More on the Taxation of Capital*

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

April 2016

Abstract

Should capital income taxes be zero in the long run, as argued by Chamley (1986) and Judd (1985)? Or should instead capital be heavily taxed as suggested by Straub and Werning (2015)? We revisit the Ramsey literature on the optimal taxation of capital and make again the case for a low, possibly zero, tax on capital income.

Keywords: capital income tax; long run; uniform taxation

JEL Codes: E60; E61; E62

*Preliminary. We would like to thank Isabel Correia, João Guerreiro, Albert Marcet. Nicolini and Teles would like to acknowledge 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 N° 49396. 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 adresses: juanpa@minneapolisfed.org, pteles@ucp.pt.

1 Introduction

Should capital income taxes be zero in the long run, as argued by Chamley (1986) and Judd (1985)? Or should instead capital be heavily taxed as suggested by Straub and Werning (2015)? We revisit the Ramsey literature on the optimal taxation of capital and make again the case for a low, possibly zero, tax on capital income.

In the problem of the optimal taxation of capital a key distinction is whether it is future capital or current capital that is being taxed, whether capital is being taxed or confiscated. This is not an easy distinction since taxing future capital also taxes current capital. If incentives to confiscate initial capital were purely temporary, focusing on the long run taxation of capital would be a way of abstracting from the confiscation of the initial stock. However, the results of Straub and Werning (2015) surprisingly show that not to be the case. An alternative way, the one we choose to take here, is to abstract from the confiscation of the initial capital stock including the indirect effects on its valuation, a strategy that Armenter (2008) also follows.

Once the initial confiscation is ruled out, then the reasons for not taxing capital income become apparent. Capital should not be taxed when consumption and labor ought to be taxed at uniform rates over time. Roughly, when consumption and labor elasticities are constant over time, consumption and labor should be taxed at constant rates over time, and therefore there is no reason to tax capital. If instead consumption or labor elasticities vary over time, then goods should be taxed at varying rates and capital taxes (or subsidies) can be a way of accomplishing this. In a steady state, elasticities are constant and that explains the results in Chamley and Judd.

Taxing future capital is a very inefficient way of confiscating current capital. This is the first main point of the paper. A very clear example of this is in Straub and Werning (2015). They give examples where the optimal policy in Chamley's representative agent model is to tax capital income at 100% forever, so that the optimal allocations converge to zero capital and zero consumption. The reason for this is to confiscate the initial private wealth in a context where the capital income tax is capped at a natural limit of 100%. Ideally only the second period capital income tax would be large, but then it may have to be larger than 100%. The third best policy would then use the whole term structure, taxing capital forever at its limit. This confiscatory policy, as shown here, can be replaced with a sizeable gain, by a relatively high consumption tax in the second period. The reason the consumption tax can do better than the capital income tax is that it is not subject to the same limit as the capital tax. Both Chamley and Judd and Straub and Werning consider only labor and capital income taxes. When both labor and consumption taxes are used, consumption taxes in the second period can achieve the desired confiscation, and zero capital income taxes are again optimal in the steady state.

A better way to confiscate than using future consumption taxes is to tax the initial capital stock directly. It is not clear why the same purpose, of raising a particular levy, should be achieved in a less efficient way with future consumption taxes, or even less efficiently through future capital income taxes. For this reason, we impose limits not on the initial tax rates but on the resulting confiscation, including indirect valuation effects. Once this assumption is made, the conditions for not taxing capital are conditions for uniform taxation of consumption and labor.

The second main point of the paper is to explain why capital should or should not be taxed, once confiscation is ruled out, relating it to the optimal taxation principles in the public finance literature of Diamond and Mirrlees (1971) and Atkinson and Stiglitz (1972). Taxing capital imposes differential taxation on consumption and labor over time. When is that desirable?

Suppose it was the case that consumption in different periods should be taxed at a common rate, and similarly, labor in different periods should also be taxed at a common rate. Then, it would not be optimal to use the capital tax to differentially tax the different goods. The conditions for uniform taxation of consumption and labor in different periods are conditions similar to the ones in Atkinson and Stiglitz (1972). In Atkinson and Stiglitz, if preferences are separable in labor and homothetic in the consumption goods, then the consumption goods should be taxed at a uniform rate. Similarly, a generalization of Atkinson and Stiglitz to a model with one consumption good and multiple labor types, will also prescribe common taxation of labor if preferences are separable in the consumption good and homothetic in the types of labor.

Additively separable preferences that are isoelastic in both consumption and labor are preferences that are separable and homothetic in both consumption and labor. For those preferences, abstracting from the initial confiscation, consumption and labor should be taxed at a constant rate and capital should not be taxed. One simple intuition is the partial equilibrium one that goods should be taxed according to their demand elasticity. When elasticities are the same, so should taxes be. In the steady state, even if preferences are not isoelastic, elasticities are constant because allocations are constant. That is probably the clearest intuition for why capital should not be taxed in the steady state. Away from the steady state, and abstracting from the initial confiscation, to the extent that elasticities may be fairly stable, optimal taxation of capital income will be close to zero.

