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

Barbara Annicchiarico[†] Valentina Antonaroli[‡] Alessandra Pelloni[§]

May 2021

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. 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.

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

JEL classification: *O41, E62, H21.*

*We are very grateful to two anonymous referees and to the co-editor, Shankha Chakraborty, for their excellent and constructive comments. 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), Saez and Zucman (2019) and Bastani and Waldenström (2020). On the other hand, traditional arguments against capital taxation have 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. 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. Unprecedented steps in this direction have been taken by the Biden Administration by proposing a plan to tax negotiators from 135 countries at the OECD to force companies to pay taxes where their revenues are earned, not where the profits can be shifted to.[1]

In this paper, we propose an argument in favor of capital income taxation based purely on efficiency grounds and meant to complement arguments based on equity grounds. To work out our argument we study the trade-off between taxing capital income and taxing labor income through the lenses of innovation-led growth theory. According to this theory, in all its variants, the engine of growth is the discovery of ideas for new processes and products.[2] It is technically possible for any number of people to use an idea at the same time: an important implication of the non rivalry of ideas is the "market size" effect. A higher level of economic activity will allow ideas to be used more intensively and this will increase efficiency, because the cost of each discovery per user will be lower. 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 the "market size" effect can be 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 may be efficient has already been shown in models with a strong scale effect. Pelloni and Waldmann (2000) and Amano et al. (2009) show that taxing capital can be useful 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

[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.

[2]For a long list of references on empirical evidence in favor of the link between R&D, innovation and growth, see the survey in Mohnen (2019).

on have been questioned exactly because of the increasing relationship between growth and labor they feature.[3] One could be justified in believing that the result of positive capital taxation obtained in these models is made possible by the debated positive link between growth and labor they incorporate. Our analysis 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.[4] A second generation endogenous growth approach has evolved from 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.[5] Our analysis incorporates this insight in the prototypical Schumpeterian vertical innovation model proposed by Grossman and Helpman (1991), simply by assuming that innovation in a line of products is more complex and costly depending on the size of the line itself, measured by generated value added, as in Barro and Sala-i-Martin (2004).

In the so-called horizontal innovation models, R&D leads to the invention of new goods. The observation from which vertical innovation models start is simply that most growth innovations consist in the increase in quality or in the reduction in cost of existing goods, rather than in the invention of new goods. Vertical innovation models are often labeled "Schumpeterian" because they have made operational the Schumpeterian notion of "Creative Destruction". In these models there is a positive association between innovation and higher turnover rates which is borne out by micro evidence on firms dynamics.[6] The association is also important for normative analysis. When new innovations replace old technologies, so that there are winners and losers, a rich set of spillovers arise which are relevant for optimal taxation.[7]

[3]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. (2020) 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.

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

[5]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. For evidence on long-run trends in R&D and TFP favoring fully endogenous "Schumpeterian" growth theory over semi-endogenous growth theory, see Ha and Howitt (2007).

[6]Foster et al. (2001) offer influential evidence on the importance for TFP growth of replacement of less productive with more productive plants.

[7]Aghion et al. (2014) provide an outline of Schumpeterian growth theory as an integrated framework, that has been widely confronted with data, for understanding various aspects of growth not fully captured by alternative theories. They show how the emphasis of the theory on the reallocation of resources among incumbents and new entrants leads to predictions (e.g. on the size, age and distribution of firms, as well as on the impact on growth of macro determinants, such as rule of law, openness and proximity to the technological frontier) that are widely supported by micro and macro data. Using long-run panel data for OECD countries Madsen and Timol (2011) find evidence of convergence in manufacturing productivity

In our model, the economy is populated by an infinitely lived representative household deriving utility from consumption and leisure. In the final good sector a representative perfectly competitive firm uses labor and differentiated intermediates as production inputs. Intermediates are instead produced by imperfectly competitive firms using final goods. A first externality arises because monopoly power means that the price of each intermediate is higher than its marginal cost. This "market power" distortion is familiar from the standard analysis of monopoly and arises in horizontal innovation models as well.

Each intermediate has a quality ladder along which improvements occur thanks to R&D activity. Notably, in the Schumpeterian models growth involves both positive and negative externalities. The first is the "incomplete appropriability" externality: namely, any new innovation raises aggregate productivity for ever; however of the whole social surplus so generated the innovator captures only the part corresponding to the profits from her/his innovation, to the exclusion of the part going to labor income. Moreover, profit flows stop when the next innovation occurs. This externality occurs in horizontal innovation models as well but to a lesser degree because profits from an invention accrue to the inventor for ever. By contrast, in Schumpeterian models, the introduction of a leading-edge technology involves a negative externality as the rents of the previous innovator are canceled. Aghion and Howitt (1992) define this the "business stealing" effect of innovation. Another externality is represented by the fact that the amount of government expenditures is increasing in income, a circumstance not taken into account by agents when making their economic choices. We label this externality the "weight of government" distortion.

We follow the Ramsey approach in limiting the tax instruments available: in particular, income taxes are linear and time-invariant, while the government budget is balanced at all times.

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 for the final good 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. The total quantity of rents that can be captured by successful innovators also goes up: indeed in first generation endogenous growth models implies a positive link between growth and labor. In these models decreasing the tax on labor could increase growth in spite of the concurrent increase in the capital tax. However, in our model the scale effect has been eliminated. In fact, more employment also means that additional research input is needed for innovating, so that the two effects of more profitability of innovations through higher demand on the one hand and lower profitability, due to higher costs of doing research on the other hand, cancel out in equilibrium, and the growth rate is unaffected by employment. This means that increasing the tax rate on capital will always compress growth and worsen the "incomplete appropriability" distortion.[8] However, this growth compression reduces the negative externality represented by the "business stealing" effect: in general, through this effect, in a Schumpeterian model, a tax on saving will have a lower social cost in terms of reduced growth than it would in a horizontal innovation model. Finally, capital taxes will decrease the "weight of government" distortion.

driven by domestic R&D and international R&D spillovers as predicted by Schumpeterian models. For further evidence buttressing Schumpeterian theory, see Acemoglu et al. (2018) and Kogan et al. (2017).

[8]For empirical evidence on the relationship between capital taxes and growth see Djankov et al. (2010).

Summing up, shifting the tax burden from labor to capital will reduce the social welfare losses implied by the distortions due to market power, creative destruction and the weight of government, while the welfare loss from the appropriability of inventions distortion will go up. For reasonable parameters' values we find that social welfare is maximized when the optimal tax rate on capital income, though lower than the rate on labor income, is not only positive but sizable: about 24 percent in our baseline case.

Generally, in papers on taxation and growth the focus is on growth effects, in the belief that these effects will always prevail on level effects as regards welfare calculations, bar distributional considerations. However, we show that this is not necessarily the case. To repeat, in our model even though growth is inefficiently low and decreasing in the capital tax, it can be optimal to reduce growth even more, so that households can increase current consumption.

Two further results stand out for their policy relevance. The first is that the optimal tax on capital should be higher the higher the capital's income share. This is particularly interesting given the observed increase in the share of capital income since the 1980s, as shown by Akcigit and Ates (2021). Another important result of our analysis is that a larger size of the public sector leads to a higher optimal tax rate on capital. This is relevant for policy making in advanced and emerging economies, given Wagner's law, i.e. the prediction that development will be accompanied by an increased share of public expenditure in gross national product (see Jalles 2019). Indeed looking at current economic events, the effect could become especially salient, if, as is likely, public actions implemented across countries as a response to the SARS-CoV-2 epidemics will not be totally reversed, just think of spending on those health systems whose inadequacies have been exposed during the crisis.

By a rich set of simulation exercises we show the robustness of our results when parameters change: we find that the optimal tax rate on capital is increasing in the rate of time discount, in the compensated elasticity of labor supply, in the inverse of the intertemporal elasticity of substitution in consumption and in the marginal cost of intermediates, while decreasing in the productivity of research. The tax on capital will also be higher when R&D expenses are verifiable and subsidized, when government expenditures positively affect the production function in the final sector and in the presence of public debt, as we show when we extend the model to allow for these other kinds of public outlays. Finally, to shed some light on the interplay between efficiency and equality goals in defining the optimal tax scheme we extend the model to two classes of agents, with wealth being owned by one class while the other class is composed of hand-to-mouth workers. We show that the more weight is given in the social welfare function to the representative hand-to mouth worker' welfare the more capital income will have to be taxed.

The simple structure of our model allows us to provide closed form solutions for the socially optimal and the market level of labor both of labor and growth. Since taxing capital discourages growth, while taxing labor depresses employment, if a certain parameter leads to an increase in the difference between the socially optimal and the market level of labor (rate of growth), then the tax on labor (capital) will be decreasing in that parameter.

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 as well as sensitivity and robustness checks, and Section 7 presents some extensions. The last section concludes.

2 Related Literature

This paper contributes to the literature on the optimal taxation of labor and asset income.[9] 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 already been studied in the literature.

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, 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

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

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.

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 weighting the disutility 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.[10]

[10]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 Appendix A.

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, and F_t denotes the real value of financial assets. The decision of households to finance firms is akin to the decision to accumulate physical capital in the standard Ramsey model. The government imposes a labor income tax τ_w and a capital income tax τ_r – both constant over time – and returns part of this revenue as lump-sum transfers T_t. By a tax on capital income then, as in the Chamley and Judd models, we mean a tax on income from savings.

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

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

and the transversality condition

$$\lim_{t\to\infty} C_t^{-\sigma}F_t e^{-\rho t} = 0. \tag{4}$$

See Appendix A. Optimization at an interior point implies that the marginal rate of substitution between consumption and leisure 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.

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 $\nu \in [0,1]$, whose quality at time t 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 ν 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. Quality at time 0 is normalized at one, that is $q_0(\nu) = 1$.

Profit maximization gives the demand function of the inputs, i.e. for the intermediate good ν,

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

and for labor,

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

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 ψ to β, with no loss of generality,[11] while $q_t(\nu)$ is the unconstrained monopoly price.[12]

Plugging $P_t(\nu \mid q)$ into (8), we obtain the quantity of the intermediate good,

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

Because of market power this quantity is inefficiently low: this is what we have labelled the "market power" distortion. 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.

[11]In Appendix B we derive all the analytical results for $\psi \neq \beta$. We are then able to show that the value of ψ is irrelevant per se and that what counts is the relation of this value to the value of ζ, a parameter measuring the productivity of research, which we will introduce later. These two coefficients appear in the solution of the model combined in the same function, so that only the value of this function and not the individual values of the two coefficients can be obtained by implication when calibrating the model. This means that we are free to fix one of the two parameters at any positive value less than unity, e.g. to fix $\psi = \beta$, and use the value of the aforementioned combining function to obtain the other parameter, ζ. In Section 6 we show how results change when ψ moves away from β, while keeping the value of ζ fixed.

