CEP 13-09

Capital Taxation, Intermediate Goods, and Production Efficiency

Till O. Gross

October 2013

CARLETON ECONOMIC PAPERS

Carleton UNIVERSITY

Department of Economics

1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B6

Capital Taxation, Intermediate Goods, and Production Efficiency

Till Gross *

October 23, 2013

Abstract

An important controversy in public finance is whether long-run capital taxes are optimally zero or not, with a broad variety of models supporting each case. This paper examines the question whether capital is special and if so, what the underlying principle could be that explains both types of results. I find that capital is provided without distortions in a wide class of models, i.e. that its marginal product is the same in first and second best. The conditions for this to hold are that the government is able to tax all of capital's co-factors of production separately and that capital does not enter the utility function. When individually rational behavior leads to sub-optimal capital accumulation, then capital taxes are used to implement the optimal allocation. The intuition is that capital is an intermediate good; optimal taxation seeks to tax endowments and intermediate goods do not have any endowment component.

1 Introduction

Capital taxation is an important policy issue: in the OECD countries, about 1/3 of GDP is capital income and corporate taxes generate 3-3.5% of GDP in revenues, contributing roughly 10% to total tax revenues. However, there is intense debate about how capital should be taxed – if at all.

Chamley (1986) and Judd (1985) have shown how capital taxes with infinitely lived agents and perfect commitment are optimally zero in the long run. Kocherlakota (2010) points out that this result "is startlingly robust across different formulations of preferences and technology," leading

*Department of Economics, Carleton University. I would like to thank Richard Brecher for many valuable discussions, comments, and suggestions. Also thanks to participants at a seminar at the Université du Québec à Montréal and at the Bank of Canada, as well as participants at the 2012 conference of the Société canadienne de science économique, the spring 2013 Midwest Macro Meetings, and the 2013 conference of the Association of Public Economic Theory. I would like to thank Marek Kapicka and Julian Neira for their comments, too. All errors are of course my own. Correspondence to: `till_gross@carleton.ca`. All comments and suggestions welcome!

authors such as Atkeson, Chari, and Kehoe (1999) and Mankiw, Weinzierl, and Yagan (2009) to argue that it should be used as a guideline for policy. On the other hand, "we see that the Ramsey approach is disturbingly non-robust," with abundant research showing non-zero optimal capital taxes, so Conesa, Kitao, and Krueger (2009) and Diamond and Saez (2011), for example, call for positive capital taxes.

Jones, Manuelli, and Rossi (1997) claim that "there is nothing special about capital income" and should be taxed like other input factors, whereas Ljungqvist and Sargent (2004) retort that physical capital is special. In this paper, I examine the question whether capital taxes are special and if so, what the underlying principle could be that explains both the zero-capital tax results and the various exceptions.

I find that capital is indeed special, but not in the sense that it should be taxed at a zero rate; rather, it should be provided efficiently (without distortions), as it is an intermediate good. As known from Diamond and Mirrlees (1971) for the static case, it is optimal to be on the production possibility frontier (which they call production efficiency). It follows that intermediate goods should be provided without distortions from the government's perspective. These results do not carry over to a dynamic framework, though. As I explain in more detail in the next section, if capital is the only means of transferring resources intertemporally, then being on the production possibility frontier has no implications for the allocation of capital.

Another concept is thus needed to talk about the production efficiency of intermediate goods. If an input is undistorted in the second best – I define it as the marginal product being the same as in the first best, when lump-sum taxes are available – then I say that it satisfies production efficiency. I provide a generalized proof that the marginal product of capital will be undistorted in the long-run if it does not enter the households' or the government's objective function and all of capital's co-factors of production can be taxed separately.

The intuition of why intermediate goods should not be taxed can be summarized as follows: Taxation of endowments does not generate distortions; since one of the goals of optimal taxation is to minimize distortions, tax authorities aim to tax endowments. Since it is often not feasible to tax endowments directly, it has to be done indirectly. For example, a tax on labor income indirectly taxes the time endowment. For intermediate goods, there is no endowment component, so it is not optimal to distort them if it is possible to generate tax revenues from other sources.

What distinguishes capital from other intermediate goods and often leads to non-zero capital taxes is that in most models it is the only means of transferring resources from one time period or state of nature to another. A standard intermediate input is only used by firms in production processes, so unless there is an externality firms' choices are socially optimal and no taxes are needed to correct any misallocation. On the other hand, agents can make capital accumulation decisions which are optimal from their individual perspective (e.g. to self-insure or smooth consumption over the life-cycle) but not from an aggregate point of view. Capital taxes are thus used to implement the efficient allocation of capital and not to raise revenues or redistribute resources. I analyze a wide variety of published papers on optimal capital taxation and discuss the role of capital taxes in them.

There are two other main contenders as possible principles of capital taxation which I will compare to in detail in section 5. The most common explanation for the zero capital tax result is that a constant tax on capital is equivalent to an ever-increasing tax on consumption in the future. Golosov and Tsyvinski (2008) for example describe how this violates uniform commodity taxation, as established in Atkinson and Stiglitz (1972, 1976). However, the zero-capital tax holds in many cases where the assumptions of uniform commodity taxation are not met and vice versa.

In a recent publication, Albanesi and Armenter (2012) present a general framework of intertemporal distortions and argue that the frontloading of taxes is essential, i.e. that if it is possible (it does not have to be optimal) to have all distortions covered during some finite time, then there will be no intertemporal distortions in the long run. They posit that there is something special about the intertemporal margin. I view this paper as complementary to their work in understanding capital taxation, as they focus on the absence of distortions from the individual's perspective as opposed to the government's. Albanesi and Armenter (2012) do not explain why the household's intertemporal margin is distorted in various models, most notably with overlapping generations or with an incomplete tax system.

To the best of my knowledge, there has not been a thorough investigation of capital as an intermediate good, although the idea is not new: Judd (1999) provides both the explanation of an explosive commodity tax and taxation of intermediate goods, but does not pursue it further. Kocherlakota (2010) points out the analogy between money and capital and argues that money ought not to be taxed as it is an intermediate good, but his explanation for the zero capital-tax is the equivalence to an ever-increasing consumption tax.[1]

[1]Correia (1996a) briefly mentions production efficiency as the intuition behind her results, referring to Munk

The contribution of this paper is thus twofold: (i) I develop an abstract framework and show that capital is provided without distortions in a very large class of models and under which conditions (not in utility function, co-factors of production are taxable); (ii) I explain how the concept of capital as an intermediate good can accommodate the various different results found in the literature on optimal capital taxation. It thus proposes a unified principle of capital taxation (no distortion), which can be used to inform the policy debate on how to implement specific capital taxes in practice.

In the following section, I present the proof that intertemporal intermediate goods provision is optimally undistorted. In section 3 I show how this can be applied to a variety of models from the optimal taxation literature. In section 4 I discuss other models which do not meet the assumptions laid out in section 2, including the taxation of the initial capital stock. In section 5 I discuss alternative explanations for the zero capital tax result. The final section concludes.

2 Production Efficiency of Intermediate Goods

In this section, I provide a proof of the production efficiency of intertemporal intermediate goods. I first spell out the problem of why Diamond and Mirrlees (1971) does not apply to capital. Then I provide definitions and set up the model. The framework I discuss here is general and abstract, so I briefly illustrate how one can map Chamley-Judd into it in the text; more detailed examples are in appendix A. Finally, I show a lemma and the proof.

2.1 Capital and the Production Possibility Frontier

The production efficiency theorem states that the allocation should optimally be on the production frontier; as a corollary, intermediate goods should be provided efficiently, which means in the model of Diamond and Mirrlees (1971) untaxed. Taxing an intermediate input creates a distortion in both production and the consumption of the final good, so a tax on the final good alone, which generates the same revenues, will cause less distortions. Therefore, it is always more efficient to raise tax revenues from introducing distortions on final goods only. Not distorting the allocation of

(1980), which is about taxing firms at different rates for their inputs. However, there are two parts to production efficiency: that all firms should be taxed at the same rate for their inputs and that intermediate goods should be provided efficiently. As Diamond and Saez (2011) point out, there is no heterogeneity among producers in the standard model by Chamley and Judd. One can interpret the producers of consumption in different periods as different firms, but in steady state they are all taxed at the same rate. It thus does not appear to be production efficiency in the sense of equally taxing firms that is driving the results of optimal capital taxation.

intermediate goods is a necessary condition for being on the production possibility frontier.

However, their analysis only considers the case of an intermediate good that is produced from the same set of factors of production as the final good for which it is used as an input.[2] What happens if an intermediate good is the only means of transferring resources from one set of factors to another?

Capital is such a case, as it is an output in period t produced by inputs in that period (such as labor at t), but used as an input in period $t+1$ (together with labor at $t+1$). Labor inputs at periods t and $t+1$ are obviously different and resources can be shifted from one period to the other only through capital, so it is not clear if and how the theorem of production efficiency of intermediate goods applies. The concept of being on the production possibility frontier cannot be put to use (as Diamond and Saez (2011) point out in footnote 15), since changing the capital stock invariably shifts resources from one time period to another. Using more or less capital moves the economy along the production possibility frontier, but not inside the production possibility set. There is no other common factor of production that can be adjusted to achieve an unequivocally higher production of final goods in all periods.