The application of the result in Diamond and Mirrlees (1971) that capital should not be taxed because it is an intermediate good is not straightforward in these dynamic models. In a model where labor and capital may be used to produce different goods, under the conditions of Diamond and Mirrlees, production should not be distorted and therefore intermediate goods should not be taxed. But labor in different periods in these dynamic models is not the same labor, and that makes a difference. The results of Diamond and Mirrlees cannot be directly applied.

The results are first shown for a representative agent economy (Section 2). The results extend to the case in which capital is concentrated in the hands of only a few and their weight in the social welfare function approaches zero. The only reason to tax capital in that case, as in the representative agent economy, is if price elasticities vary with the allocations along the transition. We show this in Section 3.

Other relevant literature includes Chari, Christiano, Kehoe (1994), Atkeson, Chari and Kehoe (1999), Judd (1999, 2002), Coleman (2000), Abel (2007), Lucas and Stokey (1983), Greulichy, Laczó and Marcet (2016).

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,

$$U = \sum_{t=0}^{\infty} \beta^t u\left(C_t, N_t\right).$$

satisfying the usual properties.

Government consumption G_t is exogenous. The production technology is

$$C_t + G_t + K_{t+1} - (1-\delta) K_t \leq A_t F\left(N_t, K_t\right) \tag{1}$$

where K_t is capital, A_t is an aggregate productivity parameter and the production function is constant returns to scale.

The government finances public consumption with time varying taxes on consumption τ_t^c, labour income τ_t^n, capital income τ_t^k. There is also a tax on the initial wealth l_0. The consumption taxes and the initial levy are restricted to be zero in Chamley (1986) and Straub and Werning (2015).

The flow of funds for the households can be described by

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

for $t \geq 1$, and

$$\frac{1}{1+r_1} b_1 + K_1 = (1 - l_0) \left[b_0 + \left[1 + \left(1 - \tau_0^k\right)(U_0 - \delta)\right] K_0\right] + (1 - \tau_0^n) W_0 N_0 - (1 + \tau_0^c) C_0.$$

There is also a no-Ponzi games condition. b_{t+1} are real bonds that cost $\frac{1}{1+r_{t+1}}$ and pay one unit of good in period $t+1$. r_{t+1} is the real interest rate between period t and $t+1$, W_t is the wage, U_t is the rental rate of capital. Capital income is taxed with a depreciation allowance.

The household that maximizes utility subject to the budget constraint must equate the marginal rate of substitution between consumption and labor to the real wage distorted by the consumption and the labour income tax,

$$-\frac{u_C(t)}{u_N(t)} = \frac{(1 + \tau_t^c)}{(1 - \tau_t^n) W_t}. \tag{2}$$

where $u_C(t)$ stands for $u_C(C_t, 1 - N_t)$ and $u_N(t)$ for $u_N(C_t, N_t)$. The optimal decision on bonds and capital requires

$$\frac{u_C(t)}{(1 + \tau_t^c)} = (1 + r_{t+1}) \frac{\beta u_C(t+1)}{\left(1 + \tau_{t+1}^c\right)}, \tag{3}$$

$$1 + r_{t+1} = 1 + \left(1 - \tau_{t+1}^k\right)(U_{t+1} - \delta). \tag{4}$$

The price of the good must equal marginal cost,

$$1 = \frac{W_t}{A_t F_n(t)} = \frac{U_t}{A_t F_k(t)}, \tag{5}$$

with $F_n(t) \equiv F_n\left(\frac{K_t}{N_t}\right)$ and $F_k(t) \equiv F_n\left(\frac{K_t}{N_t}\right)$.

These marginal conditions can be written as

$$-\frac{u_C(t)}{u_N(t)} = \frac{(1+\tau_t^c)}{(1-\tau_t^n) A_t F_n(t)}. \tag{6}$$

$$\frac{u_C(t)}{(1+\tau_t^c)} = \left[1+\left(1-\tau_{t+1}^k\right)\left[A_{t+1}F_k(t+1)-\delta\right]\right] \frac{\beta u_C(t+1)}{\left(1+\tau_{t+1}^c\right)}, \tag{7}$$

with the intertemporal condition for labor being

$$\frac{u_N(t)}{\beta u_N(t+1)} = \frac{(1-\tau_t^n)}{\left(1-\tau_{t+1}^n\right)} \frac{A_t F_n(t)}{A_{t+1} F_n(t+1)} \left[1+\left(1-\tau_{t+1}^k\right)\left[A_{t+1}F_k(t+1)-\delta\right]\right]$$

which will make it apparent how the different taxes distort the allocation relative to the first best

The first best is the allocation that maximizes utility taking into account only the resource constraints (1). The resulting efficient allocation would be described by the usual marginal conditions

$$-u_N(t) = A_t F_n(K_t, N_t) u_C(t), \tag{8}$$

$$u_C(t) = \beta u_C(t+1)\left[A_{t+1}F_k(t+1)+1-\delta\right] \tag{9}$$

and

$$C_t + G_t + K_{t+1} - (1-\delta) K_t = A_t F(t). \tag{10}$$

The budget constraint of the households can be written as

$$\sum_{t=0}^{\infty} q_t \left[(1+\tau_t^c) C_t - (1-\tau_t^n) W_t N_t\right] = \\ (1-l_0)\left[b_0 + \left[1+\left(1-\tau_0^k\right)(U_0-\delta)\right] K_0\right] \tag{11}$$

