Optimal Capital Taxation Revisited[*]

V. V. Chari

University of Minnesota and Federal Reserve Bank of Minneapolis

Juan Pablo Nicolini

Federal Reserve Bank of Minneapolis

Universidad Di Tella, Universidad Autonoma de Barcelona

Pedro Teles

Banco de Portugal, Catolica Lisbon SBE, CEPR

October 2016

Abstract

We revisit the question of how capital should be taxed. We argue that if we allow governments to use the kinds of tax instruments widely used in practice, for preferences that are standard in the macroeconomic literature, it is optimal never to distort capital accumulation. The results also hold with heterogenous agents and in environments with partial commitment. We show that the results in the literature that lead to the presumption that capital ought to be taxed for some length of time arise because the tax system is restricted.

Keywords: capital income tax; long run; uniform taxation

JEL Codes: E60; E61; E62

1   Introduction

How should capital be taxed? How should it be taxed in the long run and along the transition? An influential literature on the optimal Ramsey taxation of capital, as in Chamley (1986) and Judd (1985), compares a tax on labor income with a particular tax on capital income that is capped at some level. The common result is that capital should be taxed at its maximum level initially and for a number of periods, and that in the steady state, capital should not be taxed. In more recent work, in Bassetto and Benhabib (2006) and Straub and Werning (2015), it is shown that full taxation of capital can last forever.[1] This literature leads to the presumption that capital taxes should be high for some length of time.

In this paper we take the same Ramsey approach to the optimal taxation of capital, in that the tax system is exogenously given, but enlarge the set of instruments to include other taxes widely used in practice in developed economies such as dividend, consumption and wealth taxes. We refer to a tax system with this enlarged set of tax instruments as a *rich tax system*. As is well known many tax policies yield the same wedges and the theory pins down distortions in choices. Following the public finance literature we refer to these distortions as wedges.

The main question we address in this paper is Does the Ramsey policy yield intertemporal wedges? If it does, we say future capital is taxed. If it does not, we say future capital is not taxed. We begin by studying the standard neoclassical growth model with a representative agent. We show that with a rich tax system, capital should not be taxed in the steady state. Along the transition capital may be taxed or subsidized. We then consider a class of economies that are standard in the macroeconomics literature and show that with these preferences, future capital should never be taxed, except possibly for a one period transition. We then consider heterogeneous agent economies in which agents differ in their initial wealth. We show that the representative agent results also hold in those economies with heterogeneous agents. Our results differ from those in the literature described above because we allow for a rich tax system, while that literature considers a restricted tax system.

A central feature of the literature on optimal taxation is that, absent other restrictions, factors in fixed supply should be taxed away completely. This feature implies

---

[1]Other relevant literature includes Chari, Christiano, Kehoe (1994), Atkeson, Chari and Kehoe (1999), Judd (1999, 2002), Coleman (2000), Lucas and Stokey (1983).

that in the growth model, with a representative agent, the initial capital, as well as holdings of government bonds, should be taxed, possibly at rates in excess of 100% in order to fund government spending. In the Ramsey literature it is conventional to impose restrictions on such taxes on initial wealth. In part of our analysis, we depart from this literature by imposing a restriction on the value of initial wealth, rather than on the taxes themselves (see Armenter (2008) for an analysis with such a restriction). With standard preferences and such a restriction, we show that future capital should never be taxed. We go on to show that with the conventional restriction in the Ramsey literature, capital accumulation in the very first period is distorted and is undistorted thereafter.

We also analyze restricted tax systems like those in the literature described earlier. With taxes on capital and labor only and with capital taxes restricted to be less than 100%, we recover the results in Chamley and Straub and Werning that capital should be taxed for some, possibly infinite, length of time. We argue that the key force underlying these results is that such taxation reduces the value of initial wealth and therefore represents an attempt to confiscate initial wealth indirectly given that the government is not allowed to confiscate that wealth directly.

The conventional view of Ramsey equilibrium in dynamic economies is that the government chooses policies in period zero and commits to these policies thereafter. Our results hold under this conventional view. We go on to show that our results hold under an alternative view. In this alternative view, the government in each period has partial commitment in the sense that it can commit to one period returns on assets in utility terms. That is, the government in the following period is free to choose policies as it desires but must respect the previously committed return constraints.

The idea behind this notion of partial commitment begins by noting that returns on assets must satisfy intertemporal Euler equations on the equilibrium path. This notion of partial commitment requires that returns on assets must satisfy those conditions both on and off the equilibrium path. Requiring governments in any period to satisfy these restrictions allows governments to choose policies as they see fit but ensures that both on an off the equilibrium path the policies do not induce agents to regret their past choices. In particular, governments can deviate from equilibrium choices, but these deviations must have the property that agents do not wish that they could go back to the previous period and change the previous choices.

We show that the Markov equilibrium in this set up coincides with the commitment

equilibrium with restrictions on the value of wealth. Thus partial commitment provides one rationalization for studying Ramsey equilibria with initial wealth restrictions.

We go on to study partial commitment to future taxes, rather than returns on assets. This analysis is motivated by the analysis of Ramsey problems with exogenously specified initial taxes. We argue that Markov equilibria do not coincide with the commitment equilibrium. The reason is that commitment outcomes face a time inconsistency problem as in Lucas and Stokey (1983). In particular with single period debt, the government has a strong incentive to choose policies so as to reduce the value of inherited debt. These incentives imply that Markov equilibria do not coincide with the commitment equilibrium.

We briefly analyze an economy with heterogeneous agents and show that our representative agent results hold in such economies. An interesting feature of heterogeneous agent economies is that even if initial policies are unrestricted, the Ramsey equilibrium could distort intratemporal margins in order to achieve redistributive roles (see Werning, 2007).

Finally we relate our results to the results on uniform commodity taxation (Atkinson and Stiglitz, 1972). Standard preferences used in the macroeconomic literature 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 final consumption and labor aggregates. In this recast economy, we show that Diamond and Mirrlees (1971) production efficiency theorem can be extended to obtain that it is optimal not to 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, future capital should never be taxed.

In this paper we prove some new results, bring together well known results in a unified framework, clarify the relation among these results, and relate the results on dynamic taxation to those on uniform commodity taxation. The new results in this paper are on the equivalence between Markov equilibria and commitment equilibria, the insights on the desire to tax capital heavily in restricted tax systems, and on clarifying the exact relation with the production efficiency literature. The results that with a rich tax system, and with standard macro preferences it is optimal not to tax future capital after a one period transition is in Chari and Kehoe (1999), Zhu (1992), Werning (2007). The heterogeneous agent results are in Werning (2007).

4

2 A representative agent economy

The model is the deterministic neoclassical growth model with taxes. The preferences of a representative household are over consumption c_t and labor n_t,

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

satisfying the usual properties.

The production technology is described by

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

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

We now describe a competitive equilibrium with taxes in which the government finances public consumption and initial debt with (possibly) time varying proportional taxes. We allow for a *rich tax system* that includes taxes on consumption τ_t^c, labour income τ_t^n, capital income τ_t^k, dividends τ_t^d, a tax on initial wealth, l_0 and a nonnegative lump sum transfer T_0, in period zero.

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 trades shares of the firm, and receives dividends. In Appendix 1 we describe an alternative, more widely used decentralization in which the household owns the capital stock and firms rent capital from the household. 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. We now describe the problems of the firm and the household and define a competitive equilibrium.

Firm The representative firm produces and invests in order to maximize the present value of dividends, net of taxes, $\sum_{t=0}^{\infty} q_t (1 - \tau_t^d) d_t$, where q_t is the price of one unit of the good produced in period t in units of the good in period zero and τ_t^d are dividend taxes. Dividends, d_t, are given by

$$d_t = F(k_t, n_t) - w_t n_t - \tau_t^k [F(k_t, n_t) - w_t n_t - \delta k_t] - [k_{t+1} - (1 - \delta)k_t] \tag{3}$$

where w_t is the pre-tax wage rate, τ_t^k is the tax rate on capital income net of depreciation.