A different concept than being on the production frontier is needed for capital. What I propose is whether the allocation of a factor is distorted or not. When a factor allocation is distorted it is in order to raise revenues, whereas it is undistorted so as to maximize production, irrespective of government revenue requirements. I thus define that a factor allocation satisfies production efficiency if it is undistorted. It is easy to define how an intermediate good which shares some common inputs with the final good is undistorted. The common factor (for example labor) can either be used directly as an input for the final good or indirectly by producing the intermediate good first and then using this to produce the final good. The marginal product of the common factor has to be the same producing the final good directly or indirectly, otherwise one could produce more output with less inputs by substituting towards the activity with the higher marginal product.

For a good like capital, this rule does not apply, as there are no common factors. However, the first-best allocation is obviously undistorted. Therefore, if the marginal product of an input is the same in first and second best, it has to be undistorted, too. This is a stronger requirement than the one discussed in the previous paragraph.

[2] Also see Acemoglu, Golosov, and Tsyvinski (2008), who study both classic intermediate goods as in Diamond and Mirrlees (1971) and capital.

I would like to emphasize that undistorted does not imply untaxed. It is obvious that if an intermediate good has a negative externality, then taxes are necessary to make the individually rational behavior consistent with aggregate optimality. Similarly, individually optimal capital accumulation decisions are often not socially optimal, requiring a corrective tax. I discuss this in detail in section 3.

Definition 1 *An intermediate good is defined as any commodity which (i) does not enter any of the households' utility functions nor the government's objective function, (ii) is only available as an output of one or more goods, and (iii) is used as an input in the production of at least one good.*

Definition 2 *The allocation of an input satisfies production efficiency if it is undistorted. A sufficient condition for an undistorted allocation is if a factor's marginal product is the same as in the first best.*

Why not use the notion that a factor is undistorted when the marginal rate of substitution is equal to the marginal rate of transformation? Multiple problems arise. When households are heterogeneous (and face borrowing constraints, for instance), should one evaluate the marginal rate of intertemporal substitution of each household or the aggregate (and if so, how does one weight each individual in the aggregate)? If there are externalities in production or consumption, how do they factor in? What if the government is not (only) maximizing the utility of households? How would one define the government's marginal rate of substitution? The above definition of "undistorted" is independent of these issues and can therefore be applied more broadly.

2.2 Allocation of Intermediate Goods

The government is maximizing its objective function

$$V(X_{-K}), \tag{1}$$

where X_{-K} is a vector of allocations not including intermediate good K.[3] Optimization is subject to the government budget constraint(s). It can perfectly commit to its policy.[4] I assume that the

[3]The vector of allocations X consists for each product j of the individual consumption of each household $i \in I$ $\{c_{ij}\}_I$, of government consumption g_j, and inputs $\{x_{j,\hat{j}}\}_{M_j}$. I specify these terms below.

[4]In section 4.2 I discuss the case of imperfect commitment.

government's objective function is such that it allows for an interior solution for the intermediate good K and that resources are always valued.

The government can issue bonds and set taxes on at least a subset of goods. As in Diamond and Mirrlees (1971), I distinguish between producer prices p (before tax) and consumer prices $\tilde{p}$ (after tax). When the government is not able to freely choose after-tax prices through appropriate taxes, a set of constraints is imposed on these prices. Let p_1 and $\tilde{p}_1$ be the unconstrained prices and p_2 and $\tilde{p}_2$ the constrained ones, so that

$$\tau(p_2, \tilde{p}_2) \geq 0. \tag{2}$$

Example: The objective function $V(X_{-K})$ in Chamley-Judd is the representative household's discounted lifetime utility. Prices p are the wage w and interest rate r. The government has to finance exogenous expenditures via bonds and taxes on w and r. It can thus freely choose after-tax prices $\tilde{p}$, usually with the exception of capital taxes at time zero; for instance, one could assume that they are zero so that $\tau(p_2, \tilde{p}_2) \geq 0$ is given by $r_0 = \tilde{r}_0$.

Assumption 1 (Government Bonds) *The government can issue non-productive bonds B, which are a perfect substitute for an intertemporal intermediate good K from the household's perspective.*

Assumption 2 (Tax System) *All co-factors of production of intermediate good K may be taxed separately by the government.*[5]

Households aim to maximize their utility, which is also independent of intermediate good K, subject to budget constraint(s). I assume that households also always value resources, so that the household budget constraint is always binding. Firms maximize profits. Both households and firms act competitively, taking prices and taxes as given. They observe the government's policy and react subsequently. Instead of choosing its tax and fiscal policy instruments and evaluating policy through the private sector's reaction functions, the government can directly choose allocations in the economy if it takes into account households' and firms' optimality conditions. These constraints depend on the informational restrictions and/or tax instruments at the government's disposal.

[5]This does not rule out home-production, for instance: the goods produced at home are not taxable and its prices are captured in p_2. As long as home production does not use the same capital as the market, it does not affect results. Unobserved effort and ability are also unproblematic, as long as the government can tax effective labor, which might depend somehow on effort, ability, and hours; when these factors independently affect the marginal product of capital, however, then the assumption is violated.

Each household of type $i \in I$, I being the set of types of households, cares only for its total asset position $a = k + b$, which is equal to its holdings of the intermediate good plus the government substitute. The lower-case letters stand for individual holdings, whereas the upper-case letter A stands for the total distribution. Household decisions are only based on after-tax prices. The households' optimality conditions can be captured in the constraint set

$$\Omega_H(X_{-K}, A, \tilde{p}) \geq 0. \tag{3}$$

Example: Optimal household behavior in Chamley-Judd requires the budget constraint, the Euler equation, the labor-leisure trade-off, and the transversality condition to hold.

Competitive firms produce output of good $j \in J$, with J being the set of all products, according to a continuously differentiable production function $F_j(X)$ with constant returns to scale. The production function satisfies the Inada conditions for all productive inputs. Producer prices are then some function of the allocation X

$$p - f(X) = 0. \tag{4}$$

Example: Profit-maximization in Chamley-Judd implies that wages and the interest rate are equal to the marginal products of labor and capital, respectively.

Finally, there is a resource constraint for each product $j \in J$ (one could also call it market clearing for each good):

$$F_j(X) = C_j + g_j + \sum_{\hat{j} \in M_j} x_{j,\hat{j}}. \tag{5}$$

Output $F_j(X)$, as a function of the vector of allocations X, has to equal total consumption of the product: C_j is the aggregate consumption of all households of product j, g_j is government consumption of product j, and $\sum_{\hat{j} \in M_j} x_{j,\hat{j}}$ is the sum of all inputs out of product j, with M_j as the set of all inputs produced from good j.[6] Each product can potentially be used as multiple inputs and each input can be used for multiple products.[7] When the resource constraints for all products and the budget constraints of all households are satisfied, then the government budget constraints hold by Walras' Law and can thus be dropped from the control problem. Denote by $F(X) = 0$ the

[6] It is also implicitly assumed that all goods are valuable and that free disposal of goods is never used.

[7] This formulation allows for a stochastic production function, where production in each state is considered a different good.

set of resource constraints for all goods. *Example: The resource constraint in Chamley-Judd simply states that total output in each period (each good j corresponds to one time-period t) has to equal the sum of private consumption, government expenditures, and capital investment.*

The government's problem can thus be stated as:

$$\begin{aligned} &\max V(X_{-K}) \\ &\text{s.t.}\, \Omega_H(X_{-K}, A, \tilde{p}) \geq 0 \\ &\quad \tau(p_2, \tilde{p}_2) \geq 0 \\ &\quad p - f(X) = 0 \\ &\quad F(X) = 0. \end{aligned} \tag{6}$$

The resource constraint is always binding under the assumption of non-satiation and possible transfers from the government to agents; let θ_j be the government's Lagrange multiplier for good j. For all producer prices in p_1, the Lagrange multiplier for $p - f(X) = 0$ is zero, as these prices do not appear anywhere else in the problem. Producer prices in p_2, however, appear both in the set of tax constraints $\tau(\cdot)$ and in the set of firm optimality conditions. But by assumption 2, these prices are not affected by input K. The binding constraint can thus be written as $p_2 - f(X_{-K}) = 0$. Since K is an intermediate good and not part of the objective function or a household's utility function, the first-order condition with respect to K will therefore only involve the resource constraints. Lemma 1 follows:

Lemma 1 *The allocation of an intermediate good, for which all co-factors of production can be taxed independently, is determined by:*

$$\sum_{j \in J} \theta_j \frac{\partial F_j(X)}{\partial x_{j,\hat{j}}} = \theta_{\hat{j}}, \tag{7}$$

Note that while the condition itself is completely independent of the government objective function or any other constraints, the values of the resource Lagrange multipliers typically depend on other constraints. Lemma 1 thus implies that production efficiency for the intermediate good holds whenever the Lagrange multiplier ratios $\theta_j/\theta_{\hat{j}}$ are the same in first and second best, which is generally difficult to assess.

2.3 Conditions for Production Efficiency

In this section I show three relevant cases in which one can give a definitive answer to that problem: Capital satisfies production efficiency in any steady state, on average in any stationary equilibrium, or on average over an infinitely long horizon, given that all factors of production are taxed independently.