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. The initial levy is restricted regardless of the taxes that are used to obtain it. $\mathcal{W}_0$ is the exogenous level of initial wealth that the household can keep, measured in units of utility,

$$u_C(0)(1-l_0)\left[b_0+\left[1+\left(1-\tau_0^k\right)(U_0-\delta)\right]K_0\right]=\mathcal{W}_0 \tag{12}$$

The implementability condition can then be written as

$$\sum_{t=0}^{\infty}\beta^t\left[u_C(t)C_t+u_N(t)N_t\right]=\mathcal{W}_0. \tag{13}$$

The implementability condition (13) together with the resource constraints (1) 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 $\tau_t^k \leq 1$, and $\tau_t^n \leq 1$ for all t. These restrictions will not be binding in this set up.

The other equilibrium condition, other than (13) and (1), are satisfied by other variables

$$-\frac{u_C(t)}{u_N(t)}=\frac{(1+\tau_t^c)}{(1-\tau_t^n)W_t}, \tag{14}$$

determines τ_t^n

$$\frac{u_C(t)}{1+\tau_t^c}=(1+r_{t+1})\frac{\beta u_C(t+1)}{1+\tau_{t+1}^c}, \tag{15}$$

determines r_{t+1}

$$1=\frac{W_t}{A_tF_n(t)}, \tag{16}$$

determines W_t

$$\frac{W_t}{A_tF_n(t)}=\frac{U_t}{A_tF_k(t)}, \tag{17}$$

determines U_t

$$\frac{u_C(t)}{(1+\tau_t^c)}=\frac{\beta u_C(t+1)}{\left(1+\tau_{t+1}^c\right)}\left[1-\delta+\left(1-\tau_{t+1}^k\right)U_{t+1}\right],$$

determines τ_{t+1}^c, given τ_t^c. The constraint (12) will be satisfied with l_0.

This implementation does not use τ_t^k for all t and τ_0^c. 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.

Capital taxes here are redundant instruments. So what does it mean that capital should not be taxed? The sense in which capital is not taxed is that the intertemporal margins are not distorted.

We now characterize the optimal allocations. For simplicity, we assume additive separability in preferences. The marginal conditions for a solution of the Ramsey problem follow.

$$\frac{u_C(t)}{\beta u_C(t+1)} = \frac{1+\varphi(1-\sigma(t+1))}{1+\varphi(1-\sigma(t))}\left[A_{t+1}F_k(t+1)+1-\delta\right], \, t \geq 0 \tag{18}$$

and

$$-\frac{u_N(t)}{u_C(t)} = \frac{1+\varphi(1-\sigma(t))}{1+\varphi(1+\sigma^n(t))}A_tF_N(t), \, t \geq 0, \tag{19}$$

where $\sigma(t) = -\frac{u_{CC}(t)C_t}{u_C(t)}$ and $\sigma^n(t) = \frac{u_{NN}(t)N_t}{u_N(t)}$.

The intertemporal Ramsey condition for labor is

$$\frac{u_N(t)}{\beta u_N(t+1)} = \frac{1+\varphi\left(1+\sigma^n_{t+1}\right)}{1+\varphi\left(1+\sigma^n_t\right)}\frac{A_tF_N(t)}{A_{t+1}F_N(t+1)}\left[A_{t+1}F_k(t+1)+1-\delta\right], \, t \geq 0 \tag{20}$$

The term$\frac{A_tF_N(K_t,N_t)}{A_{t+1}F_N(K_{t+1},N_{t+1})}\left[A_{t+1}F_k(K_{t+1},N_{t+1})+1-\delta\right]$ is a marginal rate of transformation of labor between time t and time $t+1$.

If $\varphi = 0$, the first best would be achieved with the intertemporal marginal rates of substitution being equal to the marginal rates of transformation. In the second best, the intertemporal marginal conditions of the first best still hold if the price elasticity of consumption and labor are constant, $\sigma(t) = \sigma$, and $\sigma^n_t = \sigma^n$. Even in the second best it is optimal not to distort the consumption and labor intertemporal margins.

For preferences that have constant elasticity of consumption, $\sigma(t) = \sigma$, it is optimal not to distort the intertemporal margin for consumption from period zero on, so that, from (7), the consumption tax can be set constant over time . The capital tax would be set to zero. Labor taxes would move over time depending on the labor elasticity $\sigma^n(t)$. Instead, if preferences have constant elasticity for labor, $\sigma^n(t) = \sigma^n$, it is optimal not to distort the intertemporal margin for labor, also from period zero on. The labor tax could be set constant over time (possibly zero) and the capital tax should then also be set to zero. The consumption taxes would move over time depending on the consumption elasticity $\sigma(t)$. If both elasticities are constant, uniform taxation is optimal and capital should not be taxed, provided neither consumption nor labor taxes move over time.

