Investment Subsidies and Redistributive Capital Income Taxation in a Neoclassical Growth Model*

Günther Rehme

Technische Universität Darmstadt†

April 13, 2021

Abstract

How do investment subsidies bear on pure redistribution when coupled with capital income taxes? In a heterogeneous-agent, neoclassical growth framework it is found that in the short run and absent optimizing behaviour investment subsidies are good for growth but bad for redistribution. But when all agents and the government act time-consistently and optimally for the long run, the tax scheme does not distort accumulation anymore. This holds regardless of social preferences. For the feedback Stackelberg equilibrium I find that (pure) redistribution can go either way and capital income taxes are nonzero in the long run optimum, depending on the social weight of those who receive redistributive transfers, the distribution of pre-tax factor incomes, and the intertemporal elasticity of substitution. It is argued that investment subsidies may be an important indirect tool for redistribution.

KEYWORDS: Economic Growth, Redistribution, Investment Subsidies, Capital Income Taxes, Stackelberg, Time Consistency

JEL classification: O41, H21, D33, E6, C73

*I have benefitted from feedback by Santiago Caram, Kithsiri Ehelepola, Sabine Eschenhof-Kammer, Christian Keuschnigg, Enrique Neder, Marco Runkel, Mar Delgado Téllez as well as from presentations at the IIPF Congress in Maastricht 2008. the Econometric Society Meetings in Rio de Janeiro 2008 and Tokyo 2009, and the Center for Training and Research in Public Finance and Policy (CTRPFP), Kolkata, 2015, the Arnoldshain Seminar XV "The EU and Latin America Facing Globalization" in Vienna, 2017, the 10th International Research Conference (IRC) at the Central Bank of Sri Lanka in Colombo, 2017, the 11th VEAM Meeting in Hanoi, 2018, and the XXVI Encuentro de Economía Pública in Oviedo, 2019. Of course, all errors are my own.

†*Correspondence* TU Darmstadt, Department of Economics, FB 1/VWL 1, Hochschulstraße 1, D-64289 Darmstadt, Germany. *phone:* +49-6151-16-57258; *fax:* +49-6151-16-57257; *e-mail:* rehme@vwl.tu-darmstadt.de

1 Introduction

One realistic fiscal policy reaction during economic crises has been to institute measures that fight a fall in investment activity. Many of these measures allow for writing off a particular percentage of the investment outlays against the investors' tax bill, especially capital income taxes.[1]

The fact that such fiscal measures promote investment is well known. However, less well-known are the distributional consequences of them. If redistribution is financed out of taxes, then allowing investors to write off some of their outlays against their tax bill reduces net tax revenues and that seems to have negative consequences for redistribution. On the other hand, investment promotion stimulates economic growth and that may be good for redistribution, especially in the long run. That is the problem that is analyzed in this paper. More precisely, a capital income tax scheme coupled with investment subsidies is considered to analyze their effect on long-run economic growth and redistribution. The analysis is set in a two-class, neoclassical growth framework.

In the literature on optimal taxation Judd (1985) and Chamley (1986) have shown that capital income taxes are no good instruments for pure redistribution in a neoclassical growth framework. Their finding is that optimally capital income taxes should be zero in the long run.[2]

The intuition for the result is intriguing. Even workers who may not own capital and may, therefore, not accumulate resources might benefit more from higher steady state wages resulting from nondistorted accumulation with zero taxes than having redistributive transfers now at the expense of a lower steady state capital stock and so wages in the long run.

The authors then contemplated other capital income policy packages, including consumption taxes, and basically found the same result as in, for instance, Judd (1999). However, the result that capital income taxes are no good instruments for redistribution need not always hold. The optimal capital income tax rate may be nonzero in the Judd-Chamley and other growth contexts as has been shown by many contributions.[3]

[1]See, for example, Asen (2020) for recent data on capital allowances in OECD countries. Her study suggests a wide array of differences among these countries.

[2]Sargent and Ljungqvist (2004), p. 487, call this a "celebrated result". Similar results have been obtained by many authors as, for example, by Lucas (1990).

[3]See, for example, Kemp, van Long, and Shimomura (1993), Aiyagari (1995), Uhlig and Yanagawa (1996), Lansing (1999) Grüner and Heer (2000), Chamley (2001), Domeij and Heathcote (2004), Werning (2007), Conesa, Kitao, and Krueger (2009), Zhang, Davies, Zeng, and McDonald (2008), Selim (2010), Saez (2013), Reinhorn (2018), Straub and Werning (2020), Conesa and Domínguez (2020)

The present paper relates to these findings and develops interesting novel features of those insights. In particular, I consider a policy package in a Markov-perfect environment whereby investment subsidies are coupled with capital income taxes.[4] The tax revenues are used for investment subsidies and for pure (unproductive) redistribution from the accumulated to the non-accumulated factor of production. That governments redistribute resources but also subsidize investment appears to be a pervasive phenomenon in most countries. Hence, these realistic features may justify the policy package under consideration.

It is shown that coupling capital income taxes with investment subsidies to finance pure redistributive transfers to the non-accumulated factor of production ("workers") may also imply a nondistortionary policy package, similar to a consumption tax on "capitalists". Thus, the present paper relates to the finding of e.g. Jones, Manuelli, and Rossi (1997) who show in a *representative agent* framework that an investment subsidy can offset the growth distortion associated with a capital income tax and that a consumption tax is the optimal second best policy. A similar point was also made by Kaldor (1955) and Fisher (1937) who basically proposed that taxable "income" should be "income after savings are taken out". See Fisher (1937), p. 54. In the present paper, however, we contemplate a (simple) heterogeneous agent framework with a capital-income-cum-investment-subsidy tax scheme.

Analyzing that tax scheme then yields the following. For arbitrary, i.e. not necessarily optimizing behaviour of the agents (capital owners and workers) and the government granting more investment subsidies is generally good for economic growth, but bad for redistribution. In turn, higher capital income taxes are bad for growth, but good for redistribution. Furthermore, it is shown that when there is a sudden drop in the real return to capital, as happens most often during economic crises, the government should increase the investment subsidies or cut the capital income tax rate, if it wishes to stabilize the real return to investment. These results would correspond to what one usually expects.

But when the agents and a benevolent government that represents the weighted interests of the workers and the capitalists act optimally, it turns out that in a feedback Stackelberg equilibrium the government would mostly find it optimal to choose a subsidy policy that does not distort accumulation. The reason is that it would remove the

---

and others.

[4] As Long (2010) I use the terms feedback strategy and Markov-perfect strategy interchangeably. The associated equilibrium concepts are, thus, alternative names for the same concept.

distorting effect policy has on capital accumulation. Thus, the paper shows that, in relation to the investors' tax bill, a proportionally equal (full) expensing of investment outlays is optimal for the long run.[5]

That finding is not really new and has, for example been obtained in a partial equilibrium context by Samuelson (1964), Hall and Jorgenson (1967) and Hall and Jorgenson (1971).[6] Thus, the paper generalizes the optimality result to a simple dynamic heterogeneous agent, two-class model with potential distributional conflicts. The reason is that the result derived in the present framework does not depend on the social weights attached to the interests of different factor owners. Even an entirely pro-labour government would choose a nondistortionary tax-subsidy policy, even though this could mean less tax revenues and so less redistributive transfers.

An important implication of the result is that under the tax scheme analyzed here efficiency and equity concerns can separated from one another. Unlike for the Judd-Chamley results efficiency does not requires one, and only one long-run, optimal redistributive policy choice. Thus, in the model granting investment subsidies potentially serves as an important *indirect* redistribution device, because transfers ultimately depend on the capital income tax rate chosen.

In fact, when a benevolent government pursues such a nondistortionary policy that seems to benefit everybody, it turns out that capital income taxes are not always zero in the long-run optimum. In fact, they may be positive or negative, implying that (pure) redistribution can go either way in the model. That means that redistribution towards capital or to labour can be optimal choices for the long run, depending on the model's parameters.

In summary, I find for the feedback Stackelberg equilibrium that capital income taxes are optimally nonzero and that pure redistribution can go either way. The results depends on very intuitive conditions. As one might expect from actual taxation by governments the optimal choice of capital income taxes in the long run depends on the social weight of those who receive redistributive transfers, the distribution of pre-tax income among individuals, and the intertemporal elasticity of substitution.

[5] In that sense the assumption of full expensing of investment outlays in relation to the capital income tax bill made in Rehme (1995), and Rehme (2002) which provided verbal arguments why this is optimal in a general equilibrium, endogenous growth framework, is endogenized in this paper and found to be optimal in the present general equilibrium, neoclassical growth framework. For papers that find a related result see, for example, Davies, Zeng, and Zhang (2009).

[6] In a representative agent, dynamic general equilibrium framework Abel (2007) has established the same result.

In a numerical simulation based on U.S. data the theoretical results are given a quantitative flavor. In particular, the simulation highlights the important relationship between the strength of social preferences and positive (negative) capital income taxes and their effect on (pure) redistribution.

The main message of the present paper is that investment subsidies, when coupled with capital income taxes, are important redistributive devices in a long-run optimum.

The paper is organized as follows: Section 2 presents the model. Section 3 analyzes the optimality for tax rates in long-run equilibrium. Section 5 provides concluding remarks.

2 The Model

The model is set in continuous time and the following conventions are used. A variable m functionally depending on another variable z is denoted by $m = m[z]$, that is, square brackets $[\cdot]$ denote a functional dependence. For all variables that are continuous functions of time the subscript t denotes their dependence on time. Thus, we define $s_t \equiv s[t]$ for some variable s depending on time t. As is common, the change of a variable depending on time, i.e. $\partial s_t/\partial t$, is denoted by $\dot{s}_t$. In contrast, a change in a variable z with its effect on m are interchangeably denoted by m_z or $\partial m/\partial z$. For a simple derivative I use the convention $m' \equiv m'[z] \equiv dm/dz$.

The economic environment is taken to be the following. The economy consists of a government, identical competitive firms and two types of infinitely-lived, equally patient and price taking individuals called workers and capitalists. All agents derive utility from the consumption of a homogenous, malleable good. For simplicity, we normalize the population so that the group of capitalists and workers can be treated as one individual each.

The model abstracts from population growth, uncertainty, and technological progress. We assume the workers supply one unit of (raw) labour inelastically and do not save or invest.[7] Thus, all the wealth is concentrated in the hands of the capitalists who supply the services from their (human and non-human) wealth to competitive firms.

[7] The assumption may be rationalized by imposing transaction costs on the workers when borrowing small amounts. Thus, the model uses the commonly used framework of Kaldor (1956) and Pasinetti (1962), which is also employed by Judd (1985) and Lansing (1999).

2.1 Capitalists

At each period the *capital owners* choose how much of their income to consume or invest, and they take prices and policy as given. We assume that capital is *broadly* defined and includes human capital. See Mankiw, Romer, and Weil (1992). This assumption implies that the model also captures distributional problems between owners of physical and human capital on the one hand and unskilled workers on the other. The details following from this assumption are set out in appendix A. For simplicity the term *capital* will always refer to *broad capital* in the rest of the paper.[8]

The capital owners' instantaneous budget constraint is given by

$$c_t + i_t = r_t k_t - T_t \quad \text{and} \quad i_t = \dot{k}_t. \tag{1}$$

Thus, the capitalists derive income from renting their physical and human capital, k_t, to competitive firms at the rate r_t. In addition they have to pay taxes to the government. Their exact tax burden is specified below. We will abstract from depreciation of broad capital and implicitly assume that the capital stocks decay at the same rate which is taken to be zero. Thus, gross investment equals net investment $i_t = \dot{k}_t$ This simplifies the analysis and also follows the authors on broad capital mentioned above. The capitalists' income net of taxes is then spent on consumption c_t and investment i_t.