I refer to capital as an intermediate input, which is produced in time period t and used as an input in period $t+1$. Assume that the government discounts the future at a constant rate β. If all factors of production can be taxed separately and there is no aggregate uncertainty, then equation 7 from Lemma 1 can be written as:

$$\theta_t = \beta\theta_{t+1} F_K(t+1). \tag{8}$$

The term $F_K(t+1)$ refers to the marginal product of capital at time $t+1$. It should be noted that government bonds play an important role: households' asset holdings are then separate from capital used in production. I will further discuss this assumption in specific cases below.

In a steady state, the Lagrange multiplier for resources at any two periods is the same, so $\theta_t = \theta_{t+1}$, from which it follows that $1 = \beta F_K$.[8] This rule is the same in first and second best (the ratio of θ_{t+1}/θ_t is equal to one in either case), therefore capital satisfies production efficiency in any steady state:

Proposition 1 *If capital is an intermediate good for which all co-factors of production can be taxed independently, then it will be provided according to production efficiency in any steady state.*

This can be generalized to the case of aggregate uncertainty, where production efficiency holds on average. Let S be the set of states, $\mu(s^t)$ be the probability of experiencing a history of states s^t, and $\mu(s_{t+1}|s_t)$ be the time-invariant transition probability from state s_t to state s_{t+1}. Assume that the government maximizes expected payoffs. Equation 7 then becomes:

$$\theta(s^t) = \beta \sum_S \theta(s^{t+1})\mu(s^{t+1}|s^t)F_K(s^{t+1}) \tag{9}$$

$$\Leftrightarrow \theta(s^t)\mu(s^t) = \beta \sum_S \theta(s_{t+1}, s^t)\mu(s_{t+1}|s^t)F_K(s_{t+1}, s^t). \tag{10}$$

[8] This extends easily to a balanced growth path.

In a stationary equilibrium the expected value of resources at time t and $t+1$ is equal:[9]

$$\sum_{S^t} \theta(s^t)\mu(s^t) = \sum_{S^{t+1}} \theta(s^{t+1})\mu(s^{t+1}) \tag{11}$$

$$\Leftrightarrow \sum_{S^t} \theta(s^t)\mu(s^t) = \sum_{S^t}\sum_{S} \theta(s_{t+1}, s^t)\mu(s_{t+1}|s^t). \tag{12}$$

Rewriting equation 10 by summing over all histories s^t, one obtains

$$\sum_{S^t} \theta(s^t)\mu(s^t) = \beta \sum_{S^t}\sum_{S} \theta(s_{t+1}, s^t)\mu(s_{t+1}|s^t)F_K(s_{t+1}, s^t), \tag{13}$$

and combining with equation 12 this yields

$$\sum_{S^t}\sum_{S} \theta(s_{t+1}, s^t)\mu(s_{t+1}|s^t)(F_K(s_{t+1}, s^t) - 1/\beta) = 0. \tag{14}$$

Production efficiency thus holds on average, as the above equation is valid both for first and second best. Only the weights vary potentially. The next proposition formalizes the result:

Proposition 2 *If capital is an intermediate good for which all co-factors of production can be taxed independently, and the government maximizes expected payoffs, then capital will on average be provided according to production efficiency in any stationary equilibrium.*

Finally, if one assumes that the value of government funds is strictly positive and bounded above, then its average growth rate $g_t = \theta_{t+1}/\theta_t - 1$ tends to zero as the horizon goes to infinity, and therefore the average distortion, too:

$$\lim_{T\to\infty} \frac{1}{T}\sum_{i=t}^{T} g_i = \lim_{T\to\infty} \frac{1}{T}\sum_{i=t}^{T} \frac{1/\beta - F_K(i)}{F_K(i)} = 0. \tag{15}$$

Proposition 3 *In the deterministic case, if capital is an intermediate good for which all co-factors of production can be taxed independently, then capital will on average be provided according to production efficiency over an infinitely long horizon.*

Why is capital provided efficiently only in the long run? First, in the short run the initial capital

[9] This is equivalent to the steady-state notion of equal Lagrange multipliers $\theta_t = \theta_{t+1}$, which is $E_0\theta_t = E_0\theta_{t+1}$ in a stationary equilibrium.

stock is fixed, so that initial capital taxes are lump-sum taxes (due to not modeling expectations before time zero or agents reacting to being expropriated). I discuss this further in section 4.2. Second, if resources are more valuable to the government in one period than another, then capital taxes can be used to shift resources to the period with a high value. In the long run, periods of high and low values of resources cancel out on average (if it kept decreasing or increasing, then it would go to zero or infinity).

3 Capital Taxes and Production Efficiency

In this section, I examine how the above principle of production efficiency of intermediate goods applies to models in the capital taxation literature: infinite-dynasties as in Chamley (1986) versus overlapping generations as in Erosa and Gervais (2002) and idiosyncratic productivity shocks as in Aiyagari (1995) or Golosov, Kocherlakota, and Tsyvinski (2003). I explain how individually optimal behavior may lead to non-optimal capital accumulation in the aggregate and capital taxes are thus needed to implement the optimal allocation. The appendix shows how these papers can be mapped into the present framework (along with two others on unemployment).

3.1 Infinitely-lived Dynasties vs. Overlapping Generations

In the baseline setup in Chamley (1986), there is one representative infinitely-lived household, markets are perfectly competitive, and the benevolent government needs to finance a stream of exogenously given expenditures through proportional taxes on capital and labor (which are the only factors of production). The government can also use one-period bonds to smooth its expenditures over time. The assumptions necessary for Lemma 1 thus hold and production efficiency is achieved in steady state. Optimal capital taxes are then zero since households value assets in exactly the same way as the government values capital – for the return it generates in the next period. There is no consumption smoothing element in steady state and household and government discount factors are equal.

Following the notation above, let β denote the discount factor, $u_c(t)$ the marginal utility of consumption at time t, $\tilde{R}_t$ the household's rate of return on assets, θ_t is the government's value of resources, and $F_K(t)$ the marginal product of capital (including principal and depreciation). The household's and government's Euler equations are

$$u_c(t) = \beta u_c(t+1)\tilde{R}_{t+1} \tag{16}$$

$$\theta_t = \beta\theta_{t+1}F_K(t+1), \tag{17}$$

and therefore in steady state $\tilde{R} = F_K$. Constant returns to scale and perfect competition imply that the pre-tax return on assets is equal to the marginal product of capital and hence taxes on capital are optimally zero.

The picture changes when agents are not infinitely lived, as in Erosa and Gervais (2002). Assume for simplicity that population growth is zero and that utility is additively separable over time, with γ as an individual's time-discount factor. Each agent lives for J periods with a productivity profile $\{x_j\}_{j=1}^{J}$.[10] Lemma 1 and production efficiency in steady state hold in this case, but agents value assets beyond returns: if the life-cycle productivity profile is non-degenerate, then agents use assets to smooth consumption over their life-cycle. For example, let each agent live for two periods, with a high productivity when young and a lower one when old. Then agents save when young, even if the rate of return is lower than the inverse of the discount factor.

An individual's Euler equation of type j and its sum in the aggregate are given by

$$u_c(t,j) = \gamma u_c(t+1, j+1)\tilde{R}_{t+1} \tag{18}$$

$$\sum_{j=1}^{J-1} u_c(t,j) = \gamma\tilde{R}_{t+1}\sum_{j=1}^{J-1} u_c(t+1, j+1). \tag{19}$$

It is clear that if the discount rates of individuals and the government are different $\gamma \neq \beta$, then it is optimal to impose a tax (or subsidy, of course). But even if discount rates are equal, it is not guaranteed that $\sum_{j=1}^{J-1} u_c(t,j) = \sum_{j=1}^{J-1} u_c(t+1, j+1)$. In Chamley-Judd with infinite horizons, the consumption profile is the same across time periods in steady state; with overlapping generations, the cohorts who save are different from the cohorts who receive these savings. It follows that generally $\gamma\tilde{R}_{t+1} \neq 1$ and there are intertemporal distortions at the individual level (although not for society, since production efficiency of capital holds).

It is of course ordinarily not possible to characterize the optimal use of any specific tax, it is

[10] If labor inputs from different cohorts are of the same type (i.e. total effective labor is $N = \sum_J x_j n_j$, where n_j is labor supply of an individual of age j), then the government can tax all factors of production, even if it does not have access to age-dependent taxes.

the tax system as a whole that is relevant. However, it does seem natural to talk about capital and labor taxes and in line with a lot of the literature, I will refer to these specific taxes and not only allocations.[11] Erosa and Gervais (2002) show that capital taxes are generally non-zero: the government would like to tax individual labor supply when its income elasticity is relatively low and non-zero capital taxes limit the households' ability to substitute labor intertemporally.[12]

3.2 Idiosyncratic Productivity Shocks

When agents face uncertainty about their income, they use asset holdings to smooth their consumption profile. They will therefore value assets not only for the returns that they generate in the future, but also as an insurance device (unless they have access to a perfect insurance scheme). In the framework by Aiyagari (1995), infinitely-lived agents face uninsurable idiosyncratic income shocks and borrowing constraints. The government is also not able to offer insurance, as it can only use linear taxes on labor and capital. This implies that agents cannot smooth consumption perfectly (unless they have infinite assets) and that consumption in each period depends on the shock received. It thus follows that in steady state the average expected marginal utility next period is higher than the average marginal utility this period (for a strictly concave utility function):

$$\sum_I u_c(t,i,s) < \sum_I \sum_{S'} \mu(s'|s) u_c(t+1,i,s'). \tag{20}$$

$u_c(t,i,s)$ is the marginal utility of individual i at time t and history of shocks s. $\mu(s'|s)$ is the probability of moving to history s' conditional on history s. Now it is clear that in order to implement the modified golden rule, i.e. $1 = \beta F_K$, the government has to levy a tax on capital

[11] When consumption taxes for instance are also available along capital and labor taxes, then one of the taxes is redundant and the same allocation could be implemented without one of the three. It is therefore arbitrary from a technical point of view to focus on capital and labor taxes.

