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

- Preliminary version -

Günther Rehme

Technische Universität Darmstadt†

November 27, 2017

Abstract

In this paper it is analyzed how 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. They may, however, stabilize the investment return in a recession. But when the agents and the government act optimally for the long run the investment subsidies should be such that the tax scheme does not distort accumulation anymore. This holds regardless of social preferences. I find that redistribution and so capital income taxes may be 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: Growth, Redistribution, Investment Subsidies, Capital Income Taxes

JEL classification: O41, H21, D33

*I thank Tony Atkinson, James Mirrlees, and Robert Waldmann for inspiring comments on a project that stimulated this research. I have also benefitted from comments by Santiago Caram, Sabine Eschenhof-Kammer, Christian Keuschnigg, Enrique Neder and Marco Runkel, 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 conference “Fiscal Issues, Reforms and Development Policies” at the Center for Training and Research in Public Finance and Policy (CTRPFP) at the Center for Studies in Social Sciences Calcutta (CSSSC), Kolkata, 2015, and the Arnoldshain Seminar XV “The EU and Latin America Facing Globalization” in Vienna, 2017. 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 fiscal policy response of most governments during the last financial and economic crisis has been to institute measures that fight a fall in investment activity which has been accompanying the economic slowdown. Most of these measures allow for writing off a particular percentage of the investment outlays against the investors' tax bill, especially capital income taxes. A particular form of such investment promoting fiscal policies has, for example, been to increase accelerated capital depreciation allowances. All these forms of investment promotion have, in one form or another, been increased significantly in, for example, the United States, the United Kingdom, France and Germany, but also in other OECD and non-OECD countries in the years 2008/9, making it possible in some cases even to fully expense the total investment outlays.[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

[1] For the U.S. see the Economic Stimulus Act of 2008 (Treasury Department, Tres. Reg. Sec. 1.168(k)-1), for Germany, see "Konjunkturpaket 2008" (Bundesministerium der Finanzen, November 2008), for France see "La Loi de Finance 2009", December 2008; for the United Kingdom see "Fiscal Act 2008". See also the "Provisional Measure 428", May 2008, as part of Brazil's new Productive Development Policy (PDP), and Russia's measures according to which the profit tax base will "decrease for companies investing in capital assets as the immediately recoverable depreciation allowance is raised from 10% to 30% of the asset cost." See "Federal Budget of Russia", November 2008.

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 other growth contexts as has been shown by many contributions. See, for example, Kemp et al. [1993], Aiyagari [1995], Rehme [1995], Uhlig and Yanagawa [1996], Lansing [1999] Grüner and Heer [2000], Chamley [2001], Saez [2002], Domeij and Heathcote [2004], Abel [2007], Werning [2007], Conesa et al. [2010], Selim [2010], and others.

The present paper relates to these findings. In particular, I consider a policy package whereby investment subsidies are coupled with capital income taxes. 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")

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

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 et al. [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 the government would mostly find it optimal to choose a subsidy policy that does not distort accumulation. The reason is that that would remove the distorting effect policy has on capital accumulation. Thus, the paper shows that, in relation to the investors' tax bill, full expensing of investment outlays is optimal for the long run.[3] That finding

[3] In that sense the assumption of full expensing of investment outlays in relation to the capital income tax bill made in Rehme [1995], Rehme [1998], ch. 2, and Rehme [2002] which provided verbal arguments why this may be optimal in a general equilibrium, endogenous growth framework, is endo-

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]. In a representative agent, dynamic general equilibrium framework Abel [2007] has recently established the same result. 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. Thus, in the model granting investment subsidies serves as an important *indirect* redistribution device, because transfers ultimately depend on the capital income tax rate chosen.

When taking governments - no matter which clientele a benevolent government represents - to pursue such a nondistortionary policy that seems to benefit everybody, it turns out that capital income taxes may not always be optimally zero in the long run. Rather, I find that capital income taxes may optimally be nonzero for redistribution. This 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.

In a numerical simulation using calibrated parameters for the U.S. from the business cycle literature the theoretical results are given a quantitative flavor. In particular, the simulation highlights the important relationship between the strength of social preferences and possibly positive capital income taxes.

---

genized in this paper and found to be optimal in the present general equilibrium, neoclassical growth framework. For recent papers that find a related result see, for example, Rehme [2007] and Davies et al. [2009].

The main message of the present paper is that investment subsidies, when coupled with capital income taxes, may well be important redistributive devices in the long run, if the agents and the government behave optimally.

The paper is organized as follows: Section 2 presents the model. Section 3 analyzes the optimality for tax rates in long-run equilibrium. Section 4 provides concluding remarks.

2 The Model

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 form the consumption of a homogenous, malleable good. The population is normalized so that the number of each type equals one. The model abstracts from uncertainty, technological progress, population growth and depreciation. The latter implies that aggregates are really defined in net terms which has no consequence for the price-taking, market clearing logic of the model. The workers supply one unit of unskilled labour inelastically and do not save or invest.[4] Thus, all the wealth (human and physical) is concentrated in the hands of the capitalists who supply the services from their wealth to competitive firms.

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 et al. [1992]. This assumption implies

[4]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].

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.[5]

The capital owners' instantaneous budget constraint is given by

$$c_t + i_t = (1-\theta_t) r_t k_t + p_t i_t \quad \text{and} \quad i_t = \dot{k}_t.$$

Thus, the capitalists derive income from renting their physical and human capital, k_t, to competitive firms at the rate r_t. Gross rental income is taxed at the rate θ_t and a fraction p_t of investment undertaken, i_t, is subsidized by the government. For simplicity investment is net of depreciation and we abstract from the latter altogether. By assumption $\theta_t, p_t \in [0,1]$. Thus, investment subsidies are $p_t i_t$. The capitalists' consumption c_t depends on their after-tax capital income minus after-tax investment.[6]

Rearranging the capital owners solve

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

$$s.t. \qquad \dot{k}_t = \left(\tfrac{1-\theta_t}{1-p_t}\right) r_t k_t - \tfrac{c_t}{1-p_t} \tag{1}$$

$$k(0) = \text{given}, \quad k(\infty) = \text{free}. \tag{2}$$

where ρ is the constant rate of time preference, common to all agents. The instanta-

[5]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 et al. [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.