Note that 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. (In a steady state of the competitive equilibrium it is possible to show that dividends are always positive).

Let the interest rate between periods t and $t + 1$ be defined by

$$\frac{q_t}{q_{t+1}} \equiv 1 + r_t, \text{ with } q_0 = 1. \tag{4}$$

The first order conditions to the firm's problem are

$$F_{n,t} = w_t \tag{5}$$

$$1 + r_t = \frac{\left(1 - \tau_{t+1}^d\right) \left[1 + \left(1 - \tau_{t+1}^k\right) \left(F_{k,t+1} - \delta\right)\right]}{1 - \tau_t^d} \tag{6}$$

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

Substituting for d_t from (3) and using (4) − (6) it is easy to show that the present discounted value of dividends is given by

$$\sum_{t=0}^{\infty} q_t \left(1 - \tau_t^d\right) d_t = \left(1 - \tau_0^d\right) \left[1 + \left(1 - \tau_0^k\right) \left(F_{k,0} - \delta\right)\right] k_0. \tag{7}$$

Household The flow of funds constraint in period t for the household is given by

$$\frac{1}{1 + r_t} b_{t+1} + p_t s_{t+1} = b_t + p_t s_t + (1 - \tau_t^d) d_t s_t + (1 - \tau_t^n) w_t n_t - (1 + \tau_t^c) c_t \tag{8a}$$

for $t \geq 1$, and for period 0

$$\frac{1}{1+r_0}b_1 + p_0 s_1 \;=\; (1-l_0)\left[b_0 + p_0 s_0 + (1-\tau_0^d)d_0 s_0\right] + \tag{9}$$

$$(1-\tau_0^n)\,w_0 n_0 - (1+\tau_0^c)\,c_0 + T_0. \tag{10}$$

where b_{t+1} denotes holdings of government debt that pay one unit of consumption in period $t+1$, s_{t+1} denotes the household's holdings of the shares of the firm, p_t is the price per unit of firm's shares in units of the good in period t.[2] Note that the price p_t is the price of shares after dividends have been paid in period t. The initial conditions are given by b_0 and $s_0 = 1$.

The household's problem is to maximize utility (1), subject to $(8a),(9)$ and a no-Ponzi scheme condition, $\lim_{T\to\infty} q_{T+1} b_{T+1} \geq 0$.

The first order conditions of the household's problem include

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{(1+\tau_t^c)}{(1-\tau_t^n)\,w_t}, \; t \geq 0. \tag{11}$$

$$\frac{u_{c,t}}{(1+\tau_t^c)} = (1+r_t)\frac{\beta u_{c,t+1}}{(1+\tau_{t+1}^c)}, \; t \geq 0, \tag{12}$$

and

$$1+r_t = \frac{p_{t+1} + (1-\tau_{t+1}^d)\,d_{t+1}}{p_t}, \tag{13}$$

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

The transversality condition implies that the price of the stock equals the present value of future dividends,

$$p_t = \sum_{s=0}^{\infty} \frac{q_{t+1+s}}{q_t}\left(1-\tau_{t+1+s}^d\right)d_{t+1+s}, \tag{14}$$

Using the no-Ponzi scheme condition, the budget constraints of the household, $(8a)$

---

[2] Note 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.

and (15), can be consolidated into a single budget constraint,

$$\sum_{t=0}^{\infty} q_t \left[(1 + \tau_t^c) c_t - (1 - \tau_t^n) w_t n_t\right] \leq (1 - l_0) \left[b_0 + p_0 s_0 + (1 - \tau_0^d) d_0 s_0\right] + T_0. \quad (15)$$

Substituting for the price of the stock from (14) for $t = 0$, and using (7) as well as $s_0 = 1$, the budget constraint can be written as

$$\sum_{t=0}^{\infty} q_t \left[(1 + \tau_t^c) c_t - (1 - \tau_t^n) w_t n_t\right] \leq W_0 + T_0 \quad (16)$$

where the initial wealth of the household, excluding the lump sum transfer, is given by

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

A *competitive equilibrium* for this economy consists of a set of allocations $\{c_t, n_t, k_{t+1}, b_{t+1}, s_{t+1}, d_t\}$, prices $\{q_t, p_t, w_t\}$, and policies $\{\tau_t^c, \tau_t^n, \tau_t^d, \tau_t^k, l_0, T_0\}$, given $\{k_0, b_0, s_0\}$ such that households maximize utility subject to their constraints, firms maximize value and markets clear in that resource constraints (2) are satisfied and the market for shares clears in that $s_t = 1$ for all t.

Note that we have not explicitly specified the government's budget constraint because it is implied by the household's budget constraint and market clearing.

We find it convenient to 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 best competitive equilibrium and the *Ramsey allocation* is the associated implementable allocation.

Given our focus on the extent to which optimal tax systems distort intertemporal decisions, we begin by providing a partial characterization of the distortions introduced by taxes. To obtain this characterization consider the first order conditions associated with the equilibrium when lump-sum taxes are available. These are given by

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1}{F_{n,t}}, \quad (17)$$

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = 1 + F_{k,t+1} - \delta, \quad (18)$$

$$\frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{F_{n,t}}{F_{n,t+1}} \left[1 + \left[F_{k,t+1} - \delta\right]\right], \tag{19}$$

and the resource constraints (2). We have explicitly characterized the intertemporal labor margin in (19) because we will be interested in understanding when it is optimal not to distort this margin.

In the growth model with distorting taxes we can combine the first order conditions of the household and the firm to obtain that taxes introduce wedges in those first order conditions as follows

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{(1 + \tau_t^c)}{(1 - \tau_t^n) F_{n,t}}, \tag{20}$$

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = \frac{\left(1 - \tau_{t+1}^d\right)\left(1 + \tau_t^c\right)}{\left(1 - \tau_t^d\right)\left(1 + \tau_{t+1}^c\right)} \left[1 + \left(1 - \tau_{t+1}^k\right)\left[F_{k,t+1} - \delta\right]\right], \tag{21}$$

and

$$\frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{\left(1 - \tau_{t+1}^d\right)\left(1 - \tau_t^n\right)}{\left(1 - \tau_t^d\right)\left(1 - \tau_{t+1}^n\right)} \frac{F_{n,t}}{F_{n,t+1}} \left[1 + \left(1 - \tau_{t+1}^k\right)\left[F_{k,t+1} - \delta\right]\right] \tag{22}$$

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. In this sense, a constant dividend tax is equivalent to a levy on the initial capital stock. 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. Indeed, as shown below, many tax systems can implement the same allocations.

A competitive equilibrium has *no intertemporal distortions in consumption* from period s onwards if the first order conditions (18) and (21) coincide for all $t \geq s$. Similarly, a competitive equilibrium has *no intertemporal distortions in labor* from period s onwards if the first order conditions (19) and (22) coincide for all $t \geq s$. Finally, a competitive equilibrium has *no intertemporal distortions* from period s onwards if it has no such distortions for both consumption and labor.

Implementability In order to characterize the Ramsey equilibrium we begin by characterizing the set of implementable allocations. In order to do so, we substitute prices and taxes from the first order conditions for the household into the household's budget constraint (16), and use the nonnegativity of the lump sum transfers, to obtain

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t} c_t + u_{n,t} n_t \right] \geq \mathcal{W}_0. \tag{23}$$

where

$$\mathcal{W}_0 = \frac{u_{c,0}}{(1 + \tau_0^c)} W_0. \tag{24}$$

Thus, any implementable allocations together with initial conditions and period zero policies must satisfy (67) and the resource constraints (2). We now show that the converse also holds. Specifically, consider an arbitrary allocation, that together with initial conditions and period zero policies, satisfies (67) and (2). We will show that this allocation is implementable. To do so, we construct the remaining elements of the allocation, prices and policies and show that all the conditions of a competitive equilibrium are satisfied.