In the steady state, the elasticities are constant. It is not optimal to distort either intertemporal margin. If τ_t^c and τ_t^n are kept constant, then it is optimal not to tax capital in the steady state.

2.1 The initial confiscation

We now relate these results to Chamley (1986) and Judd (1985), as well as Straub and Werning (2015). In order to do that, we allow for valuation effects on the initial confiscation. We assume that the household has to keep at least (positive) $\mathcal{W}_0$ initial wealth, in units of goods. In period zero, the restriction the Ramsey planner faces is

$$\frac{(1-l_0)}{(1+\tau_0^c)}\left[b_0+\left[1+\left(1-\tau_0^k\right)\left(A_0F_k\left(0\right)-\delta\right)\right]K_0\right]\geq\mathcal{W}_0$$

For any τ_0^c and τ_0^k, l_0 can satisfy this. The implementability restriction can be written as

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

The Ramsey conditions imply

$$\frac{u_C\left(0\right)}{\beta u_C\left(1\right)}=\frac{1+\varphi\left(1-\sigma\left(1\right)\right)}{1+\varphi\left(1-\sigma\left(0\right)+\frac{\sigma(0)\mathcal{W}_0}{C_0}\right)}\left[1-\delta+A_1F_k\left(1\right)\right]$$

This can be satisfied with a relatively high consumption tax in period one, τ_1^c. The capital income tax may not be used for this purpose, because of the restriction that it cannot be larger than 100%.

The intuition for this result is that the lower the valuation of consumption at time zero, the lower the value of initial wealth for the household. This means that the relative value of revenues from distortionary taxation is higher: How can $u_C\left(0\right)$ be made small? By having consumption at zero be high. Consumption at zero is high if the real rate is low and a high expected consumption tax at time one can make the real rate low. A high capital income tax also makes the real rate low, but capital income cannot be taxed more than 100%. With constant elasticities τ_1^c will be higher than τ_0^c, and taxes can be kept constant after that. The tax on capital will be always zero.

There are two reasons why zero capital taxation may only be achieved in the steady state. One reason is that constant elasticities are a feature of constant allocations. The other is that there are confiscatory motives in the short run. In this set up with both labor and consumption taxes, the confiscatory motives take only one period to be dealt with. That would not be the case without consumption taxes.

In a similar set up but without consumption taxes, Straub and Werning (2015) give examples where capital taxes are taxed at the maximum rate of 100% forever. The economy does not converge to the steady state with positive consumption and capital. Both consumption and capital converge to zero. The reason for it is that as in Chamley, in Straub-Werning, there are only labor income taxes, no consumption taxes. Without consumption taxes, τ_{t+1}^{k} would have to satisfy

$$\frac{u_C(t)}{(1+\tau_t^c)} = \frac{\beta u_C(t+1)}{\left(1+\tau_{t+1}^c\right)}\left[1+\left(1-\tau_{t+1}^k\right)(U_{t+1}-\delta)\right]$$

and the restriction $\tau_{t+1}^{k} \leq 1$ could be binding. It turns out that, as shown by Straub-Werning, the restriction may be binding not only in period one, but in all future periods. The intuition is that in order to get $u_C(0)$ to be low, real rates must be low. Since the capital income tax is capped, it is not enough to use the real rate between periods zero and one. The whole term structure needs to be used.

The alternative solution, once consumption taxes are allowed for, is a relatively high consumption tax in period one. In the isoleastic case, consumption taxes could remain constant from period one on. Consumption and capital would converge to the steady state, and welfare would be considerably higher.

The result in Straub and Werning (2015) is an interesting example of an incomplete set of policy instruments. The consumption tax that could be redundant in other contexts, is crucial in their problem for optimal policy and welfare.

2.2 Why shouldn't capital be taxed in the long run?

Why shouldn't capital be taxed in the long run? Is it because capital is an intermediate good? If that was the reason then capital should not be taxed in the short term as well. In this section we interpret the result of the zero taxation of capital in the long run. It is based on an extension of the result of uniform taxation of Atkinson and Stiglitz (1972).

We start by considering a model with one good that cannot be taxed (leisure) and that can be transformed into multiple consumption goods with a linear technology.

The consumption goods can be taxed with differential consumption taxes, and with a common labor tax. This is the set up in Atkinson and Stiglitz (1972). If the utility function is separable in labor and homothetic in the consumption goods, then it is optimal to impose the same wedge between each consumption good and labor, meaning that there is no distortion across the different consumption goods.[1]

The second step is to consider a model with multiple labor types and one consumption good. Each of the types of labor can be transformed into the consumption good with a linear technology. The different types of labor can be taxed at a common rate by a common consumption tax or at different rates by differential labor taxes. Under separability and homotheticity in the types of labor, it is optimal to tax the different labor types at the same rate.

