1}\}$ and τ_0^k to maximize utility subject to the resource constraints (2), the implementability constraint (27), and (30). We follow the literature in setting $l_0 = \tau_0^c = \tau_0^d = 0$. In addition, we assume that $\bar{\tau}$ is not so high that the government can finance the present value of expenditures and the initial debt purely with the tax on capital income in period zero.

The constraint (30) may bind for a finite number of periods as in Chamley (1986) or forever as in Straub and Werning (2015). Straub and Werning (2015) set $\bar{\tau}$ to be 100% and show that the optimal solution for particularly high levels of initial debt may be to have the capital income tax set at 100% forever.

To obtain some intuition for these results, notice that the planner has a strong incentive to make $u_{c,0}$ small so as to reduce the right side of the implementability constraint, (27). Since this right side can also be reduced by confiscating capital, we refer to this incentive as the *confiscation motive*. In determining the optimal tax rates, the planner trades off the gains from the confiscation motive against the losses from intertemporal distortions. To understand this trade-off, consider the intertemporal condition

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = 1 + (1 - \tau_{t+1}^k) (F_{k,t+1} - \delta).$$

Given $u_{c,1}$, the confiscation motive provides an incentive to make τ_1^k large to reduce $u_{c,0}$. If

the confiscation motive is sufficiently strong, the bound on τ_1^k is met. In this case, the planner has an incentive to make $u_{c,1}$ small to further reduce $u_{c,0}$, thereby confiscating initial wealth to a greater extent. Fixing $u_{c,2}$, $u_{c,1}$ in turn can be made small by making τ_2^k large. Again, if the confiscation motive is sufficiently strong, the upper bound will be met. This recursion suggests that the Ramsey solution will have capital taxes be at the upper bound for a length of time. If the initial debt is sufficiently large, the confiscation motive is very strong, and the length of time could be infinite as pointed out by Straub and Werning (2015).

With, say, dividend taxes, it is possible to reduce $u_{c,0}$ relative to $u_{c,1}$ to an arbitrary extent without distorting intertemporal decisions from period one onward. That is, the after-tax interest rate between period zero and period one can be made negative. With capital income taxes bounded by 100%, $u_{c,t}$ can be reduced relative to $u_{c,t+1}$ only to a limited extent. That is, the after-tax interest rate between any two periods can be reduced to zero at most. The confiscation motive makes it desirable to flatten the entire term structure to zero. If this motive is sufficiently strong, then capital taxes will be 100% forever.

In sum, the results in the literature arise from restrictions on the tax system. These restrictions exclude a multitude of commonly used taxes.

C. Partial commitment equilibria

The notion of a Ramsey equilibrium is developed in an environment in which in period zero, the government commits to an infinite sequence of policies. Here we consider an alternative institutional framework in which the government has partial commitment. We develop a notion of equilibrium for such an environment, referred to as a *partial commitment equilibrium*. In this environment, in any period, governments lack full commitment in the sense that they cannot specify the entire sequence of policies that will be chosen in the future. They do have the ability to constrain the set of policies in the subsequent period. We first consider constraints on one period ahead value of the wealth in utility terms. We then consider constraints on one period ahead policies.

To set the stage for the environments with partial commitment, consider first environments in which the history of past promises is irrelevant. In these environments, it is well known that Ramsey outcomes are typically time inconsistent. For example, suppose

that l_0, τ_t^k, τ_t^d are all restricted to be less than 100%. Here the Ramsey outcome when history is irrelevant is to tax the initial wealth completely and commit not to do so in the future. Clearly the government in period 1 will pursue a policy of taxing wealth away completely and private agents will adjust their wealth accumulation decisions accordingly. Thus, the Ramsey outcome is typically time inconsistent and some form of commitment is needed if Ramsey outcomes are to be time consistent.

Partial commitment to value of wealth We begin by showing that the Ramsey problem can be written in a recursive form. To do so, note that the implementability constraint can be equivalently written as a sequence of implementability constraints of the form

$$(31) \quad \beta \mathcal{W}_{t+1} + u_{c,t} c_t - u_{n,t} n_t = \mathcal{W}_t.$$

together with the limiting condition $\lim_{T \rightarrow \infty} \beta^T \mathcal{W}_{T+1} = 0$. The Ramsey problem is now to maximize utility (1) subject to the sequence of implementability constraints (31) and the resource constraints. Standard dynamic programming arguments as in Stokey and Lucas (1989) imply that this Ramsey problem can be written recursively as

$$(32) \quad V_t(k, \mathcal{W}) = \text{Max } u(c, n) + \beta V_{t+1}(k', \mathcal{W}')$$

subject to

$$(33) \quad c + g_t + k' - (1 - \delta)k \leq F(n, k) \text{ and}$$

$$(34) \quad \beta \mathcal{W}' + u_c c - u_n n = \mathcal{W}.$$

Note that value functions are indexed by time because government expenditures may depend on time.

Consider now an environment with partial commitment in that the government in each period chooses current policies and the value of wealth in utility terms for the following period. The government in the current period must respect the value of wealth that the

previous government has chosen. We develop a notion of a Markov equilibrium with partial commitment on returns, which we call a *nonconfiscatory equilibrium*. The state of the economy in period t is given by $s = \{k, \mathcal{W}\}$. It is convenient and without loss of generality to think of the government in period t as choosing allocations, policies, and prices directly in that period.

