RCEA Network

Rimini Centre for Economic Analysis
Working Paper Series

wp 20-20

Optimal Factor Taxation in A Scale Free Model of Vertical Innovation

Barbara Annicchiarico
Valentina Antonaroli
Alessandra Pelloni

Optimal Factor Taxation in A Scale Free Model of Vertical Innovation*

Barbara Annicchiarico† Valentina Antonaroli‡ Alessandra Pelloni§

June 2020

Abstract

The objective of the paper is to study how the tax burden arising from an exogenous stream of public expenditures and transfers should be distributed between labor and capital in a scale-less endogenous growth model, where the engine of growth are successful innovations. Our laboratory is a prototypical quality ladder model with a labor/leisure choice where R&D productivity is decreasing in the size of the economy. This decreasing productivity removes scale effects, which are a controversial prediction of first-generation endogenous growth models. Our contribution is to show that even when labor supply has no effects on growth in the long run, it will still be optimal to tax capital, for reasonable parametrizations of the model. This is true even if the long-run growth rate decreases, with respect to the initial situation in which capital income is not taxed.

Keywords: *Endogenous growth, Scale effects, Capital Income Taxation, Welfare effect.*

JEL classification: *O41, E62, H21*

*The paper has been presented at seminars at the Department of Economics, Free University of Berlin February 2019, and at the X Workshop IBEO - Institutions, Individual Behavior and Economic Outcomes - Alghero July 2019. We wish to thank all participants for very useful comments. and suggestions.

†Corresponding author: Department of Economics and Finance, University of Rome "Tor Vergata"; Via Columbia 2, Roma, 00133, e-mail address: barbara.annicchiarico@uniroma2.it.

‡Department of Economics and Finance, University of Rome "Tor Vergata"; e-mail address: valentina.antonaroli@gmail.com.

§Department of Economics and Finance, University of Rome "Tor Vergata", e-mail address: alessandra.pelloni@uniroma2.it.

1 Introduction

Shifting the tax burden from labor to capital is attracting more and more interest among economists and policy makers as a way to contrast the increased inequality in income and wealth distribution observed in many countries in the last decades. See e.g. Piketty and Saez (2013), Bastani and Waldenström (2018) and Saez and Zucman (2019). On the other hand, the traditional argument against capital taxation has been developed along two main lines. The first, started with Chamley (1985) and Judd (1985), who show that in deterministic infinite-horizon settings taxes on capital income should be avoided in the long run because they introduce distortions in consumption versus saving decisions that pile up over time, provided authorities have access to a commitment technology. The second, started with Atkinson and Stiglitz (1976) life cycle model with inequality in earnings ability, where progressive income taxation is more efficient than differential commodity taxation and therefore capital income taxation.

In coherence with these latter views, statutory rates on capital income have fallen sharply in OECD countries since the late 1980s, both at the corporate and at the personal level, and little effort has been made for the international coordination of tax policy which effectively taxing capital would need.[1]

In this paper, we study the trade off between taxing capital income and taxing labor income through the lenses of knowledge-driven growth theory. Specifically, we use a scale-less quality ladder growth model, where productivity-enhancing innovations result from profit-motivated R&D investment. This leads us to an argument in favor of capital income taxation that is based purely on efficiency grounds and that can complement those based on equity grounds.

According to innovation-led growth theory, in all its variants, the engine of growth is the discovery of ideas for new processes and products. It is technically possible for any number of people to use an idea at the same time: using an idea more intensively, therefore, increases efficiency, because the cost of its discovery per user will be lower. A higher level of economic activity will allow this more intense use. Shifting the burden of taxation from labor to capital income can then increase efficiency, because taxation of labor income may deter employment, thus reducing the static benefits deriving from the use of already existing ideas.

Our specific contribution is to show that this mechanism is powerful enough to obtain the result of a welfare enhancing positive shift of the tax burden from labor to capital, even in a model in which the rate of growth is not increasing in the scale of the economy.

That reducing the tax rate on labor income, while increasing that on capital income,

[1] Financial globalization and international tax competition have contributed to the decline in capital taxation because with free capital flows each country faces a highly elastic capital tax base. Braulke and Corneo (2004) show that, even with perfect capital mobility, there will always be a country that benefits from introducing a tax on capital.

may be efficient has in fact already been shown in the literature using models with a strong scale effect. Pelloni and Waldmann (2000) and Amano et al. (2009) show that taxing capital can be efficient in a model with learning by doing *à la* Romer (1986), while Aghion et al. (2013) show the same in a model of vertical innovation. Finally, Chen et al. (2017) and Long and Pelloni (2017), using models of horizontal innovation, show that capital taxation may increase, respectively, growth and welfare.

However, the first generation endogenous growth models, these tax analyses are based on, have been questioned exactly because of the increasing relationship between growth and labor they feature. For example, in an economy with increasing population these models would predict an ever increasing per capita GDP growth rate. Moreover, as pointed out by Jones (1995) firstly and insisted on by Bloom et al. (2017) more recently, in the data the resources devoted to R&D in advanced countries appear to increase steadily, while the rate of GDP per capita growth shows no trend.[2]

One could be justified in believing that the result of positive capital taxation obtained in the first-generation endogenous growth models mentioned above is made possible by the debated positive link between growth and labor they incorporate. Our simple model proves this belief to be wrong. Since Jones (1995) has raised the issue, theorists have begun to think of ways to endogenize R&D and innovation in growth models without giving rise to the strong scale effect. The first approach, proposed by Jones (1995) himself, is the semi-endogenous solution, according to which the more technical knowledge has been already accumulated, the more R&D effort it takes to increase technical knowledge to a given percentage.[3]

An alternative fully endogenous growth approach has also been characterized by the insight that what matters for productivity growth is not the total amount of R&D resources, but rather its share in terms of GDP.[4] In the words of Cozzi (2017, p. 28), "if each representative worker spends a larger fraction of his/her day researching – and the rest of the day producing in manufacturing – the growth rate of its productivity will be higher." Our model incorporates this insight, simply by assuming that innovation in a

[2]Chu and Cozzi (2019) distinguish a "market size effect", under which the growth rate of technologies is increasing in the amount of labor that *uses* the technologies, and a "scale effect" under which the growth rate of technologies is increasing in the amount of labor devoted to creating the new technologies. The two effects are related because in a larger economy, *coeteris paribus*, both effects arise.

[3]Well known characteristics of these models are that the rate of TFP growth is increasing in population growth and is independent of tax policy parameters. However, these controversial features may not arise when human capital accumulation and/or returns to specialization are considered. See e.g. Bucci (2015) and Strulik et al. (2013).

[4]Many mechanisms have been proposed in the literature to explain why this maybe the case. The assumptions differ widely, implying various limitations to each particular model. The early papers by Smulders and Van de Klundert (1995), Peretto (1998), Young (1998), Dinopoulos and Thompson (1999) and Howitt (1999) allow, in different ways, the number of varieties of goods to expand with population, so that research efforts dilute across more varieties. According to Matsuyama (1992) what count is research outlays per entrepreneurial capacity, while Zeira (2011) develops a model with patent races. Sequeira et al. (2018) introduce a knowledge production function featuring complexity effects in R&D, while Cozzi (2001) and Cozzi and Spinesi (2006) focus on the interaction between copying and inventing.

line of products is more complex and costly depending on the size of the line itself, as measured by generated value added, as in Barro and Sala-i Martin (2004).

In our model, the economy is populated by an infinitely lived representative household deriving utility from consumption and leisure. In the final good sector perfectly competitive firms use labor and differentiated intermediates as production inputs. Each intermediate has a quality ladder along which improvements occur thanks to R&D activity. This activity leads to monopoly power in producing the good to the successful innovator. The price of each intermediate is then higher than its marginal cost. This "market power" distortion is familiar from the standard analysis of monopoly. Each new discovery in a certain line erases the value of the previous discovery: this is known in the literature as a "creative destruction" or "business stealing" effect. This effect represents an externality because the decision to invest is based on the present discounted value of the innovator's profits only, while the social value of the innovation factors in the loss represented by the value of the previous innovation. This creates the theoretical possibility that the rate of innovation is too high. A third market failure stems from the incomplete appropriability of the social surplus from a new invention. Part of this social surplus goes to labor income and is not taken into account when a potential innovator, motivated by future profits, decides whether to engage in R&D. The private value of an innovation falls short of its social value: the rate of innovation is too low.

We follow the Ramsey approach in restricting the tax instruments available: in particular we only consider linear time-invariant income taxes. The government raises a constant fraction of income as fiscal revenue, a constant fraction of this revenue is transferred back to consumers lump sum and the government budget is balanced at all times. The tax rates must adjust endogenously.[5] A fourth distortion in the model is represented by government expenditures whose amount depend on income. This is not taken into account by agents when deciding how much to work and save. We label this externality the "weight of government" distortion.

The tax mix affects all of the four distortions discussed above. A tax shift from labor to capital will encourage employment, so that the demand and therefore the production for each intermediate goes up. This increases social welfare because, due to the "market power" distortion, the production of intermediates is too low. A larger extent of the market for a new good means that the cost of a blueprint can be shared by more users. Decreasing the tax on labor could in theory also reduce the "appropriability" distortion and increase growth, even if the tax on capital goes up. Specifically, more employment increases the total quantity of rents that can be captured by successful innovators. However, in our model more employment also means that additional research input is needed

[5] By a tax on capital income, as in the Chamley and Judd models, we mean a tax on income from savings. In our model there is no capital in the physical sense and wealth consists in the value of firms shares. The decision of households to finance firms is akin to the decision to accumulate physical capital in the standard Ramsey model.

for innovating, as seen above, so that the two effects cancel out in equilibrium and the pre-tax growth rate is unaffected. This means that increasing the tax rate on capital will always compress growth and worsen the appropriability distortion.

Finally, both higher labor and capital taxes will decrease the weight of government friction, but to a different extent. Actually, following a tax shift from labor to capital, households will benefit from higher current consumption, but will also suffer less leisure time and lower consumption growth.

Thanks to the simple structure of our model, it is easy to calculate the optimal tax rate on capital income for reasonable parameters' values. This rate is not only positive but sizable: about 24 percent in our baseline case. The optimal tax rate on capital income will, however, be generally lower than the one bearing on labor income. In particular, we find that optimal taxation on capital should be higher, the higher the fiscal revenue to be raised as a percentage of income and the higher the capital's income share. This last finding is particularly relevant given the observed increase in the share of capital income since the 1980s, as shown by Akcigit and Ates (2019). Furthermore, we find that the optimal tax rate on capital is increasing in the compensated elasticity of labor supply and in the time discount rate, and decreasing in the intertemporal elasticity of substitution in consumption. Another interesting result of our analysis is that a larger size of the public sector, whether due to public consumption or higher transfers to households, leads to a higher optimal tax rate on capital.

Generally, in papers on taxation and growth the focus of tax experiments is often on GDP growth, in the belief that growth effect will always prevail on level effects as regards welfare calculations, bar distributional considerations. However, we show that this is not necessarily the case. Indeed in calibrated examples, our counterintuitive result arises for parameters' values well within the range of selections adopted in other settings in public finance, quantitative growth theory and business cycle analysis. To repeat, in our model it can be optimal to reduce growth even if growth is too low to begin with, when initial consumption is also too low.