[12]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.

Note that (12) says that a higher labor supply means a higher quantity of each intermediate good, and thus higher profits in equilibrium. This "market size" generates 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 line ν with initial quality q at time t. The value of a firm in this line of product, denoted by $V_t(\nu \mid q)$, obeys the following standard arbitrage condition:

$$r_t = \frac{\dot{V}_t(\nu \mid q)}{V_t(\nu \mid q)} - p_t(\nu \mid q) + \frac{\pi_t(\nu \mid q)}{V_t(\nu \mid q)}. \tag{13}$$

This means that the rate of capital gains $\dot{V}_t(\nu \mid q)/V_t(\nu \mid q)$ minus the expected rate of capital loss $p_t(\nu \mid q)$ plus the rate of profit $\pi_t(\nu \mid q)/V_t(\nu \mid q)$ must equal the market rate of return r_t. $p_t(\nu \mid q)$ is the expected rate of capital loss in this line, because when 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. At the same time a new discovery expands the production possibility frontier of the economy for ever, in a measure corresponding to the rise in GDP, part of which goes to labor income. However, a potential innovator decides whether to engage in R&D on the basis of future profits only, thus neglecting part of the social surplus from the new invention. This "incomplete appropriability" externality represents a third market failure in the model.

We now describe how new machine vintages are invented by R&D, i.e. specify the technology for innovating. If a firm spends $z_t(\nu \mid q)$ units of the final good for research in line ν when quality is at level q, then it generates a flow rate

$$p_t(\nu \mid q) = \frac{\zeta z_t(\nu \mid q)}{L_t q_t(\nu)}, \ \zeta > 0, \tag{14}$$

of innovation.[13] The innovation leads to a new rung in the quality ladder in line ν where quality moves from $q_t(\nu)$ to $\lambda q_t(\nu)$. The flow rate of innovation in a sector ν 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).[14] The idea is that innovation tends to be more expensive in a large market, since the costs associated with the discovery, the development and

[13] This "lab-equipment" assumption just means that the same technology is used for producing innovations and for producing material goods.

[14] As already mentioned the specification in (14) follows Barro and Sala-i-Martin (2004, chapter 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

the marketing of new technologies are then higher.[15] 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 toward large firms and they can self-protect through the creation of patent thickets (see Akcigit and Ates 2021).

We assume that there is free-entry into the research market and 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 ν that has quality $q\lambda^{-1}$ at time t must then be zero: $p_t(\nu \mid q\lambda^{-1})V_t(\nu \mid q) - z_t(\nu \mid q\lambda^{-1}) = 0$. If initial quality is $q\lambda^{-1}$ the technological assumption in (14) implies $p_t(\nu \mid q\lambda^{-1}) = \frac{\zeta z_t(\nu|q\lambda^{-1})}{L_t q_t(\nu)\lambda^{-1}}$. Combining the last two equations we arrive to:

$$V_t(\nu \mid q) = \frac{L_t q_t(\nu)}{\zeta\lambda}. \tag{15}$$

We consider now the entrepreneur's choice of industry in which to target R&D efforts, and the optimal scale of those efforts. We have seen that the profit flows are proportional to the quality level in each industry (see 12). Since the R&D costs to be incurred to reach a given arrival rate of innovation in an industry are also proportional to the quality level of the industry, we infer that an entrepreneur will be indifferent as to the industry in which she/he devotes her/his R&D efforts provided that she/he expects her/his prospective leadership position to last equally long in each one.[16]

Formally, consider that for the monopolist in line ν quality q is a given, so that from (15) $\frac{\dot{V}_t(\nu|q)}{V_t(\nu|q)} = \frac{\dot{L}_t}{L_t}$. At the same time, from (12) and (15) we have:

$$\pi_t(\nu \mid q)/V_t(\nu \mid q) = (1-\beta)\zeta\lambda. \tag{16}$$

Hence the arbitrage condition (13) becomes:

$$r_t = \dot{L}_t/L_t - p_t + (1-\beta)\zeta\lambda, \tag{17}$$

where we have dropped the arguments in $p_t(\nu \mid q)$ since the equation shows that the arrival rate of innovation is the same across all lines.[17]

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).

[15]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.

[16]Barro and Sala-i-Martin (2004, chapter 7) explain that the hypothesis that the probability of innovation in each sector is equal across sectors in equilibrium is necessary for endogenous growth to obtain. If the expected reward from innovating increased more steeply than the cost of innovating with the quality level of a sector, then the more advanced sectors would grow faster and, in general, the aggregate rate of growth would increase over time, as the average quality would increase. If the increase in costs effect dominated, we would have decreasing returns in the R&D sector and the rate of growth would decline over time. Endogenous steady-state growth requires that the reward and cost effects cancel out.

[17]Notice, however, that innovation will not be uniform across sectors. There will be an evolving

3.4 Government

To close the model, we assume that government expenditures G equal a fixed fraction g of gross output, another fraction of which is transferred back to families as lump-sum transfers:

$$G_t = gY_t, \tag{18}$$

and

$$T_t = \tau Y_t, \tag{19}$$

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.[18] Our assumption of an exogenously given flow of government disbursements comes from the standard Chamley-Judd analysis and is made mainly for convenience and for easy comparability of our results.[19] 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.[20]

It should be noted that the endogeneity of government expenditures with respect to income is not taken into account by households. This is the "weight of government" distortion of the model.

We rule out a market for government bonds and assume that the government runs a balanced budget at all times.[21] The assumptions of constant tax rates and of a balanced government budget are adopted as they greatly simplify computations and make the intuition of results more transparent.[22] Moreover, though restrictive, these assumptions do not seem unreasonable in a model with no aggregate uncertainty and no room for countercyclical policies. Indeed, complicated time dependent policies, as those studied in the literature started by Chamley (1985) and Judd (1985), do not appear plausible or empirically relevant.[23]

Under the above assumptions the budget constraint of the government immediately

distribution of product qualities, with individual products constantly swapping relative positions within that distribution.

[18]Consider all total value-added; this is given, using (23), by $Y - X = Y(1-\beta^2)$. Then, it is straightforward that imposing $0 < G < Y(1-\beta^2)$ gives us the restriction $0 < g < 1-\beta^2$.

[19]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.

[20]That government consumption does not directly affect the choices of agents 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. In Section 7 we will consider the case of G entering the production function in the final sector. In an other extension public money will subsidise R&D.

[21]Given the long-run perspective of our analysis, a balanced budget rule may represent a fair approximation of a fiscal conduct consistent with the sustainability of public finances.

[22]These assumptions also make our results easier to compare with those obtained in our closer antecedents in the literature, from Aghion et al. (2013) to Chen et al. (2019) with which we share the assumptions.

[23]Chamley and Judd prescribed taxes at confiscatory rates in the short run, when capital is in quasi-fixed amount and at zero rates in the long run. Piketty (2010, p. 826) then asks: "How can one apply these results to the real world? Is today the short run or the long run?" Actually the implementation of these schemes would require a commitment technology not observed in reality and whose use could be difficult to reconcile with the basic democratic principle of alternation in political power.

follows

$$G_t = w_t \tau_w L_t + r_t \tau_r F_t - T_t. \tag{20}$$

3.5 Aggregation and Market Equilibrium

We now define the aggregate quality index Q_t as a combination of the various quality improvements:

$$Q_t = \int_0^1 q_t(\nu) \mathrm{d}\nu. \tag{21}$$

Using equations (11) and (21) into the production function for the final good in (6), we get total output in the final good sector:

$$Y_t = \frac{1}{\beta} Q_t L_t \tag{22}$$

which shows how increases in quality affect aggregate output. Let $X_t = \int_0^1 \psi q_t(\nu) x_t(\nu \mid q) d\nu$ denote the aggregate expenditure on final goods used to produce intermediate goods, then using (11) and (21), and recalling the normalization $\psi = \beta$ we have:

$$X_t = \beta Q_t L_t. \tag{23}$$

From (22) and (23) we have a relation between intermediate goods and final output, i.e.

$$X_t = \beta^2 Y_t. \tag{24}$$

From (9) and (22) we can find an expression for the wage as a function of Q_t

$$w_t = \frac{(1-\beta) Q_t}{\beta}. \tag{25}$$

From (12) and (21) total profits of the intermediate goods sector are given by

$$\Pi_t = \int_{\nu=0}^1 \pi_t(\nu \mid q) \mathrm{d}\nu = (1-\beta) Q_t L_t. \tag{26}$$

Let $Z_t = \int_0^1 z_t(\nu \mid q) \mathrm{d}\nu$ denote the aggregate spending on R&D of firms starting from level of quality q, then using (14) and (21) we get

$$Z_t = \int_0^1 z_t(\nu \mid q) \mathrm{d}\nu = \frac{p_t}{\zeta} Q_t L_t. \tag{27}$$

Finally, 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 (21), we can write:

$$F_t = \frac{L_t Q_t}{\lambda \zeta} \tag{28}$$

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

In equilibrium, the following market clearing conditions for final goods, where investment is represented by Z_t, and labor are satisfied:

$$Y_t - X_t = C_t + Z_t + G_t, \tag{29}$$

$$H_t = L_t. \tag{30}$$

We now want to obtain a relationship between the two tax schedules in equilibrium. Consider the government's budget constraint in (20) then, using condition (28) and the definitions for G_t and T_t, we can write:

$$\tau_w = \frac{(g+\tau)Y_t - r_t \tau_r F_t}{w_t L_t}. \tag{31}$$

Next, use (9) for $w_t L_t$, and combine equations (22) and (28), so we can use $F_t = \beta Y_t / \lambda\zeta$. We end up with the following:

$$r = \left(\frac{g+\tau}{1-\beta} - \tau_w \right) \frac{\lambda\zeta}{\tau_r \beta}(1-\beta). \tag{32}$$

From this equation we can see that if policy variables do no change over time r 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), (23) and (25); the government budget constraint (20); the market clearing conditions for wealth (28), for final goods (29) and for labor (30).