Let $\hat{V}_{t+1}(s')$ denote the continuation value induced by the choices of the government in future periods. The government's problem in period t is to solve

$$(35) \quad \hat{V}_t(k, \mathcal{W}) = \text{Max } u(c, n) + \beta \hat{V}_{t+1}(k', \mathcal{W}')$$

subject to (33) and (34). Let $\hat{h}_t(s)$ denote the solution to (35).

A Markov equilibrium with partial commitment, a nonconfiscatory equilibrium, consists of value functions $\hat{V}_t(s)$ and policy functions $\hat{h}_t(s)$, which solve (35) for all s and t .

Suppose now that $\hat{V}_{t+1}(s') = V_{t+1}(s')$, that is, the government in period t believes that the governments from period $t+1$ onward will follow the Ramsey plan. Then, since (32) coincides with (35), the government in period t will find it optimal to choose the Ramsey plan as well.

We have proved the following proposition.

Proposition 5: (Partial commitment is full commitment) The Ramsey equilibrium with wealth restrictions is a Markov equilibrium with partial commitment.

In our view, an attractive feature of these results is that even if governments in the previous periods have, for whatever reason, not pursued Ramsey policies, current governments will follow Ramsey policies as long as they believe future governments will do so as well.

An alternative environment with partial commitment can also be used to establish equivalence between Ramsey and Markov equilibria. This alternative environment builds on Kydland and Prescott's (1980) method for computing Ramsey outcomes. They show that a Ramsey equilibrium could be characterized recursively starting in period one, with the addition of a state variable. This state variable represents promised marginal utilities or, alternatively, returns. This formulation can also be used to establish equivalence. This alternative formulation is more convenient for establishing equivalence between Ramsey and

Markov equilibria in environments with heterogeneous agents.⁸

Partial commitment to instruments Some of the Ramsey literature, starting with Lucas and Stokey (1983) and Chari and Kehoe (1999), restricts period zero policies, as opposed to our restrictions here on period zero wealth in utility terms. In environments with partial commitment, one interpretation of this formulation is that governments can commit to policies one period ahead. We argue that partial commitment to policies alone typically cannot implement the Ramsey outcomes. Specifically, consider an alternative form of partial commitment in which the government in period t chooses a subset of policies, $\{\tau_{t+1}^c, \tau_{t+1}^k, \tau_{t+1}^d\}$, that will be implemented in period $t+1$. The government in any period t is free to choose the labor income tax, τ_t^n . In Chari, Nicolini, and Teles (2018), we formally define Markov equilibria for this environment. Here we simply note that an extensive literature has shown that Markov and Ramsey equilibria typically do not coincide for environments like the ones considered here.

Chari and Kehoe (1993) characterize Markov equilibria in the Lucas and Stokey (1983) environment and show that Markov and Ramsey equilibria do not coincide. Klein, Krusell, and Ríos-Rull (2008) characterize Markov equilibria in environments similar to ours, with partial commitment to instruments, and show that these equilibria do not coincide with Ramsey equilibria. The results in these papers imply that Markov outcomes are in general different from commitment outcomes. Together with our results on partial commitment on wealth, this result shows that the nature of partial commitment plays a crucial role in determining whether Markov equilibria coincide with commitment equilibria.

3. Heterogeneous agents model

Here we briefly discuss extending our results to heterogeneous agents models as in Judd (1985). The analysis here is closely related to that in Werning (2007). This extension allows us to consider redistributive motives for taxation as well as the need to raise funds to finance public goods. We only consider standard macro preferences and wealth constraints. If the parameters of preferences are the same across agents, our result that it is optimal to

⁸Details are available upon request.

not tax capital continues to hold. If they are not the same and taxes cannot be type-specific, it may be optimal to introduce intertemporal distortions. If taxes can be type-specific, zero capital taxation is optimal even if preferences differ across agents.

We consider an economy with an equal measure of two types of agents, 1 and 2. The social welfare function is

$$\theta U^1 + (1 - \theta) U^2$$

with weight $\theta \in [0, 1]$. The individual preferences are assumed to be the standard preferences allowing for possibly different elasticities for the two types of agents,

$$U = \sum_{t=0}^{\infty} \beta^t \left[\frac{(c_t^i)^{1-\sigma^i} - 1}{1 - \sigma^i} - \eta^i (n_t)^{\psi^i} \right].$$

The resource constraints are

$$c_t^1 + c_t^2 + g_t + k_{t+1} - (1 - \delta) k_t \leq A_t F(n_t^1 + n_t^2, k_t),$$

where $k_t = k_t^1 + k_t^2$.

The taxes are the ones in the rich tax system considered in the representative agent economy that includes taxes on consumption τ_t^c , labor income τ_t^n , capital income τ_t^k , dividends τ_t^d , and a tax on initial wealth, l_0 . Note that we do not allow for the taxes to differ across agents. We assume for now that agents cannot be taxed differently based on type.