Transitional dynamics effects may be important for taxation: optimizing a long-run economic position is different from optimizing over the entire dynamic path. However we show that although our model incorporates features that may give rise to indeterminacy, here the economy always follows a unique unstable balanced growth path (BGP).

The rest of the paper is organized as follows. The next section is devoted to the related literature on optimal taxation, section 3 characterizes the model and describes the balanced growth path, section 4 shows how to find the Ramsey planner's solution, section 5 presents the social planner's solution, section 6 reports numerical calculations and some sensitivity and robustness checks, and the last section concludes.

2 Related Literature

This paper contributes to the literature on the optimal taxation of labor and asset income.[6] This is a complex problem treated in many sub-literatures. What follows is not a complete overview of the different approaches and findings, but a highly selective drawing of some key policy inferences. The basic aim is showing that our recommendation of a positive capital taxation is based on assumptions that differ from those whose implications have been studied in the works we consider in what follows.

The classic results by Chamley and Judd have been recently reconsidered by Straub and Werning (2020), who show that the results hold only on very restrictive conditions. In particular, in the Judd model, with two classes of agents, an intertemporal elasticity of substitution bigger than one is needed, while in the Chamley representative agent model the result holds with recursive non additive utility but then zero wealth and zero labor taxes will obtain asymptotically. In infinite horizon models, taxing capital can increase social welfare if the economy has an informal sector or if there is shifting between labor and capital income (e.g. Correia 1996, Peñalosa and Turnovsky 2005 and Reis 2011). With imperfectly competitive product or capital markets, or with search costs in the labor market, it may be optimal to tax or subsidize capital income (e.g Guo and Lansing 1999 and Chamley 2001). Finally, grounds for taxing capital are public expenditure in the utility or in the aggregate production function (e.g. Martin 2010 and Ben-Gad 2017).

The Atkinson and Stiglitz (1976) result only obtains with no complementarity between leisure and consumption and homogeneity of consumption sub-utility. Capital income taxation becomes optimal if future earnings are uncertain and insurance markets are missing, if low ability people have a higher rate of time discount or a lower return on their investments or if inequality in life-time resources is also due to bequests (Cremer et al. 2003, Jacobs and Schindler 2012 and Piketty and Saez 2013). In dynamic Mirrlees models, where the set of policy instruments is constrained only by informational frictions, it is optimal to distort savings decisions to improve the incentive to work, when agents' abilities are stochastic and have private information (see Golosov et al. 2006).

In OLG models taxing capital may be a way to redistribute resources across cohorts or to mimic unavailable age-dependent taxes (e.g. Erosa and Gervais 2002, Conesa et al. 2009 and Bastani et al. 2013), while Pirttilä and Tuomala (2001) show that a positive tax on capital income is desirable if an increase in investment leads to more labor income inequality.

Representative agent endogenous growth models with no market failures have been used to support the case against capital taxation on the grounds of its adverse effects on growth. See e.g. the survey in Jones and Manuelli (2005). Taxing capital may, however,

[6] Indeed in our model in which the rate of returns are not uncertain and are equal across household, taxes on the stock of capital or from the income stream it generates are equivalent.

be efficient when government spending enters the utility or production function, a result echoing the analogous one obtained in exogenous growth models (e.g. Baier and Glomm 2001, Park and Philippopoulos 2004, Chen and Lee 2007 and Marrero and Novales 2007). In models in which human capital can be accumulated without bound and transmitted from one generation to the next as physical capital can, the difference between the two forms of capital tends to disappear and so does the rationale for a difference between the tax treatment between them (see e.g. de Hek 2006, Chen et al. 2011 and Chen and Lu 2013): taxing labor income, i.e. income from human capital, becomes as bad for growth as taxing income from savings.

Our paper is closely related to Zeng and Zhang (2002), whose positive analyses of taxation in non-scale Schumpeterian R&D models show that long-run growth is independent of both consumption and labor-income taxes, and negatively affected by capital income taxes. Peretto (2007) shows that, under restricting assumptions, a tax on dividends, by distorting the firms' choices between investing in quality improvements and investing in new products, may, differently from taxes on capital gains and taxes on corporate incomes, increase growth and therefore welfare. A positive capital taxation result is obtained by Chen et al. (2019) in a model of semi-endogenous growth. As is standard in semi-endogenous models, in Chen et al. (2019) long-run growth is policy invariant. This implies that taxing capital is not bad for growth in their model and makes the result somewhat less surprising than ours. Our work concentrates on the general equilibrium effects of taxation.

3 The Model

We extend the prototypical Schumpeterian framework of Grossman and Helpman (1991), in the de-scaled version proposed by Barro and Sala-i Martin (2004, chapter 7), by considering a labor/leisure choice for identical households and by introducing taxation on labor and capital income. The government collects tax revenues to finance public consumption and transfers for households.

3.1 Households

There exists a continuum of length one of infinitely lived identical households. Each household has preferences represented by

$$U_t = \int_{t=0}^{\infty} e^{-\rho t} \frac{C_t^{1-\sigma} \left(1 - k(1-\sigma) H_t^{1+\frac{1}{\eta}}\right)^{\sigma} - 1}{1-\sigma} \mathrm{d}t, \tag{1}$$

where C_t is consumption, H_t are hours of work, $\rho \in (0,1)$ is the rate of time preference, σ is the inverse of the inter-temporal elasticity of substitution in consumption, $\eta > 0$ is the Frisch elasticity of labor supply and $k > 0$ is a parameter measuring the dis-utility from labor. A restriction that must hold is $k(1-\sigma) < 1$, otherwise the marginal utility of consumption could be negative for high values of H. The restriction will always hold if $\sigma > 1$, that is what we assume. The utility function is strictly concave and is such that the conditions for the non-satiation of consumption and leisure $l = 1 - H$ are satisfied, ensuring l and C are goods.[7]

The representative household chooses consumption and labor to maximize (1), subject to the instantaneous budget constraint

$$\dot{F}_t = r_t(1-\tau_r)F_t + w_t(1-\tau_w)H_t - C_t + T_t, \tag{2}$$

where the interest rate r_t and the wage w_t are taken as given. F_t is the real value of financial assets. The government imposes a labor-income tax τ_w and a capital-income tax τ_r – both stationary – and returns part of this revenue as lump-sum transfers T_t.

Solving the maximization problem gives the optimal time path of consumption and labor:

$$\frac{\dot{C}_t}{C_t} + \left[\frac{k(1-\sigma)(1+\eta)H_t^{\frac{1}{\eta}}}{\eta\left(1-k(1-\sigma)H_t^{1+\frac{1}{\eta}}\right)}\right]\dot{H}_t = \frac{r_t(1-\tau_r)-\rho}{\sigma}, \tag{3}$$

and the transversality condition

$$\lim_{t\to\infty} \mu_t F_t e^{-\rho t} = 0, \tag{4}$$

where μ_t is the Lagrange multiplier associated with constraint (2) in the current value Hamiltonian of the representative household. See appendix. Optimization at an interior point implies that the marginal rate of substitution between consumption and labor equals their relative price, i.e.:

$$\frac{w_t(1-\tau_w)}{\sigma k C_t} = \frac{(1+\eta)H_t^{\frac{1}{\eta}}}{\eta\left(1-k(1-\sigma)H_t^{1+\frac{1}{\eta}}\right)}. \tag{5}$$

The Euler Equation in (3) shows that the capital-income tax affects negatively the consumption growth rate. Similarly, (5) shows how a higher labor-income tax, by lowering the after-tax wage, can raise leisure relatively to consumption.

[7] This specification for preferences was first proposed by Trabandt and Uhlig (2011). It is consistent with long-run growth and with a constant Frisch elasticity of labor supply. For a proof of the strict concavity in C_t and l_t of the instantaneous utility function in (1), see the appendix.

3.2 Final Good Firms

In this economy there are a perfectly competitive final good sector and an imperfectly competitive intermediate goods sector. In the final sector identical competitive firms (normalized to one for simplicity) produce the final good Y_t, taken to be the numéraire, with the following production function:

$$Y_t = \frac{L_t^{1-\beta}}{\beta} \int_{\nu=0}^{1} q_t(\nu) x_t(\nu \mid q)^{\beta} \mathrm{d}\nu, \ \beta \in (0,1), \tag{6}$$

where L_t is labor and $x_t(\nu \mid q)$ the quantity of the intermediate good in line $v \in [0,1]$, whose quality is $q_t(\nu)$. We have:

$$q_t(\nu) = \lambda^{n_t(\nu)} q_0(\nu), \tag{7}$$

where $\lambda > 1$ represents the quality-step size between successive innovations in each line and $n_t(\nu)$ is the number of innovations in line v having occurred between time 0 and time t. Only the highest grade of intermediates that is currently available in each sector will actually be produced. Profit maximization gives the demand function of the inputs, i.e. for the intermediate good v,

$$x_t(\nu \mid q) = \left(\frac{q_t(\nu)}{P_t(\nu \mid q)} \right)^{\frac{1}{1-\beta}} L_t, \tag{8}$$

and of labor,

$$w_t = \frac{(1-\beta) Y_t}{L_t}, \tag{9}$$

where $P_t(\nu \mid q)$ is the price of the intermediate good. Since the representative firm in the final output sector is competitive and subject to constant returns to scale, profits are zero in equilibrium.

3.3 Intermediate Good Firms

In the intermediate goods sector, each industry is temporarily dominated by an industry leader until the arrival of the next innovation, when its owner becomes the new industry leader. The marginal cost of production of an intermediate good, once it has been invented, is given by $\psi q_t(\nu)$ units of the final good, with $0 < \psi < 1$. Firms in the R&D sector must decide how much to invest in R&D. Successful researchers will set the price at which to sell their invented goods to final-output firms.

3.3.1 Production

The innovator maximizes profits $\pi_t(\nu \mid q) = [P_t(\nu \mid q) - \psi q_t(\nu)]x_t(\nu \mid q)$ at each point in time, which gives the optimal price:

$$P_t(\nu \mid q) = q_t(\nu), \tag{10}$$

where we have normalized $\psi = \beta$ and $q_t(\nu)$ is the unconstrained monopoly price.[8] Plugging $P_t(\nu \mid q)$ into (8), we obtain the quantity of the intermediate good,

$$x_t(\nu \mid q) = L_t. \tag{11}$$

Substituting the above into the profit function yields:

$$\pi_t(\nu \mid q) = (1-\beta)q_t(\nu)L_t, \tag{12}$$

which shows that profits received by inventors of higher quality products will be larger.