Since multiple tax systems can implement the same allocation, for simplicity we begin by considering the case where $\tau_t^k = 0$ for $t \geq 1$ and $\tau_t^c = \tau_0^c$ for all t. The wage rates w_t are pinned down by (5), and the tax rate on labor τ_t^n is pinned down by (20). Given τ_0^d, the time path of dividend taxes is pinned down by (21) while the time path of consumption prices q_t for $t \geq 1$ is determined by (4) and (6), given $q_0 = 1$. Finally, (14) determines the stock prices p_t, the household's flow of funds determines debt holdings b_{t+1} and dividends d_t are given by (3). It is immediate that these allocations satisfy all the marginal conditions for households and firms. The lump sum transfers are chosen to satisfy the household budget constraint. Thus, the so constructed allocation, prices and policies are a competitive equilibrium.

We summarize this discussion in the following proposition

Proposition 1: (Characterization of the implementable allocations). Any implementable allocation satisfies the implementability constraint (67) 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 $(\tau_0^c, \tau_0^d, \tau_0^k, l_0)$, satisfies (67) and (2), it is implementable.

We emphasize that each implementable allocation can be implemented in numerous ways. For example, consider a tax system which arbitrarily specifies a sequence of taxes on capital income $\widetilde{\tau}_t^k$. The other taxes can be constructed using a similar procedure to the one described above. Alternatively, if taxes on dividends are set to zero, time varying taxes on consumption and labor can be chosen to implement the same allocations.

Given that capital taxes here are redundant instruments, what does it mean that capital should not be taxed? In our view, the relevant question is whether it is optimal to have no intertemporal distortions.

Next we consider restrictions on tax rates. One common practice is to impose an upper bound on the capital income tax. One justification for this upper bound is that the tax revenue ought not to exceed the base, so that $\tau_t^k \leq 1$. Such restrictions are imposed in Chamley (1986) and Judd (1985), Bassetto and Benhabib (2006) or Straub and Werning (2015). These restrictions do not affect the set of implementable allocations because with a rich tax system there are alternative taxes.

Note that analogous restrictions on labor and dividend taxes such as $\tau_t^n \leq 1$, $\tau_t^d \leq 1$ do not restrict the set of implementable allocations. This result follows immediately from inspecting (20) and (21).

2.1   Ramsey equilibrium

Given Proposition 1, it follows that the Ramsey allocation together with period zero policies maximizes utility subject to (67) and (2). Suppose that policies are unrestricted in the sense that any one of the taxes on wealth, dividends, capital income can be greater than 100%. Then, in a Ramsey equilibrium it is possible to set $\mathcal{W}_0$ to any arbitrary value. It immediately follows that it is possible to implement the lump-sum tax allocation as the Ramsey equilibrium and we have the following proposition.

Proposition 2: (No distortions ever) If period zero policies are unrestricted, then the Ramsey equilibrium coincides with the lump-sum tax allocation.

Suppose now that policies and initial conditions are restricted 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

$$\mathcal{W}_0 = \bar{\mathcal{W}},$$

which we refer to as the *wealth restriction in utility terms*. One example of such a restriction is that taxes on wealth, dividends and capital income cannot exceed 100% and that the initial debt, b_0, is positive. Then, since it is possible to set the tax on wealth equal to 100%, then $\bar{\mathcal{W}}$ is zero.

With this restriction, policies, including initial policies, can be chosen arbitrarily but the household must receive a value of initial wealth in utility terms of $\bar{\mathcal{W}}$ (see

Armenter (2007) for an analysis with such a restriction). We show below that this outcome is the equilibrium outcome for an environment with partial commitment.

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

$$-\frac{u_{c,t}}{u_{n,t}} = \frac{1 + \varphi\left[1 + \sigma_t^n - \sigma_t^{nc}\right]}{1 + \varphi\left[1 - \sigma_t - \sigma_t^{cn}\right]} \frac{1}{F_{nt}}, \, t \geq 0 \tag{25}$$

$$\frac{u_{c,t}}{\beta u_{c,t+1}} = \frac{1 + \varphi\left[1 - \sigma_{t+1} - \sigma_{t+1}^{cn}\right]}{1 + \varphi\left[1 - \sigma_t - \sigma_t^{cn}\right]} \left[1 + F_{k,t+1} - \delta\right], \, t \geq 0 \tag{26}$$

$$\frac{u_{n,t}}{\beta u_{n,t+1}} = \frac{1 + \varphi\left(1 + \sigma_{t+1}^n - \sigma_{t+1}^{nc}\right)}{1 + \varphi\left(1 + \sigma_t^n - \sigma_t^{nc}\right)} \frac{F_{n,t}\left[1 + F_{k,t+1} - \delta\right]}{F_{n,t+1}}, \, t \geq 0 \tag{27}$$

together with the constraints. Here,

$$\sigma_t = -\frac{u_{cc,t}c_t}{u_{c,t}}, \, \sigma_t^n = \frac{u_{nn,t}n_t}{u_{n,t}}, \, \sigma_t^{nc} = -\frac{u_{nc,t}c_t}{u_{n,t}}, \, \sigma_t^{cn} = -\frac{u_{cn,t}n_t}{u_{c,t}},$$

and φ is the multiplier of the implementability condition.

Comparing these conditions $(25) - (27)$ with the related conditions with lump-sum taxes $(17) - (19)$ it is clear that the optimal wedges depend on own and cross elasticities of consumption and labor. If those elasticities are constant, it is optimal not to have intertemporal distortions. Note that in this case intratemporal wedges are constant and in general positive.

Note that conditions (26) and (27) imply that if elasticities are not constant over time, it is optimal to have intertemporal distortions, but whether it is optimal effectively to tax or subsidize capital accumulation depends on whether elasticities are increasing or decreasing over time.

Note also 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 proposition.

Proposition 3: (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

$$U = \sum_{t=0}^{\infty} \beta^t \left[\frac{c_t^{1-\sigma} - 1}{1 - \sigma} - \eta n_t^{\psi} \right]. \tag{28}$$

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

Proposition 4: (No intertemporal distortions ever) Suppose that preferences are given by (28) and the wealth restriction must be satisfied. Then, the Ramsey solution has no intertemporal distortions for all $t \geq 0$.

Note that the preferences above are separable and homothetic in both consumption and labor (In Appendix 2 we show that they are the only time separable preferences with those properties). We use these properties to provide intuition for the results in Section 4 below, where we relate them to results on uniform commodity taxation and production efficiency.

The Ramsey equilibrium characterized in Proposition 3 can be implemented in a variety of ways. In one implementation, the initial wealth tax rate l_0 is chosen to satisfy the wealth constraint, consumption is taxed at a constant rate over time and all other taxes are set to zero. In an alternative implementation, initial wealth is taxed, labor is taxed at a constant rate and all other taxes are set to zero.

Consider the particular case in which $\bar{\mathcal{W}} = 0$. In this case, the Ramsey policy effectively confiscates all of the household's initial wealth. If this initial wealth is large enough relative to the present value of government expenditures, it is possible to implement the lump-sum tax allocation. If it is not, then taxes are used in all periods to finance the remaining present value of government expenditures. Note that even in this case, Proposition 4 says that it is optimal not to distort intertemporal decisions. That is, in this case one implementation of the Ramsey equilibrium is to tax consumption at a constant positive rate over time and set all taxes, other than the taxes on initial wealth, to zero.