The implementability conditions can be written as

$$(36) \quad \sum_{t=0}^{\infty} \beta^t (u_{c,t}^i c_t^i + u_{n,t}^i n_t^i) = \bar{W}^i,$$

for $i = 1, 2$, where \bar{W}^i denotes the wealth restriction of agent i .⁹

Since the taxes must be the same for the two agents, an implementable allocation must

⁹We assume that the feasible set is non-empty. This assumption is satisfied only for some pairs of exogenously specified wealth.

also satisfy the following marginal conditions:

$$\frac{u_{c,t}^1}{u_{c,t}^2} = \frac{u_{n,t}^1}{u_{n,t}^2} \text{ and } \frac{u_{c,t}^1}{u_{c,t}^2} = \frac{u_{c,t+1}^1}{u_{c,t+1}^2}.$$

These conditions can be written as

$$(37) \quad u_{c,t}^1 = \gamma u_{c,t}^2 \text{ and } u_{n,t}^1 = \gamma u_{n,t}^2,$$

where γ is some endogenous number.¹⁰

Let φ^1 and φ^2 be the multipliers of the two implementability conditions, (36) for $i = 1, 2$, and let λ_t be the multiplier of the resource constraint. The first-order conditions for $t \geq 0$ imply¹¹

$$(38) \quad u_{c,t}^2 \frac{\gamma [\theta + \varphi^1 (1 - \sigma^1)] \frac{\sigma^2}{c_t^2} + [(1 - \theta) + \varphi^2 (1 - \sigma^2)] \frac{\sigma^1}{c_t^1}}{\frac{\sigma^2}{c_t^2} + \frac{\sigma^1}{c_t^1}} = \lambda_t \text{ and}$$

$$(39) \quad u_{n,t}^2 \frac{\gamma [\theta + \varphi^1 (1 + \psi^1)] \frac{\psi^2}{n_t^2} + [(1 - \theta) + \varphi^2 (1 + \psi^2)] \frac{\psi^1}{n_t^1}}{\frac{\psi^2}{n_t^2} + \frac{\psi^1}{n_t^1}} = -\lambda_t F_{n,t},$$

which together with

$$-\lambda_t + \beta \lambda_{t+1} (f_{k,t+1} + 1 - \delta) = 0$$

imply that, if elasticities are equal, $\sigma^1 = \sigma^2 = \sigma$ and $\psi^1 = \psi^2 = \psi$, capital should never be taxed. To see this, notice that, from (37), c_t^1 must be proportionate to c_t^2 , $c_t^1 = \gamma^{-\frac{1}{\sigma}} c_t^2$, and n_t^1 must also be proportionate to n_t^2 , $n_t^1 = (\gamma)^{\frac{1}{\psi}} n_t^2$. It then follows that the terms multiplying the marginal utilities on the left-hand side of (38) and (39) are time invariant.

Note that if the elasticities would differ across the different agents, the allocations would not be proportionate and the result would not hold. In this case it would be optimal to impose intertemporal distortions. The reason behind this result will become apparent in

¹⁰See also Greulich, Laczó, and Marcat (2016). Werning (2007) also computes optimal taxes with heterogeneous agents taking advantage of this proportionality of marginal utilities.

¹¹See Appendix D for the derivation.

the next section on the relation of our results to the optimality of production efficiency in Diamond and Mirrlees (1971).

We summarize the discussion in the following proposition.

Proposition 6: (Intertemporal distortions in heterogeneous agent economies) Suppose that preferences for all types of agents are in the class of standard macroeconomic preferences. If all agents have the same preferences, then the Ramsey equilibrium has no intertemporal distortions for all $t \geq 0$. If the preferences are different in that $\sigma^1 \neq \sigma^2$ or $\psi^1 \neq \psi^2$, then the Ramsey equilibrium has intertemporal distortions.

This proposition shows that with standard and identical preferences, allowing for redistributive concerns does not overturn the result that, with a rich tax system, capital should not be taxed. If the preferences of agents are different, then intertemporal distortions are typically optimal. If we allow for type specific tax rates, then we drop (37) as a constraint on the Ramsey problem and it is straightforward to show that capital should not be taxed. In the section below, we relate all these results to the results in the literature on uniform taxation and production efficiency.

4. Uniform taxation and production efficiency

In this section, we connect our results to the results on production efficiency and uniform taxation. To develop these connections, we set up an alternative economy, which we call an *intermediate goods economy*, that seems different at face value but turns out to be equivalent to the one considered above. In this alternative economy, the representative household consumes a single final good denoted by C and supplies a single final labor input denoted by N . Preferences for the households are given by

$$(40) \quad U(C, N) = \frac{C^{1-\sigma} - \frac{1}{1-\beta}}{1-\sigma} - \eta N^\psi.$$

The economy has three types of technologies. The first one is given by the resource constraint (2). We refer to the representative firm that operates this technology as the *capital accumulation firm*. This firm produces intermediate goods c_t , hires intermediate labor inputs n_t , and accumulates capital according to the technology (2). The second technology, operated by the *consumption firm*, produces the final good C using the intermediate goods c_t according to

the constant returns to scale technology given by

$$(41) \quad C = \mathcal{C}(c_0, c_1, \dots) = \left[\sum_{t=0}^{\infty} \beta^t c_t^{1-\sigma} \right]^{\frac{1}{1-\sigma}}.$$

The third type of technology, operated by the *labor firm*, produces the intermediate labor inputs using final labor according to the constant returns to scale technology given by

$$(42) \quad N = \mathcal{N}(n_0, n_1, \dots) = \left[\sum_{t=0}^{\infty} \beta^t n_t^\psi \right]^{\frac{1}{\psi}}.$$

In terms of the tax system, we assume that the government can levy a tax on the final consumption good, τ^c , and on final labor denoted by τ^n . In addition, we assume that the government can levy taxes on all intermediate goods, denoted by τ_t^c and τ_t^n . We retain the dividend taxes and the capital income taxes levied on the capital accumulation firm, as well as the initial levy l_0 . We do not impose any taxes on the profits of either the consumption firm or the labor firm because these profits will be zero in equilibrium.