In this model technical progress consists in the expansion of the quality-ladder positions $q_t(\nu)$ in the various sectors. We define the aggregate quality index Q_t as a combination of the various quality improvements:

$$Q_t = \int_0^1 q_t(\nu)\mathrm{d}\nu. \tag{13}$$

Using equations (11) and (13) into the production function for the final good in (6), we get total output in the final good sector:

$$Y_t = \frac{1}{\beta}Q_t L_t, \tag{14}$$

which shows how increases in quality affect aggregate output. From (11) aggregate expenditure on intermediates is

$$X_t = \beta Q_t L_t. \tag{15}$$

From (14) and (15) we have a relation between intermediate-goods and final output, i.e.

$$X_t = \beta^2 Y_t. \tag{16}$$

[8] This implies we are in a so called "drastic innovation" regime in which the following restriction holds: $\lambda \geq (1/\beta)^{\beta/(1-\beta)}$. If innovations were not drastic limit prices would be chosen. This will not change the structure of the model, but could affect its calibration and change results from a quantitative point of view.

From (14) and (9) we can find an expression for the wage as a function of Q_t

$$w_t = \frac{(1-\beta)Q_t}{\beta}. \tag{17}$$

Finally, from (12) and (13) total profits of the intermediate good sector are given by

$$\Pi_t = \int_{\nu=0}^{1} \pi_t(\nu \mid q)\mathrm{d}\nu = (1-\beta)Q_t L_t. \tag{18}$$

Note that (12) as well as (18) say that a higher labor supply means a higher quantity of each intermediate good, and thus higher profits in equilibrium. This implies that there is an externality to labor in the economy: in deciding how much labor to supply, workers will not take into account this positive effect on profits. Looked at from another point of view, there is an aggregate demand externality in this economy. The demand for each intermediate and the profits obtainable from its production are larger the larger the size of the whole economy, measured by L_t, is. A tax program leading to a higher level of economic activity can therefore increase welfare by reducing the inefficiency due to monopolistic conditions.

3.3.2 R&D Activity

Households keep a diversified portfolio of all firms in the economy, so they only care about the net present value of the expected profits of a firm. Let $p_t(\nu \mid q)$ denote the arrival rate of innovation in this line of products. The value of a firm in line v, with quality q at time t, denoted by $V_t(\nu \mid q)$, then obeys the following standard arbitrage condition:

$$r_t = \frac{\dot{V}_t(\nu \mid q)}{V_t(\nu \mid q)} - p_t(\nu \mid q) + \frac{\pi_t(\nu \mid q)}{V_t(\nu \mid q)}. \tag{19}$$

This means that capital gains $\dot{V}_t(\nu \mid q)$ minus the expected capital loss $p_t(\nu \mid q)V_t(\nu \mid q)$ plus profits $\pi_t(\nu \mid q)$ must equal the market rate of return r_t. In fact when this innovation occurs the existing monopolist loses its position. This is the already mentioned "business stealing effect", which represents a second market failure in the model and makes it theoretically possible for the rate of innovation to be inefficiently high, for society as a whole.

The innovation occurs because competitive entrepreneurs invest in R&D. We assume that there is free-entry into the research market. Specifically, the arrival rate of innovation in line v with quality $q\lambda^{-1}$ is given by:

$$p_t(\nu \mid q) = \frac{\zeta z_t(\nu \mid q)}{L_t q_t(\nu)}, \ \zeta > 0 \tag{20}$$

where z_t are units of final good spent in R&D. This "lab-equipment" assumption just means that the same technology is used for producing innovations and for producing material goods.

The probability of successful innovation in a sector v increases with R&D expenses *relative* to the size of the sector, measured by value added (profits) $(1-\beta)L_t q_t(\nu)$, from (12).[9] The idea is that innovation tends to be more expensive in a large market, since the costs associated with the discovery, the development and the marketing of new technologies are higher.[10] The literature on reduced business dynamism in the US has suggested other reasons why it may be more difficult for entrants to displace large firms: large firms can better exploit data-network effects, the regulatory framework may be more lenient towards large firms, and they can self-protect through the creation of patent thickets (see Akcigit and Ates 2019).

Assuming that spending on R&D is always positive, the net expected return per unit of time on spending z_t in R&D in a line v that has quality $q\lambda^{-1}$ at time t must then be zero: $p_t(\nu \mid q)V_t(\nu \mid q) - z_t(\nu \mid q) = 0$, or using (20):

$$V_t(\nu \mid q) = L_t q_t(\nu)/\zeta\lambda. \tag{21}$$

Recalling that F_t is the market aggregate value of firms in the intermediate sector $F_t = \int_0^1 V(\nu \mid q)\mathrm{d}\nu$ and using (13), we can then write:

$$F_t = \frac{L_t Q_t}{\lambda\zeta}, \tag{22}$$

which tells us that F_t is increasing in labor and in the aggregate quality of goods.

[9] As already mentioned the specification in (20) follows Barro and Sala-i Martin (2004, chap. 7). It is also used in Chu and Cozzi (2014) and Cozzi (2017), with the difference that in these papers the R&D technology is "knowledge driven" rather than "lab equipment", so that the probability of success in a sector depends on the ratio between labor in R&D in the sector and labor in the aggregate. We recall the principle of "equivalent invention" due to Gilfillan (1935), described by Young (1998). When the profitability of finding a solution to a problem goes up, the productivity of a researcher is reduced by the possibility that other researchers find the same solution or analogous solutions at around the same time. The increased variety of technologies developed at the same time for reaching the same goal (inventing similar goods) absorbs an increased research input. The equilibrium level of R&D goes up, without being associated with an increase in the rate of product quality improvement. Empirically, the exit rate of firms (turnover) is negatively correlated with average firm size in most countries (see e.g. Bartelsman et al. 2005).

[10] See e.g. Sequeira et al. (2018, p. 128) who refer to the "costs pertaining to the construction of prototypes and samples, new assembly lines and training of workers, and generic coordination, organizational, marketing, and transportation costs" that tend to be higher in large markets. In particular, in Sequeira et al. (2018) the rate of growth of knowledge depends positively on R&D labor and negatively on aggregate labor raised to an index of complexity. This index is increasing in the level of knowledge, so that asymptotically the model is no-scale.

3.4 Government

To close the model, we assume that government expenditure G equals a fixed fraction g of gross output, another fraction of which is transferred back to families:

$$G_t = gY_t, \tag{23}$$

and

$$T_t = \tau Y_t. \tag{24}$$

We rule out a market for government bonds and assume that the government runs a balanced budget. Part of the revenue from income taxes is transferred back to households. Our assumption of given g and τ is made mainly for convenience.[11] However, the public expenditure components that might be seen as exogenous in actual economies (from public wages, the payments of interest on public debt, etc.) are far from zero and have remained fairly stable, as a percentage of output, over the last decades.[12]

$$G_t = w_t \tau_w L_t + r_t \tau_r F_t - T_t, \tag{25}$$

with the restriction $0 < g < 1-\beta^2$ in place so to rule out the possibility for the government to confiscate all total value-added in the economy.[13]

3.5 Market equilibrium

In equilibrium, the following market clearing conditions for final goods, where investment is represented by $Z_t = \int_0^1 z_t(\nu \mid q)\mathrm{d}\nu$, and labor are satisfied:

$$Y_t - X_t = C_t + Z_t + G_t, \tag{26}$$

$$H_t = L_t. \tag{27}$$

We want to obtain a relationship between the two tax schedules in equilibrium. Consider the government's budget constraint in (25) and, using condition (22) and the definitions

[11]Indeed, in a unbounded growth model to make government grow in the long run at the same rate as the rest of the economy is indeed the only possibility: growth at a slower rate would make its role negligible asymptotically, while growth at a higher rate would make it violate the resource constraint of the economy. Similar considerations apply to transfers.

[12]To streamline our analysis we have assumed that government consumption does not directly affect the choices of agents. This does not per se imply that G is pure waste. For instance the production function in the final sector could be Leontief: $Y = \min\left[G/g, \frac{L^{1-\beta}}{\beta}\int_{\nu=0}^{1} q(\nu)x(\nu \mid q)^{\beta}\mathrm{d}\nu\right]$. Of course this would be an extreme assumption as regards the role of public services in economic activity.

[13]Consider all total value-added; this is given, using (15), by: $Y - X = Y(1-\beta^2)$. Then, it is straightforward that imposing $0 < G < Y(1-\beta^2)$ gives us the restriction $0 < g < 1-\beta^2$.

for G_t and T_t, we can write:

$$\tau_w = \frac{(g+\tau)Y_t - r_t \tau_r F_t}{w_t L_t}. \tag{28}$$

Next, use (9) for $w_t L_t$, and combine equations (22) and (14), so we can use $F_t = \beta Y_t / \lambda\zeta$. We end up with the following:

$$\tau_w = \frac{\tau+g}{1-\beta} - \frac{\beta r_t}{\lambda\zeta(1-\beta)}\tau_r. \tag{29}$$

From this equation we can see that if policy variables do no change over time r_t will not change either.

In a competitive equilibrium, individual and aggregate variables are the same, and prices and quantities are consistent with the efficiency conditions (2), (3), (4), and (5) for households; the profit-maximization conditions for firms in both the final goods sector, (8) and (9), and in the intermediate goods sector, (10), (15) and (17); the government budget constraint (25); the market clearing conditions for wealth (22), for final goods (26) and for labor (27).

In appendix we prove that the economy will always be in a balanced growth equilibrium, which is unique. This is interesting because it is known that, even in models like ours with only one state variable, indeterminacy may arise due to preferences non separable over labor and consumption and/or destabilizing fiscal policy.[14]

$$L = \left(\frac{\frac{\frac{1-\beta}{\beta}(1-\tau_w)\lambda\zeta\eta}{\sigma k(1+\eta)}}{\frac{1-\beta}{\beta}(1-\tau_w)\lambda\zeta\frac{\sigma+\eta}{\sigma(1+\eta)} - r(1-\tau_r)\left(\frac{1}{\sigma}-1\right) + \frac{\rho}{\sigma}\tau\frac{1}{\beta}\lambda\zeta} \right)^{\frac{\eta}{1+\eta}}. \tag{30}$$

3.6 Balanced Growth Path

From (21) we know that along a BGP, when L is given, for given q, the value of a firm in a given line v will be fixed, so from (19) we arrive to

$$V(\nu \mid q) = \frac{\pi(\nu \mid q)}{r + p(\nu \mid q)}. \tag{31}$$

Substituting in this equation the value for $V(\nu \mid q)$ from (21) and for $\pi(\nu \mid q)$ from (12) we obtain:

$$r + p(\nu \mid q) = \lambda\zeta(1-\beta). \tag{32}$$

14 Pelloni and Waldmann (2000), Mino (2001), Palivos et al. (2003), Park (2009), Chen and Lee (2007), and Wong and Yip (2010), among others, have shown that taxing and government spending as well as the non separability of preferences over consumption and leisure (which we assume) may cause indeterminacy in endogenous growth models.