2.2 Firms

The *firms* operate in a perfectly competitive environment and maximize profits. The capital owners rent capital to and demand shares of the firms, which are collateralized one-to-one by capital. The markets for assets, capital and labour clear at each point in time so that the firms face a path of uniform, market clearing rental rates for *broad capital* and labour, r_t and w_t. Given perfect competition the firms rent broad capital and hire labour in spot markets in each period. Output serves as numéraire and its price set equal to 1 at each date, implying that the price of (broad) capital, k_t, in terms of overall consumption stays at unity.

Aggregate production is constant returns to scale in broad capital and labour inputs.

[8] One easily verifies that the paper's results would also carry over when working with a narrow concept of capital. But as has, for instance, been pointed out by Mankiw, Romer, and Weil (1992) or Barro and Sala–i–Martin (2004), ch.2, convergence of growth rates across countries would require a larger capital share than the one conventionally used in growth research.

Since the labour input equals one, k_t can also be interpreted as the capital-labour ratio. The production function $f[k_t]$ for the representative firm is assumed to be increasing and strictly concave in k_t with $\lim_{t\to\infty} f'[k_t] = 0$ and $\lim_{t\to 0} f'[k_t] = \infty$. Profit maximization implies

$$r_t = f'[k_t] \quad \text{and} \quad w_t = f[k_t] - r_t \cdot k_t \tag{2}$$

and perfect competition and the free entry and exit of firms means that profits, $f[k_t] - r_t k_t - w_t$, are zero.

Recall k_t is broadly defined so that the share of (broad) capital, denoted by α and satisfying $\alpha = (r_t k_t)/f(k_t)$, is generally larger than one half. Hence, the capitalists (owners of human and physical capital) have a higher gross income share than the low-skilled workers.[9]

2.3 Workers

The (unskilled) *workers* do not invest and are not taxed by assumption. They supply one unit of raw labour inelastically at each date and derive utility from consuming their entire wage and transfer income. Their total income x_t depends on wage income, w_t, and lump-sum transfers granted by the government, TR_t, i.e.

$$x_t = w_t + TR_t. \tag{3}$$

Their intertemporal utility is given by $\int_0^\infty v[x_t]\, e^{-\rho t} dt$ where $v[x_t]$ need not be the same as that of the capitalists, but it is also assumed to satisfy $v' > 0, v'' < 0$ and the conditions $\lim_{x_t\to\infty} v' = 0$ and $\lim_{x_t\to 0} v' = \infty$ where $v' = \frac{dv[x_t]}{dx_t}$ and $v'' = \frac{d^2 v[x_t]}{dx_t^2}$.

2.4 Government

As in Judd (1985) and Lansing (1999), I rule out a market for government bonds. A missing bond market may be a justified for an analysis that focuses on the long-run

[9]For example, Barro and Sala–i–Martin (2004), p.496, find that the estimates for convergence according to the Solow model "... are around 2–3 percent per year in the various contexts. This slow speed of convergence implies that it takes 25–35 years to eliminate one-half of an initial gap in per capita incomes. This behavior deviates from the quantitative predictions of the neoclassical growth model if the capital share is close to one-third. The empirical evidence is, however, consistent with the theory if the capital share is around three-quarters."

and where some form of Ricardian Equivalence holds. Thus, the paper concentrates on pure redistribution financed by real resources in a two-class model, and abstracts from the intricate issues associated with an analysis of the link between public debt, pure redistribution and capital accumulation.

By assumption the government collects income taxes on capital at the rate θ_t to grant an investment subsidy $(p_t \dot{k}_t)$ to the capital owners and use the remaining resources T_t for lump-sum transfers TR_t to the workers.[10] The government runs a balanced budget and pursues an optimization problem formulated below. The government budget is given by

$$TR_t = \theta_t r_t k_t - p_t \dot{k}_t = T_t. \tag{4}$$

Therefore the returns to human and physical capital are taxed equally. Capital income taxes in this model are then really an equal tax on all the returns from accumulated factors of production. By arbitrage these returns are the same for the capital stocks. Hence, in terms of the tax package considered a capital-income-cum-investment-subsidy-tax (CICIST) scheme is contemplated.

Notice that no strong restrictions on the domain of θ_t and p_t are invoked at this stage of the analysis. Thus, it is possible that with negative p_t the investment subsidy is an investment tax really, and a negative θ_t would imply a capital income subsidy. But it turns out that both p_t and θ_t cannot be larger than one. This is true in the equilibria below and assumed for meaningful derivations in this model from now on.[11]

[10] As $c_t = (1-\theta_t) r_t k_t - \dot{k}_t + p_t \dot{k}_t$, the term $p_t \dot{k}_t$ may be interpreted as a form of politically determined capital depreciation allowance which is directly and positively related to the amount invested. See, for instance, Jones, Manuelli, and Rossi (1997) and Guo and Lansing (1999) who show that investment subsidies may take the form of accelerated depreciation. The consequences for an optimal tax policy with the special design of investment subsidies in the form of accelerated depreciation allowances is analyzed in detail in a companion paper.

[11] Thus, the support of θ is $\theta \in (-\infty, 1)$ which is simply assumed to hold from now on. An earlier version of this project restricted to choice of θ to be non-negative. See Rehme (2011). Thus, allowing for a larger potential solution space offers interesting new insights in the present paper.

2.5 The Private Sector.

Inserting the tax T_t to be paid by the capital owner in equation (4) into equation (1) and rearrangement of the resulting expression yields the capital owner's budget constraint

$$\dot{k}_t = \frac{(1-\theta_t)\; r_t\; k_t - c_t}{(1-p_t)}. \tag{5}$$

From that one can then state the *capital owner's* problem as

$$\max_{c_t} \int_0^\infty u[c_t]\; e^{-\rho t} dt$$

subject to the budget constraint in equation (5) and a given initial capital stock $k(0) = k_0$.

We solve that intertemporal problem by dynamic programming, which always yields time-consistent solutions. Thus, the capitalist's problem is then

$$V_c[k_0] \equiv \max_{c_t} \int_0^\infty u[c_t] e^{-\rho t} dt \quad \text{s.t. } \dot{k}_t = \frac{(1-\theta_t)\; r_t\; k_t - c_t}{(1-p_t)}, \quad k(0) = k_0,$$

where the capital owner takes policy as given.

For this autonomous problem the *value function* of the capital owners, indexed by subscript c, is simply denoted by $V_c(k_t)$, with the understanding that it depends on time as well, and must satisfy the *Hamilton-Jacobi-Bellman (HJB)* equation

$$\rho V_c[k_t] = \max_{c_t} \left\{ u[c_t] + V_c'[k_t] \left(\frac{(1-\theta_t) r_t k_t - c_t}{(1-p_t)} \right) \right\},$$

where the value function is assumed to be twice continuously differentiable, and the usual initial as well terminal boundary point conditions are imposed. For convenience let $V_c'[k_t] \equiv \partial V_c[k_t]/\partial k_t$ and $V_c''[k_t] \equiv \partial^2 V_c[k_t]/\partial k_t^2$ and note these expressions are also functions of time t.[12]

As is well-known, the first order necessary conditions for this problem involve the

[12]This form of the HJB equation for our autonomous problem with exponential discounting follows, for example, Kamien and Schwartz (1991), sec. 21, or Acemoglu (2009), ch. 7.

following equations to be met,

$$u'[c_t] = \frac{V_c'[k_t]}{1-p_t} \tag{6}$$

$$\rho V_c'[k_t] = V_c'[k_t]\left(\frac{1-\theta_t}{1-p_t}\right) r_t + V_c''[k_t]\left(\frac{(1-\theta_t) r_t k_t - c_t}{(1-p_t)}\right) \tag{7}$$

plus an initial and a transversality condition.

Important here is equation (6), because it implicitly defines the capital owners' optimal *decision rule*. The latter is of the *feedback* type, because, in the most general case, it depends only on the current state k, and the current policy variables θ and p. Thus, in the optimum for the capitalist we have in the the most general case that $c_t = f[k_t; p_t, \theta_t]$. Notice, however, that here the *decision rule* is not directly dependent on θ_t. In fact, at time t it does not depend the the current tax rate so that in the present case $c_t = c[k_t; p_t]$ really captures the capital owner's optimal decision rule.

It is important to reiterate that the capitalist follows a rule for the state of the problem, that is, a rule following the state variable k_t. The policy variables θ_t and p_t are parameters and beyond the control of the capital owner. Then[13]

$$\frac{dV_c'[k_t]}{dt} = V_c''[k_t] \cdot \dot{k}_t = V_c''[k_t]\left(\frac{(1-\theta_t) r_t k_t - c_t}{(1-p_t)}\right).$$

Thus, we get

$$\frac{\dot{V}_c'[k_t]}{V_c'[k_t]} = \rho - \left(\frac{1-\theta_t}{1-p_t}\right) r_t \tag{8}$$

which is the *Euler* equation, showing how agents evaluate the evolution of their capital stock in terms of their welfare.

Equation (6) implies $u'[c_t](1-p_t) = V'[k_t]$. Taking time derivatives yields

$$u''[c_t] \cdot \dot{c}_t \cdot (1-p_t) - u'[c_t] \cdot \dot{p}_t = \dot{V}_c'[k_t].$$

[13] Notice that under the optimal choice and state variable solutions there is a link between a Hamiltonian and the HJB. Furthermore, in the general case $V_c'[k_t] = \lambda_t$, where λ_t is commonly interpreted as the costate variable, representing the *shadow* value of more capital. Of course, the same interpretation applies to $V_c'(k_t]$ for the optimal k_t in the dynamic programming problem. Thus, equation (8) represents the evolution the shadow value of more capital.

Dividing by $u'[c_t](1-p_t) = V'_c[k_t]$ and using equation (8) establishes

$$\begin{aligned}
\frac{u''[c_t]}{u'[c_t]\cdot(1-p_t)}\cdot\dot{c}_t\cdot(1-p_t) - \frac{u'[c_t]\cdot\dot{p}_t}{u'[c_t]\cdot(1-p_t)} &= \frac{\dot{V}'_c[k_t]}{V'_c[k_t]} = \rho - \left(\frac{1-\theta_t}{1-p_t}\right) r_t \\
\dot{c}_t\cdot\left(\frac{u''[c_t]}{u'[c_t]}\cdot\frac{c_t}{c_t}\right) - \frac{\dot{p}_t}{1-p_t} &= \rho - \left(\frac{1-\theta_t}{1-p_t}\right) r_t \\
\frac{\dot{c}_t}{c_t}\cdot\left(\frac{u''[c_t]}{u'[c_t]}\cdot c_t\right) &= \rho - \left(\frac{1-\theta_t}{1-p_t}\right) r_t + \frac{\dot{p}_t}{1-p_t} \\
\frac{\dot{c}_t}{c_t} &= \eta\left\{\left(\frac{1-\theta_t}{1-p_t}\right) r_t - \frac{\dot{p}_t}{1-p_t} - \rho\right\}
\end{aligned}$$

where $\eta[c_t] \equiv -\frac{u'[c_t]}{c_t\cdot u''[c_t]} > 0$ denotes the capitalists' intertemporal elasticity of substitution.

Lastly, one has to check for the terminal, that is, the transversality condition of the optimum. To that end and for our case we simply require

$$\lim_{t\to\infty} V'[k_t]\cdot k_t\cdot e^{-\rho t} = 0$$

which is satisfied because k_t approaches a positive constant in the steady state and from equation (6) we know that $V'[k_t] = u'[c_t](1-p_t)$ which is finite and positive for $c_t > 0$, $p < 1$. But then the transversality condition is indeed satisfied.[14]

Consequently, the time-consistent household optimum is characterized by

$$\dot{k}_t = \frac{(1-\theta_t)r_t k_t - c}{(1-p)}, \tag{9}$$

$$\dot{c}_t = c_t\cdot\eta[c_t]\cdot\left[\left(\frac{1-\theta_t}{1-p_t}\right) r_t - \frac{\dot{p}_t}{1-p_t} - \rho\right]. \tag{10}$$

In general the consumption growth rate is not necessarily constant, when policy changes, that is, when p_t and θ_t move over time. That is interesting, but most private sector agents expect the tax code and policy not to change systematically over time. For realism's sake that is what is assumed below.