The households' problem is to maximize (40) subject to the budget constraint

$$(43) \quad p(1 + \tau^C)C - w(1 - \tau^N)N \leq (1 - l_0) V_0,$$

where p and w denote the prices of final consumption and labor in units of the consumption good in period zero, and V_0 is the value of initial wealth in units of goods.

The consumption firm's problem is to maximize

$$(44) \quad pC - \sum_{t=0}^{\infty} q_t(1 + \tau_t^c)c_t$$

subject to (41) and the labor firm's problem is to maximize

$$(45) \quad \sum_{t=0}^{\infty} q_t(1 - \tau_t^n)w_t n_t - wN$$

subject to (42).

The capital accumulation firm's problem is the same as in our benchmark model. A competitive equilibrium is defined in the standard fashion.

Next, we show that the competitive equilibria in the two economies coincide in the intermediate goods economy and in the growth economy with distorting taxes. To do so, we first show that we can rewrite the equilibrium in the intermediate goods economy as an equilibrium in the growth economy with taxes by incorporating the decisions of the consumption firm and the labor firm directly into the households' problem.

Consider the budget constraint of the households. Given that in any competitive equilibrium, profits are zero for the consumption and the labor firm, we can substitute $pC = \sum_{t=0}^{\infty} q_t(1 + \tau_t^c)c_t$ from (44) and $wN = \sum_{t=0}^{\infty} q_t(1 - \tau_t^n)w_t n_t$ from (45) into (43) to obtain a budget constraint of the form (11) in the growth economy with taxes. The only difference is the presence of the tax on final consumption and final labor, which amounts to rescaling the consumption and labor income taxes in the original economy.

Substituting from (42) and (41) into (40), we see that the households' utility function in the rewritten intermediate goods economy is the same as in the growth economy with taxes.

To establish that the converse holds, note that we can set up the households' problem in the growth economy with taxes as a two-stage problem of first choosing an aggregate value for consumption and labor and then choosing the disaggregated levels of consumption and labor to achieve the desired value for final consumption and labor. Thus, the equilibria in the two economies coincide.

In the intermediate goods economy, the Ramsey problem is to maximize (40) subject to the implementability constraint

$$(46) \quad U_C C + U_N N = \bar{W},$$

where \bar{W} denotes the exogenously specified bound on the value of wealth the planner must deliver and is given by

$$\bar{W} = \frac{U_C}{1 + \tau^C} (1 - l_0) V_0$$

and the requirement that the allocation is in the *production set* given by (41) and (42) with inequalities and (2).

Next, when we apply the same logic as in Diamond and Mirrlees (1971) to our intermediate goods economy, it is immediate that the solution to the Ramsey problem must satisfy production efficiency. An implication of this result is that it is possible to implement the Ramsey equilibrium with no taxation of intermediate goods.

Thus, setting these taxes to zero, we can combine the first-order conditions for the capital accumulation firm and the consumption firm to obtain

$$(47) \quad \frac{\mathcal{C}_{ct}}{\mathcal{C}_{ct+1}} = 1 - \delta + F_{k,t+1}.$$

Similarly, combining the first-order conditions for the capital accumulation firm and the labor firm, we obtain

$$(48) \quad \frac{\mathcal{N}_{nt}}{\mathcal{N}_{nt+1}} = \frac{F_{n,t}}{F_{n,t+1}} (1 - \delta + F_{k,t+1}).$$

Condition (47) equates the rates at which c_t is transformed into c_{t+1} through the composite C to the rate at which c_t is transformed into c_{t+1} through capital, k_{t+1} . Condition (48) is the analog for labor in consecutive periods. Note that

$$(49) \quad \frac{\mathcal{C}_{ct}}{\mathcal{C}_{ct+1}} = \frac{c_t^{-\sigma}}{\beta c_{t+1}^{-\sigma}} = \frac{u_{c,t}}{\beta u_{c,t+1}}.$$

A similar expression holds for the intermediate labor inputs.

The intermediate goods economy and our benchmark economy are obviously equivalent. This observation, together with (47) and (49), along with the analogous conditions for labor, imply that it is optimal to not tax capital in the benchmark economy. We summarize this discussion in the following proposition, which is the equivalent to Proposition 3 in the intermediate goods economy.

Proposition 3’: (No intertemporal distortions ever) Suppose that preferences and technologies are given by (40), (41), and (42) and the wealth restriction must be satisfied. Then, the Ramsey solution has no intertemporal distortions for all $t \geq 0$.

Consider next the case in which rather than imposing a wealth constraint, initial policies are restricted so that rents cannot be fully taxed away. Here, because rents are not fully taxed away, production efficiency is no longer optimal. Nevertheless, it is straightforward to show that an analog of Proposition 5 holds in the intermediate goods economy in the sense that it is not optimal to distort intermediate goods decisions from period one onward. It is optimal to distort intermediate goods in period zero relative to all other intermediate goods in order to partially tax rents.

The intermediate goods economy helps to clarify circumstances in which it is optimal to not have intertemporal distortions. We have shown that it is optimal to not have such distortions when the underlying economy can be represented as a constant returns to scale economy in the production of intermediate goods. In this sense, we have shown an equivalence between the principle that intermediate goods taxation is undesirable and intertemporal distortions should not be introduced.

In the process of doing so, we have shown that the celebrated result of Atkinson and Stiglitz (1972), that uniform commodity taxation is optimal when preferences are homothetic and separable, follows from the production efficiency result of Diamond and Mirrlees (1971). Thus, our result of no intertemporal distortions is very closely connected to the uniform commodity taxation result.