Using (5), (18), (22), (23) and (27), given (9) and (30), into the resource constraint (29), so as to express all variables in terms of $Q_t L_t$, and then simplifying we obtain an expression of p_t as a function of L_t:

$$p_t = \left(1 - \beta^2 - g - \frac{\eta \left(1 - k(1-\sigma) L_t^{1+\frac{1}{\eta}} \right) (1-\beta)(1-\tau_w)}{\sigma k (1+\eta) L_t^{1+\frac{1}{\eta}}} \right) \frac{\zeta}{\beta}. \tag{33}$$

Plugging (32) and (33) into (17) we obtain the following non linear differential equation for L_t :

$$\dot{L}_t = a L_t - b L_t^{-\frac{1}{\eta}}, \tag{34}$$

where
$a \equiv \left[1 - \beta^2 - g + \frac{\eta(1-\sigma)(1-\beta)(1-\tau_w)}{\sigma(1+\eta)} + \left(\left(\frac{g+\tau}{1-\beta} - \tau_w \right) \frac{1}{\tau_r} - \beta \right) \lambda(1-\beta) \right] \frac{\zeta}{\beta}$,
$b \equiv \frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)} \frac{\zeta}{\beta} > 0$.

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:

$$L = \left(\frac{b}{a} \right)^{\frac{\eta}{1+\eta}}. \tag{35}$$

which is clearly uniquely pinned down by the parameters of the model and by the policy variables. Of course the model will only be well specified if L belongs to the interval $(0,1)$. In particular, for L to be positive we need a positive a, since b is always positive.

The restriction $a > 0$ helps in establishing the dynamic nature of the fixed point of (34). From (34) we can calculate:

$$\frac{d\dot{L}_t}{dL_t}_{|L_t=L} = a + \frac{1}{\eta} b L^{-\frac{1}{\eta}-1} > 0. \tag{36}$$

The differential equation (34) is therefore unstable which implies that in equilibrium L will be constant at the value (35). It is easy to infer that the whole economy will then be on its BGP at all times. In fact, if L is constant over time, so will p, given (33), while from (3) we deduce that consumption will grow at a constant rate which we denote by γ. From (5) we see that w will grow at the same constant rate as C, and so, by (9), will Q, as defined by (21). Indeed Y, X, Z, Π, F, G and T will all grow at the same rate, as they are linear functions of Q. That this economy will always be in a unique balanced growth equilibrium 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.[24]

To pin down the common rate of growth, γ, we proceed as follows. From (7) we know that $q_t(\nu)$ is a function of $n_t(\nu)$. From the analysis so far n has a Poisson distribution with intensity pt so that the probability of exactly n improvements from 0 to t is

$$f(n,t) = \frac{(pt)^n e^{-pt}}{n!}. \tag{37}$$

Since we have shown that in equilibrium this intensity is common to all products, $f(n,t)$, by the law of large numbers, represents the measure of products that are improved n times during an interval of time of length t. From (21) we then have:

$$Q_t = \int_0^1 q_t(\nu) \mathrm{d}\nu = \sum_{n=0}^{\infty} f(n,t)\lambda^n = e^{-pt} \sum_{n=o}^{\infty} \frac{(\lambda pt)^n}{n!} = e^{(\lambda-1)pt}. \tag{38}$$

In Appendix A we give a more detailed derivation. So even if the quality of each product can change discretely over time, Q_t is a continuous function whose rate of growth is given by:

$$\gamma = \frac{\dot{Q}_t}{Q_t} = (\lambda - 1)p. \tag{39}$$

From (15) we know that along a BGP, when L is given, for given q, the value of a firm in a given line ν will be fixed, so from (13) we arrive to

$$V(\nu \mid q) = \frac{\pi(\nu \mid q)}{r+p}. \tag{40}$$

Substituting in this equation the value for $V(\nu \mid q)/\pi(\nu \mid q)$ from (16) we obtain:

$$r + p = \lambda\zeta(1-\beta). \tag{41}$$

From (3), remembering that consumption grows at rate γ, we can write:

$$\gamma = \frac{r(1-\tau_r) - \rho}{\sigma}. \tag{42}$$

[24]See Pelloni and Waldmann (2000), Mino (2001), Palivos et al. (2003), Park (2009), Chen and Lee (2007), and Wong and Yip (2010), among others.

From this formula we can already see that the growth rate is positively related to the after-tax interest rate on capital. From (39), (41) and (42) we then get:

$$r = \frac{\rho + \sigma\lambda\zeta(1-\beta)(\lambda-1)}{1-\tau_r+\sigma(\lambda-1)}, \tag{43}$$

and

$$\gamma = \frac{(\lambda-1)\left[\lambda\zeta(1-\beta)(1-\tau_r)-\rho\right]}{1-\tau_r+\sigma(\lambda-1)}, \tag{44}$$

as well as

$$p = \frac{\lambda\zeta(1-\beta)(1-\tau_r)-\rho}{1-\tau_r+\sigma(\lambda-1)}. \tag{45}$$

Just looking at (44) we can see that the growth rate is decreasing in ρ, the rate of time discount, and in σ, the inverse of the intertemporal elasticity of substitution, while 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 (4) we must impose $r(1-\tau_r) > \gamma$.[25]

Combining (32), (35) and (43) we can write:

$$L = \left(\frac{1}{\frac{\sigma k(1+\eta)}{\eta(1-\tau_w)}\left[1-\frac{g}{1-\beta}+\beta\frac{\frac{\rho}{\zeta(1-\beta)}+(\sigma-1+\tau_r)(\lambda-1)}{1-\tau_r+\sigma(\lambda-1)}\right]-k(\sigma-1)} \right)^{\frac{\eta}{1+\eta}}. \tag{46}$$

By substituting the BGP expression for r into (32), we can compute the relationship between the two tax schedules,

$$\tau_w = \frac{\tau+g}{1-\beta} - \frac{\beta\left[\rho+\sigma\lambda\zeta(1-\beta)(\lambda-1)\right]}{\lambda\zeta(1-\beta)\left[1-\tau_r+\sigma(\lambda-1)\right]}\tau_r. \tag{47}$$

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. We want to express the welfare function in terms of the tax rate on capital in order to find its maximum value. Neglecting the constant term, (1) can be expressed as

$$W_0 = \int_{t=0}^{\infty} e^{[\gamma(1-\sigma)-\rho]t}\frac{C_0^{1-\sigma}\left(1-k(1-\sigma)L^{1+\frac{1}{\eta}}\right)^{\sigma}}{1-\sigma}\mathrm{dt}, \tag{48}$$

[25]The condition for $\gamma > 0$ in (44) is $\rho < \lambda\zeta(1-\beta)(1-\tau_r)$.

where C_0 is consumption at time 0. Solving the integral in (48), we obtain

$$W_0 = \frac{1}{1-\sigma} \frac{C_0^{1-\sigma} \left(1 - k(1-\sigma)L^{1+\frac{1}{\eta}}\right)^{\sigma}}{\rho - \gamma(1-\sigma)}, \tag{49}$$

where we have used the fact that $\gamma(1-\sigma) - \rho < 0$, implied from the transversality condition (4). Indeed the inequality always holds if $\gamma > 0$ since we are assuming that $\sigma > 1$.[26]

From (5) we have that:

$$C_0 = \frac{\eta Q_0(1-\beta)(1-\tau_w)\left(1 - k(1-\sigma)L^{1+\frac{1}{\eta}}\right)}{\beta\sigma k(1+\eta)L^{\frac{1}{\eta}}}, \tag{50}$$

where we have used (25) to eliminate w. C_0 can be expressed in terms of τ_r by using (35) to eliminate L and by using (47) to eliminate τ_w. Finally, γ can be expressed in terms of τ_r by using (44). 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. 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. In particular we want to understand if the market growth rate is too high or too low from a social point of view. As already mentioned, 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,

$$\dot{Q}_s = \frac{\zeta(\lambda - 1)Z_s}{L_s}. \tag{51}$$

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

[26]For households:

$$\lim_{t\to\infty} C_t^{-\sigma} F_t e^{-\rho t} = 0.$$

economy, adjusted for the size of employment L_s. In the centralized economy, the following condition for each intermediate holds:

$$x_s(\nu, t \mid q) = \beta^{\frac{1}{\beta-1}} L_s \tag{52}$$

which is the social planner equivalent of the decentralized quantity in (8). Using the definition of aggregate spending on equipments, we have

$$X_s = \beta^{\frac{\beta}{\beta-1}} Q_s L_s. \tag{53}$$

Monopoly pricing implies that the privately chosen quantity (equation 23) is smaller than the socially chosen amount, since $\beta^{\frac{\beta}{\beta-1}} > \beta$. Final output in equilibrium can be computed from (6) and (53) and is expressed as:

$$Y_s = \beta^{\frac{1}{\beta-1}} Q_s L_s. \tag{54}$$

For given Q_s and labor, the level of output in the decentralized economy is lower than the optimal social value (recall equation 22), since $\beta^{\frac{1}{\beta-1}} > 1/\beta$. 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:

$$\dot{Q}_s = \frac{\zeta(\lambda-1)}{L_s} \left(\beta^{\frac{1}{\beta-1}} Q_s L_s - \beta^{\frac{\beta}{\beta-1}} Q_s L_s - g\beta^{\frac{1}{\beta-1}} Q_s L_s - C_s \right). \tag{55}$$

Solving the optimization problem of the social planner and combining the necessary conditions for a maximum, we obtain an expression for optimal employment:

$$L_s = \left(\frac{\eta}{k(\sigma+\eta)} \right)^{\frac{\eta}{\eta+1}} \tag{56}$$

and for the growth rate of consumption:

$$\frac{\dot{C}_s}{C_s} = \gamma_s = \frac{\zeta\beta^{\frac{1}{\beta-1}}(\lambda-1)(1-\beta-g)-\rho}{\sigma}. \tag{57}$$

See Appendix A for details.

Notice that L_s does not depend on technological parameters. This is because for preferences to be consistent with unbounded growth we need the substitution and income effects of an increase in the marginal productivity of labor to cancel out. Nor does L_s depend on the rate of time discount, since by construction the level of employment does not effect growth.

The social optimum growth rate, γ_s is increasing in ζ and λ, the two parameters determining the productivity of research, and decreasing in all other parameters.

Recall the market growth rate in (44). We want to compare γ and γ_s, but it 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 parameterize the model in order to examine quantitatively the effects of increasing the capital income tax rate on labor supply, economic growth and social welfare.

A caveat is in order. As is always the case when working with parsimonious models like ours, simulations are aimed at elucidating new theoretical findings and at getting a first assessment of their relevance. The knowledge acquired through a clearly understood analytical structure can then be used in the development of larger scale models that, after being matched to the data of a particular economy or groups of economies, can help in the formulation of concrete policies.[27] It should be noted that this is usually the approach adopted in optimal tax analyses, where relevant features of the economy are modeled separately because the use of settings with many complications would make it difficult to isolate the underlying mechanisms.