This shows that the arrival rate of innovation will be the same along all lines. It follows from (20) that:

$$Z = \int_0^1 z(\nu \mid q)\mathrm{d}\nu = \frac{p}{\zeta}QL. \tag{33}$$

In a BGP the supply of labor will be constant, so (5) implies that consumption and wages will grow at the same rate, but since the wage rate is proportional to the aggregate quality index by (9), then w, C and Q all grow at the same rate, which we define γ:

$$\frac{\dot{C}}{C} = \frac{\dot{Q}}{Q} = \gamma. \tag{34}$$

Using (3) and then (32) we can write:

$$\gamma = \frac{r(1-\tau_r)-\rho}{\sigma} = \frac{\left[\lambda\zeta(1-\beta)-p\right](1-\tau_r)-\rho}{\sigma}. \tag{35}$$

From this formula we can already see that the growth rate is positively related to the after-tax interest rate on capital and negatively to ρ and σ. However, this formula is not sufficient for the final solution, because p is determined endogenously: to pin it down we need to find the growth rate for Q. Using the Law of Large Numbers, in an interval of time $\triangle t$, $p \cdot \triangle t$ sectors will be innovating once, rising the quality of their goods by λ, while the rest will not. Going to the limit as $\triangle t$ goes to 0 brings us to

$$\gamma = \frac{\dot{Q}}{Q} = p(\lambda-1). \tag{36}$$

From (32), (35), and (36) we then get:

$$r = \frac{\rho + \sigma\lambda\zeta(1-\beta)(\lambda-1)}{1-\tau_r+\sigma(\lambda-1)}, \tag{37}$$

and

$$\gamma = \frac{(\lambda-1)\left[\lambda\zeta(1-\beta)-\rho/(1-\tau_r)\right]}{1+\sigma(\lambda-1)/(1-\tau_r)}, \tag{38}$$

as well as

$$p = \lambda\zeta(1-\beta) - \frac{\rho + \sigma\lambda\zeta(1-\beta)(\lambda-1)}{1-\tau_r+\sigma(\lambda-1)} \tag{39}$$

Just looking at (38) we can see that the growth rate is a decreasing function of ρ, the rate of time discount, and of σ, the inverse of the intertemporal elasticity of substitution, while obviously increasing in the productivity of research, ζ and in the size of the quality step λ. τ_r can be seen to have a negative effect on growth. Note that the growth rate does not depend on L. That is, the model, as anticipated, does not feature "strong" scale effects. In order for the free-entry condition to hold with equality, γ must be positive and, in order to satisfy the transversality condition, we must impose $r(1-\tau_r) > \gamma$.[15]

Finally, by using the BGP expression for r into (29), we can compute the relationship between the two tax schedules,

$$\tau_w = \frac{\tau + g}{1-\beta} - \frac{\beta\left[\rho + \sigma\lambda\zeta(1-\beta)(\lambda-1)\right]}{\lambda\zeta(1-\beta)\left[1-\tau_r+\sigma(\lambda-1)\right]}\tau_r. \tag{40}$$

Again it is easy to see that there is an inverse relationship between the two tax rates.

4 Optimal Tax Analysis

As is typical in optimal tax theory we posit that the tax system should maximize a social welfare function (here naturally identified with the utility of the representative dynasty), taking into account how agents react to taxes, i.e. considering the general equilibrium conditions of the economy as constraints of the maximization problem.

Given the growth rate γ and the labor supply L found above, it is possible to calculate maximum utility W. We want to express the welfare function in terms of the tax rate on capital in order to find its maximum value. Utility in (1) can be written as

$$W = \int_{t=0}^{\infty} e^{[\gamma(1-\sigma)-\rho]t} \frac{C_0^{1-\sigma}\left(1-k(1-\sigma)L^{1+\frac{1}{\eta}}\right)^{\sigma} - 1}{1-\sigma}\mathrm{d}t, \tag{41}$$

where C_0 is consumption at time 0.

Solving the integral in (41), we obtain

$$W = \frac{1}{1-\sigma}\left[\frac{C_0^{1-\sigma}\left(1-k(1-\sigma)L^{1+\frac{1}{\eta}}\right)^{\sigma}}{\rho-\gamma(1-\sigma)}\right]. \tag{42}$$

From (5) we have that:

$$C_0 = \frac{\eta Q_0(1-\beta)(1-\tau_w)\left(1-k(1-\sigma)L^{1+\frac{1}{\eta}}\right)}{\beta\sigma k(1+\eta)L^{\frac{1}{\eta}}} \tag{43}$$

where we have used (9) to eliminate w. C_0 can be expressed in terms of τ_r by using (30) to eliminate L and by using (40) to eliminate τ_w. Finally, γ can be expressed in terms of τ_r by using (38). With these substitutions W can be written as a differentiable function of one variable τ_r. By calculating the first derivative of the function and equating it to zero we would get an equation in τ_r, which could be solved to give us the optimal taxes. It is much quicker to calculate the optimal τ_r numerically, by a simple search algorithm.

[15] The condition for $\gamma > 0$ in (38) is $\rho < \lambda\zeta(1-\beta)(1-\tau_r)$.

In our quantitative exercise, we will show that for specifications of tastes and technology parameters often used in calibration exercises, it is possible by raising the tax on capital above zero to reduce growth and yet induce welfare improvements. This can happen even if the before-tax market equilibrium growth rate is inefficiently low, as will be shown in the next section. Such result goes against the widely held belief that, by compounding, growth effects always prevail over level effects.

5 Social Planner's Solution

In this section, we will analyze the social planner's problem. Our main aim is to understand if the market growth rate is too high or too low from a social point of view. As is well known in Schumpeterian models, growth can be too high due to the business stealing effect or too low due the appropriability effect. In our model there is a third externality generated by the link between growth and public expenditures. Variables keep the same meaning as in the market economy and the index s characterizes variables in the socially planned economy. The social planner seeks to maximize the representative household's utility subject to the economy's resource constraint, $Y_s = C_s + X_s + Z_s + G_s$ and to the R&D technology,

$$\dot{Q}_s = \frac{\zeta(\lambda - 1)Z_s}{L_s}. \tag{44}$$

Given our assumptions on the production function of new ideas, Z_s, the aggregate amount of R&D effort, leads to an increase in quality $\lambda - 1$ at the flow rate of ζ in the centralized economy, adjusted for the size of employment L_s. In the centralized economy, the following condition for each intermediate holds:

$$x_s(\nu, t \mid q) = \beta^{\frac{1}{\beta-1}} L_s, \tag{45}$$

which is the social planner equivalent of the decentralized quantity in (8). Aggregating, we have

$$X_s = \beta^{\frac{\beta}{\beta-1}} Q_s L_s. \tag{46}$$

Monopoly pricing implies that the privately chosen quantity (equation 15) is smaller than the socially chosen amount, since $\beta^{\frac{\beta}{\beta-1}} > \beta$. Final output in equilibrium can be computed from (6) and (46) and is expressed as:

$$Y_s = \beta^{\frac{1}{\beta-1}} Q_s L_s. \tag{47}$$

For given Q_s and labor, the level of output in the decentralized economy is lower than the optimal social value (recall equation 14), since $\beta^{\frac{1}{\beta-1}} > 1/\beta$. Equation (47) implies that the social planner's growth rate of Y_s equals the growth rate of Q_s. The social planner decides on the optimal paths of the control variables C_s and L_s and of the state variable Q_s, given the constraint:

$$\dot{Q}_s = \frac{\zeta(\lambda-1)}{L_s}\left(\beta^{\frac{1}{\beta-1}}Q_sL_s - \beta^{\frac{\beta}{\beta-1}}Q_sL_s - g\beta^{\frac{1}{\beta-1}}Q_sL_s - C_s\right). \tag{48}$$

Besides the above constraint, necessary conditions for a maximum are:

$$(i)\ C_s^{-\sigma}\left(1-k(1-\sigma)L_s^{1+\frac{1}{\eta}}\right)^{\sigma} = \frac{\mu_s\zeta(\lambda-1)}{L_s};$$

$$(ii)\ C_s^{-\sigma}\left(1-k(1-\sigma)L_s^{1+\frac{1}{\eta}}\right)^{\sigma-1} k\left(\sigma-1\right)\left(1+\frac{1}{\eta}\right)L_s^{\frac{1}{\eta}} = \frac{\mu_s\zeta(\lambda-1)}{L_s^2};$$

$$(iii)\ \mu_s\beta^{\frac{1}{\beta-1}}\zeta(\lambda-1)(1-\beta-g) = \rho\mu_s - \dot{\mu}_s;$$

$$(iv)\ \lim_{t\to\infty} e^{-\rho t}\mu_s Q_s = 0,$$

where the Lagrange multiplier μ_s applies to the constraint of the underlying current-value Hamiltonian for the social planner's problem. Combining (i) and (ii) one obtains the expression for optimal employment:

$$L_s = \left(\frac{\eta}{k\left(\sigma-1\right)}\right)^{\frac{\eta}{\eta+1}}. \tag{49}$$

L_s is then constant through time. Also from (iii) one see that μ_s grows at a constant rate. By combining (i) and (ii), the growth rate of consumption in the centralized economy is easily seen to be:

$$\frac{\dot{C}_s}{C_s} = \gamma_s = \frac{\zeta\beta^{\frac{1}{\beta-1}}(\lambda-1)(1-\beta-g)-\rho}{\sigma}. \tag{50}$$

Recall the market growth rate in (38). We want to compare γ and γ_s, but is not easy to understand which one is larger, because both are a function of many other parameters. In particular, we find that the difference depends on the values of λ and β. Given our choices of parameters, we will show that the socially optimal growth rate is definitely higher than the market one, when τ_r is set to zero. In other words, the positive externality to growth quantitatively prevails over the negative ones. So our result of positive capital income taxation is not driven by the fact that taxing capital is an obvious way to lower an inefficiently high rate of growth.

6 Quantitative Analysis

In this section, we will parametrize the model in order to examine quantitatively the effects of increasing the capital-income tax-rate on labor supply, economic growth and social welfare.

Obviously this model is not rich enough in number of variables to match the data for advanced economies well. So our quantitative analysis aims at clarifying possible mechanisms of action of policy not considered before in the literature rather than at finding numerical results.

To calibrate the model we set values for the 10-tuple $\{\beta,\ \sigma,\ \rho,\ \eta,\ L,\ p,\ \gamma,\ \tau,\ g,\ \tau_r\}$, and the implied ones by the 5-tuple $\{\lambda,\ k,\ \zeta,\ \tau_w,\ r\}$. The benchmark calibration is shown in Table 1.

6.1 Parametrization

In our model, per capita GDP growth and TFP growth coincide. Using Eurostat and AMECO data for the period 1995-2018, we therefore choose an intermediate value between the two for the European Union (EU 28) countries, $\gamma = 0.019$. The calibration of the parameters of the utility function follows from Trabandt and Uhlig (2011): the Frisch elasticity is assumed to be constant, such that $\eta = 1$; the inter-temporal elasticity of substitution of consumption is set to 0.5, corresponding to $\sigma = 2$; the parameter measuring the dis-utility from labor, k, is left free to vary, in order to keep the level of labor around its steady-state value. According to the Labour Force Survey of Eurostat the average EU 28 worker used about 21% of her (his) time endowment to work. We set the level of labor supply to L accordingly. The time discount rate is set to $\rho = 0.03$, in line with the literature, where we find a value that varies between 1% and 5%.