A. Production efficiency in a heterogeneous agents economy

Consider now developing an intermediate goods economy that is equivalent to our heterogeneous agent economy. Here we think of the intermediate goods economy as producing two distinct types of final consumption goods denoted by C^i for $i = 1, 2$. For simplicity, we assume the economy utilizes one common type of final labor, denoted by N . The preferences for households of type i are given by

$$(50) \quad U(C^i, N^i) = \frac{(C^i)^{1-\sigma^i} - \frac{1}{1-\beta}}{1-\sigma^i} - \eta (N^i)^\psi,$$

where N^i denotes the amount of the common final labor supplied by type i . The technologies for the capital accumulation and the labor firm are the same as before. Each consumption

good is produced by its own constant returns to scale technology given by

$$(51) \quad C^i = \mathcal{C}^i(c_0^i, c_1^i \dots) = \left[\sum_{t=0}^{\infty} \beta^t (c_t^i)^{1-\sigma^i} \right]^{\frac{1}{1-\sigma^i}}$$

for $i = 1, 2$.

The tax system is the same except that now we require that the tax rate on final consumption goods must be the same for the two types.

Households of type i maximize (50) subject to the budget constraint

$$(52) \quad p^i(1 + \tau^C)C^i - w(1 - \tau^N)N^i \leq (1 - l_0) V_0^i,$$

where p^i denotes the price of the final consumption good of type i in units of the consumption good in period zero.

The consumption firm of type i maximizes

$$(53) \quad p^i C^i - \sum_{t=0}^{\infty} q_t(1 + \tau_t^c)c_t^i$$

subject to (51).

The other firms solve the same problems as in the representative agent economy. A competitive equilibrium is defined in the standard fashion.

Next, we show that the competitive equilibria in the intermediate goods economy and the heterogeneous agents economy with type-independent taxes coincide. Recall that in the representative agent model, we used zero profits for consumption goods producers and replaced the pretax value of consumption in the households' budget constraint in the intermediate goods economy with the value of the intermediate goods. Following the same procedure, we can use (53) and write (52) as

$$(1 + \tau^C) \sum_{t=0}^{\infty} q_t(1 + \tau_t^c)c_t^i - w(1 - \tau^N)N^i \leq (1 - l_0) V_0^i.$$

Notice that this budget constraint coincides with the budget constraints in the heterogeneous

agents economy except for rescaling. The rest of the argument that the equilibria coincide is the same as in the representative agent economy.

Consider now the Ramsey problem in the intermediate goods economy. Since the tax rates on both consumption goods must be the same, the Ramsey problem has an additional constraint that can be written as

$$(54) \quad \frac{U_{C^i}^i C_{c_t}^i}{-U_{N^i}^i N_{n_t}^i} = \frac{U_{C^j}^j C_{c_t}^j}{-U_{N^j}^j N_{n_t}^j}, t \geq 0.$$

The Ramsey problem is now to maximize utility subject to the implementability constraint, the requirement that the allocation is in the production set and (54). The Ramsey problem in the intermediate goods economy with type-dependent taxes simply drops (54).

We now have the analog of Proposition 6.

Proposition 6’: (No intertemporal distortions in heterogenous agents economies) Suppose that preferences for all types of agents are in the class of standard macroeconomic preferences. If all agents have the same preferences, then the Ramsey equilibrium has no intertemporal distortions for all $t \geq 0$.

Remark: If the preferences of the two agents are different, in the sense that σ^1 is different from σ^2 , the result in Proposition 6’ does not hold. This result does hold if the two types of final consumption goods could be taxed at different rates. This result is consistent with the result on the optimality of production efficiency in Diamond and Mirrlees (1971) that requires that all final goods may be taxed at possibly different rates.

5. Concluding remarks

We have extended the production efficiency theorem to environments in which taxation of pure rents is limited but the Ramsey planner faces a wealth constraint. Our notion of zero taxation of capital is stronger than production efficiency but follows from it for the kinds of preferences that are standard in the macroeconomics literature. Using these two results, we have shown that in standard macroeconomic models, production efficiency requires that capital taxation is inefficient if the planner has access to a rich tax system.

We have also argued that setting up Ramsey problems with wealth constraints has the desirable feature that, with partial commitment, the Ramsey outcomes are time consistent.

References

- [1] Abel, Andrew B., 2007, “Optimal Capital Income Taxation,” Working Paper 13354, National Bureau of Economic Research.
- [2] Armenter, Roc, 2008, “A Note on Incomplete Factor Taxation,” *Journal of Public Economics* 92 (10-11), 2275–2281.
- [3] Atkeson, Andrew, V. V. Chari, and Patrick J. Kehoe, 1999, “Taxing Capital Income: A Bad Idea,” *Federal Reserve Bank of Minneapolis Quarterly Review* 23 (3), 3–18.
- [4] Atkinson, Anthony B. and Joseph E. Stiglitz, 1972, “The Structure of Indirect Taxation and Economic Efficiency,” *Journal of Public Economics* 1 (1), 97–119.
- [5] Bassetto, Marco, and Jess Benhabib, 2006, “Redistribution, Taxes, and the Median Voter,” *Review of Economic Dynamics* 9 (2), 211–223.
- [6] Chamley, Christophe, 1986, “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives,” *Econometrica* 54 (3), 607–622.
- [7] Chari, V.V. and Patrick J. Kehoe, 1993, “Sustainable Plans and Debt,” *Journal of Economic Theory* 61 (2), 230–261.
- [8] Chari, V. V., and Patrick J. Kehoe, 1999, “Optimal Fiscal and Monetary Policy,” *Handbook of Macroeconomics*, vol. 1C, edited by John Taylor and Michael Woodford, North-Holland.
- [9] 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), 617–652.
- [10] Chari, V.V., Juan Pablo Nicolini and Pedro Teles, 2018, “Optimal Capital Taxation Revisited,” Working Paper 752, Federal Reserve Bank of Minneapolis.
- [11] Coleman, Wilbur J. II, 2000, “Welfare and Optimum Dynamic Taxation of Consumption and Income,” *Journal of Public Economics* 76 (1), 1–39.
- [12] Diamond, Peter A. and James A. Mirrlees, 1971, “Optimal Taxation and Public Production I: Production Efficiency,” *American Economic Review* 61 (1), 8–27.