Suppose next that policies are restricted in that the household must keep at least $\mathcal{V}_0$ initial wealth, but now in units of goods rather than in utility terms. This *wealth restriction in goods units* implies that the constraint faced by the Ramsey planner on the confiscation of initial wealth is

$$\frac{W_0}{1 + \tau_0^c} \geq \bar{\mathcal{V}}.$$

The implementability constraint can then be written as

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t} c_t + u_{n,t} n_t \right] \geq u_{c,0} \bar{\mathcal{V}}.$$

The problem is the same as before except for the term on the right hand side. The Ramsey conditions are the same as before in (25), (26) and (27), for all $t \geq 1$. The conditions for period zero are different. The intertemporal condition for consumption between periods zero and one, for example, is now

$$\frac{u_{c,0}}{\beta u_{c,1}} = \frac{1 + \varphi \left(1 - \sigma_1 + \sigma_1^{cn} \right)}{1 + \varphi \left(1 - \sigma_0 + \sigma_0^{cn} - \frac{\sigma_0 \bar{\mathcal{V}}}{c_0} \right)} \left[1 + F_{k,1} - \delta \right]. \tag{29}$$

and the intratemporal condition at time zero is

$$-\frac{u_{c,0}}{u_{n,0}} = \frac{1 + \varphi \left[1 + \sigma_0^n - \sigma_0^{nc} + \sigma_0^{nc} \frac{\bar{\mathcal{V}}}{c_0} \right]}{1 + \varphi \left[1 - \sigma_0 + \sigma_0^{cn} - \sigma_0 \frac{\bar{\mathcal{V}}}{c_0} \right]} \frac{1}{F_{n,0}} \tag{30}$$

Since the Ramsey conditions for $t \geq 1$ are unaffected, as before, wether it is optimal effectively to tax or subsidize capital accumulation depends on wether elasticities are increasing or decreasing over time.

With standard macro preferences, since elasticities are constant over time, it is optimal to have no intertemporal distortions, from period one onwards. Consider now intertemporal distortions in period zero. With standard macro preferences $\sigma_1 = \sigma_0$ and the cross elasticities are zero, so that if $W_0 > 0$, (29) implies that

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

Thus, it is optimal to effectively tax capital accumulation in period zero, or subsidize the consumption good in period zero, relative to consumption in future periods. One intuition for this result is as follows. The household is entitled to an exogenous amount of wealth in period zero. The Ramsey planer finds it optimal to reduce the value of this wealth in utility terms. This value can be reduced 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 5: (No intertemporal distortions after one period) Suppose preferences satisfy (28) and the wealth restriction in goods units must be satisfied. Then the Ramsey solution has no intertemporal distortions for all $t \geq 1$. If $\bar{\mathcal{V}} > 0$, it is optimal to effectively tax capital accumulation from period zero to period one.

In Section 4 below we relate this result to results on uniform commodity taxation and production efficiency.

The Ramsey allocation can be implemented as follows. Set the initial tax rate on wealth to satisfy the wealth restriction. Set capital income and consumption taxes to zero in all periods. Set the labor income tax to satisfy (30), (25) and (??). Specifically, set the labor income tax to

$$
1 - \tau_0^n = \frac{1 + \varphi \left[1 - \sigma + \frac{\sigma \bar{\mathcal{V}}}{c_0} \right]}{1 + \varphi \left[1 + \sigma^n \right]},
$$

in period zero and to

$$
1 - \tau^n = \frac{1 + \varphi \left[1 - \sigma \right]}{1 + \varphi \left[1 + \sigma^n \right]}
$$

in all future periods. Set the dividend tax to zero in period zero and then to a constant value thereafter. This constant value τ^d must satisfy (29) and (??) so that its value is given by

$$
\frac{1 + \varphi \left(1 - \sigma \right)}{1 + \varphi \left(1 - \sigma + \frac{\sigma \bar{\mathcal{V}}}{c_0} \right)} = 1 - \tau^d.
$$

Note that under this implementation the tax rate on dividends is always less than one and is positive if $\mathcal{V}_0$ is positive. An alternative implementation uses the consumption tax rather than the dividend tax. The disadvantage of this implementation is that in order to satisfy (25) and (??) the required tax on labor income might have to be negative, to compensate the effect of the higher consumption tax after period 1. The dividend tax implementation has the advantage that this tax affects only intertemporal decisions so that the labor income tax can be chosen to satisfy the intratemporal condition. The consumption tax has the disadvantage that it affects inter and intratemporal decisions. The dividend tax has the disadvantage that, as we remarked earlier, the base on which it is levied could be negative, so that the tax would constitute a subsidy to the firm.

The standard implementation in the literature uses capital and labor income taxes

and sets consumption and dividend taxes to zero. A disadvantage of this implementation is that the capital income tax may have to be greater than 100% to implement the Ramsey allocation. Given this disadvantage, the literature typically imposes an additional restriction that the tax rates on capital income cannot exceed some upper limit $\bar{\tau}$. This restriction implies the following additional constraint to the Ramsey problem

$$\frac{\frac{u_{c,t}}{\beta u_{c,t+1}} - 1}{F_{k,t+1} - \delta} \geq 1 - \bar{\tau}.$$

This restriction may bind for a number of periods as in Chamley (1986) or forever as in Straub and Werning (2015). Straub and Werning allow the maximum tax rate to be 100% and show that the optimal solution for particularly high levels of initial debt may have the capital income tax be set at 100% forever. One intuition for the Straub and Werning finding is that by taxing capital income forever, real interest rates are zero forever and that is the way consumption in period zero can be increased the most, reducing the value of the good in the initial period. Giving that the initial real rate cannot be below zero, the whole term structure is flattened down to zero.

To see this more clearly, notice that the planner has a strong incentive to make $u_{c,0}$ small so as to reduce the value of initial wealth. We refer to this incentive as the confiscation motive. The planner however must respect the intertemporal conditions, with restricted taxes,

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

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 reduce $u_{c,0}$, thereby confiscating initial wealth. 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, and then zero. 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.

With a rich tax system, and with fixed initial policies, the confiscation motive is satisfied in one period by levying a sufficiently high dividend tax in period one, as is

16

apparent from (21). The advantage of the dividend tax is that it effectively allows the tax to apply to a larger base than the capital income tax. This advantage can be seen by inspecting (21) evaluated in period zero with zero consumption taxes and zero dividend taxes in period zero. Notice that if the dividend tax in period one was used fully at 100% the gross return on capital would be zero. In contrast, full taxation of capital income with $\tau_1^k = 1$, can only reduce the net return to zero, provided $F_{k,1} - \delta \geq 0$.

2.2 Partial commitment equilibria

The notion of 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 our 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 consider two kinds of constraints.

In the first kind, the government in any period can commit to the one period returns on assets in utility terms. The government in the following period is free to choose policies as it wishes but must respect the previously committed return constraints. In the second kind, the government in any period can commit to (a subset of) policies in the following period. We show that with the first kind of partial commitment the equilibrium coincides with that under full commitment with constraints on the initial value of wealth. With the second kind of partial commitment, equilibrium outcomes do not coincide with those under full commitment with constraints on initial policies.

Consider the environment with partial commitment on returns. In order to develop our notion of partial commitment in this environment consider the intertemporal Euler equations for bonds and capital from period $t-1$ to period t.

$$\frac{u_{c,t-1}}{\beta\left(1+r_{t-1}\right)\left(1+\tau_{t-1}^c\right)} = \frac{u_{c,t}}{\left(1+\tau_t^c\right)}, \tag{31}$$

$$\frac{u_{c,t-1}\left(1-\tau_{t-1}^d\right)}{\beta\left(1+\tau_{t-1}^c\right)} = \frac{u_{c,t}\left(1-\tau_t^d\right)\left[1+\left(1-\tau_t^k\right)\left(F_{k,t}-\delta\right)\right]}{\left(1+\tau_t^c\right)}. \tag{32}$$

Let $\lambda_{1,t}$ denote the right side of (31), and $\lambda_{2,t}$ denote the right side of (32). With partial

17

commitment, the government in period $t-1$ chooses period $t-1$ policies as well as $\lambda_{1,t}$ and $\lambda_{2,t}$. The government in period t can choose any policies but they must have the property that the induced allocations and policies must satisfy the constraints on returns

$$\lambda_{1,t} = \frac{u_{c,t}}{(1+\tau_t^c)} \tag{33}$$

and

$$\lambda_{2,t} = \frac{u_{c,t}\left(1-\tau_t^d\right)\left[1+\left(1-\tau_t^k\right)\left(F_{k,t}-\delta\right)\right]}{(1+\tau_t^c)}. \tag{34}$$

The government in period t chooses period t policies as well as $\lambda_{1,t+1}$ and $\lambda_{2,t+1}$, to constrain policies in period $t+1$.[3] We assume $\lambda_{1,0}$ and $\lambda_{2,0}$ are given.