The implied value for the steady-state interest rate r is then 0.0680, close to the average real return on the stock market over the last century (7%). The production parameter is set to $\beta = 0.65$, such that the share of labor ($s_L = 1/(1+\beta)$) is calibrated to be 60%, consistently with most literature. Empirical estimates for the monopoly price range between 1.05 and 1.5, so the mark-up in this model is calibrated to be $\mu = 1/\beta = 1.5$. The quality-step λ must be not lower than $(1/\beta)^{\beta/(1-\beta)}$, so as to ensure that an innovator will set the unconstrained monopoly price. We set the value of the success rate p at 0.014, so that given γ and from (36), the implied value for the quality-step λ results larger than $(1/\beta)^{\beta/(1-\beta)}$.

Using Eurostat data for EU 28, the benchmark values of the ratio of government expenditure to GDP, $G/(Y-X)$, is set at 21 percent and of the ratio of government transfers to GDP, $T/(Y-X)$, is calibrated to be 20 percent. Notice that GDP is equal to $Y - X = Y(1-\beta^2)$, where $\beta^2 = 42\%$ represents the ratio of intermediates to final output. Then, using the fact that $G = gY$ and $T = \tau Y$, the values for g and τ immediately follows. Finally, the marginal capital-income tax rate is fixed at a starting value of $\tau_r = 0$, so that we can isolate the effects of introducing the tax on capital in the model. In the baseline calibration the socially optimal growth rate γ_s is 0.037, that is two times larger than the market one.

We calculate the values for the remaining parameters as follows. The R&D parameter ζ, given r, p, λ and β, is obtained from equation (32) and is equal to 0.0994 in our model. The benchmark value for the labor-income tax rate corresponding to $\tau_r = 0$ is computed from (29) and its starting value is found to be $\tau_w = 0.6765$.

6.2 Results

We now calculate the optimal tax structure. To do so, we simply need to maximize the expression for W in (42) with respect to τ_r. In Table 2, we report the optimal τ_r (and τ_w), as well as the optimal values for growth and labor supply in the baseline model. Comparing the optimal values in the second column with the ones in steady-state, which correspond to a zero tax on capital scenario, we are able to evaluate the impact of the tax program. We immediately see that the capital-income tax rate associated with maximum utility is positive and quite sizable.

We now recapitulate the chain of effects put in motion by shifting the tax burden from labor to capital. First of all, labor supply will increase, because of a positive substitution effect, not offset by a negative income effect. The increased labor supply will push up the demand for the intermediates. This reduces the monopoly power distortion and is an efficiency gain. However, the after-tax interest rate will decrease, so the rate of growth will decrease, as indeed we know from (38). This decrease in growth would reduce the business stealing and the weight of government negative externalities, so could be a further gain in efficiency. However, we know from our analysis of the social planner solution that the market equilibrium, even without a tax on capital (always growth decreasing in this model), generates an inefficiently low growth rate. This shows that growth was inefficiently high to begin with cannot be the key to our result that a positive effect on welfare is consistent with a negative effect on growth. This is interesting, because there exists a general shared view that growth effects tend to prevail over level effects, by compounding and that, therefore, when growth is too low it cannot be efficient to reduce it further. However, what we find contradicts this consensus view.

Finally, we consider welfare gain from adopting the optimal tax mix. The welfare gain is defined as the fraction of additional consumption that individuals would need, in perpetuity, under the alternative policy in order to be indifferent between the two policies. Under our baseline calibration the welfare gain is equivalent to a permanent increase in consumption of 3.93 percent per year.

6.3 Sensitivity Analysis

To clarify the role of the various parameters in determining the value of the optimal capital income tax rate, we perform some sensitivity analysis. We calculate how the optimal tax on capital-income changes when changing the parameters β, σ, ρ and η, leaving all the other unchanged. Of course this means the implied initial steady-state values for τ_w, γ and L will change accordingly. We also consider different sizes of the public sector, by changing transfers and public spending. Results are reported in Table 3, where the optimal values of the baseline model are reported in the first row, while values in parentheses refer to the steady state values of τ_w, γ and L at zero capital income tax rate. To check the overall robustness of our results we also undertake the same sensitivity exercise by varying each parameter or variable at a time along with the two implied parameters k and ζ, so as to keep the growth rate of the economy γ and employment L at their steady state values.

We will now attempt to interpret intuitively the effects summarized in Tables 3 and 4, and represented in Figures 1-5.

6.3.1 Effect of β

Let's start by looking at the role of β. Agents when choosing how much to work, take all aggregate variables as given. The increase in the demand for intermediates through higher employment has a first order positive effect on profits, which, due to the monopoly power distortion, are increasing in the size of the market. This spillover from labor to profits is increasing in β because the ratio between labor income and profits is $1/\beta$. See also Figure 1, where we show how the optimal tax mix changes with the labor income share $1/(1+\beta)$. Reducing the tax on labor pushes labor and the related positive spillover up. This positive spillover will be higher the higher is β. These mechanisms explain why for higher β the optimal tax mix delivers a relatively higher tax on capital income.

6.3.2 Effect of σ

A higher σ calls for a higher capital tax rate. See Figure 2. Shifting the tax incidence from labor to capital increases labor and current consumption. At the same time it lowers growth – the less the higher is σ – and thus worsens the dynamic inefficiency in the model, due to the appropriability failure. However, a higher elasticity of the marginal utility of consumption means consumers care relatively more about the current increase in consumption, (which is lower than future consumption in a growing economy) than about the decrease in future consumption (which is higher). So, when the current consumption is increased along with employment, this increment is given more weight than the future loss.

6.3.3 Effect of ρ

The optimal τ_r is increasing in the rate of time preference ρ, as shown in Figure 3. It is not difficult to understand intuitively why. A lower τ_w means higher labor and more utility thanks to higher consumption now and less in the future, as growth will be slower because of the higher τ_r. However, the future is discounted more heavily with higher discount rate ρ.

6.3.4 Effect of η

A higher Frisch elasticity induces a higher optimal capital tax rate. See Figure 4. This is because the advantage of pushing up the tax on capital and down the tax on labor is bigger the bigger the increase in labor the tax program is able to induce: a bigger increase in labor will mean a larger increase in the demand size for intermediates and therefore a larger reduction in the monopoly power distortion.

6.3.5 Effects of T and G

Another interesting result is that a higher ratio of public outlays, whether due to higher transfers or higher public consumption, is associated with a higher optimal $\tau_r\ /\tau_w$, as Figure 5 shows. The figure shows the effect of higher public expenditure, but the effect of higher transfers is very similar.

The intuitive explanation for this effect is that as the fiscal revenue to be raised increases, increasing τ_r becomes less and less distortionary compared to increasing τ_w. In fact the rate of growth is a linear function of the tax rate on capital, while labor declines at an increasing rate when τ_w increases.

6.4 Robustness

In this section we conduct one further experiment introducing a "dilution effect". We allow for the variety of goods to increase with employment, while the effectiveness of the research effort on each variety decreases when it is spread more thinly over a larger number of varieties as in Smulders and Van de Klundert (1995), Peretto (1998), Young (1998), Dinopoulos and Thompson (1999) and Howitt (1999).

In the final sector the production function (6) is replaced by:

$$Y_t = \frac{L_t^{1-\beta}}{\beta} \int_{\nu=0}^{N} q_t(\nu) x_t(\nu \mid q)^{\beta} \mathrm{d}\nu,\ \beta \in (0,1), \tag{51}$$

so that the variety of products is not a given but can change over time. The arrival rate of innovation in line v with quality $q\lambda^{-1}$ is given by:

$$p_t(\nu \mid q) = \frac{\zeta z_t(\nu \mid q)}{N_t q_t(\nu)}, \zeta > 0. \tag{52}$$

The probability of successful innovation in a sector v increases with R&D expenses and decreases with quality q_t and with the mass of firms/products N due to a dilution effect. The creation of new varieties of goods is an externality to labor, due to "learning by doing" or rather "inventing by doing". This follows Dinopoulos and Thompson (1999).[16] We also assume that N is equal to L at all times. In appendix we show that there is a positive impact of labor on initial consumption and therefore on the whole path of consumption. The optimal tax on labor income is now approximately 34 percent while the tax on capital income is 53 percent, as reported in Table 5. It is not difficult to understand intuitively why the tax on capital is considerably higher than in the base model. The positive impact of more labor on the level of consumption is due to the fact that a larger N increases the productivity of labor, *coeteris paribus*. In fact, when labor goes up, aggregate income goes up, not only because more labor is used, but also because it is more productive thanks to the simultaneous increase in N. On the other hand, the negative dilution effect on research productivity, due to a higher L, is the same as in the base model, the only difference being that it works through N, rather than being direct.

7 Conclusions

The objective of the paper has been to study how the tax burden should be distributed between factor incomes in a scale-less Schumpeterian model of endogenous growth. We focused on the effects of labor and capital-income taxation on labor supply and growth. In the model there are four inefficiencies: market power of firms, implying monopoly pricing as a mark-up over marginal cost; an appropriability problem, implying that technological advances generate a social surplus higher than the cost of their discovery; a business-stealing effect, implying that part of the reward from successful R&D has no social value; finally, GDP has a cost in term of public expenditures not taken into account by private agents. A tax program combining a tax rate on capital income with one on labor income will affect these four inefficiencies.

For standard parameters' values, the results imply that the optimal tax rate on capital will in general be positive. Starting from a zero capital tax-rate, the optimal value associated with maximum utility is calculated to be around 24 percent in our baseline calibration. It is interesting that in this model, a positive capital-income tax can be optimal, even if growth would be higher when capital income is not taxed, and even if growth is suboptimal with a zero capital tax.

We have done some sensitivity analysis to understand the role of the various parameters. We find that the optimal tax on capital-income is increasing in the rate of time discount, the Frisch elasticity and the size of the public sector, and decreasing in the inter-temporal elasticity of substitution of consumption and the income share of labor.

[16] The creation of new varieties is costless, with the quality level associated with a variety j created at time t equal to the average quality at time t. The probability of inventing a new variety is linear in L and a fraction m of varieties disappear per unit of time, so that $dN/dt = hL - mN$. It is then for simplicity assumed that in steady state: $N = \frac{h}{m}L$ and h/m is normalized to one without any loss of generality.

When the model is extended, so that employment leads to an increase in the variety of available goods, the tax on capital will be much higher.

Papers on taxation and growth often report the effects of the tax experiments only on growth, and not also on welfare. When market equilibrium growth rate is lower than optimal, there is an implicit presumption that, through compounding, growth effects always prevail over level effects. In the model that we present, growth is inefficiently low in the absence of the tax on capital, however, when the introduction of the tax lowers growth, there is still a positive welfare effect.

In future research, we plan to extend the model by allowing for heterogenous agents, with different propensities to save and human capital endowments, and for aggregate uncertainty and study the trade-offs for policy analyzed in this paper in more general settings.

Appendix

Utility Function

In what follows we show why the restriction $1 + k(\sigma - 1) > 0$, not just $k > 0$, as asserted in Trabandt and Uhlig (2011), is necessary for the following function u:

$$u = \frac{C^{1-\sigma}\left(1 + k(\sigma-1)(1-l)^{1+\frac{1}{\eta}}\right)^{\sigma}}{1-\sigma}, \tag{A-1}$$

to be a well defined instantaneous utility function. Note that we have dropped the time index for simplicity.