- [13] Greulichy, Katharina, Sarolta Laczó, and Albert Marcet, 2016, “Pareto-Improving Optimal Capital and Labor Taxes,” mimeo, Barcelona GSE.
- [14] Judd, Kenneth L., 1985, “Redistributive Taxation in a Simple Perfect Foresight Model,” *Journal of Public Economics* 28 (1), 59–83.
- [15] Klein, Paul, Per Krusell, and Jose-Víctor Ríos-Rull, 2008, “Time-Consistent Public Policy,” *Review of Economic Studies* 75, 789–808.
- [16] Kydland, Finn E. and Edward C. Prescott, 1980, “Dynamic Optimal Taxation, Rational Expectations and Optimal Control,” *Journal of Economic Dynamics and Control* 2, 79–91.
- [17] Lucas, Robert E., Jr. and Nancy L. Stokey, 1983, “Optimal Fiscal and Monetary Policy in an Economy without Capital,” *Journal of Monetary Economics* 12 (1), 55–93.
- [18] Pratt, John W., 1964, “Risk aversion in the small and in the large,” *Econometrica*. 32 (1–2), 122–136.
- [19] Ramsey, Frank P., 1927, “A Contribution to the Theory of Taxation,” *Economic Journal* 37 (145), 47–61.
- [20] Stokey, Nancy L., Robert E. Lucas, Jr., with Edward C. Prescott, 1989, *Recursive Methods in Economic Dynamics*, Harvard University Press.
- [21] Straub, Ludwig and Iván Werning, 2015, “Positive Long Run Capital Taxation: Chamley-Judd Revisited,” mimeo, MIT.
- [22] Werning, Iván, 2007, “Optimal Fiscal Policy with Redistribution,” *Quarterly Journal of Economics* 122 (3), 925–967.
- [23] Zhu, Xiaodong, 1992, “Optimal Fiscal Policy in a Stochastic Growth Model,” *Journal of Economic Theory* 58 (2), 250–289.

6. Appendix

In this appendix, we describe some results not included in the main text

A. An alternative decentralization

Here we analyze an economy that is equivalent to the one we study but that is closer to the standard setup used in the literature. This economy has households accumulating capital and renting it out to firms.

The household owns the capital stock and rents it to a representative firm every period at rate u_t . The capital income tax τ_t^k is paid by the representative firm, and the dividend tax, τ_t^d , is a tax on capital income net of gross investment paid by the household. This dividend tax resembles the tax proposed by Abel (2007) as a way of collecting lump-sum revenue from the taxation of the initial capital stock.

The flow of funds for the households for $t \geq 1$ can then be written as

$$\begin{aligned} \frac{1}{1+r_t} b_{t+1} + k_{t+1} &= b_t + (1-\delta)k_t + u_t k_t - \tau_t^d [u_t k_t - (k_{t+1} - (1-\delta)k_t)] \\ &\quad + (1-\tau_t^n) w_t n_t - (1+\tau_t^c) c_t. \end{aligned}$$

In the initial period, the constraint is

$$\begin{aligned} \frac{1}{1+r_0} b_1 + (1-\tau_0^d) k_1 &= (1-l_0) [b_0 + (1-\tau_0^d)(1-\delta)k_0 + (1-\tau_0^d)u_0 k_0] \\ &\quad + (1-\tau_0^n) w_0 n_0 - (1+\tau_0^c) c_0. \end{aligned}$$

The marginal conditions are (6), (7), and

$$(55) \quad 1+r_t = \frac{(1-\tau_{t+1}^d)[u_{t+1} + 1-\delta]}{1-\tau_t^d}.$$

The representative firm maximizes profits

$$\Pi_t = F(k_t, n_t) - w_t n_t - u_t k_t - \tau_t^k (u_t - \delta) k_t.$$

The capital income tax τ_t^k is a tax on sales minus wages and depreciation of capital, which is the model counterpart to a profit tax.

The price of the good must equal the marginal cost,

$$(56) \quad 1 = \frac{w_t}{F_n(t)} = \frac{u_t + \tau_t^k (u_t - \delta)}{F_k(t)}.$$

These marginal conditions can be written as (15), (16), and (17).

The households' budget constraint can be written as

$$\begin{aligned} & \sum_{t=0}^{\infty} q_t [(1 + \tau_t^c) c_t - (1 - \tau_t^n) w_t n_t] = \\ & (1 - l_0) [b_0 + (1 - \tau_0^d) (1 - \delta) k_0 + (1 - \tau_0^d) u_0 k_0], \end{aligned}$$

where $q_t = \frac{1}{(1+r_0)\dots(1+r_{t-1})}$ for $t \geq 1$, with $q_0 = 1$. This uses the no-Ponzi-scheme condition $\lim_{T \rightarrow \infty} q_{T+1} b_{T+1} \geq 0$.