Finally, the model we are interested in is a model with multiple consumption goods and labor types in which each labor can be transformed into the different consumption goods using capital, and each consumption good can be transformed into the different labor types, also using capital. Suppose the utility function is separable between the consumption goods and the types of labor. If it is homothetic in the consumption goods, then it is optimal not to distort across consumption goods, imposing a common distortion between each labor and the different consumption goods. One implementation is to tax all the consumption goods at the same rate, not tax capital, and tax each labor (possibly) at a different rate. If the homotheticity is in the types of labor, then it is optimal not to distort across the different labor types, imposing the same distortion across each consumption good and the different types of labor. One implementation is with common labor taxes and a zero tax on capital. Consumption taxes could be different across goods. With homotheticity of both consumption goods and types of labor, then it is optimal to have both uniform consumption taxes and uniform labor taxes .

The reason why this model with separable and homothetic preferences is interesting is that in the steady state, it is as if preferences were separable and homothetic in consumption and labor. Thus in the steady state it is not optimal to impose distortions across consumption goods and across types of labor. It is not optimal to impose

[1]The result of uniform taxation of Atkinson and Stiglitz (1972) is a corollary of the result in Diamond and Mirrlees (1971) that intermediate goods should not be taxed. In fact, under the assumptions of separability and homotheticity, it is possible to write preferences as functions of a composite good that is a constant returns to scale function of the different consumption goods. Under those conditions, it is optimal to tax only the composite good.

intertemporal distortions, so that taxing capital is not optimal.

In a dynamic model with time separable preferences, the preferences that satisfy Atkinson and Stiglitz's conditions for uniform taxation, are additively separable between consumption and labor and have constant elasticity in both consumption and labor. One intuition for the result that capital should not be taxed is related to the standard principle that goods should be taxed according to their price elasticities. Because elasticities are constant over time, taxes should not distort across goods (and across types of labor), and therefore capital should not be taxed. In the steady state, those elasticities are constant, because consumption and labor are constant.

The initial confiscation of capital is another reason to tax capital in the long run, but we are abstracting from it here. Chamley (1986) and Judd (1985) had hoped that would have no long run effects, but Straub and Werning (2015) have shown their hope was unfounded. By adding consumption taxes, we recover the transitory nature of those effects.

We now formalize these arguments by reviewing the results on uniform taxation.

Uniform commodity taxation The model in Atkinson and Stiglitz (1972) has one good that cannot be taxed, leisure, and that can be transformed into multiple consumption goods with a linear technology. The consumption goods can be taxed with differential consumption taxes, and with a common labor tax. If the utility function is separable in labor and homothetic in the consumption goods, then it is optimal to impose the same wedge between each consumption good and labor. This means that uniform consumption taxation is optimal, that no distortion should be imposed across the different consumption goods. The analysis in Atkinson and Stiglitz is static. In a dynamic model without capital, labor today cannot be transformed into consumption tomorrow. It is not obvious what uniform taxation means there. But with capital, labor today can be transformed into consumption tomorrow. Uniform consumption taxation means that capital should not be taxed in the production of future consumption (provided consumption in different periods is taxed at the same rate).

We now review the result of uniform consumption taxation. The economy has a representative household with preferences over N consumption goods, C_t, $t = 1, ..., N$, and labor, n. It is separable in the consumption goods and labor, and homothetic[2] in

[2]Monotonic transformation of an homogeneous function.

the consumption goods, so that it can be written as

$$U\left(\mathcal{C}\left(C_1, C_2, ..., C_N\right), n\right) \tag{21}$$

where $\mathcal{C}$ is homogeneous of degree k. Each consumption good $t = 1, ..., N$, is produced with labor n, according to

$$C_t + g_t = A_t n_t,\ t = 1, ..., N. \tag{22}$$

with

$$\sum_{t=1}^{N} n_t = n. \tag{23}$$

g_t is government consumption of good t.

Notice that even if consumption is indexed by t, the model is static. How else could time today be used directly to produce consumption tomorrow, other then through capital accumulation?

In the first best, that maximizes (21) subject to (22) and (23), there are no wedges, so that

$$-\frac{U_C \mathcal{C}_{C_t}}{U_n} = \frac{1}{A_t},\ t = 1, ..., N,$$

and

$$\frac{\mathcal{C}_{C_t}}{\mathcal{C}_{C_{t+1}}} = \frac{A_{t+1}}{A_t},\ t = 1, ..., N-1.$$

Notice that the second condition can be interpreted as a condition of productive efficiency. It equates the marginal rates of technical substitution of C_t for C_{t+1} in the production of the composite $C = \mathcal{C}\left(C_1, C_2, ..., C_N\right)$ and in the production of labor.