Given (A-1) we have that iff $1 + k(\sigma-1)(1-l)^{1+\frac{1}{\eta}} > 0$ (always true for $\sigma > 1$), then the Hessian is negative definite. To simplify notation let $\mathbf{\Omega} \equiv 1 + k(\sigma-1)(1-l)^{1+\frac{1}{\eta}}$, so that under the assumption that $\mathbf{\Omega} > 0$ we have

$$\begin{aligned}
&u_c = C^{-\sigma}\mathbf{\Omega}^{\sigma} > 0,\\
&u_{cc} = -\sigma C^{-\sigma-1}\mathbf{\Omega}^{\sigma} < 0,\\
&u_l = \left(1+\tfrac{1}{\eta}\right)\sigma k C^{1-\sigma}\mathbf{\Omega}^{\sigma-1}(1-l)^{\frac{1}{\eta}} > 0,\\
&u_{ll} = -\left(1+\tfrac{1}{\eta}\right)\sigma k C^{1-\sigma}\mathbf{\Omega}^{\sigma-2}(1-l)^{\frac{1}{\eta}}\left[\left(1+\tfrac{1}{\eta}\right)(\sigma-1)^2 k(1-l)^{\frac{1}{\eta}} + \mathbf{\Omega}\tfrac{1}{\eta}(1-l)^{-1}\right] < 0,\\
&u_{lc} = (1-\sigma)\left(1+\tfrac{1}{\eta}\right)\sigma k C^{-\sigma}\mathbf{\Omega}^{\sigma-1}(1-l)^{\frac{1}{\eta}} < 0,\\
&u_{ll}u_{cc} - u_{lc}^2 = \left(1+\tfrac{1}{\eta}\right)\sigma^2 k C^{-2\sigma}\mathbf{\Omega}^{2\sigma-1}(1-l)^{\frac{1}{\eta}-1}\tfrac{1}{\eta} > 0.
\end{aligned}$$

This proves that iff $1 + k(\sigma-1)(1-l)^{1+\frac{1}{\eta}} > 0$, the Hessian is negative definite.

Households Maximization Problem and Frisch Elasticity

The current-value Hamiltonian $\mathcal{H}$ and the first-order-conditions for the consumer are:

$$\mathcal{H}_t = \left(\frac{C_t^{1-\sigma}\left(1-k(1-\sigma)H_t^{1+\frac{1}{\eta}}\right)^{\sigma}-1}{1-\sigma} \right) + \lambda_t \;\left[r(1-\tau_r)F + w(1-\tau_w)H - C + tY \right];$$

$(i)\ \frac{\partial \mathcal{H}}{\partial C} = 0 \Rightarrow C^{-\sigma}\left(1-k(1-\sigma)H^{1+\frac{1}{\eta}}\right)^{\sigma} = \lambda;$

$(ii)\ \frac{\partial \mathcal{H}}{\partial H} = 0 \Rightarrow k\sigma\frac{(1+\eta)}{\eta}C^{1-\sigma}\left(1-k(1-\sigma)H^{1+\frac{1}{\eta}}\right)^{\sigma-1} H^{\frac{1}{\eta}} = \lambda w(1-\tau_w);$

$(iii)\ \frac{\partial \mathcal{H}}{\partial F} = \rho\lambda_t - \dot{\lambda}_t \Rightarrow \lambda_t r(1-\tau_r) = \rho\lambda_t - \dot{\lambda}_t;$

$(iv)\ \lim_{t\to\infty} e^{-\rho t}\lambda F = 0.$

From (i), obtain

$$C^{1-\sigma} = \left[\left(1-k(1-\sigma)H^{1+\frac{1}{\eta}}\right)\lambda^{-1/\sigma}\right]^{1-\sigma}, \tag{A-2}$$

and use it into (ii),

$$k\sigma\frac{(1+\eta)}{\eta}H^{\frac{1}{\eta}} = w(1-\tau_w)\lambda^{1/\sigma}, \tag{A-3}$$

from which we can compute the Frisch elasticity:

$$\epsilon_F = \frac{dw}{dH}\frac{w}{H} = \frac{\lambda^{1/\sigma}(1-\tau_w)}{\frac{1}{\eta}\frac{1+\eta}{\eta}k\sigma H^{\frac{1}{\eta}-1}}\frac{w}{H} = \eta. \tag{A-4}$$

Dynamic Properties of the Equilibrium

In this section we show that the model has a unique steady state solution with no transitional dynamics.

Using (22) to express F_t in terms of L_t and Q_t, (2) can be rewritten as:

$$\frac{\dot{L}_t}{L_t} + \frac{\dot{Q}_t}{Q_t} = r(1-\tau_r) + w_t(1-\tau_w)\frac{1}{Q_t}\lambda\zeta - \frac{C_t}{L_tQ_t}\lambda\zeta + \frac{T_t}{L_tQ_t}\lambda\zeta, \tag{A-5}$$

where we have dropped the time index from the interest rate, because from (29) we know it to be constant through time. Using (5), (17), (24) and (14) and simplifying we obtain:

$$\frac{\dot{L}_t}{L_t} + \frac{\dot{Q}_t}{Q_t} = r(1-\tau_r) + \frac{(1-\beta)}{\beta}(1-\tau_w)\lambda\zeta\frac{(\sigma+\eta)\,kL_t^{\frac{1}{\eta}+1} - \eta}{\sigma k(1+\eta)L_t^{\frac{1}{\eta}+1}} + \tau\frac{1}{\beta}\lambda\zeta. \tag{A-6}$$

Differentiating (5) with respect to time we get:

$$\frac{\dot{C}_t}{C_t} = \frac{\dot{Q}_t}{Q_t} - \left[\frac{1+\eta k(1-\sigma)L_t^{\frac{1+\eta}{\eta}})}{1-k(1-\sigma)L_t^{\frac{1+\eta}{\eta}}}\right]\frac{\dot{L}_t}{\eta L_t}. \tag{A-7}$$

Combining (A-7) and (3) we eliminate $\frac{\dot{C}_t}{C_t}$ and obtain:

$$\frac{\dot{Q}_t}{Q_t}-\left[\frac{1+\eta k(1-\sigma)L_t^{\frac{1+\eta}{\eta}}}{1-k(1-\sigma)L_t^{\frac{1+\eta}{\eta}}}\right]\frac{\dot{L}_t}{\eta L_t}+\left[\frac{k(1-\sigma)(1+\eta)L_t^{\frac{1}{\eta}}}{\eta\left(1-k(1-\sigma)L_t^{1+\frac{1}{\eta}}\right)}\right]\dot{L}_t=\frac{r(1-\tau_r)-\rho}{\sigma}.$$

Combining the above equation with (A-6), we eliminate $\frac{\dot{Q}_t}{Q_t}$ and, after rearranging, get the following differential equation in L_t :

$$\dot{L}_t=\frac{\mathbf{A}(L_t)}{\mathbf{B}(L_t)}, \tag{A-8}$$

where $\mathbf{A}(L_t)\equiv\eta\mathbf{C}(L_t)\mathbf{D}(L_t)L_t$,with $\mathbf{C}(L_t)\equiv a+bL_t^{-\frac{1+\eta}{\eta}}$, $a\equiv-\frac{1-\beta}{\beta}(1-\tau_w)\lambda\zeta\frac{(\sigma+\eta)}{\sigma(1+\eta)}+r(1-\tau_r)\left(\frac{1}{\sigma}-1\right)-\frac{\rho}{\sigma}-\tau\frac{1}{\beta}\lambda\zeta$ and $b\equiv\frac{\frac{(1-\beta)}{\beta}(1-\tau_w)\lambda\zeta\eta}{\sigma k(1+\eta)}$, $\mathbf{D}(L_t)\equiv 1-k(1-\sigma)L^{\frac{1+\eta}{\eta}}>0$, $\mathbf{B}(L_t)\equiv(1+\eta)\left[k(1-\sigma)(1+\eta)L_t^{\frac{1}{\eta}+1}-1\right]<0$, since we set $\sigma>1$.

It is easy to prove the existence and uniqueness of a steady state solution. We indicate state variables by dropping time indices. Along a BGP $\mathbf{A}(L)=0$. Since $L>0$ and $\mathbf{D}>0$ this implies $\mathbf{C}=0$ hence:

$$L=\left(-\frac{b}{a}\right)^{\frac{\eta}{1+\eta}}. \tag{A-9}$$

which is clearly uniquely pinned down by the parameters and corresponds to (30) in the text. Of course the model will only be well specified if these parameters make L belong to the interval $(0,1)$. To prove determinacy we need to show that $d\dot{L}_t/dL_t>0$ around the BGP equilibrium.

Since along a BGP $\mathbf{A}(L)=0$, from (A-8) we write:

$$d\dot{L}_t/dL_t|_{L_t=L}=\frac{\mathbf{A}'(L)}{\mathbf{B}(L)}. \tag{A-10}$$

For determinacy we need to prove that this derivative is strictly positive. We have $\mathbf{A}'(L)=\mathbf{C}'(L)\mathbf{D}(L)L$, since in steady state $\mathbf{C}(L)=0$.As both $\mathbf{D}(L)>0$ and $L>0$, $\mathbf{A}'(L)$ signs as $\mathbf{C}'(L)$. Since $\mathbf{C}'(L)=-\frac{1+\eta}{\eta}bL^{-\frac{1+2\eta}{\eta}}$ is negative, $\mathbf{A}'(L)$ is negative. Since $\mathbf{B}$ is negative, from (A-9), we have obtained our proof.

Robustness: The economy with "Dilution Effect"

The aggregate quality index Q_t is now given by:

$$Q_t=\int_0^N q_t(\nu)\mathrm{d}\nu, \tag{A-11}$$

so that equation (A-11) replaces equation (13). $N = L$ implies that in equilibrium equation (20)and equation (52) are equivalent and that, as a consequence (21) is still valid.

Recalling that F_t is the market aggregate value of firms in the intermediate sector and using (A-11), we can then write $F_t = \int_0^N V(\nu \mid q)\mathrm{d}\nu$ so that (22) is still valid as well.

The analysis of the dynamics and of the steady state of the model in section 3 is conducted using equations which are still valid when $N = L$ so all the results are still valid. However, the optimal tax policy is now different. In fact while equation (43) still holds, now $Q_0 = \int_0^N q_0(\nu)\mathrm{d}\nu$ is increasing in $L = N$. Assuming symmetry with $q_0(\nu) = q_0$, we have $Q_0 = Nq_0 = Lq_0$, so that initial consumption is increasing in L.

References

Aghion, P., Akcigit, U., and Fernández-Villaverde, J. (2013). Optimal capital versus labor taxation with innovation-led growth. *National Bureau of Economic Research*, Working Paper No. 19086.

Akcigit, U. and Ates, S. T. (2019). Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory. *National Bureau of Economic Research, Inc*, Working Paper No. 25755.

Amano, D., Itaya, J.-i., and Mino, K. (2009). The effects of factor taxation in dynamic economies with externalities and endogenous labor supply. In Kamihigashi, T. and L., Z., editors, *International Trade and Economic Dynamics*, pages 361–382. Springer.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.