The marginal conditions of the households and firms can be used to write the budget constraint as an implementability condition, which will be written as

$$\sum_{t=0}^{\infty} \beta^t [u_c(t) c_t + u_n(t) n_t] = \frac{(1 - l_0)}{1 + \tau_0^c} \left[b_0 + (1 - \tau_0^d) k_0 + \frac{(1 - \tau_0^d) (F_k(0) - \delta)}{1 + \tau_0^k} k_0 \right].$$

The initial confiscation is restricted regardless of the taxes that are used to obtain it. We define \mathcal{W}_0 to be the exogenous level of initial wealth that the households can keep, measured in units of utility at time 0,

$$(57) \quad u_c(0) \frac{(1 - l_0)}{1 + \tau_0^c} \left[b_0 + (1 - \tau_0^d) k_0 + \frac{(1 - \tau_0^d) (F_k(0) - \delta)}{1 + \tau_0^k} k_0 \right] = \mathcal{W}_0.$$

The implementability condition can then be written as

$$(58) \quad \sum_{t=0}^{\infty} \beta^t [u_c(t) c_t + u_n(t) n_t] = \mathcal{W}_0.$$

The implementability condition (58) and the resource constraints (2) are the only equilibrium restrictions on the sequences of consumption, labor, and capital. The taxes have the natural restriction that the tax revenue does not exceed the base, so that $\tau_t^k \leq 1$ for all t . This restriction will not be binding in this setup.

The other equilibrium conditions, other than (58) and (2), are satisfied by other variables. The variable τ_t^n is determined by

$$-\frac{u_c(t)}{u_n(t)} = \frac{(1 + \tau_t^c)}{(1 - \tau_t^n) w_t};$$

the variable r_t is determined by

$$\frac{u_c(t)}{1 + \tau_t^c} = (1 + r_t) \frac{\beta u_c(t+1)}{1 + \tau_{t+1}^c};$$

the variable w_t is determined by

$$1 = \frac{w_t}{F_n(t)};$$

the variable u_t is determined by

$$\frac{w_t}{F_n(t)} = \frac{u_t + \tau_t^k (u_t - \delta)}{F_k(t)};$$

and either τ_{t+1}^c given τ_t^c or τ_{t+1}^d given τ_t^d is determined by

$$u_c(t) = \beta u_c(t+1) \frac{(1 + \tau_t^c) (1 - \tau_{t+1}^d)}{(1 + \tau_{t+1}^c) (1 - \tau_t^d)} [u_{t+1} + (1 - \delta)].$$

The constraint (57) will be satisfied with l_0 .

This implementation does not use τ_t^k and τ_t^d for all t as well as τ_0^c or else τ_t^k and τ_t^c for all t and τ_0^d . They are redundant instruments. It follows that the restriction that $\tau_t^k \leq 1$

will not bind. This also means that the capital tax can always be set equal to zero.

B. Extension to uncertainty

Consider an economy where government expenditures and total factor productivity are stochastic. The exogenous state of the economy is given by s_t and takes on a finite set of values. The history of the exogenous state in period t is $s^t = (s_0, s_1, \dots, s_t)$. The probability of a history s^t is $\pi(s^t)$. Allocations and prices are now functions of the history. For example consumption, labor, and capital allocations are given by $c_t(s^t)$, $n_t(s^t)$, $k_{t+1}(s^t)$.

The preferences of a representative household are given by

$$(59) \quad U = \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi(s^t) u(c_t(s^t), n_t(s^t)),$$

satisfying the usual properties.

The production technology is described as before, by

$$(60) \quad c_t(s^t) + g_t(s^t) + k_{t+1}(s^t) - (1 - \delta) k_t(s^{t-1}) \leq A_t(s^t) F(n_t(s^t), k_t(s^{t-1})).$$

The implementability constraint is given by

$$(61) \quad E_0 \sum_{t=0}^{\infty} \beta^t [u_{c,t}(s^t) c_t(s^t) + u_{n,t}(s^t) n_t(s^t)] = \mathcal{W}_0.$$

We impose the following restriction on the Ramsey problem:

$$\mathcal{W}_0 \geq \bar{\mathcal{W}}.$$

The Ramsey problem is to maximize (59) subject to (60) and (61). The first-order conditions are

$$(62) \quad -\frac{u_{c,t}}{u_{n,t}} = \frac{1 + \varphi(1 + \sigma_t^n - \sigma_t^{nc})}{1 + \varphi(1 - \sigma_t - \sigma_t^{cn})} \frac{1}{F_{nt}}, \text{ for all states, } t \geq 0$$

$$(63) \quad u_{c,t} = \beta E_t \left[u_{c,t+1} \frac{1 + \varphi (1 - \sigma_{t+1} - \sigma_{t+1}^{cn})}{1 + \varphi (1 - \sigma_t - \sigma_t^{cn})} (1 + F_{k,t+1} - \delta) \right], \text{ for all states, } t \geq 0$$

where we have suppressed the dependence on the history for notational convenience and E_t denotes the conditional expectation in history s^t . Consider now preferences that are standard in the macroeconomics literature. In this case, the cross elasticities are zero and the own elasticities are constant, so (63) can be written as

$$u_{c,t} = \beta E_t [u_{c,t+1} (1 + F_{k,t+1} - \delta)], t \geq 0.$$

It follows that the Ramsey solution has no intertemporal distortions for all $t \geq 0$.