The second best equilibrium Taxes are assumed to be on the consumption goods τ_t^C and on labor τ^n. There is also a levy l. Total time cannot be taxed, and neither can leisure. The household has budget constraint

$$\sum_{t=1}^{N} \left(1 + \tau_t^C\right) p_t C_t - \left(1 - \tau^n\right) wn = (1 - l) B \tag{24}$$

where p_t is the price of C_t in units of the composite good C. The wage rate w is also measured in units of C, and B are the liabilities of the government, also in units of C. The marginal conditions are

$$-\frac{U_n}{U_C \mathcal{C}_{C_t}} = \frac{(1-\tau^n)\, w}{(1+\tau_t^C)\, p_t},\ t = 1, ..., N. \tag{25}$$

One representative firm behaving competitively produces the consumption good C_t using labor n_t. Profits are $\Pi_t = p_t C_t - w n_t$. The marginal conditions are

$$\frac{w}{p_t} = A_t,\ t = 1, ..., N.$$

It follows that

$$-\frac{U_n}{U_C \mathcal{C}_{C_t}} = \frac{(1-\tau^n)}{(1+\tau_t^C)} A_t,\ t = 1, ..., N.$$

A common consumption tax distorts the margin between consumption and labor, while differential consumption taxes distort across consumption goods.

In order to solve the Ramsey problem, it is useful to write the budget constraint, (24), using the marginal conditions, (25) and $-\frac{U_n}{U_C} = (1-\tau^n)\, w$, as

$$\sum_{t=1}^{N} U_C \mathcal{C}_{C_t} C_t + U_n n = U_C\,(1-l)\, B \tag{26}$$

We impose

$$U_C\,(1-l)\, B \geq \mathcal{W}_0$$

At the optimum it holds with equality, so

$$\sum_{t=1}^{N} U_C \mathcal{C}_{C_t} C_t + U_n n = \mathcal{W}_0 \tag{27}$$

This implementability condition together with the resource constraints are the only conditions restricting the equilibrium variables C_t, $t = 1, ...N$, and n.

The Ramsey problem is to maximize (21) subject to (27) and the resource constraints (22) and (23). The first order conditions are, for $j = 1, ..., N$,

$$U_C \mathcal{C}_{C_j} + \phi \left[U_{CC} \mathcal{C}_{C_j} \sum_{t=1}^{N} \mathcal{C}_{C_t} C_t + U_C \sum_{t=1}^{N} \mathcal{C}_{C_t C_j} C_t + U_C \mathcal{C}_{C_j} + U_{nC} \mathcal{C}_{C_j} n \right] = \lambda_j \tag{28}$$

and

$$-U_n + \phi \left[-U_{Cn} \sum_{t=1}^{N} \mathcal{C}_{C_t} C_t + U_{nn} n + U_n \right] = \lambda_j A_j. \tag{29}$$

The fact that $\mathcal{C}$ is homogeneous of degree k implies that $\sum_{t=1}^{N} \mathcal{C}_{C_t} C_t = kC$ and $\sum_{t=1}^{N} \mathcal{C}_{C_t C_j} C_t = (k-1) \mathcal{C}_{C_j}$. The marginal conditions of the Ramsey problem can then be written as

$$U_C \mathcal{C}_{C_j} + \phi \left[U_{CC} \mathcal{C}_{C_j} kC + U_C (k-1) \mathcal{C}_{C_j} + U_C \mathcal{C}_{C_j} + U_{nC} \mathcal{C}_{C_j} n \right] = \lambda_j \tag{30}$$

$$-U_n + \phi \left[-U_{Cn} kC + U_{nn} n + U_n \right] = \lambda_j A_j \tag{31}$$

so that

$$\frac{U_C \mathcal{C}_{C_j}}{-U_n} = \frac{1 + \phi \left[k \frac{U_{Cn} C}{U_n} - \frac{U_{nn} n}{U_n} - 1 \right]}{1 + \phi \left[\frac{k U_{CC} C}{U_C} + k + \frac{U_{nC}}{U_C} n \right]} \frac{1}{A_j} \tag{32}$$

It follows that it is optimal to tax all the consumption goods at the same rate, meaning that there are no distortions across consumption goods, as in the first best.

Uniform labor taxation In a dynamic model, there are multiple consumption goods and multiple types of labor. We can consider a static model with multiple labor types and one consumption good. Each labor can be used to produce the consumption good with a linear technology. The different labor types can be taxed at a common rate by a common consumption tax or at different rates by differential labor taxes. With separability and homotheticity in the types of labor, it is optimal to tax the different labor types at the same rate, meaning that uniform labor taxation is optimal. Again, capital is the way consumption today can be transformed into labor tomorrow. With separability and homotheticity in labor types, capital should not be taxed in the production of future labor. All labors should be taxed at the same rate and capital should not be taxed.

Suppose the economy now has a single consumption good that can be produced with different types of labor. Preferences are

$$U\left(C, \mathcal{H}\left(n_1, n_2, ..., n_N\right)\right)$$

with $H = \mathcal{H}\left(n_1, n_2, ..., n_N\right)$. The utility function is separable in the consumption good and the labor types and homothetic in the labor types. $\mathcal{H}$ is constant returns to scale. The consumption good can be produced with labor n_t, according to

$$c_t = A_t n_t,\ t = 1, ..., N,$$

with

$$C = \sum_{t=1}^{N} c_t.$$

In the first best, the marginal conditions are

$$-\frac{U_C}{U_H \mathcal{H}_{n_t}} = \frac{1}{A_t},\ t = 1, ..., N,$$

and

$$\frac{\mathcal{H}_{n_t}}{\mathcal{H}_{n_{t+1}}} = \frac{A_t}{A_{t+1}},\ t = 1, ..., N-1.$$

The second condition can be interpreted as a condition of productive efficiency. It equates the marginal rates of technical substitution of n_t for n_{t+1} in the production of the composite $\mathcal{H}\left(n_1, n_2, ..., n_N\right)$ and in the production of consumption.