When the factor and goods markets clear the representative (unskilled) worker's

[14] That the particular transversality condition used above is applicable, when k approaches a constant value in the steady state, is based on, for example, Acemoglu (2009), theorem 7.12 and 7.13.

income and so consumption is

$$x_t = w_t + TR_t = f[k_t] - r_t k_t + \theta_t r_t k_t - p_t \dot{k}_t$$

where equations (2) and (4) have been used. In equilibrium the overall resource constraint is such that the agents satisfy their budget constraints. By substitution of equation (5) into the expression for x_t above one then obtains[15]

$$\begin{aligned} x_t =& f[k_t] - (1-\theta_t) r_t k_t - p_t \left(\frac{1-\theta_t}{1-p_t} \right) r_t k_t + \frac{p_t \, c_t}{1-p_t}, \\ =& f[k_t] - \left(\frac{1-\theta_t}{1-p_t} \right) r_t k_t + \frac{p_t \, c_t}{1-p_t}. \end{aligned} \tag{11}$$

As a consequence in equilibrium the income of the representative unskilled worker is increasing in the consumption of the owner of broad capital, for given p_t.

2.6 Arbitrary Behaviour and Policy

Suppose the agents and the government act in arbitrary (unspecified) ways at a particular point in time.[16] Then it is straightforward to show the following for instantaneous per-period reactions:

1. An increase in investment subsidies does not seem to raise the workers' after-tax wages and so consumption. This appears to make p_t a non-effective redistributive policy tool.
2. An increase in the capital income tax net of investment taxes seem to be a positive (pure) redistribution device as they raise after-tax wages.
3. Higher capital income taxes seem to imply lower capital investment.
4. An increase investment subsidies may imply more capital investment and so more capital accumulation.

For the associated algebra see Appendix B. These results apply when behaviour and policy are unspecified and when agents and the government do not necessarily

[15] The second line of this equation follows when one multiplies the second term on the right hand side in the first line by $(1-p)/(1-p)$, collects terms and rearranges.

[16] One may, of course, argue that relative to an infinite time horizon a point in time may be longer than imagined if the time unit is chosen accordingly. For a similar approach see Abel (2007), p. 14.

optimize. These results hold for the short-run. When introducing optimizing behaviour below, the above results need qualification, once one looks at optima for the long run.

2.7 Non-Distortion of Accumulation

One important consequence of the Judd (1985) and Chamley (1986) result that capital income taxes be optimally zero in the long run is that the capital accumulation process is not disturbed by political interference in that case.

In the model the impact of any accumulation distortion can be inferred from the Euler equation, equation (8). We will rerender it here where the shadow value of more capital is denoted by $V_c'[k_t]$,

$$-V_c'[k_t]\left(\frac{1-\theta_t}{1-p_t}\right)r_t+\rho V_c'[k_t]=\dot{V}_c'[k_t].$$

This equation shows how agents evaluate the evolution of the state variable k_t in terms of their welfare, measured by the evolution of the shadow price $V_c'[k_t]$, which then leads them to pursue a particular accumulation programme. Policy would in general distort this evaluation which is captured by the term $\frac{1-\theta_t}{1-p_t}$.

The government does *not* distort this evaluation in a long-run equilibrium with $\dot{V}_c'[k_t]=0$ when $\theta_t=p_t,\forall t$. The crucial question is whether $\theta_t=p_t=0$ is the optimal solution, because nonnegative values for tax variables are also possible. In this paper that is analyzed that in detail below.

If $p_t=\theta_t$, then the tax arrangement with non-zero θ_t reduces to a tax on the capital owner's consumption. As is well known, consumption taxes are not distorting accumulation. To see this consider equation (5) where $\dot{k}_t=r_tk_t-\frac{c_t}{1-\theta_t}$ when $p_t=\theta_t$. The taxes are then tantamount to taxing c_t. In that sense, a policy with $p_t=\theta$ is equivalent to synthetic, non-distorting consumption tax on the capital owners.

3 The Optimal Long-Run Investment Subsidy and Capital Income Tax

Consider a benevolent government that respects the private sector's problem and/or its optimality conditions and represents the agents' interests by attaching weights to their welfare. By assumption the government is as impatient as the private sector and so has the same rate of time preference as the agents.

Let $\gamma \in (0, \infty)$ represent the social weight attached to the welfare of the representative worker, $v[x_t]$, relative to that of the capitalist, $u[c_t]$.[17] If $\gamma \to 0$, the government is only concerned about the representative capitalist, whereas it only cares about the representative worker when $\gamma \to \infty$. By assumption the government inherits a capital income tax rate $\theta(0)$ that is less than one and takes this as given at time zero. This makes the tax problem nontrivial.[18]

The government keeps the agents on their respective supply and demand curves, and chooses a policy that can be realized as a competitive equilibrium *in quantities*. Thus, the government is taken to choose its policy instruments, but lets the market determine the path of the (pre-tax) return on capital. The solution of the government's problem is then compatible with a private ownership competitive equilibrium *in quantities*.[19] Hence, I relate to the dual approach for solving Ramsey tax problems as described in, for example, Sargent and Ljungqvist (2004), ch. 15.3. Similar approaches are used by Judd (1985), Judd (1999), and Lansing (1999).

Furthermore, from now on time subscripts are dropped for convenience whenever it is clear that a particular variable depends on time, and variables that attain a long-run equilibrium position such as a steady state are indexed by an asteristk $(^*)$.

Unless stated otherwise the government's problem is to choose a policy pair (θ, p), where the choice of θ is in principle unrestricted. Furthermore, in line with most tax provisions and as the focus of the analysis is on the long run, I only concentrate on optimal long-run policies that are constant over time. Thus, the focus is on situations where $\dot{\theta} = \dot{p} = 0$, because that is what most agents expect to be the case for the long-run.

Consider now a dynamic game between the private sector and the government. The latter moves first by announcing a policy which the follower, that is, the private sector takes into account when making its decision. Thus, the government is the Stackelberg leader, and the private sector is the Stackelberg follower. For an introduction into

[17]The model's normalization implies that we consider a representative (broad) capital owner and (unskilled) worker each. As a stronger microfoundation of the political process is beyond the scope of the paper, I follow the common procedure to attach fixed (exogenous) weights on the representative agents' welfare. For a similar setup see, for example, Lansing (1999), p. 432.

[18]The assumption rules out taxing the initial capital stock via a so-called capital levy that would constitute a lump sum tax, since initial capital is in fixed supply. See Judd (1985), and Chamley (1986) or, for example, Sargent and Ljungqvist (2004), ch. 15.3.

[19]This builds on Jones (1965), Atkinson and Stiglitz (1989), lec. 6, and Turnovsky (2000), ch. 12.6. Here I follow Turnovsky's setup.

these kinds of games see, for example, Dockner, Jorgensen, Long, and Sorger (2000) and Long (2010).

Judd and Chamley focussed on open-loop-Stackelberg equilibria in their setups. The equilibria of such games do require the assumption that the government can commit itself to a policy announced at the outset of the game. As is known, that may yield time inconsistency issues. See, for example, Kemp, van Long, and Shimomura (1993), Xie (1997) and others.[20]

In turn and building on the contribution by Kemp, van Long, and Shimomura (1993), consider now a game in which the government plays a feedback (Markov-perfect) strategy $\theta = \theta(k)$ and $p = p(k)$, which is time-consistent by construction. The representative capitalist treats θ, p and r as known and given functions of time. The equilibria of such games do not require the assumption that the government can commit itself to a policy announced at the outset of the game.

Following Kemp, van Long, and Shimomura (1993), p. 421, assume that the private sector optimum in equations (5) and (10), that is,

$$\dot{k} = \frac{(1-\theta)rk - c}{(1-p)} \quad \text{and} \quad \dot{c} = c \cdot \eta[c] \cdot \left[\left(\frac{1-\theta}{1-p} \right) r - \frac{\dot{p}}{1-p} - \rho \right]$$

forms a system that has a unique stationary saddle point. The stable path through that point, say $c[k]$, is the equilibrium consumption path of the capitalist. Thus, the optimal (feedback) decision rule of the capital owner *most generally* satisfies $c = c[k; \theta, p]$ where policy is taken parametrically by the household and policy is constant, i.e. $\dot{\theta} = \dot{p} = 0$.[21]

The dynamic problem facing the government is then

$$V_g(k_0) \equiv \max_{\theta, p} \int_0^\infty \{\gamma v[x] + u[c[k; \theta, p]]\} e^{-\rho t} dt \tag{12a}$$

$$s.t. \quad v[x] = v\left[f(k) - \left(\frac{1-\theta}{1-p} \right) rk + \frac{p \cdot c[k; \theta, p])}{1-p} \right] \tag{12b}$$

$$\dot{k} = \frac{(1-\theta)rk - c[k; \theta, p]}{(1-p)} \quad \text{and} \quad k(0) = k_0 \tag{12c}$$

[20]Clearly, the classic paper for the problem of time-inconsistency is Kydland and Prescott (1977). For an analysis of the open-loop setup of the problem under study see, for example, Rehme (2007).

[21]Note that in our present setup the capital owners' decision rule at time t does not depend on θ at time t. To keep the arguments as general as possible I formulate the general problem first and then use the specific decision rule when analyzing any optima.

For this autonomous problem it is known that the value function of the government, denoted by $V_g(k)$, satisfies the Hamilton-Jacobi-Bellman (HJB) equation[22]

$$\rho V_g(k) = \max_{\theta,p}\left\{\gamma v\left[f(k) - \left(\frac{1-\theta}{1-p}\right)rk + \frac{p\cdot c[k;\theta,p])}{1-p}\right]\right.$$
$$\left. + u[c[k;\theta,p]] + V_g'(k)\left(\frac{(1-\theta)rk - c[k;\theta,p]}{(1-p)}\right)\right\}. \quad (13)$$

The solution to (12) then satisfies

$$\begin{aligned}\theta[k] &\equiv \arg\max_{\theta}\left\{\gamma v[x] + u[c[k;\theta,p]] + V_g'(k)\left(\frac{(1-\theta)rk - c[k;\theta,p]}{(1-p)}\right)|p\right\}\\ p[k] &\equiv \arg\max_{p}\left\{\gamma v[x] + u[c[k;\theta,p]] + V_g'(k)\left(\frac{(1-\theta)rk - c[k;\theta,p]}{(1-p)}\right)|\theta\right\}\end{aligned}$$

where $v[x]$ is given by equation (12b). Thus, recalling Basar and Olsder (1995), p. 227/8, it is acceptable to consider $(c[k], \theta[k], p[k])$ as a feedback equilibrium with the government as the leader. As $\theta[k], p[k]$ are derived from the Hamilton-Jacobi-Bellman equation (13), the equilibrium is time-consistent.

I now simplify the analysis by assuming that $\eta[c]$ is a (positive) constant η. Again we look for constant optimal policies. A tilde over a variable will denote that we look for a solution that depends on the feedback rule $\tilde{c} = \tilde{c}[k, \theta, p]$ or, expressed in reduced form, simply $\tilde{c} = \tilde{c}[k]$. Thus, the capitalists' consumption follows a rule and is not a state variable anymore in the subsequent problem.