[6]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 et al. [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 to this one. See Rehme [2009].

neous utility function $u[c_t]$ satisfies the usual properties $u' > 0, u'' < 0$ and $\lim\limits_{c_t \to \infty} u' = 0$ and $\lim\limits_{c_t \to 0} u' = \infty$ where $u' = \frac{du[c_t]}{dc_t}$ and $u'' = \frac{d^2 u[c_t]}{dc_t^2}$. The current value Hamiltonian for this problem is

$$H = u[c_t] + \lambda_t \left(\left(\frac{1-\theta_t}{1-p_t} \right) r_t k_t - \frac{c_t}{1-p_t} \right)$$

and the necessary first order conditions for its maximization are

$$H_c : \quad u' - \tfrac{\lambda_t}{1-p_t} = 0 \tag{3a}$$

$$H_k : \quad -\lambda_t \left(\tfrac{1-\theta_t}{1-p_t} \right) r_t + \rho \lambda_t = \dot{\lambda}_t \tag{3b}$$

plus the transversality condition $\lim\limits_{t \to \infty} k_t \lambda_t e^{-\rho t} = 0$ and the requirement that equation (1) holds.[7] The (non-negative) co-state variable λ_t represents the capital owners' shadow price of an additional unit of capital in terms of utility.

2.2 Workers

The (unskilled) *workers* do not invest and are not taxed by assumption.[8] They supply one unit of 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.

$$x_t = w_t + TR_t. \tag{4}$$

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

[7] As H is concave in c_t and k_t, the necessary conditions are also sufficient.

[8] The working population is normalized so that there is one worker and one capitalist.

conditions $\lim\limits_{x_t \to \infty} v' = 0$ and $\lim\limits_{x_t \to 0} v' = \infty$ where $v' = \frac{dv[x_t]}{dx_t}$ and $v'' = \frac{d^2 v[x_t]}{dx_t^2}$.

2.3 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. 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 $\lim\limits_{t \to \infty} f'(k_t) = 0$ and $\lim\limits_{t \to 0} f'(k_t) = \infty$. Profit maximization implies

$$r_t = f'(k_t) \tag{5}$$

$$w_t = f(k_t) - f'(k_t)k_t \tag{6}$$

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 that k_t is broadly defined. In that case the share of (broad) capital will in general be larger than one half so that the capitalists (owners of human and physical capital) have higher gross income than the low-skilled workers. For example, Barro and Sala–i–Martin [2004], ch. 2.6.6, assume that the share of broad capital is roughly

0.75.

2.4 Government

Following Judd [1985] and Lansing [1999], I rule out a market for government bonds and assume that the government can commit itself to following a tax-transfer policy announced at $t = 0$. The government chooses paths of the capital income tax rate, θ_t, the fraction of investment the government wishes to subsidize, p_t, and the transfers to the workers, TR_t, to maximize a weighted sum of the agents' lifetime utilities, subject to the behaviour of the private sector in an equilibrium and the condition that its budget be balanced at each point in time

$$TR_t = \theta_t r_t k_t - p_t \dot{k}_t. \tag{7}$$

Thus, the government collects income taxes on capital to grant an investment subsidy ($p_t \dot{k}_t$) to the capital owners and use the remaining resources for lump-sum transfers to the workers. Thus, by assumption the returns to human and physical capital are taxed equally. Capital income taxes in this model are then really an equal tax on all returns from accumulated factors of production. Hence, in terms of the tax package considered we contemplate a capital-income-cum-investment-subsidy-tax (CICIST) scheme.

2.4.1 Arbitrary Private Sector Behaviour and Investment Return Stabilization

Suppose we consider the effects of policy changes when the agents and the government satisfy their budget constraints but otherwise act in unspecified ways at a particular point in time. By arbitrary behaviour it is meant that the capital owners and the government have not necessarily solved for their optimal decisions yet. However, it is not assumed to be absolutely unspecified since we assume that the agents at least obey their budget constraints. The assumed nonsatiation in utility clearly implies that.

In the model agents face given price paths. Although the setup is not really capturing business cycle phenomena we can still get a flavour what investment subsidies entail in an economic downturn. The latter most often entails a significant drop in the real return to capital r_t. A standard mechanism to model such a drop is to argue that there is a shock to technology so that the marginal product of capital falls where the latter equals the real return on capital in equilibrium. Here we leave that (i.e. such equilibrium responses due to changes in economic fundamentals) outside the model and simply argue that for some reason (e.g. the financial markets, animal spirits or whatever) the real return of capital falls due to some process that is outside the model. We will look at the model's implications for one policy response due to such a drop in r_t.[9] More precisely, we consider a government that wishes to 'stabilize' the (after-tax) real return on investment by means of the two policy instruments considered in this paper.

By equation (1) one easily verifies that the (after-tax) real return on investment is given by

$$R_t \equiv \frac{(1-\theta_t)}{(1-p_t)} \cdot r_t. \tag{8}$$

Suppose there is a significant drop in r_t and the government reacts to that by changing p_t or θ_t in order to keep R_t constant. Heroically, we assume that the government can react immediately, once it perceives the drop in r_t. For such a government we would then have

$$dR_t = 0 = R_r\, dr_t + R_p\, dp_t + R_\theta\, d\theta_t.$$

[9]Thus, we assume that the price path r_t features a particular point in time where there is a sharp or noticeable drop in the real return on capital.

If the government counteracts the drop in r_t only by changing the investment subsidies, keeping the (path of the) tax rates θ_t and R_t constant, one finds that

$$\frac{dp_t}{dr_t} = -\frac{(1-p_t)}{r_t} \tag{9}$$

which is negative for $\theta_t, p_t \in (0,1)$ and $r_t > 0$ which I assume to hold. Thus, if the real return on capital r_t *falls*, i.e. when ($dr_t < 0$), and the government wants to 'stabilize' the real return on investment, it must *increase* the investment subsidies.

In turn, if the government fixes (the path of) p_t and adjusts the tax rates one gets

$$\frac{d\theta_t}{dr_t} = \frac{(1-\theta_t)}{r_t} \tag{10}$$

which is positive under the assumptions made. Thus, a *drop* in r_t requires a tax cut, i.e. *lower* θ_t.

Hence, if the government wishes to to 'stabilize' the real return to investment when there is a *drop* in the real return to capital r_t, then it should *increase* investment subsidies p_t or *cut* the capital income tax rate θ_t by compensating amounts, respectively.

To get a feeling for the magnitudes involved consider the following scenario. Suppose the real return before the downturn is around five percent, i.e. $r_t = 0.05$. That seems to be a number that is broadly in line with the long-run real return in the United States. Let the capital income tax rate be $\theta_t = 0.35$ as is approximately the case in the U.S. now. Furthermore, let all the current capital investment subsidies equal a conservative $p_t = 0.25$. Now suppose the real return drops by 50 percent.[10] Then

$$\frac{dp_t}{p_t} = \left(\frac{-(1-p_t)}{p_t}\right) \cdot \frac{dr_t}{r_t} = (-0.75/0.25) \cdot (-0.50) = 1.50.$$

[10]Most commentators argue that, for instance, during the Great Depression the real return dropped by about 80 percent. Thus, assuming a drop of 50 percent is considering a situation that is less severe than that in the 1930s.