[12] Conesa, Kitao, and Krueger (2009) calibrate a life-cycle model to the United States and find a large optimal capital tax rate – 36% in their preferred specification.

income[13], since the individual's Euler equation of type i and its sum in the aggregate are

$$u_c(t,i,s) = \beta \tilde{R}_{t+1} \sum_{S'} \mu(s'|s) u_c(t+1,i,s') \tag{21}$$

$$\sum_I u_c(t,i,s) = \beta \tilde{R}_{t+1} \sum_I \sum_{S'} \mu(s'|s) u_c(t+1,i,s'). \tag{22}$$

In Golosov, Kocherlakota, and Tsyvinski (2003), agents also face uninsurable idiosyncratic income shocks. The government may use any form of taxation, but cannot observe agents' ability. In order to incentivize the more productive people to work (as opposed to shirking and mimicking the less productive workers while enjoying more leisure), the government has to reward the more able by granting them higher consumption. The same logic as before in the Aiyagari economy thus holds and an intertemporal wedge for the individual is needed to implement the modified golden rule. More formally, the inverse Euler equation is generally optimal in this type of economy:

$$\frac{1}{u_c(t,i,s)} = \frac{1}{\beta F_K(t+1)} \sum_{S'} \frac{\mu(s'|s)}{u_c(t+1,i,s')}. \tag{23}$$

The inverse of the marginal utility is the resource cost of providing that utility, so the average for all agents has to be equal across time periods in steady state:

$$\sum_I \frac{1}{u_c(t,i,s)} = \sum_I \sum_{S'} \frac{\mu(s'|s)}{u_c(t+1,i,s')}. \tag{24}$$

The marginal product of capital is thus undistorted, but a capital tax is necessary to achieve it, as consumption depends on the shock and therefore $\sum_I u_c(t,i,s) < \sum_I \sum_{S'} \mu(s'|s) u_c(t+1,i,s')$.

4 Departures from Production Efficiency

In this section I discuss cases where production efficiency does not hold and explain what causes these deviations in terms of violations of the assumptions in Lemma 1. I also briefly examine the importance of government bonds.

[13] Chamley (2001) argues that a violation of the modified golden rule is irrelevant for the evaluation of the efficiency of capital income taxation in the long-run. He assumes that the exogenous rate of return is below $1/\beta$ and that there are no government bonds, so a steady state with a constant value of government funds as in Aiyagari (1995) does not exist.

4.1 Incomplete Tax System

If intermediate good K affects the price of a good which cannot be taxed separately, i.e. which is part of vector p_2, then the first-order condition with respect to K contains an additional term and Lemma 1 fails to hold.

A well-known exception to the Chamley-Judd result is an incomplete tax system. Jones, Manuelli, and Rossi (1997) show that restricting tax rates leads to non-zero capital taxes: examples include a binding cap on the tax of pure rents, the inability to differentiate two different types of labor, or having only one income tax for both capital and labor income. Correia (1996a) shows a similar result for the case of a production factor which non-trivially interacts with capital (i.e. cross-derivatives in the production function are non-zero) that cannot be taxed. Reis (2011) analyzes an economy where entrepreneurial and capital income are indistinguishable and finds that labor taxes that are higher than capital taxes, but the latter are still positive. What separates all of these findings from the results presented in the previous section is that production efficiency does not hold and capital taxes are used to raise revenues, since the assumptions for Lemma 1 are not met. Similarly, if there is unobservable entrepreneurial (Albanesi, 2006) or investment effort which affects returns to capital, then the tax system is incomplete and production efficiency no longer holds.

4.2 Taxation of the Initial Capital Stock

If good K is an endowment, then its value enters the household's optimality conditions.[14] The first-order condition with respect to K thus contains additional terms outside of the resource constraint and Lemma 1 fails to hold.

Capital at time zero is clearly not an intermediate good as it is not produced inside the model framework (unless the timeless perspective, as proposed by Woodford (1999) for a related problem in monetary policy is applied). It should thus be fully taxed. If the capital tax at time zero is restricted, then taxes at subsequent periods can be used to indirectly tax the initial capital stock.[15] A large tax on capital at time one will of course deter investment at time zero and increase

[14]There are no government bonds for endowments, so their values always enter the household problem.

[15]Abel (2007) finds that taxing capital can generate significant tax revenues even in steady state when coupled with investment tax credits. The household's Euler equation is undistorted at any point and the tax credits for investment are lower than the returns on it. However, this proposed policy is taxation of the initial capital stock in disguise. The tax credit has to be paid one period before the capital taxes are collected; tracing this chain back, it reveals that all tax revenues date back to time zero, as there are no tax credits for the already existing capital stock.

consumption, but the curvature of the utility function limits this. The exceptions are the case of a quasi-linear utility function in consumption or a small open economy without residential taxes – when the net returns to capital are linear and independent of capital taxes (Gross, 2013a).[16]

4.3 (Human) Capital in the Utility Function

If a good K affects the government's and/ or household's objective function, then the first-order condition with respect to K contains additional terms outside of the resource constraint and Lemma 1 fails to hold.

When capital is in the utility function, it is obviously no longer an intermediate good. This raises a related question about human capital. Judd (1999) argues that it should be treated the same as physical capital; Jones, Manuelli, and Rossi (1997) find that when labor is used to generate human capital, then it should also be exempted from taxes. However, as Ljungqvist and Sargent (2004) point out, this result is due to their specification of the human capital accumulation process, which makes raw labor disappear from the implementability constraint. In other words, the setting by Jones, Manuelli, and Rossi (1997) ensures that raw labor is only relevant for human capital accumulation. A tax on labor therefore taxes human capital accumulation, similarly to capital taxes. The government cannot tax any endowments besides the initial one and thus sets optimal long-run taxes to zero, if possible. If not (for example because one does not allow for government bonds), then taxes on both intermediate goods have to be levied. If human capital conferred utility directly, on the other hand, then production efficiency would not apply to it.

4.4 A Note on Government Bonds

What is the importance of government bonds for production efficiency of intertemporal intermediate goods? They are necessary for Lemma 1, since they ensure that the government can smooth distortions over time without using capital. In steady state, this role ceases to be of importance anymore, since distortions will be the same across periods. On the other hand, in a stationary equilibrium, they still play an important part in transferring government funds between states of

[16] With imperfect commitment, as in Klein and Ros-Rull (2003), from the current government's perspective, the present capital stock is taken as given and the effects of capital taxes on past accumulation decisions (through expectations) are not considered. The government thus perpetually aims to tax what it perceives as an endowment, the current capital stock, but what is actually an intermediate good. As is well known, these taxes are inefficiently high.

nature and are necessary for production efficiency, see for instance Chari, Christiano, and Kehoe (1994) and Farhi (2010).

In an overlapping generations framework, government bonds can be important to implement the optimal capital stock. For example, Erosa and Gervais (2002) analyze an economy with weak separability of labor and age-dependent taxes. It follows that the government can induce a perfectly smooth consumption profile, which then requires zero capital taxes to obtain the optimal capital stock. Since agents are born with zero wealth and have no reason to save or borrow, the aggregate household wealth is zero and the entire capital stock has to be owned by the government.

Another interesting example is the case discussed in Piketty and Saez (2012). Assume a life-cycle model where each agent lives for one period and intrinsically values bequests and wealth at the end of life. With government bonds, the government can treat assets as a separate variable from capital and is thus able to tax wealth and bequests while at the same time ensuring production efficiency of capital. Government bonds are thus necessary to implement the optimal allocation of capital – if there were no bonds, capital would be directly present in the utility function and the assumptions necessary for Lemma 1 would be violated. See appendix B for details.

5 Alternative Proposals for Principles of Capital Taxation

Several other explanations have been brought up for the zero capital tax-result, besides production efficiency of intermediate goods. One of them is the infinitely elastic supply of capital in steady state, but Judd (1985) had already shown that the discount rate can be endogenous. Judd (1999) proves for a deterministic economy that the assumption of a steady state is not necessary, but that results hold on average over a long horizon. Two other explanations are more prominent, so I turn to them in more detail in this section.

5.1 Uniform Commodity Taxation

Corlett and Hague (1953) showed that commodities which are more complementary with leisure should be taxed at a higher rate than goods that are less complementary. In line with this result, Atkinson and Stiglitz (1976) proved that even when there is a motive for distribution among different types of agents, all commodities should be taxed uniformly when they are all equally complementary with leisure. If time is the only endowment, then taxing leisure and labor equally amounts to non-distortionary taxation. If leisure cannot be taxed directly, then taxing it indirectly through consumption goods which are complementary to it is the next-best alternative. So how does this relate to capital taxation?