Second best equilibrium Taxes are assumed to be on the consumption good τ^C and on labor types τ_t^n. There is also a levy l. The household budget constraint is

$$\left(1+\tau^C\right) C - \sum_{t=1}^{N} \left(1-\tau_t^n\right) w_t n_t = (1-l)\, B \tag{33}$$

The marginal conditions are

$$-\frac{U_H \mathcal{H}_{n_t}}{U_C} = \frac{\left(1-\tau_t^n\right) w_t}{1+\tau^C},\ t = 1, ..., N. \tag{34}$$

One representative firm behaving competitively produces c_t using n_t. It follows that

$$w_t = A_t,\ t = 1, ..., N.$$

and

$$-\frac{U_H \mathcal{H}_{n_t}}{U_C} = \frac{(1-\tau_t^n)}{(1+\tau^C)} A_t,\ t = 1, ..., N.$$

In order to solve the Ramsey problem, it is useful to write the budget constraint, (24), using the marginal conditions, (25) and $-\frac{U_n}{U_C} = (1-\tau^n)\, w$, as

$$U_C C + \sum_{t=1}^{N} U_H \mathcal{H}_{n_t} n_t = U_C \frac{(1-l)\, B}{1+\tau^C} \tag{35}$$

We impose that

$$U_C \frac{(1-l)\, B}{1+\tau^C} \geq \mathcal{W}_0$$

At the optimum it holds with equality, so

$$U_C C + \sum_{t=1}^{N} U_H \mathcal{H}_{n_t} n_t = \mathcal{W}_0 \tag{36}$$

This implementability condition together with the resource constraints are the only conditions restricting the equilibrium variables n_t, $t = 1, ...N$, and C. The first order conditions of the Ramsey problem are for $j = 1, ..., N$,

$$U_C + \phi \left[U_{CC} \mathcal{C} + U_C + U_{HC} \sum_{t=1}^{N} \mathcal{H}_{n_t} n_t \right] = \lambda_j \tag{37}$$

$$-U_H \mathcal{H}_{n_j} + \phi \left[U_{CH} \mathcal{H}_{n_j} C + U_{HH} \mathcal{H}_{n_j} \sum_{t=1}^{N} \mathcal{H}_{n_t} n_t + U_H \sum_{t=1}^{N} \mathcal{H}_{n_t n j} n_t + U_H \mathcal{H}_{n_j} \right] = \lambda_j A_j \tag{38}$$

The fact that $\mathcal{H}$ is CRS implies that $\sum_{t=1}^{N} \mathcal{H}_{n_t} n_t = kH$ and $\sum_{t=1}^{N} \mathcal{H}_{n_t n j} n_t = (k-1)\, \mathcal{H}_{nj}$. This means that the marginal conditions of the Ramsey problem imply

$$\frac{U_C}{-U_H \mathcal{H}_{n_j}} = \frac{1 + \phi \left[-\frac{U_{CH} C}{U_H} - \frac{k U_{HH} H}{U_H} - k \right]}{1 - \phi \left[-\frac{U_{CC} \mathcal{C}}{U_C} - 1 - \frac{k U_{HC} H}{U_C} \right]} \frac{1}{A_j} \tag{39}$$

It is optimal to tax all the labor types at the same rate, meaning that there are no distortions across different types of labor, as in the first best.

The dynamic model with capital We are not exactly interested in the static models above with multiple consumption goods or labor types, but rather in a dynamic model with capital, multiple consumption goods and multiple labor types. As it turns out, the results above can be used to explain under what conditions capital should not be taxed in the dynamic model. Capital is the way consumption today can be transformed into consumption tomorrow, and labor today into labor tomorrow. Under the conditions of separability and homotheticity of both consumption goods and labor types, it is optimal not to distort the intertemporal margins. Under those conditions, capital should not be taxed.

In the dynamic model, time additive preferences that are separable and homothetic are additive and constant elasticity,

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

In the dynamic model, with constant elasticity in consumption, the optimal policy has uniform taxation in consumption with zero capital taxation and variable labor taxes. No distortion is imposed in the intertemporal consumption margin. Instead, distortions are imposed in the intertemporal labor margin. If, instead, the labor elasticity is constant, then optimal policy has uniform labor taxation with zero capital taxation and variable consumption taxes. Again, no distortion is imposed in the intertemporal labor margin, but the distortion in the intertemporal consumption margin.

The optimal tax rule is to tax less the goods with more elastic demands. In the steady state, elasticities are constant, and therefore all consumption goods and all labor types ought to be taxed at the same rate. And capital income should not be taxed. Away from the steady state, in general it is optimal to tax the different goods at different rates, or to tax (or subsidize) capital.

Along the transition the question of whether capital should be taxed is a quantitative question. Quantitatively it may still be approximately optimal not to tax capital.