Hence, p_t should be more than doubled, that is, it should be raised from 0.25 to 0.625. Similarly,

$$\frac{d\theta_t}{\theta_t} = \left(\frac{(1-\theta_t)}{\theta_t}\right) \cdot \frac{dr_t}{r_t} = (0.65/0.35) \cdot (-0.50) = -0.93.$$

Thus, the tax rate should be reduced to almost zero.[11]

Therefore, the model suggests that many countries seem to have indeed pursued something like the objective to 'stabilize' the real return to investment in the current economic crisis, especially by adjusting p_t. This is because in most situations it is a lot easier to adjust investment subsidies than changing tax rates, in particular statutory tax rates that often require relatively longer processes of political debates and legislation.

Of course, one should be aware of the fact that the policy objective of 'stabilizing' the investment return would also imply a reverse reaction of p_t and θ_t when there is a sudden boost in r_t that may accompany an economic upswing. Furthermore, the result applies *only* to the investment return. Other 'stabilization' objectives such as e.g. transfer or tax-revenue 'stabilization' may imply other responses as regards p_t and θ_t.

2.4.2 Non-Distortion of Accumulation

One important consequence of the result that capital income taxes be optimally zero is that the capital accumulation process will not be disturbed by political interference. The impact of accumulation distortion can be inferred from the Euler equation in (3b). It shows how agents evaluate the evolution of the state variable k_t in terms of their welfare. This then leads them to a particular accumulation programme. Policy would

[11] As regards investment subsidies these numbers and calculated reactions do not seem to be too unrealistic with respect to what some governments have done in the current crisis. For example, the U.S. and France pushed one form of investment subsidy, namely accelerated capital depreciation allowances up to well over 50 percent in reaction to the crisis in 2008/9.

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 $\dot{\lambda} = 0$ in (3b) when $\theta_t = 0, p_t = 0, \forall t$. This is basically what the results in Chamley [1986] and Judd [1985] imply. But another nondistortionary policy is possible, namely when $\theta_t = p_t$. Whether nondistortionry $\theta_t = p_t$ implies zero tax rates will be the focus of the analysis below.[12]

2.4.3 The workers' consumption

When the factor input and goods markets are in equilibrium the workers' income is given by

$$x_t = w_t + TR_t = f(k_t) - r_t k_t + \theta_t r_t k_t - p_t \dot{k}_t \tag{11}$$

In equilibrium the overall resource constraint is such that the agents satisfy their budget constraints. Substitution of (1) into (11) one then obtains

$$x_t = f(k_t) - r_t k_t + \theta_t r_t k_t - p_t \left(\frac{1-\theta_t}{1-p_t}\right) r_t k_t + \frac{p_t c_t}{1-p_t}.$$

This expression can then be simplified to

$$x_t = f(k_t) - \left(\frac{1-\theta_t}{1-p_t}\right) r_t k_t + \frac{p_t c_t}{1-p_t}. \tag{12}$$

Thus, in equilibrium the total income of the workers, which equals their consumption, is increasing in the consumption of the capital owners.

[12]This assumption nests a setup with just θ and no p_t and perhaps finding that the optimal θ_t is then zero in the long run. Notice that this policy package is tantamount to a tax on the capitalists' consumption. However, it is implemented as an income tax scheme and, thus, different.

2.4.4 Arbitrary Behaviour and Policy

Again assume that 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 we then obtain

$$\frac{dx_t}{dp_t}_{|\theta_t, c_t} = \frac{c_t - (1-\theta_t)r_t k_t}{(1-p_t)^2}. \tag{13}$$

Concentrating on non-negative growth of the capital stock, we must have $c_t \leq (1 - \theta_t)r_t k_t$ so that it seems that $\frac{\partial x_t}{\partial p_t}_{|\theta_t, c_t} \leq 0$ in general. Thus, 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

$$\frac{dx_t}{d\theta_t}_{|p_t, c_t} = \frac{r_t k_t}{(1-p_t)^2} \tag{14}$$

which is positive for given $0 \leq p_t \leq 1$, $c_t \geq 0$ and $r_t > 0$. Thus, capital income taxes net of investment subsidies seem to be a positive redistribution device because they seem to 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 (1). Here we 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 policies in the interior of the unit interval, i.e. $\theta_t, p_t \in (0, 1)$, 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 there is positive taxation of capital income and net income of the capital owners is greater than their consumption.

Given the reactions of net investment and the after-tax wages due to changes in arbitrary θ_t and p_t we may summarize the above results as follows: When the agents and the policy maker obey their budget constraints and otherwise act in arbitrary (possibly non-optimal) ways at a particular point in time and a capital-income-cum-investment-subsidy tax scheme (CICIST) is used, an increase in the capital income tax rate θ_t does not appear to raise net investment, $\dot{k}_t$, but generally increases after-tax wages. In turn, an increase in the investment subsidies p_t generally raises net investment, if $\theta_t > 0$, and generally lowers after-tax wages. Thus, when behaviour and policy are unspecified, for instance, when agents and the government do not necessarily optimize, investment subsidies do not seem to be a good redistribution device in the short run.

These results hold for the short-run. Below we introduce optimizing behaviour and find that the above results need qualification, once we introduce optimizing behaviour and look at the long run.

3 The Long-Run Optimal Capital Income Tax

A benevolent government respects the private sector optimality conditions, keeps the agents on their respective supply and demand curves, and chooses a policy that can be realized as a competitive equilibrium.[13] The government solves the following prob-

[13]Similar approaches that use a two-class set following, for example, Kaldor [1956] are used by Judd [1985], Judd [1999], and Lansing [1999].

lem[14]

$$\max_{k,c,\theta,p,\lambda} \int_0^\infty \left\{ \gamma\, v \left[f(k) - \left(\frac{1-\theta}{1-p} \right) rk + \frac{p\, c}{1-p} \right] + u[c] \right\} e^{-\rho t} dt$$

s.t.

$$u'(c) - \tfrac{\lambda}{1-p} = 0 \tag{15a}$$

$$-\left(\tfrac{1-\theta}{1-p}\right) r\lambda + \rho\lambda = \dot{\lambda} \tag{15b}$$

$$\left(\tfrac{1-\theta}{1-p}\right) rk - \tfrac{c}{1-p} = \dot{k} \tag{15c}$$

$$\theta, p \geq 0 \text{ and } \lim_{t\to\infty} \lambda k e^{-\rho t} = 0 \tag{15d}$$

where $\gamma \in (0,\infty)$ represents the social weight attached to the welfare of the workers. If $\gamma \to 0$, the government is only concerned about the capitalists, whereas it only cares about the workers when $\gamma \to \infty$. The current value Hamiltonian for this problem is given by

$$\mathcal{H} = \gamma v[\cdot] + u[c] + \mu_1(u' - \tfrac{\lambda}{1-p}) + q_1\lambda \left(-\left(\tfrac{1-\theta}{1-p}\right) r + \rho \right) + q_2 \left(\left(\tfrac{1-\theta}{1-p}\right) r\, k - \tfrac{c}{1-p} \right)$$

where q_1 is the *social* marginal value of the *private* marginal value λ which measures how valuable more capital is in terms of utility. Furthermore, q_2 is the *social* marginal value of more capital k. The shadow price μ_1 measures how to keep the capital owners on their demand curve.