In the baseline models used by Chamley and Judd, there is only a single consumption good per period, but consumption in period t and in period $t+s$ are of course distinct commodities. If R is the return on capital, τ^k is the tax rate on capital, both time-invariant, and there are no commodity taxes, then the price ratio of these two consumption goods is

$$\frac{p_{t+s}}{p_t} = \frac{1}{R^s(1-\tau^k)^s}. \tag{25}$$

Alternatively, if instead of capital taxes there are consumption taxes τ_t^c, one can express the price ratio as

$$\frac{p_{t+s}}{p_t} = \frac{1+\tau_{t+s}^c}{R^s(1+\tau_t^c)}. \tag{26}$$

The two expressions are equal if $\frac{1+\tau_{t+s}^c}{1+\tau_t^c} = (1-\tau^k)^{-s}$, so a constant capital tax is in this sense equivalent to an exploding consumption tax,[17] where the ratio of future to current consumption taxes increases exponentially as the time difference between the two periods grows larger. This seems to violate the principle of uniform commodity taxation so strongly that it cannot be optimal, see for instance Golosov and Tsyvinski (2008).

There are a few conceptual problems with this very intuitive explanation, though. First of all, Atkinson and Stiglitz (1976) are referring to several different commodities and one leisure good. In the infinite horizon setting considered by Chamley and Judd, there is one leisure good (and therefore a different endowment) in every period. While one could readily think that different commodities in the same period and thus for the same leisure good (or time-endowment) should be taxed at the same rate, there is no reason to assume that this should transfer to different commodities that are connected to different leisure goods. Indeed, Golosov, Kocherlakota, and Tsyvinski (2003) show that commodities should be taxed uniformly in the same period, but not across periods.

Furthermore, in the above example it is implicitly assumed that the return on capital R is independent of the capital tax; however, in the steady state of the baseline model of both Chamley and Judd, the return net of taxes is equal to the inverse of the discount factor (which I call β).

[17] The tax on consumption also acts as a tax on labor, so constant capital taxes are not equal to increasing consumption taxes on all dimensions.

Therefore, the steady-state price ratio is always $p_{t+s}/p_t = \beta^s$, independent of capital taxes. Of course, the marginal rate of transformation is still distorted when there are capital taxes, but this points to an inefficiency in production.

The prediction of the uniform commodity taxation argument is thus not clear: should capital taxes be zero or should the marginal rate of substitution and transformation be equal across consumption goods? And under what conditions? The exploding consumption tax argument is so general that it would speak against any form of capital taxation, no matter what the circumstances. Uniform commodity taxation, on the other hand, requires specific assumptions to hold, which are weak separability of consumption goods and leisure and non-linear labor income taxation.

Weak separability of consumption and leisure is not a necessary or sufficient condition in most cases, see for instance Chamley (1986).[18] Non-linear labor income taxation is generally not a necessary or sufficient condition either. If infinitely lived agents differ in their productivities and initial endowments but can only be taxed linearly, then it is still optimal to implement zero capital taxes in steady state, see for example Judd (1985). If capital is part of the utility function, then it is still optimal to tax capital when leisure is strongly separable and the government has recourse to non-linear labor taxation. There is thus no clear link between uniform commodity taxation and capital taxation.

5.2 Frontloading

The main idea behind the frontloading principle proposed by Albanesi and Armenter (2012) is that there is something special about the intertemporal margin, which was also mentioned by Jones, Manuelli, and Rossi (1997). According to this idea, intertemporal distortions will be compounded over time, so that it is preferable to have intratemporal distortions and/or have substantial intertemporal distortions for a limited time, but get rid of them in the long run. Hence the name frontloading. Albanesi and Armenter (2012, page 1) formulate a general condition, for which this principle holds: "If there exists an admissible allocation that converges to the first best steady state, then all intertemporal distortions are temporary in the second best."

In other words, if it is possible, although probably not optimal, for the government to accumulate enough resources in a finite amount of time to finance its expenditures for the rest of time, then it

[18] An exception is Erosa and Gervais (2002), who find zero optimal capital taxes under weak separability and age-dependent taxes.

will never impose any permanent intertemporal distortions. The latter is defined "as a wedge in the [household's] Euler equation for consumption." This explanation has the advantage that it states its predictions and the sufficient condition for it to hold very clearly. However, the problem arises that "the condition is often stronger than required for the result to hold in a specific application. (p.3)" The authors ascertain that "the logic of our results holds beyond its strict mathematical confines. (p.3)" I fully agree that in order to provide a very general proof, the assumptions are so general, that some cases are not formally covered anymore, even though the predicted result still holds. However, the question remains why for example government bonds are so central to both their proof and their logic, but not important at all for some steady-state results in capital taxation.

When one of the factors of production cannot be taxed, as in Correia (1996b), then it is optimal to tax capital in the long run. The reason is that the government tries to indirectly tax the untaxable factor through capital taxes. For example, if land is in perfectly inelastic supply but untaxable, then it generates rents, which the government would like to capture. The taxes on capital then depend on how much capital contributes to land rents: if more capital leads to higher rents, taxes are positive, but if more capital results in lower rents, then capital will be subsidized. While these taxes are second-best optimal, they do nonetheless distort the intertemporal margin. Formally, the framework by Albanesi and Armenter (2012) does not capture this case, as the assumption on the production function is of constant returns to scale in only two factors, capital and labor.

Nonetheless, one might wonder why the idea of frontloading does not apply in this case. It is definitely possible to accumulate enough assets so that the economy converges to the first-best. So then why are capital taxes optimally non-zero? Albanesi and Armenter (2012) argue in footnote 20 on page 14 that "The restrictions [...]do not rule out Ramsey models with incomplete factor taxation, such as Correia (1996) and Jones, Manuelli and Rossi (1997). These restrictions can be formulated [...] by including an additional constraint at date $t = 0$ that prevents the government from manipulating the present value of assets at date $t = 0$. See Armenter (2008) for a discussion." When there is a production factor in fixed supply, Armenter (2008) argues that the steady-state tax on capital is in fact a tax that imperfectly mimics a tax on the initial wealth at time zero, which is a lump-sum tax. He then shows how long-run capital taxes are optimally zero again if the government may not change the value of assets at time zero.

The value of an asset at time zero is equal to the discounted stream of future revenues it generates. The government would indeed like to capture these rents and imposes capital taxes to do so indirectly. If a constraint makes it impossible to tax the asset, then capital taxes are of course not going to be imposed.

However, capital taxes are non-zero even if the untaxable production factor is not in perfectly inelastic supply, as shown by Correia (1996b). In fact, if the government is not allowed to tax labor after some date $t \geq 0$, then it is optimal to have non-zero capital taxes even if it is possible to accumulate enough assets so that the economy converges to the first-best. I show this in appendix C. Such capital taxes are therefore not simply aimed to capture initial asset wealth, but rather to indirectly tax returns which may not be directly taxed – in line with the predictions of the production efficiency of intermediate goods.

Albanesi and Armenter (2012) provide an incredibly comprehensive framework showing when it is not optimal for a government to distort individuals' intertemporal margin. However, I believe that for optimal taxation, the government's intertemporal margin is the important one, which is not distorted when enough tax instruments are available. If individually rational household behavior results in suboptimal capital accumulation, then it is optimal to distort the household's intertemporal margin.

6 Conclusion

This paper presents a general framework to analyze optimal capital taxation. It shows that if capital is an intermediate good and all co-factors of production can be independently taxed, then it is optimal in the second best to set the marginal product of capital as in the first best (in a steady state, stationary equilibrium, or long-run average).[19] Distorting intermediate goods is generally not optimal since it represents a distortion for the final good as well. The same tax revenues can be levied by distorting only final goods at a lower efficiency cost.

The approach presented in this paper unifies many diverging results in the literature on optimal capital taxation. What makes capital special is its undistorted marginal product, a common feature in most of these models. The capital tax that implements it differs according to the modeling assumptions: Capital taxes are zero when households' capital accumulation is first-best without taxes, such as in the standard neo-classical model used by Chamley and Judd. When the discount factors of the government and of agents differs, when agents save to self-insure against idiosyncratic

[19] This also applies to open economies: I show in Gross (2013b) that if it is optimal to have an undistorted capital allocation in a closed economy, then it is also optimal for an open economy, whether it is small or large.

shocks, or when they save to smooth consumption over their lifetimes with a non-trivial earnings profile, then taxes are needed to align the individually rational savings decisions with aggregate production efficiency requirements. When the tax system is incomplete, capital taxes are used to affect the returns of the untaxable factor and capital is no longer provided efficiently. When capital is not an intermediate good but also features in the utility function, then its allocation is also distorted.

For future research, it would be interesting to estimate what drives household savings, for example to self-insure against health or income shocks, for status reasons, for bequests, retirement savings etc. If one could also get a grip on how far current tax systems are impeded in taxing different inputs at different rates, then it could be possible to evaluate if capital taxes should be raised to raise revenues or not and what the optimal tax rates could be. As it currently stands, estimates of optimal capital tax rates are heavily model-driven, depending on the set of model characteristics and assumptions (comparing for instance Atkeson, Chari, and Kehoe (1999) and Conesa, Kitao, and Krueger (2009)).