Baier, S. L. and Glomm, G. (2001). Long-run growth and welfare effects of public policies with distortionary taxation. *Journal of Economic Dynamics and Control*, 25(12):2007–2042.

Barro, R. and Sala-i Martin, X. (2004). *Economic Growth.* Cambridge MA.: The MIT Press.

Bartelsman, E., Scarpetta, S., and Schivardi, F. (2005). Comparative analysis of firm demographics and survival: evidence from micro-level sources in OECD countries. *Industrial and Corporate Change*, 14(3):365–391.

Bastani, S., Blomquist, S., and Micheletto, L. (2013). The welfare gains of age-related optimal income taxation. *International Economic Review*, 54(4):1219–1249.

Bastani, S. and Waldenström, D. (2018). How should capital be taxed? Theory and evidence from sweden. *C.E.P.R.*, Discussion Paper No. 12880.

Ben-Gad, M. (2017). The optimal taxation of asset income when government consumption is endogenous: Theory, estimation and welfare. *Economic Inquiry*, 55(4):1689–1711.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2017). Are ideas getting harder to find? *National Bureau of Economic Research*, Working Paper No. 23782.

Braulke, M. and Corneo, G. (2004). Capital Taxation May Survive in Open Economies. *Annals of Economics and Finance*, 5(2):237–244.

Bucci, A. (2015). Product proliferation, population, and economic growth. *Journal of Human Capital*, 9(2):170–197.

Chamley, C. (1985). Efficient tax reform in a dynamic model of general equilibrium. *The Quarterly Journal of Economics*, 100(2):335–356.

Chamley, C. (2001). Capital income taxation, wealth distribution and borrowing constraints. *Journal of Public Economics*, 79(1):55–69.

Chen, B.-L., Chen, H.-J., and Wang, P. (2011). Labor-market frictions, human capital accumulation, and long-run growth: Positive analysis and policy evaluation. *International Economic Review*, 52(1):131–160.

Chen, B.-L. and Lee, S.-F. (2007). Congestible public goods and local indeterminacy: A two-sector endogenous growth model. *Journal of Economic Dynamics and Control*, 31(7):2486 – 2518.

Chen, B.-L. and Lu, C.-H. (2013). Optimal factor tax incidence in two-sector human capital-based models. *Journal of Public Economics*, 97:75–94.

Chen, P.-H., Chu, A. C., Chu, H., and Lai, C.-C. (2017). Short-run and long-run effects of capital taxation on innovation and economic growth. *Journal of Macroeconomics*, 53:207–221.

Chen, P.-h., Chu, A. C., Chu, H., and Lai, C.-C. (2019). Optimal capital taxation in an economy with innovation-driven growth. *MPRA paper*, No. 92319.

Chu, A. C. and Cozzi, G. (2014). R&D and economic growth in a cash-in-advance economy. *International Economic Review*, 55(2):507–524.

Chu, A. C. and Cozzi, G. (2019). Growth: Scale or market-size effects? *Economics Letters*, 178:13–17.

Conesa, J. C., Kitao, S., and Krueger, D. (2009). Taxing capital? Not a bad idea after all! *American Economic Review*, 99(1):25–48.

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

Cozzi, G. (2001). Inventing or spying? Implications for growth. *Journal of Economic Growth*, 6(1):55–77.

Cozzi, G. (2017). Endogenous growth, semi-endogenous growth... or both? A simple hybrid model. *Economics Letters*, 154:28–30.

Cozzi, G. and Spinesi, L. (2006). Intellectual Appropriability, Product Differentiation, And Growth. *Macroeconomic Dynamics*, 10(1):39–55.

Cremer, H., Pestieau, P., and Rochet, J.-C. (2003). Capital income taxation when inherited wealth is not observable. *Journal of Public Economics*, 87(11):2475–2490.

de Hek, P. A. (2006). On taxation in a two-sector endogenous growth model with endogenous labor supply. *Journal of Economic Dynamics and Control*, 30(4):655–685.

Dinopoulos, E. and Thompson, P. (1999). Scale effects in Schumpeterian models of economic growth. *Journal of Evolutionary Economics*, 9(2):157–185.

Erosa, A. and Gervais, M. (2002). Optimal taxation in life-cycle economies. *Journal of Economic Theory*, 105(2):338–369.

Gilfillan, S. C. (1935). *The Sociology of Invention*. Chicago: Follett.

Golosov, M., Tsyvinski, A., Werning, I., Diamond, P., and Judd, K. L. (2006). New dynamic public finance: A user's guide. *NBER Macroeconomics Annual*, 21:317–387.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, 58(1):43–61.

Guo, J.-T. and Lansing, K. J. (1999). Optimal taxation of capital income with imperfectly competitive product markets. *Journal of Economic Dynamics and Control*, 23(7):967–995.

Howitt, P. (1999). Steady endogenous growth with population and R.&D. inputs growing. *Journal of Political Economy*, 107(4):715–730.

Jacobs, B. and Schindler, D. (2012). On the desirability of taxing capital income in optimal social insurance. *Journal of Public Economics*, 96(9-10):853–868.

Jones, C. I. (1995). R & D-based models of economic growth. *Journal of Political Economy*, 103(4):759–784.

Jones, L. E. and Manuelli, R. E. (2005). Neoclassical models of endogenous growth: The effects of fiscal policy, innovation and fluctuations. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1, pages 13–65. Elsevier.

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

Long, X. and Pelloni, A. (2017). Factor income taxation in a horizontal innovation model. *Journal of Public Economics*, 154:137–159.

Marrero, G. A. and Novales, A. (2007). Income taxes, public investment and welfare in a growing economy. *Journal of Economic Dynamics and Control*, 31(10):3348–3369.

Martin, F. M. (2010). Markov-perfect capital and labor taxes. *Journal of Economic Dynamics and Control*, 34(3):503–521.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 58(2):317–334.

Mino, K. (2001). Indeterminacy and endogenous growth with social constant returns. *Journal of Economic Theory*, 97(1):203 – 222.

Palivos, T., Yip, C., and Zhang, J. (2003). Transitional dynamics and indeterminacy of equilibria in an endogenous growth model with a public input. *Review of Development Economics*, 7(1):86–98.

Park, H. (2009). Ramsey fiscal policy and endogenous growth. *Economic Theory*, 39(3):377–398.

Park, H. and Philippopoulos, A. (2004). Indeterminacy and fiscal policies in a growing economy. *Journal of Economic Dynamics and Control*, 28(4):645–660.

Pelloni, A. and Waldmann, R. (2000). Can waste improve welfare? *Journal of Public Economics*, 77(1):45–79.

Peñalosa, C. G. and Turnovsky, S. J. (2005). Second-best optimal taxation of capital and labor in a developing economy. *Journal of Public Economics*, 89(5-6):1045–1074.

Peretto, P. F. (1998). Technological change and population growth. *Journal of Economic Growth*, 3(4):283–311.

Peretto, P. F. (2007). Corporate taxes, growth and welfare in a Schumpeterian economy. *Journal of Economic Theory*, 137(1):353–382.

Piketty, T. and Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, 81(5):1851–1886.

Pirttilä, J. and Tuomala, M. (2001). On optimal non-linear taxation and public good provision in an overlapping generations economy. *Journal of Public Economics*, 79(3):485–501.

Reis, C. (2011). Entrepreneurial labor and capital taxation. *Macroeconomic Dynamics*, 15(3):326–335.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):1002–1037.

Saez, E. and Zucman, G. (2019). Progressive wealth taxation. *Brookings Papers on Economic Activity*, BPEA Conference Drafts, September 56, 2019.

Sequeira, T. N., Gil, P. M., and Afonso, O. (2018). Endogenous growth and entropy. *Journal of Economic Behavior & Organization*, 154:100–120.

Smulders, S. and Van de Klundert, T. (1995). Imperfect competition, concentration and growth with firm-specific R & D. *European Economic Review*, 39(1):139–160.

Straub, L. and Werning, I. (2020). Positive long-run capital taxation: Chamley-Judd revisited. *American Economic Review*, 110(1):86–119.

Strulik, H., Prettner, K., and Prskawetz, A. (2013). The past and future of knowledge-based growth. *Journal of Economic Growth*, 18(4):411–437.

Trabandt, M. and Uhlig, H. (2011). The Laffer curve revisited. *Journal of Monetary Economics*, 58(4):305–327.

Wong, T.-N. and Yip, C. K. (2010). Indeterminacy and the elasticity of substitution in one-sector models. *Journal of Economic Dynamics and Control*, 34(4):623 – 635.

Young, A. (1998). Growth without scale effects. *Journal of Political Economy*, 106(1):41–63.

Zeira, J. (2011). Innovations, patent races and endogenous growth. *Journal of Economic Growth*, 16(2):135–156.

Zeng, J. and Zhang, J. (2002). Long-run growth effects of taxation in a non-scale growth model with innovation. *Economics Letters*, 75(3):391–403.

Table 1: Calibration and Steady State Values at Zero Capital Income Tax Rate

Parameter	Description	Benchmark Value
ρ	Discount rate	0.03
β	Production parameter	0.65
σ	IES for consumption (inverse)	2
γ	Growth rate	0.019
L	Labor supply	0.21
η	Frisch elasticity (inverse)	1
p	Innovation success rate	0.014
τ_r	Capital-income tax rate	0
$G/(Y-X)$	Government expenditure to GDP	0.21
$T/(Y-X)$	Government transfers to GDP	0.20
τ_w	Labor-income tax rate	0.6765
k	Labor dis-utility	1.9082
r	Interest rate	0.0680
λ	Quality-step	2.3571
ζ	Parameter in R&D technology	0.0994

Table 2: Optimal Taxation

Variable	Steady-state	Optimal Taxation
τ_r	0	0.2440
τ_w	0.6765	0.5357
γ	0.019	0.0125
L	0.21	0.2401
Welfare Gain	-	0.0393

Note: the table reports the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor. The model is calibrated according to Table 1.

Table 3: Sensitivity Analysis I: one parameter changes at a time

Alternative Value	τ_r	τ_w	γ	L	Welfare Gain
baseline	0.2440 (0)	0.5357 (0.6765)	0.0125 (0.0190)	0.2448 (0.21)	0.0393
$\beta = 0.64$	0.2124 (0)	0.5386 (0.6577)	0.0141 (0.0199)	0.241 (0.2125)	0.0276
$\beta = 0.66$	0.2755 (0)	0.5329 (0.6964)	0.0109 (0.0181)	0.2392 (0.199)	0.0543
$\sigma = 1.8$	0.195 (0)	0.5669 (0.6765)	0.015 (0.0205)	0.2463 (0.2185)	0.0252
$\sigma = 2.2$	0.2844 (0)	0.5091 (0.6765)	0.0105 (0.0177)	0.2343 (0.1954)	0.0531
$\rho = 0.02$	0.1733 (0)	0.5824 (0.6765)	0.0183 (0.0227)	0.2358 (0.2116)	0.0206
$\rho = 0.04$	0.3283 (0)	0.481 (0.6765)	0.0054 (0.0153)	0.2763 (0.2008)	0.0605
$\eta = 0.8$	0.2117 (0)	0.5555 (0.6765)	0.0134 (0.0190)	0.2614 (0.232)	0.0277
$\eta = 1.3$	0.2786 (0)	0.5141 (0.6765	0.0115 (0.0190)	0.2201 (0.1808)	0.0551
$G/(Y - X) = 0.15$	0.1348 (0)	0.5021 (0.5775)	0.0155 (0.0190)	0.2428 (0.2268)	0.0089
$G/(Y - X) = 0.25$	0.3132 (0)	0.5581 (0.7425)	0.0105 (0.0190)	0.2381 (0.1887)	0.0818
$T/(Y - X) = 0.15$	0.1567 (0)	0.5058 (0.594)	0.0149 (0.0190)	0.2529 (0.2333)	0.0126
$T/(Y - X) = 0.25$	0.3282 (0)	0.565 (0.759)	0.0101 (0.0190)	0.2279 (0.1759)	0.0963

Note: the table reports the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of the main parameters and of the size of the public sector. The corresponding steady-state values at zero tax rate on capital income are reported in parentheses.