Then the current-value Hamiltonian associated with the problem formulated in equation (12), called $\mathcal{H}^F$, is then[23]

$$\mathcal{H}^F = \gamma v[\tilde{x}] + u[\tilde{c}] + \nu\cdot\tilde{\Delta} \quad \text{where} \quad \tilde{\Delta} \equiv \left[\frac{(1-\theta)\, rk - \tilde{c}[k,\theta,p]}{(1-p)}\right],$$

$\dot{k} = \tilde{\Delta}$, and $\tilde{c} = \tilde{c}[k,\theta,p]$, as well as $\tilde{x} = f[k] - \frac{1-\theta}{1-p}\cdot rk + \frac{p\tilde{c}[\cdot]}{1-p}$.

Here the control variables are θ and p. The single state variable is k.[24] The first

[22] Again, as in Kamien and Schwartz (1991), section 21, assume that $V_g(k)$ is twice continuously differentiable, or that it satisfies the conditions for the infinite horizon leading to Theorem 3.4 in Dockner, Jorgensen, Long, and Sorger (2000).

[23] Here the solution procedure in Kemp, van Long, and Shimomura (1993), sec. 4, is followed.

[24] The feedback setup eliminates c as a state variable. This is an important point because in open-loop formulations that is one reason one may obtain time-inconsistent solutions. Also the present problem is quite different from a paternalistic command (nudge) approach of the government which may also yield a time-consistent solution. That is analyzed in a companion paper.

order conditions for θ, p and k are

$$\gamma v'[\cdot] \cdot \tilde{x}_\theta + u'[\cdot] \cdot \tilde{c}_\theta + \nu \cdot \tilde{\Delta}_\theta = 0, \tag{14a}$$

$$\gamma v'[\cdot] \cdot \tilde{x}_p + u'[\cdot] \cdot \tilde{c}_p + \nu \cdot \tilde{\Delta}_p = 0, \tag{14b}$$

$$\gamma v'[\cdot] \cdot \tilde{x}_k + u'[\cdot] \cdot \tilde{c}_k + \nu \cdot \tilde{\Delta}_k = \rho\nu - \dot{\nu}. \tag{14c}$$

plus the requirements that

$$\dot{k}_t = \frac{(1-\theta)rk - \tilde{c}[k,\theta,p]}{(1-p)} \quad \text{and} \quad \lim_{t\to\infty} \nu k e^{-\rho t} = 0.$$

The co-state variable ν denotes the shadow price of capital for the government. The requirement for $\dot{k}$ is, of course, that the budget constraint be obeyed, and the second one captures the transversality condition, ruling out asymptotic left-overs or running infinite debt. The partial derivatives of $\tilde{x}, \tilde{c}$ and $\tilde{\Delta}$ are derived in Appendix C.

As we are interested in finding constant optimal policies for the long run we evaluate the first order conditions in a steady state. In such a state we must then have $\dot{k} = \dot{\nu} = 0$. For a long-run equilibrium we also require that $\dot{c} = 0$ in equation (10).

In Appendix C.2 the partial derivatives are derived that are necessary to evaluate the first order conditions in a steady state. From that one verifies that the equations (14a) to (14c) evaluated in a steady state, indexed by an asterisk, become

$$\begin{aligned} \gamma v'[\cdot]\left[\frac{rk^* + p\cdot\tilde{c}^*_\theta}{1-p}\right] + u'[\cdot]\cdot\tilde{c}^*_\theta - \nu\left[\frac{rk^* + \tilde{c}^*_\theta}{1-p}\right] &= 0 \\ \gamma v'[\cdot]\left[\frac{rk^*}{1-p}\right] - \nu\left[\frac{rk^*}{1-p}\right] &= 0 \end{aligned} \tag{15a}$$

$$\gamma v'[\cdot]\left(\frac{p\cdot\tilde{c}^*_p}{1-p}\right) + u'[\cdot]\cdot\tilde{c}^*_p - \nu\cdot\left(\frac{\tilde{c}^*_p}{1-p}\right) = 0 \tag{15b}$$

$$\begin{aligned} \gamma v'[\cdot]\left\{f' - \frac{(1-\theta)}{1-p}r + \frac{p\cdot\tilde{c}^*_k}{1-p}\right\} + u'[\cdot]\cdot\tilde{c}^*_k & \\ +\nu\cdot\left[\frac{(1-\theta)}{1-p}\cdot r - \frac{\tilde{c}^*_k}{1-p}\right] &= \rho\nu. \end{aligned} \tag{15c}$$

The second line in equation (15a) follows from the capital owners' decision rule which is independent of θ and so $c_\theta = 0$. One also verifies that in the steady state $\dot{k} = 0$ implies $\tilde{c}^* = (1-\theta)rk^*$.

The first steps of the solution to this problem are then pretty straightforward. From equation (15a) one obtains $\gamma v'[\cdot] = \nu$ in the optimum. Substituting this into equation (15b) implies that $\gamma v'[\cdot] = u'[\cdot]$ is optimal as $\tilde{c}_p^*$ is non-zero.

Using these results in equation (15c boils down to the requirement that $f'[k^*] = \rho$. But from profit maximization we know that $f' = r$. Thus, in the long-run optimum $r^* = \rho$. For a steady state equilibrium one also needs that $\dot{c} = 0$ in equation (10). But that necessitates

$$\frac{1-\theta}{1-p} \cdot r^* = \rho \Rightarrow \frac{1-\theta}{1-p} \cdot r^* = r^* \Rightarrow \frac{1-\theta}{1-p} = 1$$

which is only satisfied when $\theta = p$. At this stage of the analysis these variables can take any value $p, \theta < 1$. That then implies the following

Proposition 1 *No matter whether the government is relatively more pro-labour or pro-capital, the optimal, time-consistent (feedback Stackelberg) policy under the capital-income-cum-investment-subsidy-tax (CICIST) scheme is not to distort capital accumulation by setting $\theta = p$.*

By the proposition a government can use different distortionary instruments to offset any distortions. Here it is the coupling of the investment subsidies, which are potentially growth enhancing, with capital income taxes, which distort growth. The result implies that in the optimum the instrument mix removes the distortion. However, different optimal values for $\theta = p$ are possible and that is now analyzed in turn.

3.1 Is $\theta = 0$ with $p = 0$ an optimum?

In the long-run optimum $\gamma v'[x^*] = u'[c^*]$ where $x^* = f[k^*] - r^* k^*$ and $c^* = \rho k^*$ with $r^* = \rho$ when $\theta = p = 0$. An optimum with $\theta = 0$ as a feedback solution would then require

$$\gamma v'[f[k^*] - r^* k^*] = u'[\rho k^*]$$

where, again, k^* is determined by $f'[k^*] = r^* = \rho$. But an optimum that simultaneously satisfies $\theta = 0$ with $p = 0$ and the last equation is generically impossible.

Proposition 2 *A time-consistent, feedback Stackelberg solution with $\theta = 0 = p$ as a long-run policy optimum in a steady state equilibrium does generically not exist.*

The result, therefore, casts doubt on the general validity of the *celebrated* Judd-Chamley result and corroborates earlier findings such as Long and Shimomura (2000), Long and Shimomura (2002) and others.

3.2 Is $\theta = p \neq 0$ an optimum?

In the long-run optimum the social marginal utilities of the agents must be equated. As was shown in the previous subsection a solution with $\theta = 0$ is generically impossible. Thus, we now concentrate on solutions with $p = \theta$ being non-zero. We know the optimum must satisfy $\gamma v' = u'$. With $p = \theta$ and $f'[k^*] = r^* = \rho$ it amounts to the condition

$$\gamma v'[\tilde{x}^*] = u'[\tilde{c}^*] \quad \text{where} \quad \begin{aligned} \tilde{c}^* &= (1-\theta)r^*k^* = (1-\theta)\rho k^* \\ \text{and} \quad \tilde{x}^* &= f(k^*) - rk^* + \frac{\theta \tilde{c}^*}{1-\theta} = f(k^*) - r^*k^* + \theta\rho k^*. \end{aligned}$$

Hence, the optimal tax rate θ, when $p = \theta$, is implicitly determined by

$$\gamma v'[f(k^*) - rk^* + \theta\rho k^*] = u'[(1-\theta)\rho k^*] \tag{16}$$

and denoted by $\widetilde{\theta}^*$ where the tilde indicates the feedback solution and the asterisk that this would be a long-run optimum. Of course, the feedback solution is *time-consistent* by construction, and the optimal tax rate $\widetilde{\theta}^*$ is implicitly determined by the model's parameters.

It remains to check whether the transversality condition is met for the conditions identified so far. Notice that the transversality condition can be written as

$$\lim_{t\to\infty} \nu k e^{-\rho t} = \lim_{t\to\infty} \nu \cdot \lim_{t\to\infty} k \cdot \lim_{t\to\infty} e^{-\rho t} = 0.$$

It has been shown above that the optimum implies that ν equals $u'[\cdot]$ or $\gamma v'[\cdot]$ and is positive, because $u'[\cdot]$ and $\gamma v'[\cdot]$ are positive by assumption. In the long-run optimum we have $\gamma v'[\tilde{x}^*]$ and $u'[\tilde{c^*}]$. But then any viable optimum must rule out a tax rate such that $\tilde{c}^* = (1-\theta)\rho k^* \to 0$ implying $u'[\tilde{c^*}] \to \infty$. As a consequence for a finite ν one must have that $\widetilde{\theta}^* < 1$ in any optimum.

Similarly, any solution with $\tilde{x}^* \to 0$ must be ruled out. So $\widetilde{\theta}^*$ has to be such that $\tilde{x}^*$ is positive. That does not preclude a negative $\widetilde{\theta}^*$, however. One verifies that $\tilde{x}^* = 0$

would require a $\theta_{min} < 0$ such that $\theta_{min} \equiv \frac{(\alpha-1)f(k^*)}{\rho k^*} = \frac{(\alpha-1)r^*}{\alpha\rho} = (\alpha-1)/\alpha$ which is a negative number because $\alpha < 1$ and $r^* = \rho$ in equilibrium. Thus, for any optimum the lower bound on any optimal $\widetilde{\theta}^*$ is $\theta_{min} < 0$ so that $\widetilde{\theta}^* > \theta_{min}$.

But if the optimal $\widetilde{\theta}^*$ is bounded in the way shown above then $0 < \nu < \infty$ as $t \to \infty$ (in the long-run steady state) so that the transversality condition is indeed satisfied.

Summarizing, in the optimum the long-run $\widetilde{\theta}^* = \widetilde{p}^* \in (\theta_{min}, 1)$. Furthermore, total differentiation of equation (16) reveals that $\widetilde{\theta}^*$ is increasing in γ.[25]

Proposition 3 *In an interior, time-consistent feedback Stackelberg optimum the following holds for the steady state equilibrium.*

1. *Generically the social marginal utilities of the worker's and capitalist's consumption are equated in a long-run equilibrium, and the optimal, long-run capital income tax $\widetilde{\theta}^*$ solves*

$$\gamma v'[f(k^*) - r^*k^* + \theta\rho k^*] = u'[(1-\theta)\rho k^*].$$

   where $r^* = \rho$, and $k^* = k^*[\alpha, \rho]$.

2. *The long-run, optimal investment subsidy satisfies $\widetilde{p}^* = \widetilde{\theta}^*$.*

3. *The optimal, long-run time-consistent solution does not distort accumulation and turns the capital income tax scheme into a (synthetic) consumption tax scheme.*

4. *Depending on the strength of the reaction of the marginal utilities of consumption to taxes, the optimal, long-run capital income tax rate $\widetilde{\theta}^*$, can be positive or negative,.*

$$\widetilde{\theta}^* \in (\theta_{min}, 1) \quad \textit{where} \quad \theta_{min} = \frac{(\alpha-1)f(k^*)}{\rho k^*} = \frac{(\alpha-1)r^*}{\alpha\rho} = (\alpha-1)/\alpha < 0.$$

   and depends in an important way on the parameters characterizing the economy.