In order to understand the nature of partial commitment here, note that the Euler equations for bonds and capital, (31) and (32), will, of course, be satisfied on the equilibrium path. The spirit of this form of partial commitment is that the government must respect these intertemporal Euler equations also off the equilibrium path. This requirement is intended to capture the idea that policies must not induce regret on the part of households on their past choices. Specifically, the government in period t is free to choose policies as it wishes but since any policy must respect the intertemporal Euler equations it will not induce regret about choices made in period $t-1$. The spirit of the assumption that $\lambda_{1,0}$ and $\lambda_{2,0}$ are given is that the economy was operating in previous periods, and the choices made in those previous periods constrain the choices in period 0 as well.

Next, we develop a notion of Markov equilibrium with partial commitment on returns which we call no regrets equilibrium. It is convenient, and without loss of generality, to think of the government in period t, as choosing allocation, policies and prices directly in that period. The state of the economy in period t is given by $s_t = \{k_t, b_t, \lambda_{1,t}, \lambda_{2,t}\}$. Let $h_t(s_t)$ denote the policy function in period t which maps the state of the economy into allocations, policies, prices and $\lambda_{1,t+1}, \lambda_{2,t+1}$. The government in period t maximizes welfare taking as given the continuation value function and the policy functions in period $t+1$, subject to the marginal conditions of agents, budget constraints and market clearing conditions in period t. Specifically, the government in

---

[3]We consider the same tax instruments, except that now, in order to treat every period alike, we set the initial wealth tax to zero, $l_0 = 0$.

period t solves the following problem

$$v_t(s_t) = \max\{u(c_t, n_t) + \beta v_t(s_{t+1})\}, \tag{35}$$

subject to the period t equilibrium conditions and (33) and (34). Note that the policy function $h_{t+1}(s_{t+1})$ enters the intertemporal Euler equations in these equilibrium conditions. For example, period $t+1$ policies on consumption, labor, and the consumption tax, appear in the households bond Euler equation,

$$\frac{u_{c,t}}{(1 + \tau_t^c)} = (1 + r_t)\frac{\beta u_c(c_{t+1}(s_{t+1}), n_{t+1}(s_{t+1}))}{(1 + \tau_{t+1}^c(s_{t+1}))} \tag{36}$$

where $c_{t+1}(s_{t+1})$, $n_{t+1}(s_{t+1})$ and $\tau_{t+1}^c(s_{t+1})$ are elements of $h_{t+1}(s_{t+1})$.

A Markov equilibrium with partial commitment on returns, a *no regrets equilibrium*, consists of value functions $v_t(s_t)$ and policy functions, $h_t(s_t)$, which solve (35) for all s_t and all t.

Next we show that the Markov equilibrium outcome coincides with the Ramsey outcome with wealth constraints. Using the same logic as in our characterization result in Proposition 1, it is straightforward to show that the period t equilibrium conditions can be equivalently represented by the resource constraint and by the following period t implementability constraint,

$$\beta\lambda_{1,t+1}b_{t+1} + \beta\lambda_{2,t+1}k_{t+1} = \lambda_{1,t}b_t + \lambda_{2,t}k_t - u_{n,t}n_t - u_{c,t}c_t. \tag{37}$$

Multiplying these constraints by β^t and summing up yields the implementability constraint of the Ramsey problem. Thus the Ramsey allocation is feasible. Note that future controls do not appear in the objective function, (35), or the constraint set which includes (37). We can then use an identical argument to that in Stokey and Lucas with Prescott (1989)[4], to show that the functional equation in (35) solves the date zero sequence problem.

We have proved the following proposition.

Proposition 6: (Partial commitment is full commitment). The Markov outcome of an economy with partial commitment in returns coincides with the Ramsey outcome with wealth restriction given by $W_0 = \lambda_{1,0}b_0 + \lambda_{2,0}k_0$.

---

[4]Theorem 4.3

Kydland and Prescott (1980) propose a method to compute 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 which is the analog to $\lambda_{1,t}$ and $\lambda_{2,t}$ in our environment. The government in period zero maximizes discounted utility while being unconstrained by the added state variable. An extensive literature has exploited this recursive formulation to characterize commitment outcomes. We show here that their clever insight can be used to prove that equilibria in environments where policy makers are constrained not to induce regret coincide with equilibria with full commitment and initial wealth constraints.

A partial commitment equilibrium on instruments   Consider next an alternative form of partial commitment. In this form, the government in period t chooses a subset of policies, $\left\{\tau_{t+1}^c, \tau_{t+1}^k, \tau_{t+1}^d\right\}$ 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. The spirit of this assumption is that in the literature, as already discussed, this subset of policies is exogenously fixed in period zero. We extend this spirit to allow for partial commitment to that subset of instruments in every period. We show that the Markov equilibrium, with this form of partial commitment does not in general coincide with the Ramsey outcomes with exogenously specified initial taxes. Together with our results on partial commitment on returns, this result shows that the nature of partial commitment plays a crucial role in determining whether Markov equilibria coincide with commitment equilibria.

Consider, the implementability constraint with this form of partial commitment. It follows from (33) and (34), that the implementability constraint (37) can be rewritten as

$$\beta\widetilde{\lambda}_{1,t+1}b_{t+1} + \beta u_{c,t+1}\widetilde{\lambda}_{2,t+1}k_{t+1} = \widetilde{\lambda}_{1,t}b_t + \widetilde{\lambda}_{2,t}k_t - u_{n,t}n_t - u_{c,t}c_t, \qquad (38)$$

where

$$\widetilde{\lambda}_{1,t+1} \equiv \frac{u_{c,t+1}}{\left(1 + \tau_{t+1}^c\right)} \text{ and } \widetilde{\lambda}_{2,t+1} \equiv \frac{u_{c,t+1}\left(1 - \tau_{t+1}^d\right)\left[1 + \left(1 - \tau_{t+1}^k\right)\left(F_{k,t+1} - \delta\right)\right]}{\left(1 + \tau_{t+1}^c\right)}.$$

Notice here that $\widetilde{\lambda}_{1,t+1}$ and $\widetilde{\lambda}_{2,t+1}$ depend not only on policies chosen in the current period, $\left\{\tau_{t+1}^c, \tau_{t+1}^k, \tau_{t+1}^d\right\}$, but also on allocations and policies that will be chosen in

the next period.

A Markov equilibrium is defined analogously to the one above. The state of the economy in period t is given by $s_t = \left\{k_t, b_t, \tau_t^c, \tau_t^k, \tau_t^d\right\}$. As before let $h_t(s_t)$ denote the policy function which maps the state of the economy into allocations, the labor tax rate, prices and the period $t+1$ taxes, $\left\{\tau_{t+1}^c, \tau_{t+1}^k, \tau_{t+1}^d\right\}$. The government in period t solves the analogous problem to the one above.

Note that in a Markov equilibrium, $\widetilde{\lambda}_{1,t+1}$, for example, is given by

$$\widetilde{\lambda}_{1,t+1} = \frac{u_c\left(c_{t+1}\left(s_{t+1}\right), n_{t+1}\left(s_{t+1}\right)\right)}{\left(1 + \tau_{t+1}^c\left(s_t\right)\right)} \tag{39}$$

This equation shows the precise sense in which $\widetilde{\lambda}_{1,t+1}$ depends on the policy function which will be followed in in the next period. The government in period t takes this future policy function as given in choosing its current optimal policy. Put differently, future controls appear in the constraint set in period t. The arguments in Stokey and Lucas with Prescott (1989) no longer apply.