C. Time-separable homothetic preferences

Here we show that standard macroeconomic preferences, (24), are the only time separable preferences that are separable between consumption and labor and homothetic in consumption and labor.

Time-separable preferences that are also separable between consumption and labor must satisfy

$$\frac{\beta U_c(c_{t+1})}{U_c(c_t)} = \frac{\beta U_c(\lambda c_{t+1})}{U_c(\lambda c_t)}.$$

for all $\lambda > 0$. Differentiating with respect to λ , we have

$$U_c(c_{t+1})c_t U_{cc}(\lambda c_t) = U_{cc}(\lambda c_{t+1})c_{t+1} U_c(c_t).$$

Set $\lambda = 1$. It follows that

$$\frac{c_t U_{cc}(c_t)}{U_c(c_t)} = k$$

is independent of c_t . A standard result (see Pratt, 1964) is that $U = \frac{c^{1-\sigma}}{1-\sigma}$:

$$x f'(x) = k f(x).$$

Let $g(x) = \log f(x)$. Then

$$g'(x) = \frac{f'(x)}{f(x)}.$$

Rewrite the differential equation as

$$g'(x) = \frac{k}{x}$$

so that

$$g(x) = k \log x + C.$$

It follows that

$$\log f(x) = k \log x + C,$$

and therefore

$$f(x) = Cx^k.$$

D. Heterogeneous agents economy

Here we describe the solution of the Ramsey problem in the heterogeneous agents economy in detail.

The marginal conditions of the Ramsey problem are

$$\theta u_{c,t}^1 + \varphi^1 u_{c,t}^1 (1 - \sigma) + \mu_t^c u_{cc,t}^1 = \lambda_t, t \geq 0$$

$$(1 - \theta) u_{c,t}^2 + \varphi^2 u_{c,t}^2 (1 - \sigma) - \mu_t^c \gamma u_{cc,t}^2 = \lambda_t, t \geq 0$$

$$\theta u_{n,t}^1 + \varphi^1 u_{n,t}^1 (1 + \psi) + \mu_t^n u_{nn,t}^1 = -\lambda_t F_{n,t}, t \geq 0$$

$$(1 - \theta) u_{n,t}^2 + \varphi^2 u_{n,t}^2 (1 + \psi) - \mu_t^n u_{nn,t}^2 = -\lambda_t F_{n,t}, t \geq 0,$$

where φ^1 and φ^2 are the multipliers of the two implementability conditions, (36), and μ_t^c and

μ_t^n are the multipliers of conditions (37).

The marginal conditions for c_t^1 and c_t^2 can be used to write

$$\gamma u_{cc,t}^2 [\theta u_{c,t}^1 + \varphi^1 u_{c,t}^1 (1 - \sigma)] + u_{cc,t}^1 [(1 - \theta) u_{c,t}^2 + \varphi^2 u_{c,t}^2 (1 - \sigma)] = \lambda_t \gamma u_{cc,t}^2 + \lambda_t u_{cc,t}^1, t \geq 0$$

which can be rewritten as

$$\gamma \sigma \frac{u_{c,t}^2}{c_t^2} [\theta u_{c,t}^1 + \varphi^1 u_{c,t}^1 (1 - \sigma)] + \sigma \frac{u_{c,t}^1}{c_t^1} [(1 - \theta) u_{c,t}^2 + \varphi^2 u_{c,t}^2 (1 - \sigma)] = \lambda_t \gamma \sigma \frac{u_{c,t}^2}{c_t^2} + \lambda_t \sigma \frac{u_{c,t}^1}{c_t^1}, t \geq 0$$

Using (37), it follows that

$$u_{c,t}^2 \frac{\gamma [\theta + \varphi^1 (1 - \sigma)] \frac{\sigma}{c_t^2} + [(1 - \theta) + \varphi^2 (1 - \sigma)] \frac{\sigma}{c_t^1}}{\frac{\sigma}{c_t^2} + \frac{\sigma}{c_t^1}} = \lambda_t, t \geq 0,$$

which is condition (38) in the text.

With standard preferences, (37) implies

$$c_t^1 = \gamma^{-\frac{1}{\sigma}} c_t^2,$$

which in turn implies

$$u_{c,t}^2 \frac{\gamma [\theta + \varphi^1 (1 - \sigma)] \sigma \gamma^{-\frac{1}{\sigma}} + [(1 - \theta) + \varphi^2 (1 - \sigma)] \sigma}{\sigma \gamma^{-\frac{1}{\sigma}} + \sigma} = \lambda_t, t \geq 0.$$

For labor, using the marginal conditions for n_t^1 and n_t^2 ,

$$u_{n,t}^2 \frac{\gamma [\theta + \varphi^1 (1 + \psi)] \frac{\psi}{n_t^2} + [(1 - \theta) + \varphi^2 (1 + \psi)] \frac{\psi}{n_t^1}}{\frac{\psi}{n_t^2} + \frac{\psi}{n_t^1}} = -\lambda_t F_{n,t}, t \geq 0.$$

Similarly, using (37), we have

$$n_t^1 = (\gamma)^{\frac{1}{\psi}} n_t^2,$$

and therefore

$$u_{n,t}^2 \frac{\gamma [\theta + \varphi^1 (1 + \psi)] \psi (\gamma)^{\frac{1}{\psi}} + [(1 - \theta) + \varphi^2 (1 + \psi)] \psi}{\psi (\gamma)^{\frac{1}{\psi}} + \psi} = -\lambda_t F_{n,t}, t \geq 0,$$

which is condition (39) in the text.