6.1 Parametrization and the Optimal Tax Mix

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

In our model, per capita GDP growth and TFP growth coincide. Using Eurostat and AMECO annual 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 comes from Trabandt and Uhlig (2011): the Frisch elasticity, η, is set equal to one. This value represents an intermediate value in the range of macro and micro data estimates. The inter-temporal elasticity of substitution of consumption is set to 0.5, corresponding to $\sigma = 2$. We choose 0.18 as our benchmark value for labor L, which is consistent with data on hours worked for EU countries in 2018 according to OECD data.[28] The parameter weighting the disutility from labor, k, is then implied and equal to 2.5973. The time discount rate is set to $\rho = 0.03$, in line with the literature, where we find a value that varies between 0.01 and 0.05.

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 (0.07). The production parameter is set to $\beta = 0.65$, such that the share of labor, $1/(1+\beta)$, is calibrated to be 0.6, consistently with most literature. The implied mark-up in this model is $1/\beta = 1.5$. Empirical estimates for the mark-ups in the major EU economies range between 1.1 and 1.6. See e.g. Bundesbank (2017). 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

[27] Large scale macroeconomic models featuring endogenous innovation are actually used as laboratories to assist policy making, e.g. the QUEST III model, used by the European Commission. QUEST III is a dynamic general equilibrium model designed to match data and provide a direct link between policy interventions and market outcomes. Research and innovation mechanisms are also introduced in large scale computable general equilibrium models used for policy analysis, e.g. the RHOMOLO model designed for EU regions. These large-scale models are indeed continuously updated as improved theoretical foundations are introduced, new data become available and better estimates of the parameters are provided. For a recent attempt to identify those characteristics and properties of research and innovation that should ideally be included in macro-models to produce sound policy evaluations, see Annicchiarico et al. (2020).

[28] The fraction of time spent at work as been computed using the weighted average of worked hours using labor force statistics. See OECD (2020a) and OECD (2020b).

value of the success rate p at 0.014, so that given γ and from (39), the implied value for the quality-step λ respects the restriction.

Using Eurostat data for EU 28 for the period 2000-2018, 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. Note that GDP is equal to $Y-X=Y(1-\beta^2)$, where $\beta^2=0.42$ 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 follow. The marginal capital-income tax rate is fixed at a starting value of $\tau_r=0$, so as to better isolate the effects of introducing the tax on capital.

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 (41) and turns out to be 0.0994. The benchmark value for the labor-income tax rate corresponding to $\tau_r=0$ is computed from (32) and its starting value is found to be $\tau_w=0.6765$.

We now calculate the optimal tax structure. To do so, we simply need to maximize the expression for W in (49) 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, and the social planner solution. Comparing the optimal values in the fourth column with the ones implied by a zero tax on capital income, 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. The capital income tax rate associated with maximum utility is at 0.244, that is not far for the capital income tax rates observed in most advanced economies, although with some variability across countries.[29] Moreover, the welfare gain from adopting the optimal tax mix is equivalent to a permanent increase in consumption of 3.93 percent per year.[30] The socially optimal growth rate γ_s is 0.0378, that is about two times larger than the market one, while the corresponding labor input L_s is 0.3582. In the social planner solution the employment level is then much higher than in market equilibrium level with no tax on capital. This why the optimal tax policy must be designed to increase employment.

Another important result is that this optimal tax scheme delivers lower growth. The after-tax interest rate will decrease, so the rate of growth will decrease, as indeed we know from (44). This decrease in growth would reduce the "business stealing" and the "weight of government" negative externalities which, in principle, could make the rate of innovation excessive. However, we know from our analysis of the social planner solution that the market equilibrium, even without a tax on capital income (always growth decreasing in this model), generates an insufficient growth rate. This shows that an increase in social welfare can be consistent with a decrease in growth even when growth is compressed below an already inefficiently low value. This is interesting, because there exists a general shared view in the literature on taxation and growth that growth effects tend to prevail over level effects by compounding and that, therefore, when growth is too low it cannot be useful to reduce it further. However, what we find contradicts this consensus view.

We now explain in detail 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

[29]For instance, in the EU the average capita income tax rates estimated by McDaniel (2017) for the period 2000-2015 range from 0.17 of Germany to 0.26 of Belgium.

[30]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.

will push up the demand for the intermediates, of which too little is sold because of imperfect competition. This higher demand for the intermediates reduces the "market power" distortion and represents an efficiency gain. A tax shift from labor to capital income will also decrease the "weight of government" friction to some extent.

Finally, a lower tax on labor income could in principle also reduce the "incomplete appropriability" externality, even if the tax on capital income simultaneously increases. This is because higher employment will raise the total rents that can be captured by successful innovators (see equation 12) and therefore increase the incentives for innovating. However, in this setup, innovation tends to be more expensive in a large market. Actually, from (14) we observe that the probability of successful innovation declines with employment. The two effects cancel out in equilibrium, thus increasing the tax rate on capital income will always reduce growth and worsen the "incomplete appropriability" externality. The negative welfare effect of the growth reduction will be, however, moderated because the "business stealing" effect will decline with growth.

6.2 Sensitivity Analysis

To understand the role of the mechanisms described above in determining the value of the optimal capital income tax rate and verify the robustness of our findings, we perform a tight sensitivity analysis. We calculate how the optimal tax on capital-income changes when varying the parameters β, σ, ρ, η, ψ (normalized to β in the baseline case) and ζ leaving all the others unchanged. Of course this means the implied values for τ_w, γ and L will change accordingly. Finally, we consider different sizes of the public sector, by changing the amount of transfers and of public consumption. 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 values of τ_w, γ and L corresponding to a zero capital income tax rate.

To check the overall robustness of our results in Table 4 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 baseline values when $\tau_r = 0$.

We will now attempt to interpret intuitively the effects summarized in Tables 3 and 4, and represented in Figures 1-7, where each parameter changes at a time. In all figures the continuous lines show how the (Ramsey) optimal tax rates and the corresponding values for labor and growth change with each parameter, while the dotted lines indicate how the labor income tax rate, labor and growth, with no tax on capital, change with each parameter. In what follows "initial" will mean associated with a zero capital tax. Finally, the dashed lines represent how the socially optimal growth and labor change with each parameter.

6.2.1 Effect of β

Let's start by looking at the role of β.[31] Figure 1 shows that for a higher β the optimal tax scheme prescribes a higher ratio τ_r/τ_w. In fact, as already noted, from (46) we can see that for an increasing value of β, with no tax on capital, the level of labor (dotted line) declines, as the real wage is lower (see 25). Labor is therefore pushed farther away from its social optimum (dashed line), which, from (56), does not depend on β. On the other

[31] The value of ψ, normalized to β in the baseline case, stays unchanged in this exercise.

hand, from (44) and (57) we see that both the market and the socially optimal rates of growth decline, but the second declines more steeply so that the two get closer.[32] This is why when β is higher pushing up labor from its initial level becomes more important, even if this implies a lower growth rate.

The above results can be explained as follows. 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 "market power" distortion, are increasing in the size of the market. As reducing the tax on labor pushes labor up, from (46), the positive spillover on profits moves up up as well. This spillover from labor to profits is increasing in β, because the ratio between profits and labor income is actually β (see Appendix B).

6.2.2 Effect of σ

A higher σ calls for a higher capital tax rate. See Figure 2. Both the socially optimal and the market level of labor (with no tax on capital) decline when σ goes up, but the second declines proportionally more than the first. The opposite is true for the socially optimal and the market rate of growth (with no tax on capital income): the former declines proportionally more than the latter. Shifting the tax from labor to capital increases labor and current consumption. A higher capital tax rate lowers growth, but 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.

6.2.3 Effect of ρ

The optimal τ_r is increasing in the rate of time preference ρ as Figure 3 shows. From comparing the expressions for L and L_s (equations 46 and 56), we see that a high rate of time preference pushes the market labor down and farther away from the social optimum solution, which is unaffected by ρ. On the other hand, Figure 3 shows that for our parametrization a higher ρ reduces the distance between the market rate of growth with no tax on capital and the socially optimal value. A lower τ_w means higher labor under the optimal tax mix 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. Therefore, a higher time discounting means the optimal tax mix will be geared toward incentivizing labor supply.

6.2.4 Effect of η

A higher Frisch elasticity recommends a higher capital tax rate. This is because the advantage of pushing up the tax on capital and down the tax on labor is bigger the larger the increase in labor the tax program is able to induce. Actually, a major increase in labor will mean a larger increase in the demand size for intermediates and therefore a larger reduction in the "market power" distortion. In Figure 4 we show how a higher Frisch elasticity widens the gap between the initial (no capital income tax) and the social

[32]For the range of values considered the initial steady state growth rate is only very marginally affected by β.

optimum solution. A higher Frisch elasticity, leaving all other parameter unchanged, reduces the initial value of labor. Therefore, the higher is η the more important is to increase labor, reallocating a part of the tax burden to capital income.

6.2.5 Effect of ψ

We now consider the effects of ψ, which measures the marginal cost of producing intermediates in terms of units of final good. In the base model ψ was normalized to β, while in Appendix B the model is solved for $\psi \neq \beta$. In the appendix we show that the value of ψ is irrelevant per se and that what counts is the relation of this value to the value of ζ, which measures the productivity of research.[33] Considering the effect of changing ψ is then interesting because it shows how the optimal tax program changes when the ratio between the costs and benefits of research changes.

As is clear from Table 3 and Figure 5, a higher ψ leads to a higher optimal tax rate on capital income. A higher ψ reduces production and profits in the intermediate sector, and therefore production and the demand for labor in the final sector. A reduction in the innovation rate is also induced, as the costs of research necessary to invent better qualities of goods do not decrease, but the flow of profits from invention do decrease. As a result both the no capital tax market growth rate and labor decrease (dotted lines). The socially optimum solution for growth is also decreasing in ψ. Improving the quality of goods becomes less beneficial from a social point of view when the marginal cost of producing each unit of the newly improved good is higher. On the other hand, the socially optimal trade-off between leisure and consumption and, therefore, the optimal level of employment are not affected, because again income and substitution effects cancel out. By reducing taxes on labor income it is then possible to partially compensate for the negative effects of higher production costs in the intermediate goods sector by favoring employment.

6.2.6 Effect of ζ

A higher ζ means higher productivity of the R&D sector delivering more innovation and growth. As we can see from Figure 6 the effects are just the opposite of those observed for higher ψ. A more productive R&D sector means that it is worthwhile giving up more current consumption in exchange for higher future consumption, so that the optimal tax ratio τ_r/τ_w is decreasing in ζ.