Lucas and Stokey (1983) provide examples in production economies without capital where the Ramsey outcome is time inconsistent. Chari and Kehoe (1993) characterize Markov equilibria in that environment. Klein, Krusell, Rios-Rull (2008) characterize Markov equilibria in environments similar to ours, with partial commitment to instruments. The results in these papers imply that Markov outcomes are in general different from commitment outcomes.

3   Heterogeneous agent model

The results obtained above for the representative agent economy remain under certain conditions in economies with capital-rich and poor agents. In order to show this consider an economy with 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 across the two types of agents,

$$U = \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(c_t^i\right)^{1-\sigma^i} - 1}{1 - \sigma} - \eta^i \left(n_t\right)^{\psi^i} \right]. \qquad (40)$$

The resource constraints are

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

The taxes are the ones in the rich tax system considered in the representative agent economy that includes taxes on consumption τ_t^c, labour 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.

With heterogeneous agents it turns out that we do not need to impose constraints on the initial policies. In particular, for reasons pointed out in Werning (2007), it turns out that without constraints on wealth taxes, it may be optimal for the planner to distort intratemporal decisions.

The implementability conditions can be written as

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t}^1 c_t^1 + u_{n,t}^1 n_t^1 \right] = u_{c,0}^1 (1 - l_0) V_0^1, \qquad (41)$$

and

$$\sum_{t=0}^{\infty} \beta^t \left[u_{c,t}^2 c_t^2 + u_{n,t}^2 n_t^2 \right] = u_{c,0}^2 (1 - l_0) V_0^2, \qquad (42)$$

with $V_0^i = \left[b_0^i + \left(1 - \tau_0^d\right) \left[1 + \left(1 - \tau_0^k\right) (F_{k,0} - \delta)\right] k_0^i \right]$. Since the taxes must be the same for the two agents an implementable allocation must also satisfy the following marginal conditions

$$\frac{u_{c,t}^1}{u_{c,t}^2} = \frac{u_{n,t}^1}{u_{n,t}^2}$$

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

$$u_{c,t}^1 = \gamma u_{c,t}^2 \tag{43}$$

$$u_{n,t}^1 = \gamma u_{n,t}^2 \tag{44}$$

where γ is some endogenous number.[5]

Let φ^1 and φ^2 be the multipliers of the two implementability conditions, (41) and (42). The first order conditions for $t \geq 1$ imply[6]

$$u_{c,t}^2 \frac{\gamma \left[\theta + \varphi^1 \gamma \left(1 - \sigma^1\right)\right] \frac{\sigma^2}{c_t^2} + \left[(1 - \theta) + \varphi^2 \left(1 - \sigma^2\right)\right] \frac{\sigma^1}{c_t^1}}{\frac{\sigma^2}{c_t^2} + \frac{\sigma^1}{c_t^1}} = \lambda_t \tag{45}$$

which, together with

$$-\lambda_t + \beta \lambda_{t+1} \left[f_{k,t+1} + 1 - \delta\right] = 0$$

implies that, if elasticities are equal, $\sigma^1 = \sigma^2 = \sigma$, future capital should not be taxed from period one on.

In period zero the first order condition for consumption of type one has an additional term, so that this condition is

$$\theta u_{c,0}^1 + \varphi^1 u_{c,0}^1 \left(1 - \sigma_0^1\right) + \mu_0 u_{cc,0}^1 - \varphi^1 u_{cc,0}^1 \left(1 - l_0\right) \left(\gamma \varphi^1 V_0^1 + \varphi^2 V_0^2\right) = \lambda_0 \tag{46}$$

The first order condition with respect to l_0 is

$$-u_C^1 (0) \left(\gamma \varphi^1 V_0^1 + \varphi^2 V_0^2\right) = 0$$

Thus, the additional term is zero, so that first order condition for period zero, (46), has the same form as the ones for $t \geq 1$.

Consider next the additional restriction that the initial wealth tax has to be lower than 100%, $l_0 \leq 1$. If the solution to this problem is interior, the additional term is zero. If the solution has $l_0 = 1$, the additional term is also zero.

We summarize this discussion in the following proposition.

Proposition 6: (No intertemporal distortions in heterogenous agent economies).

---

[5] See also Greulichy, Laczó and Marcet (2016). Werning (2007) also computes optimal taxes with heterogeneous agents taking adavantage of this proportionality of marginal utilities.

[6] See Appendix... for the derivation.

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 $t \geq 0$.

Note that this proposition holds even if we impose the additional restriction that $l_0 \leq 1$.

This proposition shows that with standard and identical preferences, allowing for heterogeneity in initial wealth does not overturn the result that, with a rich tax system, future capital should not be taxed.

With heterogeneity and distributional concerns it may be optimal for the planner to distort intratemporal decisions regardless of whether the initial wealth tax is constrained to be below 100% or not. This result is in striking contrast with the result in the representative agent model. In that model, as stated in Proposition 2, if the initial wealth tax is unconstrained, the outcome coincides with the lump sum tax allocations and the intratemporal decisions are undistorted.

Consider next the case in which agents' preferences differ, in that $\sigma^1 \neq \sigma^2$. In this case it is possible to show that if the two types of agents can be taxed at different rates, it is optimal not to have intertemporal distortions. The reason is that restrictions (43) and (44) can now be dropped and the analog of proposition 6 can be proved. If the two types of agents cannot be taxed at different rates, then it may be optimal to distort intertemporal decisions.

4   Relation to production efficiency

In this Section we connect our results to results on production efficiency and uniform taxation. To develop these connections we set up an alternative economy, that 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 denoted by N. Preferences for the household are given by

$$U\left(C, N\right) = \frac{C^{1-\sigma} - \frac{1}{1-\beta}}{1-\sigma} - \eta N^{\psi}. \tag{47}$$

The economy has three types of firms. The first one is the same as the one described above. We refer to this as the *capital accumulation firm.* This firm produces interme-

24

diate goods c_t, hires intermediate labor inputs n_t, and accumulates capital according to the technology (2). The second type of firm, referred as the *consumption firm* produces the final good C using the intermediate goods c_t according to the constant returns to scale technology given by

$$C = \mathcal{C}(c_0, c_1..) = \left[\sum_{t=0}^{\infty} \beta^t c_t^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \tag{48}$$

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

$$N = \mathcal{N}(n_0, n_1..) = \left[\sum_{t=0}^{\infty} \beta^t n_t^{\psi} \right]^{\frac{1}{\psi}}. \tag{49}$$

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 good firm nor the labor firm because these profits will be zero in equilibrium.

The household's problem is to maximize (47) subject to the budget constraint

$$p(1 + \tau^C)C - w(1 - \tau^N)N \leq (1 - l_0)V_0 \tag{50}$$

where p and w denote the prices of final consumption and labor in units of the consumption good in period 0, and V_0 is the value of initial wealth in units of goods.

The consumption firm's problem is to maximize

$$pC - \sum_{t=0}^{\infty} q_t(1 + \tau_t^c)c_t \tag{51}$$

subject to (48).

The labor firm's problem is to maximize

$$\sum_{t=0}^{\infty} q_t(1 - \tau_t^n)w_t n_t - wN \tag{52}$$

subject to (49).

The capital accumulation firm's problem is the same as above. 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 and the labor firm directly into the households problem.

Consider the budget constraint of the household. 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 (61) and $wN = \sum_{t=0}^{\infty} q_t(1 - \tau_t^n)w_t n_t$ from (52) into (60) to obtain a budget constraint of the form (16) in the growth economy with taxes. The only difference is the presence of the tax on the final consumption and on final labor, which amounts to rescaling the consumption and labor income taxes in the original economy.

Substituting from (49) and (48) into (47), we see that the household's 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 household's 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 of the final values for consumption and labor. Thus, the equilibria in the two economies coincide.

In the intermediate goods economy, the Ramsey problem is to maximize (47) subject to the implementability constraint

$$U_C C + U_N N = \frac{U_C}{(1 + \tau^C)} (1 - l_0) V_0, \tag{53}$$

and the requirement that the allocation is in the *production set* given by (48) and (49)

with inequalities and (2).