Governments are unlikely to have implemented optimal policies, especially concerning taxes, so it seems impossible to determine optimal taxes by looking at the ones currently in place. In cross-country comparisons, it is difficult to identify the effect of different tax systems (and a forteriori of different specific taxes) on economic performance. It thus seems reasonable to look for optimal taxes in models. Estimates of which of the model features mentioned above are empirically relevant would thus be a significant step towards selecting the appropriate model. If this model were then to be carefully calibrated, one could deliver policy recommendations which are both empirically grounded and yet informative about optimal policy.

References

Abel, A. B. (2007): "Optimal Capital Income Taxation," NBER Working Papers 13354, National Bureau of Economic Research, Inc.

Acemoglu, D., M. Golosov, and A. Tsyvinski (2008): "Political Economy of Intermediate Goods Taxation," *Journal of the European Economic Association*, 6(2-3), 353–366.

Aiyagari, S. R. (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103(6), pp. 1158–1175.

Albanesi, S. (2006): "Optimal Taxation of Entrepreneurial Capital with Private Information," Working Paper 12419, National Bureau of Economic Research.

Albanesi, S., and R. Armenter (2012): "Intertemporal Distortions in the Second Best.," *The Review of Economic Studies*.

Armenter, R. (2008): "A note on incomplete factor taxation," *Journal of Public Economics*, 92(10-11), 2275–2281.

Atkeson, A., V. Chari, and P. J. Kehoe (1999): "Taxing capital income: a bad idea," *Quarterly Review*, pp. 3–17.

Atkinson, A. B., and J. E. Stiglitz (1972): "The structure of indirect taxation and economic efficiency," *Journal of Public Economics*, 1(1), 97–119.

——— (1976): "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 6(1-2), 55 – 75.

Brecher, R. A., Z. Chen, and E. U. Choudhri (2010): "A dynamic model of shirking and unemployment: Private saving, public debt, and optimal taxation," *Journal of Economic Dynamics and Control*, 34(8), 1392 – 1402.

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

——— (2001): "Capital income taxation, wealth distribution and borrowing constraints," *Journal of Public Economics*, 79(1), 55 – 69, ¡ce:title¿ISPE 1998: Bequests and Wealth Taxation¡/ce:title¿.

Chari, V. V., L. J. Christiano, and P. J. Kehoe (1994): "Optimal Fiscal Policy in a Business Cycle Model," *Journal of Political Economy*, 102(4), 617–52.

Conesa, J. C., S. Kitao, and D. Krueger (2009): "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, 99(1), 25–48.

Corlett, W. J., and D. C. Hague (1953): "Complementarity and the Excess Burden of Taxation," *The Review of Economic Studies*, 21(1), pp. 21–30.

Correia, I. H. (1996a): "Dynamic optimal taxation in small open economies," *Journal of Economic Dynamics and Control*, 20(4), 691–708.

——— (1996b): "Should capital income be taxed in the steady state?," *Journal of Public Economics*, 60(1), 147–151.

Diamond, P., and E. Saez (2011): "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 25(4), 165–90.

Diamond, P. A., and J. A. Mirrlees (1971): "Optimal Taxation and Public Production: I–Production Efficiency," *American Economic Review*, 61(1), 8–27.

Domeij, D. (2005): "Optimal Capital Taxation and Labor Market Search," *Review of Economic Dynamics*, 8(3), 623–650.

Erosa, A., and M. Gervais (2002): "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 105(2), 338–369.

Farhi, E. (2010): "Capital Taxation and Ownership When Markets Are Incomplete," *Journal of Political Economy*, 118(5), pp. 908–948.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70(3), 569–587.

Golosov, M., and A. Tsyvinski (2008): "Optimal Fiscal and Monetary Policy (with commitment)," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke.

Gross, T. (2013a): "Dynamic Optimal Taxation in Open Economies," Carleton University Working Paper.

——— (2013b): "Optimal Taxation and Tax Competition," Carleton University Working Paper.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1997): "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73(1), 93–117.

Judd, K. L. (1985): "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics*, 28(1), 59–83.

——— (1999): "Optimal taxation and spending in general competitive growth models," *Journal of Public Economics*, 71(1), 1–26.

Klein, P., and J.-V. Ros-Rull (2003): "Time-consistent optimal fiscal policy," *International Economic Review*, 44(4), 1217–1245.

Kocherlakota, N. (2010): *The New Dynamic Public Finance*. Princeton University Press.

Ljungqvist, L., and T. J. Sargent (2004): *Recursive Macroeconomic Theory*. MIT Press, 2nd edition edn.

Mankiw, N. G., M. Weinzierl, and D. Yagan (2009): "Optimal Taxation in Theory and Practice," *Journal of Economic Perspectives*, 23(4), 147–74.

Munk, K. (1980): "Optimal taxation with some non-taxable commodities," *Review of Economic Studies*, 47, 755–765.

Piketty, T., and E. Saez (2012): "A Theory of Optimal Capital Taxation," Working Paper 17989, National Bureau of Economic Research.

Reis, C. (2011): "Entrepreneurial Labor And Capital Taxation," *Macroeconomic Dynamics*, 15(03), 326–335.

Woodford, M. (1999): "Optimal Monetary Policy Inertia," Working Paper 7261, National Bureau of Economic Research.

A Examples

Here I provide examples of how existing models in the literature can map into the framework presented in chapter 2.

A.1 Infinitely-lived Dynasties

Assume an economy similar to Chamley (1986) and Judd (1985). The representative agent takes prices as given and maximizes lifetime utility over an infinite horizon:

$$\sum_{t=0}^{\infty} \beta^t u(c_t, l_t), \tag{27}$$

where $u(c_t, l_t)$ is a well-behaved utility function over consumption c_t and leisure l_t. $\beta \in (0, 1)$ is the discount factor. The household has one unit of time at its disposal every period, which can be used for labor n_t and leisure. The per-period budget constraint is:

$$c_t = (1 - \tau_t^n) w_t n_t + [1 - \delta + (1 - \tau_t^k) r_t] k_t - k_{t+1} + (1 + R_t) b_t - b_{t+1}. \tag{28}$$

b_t are government bonds and R_t is the interest rate on them. k_t is the amount of capital, w_t and r_t are the wage and interest rate. k_0 and b_0 are exogenously given. $0 \leq \delta \leq 1$ is the capital depreciation rate. Finally, τ_t^n and τ_t^k are the tax rates on wages and capital, respectively. Optimal behavior implies a no-arbitrage condition, that the returns on government bonds and capital must be equal after taxes,

$$1 + R_{t+1} = 1 - \delta + (1 - \tau_{t+1}^k) r_{t+1}, \tag{29}$$

as well as the familiar conditions concerning the trade-off between consumption versus leisure and consumption today versus tomorrow:

$$u_l(t) = u_c(t)(1 - \tau_t^n) w_t \tag{30}$$

$$u_c(t) = \beta u_c(t + 1)[1 - \delta + (1 - \tau_{t+1}^k) r_{t+1}]. \tag{31}$$

Subscripts refer to derivatives with respect to that variable, e.g. $u_c(t)$ is the derivative of the utility function with respect to consumption at time t.

Output is produced by a representative firm with the private inputs labor n_t and capital k_t according to a production function $h(k, n)$ with constant returns to scale that satisfies the Inada conditions. The maximization of profit, along with constant returns to scale implies zero profits and the following remunerations for the inputs:

$$r_t = f_k(k_t, n_t) \tag{32}$$

$$w_t = f_n(k_t, n_t). \tag{33}$$

The benevolent government's objective is to maximize the utility of its citizens. It needs to finance an exogenous stream of unproductive expenditures $\{g_t\}_\infty$, which converges to a constant g after some finite time to allow for a steady state. Revenue is generated by distortionary taxes on capital earnings τ_t^k and wages τ_t^n; to avoid lump-sum taxation $\tau_0^k = 0$. The government may trade in one-period bonds, with b_t denoting the total outstanding government debt. The government's per-period budget constraint is

$$g_t + b_t(1 + R_t) = \tau_t^k r_t k_t + \tau_t^n w_t n_t + b_{t+1}. \tag{34}$$

Using the no-arbitrage condition, one can eliminate R_t. Furthermore, define assets $a_t = b_t + k_t$ and after-tax prices $\tilde{r}_t = (1 - \tau_t^k) r_t$ and $\tilde{w}_t = (1 - \tau_t^n) w_t$. Adding the household's and government's budget constraint results in the national resource constraint (and using the fact that $f(k_t, n_t) = w_t n_t + r_t K_t$):

$$f(k_t, n_t) + k_t(1 - \delta) - k_{t+1} - c_t - g_t. \tag{35}$$

The government's problem is thus to maximize

$$\sum_{t=0}^{\infty} \beta^t u(c_t, 1-n_t) \tag{36}$$

$$\text{s.t.} \tilde{w}_t n_t + (1-\delta+\tilde{r}_t) a_t - a_{t+1} - c_t = 0 \quad \forall\, t \tag{37}$$

$$u_n(t) + u_c(t)\tilde{w}_t = 0 \quad \forall\, t \tag{38}$$

$$u_c(t) - \beta u_c(t+1)[1-\delta+\tilde{r}_{t+1}] = 0 \quad \forall\, t \tag{39}$$

$$f(k_t, n_t) + k_t(1-\delta) - k_{t+1} - c_t - g_t = 0 \quad \forall\, t \tag{40}$$

$$\tilde{r}_0 - f_k(k_0, n_0) = 0, \tag{41}$$

where the set of choice variables $\hat{X}$ is

$$\hat{X} = \{c_t, n_t, k_{t+1}, a_{t+1}, \tilde{w}_t, \tilde{r}_t\}_{t=0}^{\infty}. \tag{42}$$

In the context of the framework of this paper, allocations are $X = \{c_t, n_t, k_{t+1}\}_{t=0}^{\infty}$; the objective function $V(X_{-K})$ corresponds to (36); the constraint set $\Omega_H(X_{-K}, A, \tilde{p}) \geq 0$ is described by (37), (38), and (39); assets are $A = \{a_t\}_{t=0}^{\infty}$; firm optimality conditions $p - f(X) = 0$ are given by equations (32) and (33); and restrictions on tax rates $\tau(p_2, \tilde{p}_2) \geq 0$ are given by $\tilde{r}_0 = f_k(k_0, n_0)$, which satisfies the condition for Lemma 1, that all co-factors of capital are independently taxable (in steady state). The resource constraint $F(X) = 0$ is given by (40).