[14]From now on time subscripts are dropped for convenience whenever it is clear that a particular variable depends on time.

The necessary first order conditions involve

$$\mathcal{H}_k: \quad \gamma\, v'[\cdot]\left(f' - \left(\tfrac{1-\theta}{1-p}\right) r\right) + q_2 \left(\tfrac{1-\theta}{1-p}\right) r = \rho q_2 - \dot{q}_2 \tag{16a}$$

$$\mathcal{H}_c: \quad \gamma\, v'[\cdot]\tfrac{p}{1-p} + u'[\cdot] + \mu_1 u''[\cdot] - q_2 \tfrac{1}{1-p} = 0 \tag{16b}$$

$$\mathcal{H}_\theta: \quad \theta\left\{\gamma v'[\cdot]\tfrac{rk}{(1-p)} + q_1 \lambda \tfrac{r}{1-p} - q_2 \tfrac{rk}{(1-p)}\right\} = 0 \tag{16c}$$

$$\mathcal{H}_p: \quad p\left\{(\gamma v'[\cdot] - q_2)\left[\tfrac{c-(1-\theta)rk}{(1-p)^2}\right] - \lambda\left(\tfrac{\mu_1 + q_1 r(1-\theta)}{(1-p)^2}\right)\right\} = 0 \tag{16d}$$

$$\mathcal{H}_\lambda: \quad -\tfrac{\mu_1}{1-p} + q_1\left(-\left(\tfrac{1-\theta}{1-p}\right) r + \rho\right) = \rho q_1 - \dot{q}_1 \tag{16e}$$

where (16c) and (16d) have to hold with complementary slackness due to the requirement that θ and p cannot be negative.[15] Furthermore, the first order conditions require that the equations (15a), (15b) and (15c) and the transversality conditions $\lim_{t\to\infty} q_1 \lambda e^{-\rho t} = 0$ and $\lim_{t\to\infty} q_2 k e^{-\rho t} = 0$ have to hold.

At time zero, the initial value of the consumer's marginal utility is unconstrained, i.e. initial λ is unconstrained.[16] Thus, the associated costate variable q_1 at time 0 is zero, i.e. $q_1(0) = 0$. But in the model it turns out that $q_1(t) = 0$ for all t. This can be shown by the following arguments:

Let us focus on interior solutions. Equations (16c) and (16d) can be rearranged as

$$\begin{aligned} \mathcal{H}_\theta : (\gamma v' - q_2)\frac{rk}{1-p} &= -q_1\,\lambda\,\frac{r}{1-p} \\ (\gamma v' - q_2) &= -q_1\,\frac{\lambda}{k} \qquad (17) \\ \mathcal{H}_p : (\gamma v' - q_2)\frac{c-(1-\theta)rk}{(1-p)^2} &= \lambda\,\frac{\mu_1 + (1-\theta) r q_1}{(1-p)^2}. \qquad (18) \end{aligned}$$

[15]For example, one might argue that negative θ is a form of wage tax and should not be ruled out a priori. However, as can be verified from (16c) negative θ is only possible in the model when $\gamma = 0$ and the government would not really be that benevolent anymore. For the purposes of this analysis and following Judd [1985] I rule out negative θ and p.

[16]This is a standard result. It can be inferred from the fact that, as is usual, for the capital owners' problem initial λ and initial c are not restricted by an initial condition. See Chamley [1986], p. 616, or Turnovsky [2000], p. 403, for the same point.

When we substitute for $(\gamma v' - q_2)$ from (17) in (18) we obtain

$$\begin{aligned} \mathcal{H}_p : -q_1 \frac{\lambda}{k}\left(\frac{c-(1-\theta)rk}{(1-p)^2}\right) &= \lambda \frac{\mu_1 + (1-\theta)rq_1}{(1-p)^2} \\ -q_1 \frac{c}{k} &= \mu_1. \end{aligned} \tag{19}$$

Next, we substitute this expression in (16e) and get

$$q_1 \frac{c/k}{1-p} + q_1 \left(-\left(\frac{1-\theta}{1-p}\right) r + \rho\right) = \rho q_1 - \dot{q}_1.$$

This is a homogeneous, linear differential equation. Integrating from time 0 up to some time t yields

$$q_1(t) = q_1(0) e^{-\int_0^t \Delta_s ds} \quad \text{where} \quad \Delta_s \equiv \left[\frac{c/k}{1-p} - \left(\frac{1-\theta}{1-p}\right) r\right] \tag{20}$$

As $q_1(0) = 0$, we have indeed found that $q_1(t)$ is 0 for any t. Clearly, this also holds for q_1 in steady state. Thus, $q_1(t) = 0$ for all $t \in [0, \infty)$.[17]

We will now restrict the analysis to the long-run when the economy is at a steady state, balanced growth position with $\dot{k} = \dot{\lambda} = \dot{c} = \dot{q}_1 = \dot{q}_2 = 0$. Suppose the government attaches some positive weight on the workers' welfare, $\gamma > 0$, and their marginal utility is positive, $v'[\cdot] > 0$. From equation (15b) with $\dot{\lambda} = 0$ in steady state we have $\lambda\left(\rho - r\left(\frac{1-\theta}{1-p}\right)\right) = 0$ where $\lambda \geq 0$. Substituting this in (16a) implies for the

[17] An intuitive explanation for this property of the model is the following: Initially the constraints on λ and the marginal utility of consumption are nonbinding. Later it turns out that - ignoring the way how investment subsidies are financed - any p will turn out to be optimal. Given the impact of investment subsidies on consumption and accumulation of the capital owners, any policy package θ, p will lead to the nonbinding of these constraints on λ. In a sense inappropriate choices of θ and p may lead to too less or too much consumption from a social point of view. Thus, the social planner chooses a policy package of θ and p that balances these effects and attaches a zero value, $q_1 = 0$, on λ and the marginal utility of consumption for all t in the optimum.

steady state that

$$\gamma v'[\cdot]\left(f' - \frac{1-\theta}{1-p}r\right) = 0$$

must hold. But firms do not pay the capital income taxes in this model and, thus, from profit maximization we have $f' = r$. But then we must have $\theta = p$ in an optimum. But this implies that the government, no matter whether it is relatively more pro-labour or pro-capital, chooses not to distort the accumulation decision.

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

This provides an example that a government can use different distortionary instruments to offset any distortions. In the present case 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.