Suppose now, as before, that the initial wealth in utility terms, given by the right hand side of (53) must be at least as large as an exogenous value, $\bar{\mathcal{W}}$.

Next we apply the production efficiency theorem of Diamond and Mirrlees to our intermediate goods economy. This theorem asserts that if pure rents are fully taxed away, any Ramsey allocation must be at the boundary of the production set. In our context, the theorem asserts that if $\bar{\mathcal{W}} = 0$, any Ramsey allocation must be at the boundary of the production set. In the Appendix we extend this theorem to the case where $\bar{\mathcal{W}}$ is exogenously fixed.

Somewhat loosely speaking, the theorem as extended, implies that one way to implement the Ramsey allocation is to tax only final goods and not to tax intermediate goods. Thus, the theorem implies that it is optimal not to have intertemporal distortions.

More formally, the result that a Ramsey allocation must be at the boundary implies that there exist supporting prices p and w such that the allocation solves

$$\max \ pC - wN$$

subject to the requirement that the allocation is in the production set.

The first order conditions for this problem include

$$\frac{\mathcal{C}_{ct}}{\mathcal{C}_{ct+1}} = 1 - \delta + F_{k,t+1}, \tag{54}$$

and

$$\frac{\mathcal{N}_{nt}}{\mathcal{N}_{nt+1}} = \frac{F_{n,t}}{F_{n,t+1}} \left(1 - \delta + F_{k,t+1}\right). \tag{55}$$

These are conditions of production efficiency. Condition (54) 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 (55) is the analog for labor in consecutive periods. Notice that, since $\beta^t u_c(t) = U_c \mathcal{C}_{ct}$ and $\beta^t u_n(t) = U_N \mathcal{N}_{nt}$, it follows that $\frac{\mathcal{C}_{ct}}{\mathcal{C}_{ct+1}} = \frac{u_{c,t}}{\beta u_{c,t+1}}$ and $\frac{\mathcal{N}_{nt}}{\mathcal{N}_{nt+1}} = \frac{u_{n,t}}{\beta u_{n,t+1}}$. Thus, conditions (54) and (55) imply that it is optimal to have no intertemporal distortions. They also imply that intratemporal distortions are constant. We summarize this discussion in the following proposition, which is the equivalent to Proposition 4 in the intermediate good economy.

Proposition 4': (No intertemporal distortions ever) Suppose that preferences and technologies are given by (47), (48) and (49) 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 the wealth restriction is imposed in units of goods rather than in utility terms. Here, production efficiency is no longer optimal. Nevertheless, it is straightforward to show that an analogue of Proposition 5 holds in the intermediate good economy in the sense that it is not optimal to distort intermediate goods decisions from period one onwards. It is optimal to distort intermediate goods in period zero relative to all other intermediate goods in order to partially tax rents.

The intermediate good economy helps clarify circumstances in which it is optimal not to have intertemporal distortions. We have shown that it is optimal not to 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. Thus, our result of no intertemporal distortions is very closely connected to the uniform commodity taxation result.

Production efficiency in an economy without capital   How different are the results above from the ones in an economy without capital as in Lucas and Stokey (1983)? In that production economy, uniform taxation is optimal for standard macro preferences. How does that relate to production efficiency? Can we map that economy into an intermediate good economy as we did in the growth model?

In Lucas and Stokey, instead of the intertemporal technology (2), the technology is static and given by

$$c_t + g_t \leq A_t n_t. \tag{56}$$

The map between the original economy and the intermediate good economy is the same except that instead of the capital accumulation firms, we have now *production firms* that use technology (56). The conditions for production efficiency in the inter-

mediate good economy are now

$$\frac{\mathcal{C}_{ct}}{\mathcal{C}_{ct+1}} = \frac{\mathcal{N}_{nt}}{\mathcal{N}_{nt+1}} \frac{A_{t+1}}{A_t}, \tag{57}$$

instead of (54) and (55).

Because $\mathcal{C}_{ct}$ and $\mathcal{N}_{nt}$ are proportionate to $u_{c,t}$ and $u_{n,t}$, condition (57) implies that intratemporal distortions are constant over time. There are no implications for intertemporal wedges because there are no such wedges.

4.1 Production efficiency in an heterogeneous agent economy

Consider now developing an intermediate goods economy that is equivalent to our heterogenous agent economy. Here we think of the intermediate good 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 household of type i are given by

$$U\left(C^i, N^i\right) = \frac{(C^i)^{1-\sigma^i} - \frac{1}{1-\beta}}{1 - \sigma^i} - \eta \left(N^i\right)^\psi. \tag{58}$$

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

$$C^i = \mathcal{C}^i(c_0^i, c_1^i..) = \left[\sum_{t=0}^\infty \beta^t \left(c_t^i\right)^{1-\sigma^i}\right]^{\frac{1}{1-\sigma^i}}. \tag{59}$$

for $i = 1, 2$.

The tax system is the same except that we require that the tax rate on final consumption goods must be the same for the two types. We do so to show the equivalence between the intermediate goods economy and the heterogenous agent economy with type independent taxation.

Household of type i maximizes (58) subject to the budget constraint

$$p^i(1 + \tau^C)C^i - w(1 - \tau^N)N^i \le (1 - l_0) V_0^i \tag{60}$$

where p^i denotes the price of final consumption good of type i, in units of the consumption good in period 0.

The consumption firm of type i maximizes

$$p^i C^i - \sum_{t=0}^{\infty} q_t (1 + \tau_t^c) c_t^i \tag{61}$$

subject to (59).

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 heterogenous agent economy with type independent taxes coincide. A key step in this proof is that the tax rates on the final consumption goods are restricted to be the same. Recall that in the representative agent model we used zero profits for consumption goods producers and replaced for the pre-tax value of consumption in the households budget constraint in the intermediate goods economy by the value of the intermediate goods. Following the same procedure, we can use (61) and write (60) 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 heterogenous agents economy except for rescaling. Notice that if the tax rates in the two consumption goods were allowed to be different in the intermediate goods economy, the budget constraints would not coincide and the competitive equilibria would be different.

The rest of the argument that the equilibria coincide is the same as in the representative agent economy. Clearly, if we were to set-up an intermediate goods economy that coincided with the heterogenous agent economy with type-dependent taxes, we could allow the tax rates in the two final consumption goods to be different in the intermediate good economy.

Consider now the Ramsey problem in the intermediate goods economy with the same tax rate on both consumption goods. This requirement imposes an additional restriction on the Ramsey problem. Straightforward algebra shows that this constraint can be written as

$$\frac{U^i_{C^i}\mathcal{C}^i_{c_t}}{-U^i_{N^i}\mathcal{N}^i_{n_t}} = \frac{U^j_{C^i}\mathcal{C}^j_{c_t}}{-U^j_{N^i}\mathcal{N}^j_{n_t}}, \, t \geq 0. \tag{62}$$

The Ramsey problem is now to maximize utility subject to implementability constraint, the requirement that the allocation is in the production set and (62). The Ramsey problem in the intermediate good economy with type-dependent taxes simply drops (62).

We now have the analog of proposition 6.

Proposition 6': (No intertemporal distortions in heterogenous 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 agents' preferences differ, and if tax rates on consumption and labor income can be type-dependent, then the Ramsey equilibrium has no intertemporal distortions for all $t \geq 0$.

5   Concluding remarks

Should capital be taxed in the steady state, and along the transition? Once we abstract from the initial confiscation of capital, what matters are consumption and labor elasticities. In the steady state they are constant, so capital taxes should be zero in the steady state. For standard macro preferences, those elasticities are always constant, so that (future) capital should never be taxed. These results hold with heterogeneous agents.

Future capital should not be taxed for standard preferences, because taxing it imposes different taxes on different consumption goods and on labor in different periods, when instead uniform consumption and labor taxation is optimal.