[25] The total differential implies

$$v'[\cdot]d\gamma + (\gamma v''[\cdot] + u''[\cdot])\,\rho k^* d\theta = 0 \quad \Leftrightarrow \quad \frac{d\gamma}{d\theta} = -\frac{(\gamma v''[\cdot] + u''[\cdot])\,\rho k^*}{v'[\cdot]} > 0.$$

The expression is positive because $v''[\cdot]$ and $u''[\cdot]$ are negative and the other terms positive.

5. *The optimal tax rate* $\widetilde{\theta}^* = \widetilde{\theta}^*[\gamma, \alpha, \rho]$ *is higher when the social weight on the marginal utility of the workers' consumption* γ *is higher.*

The results do not depend on production externalities or any other things, the capital income taxes may be used for, except for paying for an investment subsidy scheme.

Thus, investment subsidies are good for the workers in an *indirect* way, because, with them there exists the possibility of positive redistributive effects of capital income taxes, namely when the optimal capital income tax is positive. These effects become strongest in a long-run equilibrium when the investment subsidy rate equals the capital income tax rate. In that sense investment subsidies are an indirect redistributive device in the long run.

Clearly, the result is in contrast to the model's predictions for the short run and arbitrary behaviour. In the short run p may be a bad instrument for redistribution. Hence, the effects of the policy instrument p depends on behaviour and the time horizon.

The fact that the optimal tax rate is increasing in γ appears quite realistic. Thus, as the workers's welfare gets more social weight the government would choose higher capital income taxes in this tax scheme with investment subsidies.[26]

An important implication of the model is that the long-run optimal capital income tax rate can be negative. This is, for example, the case if γ is low. Thus, political preferences are not important when considering the policy distortion on accumulation. Irrespective of γ it is optimal not to distort accumulation. But then the often identified redistribution-accumulation trade-off is separated out in this setup. As a consequence the benefits of a non-distorted economy are redistributed, and that redistribution can go either way, to the workers or the capital owners.[27]

In summary, an economy's deep structural parameters determine the optimal policy mix for the long run. Investment subsidies and non-zero capital income taxes are found to be optimal for the long-run. Redistributive taxation depends in an important way on social preferences and other deep structural variables characterizing the economy.

[26]Notice that the steady state capital stock k^* would be the same under any other capital income tax scheme for which it is shown that the long-run capital income tax should be zero. This is an important point, because overall welfare (sum of utilities) may be higher under the present tax scheme in comparison to those other capital income tax schemes.

[27]Of course, that result depends on the assumption of an inelastic labour supply, just as assumed in the previous research this paper relates to.

4 A parametrization

Proposition 3 provides a general result for the tax scheme under study. Clearly the optimal tax rate will depend on many parameters. To consider the possible effects of the latter consider an example based on the observation that it is not entirely clear why workers should evaluate a consumption good any differently than a capital owner. For that reason now assume that the two representative agents have the same form of the utility function, which is taken to be of the often-used and well-known constant intertemporal elasticity of substitution (CIES) type,

$$u[c] = \frac{c^{1-\frac{1}{\eta}} - 1}{1 - \frac{1}{\eta}} \quad \text{and} \quad v[x] = \frac{x^{1-\frac{1}{\eta}} - 1}{1 - \frac{1}{\eta}}$$

with η as the constant intertemporal elasticity of substitution. Then equation (16) boils down to

$$\gamma \left[f[k] - \rho k^* + \theta \rho k^* \right]^{-\frac{1}{\eta}} \quad = \quad \left[(1-\theta) \rho k^* \right]^{-\frac{1}{\eta}}$$

where $\rho = r^*$. This equation is not easily solvable, but clearly the optimal solution is a function of the form $\widehat{\theta} = \theta[\gamma, k^*, \eta, \rho]$.

Notice that broad capital k^* is a function of parameters too. To this end assume that the production function is of the standard type $f = Ak^\alpha$ where $0 < \alpha < 1$. Then $f' = \alpha A k^{\alpha-1}$ and $f' = r$ and so $k^* = \left(\frac{\alpha A}{r}\right)^{\frac{1}{1-\alpha}}$. Thus, in steady state the output-capital ratio is given by $f^*/k^* = A(k^*)^{\alpha-1} = A\frac{r^*}{\alpha A} = \frac{r^*}{\alpha}$. With this one can rearrange the equation above and divide by k^* to get

$$\gamma^\eta = \frac{1}{\rho} \left(f^*/k^* - r^* \right) + \theta \left(1 + \gamma^\eta \right).$$

Now notice that $f^*/k^* = r^*/\alpha$. Thus, the optimal θ, called $\widetilde{\theta}^*$, satisfies

$$\widetilde{\theta}^* = \widetilde{\theta}[\gamma, \alpha, \eta, \rho] = \left\{ \gamma^\eta + \frac{r^*}{\rho} \left(1 - \frac{1}{\alpha} \right) \right\} \left(1 + \gamma^\eta \right)^{-1} \tag{17}$$

where $r^* = \rho$. As $\frac{1}{\alpha} > 1$ the sign of the optimal tax rate $\widetilde{\theta}^*$ depends in an important way on the social weight going to the workers. As it is very unlikely that $\gamma = \left(\frac{1}{\alpha} - 1\right)^{\frac{1}{\eta}}$, the long-run capital income tax rate $\widetilde{\theta}^*$ is generically nonzero.

Corollary 1 *If the agents possess the same CIES utility functions, it is optimal to set $\widetilde{\theta}^* = \widetilde{p}^*$ in a long-run feedback Stackelberg equilibrium under a capital-income-cum-investment-subsidy-tax scheme (CICIST).*

- *The value of the optimal, long-run capital income tax rate $\widetilde{\theta}^*$ is generically non-zero.*
- *The optimal, long-run capital income tax rate is positive, i.e. $\widetilde{\theta}^* > 0$, if $\gamma > \left(\frac{1}{\alpha} - 1\right)^{\frac{1}{\eta}}$.*
- *Otherwise, the optimal long-run capital income tax rate is negative, i.e. $\widetilde{\theta}^* < 0$, if $\gamma < \left(\frac{1}{\alpha} - 1\right)^{\frac{1}{\eta}}$.*

In Appendix D it is then shown that $\widetilde{\theta}^*$ depends on the parameters as follows.

Table 1: The reaction of $\widetilde{\theta}^*$ to changes in parameters

α	ρ	$\beta_{\vert\gamma>1}$	$\beta_{\vert\gamma<1}$	γ
$+$	0	$+$	$-$	$+$

The signs are derived in Appendix D.

Corollary 2 *Under CICIST and when the agents all have the same CIES utility functions, the optimal, long-run capital income tax rate $\widetilde{\theta}^*$ in a feedback Stackelberg equilibrium is*

- *higher the higher the share of capital (α) is. Thus, the income distribution matters.*
- *higher, the more impatient (ρ) or the more willing the agents are to exchange current for future consumption (η). Thus, private-sector-preferences matter.*
- *higher, the more social weight is attached to the marginal utility of the workers' consumption (γ). Thus, public-sector-preferences matter.*

Thus, under the capital income tax scheme under consideration (CICIST) distributional and preference parameters matter for the long-run equilibrium, and that may complement the results of Judd (1985) and Chamley (1986). Importantly, the results establish that there capital income taxes are optimally non-zero in the long run when coupled with investment subsidies.

In order to get a feeling for the nature of the solutions I have conducted a numerical simulation exercise based on calibrations from the business cycle and other literature. They are presented in Appendix E.1 and reveal that, unsurprisingly, the weight γ plays a crucial role for the value of the optimal tax rate.

The simulation is able to mimic income tax rate values that apply today and in history, particularly for the U.S. Naturally the simulated numbers crucially depend on all the parameters too. In that sense the numerical exercise highlights a dependency of any non-zero capital income tax rates on social and other parameters.

5 Conclusion

This paper analyzes whether investment subsidies are bad instruments for redistribution. When coupling the latter with capital income taxes it is found that an increase in investment subsidies is a bad tool for redistribution, but good for economic growth, when the private sector and the government act non-optimally in the short run. In turn, capital income taxes are bad for economic growth and good for redistribution under these conditions. Furthermore, it is shown that allowing for more investment subsidies or setting lower capital income tax rates may stabilize the real return to investment in an economic downturn.

In contrast, for the long run and with optimizing behaviour things are quite different. The coupling of capital income taxes with investment subsidies for financing pure redistribution implies maximal investment subsidies relative to the taxes to be paid by the investors and often nonzero capital income taxes. The optimal policy package under investigation in this paper is nondistortionary for capital accumulation. This holds for any government, regardless of its social preferences. Thus, even an entirely pro-labour government would choose this nondistortionary policy in the model.

Unlike the Chamley-Judd-results suggest, the model, therefore, implies that in a feedback Stackelberg equilibrium efficiency and equity (redistribution) concerns can separated from one another. It is then found that (pure) redistribution can go either to labour to capital, implying that the optimal, long-run capital income tax can be negative or positive, depending in an important way on the model's parameters.

Furthermore, it is found that in a Markov-perfect environment, that is, in a feedback Stackelberg equilibrium redistribution and so capital income taxes are optimally nonzero in the long run. That depends on realistic conditions for taxation policy. The most important conditions identified in this paper are: (a) the social weight of those who receive redistributive transfers, (b) the distribution and so inequality in pre-tax factor incomes, (c) the intertemporal elasticity of substitution. The results imply that pure redistribution may optimally be financed by capital income taxes when using investment subsidies as an additional and complementing instrument.

Of course, several caveats apply. It would be interesting to study the convergence properties of the model and so implications for the medium run. Furthermore, in line with the literature the paper relates to no government debt was considered. Situations where that form of Ricardian equivalence would not hold also appear interesting to analyze. These and other questions are left for future research.

A Broad capital, equation (5) and the firms

The following arguments are partly based on results presented in Barro and Sala–i–Martin (2004), ch. 1.2 and 4.2. Suppose one group of agents, called the capitalists, own (aggregate) stocks of human capital H_t and physical capital K_t at time t. The latter are used in aggregate production

$$Y_t = A \cdot F(K_t, H_t, L_t) \tag{18}$$

where $F(\bullet)$ features constant returns to scale and possesses the usual properties of a neoclassical production function. Furthermore, A is a time-invariant, i.e. constant scaling factor related to the level of technology, and L_t denotes the aggregate labour input which equals the number of workers. Given constant returns to scale in K_t, H_t and L_t the output per worker can be expressed as

$$\begin{aligned} Y_t &= L_t \cdot A \cdot F(\kappa_t, h_t, 1) \\ y_t &= A \cdot f(\kappa_t, h_t) \end{aligned} \tag{19}$$

where $y_t \equiv Y_t/L_t, \kappa_t = K_t/L_t$ and $h_t = H_t/L_t$, and the function $f(\bullet)$ has a degree of homogeneity ν in κ_t and h_t that is less than one. Furthermore, the marginal products for each factor are positive and decreasing. In a competitive environment and given profit maximization it follows that

$$\frac{\partial y_t}{\partial h_t} = r_t^h, \quad \text{and} \quad \frac{\partial y_t}{\partial \kappa_t} = r_t^\kappa, \tag{20}$$

where r_t^h and r_t^κ denote the rates of return of human and physical capital, respectively. Furthermore, let w_t denote the wage rate for low-skilled labour which corresponds the marginal product of low skilled labour of the aggregate production function (18).