The nondistortionary optimal policy $\theta_t = p_t$ has the following implications: When $\theta = p$ one obtains from equation (12) that

$$x = f(k) - rk + \frac{\theta c}{1-\theta}. \tag{21}$$

Thus, the equilibrium income of the workers is increasing in the consumption of the capital owners *and* in θ, because that raises tax revenues that can be transferred to the workers raising their total income.

Next, notice that the FOC for the capital stock, that is, equation (16a) holds if $f' = r = \rho$. Thus, in the optimum the marginal product of capital must equal the return on capital which in turn must equal the time preference rate, which is constant. This condition then pins down the capital stock to $\tilde{k}$ in steady state.[18] Equation (15c) implies $c = (1-\theta)r\tilde{k} = (1-\theta)\rho\tilde{k}$ in steady state so that (16d) becomes

$$p\left\{-q_1\frac{\lambda r(1-\theta)}{(1-p)^2}\right\} = 0. \tag{22}$$

where the expression in braces would have to be zero for an interior solution. Notice that $\lambda, r, \theta, p \geq 0$. But we know that $q_1 = 0$ for any t, including points in time when the economy is in steady state. This then implies that any p would satisfy this equation.

From (16e) we have that in steady state $-\frac{\mu_1}{1-\theta} - q_1 r = 0$. As $q_1(t) = 0$ for all $t \in [0,\infty)$ we find that $\mu_1 = 0$ in steady state. Thus, $q_1 = \mu_1 = 0$. Then $q_2 = \gamma\, v'[\cdot]$ by (16c) for an interior equilibrium and substitution of this into (16b) establishes that $\gamma v' = u'$ must hold. As the capital stock is fixed at $\tilde{k}$, which depends on ρ, and as $c = (1-\theta)\rho\tilde{k}$, the latter condition boils down to finding θ such that

$$\gamma\, v'[f(\tilde{k}) - \rho\tilde{k} + \theta\rho\tilde{k}] = u'[(1-\theta)\rho\tilde{k}]. \tag{23}$$

Denote the tax rate that solves this equation by $\tilde{\theta}$. Clearly as $\gamma \to \infty$ and the government is entirely pro-labour, the LHS becomes infinite and as a consequence $\theta = 1$ would be optimal, since $\lim_{c_t \to 0} u'[\cdot] = \infty$.[19] Furthermore by taking the differential of (23) with respect to θ and γ one easily verifies that $\tilde{\theta}$ is increasing in γ.

On the other hand, if $\gamma \to 0$, then $\mu_1 = q_1 = 0$ still applies. In this case equation

[18] Thus, as $t \to \infty$ the capital stock k_t approaches some time invariant constant $\tilde{k}$. From now on the tilde will denote variables in long-run steady state equilibrium.

[19] Rehme [1995] and Rehme [2002] obtain a similar result in an endogenous growth framework where redistribution occurs via productive government input financed by a capital income tax cum investment subsidy scheme.

(16c) boils down to

$$\theta \left\{ -q_2 \frac{rk}{(1-p)} \right\} = 0 \tag{24}$$

where $q_2 > 0$ from (16b). Thus, $\theta = 0$ would be optimal by the Kuhn-Tucker conditions, as the expression in braces would then be negative.

Proposition 2 *In an interior optimum the government equates the marginal social value of the workers to the social marginal value of the capital owners. The optimal capital income tax rate in the steady state $\widetilde{\theta}^*$ solves the equalization of the social marginal welfare of the workers and the capital owners, i.e., it solves*

$$\gamma\, v'[f(\tilde{k}) - \rho\tilde{k} + \theta\rho\tilde{k}] = u'[(1-\theta)\rho\tilde{k}] \tag{25}$$

where the optimal investment subsidy p is equal to 1. The optimal capital-income-cum-investment-subsidy-tax-rate (CICIST) $\widetilde{\theta}^*$ is increasing in the social weight γ attached to the welfare of the workers. For low values of γ, the capital income tax rate is zero, but there exists a γ^* such that for $\gamma > \gamma^*$ the optimal capital income tax rate $\widetilde{\theta}^*$ is positive. If $\gamma \to \infty$, then the optimal capital income tax rate becomes 1.

Notice that this result does not depend on production externalities or any other things, the capital income taxes may be used for, except for using part of the revenue for investment subsidies.

Of course, the government does not always place so much weight on the workers. But the model implies that as the workers get more social weight the social planner would choose higher capital income taxes under the capital-income-cum-investment-subsidy-tax (CICIST) scheme.[20]

[20]Notice that $\tilde{k}$ would be the same under any other capital income tax scheme for which it is shown

It is not entirely clear why workers should evaluate a consumption good any differently from a capital owner. For that reason it is now assumed that $v[x] = u[c]$ for any $x = c$ so that the two groups have the same utility function. As I am only interested in conditions under which the capital income tax is zero in the long-run let us assume that the utility functions are of the constant relative risk aversion (CRRA) type: $u[c] = \frac{c^{1-\beta}-1}{1-\beta}$ and $v[x] = \frac{x^{1-\beta}-1}{1-\beta}$. Then (23) would be given by

$$\begin{aligned} \gamma\left(f(\tilde{k}) - (1-\theta)\rho\tilde{k}\right)^{-\beta} &= \left((1-\theta)\rho\tilde{k}\right)^{-\beta} \\ \frac{f(\tilde{k})}{(1-\theta)\rho\tilde{k}} &= \gamma^{\frac{1}{\beta}} + 1. \end{aligned}$$

As $r = \rho = f'$ the fraction $\frac{\rho\tilde{k}}{f(\tilde{k})} \equiv \alpha$ corresponds to the share of broad capital in production. The latter is usually found to be bigger than one half. See, for example, Mankiw et al. [1992]. Hence, the optimal θ is determined by

$$\tilde{\theta} = \frac{\alpha(\gamma^{\frac{1}{\beta}} + 1) - 1}{\alpha(\gamma^{\frac{1}{\beta}} + 1)} \tag{26}$$

and is increasing in the share of capital so that distribution matters. Furthermore, the optimal long-run capital income tax rate is positive as long as

$$\gamma > \left(\frac{1-\alpha}{\alpha}\right)^{\beta}. \tag{27}$$

In the macroeconomics literature it is common to argue that α, that is, the share of broad capital is bigger than one half. Furthermore, there is evidence that β, that is, the *inverse* of the intertemporal elasticity of substitution of consumption between different dates is quite large. That would imply that one would need a sufficiently large γ to

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 CICIST in comparison to those other capital income tax schemes.

obtain the result that $\tilde{\theta}$ is positive in the long run.

However, if $0 \leq \gamma < \left(\frac{1-\alpha}{\alpha}\right)^{\beta}$, then $\tilde{\theta} = 0$ would follow. This is so because $\tilde{\theta} = 0$ implies $u' = q_2 = \lambda$ by (16b) and (15a). Thus, as long as $\tilde{k}$ satisfies $r = f' = \rho$ and as long as $\gamma < \left(\frac{1-\alpha}{\alpha}\right)^{\beta}$ we have $\gamma v' < u'$ so that indeed $\tilde{\theta} = 0$ is optimal in those circumstances.