3 Distribution

The results obtained above for the representative economy remain under certain conditions in economies with capital-rich and poor agents. In order to show this consider that there are two agents, 1 and 2. The social welfare function is

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

where θ is some weight. The implementability conditions can be written as

$$\sum_{t=0}^{\infty} \beta^t \left[u_C^1(t) C_t^1 + u_N^1(t) N_t^1 \right] = \mathcal{W}_0^1, \tag{40}$$

and

$$\sum_{t=0}^{\infty} \beta^t \left[u_C^2(t) C_t^2 + u_N^2(t) N_t^2 \right] = \mathcal{W}_0^2, \tag{41}$$

as well as

$$\frac{u_C^1(t)}{u_C^2(t)} = \frac{u_N^1(t)}{u_N^2(t)}$$

$$\frac{u_C^1(t)}{u_C^2(t)} = \frac{u_C^1(t+1)}{u_C^2(t+1)}$$

that impose that tax rates must be the same for the two agents. These last conditions can be written as

$$u_C^1(t) = \gamma u_C^2(t)$$

$$u_N^1(t) = \gamma u_N^2(t)$$

where γ is some endogenous constant.[3] Finally the resource constraints are

$$C_t^1 + C_t^2 + G_t + K_{t+1} - (1-\delta) K_t \leq A_t F\left(N_t^1 + N_t^2, K_t\right).$$

We can derive first order conditions, assuming separability. Let φ^1 and φ^2 be the multipliers of the two implementability conditions, (40) and (41). Then the marginal

[3] See also Greulichy, Laczó and Marcet (2016).

conditions imply

$$u_C^2(t)\,\frac{\gamma\left[\theta+\varphi^1\left(1-\sigma_t^1\right)\right]\frac{\sigma_t^2}{C_t^2}+\left[(1-\theta)+\varphi^2\left(1-\sigma_t^2\right)\right]\frac{\sigma_t^1}{C_t^1}}{\frac{\sigma_t^2}{C_t^2}+\frac{\sigma_t^1}{C_t^1}}=\lambda_t$$

and

$$-\lambda_t+\beta\lambda_{t+1}\left[f_k(t+1)+1-\delta\right]=0$$

For the isoelastic case, it is still optimal not to distort the intertemporal margins., so that capital should not be taxed.

Once there is no room to affect the initial confiscation, and without transfers across agents, there is not much distribution that can be done. We now allow for differential transfers across agents, but restrict them to be positive. The implementability conditions (40) and (41) now become

$$\sum_{t=0}^{\infty}\beta^t\left[u_C^1(t)\,C_t^1+u_N^1(t)\,N_t^1\right]=\mathcal{W}_0^1+Tr^1.$$

$$\sum_{t=0}^{\infty}\beta^t\left[u_C^2(t)\,C_t^2+u_N^2(t)\,N_t^2\right]=\mathcal{W}_0^2+Tr^2.$$

Transfers are restricted to be positive. There is a multiplier associated with each of those constraints (ϕ^1 and ϕ^2). Now, if the multipliers of the implementability constraints of the two agents, (40) and (41), are positive, the planner would want to make negative transfers to both agents, and ϕ^1 and ϕ^2 would both be nonzero, and equal to φ^1 and φ^2. If the multiplier of one of the implementability conditions, (40) or (41), is negative (because the planner would want to make positive transfers to that agent), then transfers will be made to the point where the multiplier of the implementability is equal to the multiplier of the nonnegativity constraint, zero.

The marginal condition of the Ramsey problem becomes

$$u_C^2(t)\,\frac{\gamma\left[\theta+\varphi^1\left(1-\sigma_t^1\right)\right]\frac{\sigma_t^2}{C_t^2}+(1-\theta)\frac{\sigma_t^1}{C_t^1}}{\frac{\sigma_t^2}{C_t^2}+\frac{\sigma_t^1}{C_t^1}}=\lambda_t$$

Again, with isoelastic preferences taxing capital is not optimal.

In similar structures, Straub and Werning (2015) obtain that depending on the price elasticity parameter it may be optimal to fully tax capital. The reason for this result is again the initial confiscation. Once the initial confiscation is ruled out, the only reason to tax capital is that elasticities may be varying over time.

4 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. Capital should not be taxed because its tax is imposing a higher tax on future goods. We relate the results to the principles of optimal taxation of Atkinson and Stiglitz (1972) and Diamond and Mirrlees (1971).

The confiscation of the initial capital could be a reason to tax future capital, temporarily if consumption taxes are used. If consumption taxes are excluded capital could be fully taxed forever as in the examples of Straub and Werning (2015). In this sense, Straub and Werning is an extreme example of an incomplete set of fiscal instruments. The example is extreme in its implications for policy and welfare. Ruling out the use of future consumption taxes, implies that optimal policy may be fully taxing capital forever, resulting in asymptotic zero consumption and capital. With a moderate consumption tax in the second period of life, the initial confiscation is taken care of, and capital is not taxed in the steady state.

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

[6] 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.

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

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

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

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

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

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

[13] 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.

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