6.2.7 Effects of T and G

Another interesting result of our sensitivity exercise is that a higher ratio of public outlays, whether due to higher transfers or higher public consumption, is associated with a higher optimal τ_r/τ_w. 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.

Figure 7 shows the effect only of higher public expenditure. It is interesting how an increase in public spending reduces growth in the social optimum (dashed line), so

[33] This is why Table 4 does not report the case of changing ψ. In the second sensitivity exercise, where the scale coefficients ζ and k simultaneously adjust to keep growth and labor constant, any change in ψ is offset by a change in ζ, so that results are the same as those obtained when setting ψ equal to β.

narrowing the gap with the market solution (dotted line), which does not depend on g. This is because the social planner internalizes the "weight of government" distortion. On the other hand, the gap between the first best solution and the market solution widens as regards labor (just by comparing 46 and 56). So in the attempt of pushing up labor the optimal tax mix prescribes an increase in the ratio τ_r/τ_w when the public sector expands.

7 Extensions

In this section we extend the model in five different ways. First, we conduct one further experiment introducing a "dilution effect", where an increase in the number of varieties lowers the effectiveness of innovation activity. Second, we study the effects on the optimal tax policy of government promoting economic activity. We will study, in turn, the case of a subsidy to R&D and the case of government services entering the production function of the final good sector. Fourth, we derive the optimal tax mix in the presence of outstanding public debt. Finally, the last extension of the model features two classes of agents so we can study redistributive issues.

7.1 Dilution Effect

We now introduce a "dilution effect" into the baseline model. 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:

$$Y_t = \frac{L_t^{1-\beta}}{\beta} \int_{\nu=0}^{N_t} q_t(\nu) x_t(\nu \mid q)^{\beta} \mathrm{d}\nu, \ \beta \in (0,1), \tag{58}$$

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

$$p_t(\nu \mid q) = \frac{\zeta z_t(\nu \mid q)}{N_t q_t(\nu)}, \ \zeta > 0. \tag{59}$$

The probability of successful innovation in a sector ν 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).[34] We also assume that N is equal to L at all times. In Appendix C 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. The social planner solution for

[34] 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.

L is now much higher than its market counterpart that is why the second best policy prescribes a higher capital income tax than in the base model, while the solution for the growth rate is the same. 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.

7.2 A Subsidy to R&D

To streamline our analysis, in the baseline case we have assumed no subsidies to R&D. The assumption is not indefensible, as verifying private R&D is often difficult for public authorities. Firms may use R&D subsidies to fund other expenses, while technological innovations may stem from entrepreneurial activities different from formal R&D.

We now relax the assumption and show that capital has to be taxed more when R&D is subsidized. We assume R&D is subsidized at the rate $S_{R\&D}$, so that investing z_t units of final good in R&D at time t will cost an entrepreneur $z_t(1 - S_{R\&D})$. Appendix D reports the derivations of the model under this assumption. We show that the economy will be on a balanced growth path at all times and obtain closed-form solutions for growth and labor.

We consider the effects of a 0.1 subsidy rate under two alternative assumptions: i) that the subsidies represent additional public expenses; or ii) that the subsidies are substitutes of other expenses, so that the government intervention in terms of total government expenditure as a percentage of GDP stays at the baseline level. In the first case financing the R&D subsidy requires the increase of distortionary taxation on impact, while it does not in the second case. Table 6 reports the results.

In Appendix D we show analytically that a subsidy will always increase growth, by making investing more profitable. Using the parametrization described in Table 1, a 0.1 R&D subsidy increases the initial growth rate by 0.33 p.p. in both case i) and case ii). The subsidy reduces the "incomplete appropriability" externality and increases the "creative destruction" externality, but as the first prevails the increase in growth is a welfare gain.

The effects of the subsidy on the initial value of labor input instead depend crucially on whether it adds to public spending or not. We observe that the initial value of labor decreases in the first case, consistently with previous findings that higher public expenses reduce market labor supply. In the second case we find analytically that the initial value of labor increases. Considering the optimal tax mix, we observe that in both cases the introduction of a subsidy to R&D increases the optimal ratio τ_r/τ_w and the related welfare gains. By mitigating the "incomplete appropriability" distortion the R&D subsidy allows a tax mix that pushes labor up, so reducing the deadweight loss from market power. The effect of the subsidy on the optimal ratio τ_r/τ_w is stronger when government consumption is not simultaneously reduced, consistently with our previous findings that an increase in the size of the public sector requires a relatively higher capital income tax rate at the optimum.

7.3 G in the Production Function

We now turn to the case where public expenditure $G = gY$ affects the marginal productivity of labor and of the intermediate goods in the economy. We modify the production function of the representative firm in the final good sector from (6) to:

$$Y_t = \frac{L_t^{1-\beta}}{\beta} \left(\int_0^1 q_t(\nu) x_t(\nu)^{\frac{\beta}{\alpha}} d\nu \right)^{\alpha} G_t^{1-\alpha},\ 0 < \beta \leq \alpha < 1. \tag{60}$$

The representative firm is atomistic and thus rationally ignores the impact of its choices on G that is taken as given. This implies that the above production function is still consistent with the hypothesis of constant returns to scale at the firm level. By analogy to the baseline case, we still assume that the probability of successful innovation in a sector increases with R&D expenses relative to the size of the sector, measured by profits. This assumption implies that the initial growth rate of the economy is the same as in the baseline case. The model incorporating the new assumption on G is solved in Appendix E. In the appendix we show that the economy will be on a balanced growth path at all times and obtain closed-form solutions for growth and labor, both for the decentralized and the centralized economy.

In simulations, we consider two values for the parameters α, 0.985 and 0.95, so that the labor income share, now $(1-\beta)/(1-\alpha^{-1}\beta^2)$, is still plausible. All the other parameters stay unchanged. Results are reported in Table 7, which shows that the optimal policy mix prescribes a higher capital income tax rate than in the baseline case. It can be seen analytically that market labor is increasing in α, while its socially optimal value is decreasing in α. So when $\alpha < 1$ to realign the two values a lower tax rate on labor income will be required, *coeteris paribus*, than when $\alpha = 1$. With productive public expenditure the ratio between the marginal productivity of labor at the social level and the wage goes up: this effect makes it optimal to induce a larger increase in employment through the redistribution of the tax burden from labor to capital income than in the baseline case.

7.4 Public Debt

In this section we characterize the optimal tax mix allowing for public debt. We assume that public debt bonds are tax exempt and that public debt is a fixed proportion of GDP. In particular, we set $B_t = \kappa Y_t$, where $\kappa > 0$ so that $\kappa/(1-\beta^2)$ is the public debt-to-GDP ratio $B/(Y-X)$. See Appendix F for the full description of the modified model.

Table 8 reports the optimal tax mix for a public debt-to-GDP ratio equal to 0.6 and 0.8. We set a positive initial value for capital income taxation, contrary to what we have done so far, since it can be shown that equilibrium determinacy requires τ_r to be higher than a certain threshold level, as long as $\kappa > 0$. Clearly, the effects of debt on the optimal ratio τ_r/τ_w are positive and quantitatively important. Ours results then show that the government's ability to borrow tends to reinforce our conclusions. This confirms findings in Straub and Werning (2020) that even in the basic Ramsey model there is always a level of initial debt high enough that the tax rate on capital will stay at its upper bound for ever.

7.5 Two Classes of Agents

So far we have shown how taxing capital income may represent an efficiency gain, setting aside equality issues. However, there is of course a strict connection between the functional and the personal distribution of income because the marked skewness of wealth distribution makes the equalizing potential of the taxation of capital very important. Piketty et al. (2018) observe that the fall of corporate and estate taxes has lead to a decline in the progressivity of the US tax system in the last decades. In particular, Saez and Zucman (2020) show that the demise of the federal corporate tax in 2018 led to the regressivity of the US tax system at the very top. A simple extension of our model, fully described in Appendix G, can be used to study how the optimal tax rates depend on how egalitarian the social welfare function is.

As in the seminal Judd (1985) article we consider two classes of agents, with wealth concentrated in the hands of only one of the two classes. In a first experiment we assume that the first class of agents offer high skilled labor, while lower skilled labor is offered by hand to mouth agents. Labor income can be taxed at different rates for the two classes. We then conduct a second experiment, going back to the original assumption in Judd (1985) that labor is homogenous and only offered by hand-to-mouth agents, i.e. that wealth owners do not work.

Symbols keep the meaning they have in the base model with variables indexed by 1 if related to capitalists and by 2 if related to workers.

The first class of agents choose consumption and labor to maximize the following functional:

$$U_{1,t} = \int_{t=0}^{\infty} e^{-\rho t} \frac{C_{1,t}^{1-\sigma}\left(1 - k_1(1-\sigma)H_{1t}^{1+\frac{1}{\eta}}\right)^{\sigma} - 1}{1-\sigma} \mathrm{d}t, \tag{61}$$

subject to the instantaneous budget constraint:

$$\dot{A}_t = r_t(1-\tau_r)A_t - C_{1,t} + T_{1,t} + w_{1,t}(1-\tau_{w,1})H_{1t}, \tag{62}$$

where A_t is individual wealth.

The second class of agents maximize:

$$U_{2,t} = \int_{t=0}^{\infty} e^{-\rho t} \frac{C_{2,t}^{1-\sigma}\left(1 - k_2(1-\sigma), H_{2,t}^{1+\frac{1}{\eta}}\right)^{\sigma} - 1}{1-\sigma} \mathrm{d}t \tag{63}$$

subject to the instantaneous budget constraint:

$$C_{2,t} = w_{2,t}(1-\tau_{w,2})H_{2,t} + T_{2,t}, \tag{64}$$

Transfers are assumed for both classes to be proportional to their respective labor income.

In the production function of the final goods sector (6) we now have:

$$L_t \equiv L_{1,t}^{\alpha} L_{2,t}^{1-\alpha}, \tag{65}$$

where $L_{1,t}$ is high skilled labor, $L_{2,t}$ is low skilled labor, α is a real number in $[0,1]$, $w_{1,t}$ equals the marginal productivity of the former and $w_{2,t}$ the marginal productivity of the latter.

The Ramsey planner maximizes the following social welfare function:

$$\xi\chi U_{1,t} + (1-\chi)\, U_{2,t}. \tag{66}$$

where ξ is the weight on the first class of agents, who are a fraction χ of the population:

By varying the weight ξ in the interval $[0,1]$ one can trace out points on the constrained Pareto frontier and characterize their associated policies. When $\xi = 0$ we have a Rawlsian social welfare function and when $\xi = 1$ an utilitarian one.

Tables 9 and 10 report results under the alternative assumptions that the two classes provide labor and that only one class does.