Corollary 1 *Let the agents possess the same constant relative risk aversion utility functions. Under a capital-income-cum-investment-subsidy-tax (CICIST) scheme the optimal capital income tax rate* $\tilde{\theta}$ *is non-zero if the social planner attaches sufficient weight on the welfare of the workers* $\gamma > \left(\frac{1-\alpha}{\alpha}\right)^{\beta}$. *In contrast, if* $\gamma < \left(\frac{1-\alpha}{\alpha}\right)^{\beta}$, *then* $\tilde{\theta} = 0$ *is optimal. Hence, under CICIST the income distribution, preferences and the social weight of the workers determine whether the optimal capital income taxes are zero in the long run.*

Thus, under the (nondistorting) capital income tax scheme under consideration (CICIST) distributional and preference parameters matter and that may complement the results of Judd [1985] and Chamley [1986]. Importantly, the proposition establishes that there may be instances when capital income taxes are optimally non-zero in the long run. Here the non-zero tax result may hold even though the agents have very different utility functions or all utility functions are of the general CRRS type. Thus, even though the social planner only concentrates on the first order conditions of the private sector and does not explicitly know the agents' final decision rules, and even though he/she has freedom to choose consumption and capital independently, redistribution and so capital income taxes may optimally be non-zero in the long-run.

3.1 Numerical simulation

In order to get a feeling for the nature of the solutions let us use a numerical simulation based on calibrations from the business cycle literature. In particular, I rely on Walsh [2003], ch. 2, who bases his parameters on Cooley and Prescott [1995], p. 22, and Cooley and Hansen [1995], p. 201. For quarterly data for the United States he uses the following calibrated values for a standard money-in-the-utility function, real business cycle model.[21] For the share of broad capital I follow Barro and Sala–i–Martin [2004] and set it equal to 0.75.

Table 1 about here.

With these values it turns out that if $\gamma > \left(\frac{1-\alpha}{\alpha}\right)^{\beta} = 0.11$ the condition for positive optimal capital income tax rates is met. Table 2 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 2 about here.

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,

[21] The numbers reported for the U.S. command wide support in that literature. As the numbers are related to discrete time models, I have converted the discrete time discount rate of 0.989 to our continuous time set-up. The corresponding value of ρ was 0.011 as reported in the table.

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.

But as a consequence the theoretical arguments of this paper's model then suggest that the result that zero capital income taxes and so no pure redistribution be optimal in the long run depends on 1. the social weight attached to the workers, 2. the income share of capital in production and 3. the intertemporal elasticity of substitution.

4 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-redistribution government would choose this nondistortionary policy in the model.

Furthermore, it is found that whether or not redistribution and capital income taxes are optimally zero in the long run 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.

The findings in this paper suggest that it might not be a bad thing to keep in place, for a longer period of time, the more generous investment promoting subsidy schemes that have been introduced during the recent economic crisis.

A Broad capital, equation (1) 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

$$Y_t = A \cdot F(K_t, H_t, L_t) \tag{28}$$

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

$$\begin{aligned} Y_t &= L_t \cdot A \cdot F(\kappa_t, h_t, 1) \\ y_t &= A \cdot f(\kappa_t, h_t) \end{aligned} \tag{29}$$

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

$$\frac{\partial y_t}{\partial h_t} = r_t^h, \quad \text{and} \quad \frac{\partial y_t}{\partial \kappa_t} = r_t^\kappa, \tag{30}$$

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 (28).

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

$$r_t^\kappa = r_t^h. \tag{31}$$

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 (30) and (31) 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 (28) 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 (29). If κ_t were the measure of broad capital, then production would be describes by

$$y_t = A \cdot f(\kappa_t, h_t) = A \cdot \kappa^{\nu} \cdot \hat{f}\left(1, \frac{h_t}{\kappa_t}\right) = A \cdot \kappa^{\nu} \cdot \hat{f}\,(1, \chi) \tag{32}$$

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

$$\begin{aligned} y_t &= A \cdot f(\kappa_t, h_t) = A \cdot k_t^{\nu} \cdot k_t^{-\nu} f(\kappa_t, h_t) \\ &= A \cdot k^{\nu} \tilde{f}\left(\frac{\kappa_t}{k_t}, \frac{h_t}{k_t}\right) = A \cdot k^{\nu} \tilde{f}\left(\frac{1}{1+\chi}, \frac{\chi}{1+\chi}\right). \end{aligned} \tag{33}$$

It is then not difficult to verify that, given the homogeneity of the functions $\hat{f}$ and $\tilde{f}$, the productions function in (32) and (33) are equivalent and $\hat{f} = \hat{f}\,(1,\chi) = \tilde{f}\,(1,\chi) = \tilde{f}$ when factoring out $(1+\chi)^{-\nu}$ in $\tilde{f}$. Thus, the paper's production function corresponds to (33) 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 (1) 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 $y_t = B \cdot h_t^{\beta_1} \kappa_t^{\beta_2}$ where $0 < \beta_1 + \beta_2 < 1$, equality of the marginal products implies

$$\left(\frac{\partial y_t}{\partial h_t}=\right) \quad \beta_1 \frac{y_t}{h_t}=\beta_2 \frac{y_t}{\kappa_t} \quad \left(=\frac{\partial y_t}{\partial \kappa_t}\right).$$

Thus, $\frac{\beta_1}{\beta_2}\cdot\kappa_t = h_t$. Substituting this into the Cobb-Douglas production function yields

$$y_t = B\left(\frac{\beta_1}{\beta_2}\right)^{\beta_1}\kappa_t^{\beta_1+\beta_2}. \tag{34}$$

Now let $\beta_1+\beta_2=\alpha$ and normalize so that $B\left(\frac{\beta_1}{\beta_2}\right)^{\beta_1}=1$ 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 $h_t=\frac{\beta_1}{\beta_2}\cdot\kappa_t$, yields

$$\begin{aligned} y_t &= B\cdot k_t^{\beta_1+\beta_2}k_t^{-\beta_1-\beta_2}h_t^{\beta_1}\kappa_t^{\beta_2} = B\cdot k_t^{\beta_1+\beta_2}\left(\frac{h_t}{k_t}\right)^{\beta_1}\left(\frac{\kappa_t}{k_t}\right)^{\beta_1} \\ &= B\cdot k_t^{\beta_1+\beta_2}\left(\frac{\beta_1}{\beta_2}\right)^{\beta_1}\left(1+\frac{\beta_1}{\beta_2}\right)^{-\beta_1-\beta_2}. \end{aligned} \tag{35}$$

One easily verifies that this is equivalent to (34). Furthermore, setting $\alpha=\beta_1+\beta_2$, where β_1 represents the share of human capital and β_2 the physical capital share, scaling B so that $B=\left(\frac{\beta_1}{\beta_2}\right)^{-\beta_1}\left(1+\frac{\beta_1}{\beta_2}\right)^{\beta_1+\beta_2}$ and by the population normalization used in the main text it follows that (35) is equivalent to what is used in the latter part of the main text.