By assumption the government uses the paper's policy package on income derived from both, human and physical capital income. Then the budget constraint for the capital owners becomes

$$c_t + i_t = (1 - \theta_t)(r_t^\kappa \kappa_t + r_t^h h_t) + p_t i_t \quad \text{and} \quad i_t = \dot{\kappa}_t + \dot{h}_t.$$

Assuming the same objective of the capitalists as in the main text and rearrangement of the latter equation imply that the capitalists solve

$$\max_{c_t} \int_0^\infty u[c_t]\, e^{-\rho t} dt$$
$$s.t. \quad \dot{\kappa}_t + \dot{h}_t = \Theta \cdot (r_t^\kappa \kappa_t + r_t^h h_t) - \tfrac{c_t}{1-p_t}, \quad k_0 = \text{given}.$$

where $\Theta = \left(\frac{1-\theta_t}{1-p_t}\right)$. The first order conditions for this problem involve the following two equations for the state variables κ_t and h_t

$$-\nu_t \cdot \Theta \cdot r_t^\kappa + \rho\nu = \dot{\nu}_t \quad \text{and} \quad -\nu_t \cdot \Theta \cdot r_t^h + \rho\nu = \dot{\nu}_t.$$

From this one easily verifies that for an interior solution one must have

$$r_t^\kappa = r_t^h. \tag{21}$$

Otherwise, if $r_t^\kappa < r_t^h$ capital owners would only accumulate human capital and no physical capital, or if $r_t^\kappa > r_t^h$ capital owners would only accumulate physical capital and no human capital. Denote by r_t the return to these factors that satisfies $r_t = r_t^\kappa = r_t^h$.

Equations (20) and (21) then imply that

$$\frac{\partial y_t}{\partial h_t} = \frac{\partial y_t}{\partial \kappa_t},$$

i.e. the marginal products of physical and human capital must be equal. Given the degree of homogeneity of each marginal product there will then be clear *linear* relationship between h_t and κ_t in the optimum. To see this more clearly notice that $\frac{\partial y_t}{\partial \kappa_t} = f^1(\kappa_t, h_t)$, i.e. the marginal product will in general be a function of κ_t and h_t that is homogenous of degree $\nu - 1$ in (κ_t, h_t) where $\nu < 1$. Similarly, for $\frac{\partial y_t}{\partial h_t} = f^2(\kappa_t, h_t)$ which is homogeneous of degree $\nu - 1$ in κ_t and h_t. But as these functions are homogeneous it follows that

$$\frac{\frac{\partial y_t}{\partial \kappa_t}}{\frac{\partial y_t}{\partial h_t}} = \frac{f^1(\kappa_t, h_t)}{f^2(\kappa_t, h_t)} = \frac{h_t^{\nu_1 - 1} \cdot f^{11}(\frac{\kappa_t}{h_t}, 1)}{h_t^{\nu_1 - 1} \cdot f^{22}(\frac{\kappa_t}{h_t}, 1)} = z\left(\frac{\kappa_t}{h_t}, 1\right) = 1.$$

As the $z(\bullet)$ function depends, possibly nonlinearly, on κ_t / h_t and parameters, but equals 1, the optimal relationship between κ_t and h_t will be linear.

Suppose we can focus on h_t as function of κ_t so that $h_t(\kappa_t) = \chi \cdot \kappa_t$ where χ is some constant. If this is the case and as the production function is constant returns to scale we know by Euler's Theorem, profit maximization and perfect competition that in the optimum

$$y_t \quad = \quad r_t \cdot (\kappa_t + h_t(\kappa_t)) + w_t$$

where we have used that fact that the returns and marginal products for the two capital stocks are equal, that is, $\frac{\partial y_t}{\partial \kappa_t} = \frac{\partial y_t}{\partial h_t} = r_t$. The Euler relationship follows straightforwardly from the aggregate production function (18) and when dividing through by L_t. But then it follows that there exists a rate of return, a marginal product and an alternative formulation of the production function for the composite of the two capital stocks κ_t. To this end define broad capital as $k_t \equiv \kappa_t + h_t(\kappa_t) = \kappa_t + \chi \cdot \kappa_t$. Thus, in a setup with production featuring constant returns to scale there will be a marginal product that satisfies

$$r_t = \frac{\partial y_t}{\partial k_t} = r_t^\kappa = \frac{\partial y_t}{\partial \kappa_t} = r_t^h = \frac{\partial y_t}{\partial h_t}.$$

To see the equivalence between using a model with κ_t and h_t and one with $k_t = \kappa_t + h(\kappa_t)$ consider the production function in (19). If κ_t were the measure of broad capital, then production would be describes by

$$y_t = A \cdot f(\kappa_t, h_t) = A \cdot \kappa^\nu \cdot \hat{f}\left(1, \frac{h_t}{\kappa_t}\right) = A \cdot \kappa^\nu \cdot \hat{f}\left(1, \chi\right) \tag{22}$$

where we use the fact that $h_t = \chi \cdot k_t$. Thus, κ_t would summarize the contribution of physical and human capital in production.

For the measure employed in this paper, one can proceed similarly. Recall $k_t = \kappa_t + \chi\kappa_t$. Then

$$\begin{aligned} y_t &= A \cdot f(\kappa_t, h_t) = A \cdot k_t^\nu \cdot k_t^{-\nu} f(\kappa_t, h_t) \\ &= A \cdot k^\nu \tilde{f}\left(\frac{\kappa_t}{k_t}, \frac{h_t}{k_t}\right) = A \cdot k^\nu \tilde{f}\left(\frac{1}{1+\chi}, \frac{\chi}{1+\chi}\right). \end{aligned} \tag{23}$$

It is then not difficult to verify that, given the homogeneity of the functions $\hat{f}$ and $\tilde{f}$, the productions function in (22) and (23) are equivalent and $\hat{f} = \hat{f}\left(1, \chi\right) = \tilde{f}\left(1, \chi\right) = \tilde{f}$ when factoring out $(1+\chi)^{-\nu}$ in $\tilde{f}$. Thus, the paper's production function corresponds to (23) with $0 < \nu < 1$, which is interpreted as the share of broad capital, and A is scaled such that $A = \tilde{f}^{-1}\left(\frac{1}{1+\chi}, \frac{\chi}{1+\chi}\right)$. But then these arguments justify why one can start with the setup of broad capital k_t in the main text and simply work with equation (5) if one recalls that the paper's population normalization implies that the capitalists and workers are treated as a single agent so that $h_t = H_t$, $\kappa_t = K_t$ and $L_t = 1$.

As an example consider the Cobb-Douglas case used for the numerical simulation later in the paper. For $y_t = B \cdot h_t^{\eta_1} \kappa_t^{\eta_2}$ where $0 < \eta_1 + \eta_2 < 1$, equality of the marginal products implies

$$\left(\frac{\partial y_t}{\partial h_t}=\right) \quad \eta_1\frac{y_t}{h_t}=\eta_2\frac{y_t}{\kappa_t} \quad \left(=\frac{\partial y_t}{\partial \kappa_t}\right).$$

Thus, $\frac{\eta_1}{\eta_2}\cdot\kappa_t = h_t$. Substituting this into the Cobb-Douglas production function yields

$$y_t = B\left(\frac{\eta_1}{\eta_2}\right)^{\eta_1}\kappa_t^{\eta_1+\eta_2}. \tag{24}$$

Now let $\eta_1+\eta_2=\alpha$ and normalize so that $B\left(\frac{\eta_1}{\eta_2}\right)^{\eta_1}=1$ by a suitable choice of B. Then κ_t could be indicator of broad capital.

Instead, using $k_t=\kappa_t+h_t$ as an indicator of broad capital, where $h_t=\frac{\eta_1}{\eta_2}\cdot\kappa_t$, yields

$$\begin{aligned} y_t &= B\cdot k_t^{\eta_1+\eta_2}k_t^{-\eta_1-\eta_2}h_t^{\eta_1}\kappa_t^{\eta_2} = B\cdot k_t^{\eta_1+\eta_2}\left(\frac{h_t}{k_t}\right)^{\eta_1}\left(\frac{\kappa_t}{k_t}\right)^{\eta_1} \\ &= B\cdot k_t^{\eta_1+\eta_2}\left(\frac{\eta_1}{\eta_2}\right)^{\eta_1}\left(1+\frac{\eta_1}{\eta_2}\right)^{-\eta_1-\eta_2}. \end{aligned} \tag{25}$$

One easily verifies that this is equivalent to (24). Furthermore, setting $\alpha=\eta_1+\eta_2$, where η_1 represents the share of human capital and η_2 the physical capital share, scaling B so that $B=\left(\frac{\eta_1}{\eta_2}\right)^{-\eta_1}\left(1+\frac{\eta_1}{\eta_2}\right)^{\eta_1+\eta_2}$ and by the population normalization used in the main text it follows that (25) is equivalent to what is used in the latter part of the main text.

B Derivation: Arbitrary Behaviour and Policy

Suppose the agents and the government satisfy their budget constraints but otherwise act in arbitrary (unspecified) ways at a particular point in time. When analyzing impact changes in the investment subsidies, p_t, if the agents and the government act in unspecified ways, one obtains

$$\frac{dx_t}{dp_t}_{|\theta_t,c_t} = \frac{c_t-(1-\theta_t)r_tk_t}{(1-p_t)^2}. \tag{26}$$

For non-negative growth of the capital stock, we must have $c_t \le (1-\theta_t)r_tk_t$ so that $\frac{\partial x_t}{\partial p_t}_{|\theta_t,c_t} \le 0$ in general. From that an increase in investment subsidies does not appear to be a good redistribution device as it does not seem to raise after-tax wages. This is because higher p_t means that ceteris paribus less taxes for redistribution are collected.

In turn, an increase in taxes produces

$$\frac{dx_t}{d\theta_t}_{|p_t,c_t} = \frac{r_t k_t}{(1-p_t)^2} \tag{27}$$

which is positive. Thus, capital income taxes net of investment subsidies seem to be a positive redistribution device because they would raise the after-tax wages.

Next, consider the growth and investment effects of both policy instruments, given arbitrary behaviour and given everything else. To this end consider equation (5) to find

$$\frac{d\dot{k}_t}{d\theta_t}_{|p_t,c_t} = -\frac{-r_t k_t}{(1-p_t)} \leq 0 \quad \text{and} \quad \frac{d\dot{k}_t}{dp_t}_{|\theta_t,c_t} = \frac{(1-\theta_t)r_t k_t - c_t}{(1-p_t)^2} \geq 0.$$

Thus, for reasonable policies with $\theta_t, p_t < 1$, and given everything else (like the capitalists' consumption etc.) higher capital income taxes are bad for growth and more investment subsidies are growth enhancing when the net income of the capital owners is greater than their consumption.

These derivations may justify the statements in the main text.

C Reactions of c, x, and Δ in the closed-loop game

C.1 The general reactions

The static first order condition of the capitalists requires

$$u'[c] = \frac{V_c'[k]}{1-p} \tag{28}$$

where $V_c[k]$ denotes the value function of the capital owner. Thus, the optimal decision rule for consumption depends on the current value of the state variable k and the policy variables. Thus, the optimal feedback rule satisfies

$$\tilde{c} = \tilde{c}[k; p]$$

where from now on feedback relationships are denoted by a tilde.