Under the first assumption the lower skilled workers are the poorest half of the population (i.e. $\chi = 0.5$), and we calibrate α so that their share of national income is 0.2, which is close to the current average in European economies.[35] The optimal capital income rate is 0.27 when $\xi = 1$, higher than in a representative agent framework, and it becomes considerably higher at 0.38 when $\xi = 0$, in which case the lower skilled should not be taxed at all. These results suggest an important role of the capital tax in achieving a more equitable income distribution, even when labor income is taxed progressively.

We then move to the assumption that labor is homogenous and that capitalists do not work, i.e. we set $\alpha = 0$ and $k_1 = 0$. The optimal tax on capital income is then decreasing in ξ and in χ, for which we consider the two values 0.1 and 0.01. For all but one of the four cases we report on the tax rate is much higher, at 0.40 or more, than in a representative agent framework: however, when $\chi = 0.1$ and $\xi = 1$ it decreases to 0.24.

These results show the importance of redistributive purposes in defining the optimal tax on capital, as the incidence of the tax can be only partially shifted to labor.[36]

In our model redistribution leads to lower inequality and lower growth. However, there is scant support in the data for a negative relationship between these two variables and indeed a large literature has developed on why inequality may be itself bad for growth.[37] One important channel, likely to be especially relevant for advanced countries, is that, as predicted by classical Mancur Olson's collective action analysis, concentrated special interests organize to protect their rents, while diffuse majority interests are trumped. To include this kind of rent seeking in our analysis we focus on the cost of entry of new firms in a market, which cost is an institutional as well as a technical variable. Gutiérrez and Philippon (2019) observe a decline in the elasticity of entry with respect to Tobin's Q since the late 1990s in many industrial sectors in the US and explain this decline with antitrust authorities capture and regulations influenced by lobbying.[38] Finally, there seems to be a correlation between inequality (in particular the top 1% income shares) and lobbying intensity across US states, as documented by Aghion et al. (2021). In the light of this evidence we incorporate in the model with pure capitalists the assumption that the investment in R&D needed for entry is increasing in the proportion of net of tax income going to capitalists $\beta(1-\tau_r)/(1+\beta)$. In particular ζ in (14) is no longer a fixed parameter, but is instead the following increasing function of τ_r:

$$\zeta = \overline{\zeta}(1+\tau_r)^{\alpha_\zeta}, \overline{\zeta} > 0, \alpha_\zeta \geq 0. \tag{67}$$

The results of this further exercise are shown in the last column of Table 10. From

[35] In fact this share was 18.7 percent in Europe in 2019 in the World Inequality Database.

[36] Straub and Werning (2020) show that in the Judd model (with labor supply fixed and time varying taxes) taxes on capital will be positive in the long run, when the IES is less than 1. This is because an anticipated increase in taxes leads to higher savings today and if the tax increase is sufficiently far off in the future, then the increased savings generate a higher capital stock over a lengthy transition. Clearly, the mechanism is completely different in our analysis.

[37] For recent evidence see Berg et al. (2018), who also offer an overview of the theoretical literature on inequality and growth.

[38] For more evidence on the link between political power and the corporate sector, see Bertrand et al. (2020) for the US and Akcigit et al. (2018) for Italy.

(44) there are now two effects of opposite sign of τ_r on γ. On the one hand, the tax discourages saving but, on the other hand, it reduces rent seeking by capitalists, so that innovation/entry is less costly. When the second effect prevails, as τ_r goes up growth goes up as well. This happens in our calibrated example, for instance, when $\alpha_\zeta = 2$.

8 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 have 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 the production levels are too low; 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' decisions. 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 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 an extensive sensitivity analysis to gauge 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. When the model is extended, so that employment leads to an increase in the variety of available goods or when the government promotes economic activity, the optimal tax on capital will be much higher. We also consider an extension of the model to two classes of agents, capitalists and workers who do not save as in Judd (1985). Findings highlight the redistributional potential of capital income taxation.

Our results have important economic and institutional implications. First, given the increasing share of capital income observed in the last decades, our findings would suggest, on purely efficiency grounds, the need of a simultaneous increase of taxation on capital income. Second, a larger public sector requires a higher optimal tax rate on capital. This result is particularly relevant if we consider that in advanced economies public expenditure represents a sizable share of gross national product, and this share may further increase in the face of current and future challenges, such as an aging population ad health emergencies, as the Covid pandemic has dramatically shown. Finally, our findings offer a different perspective on the debate on tax competition versus tax coordination, warning against the risks of a "race-to-the-bottom" over capital taxes that may arise for high degree of international capital mobility.

The premise of our analysis is that it is efficient to increase the use of non rival goods (i.e. a blueprints), as their average cost per user is declining. A higher level of employment and economic activity allows this wider use: we have shown that shifting the tax burden from labor to capital, even at the cost of reducing the rate of innovation, may then make society better off. In future research, we plan to further explore the consequences of this premise for taxation policy. In particular, we intend to study the trade-off between employment and growth when distributional aspects are more fully considered.

A first avenue would be to study factor income taxation in a model in which skilled workers are employed in the research sector, while unskilled workers are employed in manufacturing, i.e. to assume that R&D is "knowledge driven", as was the case in the original horizontal innovation model by Romer (1990)

Another avenue we plan to follow is to explore optimal direct taxation in a model of directed technical change. If a progressive tax reform tends to increase labor supply this may induce firms to develop technologies that are more complementary to labor. These directed technical change effects make the optimal tax scheme more progressive. Again it would also be possible to enrich the analysis by considering skilled and unskilled labor.

We also plan to investigate the reproduction of inequality through the composition of effective demand. With high inequality, there will be more demand for luxury goods. If these goods are more capital-intensive and skilled-labor-intensive in production, inequality could feed upon itself.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, 108(11):3450–91.

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.

Aghion, P., Akcigit, U., and Howitt, P. (2014). What do we learn from Schumpeterian growth theory? In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2, pages 515–563. Elsevier.

Aghion, P., Antonin, C., and Bunel, S. (2021). *The Power of Creative Destruction: Economic Upheaval and the Wealth of Nations*. Harvard University Press.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–351.

Akcigit, U. and Ates, S. T. (2021). Ten facts on declining business dynamism and lessons from endogenous growth theory. *American Economic Journal: Macroeconomics*, 13(1):257–98.

Akcigit, U., Baslandze, S., and Lotti, F. (2018). Connecting to power: political connections, innovation, and firm dynamics. *National Bureau of Economic Research*, Working Paper No. 25136.

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.

Annicchiarico, B., Licandro, O., Mohnen, P., and Veugelers, R. (2020). Moving the frontier of macroeconomic modelling of research and innovation policy. Independent Expert Report ISBN:978-92 7620910-2, doi:10.2777/34199, EC, DG R&I, Publications Office of the European Union, Luxembourg.

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. (2020). How should capital be taxed? *Journal of Economic Surveys*, 34(4):812–846.

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.

Berg, A., Ostry, J. D., Tsangarides, C. G., and Yakhshilikov, Y. (2018). Redistribution, inequality, and growth: new evidence. *Journal of Economic Growth*, 23(3):259–305.

Bertrand, M., Bombardini, M., Fisman, R., and Trebbi, F. (2020). Tax-exempt lobbying: Corporate philanthropy as a tool for political influence. *American Economic Review*, 110(7):2065–2102.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–44.

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.

Bundesbank, D. (2017). Mark-ups of firms in selected european countries. *Monthly Report*, 53.

Capinski, M. and Kopp, P. E. (2004). *Measure, integral and probability*. Springer Science & Business Media, 2nd edition.

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.

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.

Djankov, S., Ganser, T., McLiesh, C., Ramalho, R., and Shleifer, A. (2010). The effect of corporate taxes on investment and entrepreneurship. *American Economic Journal: Macroeconomics*, 2(3):31–64.

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

Foster, L., Haltiwanger, J. C., and Krizan, C. J. (2001). Aggregate productivity growth: lessons from microeconomic evidence. In Dean, E., Harper, M., and Hulten, C., editors, *New developments in productivity analysis*, pages 303–372. University of Chicago Press.

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.

Gutiérrez, G. and Philippon, T. (2019). The failure of free entry. *National Bureau of Economic Research*, Working Paper No. 26001.

Ha, J. and Howitt, P. (2007). Accounting for trends in productivity and R&D: A schumpeterian critique of semi-endogenous growth theory. *Journal of Money, Credit and Banking*, 39(4):733–774.

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.

Jalles, J. (2019). Wagner's law and governments' functions: Granularity matters. *Journal of Economic Studies*.

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.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, 132(2):665–712.

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

Madsen, J. B. and Timol, I. (2011). Long-run convergence in manufacturing and innovation-based models. *Review of Economics and Statistics*, 93(4):1155–1171.

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.

McDaniel, C. (2017). Tax series updated february 2017. Technical Report https://www.caramcdaniel.com/researchpapers (Accessed on 30 September 2020).

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

Mohnen, P. (2019). R&D, innovation and productivity. In *The Palgrave Handbook of Economic Performance Analysis*, pages 97–122. Springer.

OECD (2020a). Hours worked (indicator. Technical Report doi: 10.1787/47be1c78-en (Accessed on 30 September 2020).

OECD (2020b). Labour force (indicator. Technical Report doi: 10.1787/ef2e7159-en (Accessed on 30 September 2020).

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. (2010). Commentary on taxation of wealth and wealth transfer. In Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., and Poterba, J. M., editors, *Dimension of Tax Design - The Mirrlees Review*, pages 825–831. Oxford University Press, Oxford.

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

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics*, 133(2):553–609.

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.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2):S71–S102.

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

Saez, E. and Zucman, G. (2020). The rise of income and wealth inequality in america: Evidence from distributional macroeconomic accounts. *Journal of Economic Perspectives*, 34(4):3–26.

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.

Uhlig, H. (1996). A law of large numbers for large economies. *Economic Theory*, 8(1):41–50.

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	0.18
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	2.5973
r	Interest rate	0.0680
λ	Quality-step	2.3571
ζ	Parameter in R&D technology	0.0994

Table 2: Optimal Taxation

Variable	Social Optimum	Initial Equilibrium	Optimal Taxation
τ_r	-	0	0.2440
τ_w	-	0.6765	0.5357
γ	0.0378	0.019	0.0125
L	0.3582	0.18	0.2098
Welfare Gain	-	-	0.0393

Note: the table reports the optimal tax mix, growth and labor for the baseline model, along with the corresponding social optimum solution and the initial equilibrium values at zero tax rate on capital income.