A.2 Borrowing Constraints

Assume an economy as in Aiyagari (1995). The notation is very similar, so sticking to the notation in this paper should still leave it easily comparable. In particular, I call the idiosyncratic productivity shocks π instead of θ, after-tax returns are denoted by a tilde instead of a bar (e.g. $\tilde{r}$ instead of $\bar{r}$), and government expenditure and debt are g and b instead of the capitalized letters. Moreover, I introduce the borrowing constraint $\epsilon \leq 0$ (instead of zero). The borrowing constraint is generally not binding for the entire population. $J(a, \pi)$ is the distribution over assets and skills and all per-capita terms are for a part of this distribution. n_t is an individual's labor supply, whereas N_t is the aggregate effective labor supply (i.e. $N = \int \pi n(a,\pi) dJ(a,\pi)$). The government's problem is

$$\sum_{t=0}^{\infty} \beta^t \{ \int_{J_t} u(c_t) + U(g_t) \} dJ_t \tag{43}$$

$$\text{s.t.} \tilde{w}_t n_t + \pi_t H(1-n_t) + (1-\delta+\tilde{r}_t) a_t - a_{t+1} - c_t = 0 \quad \forall\, t \quad \text{and} \quad \forall\,(a,\pi) \quad \text{with} \quad dJ(a,\pi) > 0 \tag{44}$$

$$\tilde{w}_t - \pi_t H'(1-n_t) = 0 \quad \forall\, t \quad \text{and} \quad \forall\,(a,\pi) \quad \text{with} \quad dJ(a,\pi) > 0 \tag{45}$$

$$u_c(t) - \beta E_t[(1-\delta+\tilde{r}_{t+1}) u_c(t+1)] = 0 \quad \forall\, t \quad \text{and} \quad \forall\,(a,\pi) \quad \text{with} \quad dJ(a,\pi) > 0 \tag{46}$$

$$a_{t+1} \geq \epsilon \quad \forall\, t \quad \text{and} \quad \forall\,(a,\pi) \quad \text{with} \quad dJ(a,\pi) > 0 \tag{47}$$

$$f(k_t, N_t) + k_t(1-\delta) - k_{t+1} - c_t - g_t = 0 \quad \forall\, t \tag{48}$$

$$\tilde{r}_0 - f_k(k_0, n_0) = 0. \tag{49}$$

The set of choice variables $\hat{X}$ is

$$\hat{X} = \{\mathbf{c_t}, \mathbf{n_t}, \mathbf{a_{t+1}}, g_t, k_{t+1}, \tilde{w}_t, \tilde{r}_t\}_{t=0}^{\infty}, \tag{50}$$

where $\mathbf{c_t}$ (and similarly $\mathbf{n_t}$ and $\mathbf{a_{t+1}}$) stand for the matrix of consumption (and market labor and next-period asset holdings) in the space of assets and productivity (a,π). In the context of the framework of this paper, allocations are $X = \{\mathbf{c_t}, \mathbf{n_t}, k_{t+1}, g_t\}_{t=0}^{\infty}$; the objective function $V(X_{-K})$ corresponds to (43); the constraint set $\Omega_H(X_{-K}, A, \tilde{p}) \geq 0$ is described by (44), (45), (39), and (47); assets are $A = \{\mathbf{a}_t\}_{t=0}^{\infty}$; firm optimality conditions $p - f(X) = 0$ are given by equations (32) and (33) (they are equal to the marginal products, exactly as in Chamley-Judd); and restrictions on tax rates $\tau(p_2, \tilde{p}_2) \geq 0$ are given by $\tilde{r}_0 = f_k(k_0, n_0)$. The resource constraint $F(X) = 0$ is given by (48).

A.3 Overlapping Generations

Assume an economy as in Erosa and Gervais (2002). As before, I will slightly modify the notation; I also assume that population and productivity growth is zero. n_t is labor supply (instead of l_t) and age-dependent productivity is π_j (instead of z_j). For clarification, U^t is the lifetime utility of a member of the cohort born at time t and $n_{t-j,j}$ for instance is the labor supply of an individual born at time $t-j$ who is j periods old; then $n_t = \sum_{j=0}^{J} \pi_j n_{t-j,j}$. I assume that the government only has access to age-independent taxes of capital and labor (extending the set of taxes does not alter the result of production efficiency). The government's problem is

$$\sum_{t=-J}^{\infty} \beta^t U^t \tag{51}$$

$$\text{s.t.} \tilde{w}_t \pi_j n_{t-j,j} + (1-\delta+\tilde{r}_t) a_{t-j,j} - a_{t-j,j+1} - c_{t-j,j} = 0 \quad \forall t \geq 0 \quad \text{and} \quad \forall j \in \{0,\ldots,J\} \tag{52}$$

$$\tilde{w}_t \pi_j U^{t-j}_{c_{t-j,j}} - U^t_{n_{t-j,j}} = 0 \quad \forall t \geq 0 \quad \text{and} \quad \forall j \in \{0,\ldots,J\} \tag{53}$$

$$U^{t-j}_{c_{t-j,j}} - U^{t-j}_{c_{t-j,j+1}} (1-\delta+\tilde{r}_{t+1}) = 0 \quad \forall t \geq 0 \quad \text{and} \quad \forall j \in \{0,\ldots,J\} \tag{54}$$

$$f(k_t, n_t) + k_t(1-\delta) - k_{t+1} - c_t - g_t = 0 \quad \forall t. \tag{55}$$

The set of choice variables $\hat{X}$ is

$$\hat{X} = \{\{c_{t-j,j}, n_{t-j,j}, a_{t-j,j+1}\}_{j=0}^{J}, k_{t+1}, \tilde{w}_t, \tilde{r}_t\}_{t=0}^{\infty}. \tag{56}$$

In the context of the framework of this paper, allocations are $X = \{\{c_{t-j,j}, n_{t-j,j}, a_{t-j,j+1}\}_{j=0}^{J}, k_{t+1}\}_{t=0}^{\infty}$; the objective function $V(X_{-K})$ corresponds to (51); the constraint set $\Omega_H(X_{-K}, A, \tilde{p}) \geq 0$ is described by (52), (53), and (54); assets are $A = \{\{a_{t-j,j}\}_{j=0}^{J}\}_{t=0}^{\infty}$; firm optimality conditions $p - f(X) = 0$ are irrelevant, as there are no restrictions on tax rates, and $\tau(p_2, \tilde{p}_2) \geq 0$ is empty. The resource constraint $F(X) = 0$ is given by (55).

A.4 New Dynamic Public Finance

Assume an economy as in Golosov, Kocherlakota, and Tsyvinski (2003), section three, (as specified in Theorem 1). Let $f(K_t, N_t)$ be the production function and total effective labor supply $N_t = \int y_t d\mu$. The government's problem in their paper (not listing the non-negativity constraints, which is non-binding for K_t) is to find the supremum of

$$\sum_{t=0}^{T} \beta^t \int U(c_t, y_t/\theta_t)\chi_1 d\mu \tag{57}$$

$$\text{s.t.} W(\sigma^* : c, y) \geq W(\sigma : c, y) \quad \forall \sigma \in \Sigma \tag{58}$$

$$f(K_t, N_t) + K_t(1-\delta) - K_{t+1} - C_t \quad \forall t. \tag{59}$$

The set of choice variables $\hat{X}$ is[20]

$$\hat{X} = \{c_t, y_t, K_{t+1}\}_{t=0}^{\infty}. \tag{60}$$

In the context of the framework of this paper, allocations are $X = \{c_t, y_t, K_{t+1}\}_{t=0}^{\infty}$; the objective function $V(X_{-K})$ corresponds to (57); the constraint set $\Omega_H(X_{-K}, A, \tilde{p}) \geq 0$ is described by (58); assets A have already been incorporated and firm optimality conditions $p - f(X) = 0$ are irrelevant here; $\tau(p_2, \tilde{p}_2) \geq 0$ is empty. The resource constraint $F(X) = 0$ is given by (59).