Table 4: Sensitivity Analysis II: k and ζ vary

Alternative Value	τ_r	τ_w	γ	L	Welfare Gain
baseline	0.2440 (0)	0.5357 (0.6765)	0.0125 (0.0190)	0.2448 (0.206)	0.0393
$\beta = 0.64$	0.2181 (0)	0.5347 (0.6577)	0.0132 (0.0190)	0.2345 (0.206)	0.0291
$\beta = 0.66$	0.2699 (0)	0.5371 (0.6964)	0.0118 (0.0190)	0.2466 (0.206)	0.0522
$\sigma = 1.8$	0.2062 (0)	0.5594 (0.6765)	0.0134 (0.0190)	0.2339 (0.206)	0.0281
$\sigma = 2.2$	0.2766 (0)	0.5149 (0.6765)	0.0117 (0.0190)	0.2458 (0.206)	0.0504
$\rho = 0.02$	0.1938 (0)	0.5695 (0.6765)	0.0147 (0.0190)	0.2325 (0.206)	0.0255
$\rho = 0.04$	0.3007 (0)	0.5029 (0.6765)	0.009 (0.0190)	0.2784 (0.206)	0.0510
$\eta = 0.8$	0.2117 (0)	0.5555 (0.6765)	0.0134 (0.0190)	0.2320 (0.206)	0.0277
$\eta = 1.3$	0.2786 (0)	0.5141 (0.6765)	0.0115 (0.0190)	0.2508 (0.206)	0.0551
$G/(Y-X) = 0.15$	0.1348 (0)	0.5021 (0.5775)	0.0155 (0.0190)	0.2206 (0.206)	0.0089
$G/(Y-X) = 0.25$	0.3132 (0)	0.5581 (0.7425)	0.0105 (0.0190)	0.2599 (0.206)	0.0818
$T/(Y-X) = 0.15$	0.1567 (0)	0.5058 (0.594)	0.0149 (0.0190)	0.2233 (0.206)	0.0126
$T/(Y-X) = 0.25$	0.3282 (0)	0.565 (0.759)	0.0101 (0.0190)	0.2668 (0.206)	0.0963

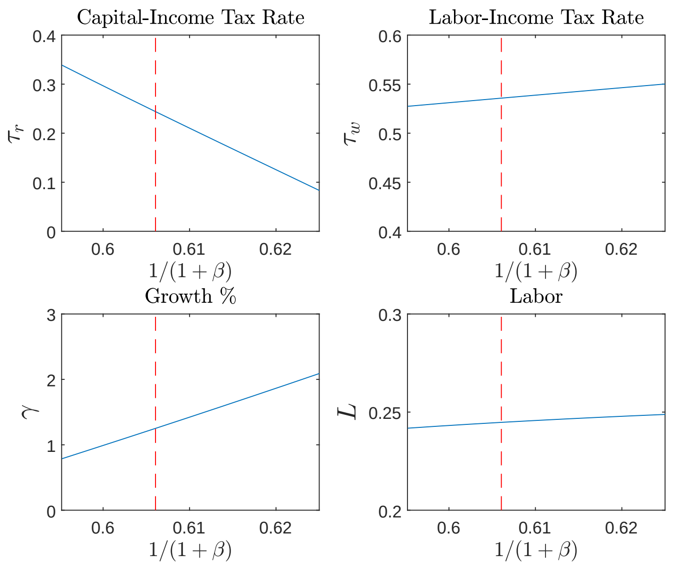
Note: the table reports the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of the main parameters and of the size of the public sector. We change each parameter or variable at a time along with the scale parameters k and ζ, so as to leave γ and L unchanged. The corresponding steady-state values at zero tax rate on capital income are reported in parentheses.

Table 5: Optimal Taxation - Dilution Effect

Variable	**Steady-state**	**Dilution Effect**
τ_r	0	0.5297
τ_w	0.6765	0.3435
γ	0.019	0.0036
L	0.21	0.2810
Welfare Gain	-	0.3131

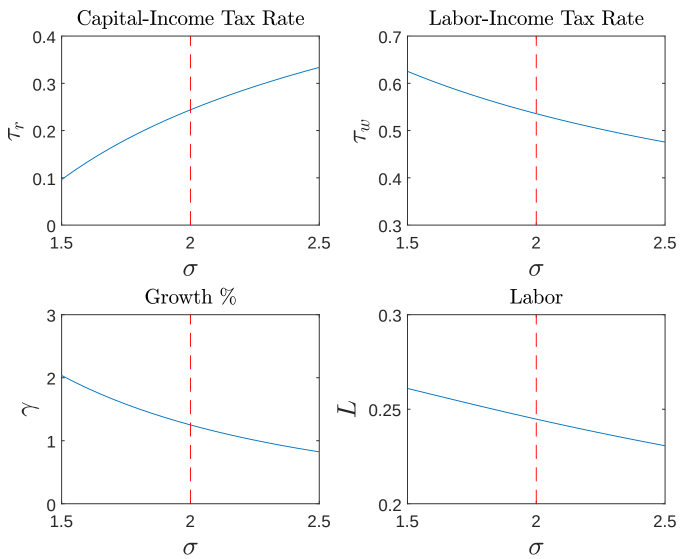
Note: the table reports the optimal tax mix, growth and labor for the model with "dilution effect". The corresponding steady-state values at zero tax rate on capital income are reported in parentheses.

Figure 1: Optimal Taxation and the Labor Income Share



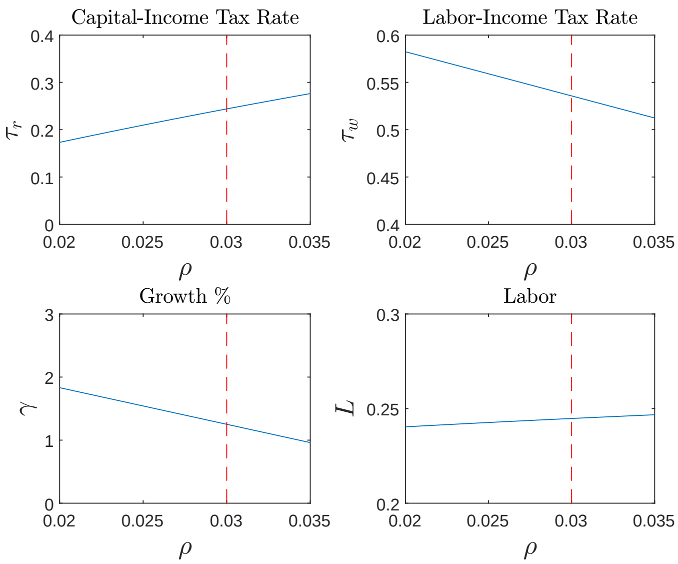
Note: the figure plots the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of $1/(1+\beta)$, the labor income share. Vertical dashed lines represent the baseline value.

Figure 2: Taxation and Labor and the Intertemporal Elasticity of Substitution



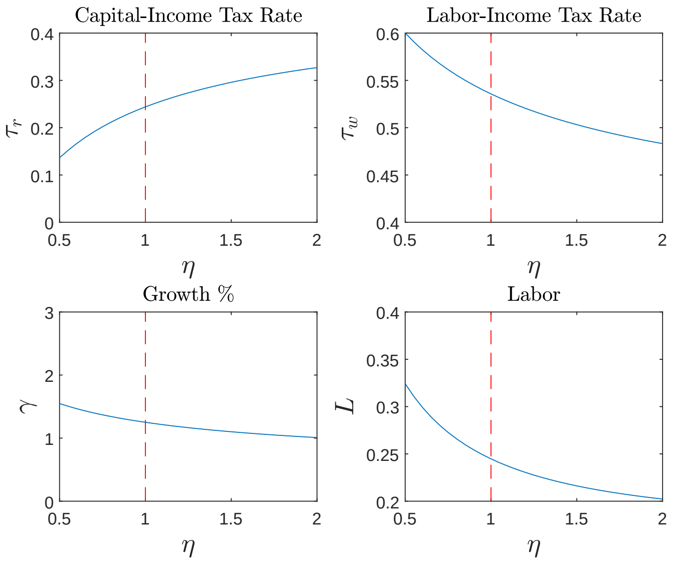
Note: the figure plots the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of σ, the inverse of the intertemporal elasticity of substitution. Vertical dashed lines represent the baseline value.

Figure 3: Taxation and Labor and the Rate of Time Preference



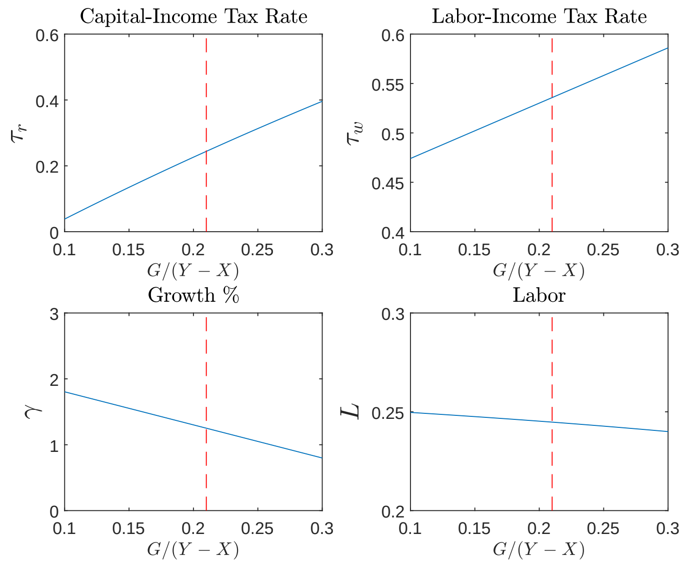
Note: the figure plots the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of ρ, the rate of time preference. Vertical dashed lines represent the baseline value.

Figure 4: Taxation and Labor and the Frisch Elasticity



Note: the figure plots the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of η, the Frisch elasticity. Vertical dashed lines represent the baseline value.

Figure 5: Taxation and Labor and Government Expenditure



Note: Note: the figure plots the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of $G/(Y - X)$, the government expenditure to GDP. Vertical dashed lines represent the baseline value.