Table 3: Sensitivity Analysis I: One Parameter Changes at a Time

	γ_S	L_S	τ_r	τ_w	γ	L	Welfare Gain
Baseline	0.0378	0.3582	0.2440 (0)	0.5357 (0.6765)	0.0125 (0.0190)	0.2098 (0.18)	0.0393
$\beta = 0.64$	0.0391	0.3582	0.2180 (0)	0.5348 (0.6577)	0.0133 (0.0190)	0.2108 (0.1853)	0.0290
$\beta = 0.66$	0.0365	0.3699	0.2697 (0)	0.5372 (0.6964)	0.0118 (0.0190)	0.2087 (0.1745)	0.0522
$\sigma = 1.8$	0.0420	0.3708	0.195 (0)	0.5669 (0.6765)	0.015 (0.0205)	0.2152 (0.1909)	0.0252
$\sigma = 2.2$	0.0344	0.3469	0.2844 (0)	0.5091 (0.6765)	0.0105 (0.0177)	0.2047 (0.1707)	0.0531
$\rho = 0.02$	0.0428	0.3582	0.1733 (0)	0.5824 (0.6765)	0.0183 (0.0227)	0.2060 (0.1849)	0.0206
$\rho = 0.04$	0.0328	0.3582	0.3283 (0)	0.481 (0.6765)	0.0054 (0.0153)	0.2131 (0.1755)	0.0605
$\eta = 0.8$	0.0378	0.3749	0.2117 (0)	0.5555 (0.6765)	0.0134 (0.0190)	0.2318 (0.2058)	0.0277
$\eta = 1.3$	0.0378	0.3444	0.2786 (0)	0.5141 (0.6765)	0.0115 (0.0190)	0.1890 (0.1552)	0.0551
$\psi = 0.6$	0.0463	0.3582	0.2157 (0)	0.5551 (0.6765)	0.0173 (0.0238)	0.2083 (0.1820)	0.0312
$\psi = 0.7$	0.0310	0.3582	0.2727 (0)	0.5150 (0.6765)	0.0087 (0.0151)	0.2113 (0.1780)	0.0484
$\zeta = 0.09$	0.0328	0.3582	0.2645 (0)	0.5211 (0.6765)	0.0097 (0.0162)	0.2109 (0.1785)	0.0457
$\zeta = 0.12$	0.0488	0.3582	0.20861 (0)	0.5598 (0.6765)	0.0187 (0.0252)	0.2079 (0.1825)	0.0292
$G/(Y-X) = 0.15$	0.0458	0.3582	0.1348 (0)	0.5021 (0.5775)	0.0155 (0.0190)	0.2122 (0.1982)	0.0089
$G/(Y-X) = 0.25$	0.0325	0.3582	0.3132 (0)	0.5581 (0.7425)	0.0105 (0.0190)	0.2081 (0.1649)	0.0818
$T/(Y-X) = 0.15$	0.0378	0.3582	0.1567 (0)	0.5058 (0.594)	0.0149 (0.0190)	0.2210 (0.2038)	0.0126
$T/(Y-X) = 0.25$	0.0378	0.3582	0.3282 (0)	0.565 (0.759)	0.0101 (0.0190)	0.1991 (0.1537)	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 initial equilibrium values at zero tax rate on capital income are reported in parentheses.

Table 4: Sensitivity Analysis II: k and ζ Vary

	γ_S	L_S	τ_r	τ_w	γ	L	Welfare Gain
Baseline	0.0378	0.3582	0.2440 (0)	0.5357 (0.6765)	0.0125 (0.0190)	0.2098 (0.18)	0.0393
$\beta = 0.64$	0.0391	0.3481	0.2181 (0)	0.5347 (0.6577)	0.0132 (0.0190)	0.2049 (0.18)	0.0291
$\beta = 0.66$	0.0365	0.3699	0.2699 (0)	0.5371 (0.6964)	0.0118 (0.0190)	0.2155 (0.18)	0.0522
$\sigma = 1.8$	0.0393	0.3511	0.2062 (0)	0.5594 (0.6765)	0.0134 (0.0190)	0.2044 (0.18)	0.0281
$\sigma = 2.2$	0.0366	0.3646	0.2766 (0)	0.5149 (0.6765)	0.0117 (0.0190)	0.2148 (0.18)	0.0504
$\rho = 0.02$	0.0364	0.3514	0.1938 (0)	0.5695 (0.6765)	0.0147 (0.0190)	0.2032 (0.18)	0.0255
$\rho = 0.04$	0.0393	0.3635	0.2805 (0)	0.5149 (0.6765)	0.0104 (0.0190)	0.2148 (0.18)	0.0510
$\eta = 0.8$	0.0378	0.3279	0.2117 (0)	0.5555 (0.6765)	0.0134 (0.0190)	0.2027 (0.18)	0.0277
$\eta = 1.3$	0.0378	0.3994	0.2786 (0)	0.5141 (0.6765)	0.0115 (0.0190)	0.2192 (0.18)	0.0551
$G/(Y-X) = 0.15$	0.0458	0.3254	0.1348 (0)	0.5021 (0.5775)	0.0155 (0.0190)	0.1927 (0.18)	0.0089
$G/(Y-X) = 0.25$	0.0325	0.3911	0.3132 (0)	0.5581 (0.7425)	0.0105 (0.0190)	0.2271 (0.18)	0.0818
$T/(Y-X) = 0.15$	0.0378	0.3163	0.1567 (0)	0.5058 (0.594)	0.0149 (0.0190)	0.1951 (0.18)	0.0126
$T/(Y-X) = 0.25$	0.0378	0.4195	0.3282 (0)	0.565 (0.759)	0.0101 (0.0190)	0.2331 (0.18)	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 initial equilibrium values at zero tax rate on capital income are reported in parentheses.

Table 5: Dilution Effect

Variable	**Social Optimum**	**Initial Equilibrium**	**Optimal Taxation**
τ_r	-	0	0.5297
τ_w	-	0.6765	0.3435
γ	0.0378	0.019	0.0036
L	0.6205	0.18	0.2810
Welfare Gain	-	-	0.3131

Note: the table reports the optimal tax mix, growth and labor for the model with "dilution effect", along with the corresponding social optimum solution and the initial equilibrium values at zero tax rate on capital income.

Table 6: Subsidy to R&D

Variable	Initial Equilibrium	Optimal Taxation
Subsidy as Additional Public Expenses		
τ_r	0	0.2854
τ_w	0.7072	0.5310
γ	0.0223	0.0139
L	0.1747	0.2130
Welfare Gain	-	0.0644
Subsidy as Substitute of Other Public Expenses		
τ_r	0	0.2550
τ_w	0.6765	0.5204
γ	0.0223	0.0149
L	0.1845	0.2138
Welfare Gain	-	0.0666

Note: the table reports the optimal tax mix, growth and labor in the presence of a 0.1 subsidy to R&D expenses, along with the corresponding initial equilibrium values at zero tax rate on capital income.

Table 7: G in the Production Function

Variable	Social Optimum	Initial Equilibrium	Optimal Taxation
$\alpha = 0.985$			
τ_r	-	0	0.2524
τ_w	-	0.6765	0.5347
γ	0.0389	0.0190	0.0123
L	0.3689	0.1810	0.2111
Welfare Gain	-	-	0.0419
$\alpha = 0.95$			
τ_r	-	0	0.2736
τ_w	-	0.6765	0.5329
γ	0.0419	0.0190	0.0117
L	0.3982	0.1837	0.2144
Welfare Gain	-	-	0.0487

Note: the table reports the optimal tax mix, growth and labor when public spending enters the production function, along with the corresponding social optimum solution and the initial equilibrium values at zero tax rate on capital income.

Table 8: Public Debt

Variable	**Initial Equilibrium**	**Optimal Taxation**
Debt Ratio $\kappa/(1-\beta^2) = 0.6$		
τ_r	0.0800	0.3003
τ_w	0.6789	0.5409
γ	0.0170	0.0109
L	0.1767	0.2062
Welfare Gain	-	0.0383
Debt Ratio $\kappa/(1-\beta^2) = 0.8$		
τ_r	0.0800	0.3179
τ_w	0.6944	0.5424
γ	0.0170	0.0104
L	0.1721	0.2052
Welfare Gain	-	0.0477

Note: the table reports the optimal tax mix, growth and labor for different debt-to-GDP ratios, along with the corresponding initial equilibrium values at a tax rate on capital income that ensures determinacy of the equilibrium.

Table 9: Two Classes: High and Low Skilled Workers

Share of Workers $1-\chi=0.5$		
Variable	**Initial Equilibrium**	**Optimal Taxation**
$\xi=0$		
τ_r	0	0.3880
$\tau_{w,1}$	0.6765	0.6611
$\tau_{w,2}$	0.6765	0
γ	0.019	0.0082
H_1	0.18	0.1988
H_2	0.18	0.2667
Unskilled/Skilled Net Income Ratio	0.1412	0.5925
Welfare Gain	-	0.2460
$\xi=1$		
τ_r	0	0.2760
$\tau_{w,1}$	0.6765	0.5699
$\tau_{w,2}$	0.6765	0.4060
γ	0.019	0.0116
H_1	0.18	0.2085
H_2	0.18	0.2261
Unskilled/Skilled Net Income Ratio	0.1412	0.2889
Welfare Gain	-	0.0573

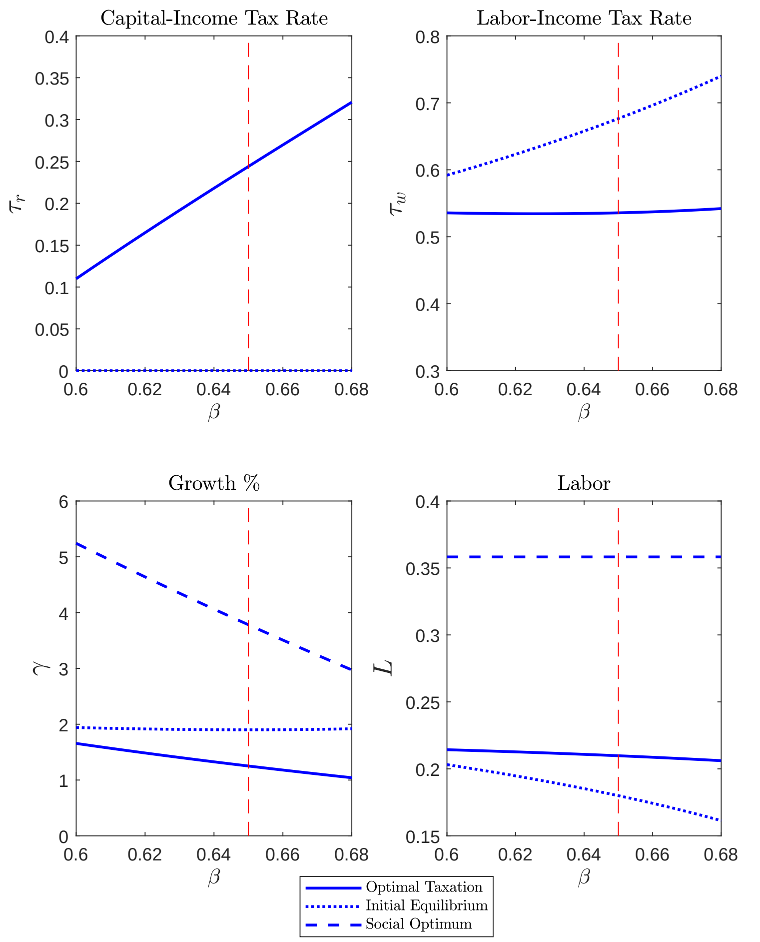
Note: the table reports the optimal tax mix, growth and individual labor supplies H_1 and H_2 for different assumptions regarding social welfare (Rawlsian v. utilitarian). k_1 and k_2 are set so that at the initial equilibrium $H_1 = H_2$=0.18.