A.5 Unemployment: Search-frictions

One could presume that involuntary unemployment invalidates production efficiency, but this is not generally the case. Domeij (2005) analyzes optimal fiscal policy in a model of unemployment due to labor market search. As before, if the government is able to tax all factors of production (and vacancies or labor market tightness is one of them), production efficiency ensues, otherwise it is violated. Taxing labor market tightness can be achieved through either a subsidy for vacancies by firms or through unemployment benefits. As a special case, when the Hosios condition holds, labor market tightness is always optimal and it is not necessary to tax (or subsidize) it, so production efficiency still applies (since the optimal tax on market tightness would be zero in that case). Domeij (2005) employs the primal approach, eliminating prices and taxes, which is a very convenient formulation in this case. I will show how it maps using the primal approach, although one could also start from the initial problem including taxes and prices. When the government is able to tax

[20]Note that c_t and y_t are mappings from histories of skills to allocations of consumption and effective labor and $C_t = \int c_t d\mu$.

vacancies or provide unemployment benefits, the problem is

$$\sum_{t=0}^{\infty} \beta^t u(c_t, 1 - n_t - s_t) \tag{61}$$

$$\text{s.t.} W_0 - \sum_{t=0}^{\infty} ((c_t - L_t)u_{1,t} - u_{2,t}s_t - u_{2,t}n_t) = 0 \tag{62}$$

$$n_{t+1} - As_t x_t^{1-\phi} - (1-\psi)n_t = 0 \quad \forall t \tag{63}$$

$$f(k_t, n_t) + k_t(1-\delta) - k_{t+1} - c_t - g_t = 0 \quad \forall t \tag{64}$$

$$\tau_0^k,\ \tau_0^n,\ \tau_0^a \text{ given} \tag{65}$$

where the set of choice variables $\hat{X}$ is

$$\hat{X} = \{c_t, x_t, s_t, k_{t+1}, n_{t+1}\}_{t=0}^{\infty}. \tag{66}$$

In the context of the framework of this paper, allocations are $X = \{c_t, x_t, s_t, k_{t+1}, n_{t+1}\}_{t=0}^{\infty}$; the objective function $V(X_{-K})$ corresponds to (36); the constraint set $\Omega_H(X_{-K}, A, \tilde{p}) \geq 0$ is described by (62) and (63); assets and firm optimality conditions have already been incorporated; and restrictions on tax rates are given by (65). The resource constraint $F(X) = 0$ is given by (40).

A.6 Unemployment: No-shirking wages

If employers pay no-shirking wages as in Brecher, Chen, and Choudhri (2010), then production efficiency prevails, since the no-shirking wage is in terms of the after-tax wage. As the household's steady-state investment decision is no different from the government's and the marginal product of capital does not feature an externality, capital taxes are optimally zero. The economy is the same as in Chamley and Judd, except that households choose whether to shirk or not and not how much to work. The paper is written in continuous time and I follow this approach and its notation here, with household assets Y (instead of X) being the exception. Replacing the government budget constraint with the household budget constraint (and having household assets instead of government debt as a state variable) does of course not change results. The government's problem is to maximize

$$\int_0^\infty e^{-\rho t}[(\mu^{-1/\theta})^{1-\theta}/(1-\theta) - \delta Z]dt \tag{67}$$

$$\text{s.t.} \dot{Y} - \tilde{r}Y - \tilde{w}Z + \mu^{-1/\theta} = 0 \tag{68}$$

$$\dot{Z} - (\tilde{w}\mu q/\delta - \rho - b - q)(1-Z) + bZ = 0 \tag{69}$$

$$\dot{\mu} - \mu[\rho - \tilde{r}] = 0 \tag{70}$$

$$\dot{K} - F(K,Z) + \mu^{-1/\theta} = 0 \tag{71}$$

$$\tilde{w} \geq 0 \tag{72}$$

$$\tilde{r} \geq 0, \tag{73}$$

where the set of choice variables $\hat{X}$ at every instant is

$$\hat{X} = \{\dot{Y}, \dot{Z}, \dot{\mu}, \dot{K}, \tilde{w}, \tilde{r}\}. \tag{74}$$

In the context of the framework of this paper, allocations are $X = \{\mu, Z, K\}$ at every instant; the objective function $V(X_{-K})$ corresponds to (67); the constraint set $\Omega_H(X_{-K}, A, \tilde{p}) \geq 0$ is described by (68), (69), and (70) at every instant; assets are $A = Y$ at every instant; firm optimality conditions $p - f(X) = 0$ are given by $w = F_Z(K,Z)$ and $r = F_K(K,Z)$ at every instant; and restrictions on tax rates $\tau(p_2, \tilde{p}_2) \geq 0$ are given by (72) and (73), none of which are binding in steady state. The resource constraint $F(X) = 0$ is given by (71) at every instant.

B Valuing wealth and bequests

Assume as in Piketty and Saez (2012) a consumer who intrinsically values wealth z and bequests q besides consumption c and leisure $1-n$, living for one period:

$$\max \quad u(c,n,z,q) \tag{75}$$

$$s.t. \quad n\tilde{w} + a\tilde{R} - c - a' \geq 0. \tag{76}$$

The after-tax wage is $\tilde{w}$ and the after-tax return on initial assets a is $\tilde{R}$. Assets next period a' are equal to wealth z and bequests equal $a'\tilde{R}$; utility can thus be rewritten as $u(c,n,a,\tilde{R})$. Assuming that the government maximizes the discounted utility of generations and has to finance an exogenous amount of government spending, the Lagrangean is then:

$$\pounds = \max \sum_{t=0}^{\infty} \beta^t u(c_t, n_t, a_t, \tilde{R}_t) \tag{77}$$

$$+\psi_t[F(K_t, n_t) - n_t\tilde{w}_t + K_t(1-\delta) - K_{t+1} + a_{t+1} - a_t\tilde{R}_t - g] \tag{78}$$

$$+\theta_t[F(K_t, n_t) + K_t(1-\delta) - K_{t+1} - c - g] \tag{79}$$

$$+\mu_t[u_n(t) + u_c(t)\tilde{w}_t] \tag{80}$$

$$+\gamma_t[u_c(t) - u_a(t)], \tag{81}$$

where β is the government's discount factor and ψ, θ, μ, and γ are the multipliers for the government budget constraint, the resource constraint, and the household's optimality conditions for labor and assets, respectively. The first-order condition for capital is simply $(\theta_t + \psi_t)(F_k(t) + 1 - \delta) = (\theta_{t+1} + \psi_{t+1})/\beta$, implying the modified golden rule in steady state. Assets on the other hand should optimally be taxed; the first-order condition implies $\psi_t \tau_t^k r_t = -u_a(t) - \mu_t(u_{na}(t) + u_{ca}(t)\tilde{w}_t) + \gamma_t(u_{aa}(t) - u_{ca}(t))$, where τ_t^k is the tax rate on assets and $r_t = F_K(t)$ is the pre-tax return.

It thus becomes apparent that there is a crucial difference in how bequests are modeled: If they affect the utility of the testator directly, then it calls for capital taxes. If on the other hand bequests are valued indirectly since they will allow the heir to afford more consumption, thereby increase the utility of the heir and thus the utility of the testator (as in the infinite dynasty setup), then bequests do not provide an additional reason to tax capital.

C Capital Taxation with Limited Labor Taxes

In this section, I show how capital taxes are generally non-zero in the long run when one of the factors of production cannot be taxed. This has of course been done before, but what I would like to emphasize in this example is that the capital taxes do not arise simply in order to tax initial assets. I therefore construct the example in such a way that the factor that is untaxable in steady state can be taxed early on, is non-accumulable, and that there is another (non- intermediate) input besides capital which can be taxed in steady state.

Assume an economy as in section A.1, except that there are two types (type 1 and type 2) of labor which can initially both be taxed, but that taxes on labor of type 1 are not available anymore after some date $T \geq 0$. The household's per-period utility function is $u(c_t, n_{1t}, n_{2t})$ and the production function is $f(k_t, n_{1t}, n_{2t})$. Assume furthermore that the government can potentially amass enough revenues in finite time to finance all future expenditures and that the economy converges to a steady state with positive consumption. The implementability constraint and resource constraints are standard, but additional constraints are in place for $t > T$ to account for the fact that taxes on labor earnings of type 1 are no longer available:

$$w_t^1 = -u_{n_1}(t)/u_c(t)\,\forall t > T. \tag{82}$$

Let the Lagrange multiplier for these constraints be $\mu_{t|t>T}$ and θ_t for the resource constraints. The first-order condition for next period's capital k_{t+1} for $t \geq T$ is then

$$\theta_t = \beta\theta_{t+1}(1 - \delta + F_k(t+1)) - \mu_{t+1}\partial w_{t+1}^1/\partial k_{t+1}. \tag{83}$$

In steady state, this implies together with the household's Euler equation

$$1 + F_k(1 - \tau^k) - \delta = 1 - \delta + F_k - \frac{\mu}{\theta}F_{kn_1} \tag{84}$$

$$\Leftrightarrow \tau^k = \frac{\mu}{\theta}\frac{F_{kn_1}}{F_k} \tag{85}$$

Capital taxes are therefore positive in steady state: The marginal product of capital is positive, the cross-derivative F_{kn} is positive for a regular production function, the value of resources θ is positive, and the value of μ is also positive, since an ϵ tax on labor of type 1 raises revenues with negligible distortions. At the same time, it is clear that the capital taxes are not used to tax the initial value of an asset.