We obtain by implicit differentiation

$$\tilde{c}_k = \frac{V_c''[k]}{u''[\cdot]} \cdot \frac{1}{1-p}, \qquad \tilde{c}_\theta = 0 \quad \text{and} \quad \tilde{c}_p = \frac{V_c'[k]}{u''[\cdot]} \cdot \frac{1}{(1-p)^2}.$$

Recall $\eta[c] \equiv -\frac{u'[\cdot]}{u''[\cdot] \cdot c}$, where denotes the intertemporal elasticity of substitution of the capital owners and is a positive number since $u''[\cdot] <$ by assumption. Using equation (28) we then get

$$\begin{aligned}
\tilde{c}_k &= \frac{V_c''[k]}{u''[\cdot]} \cdot \frac{1}{1-p} > 0 && \text{(29a)}\\
\tilde{c}_\theta &= 0 && \text{(29b)}\\
\tilde{c}_p &= \frac{V_c'[k]}{u''[\cdot]} \cdot \frac{1}{(1-p)^2} = \frac{u'[\cdot]}{u''[\cdot]} \frac{1}{(1-p)} = -\eta[\tilde{c}] \cdot \tilde{c}[\cdot] \cdot \left(\frac{1}{1-p}\right) < 0. && \text{(29c)}
\end{aligned}$$

Second, with the decision rule $\tilde{c} = \tilde{c}[k, \theta, p]$ we have

$$\tilde{x} = f[k] - \frac{1-\theta}{1-p} \cdot rk + \frac{p \cdot \tilde{c}[\cdot]}{1-p}.$$

The partial derivatives of x are given by

$$\begin{aligned}
\tilde{x}_k &= f' - \frac{1-\theta}{1-p} r + \frac{p \cdot \tilde{c}_k}{1-p} && \text{(30a)}\\
\tilde{x}_\theta &= \frac{rk}{1-p} && \text{(30b)}\\
\tilde{x}_p &= \frac{(p\tilde{c}_p + c)(1-p) - \{(1-\theta)rk - pc\}}{(1-p)^2}. \\
&= \frac{\{\tilde{c}[k,p,\theta] - (1-\theta)rk\}}{(1-p)^2} + \frac{p \cdot \tilde{c}_p}{1-p}. && \text{(30c)}
\end{aligned}$$

The signs of these derivatives are not immediately clear.

Third, $\tilde{\Delta}$ is given by

$$\tilde{\Delta} = \frac{(1-\theta)rk - \tilde{c}[k, \theta, p]}{1-p}.$$

so that its partial derivatives amount to

$$\begin{aligned}
\tilde{\Delta}_k &= \frac{1-\theta}{1-p} \cdot r - \frac{\tilde{c}_k}{1-p} && \text{(31a)}\\
\tilde{\Delta}_\theta &= -\frac{rk}{(1-p)} && \text{(31b)}\\
\tilde{\Delta}_p &= \frac{-\tilde{c}_p \cdot (1-p) + \{(1-\theta)rk - \tilde{c}[k,p,\theta]\}}{(1-p)^2}. && \text{(31c)}
\end{aligned}$$

Again, the signs of these derivatives are not clear.

C.2 Reactions in the steady state

In the steady state $\dot{\nu} = \dot{k} = 0$. Form that latter condition it follows that $\dot{k} = 0 \Rightarrow \tilde{c}^* = (1-\theta)(r^* - \delta)k^*$ where steady state variables are again marked by an asterisk. Notice that $\tilde{c}_k$ must obey (29a), and must hold in a steady state too. Furthermore, recall that η is constant in the steady state.

Thus, from the equations (28), and (29a) - (29c, noting that $-\eta < 0$ and $V_c'', u'' < 0$ and for $\theta, p < 1$, the reactions of $\tilde{c}$, $\tilde{x}$ and $\tilde{\Delta}$ in the steady state are the following.

$$\tilde{c}_k^* = \frac{V_c''[k]}{u''[\tilde{c}^*]} \cdot \frac{1}{1-p} > 0, \quad \tilde{c}_\theta^* = 0, \quad \tilde{c}_p^* = \frac{-\eta \cdot \tilde{c}^*}{1-p} < 0. \tag{32}$$

With this the partial effects on $\tilde{x}$ become

$$\tilde{x}_k^* = f'[k^*] - \frac{(1-\theta)}{1-p} r^* + \frac{p \cdot \tilde{c}_k^*}{1-p} \tag{33a}$$

$$\tilde{x}_\theta^* = \frac{r^* k^*}{1-p} \tag{33b}$$

$$\tilde{x}_p^* = \frac{p \cdot \tilde{c}_p^*}{1-p}. \tag{33c}$$

Third, the partial effects on $\tilde{\Delta}$ are then given by

$$\tilde{\Delta}_k^* = \frac{(1-\theta)}{1-p} \cdot r^* - \frac{\tilde{c}_k^*}{1-p} \tag{34a}$$

$$\tilde{\Delta}_\theta^* = -\frac{r^* k^*}{1-p} \tag{34b}$$

$$\tilde{\Delta}_p^* = -\frac{\tilde{c}_p^*}{1-p}. \tag{34c}$$

D Comparative statics of the optimal income tax rate

The entries in Table 1 are derived as follows. The optimal capital income tax rate has to satisfy

$$\gamma^\eta = \frac{1}{\rho}\left(f^*/k^* - r^*\right) + \theta\left(1 + \gamma^\eta\right),$$

where $f^*/k^* = r^*/\alpha$ and $r^* = \rho$. Thus, $\widetilde{\theta}^*$ has to satisfy

$$\gamma^\eta = (\frac{1}{\alpha} - 1) + \theta\left(1 + \gamma^\eta\right).$$

Taking total differentials implies the following.

$$\gamma: \quad \eta\gamma^{\eta-1}d\gamma = (1+\gamma^{\eta})d\theta + \theta\eta\gamma^{\eta-1}d\gamma$$
$$(1-\theta)\eta\gamma^{\eta-1}d\gamma = (1+\gamma^{\eta})d\theta \quad \Rightarrow \quad d\theta/d\gamma > 0$$

$$\alpha: \quad 0 = \left(-\frac{1}{\alpha^2}\right)d\alpha + (1+\gamma^{\eta})d\theta \quad \Rightarrow \quad d\theta/d\alpha > 0$$

$$\rho: \quad 0 = (1+\gamma^{\eta})d\theta \quad \Rightarrow \quad d\theta/d\rho = 0$$

$$\eta: \quad \ln\gamma \cdot e^{\eta\ln\gamma}d\eta = \theta\ln\gamma \cdot e^{\eta\ln\gamma}d\eta + (1+\gamma^{\eta})d\theta$$
$$(1-\theta)\ln\gamma \cdot e^{\eta\ln\gamma}d\eta = (1+\gamma^{\eta})d\theta$$
$$\Rightarrow d\theta/d\eta > 0, \forall\gamma > 1 \quad \text{or} \quad d\theta/d\eta < 0, \forall\gamma < 1$$

E The second order sufficient conditions

In order to check the sufficiency conditions of the solutions use Arrow's theorem. That requires that the Hamiltonian evaluated at the optimum solution for the choice variable (as a function of the state and costate variables) be concave in the state variable for a given co-state variable. If the concavity is strict, the solution (for the control variable) is then also the unique optimizer. Notice that in our case the optimum is restricted to the long-run, steady state. Thus, the Hamiltonian is also evaluated at the steady state optimum. For Arrows's theorem and a similar steady state problem see, for example, Weitzman (2003), ch. 3, especially p. 85-93.

The Hamiltonian, $\mathcal{H}^C$, of the government is given by

$$\mathcal{H}^C = \gamma v[x] + u[c] + \mu \cdot \Omega \quad \text{where} \quad \Omega \equiv \left(\frac{(1-\theta)(rk) - c}{1-p}\right).$$

The choice variables for $\mathcal{H}^C$ are c, θ and p. In the long-run optimum $p = \theta$ is implicitly determined from $\gamma v[\tilde{x}^*] + u[\tilde{c}^*]$. Furthermore, the first order necessary condition implies

$$\gamma v'[\tilde{x}^*] = u'[\tilde{c}^*] \quad \text{where} \quad \tilde{c}^* = (1-\theta)(r^*k^*) = (1-\theta)\rho k^*$$
$$\text{and} \quad \tilde{x}^* = f[k^*] - rk^* + \frac{\theta\tilde{c}^*}{1-\theta} = f[k^*] - r^*k^* + \theta\rho k^*.$$

From that one verifies that in the optimum the Hamiltonian $\mathcal{H}^C$ is given by $\mathcal{H}^{C*} \equiv$

$\mathcal{H}^C[k^*, c[k^*], \theta[k^*], p[k^*]]$, that is,

$$\mathcal{H}^{C^*} = \gamma v[f[k^*] - r^* k^* + \theta[k^*] \cdot \rho k^*] + u[(1-\theta[k^*])\rho k^*] + \mu \cdot 0,$$

where in the optimum p satisfies $p[k^*] = \theta[k^*]$, the tax rate $\theta[k^*]$ is implicity determined by

$$\gamma v'[f[k^*] - r^* k^* + \theta \rho k^*] = u'[(1-\theta)\rho k^*]$$

and $\Omega = 0$ in the long-run (steady state) optimum.

Arrow's theorem is satisfied if $\mathcal{H}^{C^*}$ is concave in k^*. To that end we check the sign of the second derivative of $\mathcal{H}^{C^*}$. For simplicity let $f' = f'[k^*]$ and $f'' = f''[k^*]$. From

$$\mathcal{H}_k^{C^*} = \gamma v'[\cdot] \left(f' - r^* + \frac{\partial \theta}{\partial k^*} \cdot \rho k^* + \theta \rho \right) + u'[\cdot] \left(-\frac{\partial \theta}{\partial k^*} \cdot \rho k^* + (1-\theta)\rho \right)$$

one obtains the expression for the second derivative as

$$\begin{aligned} \left(\mathcal{H}^{C^*}\right)_{kk} &= \gamma v''[\cdot] \left(f' - r^* + \frac{\partial \theta}{\partial k^*} \cdot \rho k^* + \theta \rho \right)^2 + \gamma v'[\cdot] \left(f'' + \frac{\partial^2 \theta}{\partial (k^*)^2} \cdot \rho k^* + \frac{\partial \theta}{\partial k^*} \cdot \rho + \frac{\partial \theta}{\partial k^*} \cdot \rho \right) \\ & u''[\cdot] \left(-\frac{\partial \theta}{\partial k^*} \cdot \rho k^* + (1-\theta)\rho \right)^2 + u'[\cdot] \left(-\frac{\partial^2 \theta}{\partial (k^*)^2} \cdot \rho k^* - \frac{\partial \theta}{\partial k^*} \cdot \rho - \frac{\partial \theta}{\partial k^*} \cdot \rho \right). \end{aligned}$$

This expression can be simplified to

$$\left(\mathcal{H}^{C^*}\right)_{kk} = \gamma v''[\cdot] \cdot D_1 + u''[\cdot] \cdot D_2 + \gamma v'[\cdot] \cdot f'' + (\gamma v'[\cdot] - u'[\cdot]) \cdot D_3$$

where $D_1 \equiv \left(f'[k^*] - r^* + \frac{\partial \theta}{\partial k^*} \cdot \rho k^* + \theta \rho\right)^2$, and $D_2 \equiv \left(-\frac{\partial \theta}{\partial k^*} \cdot \rho k^* + (1-\theta)\rho\right)^2$ are positive, because they are squared expressions, and $D_3 \equiv \left(\frac{\partial^2 \theta}{\partial (k^*)^2} \cdot \rho k^* + 2\frac{\partial \theta}{\partial k^*} \rho\right) \gtreqless 0$. Because $\gamma v'[\cdot] = u'[\cdot]$, the last term on the right hand side is zero, no matter what the sign of D_3 is. Furthermore, $\gamma v'[\cdot] > 0$, and $f''[k^*] < 0$ so that their product is negative. As D_1 is positive and $v''[\cdot] <$, the first term is negative. Thus, it follows that $\left(\mathcal{H}^{C^*}\right)_{kk} < 0$. Hence, the Hamiltonian $\mathcal{H}^{C^*}$ is concave in k^*, and the sufficient conditions hold for the quasi-command optimum.