B An alternative argument why $q_1(t)=0, \forall t$

For the capital owners to be on their demand curves requires $u'=\frac{\lambda}{1-p}$. so that the optimal consumption of the capital owners satisfies

$$c=c(\lambda,p).$$

Given the concavity of $u(c)$ it turns out that

$$\frac{dc}{d\lambda}=\frac{1}{u''(1-p)}<0 \quad \text{and} \quad \frac{dc}{dp}=\frac{\lambda}{u''(1-p)^2}<0. \tag{36}$$

The government then solves the following problem

$$\max_{k,\theta,p,\lambda} \int_0^\infty \left\{ \gamma\, v \left[f(k) - \left(\frac{1-\theta}{1-p} \right) rk + \frac{p\, c(\lambda,p)}{1-p} \right] + u[c(\lambda,p)] \right\} e^{-\rho t} dt$$

s.t.

$$-\left(\tfrac{1-\theta}{1-p}\right) r\lambda + \rho\lambda = \dot{\lambda} \tag{37a}$$

$$\left(\tfrac{1-\theta}{1-p}\right) rk - \tfrac{c(\lambda,p)}{1-p} = \dot{k} \tag{37b}$$

$$\theta \geq 0 \;\text{ and }\; \lim_{t\to\infty} \lambda k e^{-\rho t} = 0 \tag{37c}$$

where $\gamma \in (0, \infty)$ represents the social weight attached to the welfare of the workers. If $\gamma \to 0$, the government is only concerned about the capitalists, whereas it only cares about the workers when $\gamma \to \infty$. The current value Hamiltonian for this problem is given by

$$\mathcal{H} = \gamma v[\cdot] + u[c(\lambda,p)] + q_1\lambda \left(-\left(\tfrac{1-\theta}{1-p}\right) r + \rho \right) + q_2 \left(\left(\tfrac{1-\theta}{1-p}\right) r\, k - \tfrac{c(\lambda,p)}{1-p} \right)$$

where q_1 is the *social* marginal value of the *private* marginal value λ. Furthermore, q_2 is the *social* marginal value of more capital k.

The necessary first order conditions are

$$\mathcal{H}_k: \quad \gamma\, v'[\cdot] \left(f' - \left(\tfrac{1-\theta}{1-p}\right) r \right) + q_2 \left(\tfrac{1-\theta}{1-p}\right) r = \rho q_2 - \dot{q_2} \tag{38a}$$

$$\mathcal{H}_\theta: \quad \theta \left\{ \left((\gamma v'[\cdot] - q_2) \tfrac{rk}{(1-p)} + q_1 \lambda \tfrac{r}{1-p} \right\} = 0 \tag{38b}$$

$$\begin{aligned}\mathcal{H}_p: \quad & p \left\{ (\gamma v'[\cdot] - q_2) \left[\tfrac{c-(1-\theta)rk}{(1-p)^2} \right] + \left(\gamma v' \tfrac{p}{1-p} + u' - q_2 \tfrac{1}{1-p} \right) \tfrac{dc}{dp} \right. \\ & \left. - q_1 \tfrac{\lambda r(1-\theta)}{(1-p)^2} \right\} = 0 \end{aligned} \tag{38c}$$

$$\mathcal{H}_\lambda: \quad \left(\gamma v' \tfrac{p}{1-p} + u' - q_2 \tfrac{1}{1-p} \right) \tfrac{dc}{d\lambda} + q_1 \left(-\left(\tfrac{1-\theta}{1-p}\right) r + \rho \right) = \rho q_1 - \dot{q_1} \tag{38d}$$

where (38b) and (38c) have to hold with complementary slackness since we ruled out negative θ and p. Furthermore, the equations (15a), (15b) and (15c) and the transversality conditions $\lim_{t\to\infty} q_1 \lambda e^{-\rho t} = 0$ and $\lim_{t\to\infty} q_2 k e^{-\rho t} = 0$ have to hold.

At time zero, the initial value of the consumer's marginal utility is unconstrained, i.e. λ is unconstrained. (See Chamley [1986], p. 616, or Turnovsky [2000], p. 403, for the same point.) Thus, the associated costate variable q_1 at time 0 is zero, i.e. $q_1(0) = 0$. To show that $q_1(t) = 0$ for all t we use the following arguments:

Let us focus on interior solutions. Equation (38b) implies that

$$\begin{aligned} (\gamma v' - q_2) \frac{rk}{1-p} &= -q_1 \frac{\lambda r}{1-p} \\ (\gamma v' - q_2) &= -\frac{q_1 \cdot \lambda}{k}. \end{aligned} \tag{39}$$

Thus, for equation (38c) we obtain for an interior solution

$$\left(\gamma v'[\cdot] - q_2\right)\left[\frac{c-(1-\theta)rk}{(1-p)^2}\right] + \left(\gamma v'\frac{p}{1-p} + u' - q_2\frac{1}{1-p}\right)\frac{dc}{dp} = q_1\frac{\lambda r(1-\theta)}{(1-p)^2}$$
$$-\frac{q_1\cdot\lambda}{k}\left[\frac{c-(1-\theta)rk}{(1-p)^2}\right] + \left(\gamma v'\frac{p}{1-p} + u' - q_2\frac{1}{1-p}\right)\frac{dc}{dp} = q_1\frac{\lambda r(1-\theta)}{(1-p)^2}.$$

Now let $\Phi \equiv \gamma v'\frac{p}{1-p} + u' - q_2\frac{1}{1-p}$. Then

$$\begin{aligned}\Phi\frac{dc}{dp} &= \frac{\lambda q_1 r(1-\theta)k}{(1-p)^2 k} + \frac{\lambda q_1\left(c-(1-\theta)rk\right)}{(1-p)^2 k}\\ \Phi &= \frac{\lambda q_1\left(\frac{c}{k}\right)}{(1-p)^2}\left(\frac{dc}{dp}\right)^{-1}\end{aligned} \tag{40}$$

Given our definition of Φ equation (38d) for $\mathcal{H}_\lambda$ is given by

$$\Phi\left(\frac{dc}{d\lambda}\right) + q_1\left(-\left(\frac{1-\theta}{1-p}\right)r + \rho\right) = \rho q_1 - \dot{q}_1$$