Optimal uniform taxation is about production efficiency. We relate the results to the principles of production efficiency of Diamond and Mirrlees (1972).

Indirect confiscation of the initial capital through valuation effects could justify taxing future capital, but that reason to tax future capital lasts for one single period. This is the case indeed if a rich set of taxes is used, that includes taxes used in practice such as dividend or consumption taxes. In any case, direct confiscation is always preferable to indirect confiscation.

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, 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, 3–18.

[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 9 211-223.

[6] Chamley, Christophe,1986 "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," Econometrica 54 (3), pp. 607–622.

[7] Chari, V.V. and Patrick Kehoe, 1993, "Sustainable Plans and Debt," Journal of Economic Theory 61, 230–261.

[8] Chari, V.V., Lawrence J. Christiano, and Patrick J. Kehoe, 1994, "Optimal Fiscal Policy in a Business Cycle Model," Journal of Political Economy102 (4), 617–52.

[9] Coleman II, Wilbur John, 2000, "Welfare and optimum dynamic taxation of consumption and income," Journal of Public Economics 76 (1), 1–39.

[10] Diamond, Peter A and James A Mirrlees, 1971, "Optimal Taxation and Public Production: I– Production Efficiency," American Economic Review 61 (1), 8–27.

[11] Greulichy, Katharina, Sarolta Laczó and Albert Marcet, 2016, "Pareto-Improving Optimal Capital and Labor Taxes," mimeo, Barcelona GSE.

[12] Judd, Kenneth L., 1985, "Redistributive taxation in a simple perfect foresight model," Journal of Public Economics 28 (1), 59 – 83.

[13] Judd, Kenneth L., 1999, "Optimal taxation and spending in general competitive growth models," Journal of Public Economics 71 (1), 1–26.

[14] Judd, Kenneth L., 2002, "Capital-Income Taxation with Imperfect Competition," American Economic Review Papers and Proceedings 92 (2), 417–421.

[15] Klein, Paul, Per Krusell and Victor Rios-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, 55–93.

[18] Straub, Ludwig and Ivan Werning, 2015, Positive Long Run Capital Taxation: Chamley-Judd Revisited, mimeo, MIT

[19] Werning, Ivan, 2007, "Optimal Fiscal Policy with Redistribution," Quarterly Journal of Economics 122, 925-967.

[20] Zhu, Xiaodong, 1992, "Optimal fiscal policy in a stochastic growth model," Journal of Economic Theory 58 (2), 250-289.

6   Appendix 1: An alternative decentralization

An economy that would be equilivalent to the one we just studied but that would be closer to the standard set up used in this literature, would have the household investing in 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 more closely 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 can then be written as

$$\frac{1}{1+r_{t+1}}b_{t+1} + K_{t+1} = b_t + (1-\delta) K_t + U_t K_t - \tau_t^d \left[U_t K_t - (K_{t+1} - (1-\delta) K_t) \right]$$
$$+ (1-\tau_t^n) w_t N_t - (1+\tau_t^c) C_t,$$

for $t \geq 1$, which can be rewritten as

$$\frac{1}{1+r_{t+1}}b_{t+1} + \left(1-\tau_t^d\right) K_{t+1} = b_t + \left(1-\tau_t^d\right)(1-\delta) K_t + \left(1-\tau_t^d\right) U_t K_t$$
$$+ (1-\tau_t^n) w_t N_t - (1+\tau_t^c) C_t,$$

In the initial period the constraint is

$$\frac{1}{1+r_1}b_1 + \left(1-\tau_0^d\right) K_1 = (1-l_0) \left[b_0 + \left(1-\tau_0^d\right)(1-\delta) K_0 + \left(1-\tau_0^d\right) U_0 K_0 \right]$$
$$+ (1-\tau_0^n) w_0 N_0 - (1+\tau_0^c) C_0.$$

The marginal conditions are (11), (12), and

$$1 + r_{t+1} = \frac{\left(1-\tau_{t+1}^d\right)\left[U_{t+1} + 1 - \delta\right]}{1-\tau_t^d}. \tag{63}$$

The representative firm maximizes profits

$$\Pi_t = A_t F(K_t, N_t) - W_t N_t - U_t K_t - \tau_t^k (U_t - \delta) K_t$$

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 marginal cost,

$$1 = \frac{W_t}{A_t F_n(t)} = \frac{U_t + \tau_t^k (U_t - \delta)}{A_t F_k(t)}, \tag{64}$$

These marginal conditions can be written as (20), (21) and (??).

The budget constraint of the households can be written as

$$\sum_{t=0}^{\infty} q_t \left[\left(1 + \tau_t^c \right) C_t - \left(1 - \tau_t^n \right) W_t N_t \right] =$$
$$\left(1 - l_0 \right) \left[b_0 + \left(1 - \tau_0^d \right) \left(1 - \delta \right) K_0 + \left(1 - \tau_0^d \right) U_0 K_0 \right] \qquad (65)$$

where $q_t = \frac{1}{(1+r_1)\ldots(1+r_t)}$ for $t \geq 1$, with $q_0 = 1$.. This uses the no-Ponzi games condition $\lim_{T\to\infty} q_{T+1} b_{T+1} \geq 0$.

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

$$\sum_{t=0}^{\infty} \beta^t \left[u_C \left(t \right) C_t + u_N \left(t \right) N_t \right] = \frac{\left(1 - l_0 \right)}{1 + \tau_0^c} \left[b_0 + \left(1 - \tau_0^d \right) K_0 + \left(1 - \tau_0^d \right) \frac{A_0 F_k \left(0 \right) - \delta}{\left(1 + \tau_0^k \right)} K_0 \right]$$

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 household can keep, measured in units of utility at time 0,

$$u_C \left(0 \right) \frac{\left(1 - l_0 \right)}{1 + \tau_0^c} \left[b_0 + \left(1 - \tau_0^d \right) K_0 + \left(1 - \tau_0^d \right) \frac{A_0 F_k \left(0 \right) - \delta}{\left(1 + \tau_0^k \right)} K_0 \right] = \mathcal{W}_0 \qquad (66)$$

The implementability condition can then be written as

$$\sum_{t=0}^{\infty} \beta^t \left[u_C \left(t \right) C_t + u_N \left(t \right) N_t \right] = \mathcal{W}_0. \qquad (67)$$

The implementability condition (67) together with the resource constraints (2) are the only equilibrium restrictions on the sequences of consumption, labor and capital. There are natural restrictions on the taxes, that the tax revenue does not exceed the base, so that $\tau_t^k \leq 1$, for all t. These restrictions will not be binding in this set up.

The other equilibrium condition, other than (67) and (2), are satisfied by other variables

$$- \frac{u_C \left(t \right)}{u_N \left(t \right)} = \frac{\left(1 + \tau_t^c \right)}{\left(1 - \tau_t^n \right) W_t}, \qquad (68)$$

determines τ_t^n

$$\frac{u_C \left(t \right)}{1 + \tau_t^c} = \left(1 + r_{t+1} \right) \frac{\beta u_C \left(t + 1 \right)}{1 + \tau_{t+1}^c}, \qquad (69)$$

determines r_{t+1}

$$1 = \frac{W_t}{A_t F_n(t)}, \tag{70}$$

determines W_t

$$\frac{W_t}{A_t F_n(t)} = \frac{U_t + \tau_t^k (U_t - \delta)}{A_t F_k(t)}, \tag{71}$$

determines U_t

$$u_C(t) = \beta u_C(t+1) \frac{(1 + \tau_t^c)(1 - \tau_{t+1}^d)}{(1 + \tau_{t+1}^c)(1 - \tau_t^d)} [U_{t+1} + (1 - \delta)]. \tag{72}$$

determines either τ_{t+1}^c, given τ_t^c, or else τ_{t+1}^d given τ_t^d The constraint (66) 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 restrictions $\tau_t^k \leq 1$ will not bind. This also means that the capital tax can be set always equal to zero.