E.1 Numerical simulation

In order to get a feeling for the nature of the solutions I use a numerical simulation based on calibrations and empirical results in literature. For the share of broad capital I follow Barro and Sala–i–Martin (2004) and set it equal to 0.75. Jordá, Knoll, Kuvshinov, Schularick, and Taylor (2019) report a long-run estimate of the return on wealth of $r^* \approx 0.06$ for a weighted panel of countries and the long-run period 1870-2015. In terms of the model I set the time preference rate equal to their long-run return on wealth so $\rho = 0.06$. For a value of the elasticity of intertemporal substitution (EIS) I relate to Havranek, Horvath, Irsova, and Rusnak (2015) who find in a meta-analysis that the EIS varies greatly between countries. For simplicity I concentrate on the mean value for the U.S. which roughly equals 0.06.

To justify the exercise to vary γ parametrically notice that the reason the paper works with a high capital share is the observation that in most countries the income distribution is very skewed and only a small number of people (as a group) receive 75% of total income. Thus, a lot more people receive their income in the form of wages for their raw labour services than in income derived from human and physical, that is, *broad* capital services. For a population-based justification and relating to empirical income distributions I consider the contribution by Lansing (2015) who sets the number of workers relative to capitalists equal to nine. That seems extreme in the present setup with receivers of income from broad capital. Recently Walsh (2016), p. 21, uses data to argue that the relative number is rather 0.75 so that there are three (unskilled) workers (non-asset-holders) in relation to one capitalist. Eurostat reports data on the ”Employment rate of low skilled persons, age group 20-64” with a mean of 55,7 percent for 2019. That would suggest a number of a low-skilled-population/income-group-adjusted γ of around 1, but less than 1, if one additionally adds the non-working capital owners (asset-holders). However, to provide a robust and exact link between demographics and the social weight γ is only cursory and really beyond the scope of this paper.

Table 2: Baseline Parameter Values

α	ρ	η
0.75	0.06	0.60

These values imply that if $\gamma > \left(\frac{1-\alpha}{\alpha}\right)^{\frac{1}{\eta}} = 0.16$ the condition for positive optimal capital income tax rates is met. Table 3 reports the results of a numerical simulation for the calibrated economy varying γ only. I only concentrate on solutions where the optimal tax rate is non-zero.

Table 3: Optimal Capital Income Tax Rates $\widehat{\theta}$

γ	0.01	0.05	0.10	0.20	0.50	0.80	1
$\widehat{\theta}$	-0.25	-0.14	- 0.07	0.03	0.20	0.29	0.33

γ	2	5	10	20	50	100	1000
$\widehat{\theta}$	0.47	0.63	0.73	0.81	0.89	0.92	0.98

The numbers suggest that the social weight γ is an important determinant of the optimal tax rate. That corresponds to common intuition. Governments that give more weight to the interests of the workers seem to choose higher capital income tax rates. Recall that the exact value of γ is outside the model. It may seem that one has to require a not-so-large value of γ to obtain a non-zero and positive value of the capital income tax rate that seems realistic. Thus, even mild preference for the welfare of the low-skilled workers would call for a positive capital income tax rate in this model. Furthermore, when attaching equal weights to the welfare of the agents, that is, when $\gamma = 1$, entails an optimal capital income tax that would roughly correspond to the (highest marginal) capital income tax rate in the United States which is currently around 35 percent. Applying this logic to some European countries like France and Germany where the (highest marginal) capital income tax rate is often above 40 percent would be compatible with a larger value of γ. From the simulation it would then approximately have to equal two.

References

ABEL, A. B. (2007): "Optimal Capital Income Taxation," Working Paper 13354, NBER, Cambridge, MA.

ACEMOGLU, D. (2009): *Introduction to Modern Economic Growth*. Princeton University Press, Princeton, New Jersey.

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

ASEN, E. (2020): "Capital Cost Recovery across the OECD," Fiscal Fact 704, Tax Foundation, Washington, D.C.

ATKINSON, A. B., AND J. E. STIGLITZ (1989): *Lectures on Public Economics*. McGraw–Hill, Singapore, international, 2nd printing edn.

BARRO, R. J., AND X. SALA–I–MARTIN (2004): *Economic Growth*. MIT Press, Cambridge, Massachusetts, 2nd edn.

BASAR, T., AND G. J. OLSDER (1995): *Dynamic Noncooperative Game Theory*. Academic Press, London, 2nd edn.

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

——— (2001): "Capital Income Taxation, Wealth Distrbution and Borrowing Constraints," *Journal of Public Economics*, 79, 55–69.

CONESA, J. C., AND B. DOMÍNGUEZ (2020): "Capital Taxes and Redistribution: The Role of Management Time and Tax Deductible Investment,," *Review of Ecnomic Studies*, 37, 156–172.

CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, 99(1), 25–48.

DAVIES, J. B., J. ZENG, AND J. ZHANG (2009): "Time-consistent Taxation in a Dynastic Family Model with Human and Physical Capital and a Balanced Government Budget," *Canadian Journal of Economics*, 42, 1023–1049.

DOCKNER, E., S. JORGENSEN, N. V. LONG, AND G. SORGER (2000): *Differential Games in Economics and Management Science*. Cambridge University Press, Cambridge, U.K.

DOMEIJ, D., AND J. HEATHCOTE (2004): "On the Distributional Effects of Reducing Capital Taxes," *International Economic Review*, 45, 523–554.

FISHER, I. (1937): "Income in Theory and Income Taxation in Practice," *Econometrica*, 5, 1–55.

GRÜNER, H. P., AND B. HEER (2000): "Optimal Flat-Rate Taxes on Capital - A Re-examination of Lucas' Supply Side Model," *Oxford Economic Papers*, 52, 289–305.

GUO, J. T., AND K. J. LANSING (1999): "Optimal Taxation of Capital Income with Imperfectly Competitive Product Markets," *Journal of Economic Dynamics and Control*, 23, 967–995.

HALL, R. E., AND D. W. JORGENSON (1967): "Tax Policy and Investment Behaviour," *American Economic Review*, 57, 391–414.

——— (1971): "Application of the Theory of Optimal Capital Accumulation," in *Tax Incentives and Capital Spending*, ed. by G. Fromm, pp. 9–60. The Brookings Institution, Washington D. C.

HAVRANEK, T., R. HORVATH, Z. IRSOVA, AND M. RUSNAK (2015): "Cross-Country Heterogeneity in Intertemporal Substitution," *Journal of International Economics*, 96, 100–118.

JONES, L. E., R. E. MANUELLI, AND P. E. ROSSI (1997): "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73, 93–117.

JONES, R. W. (1965): "The Structure of Simple General Equilibrium Models," *Journal of Political Economy*, 73, 557–572.

JORDÁ, Ó., K. KNOLL, D. KUVSHINOV, M. SCHULARICK, AND A. M. TAYLOR (2019): "The Rate of Return on Everything, 1870-2015," *Quarterly Journal of Economics*, 134, 1225–1298.

JUDD, K. L. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28, 59–83.

——— (1999): “Optimal Taxation and Spending in General Competitive Growth Models,” *Journal of Public Economics*, 71, 1–26.

KALDOR, N. (1955): *An Expenditure Tax*. Allen and Unwin, London.

——— (1956): “Alternative Theories of Income Distribution,” *Review of Economic Studies*, 48(5), 83–100.

KAMIEN, M. I., AND N. L. SCHWARTZ (1991): *Dynamic Optimization, The Calculus of Variation and Optimal Control in Economics and Management*. North-Holland, Amsterdam–New York–London–Tokyo, 2nd edn.

KEMP, M. C., N. VAN LONG, AND K. SHIMOMURA (1993): “Cyclical and Noncyclical Redistributive Taxation,” *International Economic Review*, 34, 415–429.

KYDLAND, F. E., AND E. C. PRESCOTT (1977): “Rules Rather Then Discretion: The Inconsistency of Optimal Plans,” *Journal of Political Economy*, 85(3), 473–491.

LANSING, K. J. (1999): “Optimal Redistributive Capital Taxation in a Neoclassical Growth Model,” *Journal of Public Economics*, 73, 423–453.

——— (2015): “Asset Pricing with Concentrated Ownership of Capital and Distribution Shocks,” *American Economic Journal: Macroeconomics*, 7, 67–103.

LONG, N. V. (2010): *A Survey of Dynamic Games in Economics*. World Scientific Publishing, Singapore.

LONG, N. V., AND K. SHIMOMURA (2000): “Semi-Stationary Equilibrium in Leader-Follower Games,” Discussion Paper 2000s-08, CIRANO, Montreal.

——— (2002): “Redistributive Taxation in Closed and Open Economies,” in *Economic Theory and International Trade. Essays in Honour of Murray C. Kemp*, ed. by A. Woodland, pp. 104–123. Edward Elgar, Cheltenham, U.K.

LUCAS, R. E. (1990): “Supply-Side Economics: An Analytical Review,” *Oxford Economic Papers*, 42, 293–316.

MANKIW, N. G., D. ROMER, AND D. N. WEIL (1992): “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics*, 152, 407–437.

PASINETTI, L. (1962): “Rate of Profit and Income Distribution in Relation to the Rate of Economic Growth,” *Review of Economic Studies*, 29, 267–279.

REHME, G. (1995): “Redistribution, Income cum Investment Subsidy Tax Competition and Capital Flight in Growing Economies,” Working Paper ECO 95/16, European University Institute, Florence, Italy.

——— (2002): “Distributive Policies and Economic Growth: An Optimal Taxation Approach,” *Metroeconomica*, 53, 315–338.

——— (2007): “Optimal Capital Income Taxation, Investment Subsidies and Redistribution in a Neoclassical Growth Model,” Darmstadt Discussion Papers in Economics (DDPIE) 188, TU Darmstadt.

——— (2011): “Amortissement Fiscal et Redistribution dans un Modèle de Croissance Néoclassique,” *Revue de'l OFCE*, 116, 367–391.

REINHORN, L. (2018): "On optimal redistributive capital taxation," *Journal of Public Economic Theory*, 21, 460–487.

SAEZ, E. (2013): "Optimal progressive capital inncome taxes in the infinite horizon model," *Journal of Public Economics*, 97, 61–74.

SAMUELSON, P. A. (1964): "Tax Deductibility of Economic Depreciation to Insure Invariant Valuations," *Journal of Political Economy*, 72, 604–606.

SARGENT, T. J., AND L. LJUNGQVIST (2004): *Recursive Macroeconomic Theory*. MIT Press, Cambridge, Massachusetts, 2nd edn.

SELIM, S. (2010): "Optimal Taxation in a Two-sector Economy with Heterogeneous Agents," *Economics Bulletin*, 30(1), 534–542.

STRAUB, L., AND I. WERNING (2020): "Positive Long-Run Capital Taxation: Chamley-Judd Revisited," *American Economic Review*, 110, 86–119.

TURNOVSKY, S. J. (2000): *Methods of Macroeconomic Dynamics*. MIT Press, Cambridge, Mass., 2nd edn.

UHLIG, H., AND N. YANAGAWA (1996): "Increasing The Capital Income Tax May Lead To Faster Growth," *European Economic Review*, 40, 1521–1540.

WALSH, C. E. (2016): "Workers, Capitalists, Wages, Employment and Welfare," mimeo, UC Santa Cruz, Santa Cruz, CA.

WEITZMAN, M. L. (2003): *Income, Wealth, and the Maximum Principle*. Harvard University Press, Cambridge, Mass.

WERNING, I. (2007): "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, 122, 925–967.

XIE, D. (1997): "On Time Inconsistency: A Technical Issue in Stackelberg Differential Games," *Journal of Economic Theory*, 76, 412–430.

ZHANG, J., J. DAVIES, J. ZENG, AND S. MCDONALD (2008): "Optimal Taxation in a Growth Model with Public Consumption and Home Production," *Journal of Public Economics*, 92, 885–896.