In this equation we substitute for Φ from (40) to get

$$\frac{\lambda q_1\left(\frac{c}{k}\right)}{(1-p)^2}\left(\frac{dc}{dp}\right)^{-1}\left(\frac{dc}{d\lambda}\right) + q_1\left(-\left(\frac{1-\theta}{1-p}\right)r + \rho\right) = \rho q_1 - \dot{q}_1. \tag{41}$$

Now from (36) one verifies that $\left(\frac{dc}{dp}\right)^{-1}\left(\frac{dc}{d\lambda}\right) = \frac{1-p}{\lambda}$. Thus, equation (41) simplifies to

$$q_1\left[\frac{c/k}{1-p} - \left(\frac{1-\theta}{1-p}\right)r\right] = -\dot{q}_1. \tag{42}$$

This is a homogeneous, linear differential equation. Integrating from time 0 up to time t yields

$$q_1(t) = q_1(0)e^{-\int_0^t \Delta_s ds} \quad \text{where} \quad \Delta_s \equiv \left[\frac{c/k}{1-p} - \left(\frac{1-\theta}{1-p}\right)r\right] \tag{43}$$

As $q_1(0) = 0$, we have indeed found the same result as expressed in the main text. The rest of the results can be derived in a way analogous to the one presented in the paper.

References

A. B. Abel. Optimal capital income taxation. Working Paper 13354, NBER, Cambridge, MA, 2007.

S. R. Aiyagari. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy*, 103:1158–1175, 1995.

R. J. Barro and X. Sala–i–Martin. *Economic Growth*. MIT Press, Cambridge, Massachusetts, 2nd edition, 2004.

C. Chamley. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54:607–622, 1986.

C. Chamley. Capital income taxation, wealth distrbution and borrowing constraints. *Journal of Public Economics*, 79:55–69, 2001.

J. C. Conesa, S. Kitao, and D. Krueger. Taxing capital? not a bad idea after all! *American Economic Review*, 99(1):25–48, 2010.

T. F. Cooley and G. D. Hansen. Money and the business cycle. In T. F. Cooley, editor, *Frontiers of Business Cycle Research*, pages 175–216. Princeton University Press, Princeton, N. J., 1995.

T. F. Cooley and E. C. Prescott. Economic growth and business cycles. In T. F. Cooley, editor, *Frontiers of Business Cycle Research*, pages 1–38. Princeton University Press, Princeton, N. J., 1995.

J. B. Davies, J. Zeng, and J. Zhang. 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, 2009.

D. Domeij and J. Heathcote. On the distributional effects of reducing capital taxes. *International Economic Review*, 45:523–554, 2004.

I. Fisher. Income in theory and income taxation in practice. *Econometrica*, 5:1–55, 1937.

H. P. Grüner and B. Heer. Optimal flat-rate taxes on capital - a re-examination of lucas' supply side model. *Oxford Economic Papers*, 52:289–305, 2000.

J. T. Guo and K. J. Lansing. Optimal taxation of capital income with imperfectly competitive product markets. *Journal of Economic Dynamics and Control*, 23:967–995, 1999.

R. E. Hall and D. W. Jorgenson. Tax policy and investment behaviour. *American Economic Review*, 57: 391–414, 1967.

R. E. Hall and D. W. Jorgenson. Application of the theory of optimal capital accumulation. In G. Fromm, editor, *Tax Incentives and Capital Spending*, pages 9–60. The Brookings Institution, Washington D. C., 1971.

L. E. Jones, R. E. Manuelli, and P. E. Rossi. On the optimal taxation of capital income. *Journal of Economic Theory*, 73:93–117, 1997.

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

K. L. Judd. Optimal taxation and spending in general competitive growth models. *Journal of Public Economics*, 71:1–26, 1999.

N. Kaldor. *An Expenditure Tax*. Allen and Unwin, London, 1955.

N. Kaldor. Alternative theories of income distribution. *Review of Economic Studies*, 48(5):83–100, 1956.

M. C. Kemp, N. van Long, and K. Shimomura. Cyclical and noncyclical redistributive taxation. *International Economic Review*, 34:415–429, 1993.

K. J. Lansing. Optimal redistributive capital taxation in a neoclassical growth model. *Journal of Public Economics*, 73:423–453, 1999.

R. E. Lucas. Supply-side economics: An analytical review. *Oxford Economic Papers*, 42:293–316, 1990.

N. G. Mankiw, D. Romer, and D. N. Weil. A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 152:407–437, 1992.

L. Pasinetti. Rate of profit and income distribution in relation to the rate of economic growth. *Review of Economic Studies*, 29:267–279, 1962.

G. Rehme. Redistribution, income cum investment subsidy tax competition and capital flight in growing economies. Working Paper ECO 95/16, European University Institute, Florence, Italy, 1995.

G. Rehme. Essays on distributive policies and economic growth. Ph.D. Thesis, European University Institute, Florence, Italy, 1998.

G. Rehme. Distributive policies and economic growth: An optimal taxation approach. *Metroeconomica*, 53:315–338, 2002.

G. Rehme. Optimal capital income taxation, investment subsidies and redistribution in a neoclassical growth model. Discussion Paper 188, TU Darmstadt, 2007.

G. Rehme. Capital depreciation allowances and redistribution in a neoclassical growth model. Mimeo, Technische Universität Darmstadt, 2009.

E. Saez. Optimal progressive capital income taxes in the infinite horizon model. Working Paper 9046, NBER, Cambridge, MA, 2002.

P. A. Samuelson. Tax deductibility of economic depreciation to insure invariant valuations. *Journal of Political Economy*, 72:604–606, 1964.

T. J. Sargent and L. Ljungqvist. *Recursive Macroeconomic Theory*. MIT Press, Cambridge, Massachusetts, 2nd edition, 2004.

S. Selim. Optimal taxation in a two-sector economy with heterogeneous agents. *Economics Bulletin*, 30(1):534–542, 2010.

S. J. Turnovsky. *Methods of Macroeconomic Dynamics*. MIT Press, Cambridge, Mass., 2nd edition, 2000.

H. Uhlig and N. Yanagawa. Increasing the capital income tax may lead to faster growth. *European Economic Review*, 40:1521–1540, 1996.

C. E. Walsh. *Monetary Theory and Policy*. MIT Press, Cambridge, Massachusetts, 2nd edition, 2003.

I. Werning. Optimal fiscal policy with redistrbution. *Quarterly Journal of Economics*, 122:925–967, 2007.

Table 1: Baseline Parameter Values

α	ρ	β
0.75	0.011	2

Based on Walsh [2003], p. 75.

Table 2: Optimal Capital Income Tax Rates $\widetilde{\theta}$

γ	0.20	0.50	0.80	1.00	2.00	5.00
$\widetilde{\theta}$	0.08	0.22	0.30	0.33	0.45	0.59

γ	10	20	50	100	1000
$\widetilde{\theta}$	0.68	0.76	0.83	0.88	0.96