Table 10: Two Classes: Capitalists and Workers

Variable	Initial Equilibrium	Optimal Taxation	Optimal Taxation with ζ endogenous
Share of Workers $1-\chi = 0.99$			
Variable	**Initial Equilibrium**	**Optimal Taxation**	**Optimal Taxation with ζ endogenous**
$\xi = 0$			
τ_r	0	0.4191	0.6066
τ_w	0.6765	0.4219	0.3142
γ	0.019	0.0073	0.0233
H	0.18	0.2084	0.2158
Workers/Capitalists Net Income Ratio	0.0061	0.0165	0.0295
Welfare Gain	-	0.1265	0.8232
$\xi = 1$			
τ_r	0	0.3993	0.5862
τ_w	0.6765	0.4354	0.3282
γ	0.019	0.0079	0.0240
H	0.18	0.2073	0.2150
Workers/Capitalists Net Income Ratio	0.0061	0.0157	0.0276
Welfare Gain	-	0.1170	0.7716
Share of Workers $1-\chi = 0.90$			
Variable	**Initial Equilibrium**	**Optimal Taxation**	**Optimal Taxation with ζ endogenous**
$\xi = 0$			
τ_r		0.4060	0.6041
τ_w	0.6465	0.4008	0.2858
γ	0.019	0.0077	0.0234
H	0.1894	0.2141	0.2229
Workers/Capitalists Net Income Ratio	0.0667	0.1852	0.3358
Welfare Gain	-	0.1109	0.7969
$\xi = 1$			
τ_r		0.2406	0.4608
τ_w	0.6465	0.5078	0.3808
γ	0.019	0.0126	0.0268
H	0.1894	0.2054	0.2167
Workers/Capitalists Net Income Ratio	0.0667	0.1250	0.2213
Welfare Gain	-	0.0442	0.4503

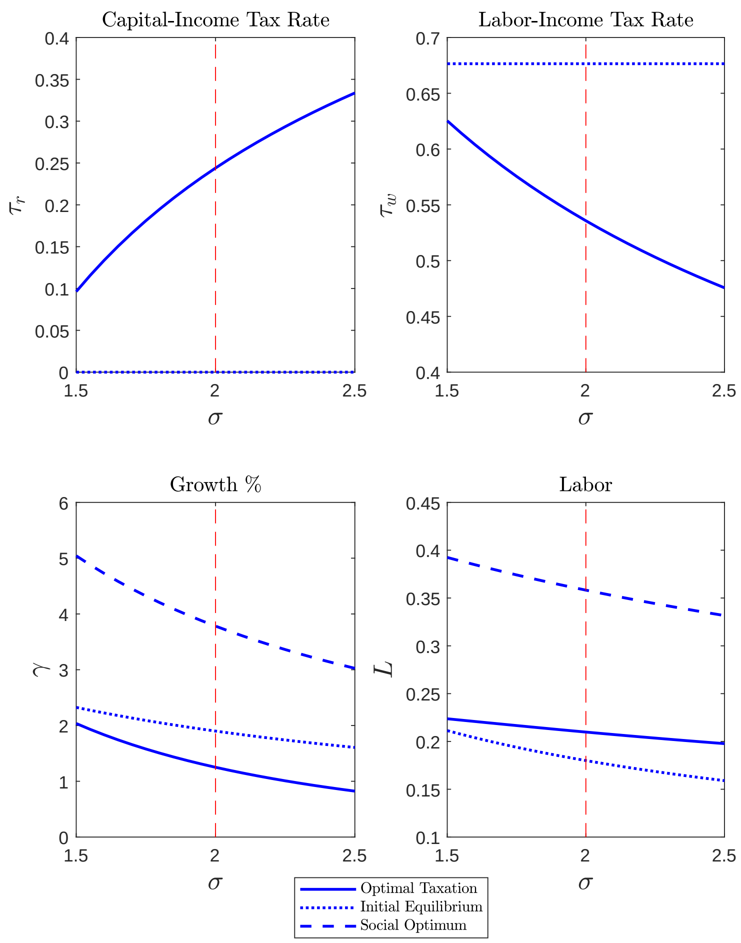
Note: the table reports the optimal tax mix, growth and individual labor supply H for different shares of workers $(1-\chi)$ and for different assumptions regarding social welfare (Rawlsian v. utilitarian). When $1-\chi = 0.99$ k is set so that at the initial equilibrium H is 0.18. The last column shows the results assuming that ζ increases in τ_r.

Figure 1: Optimal Taxation and the Production Parameter β



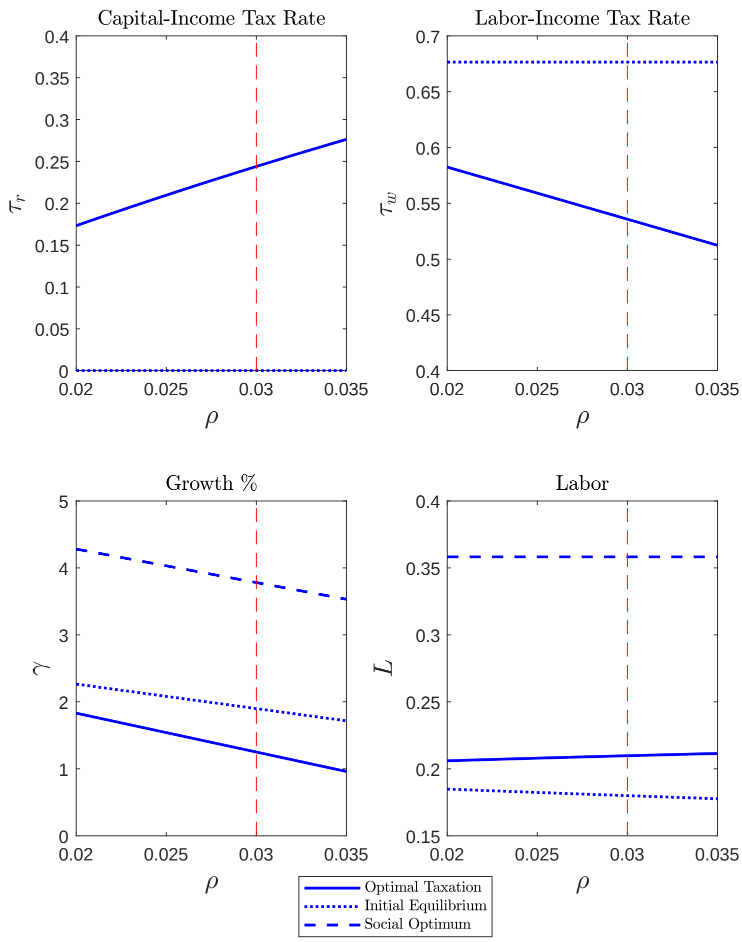
Note: continuous lines plot the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of β, a production parameter. Vertical dashed lines represent the baseline value of β. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with β, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Figure 2: Optimal Taxation and the Inverse of the Intertemporal Elasticity of Substitution σ



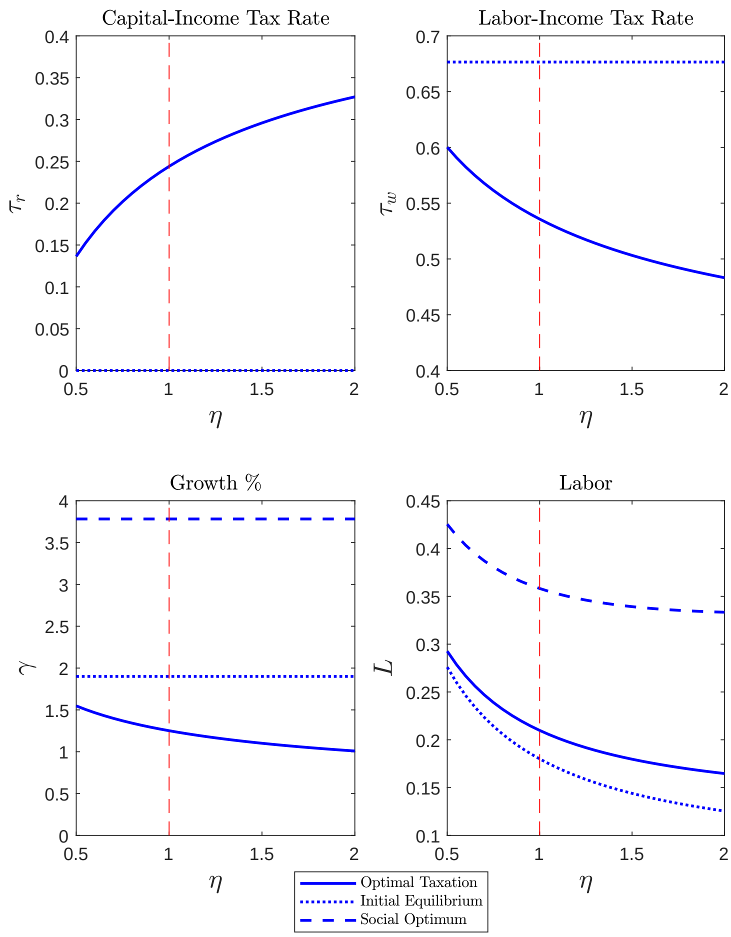
Note: continuous lines plot the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of σ, the inverse of the intertemporal elasticity of substitution. Vertical dashed lines represent the baseline value of σ. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with σ, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Figure 3: Optimal Taxation and the Rate of Time Preference ρ



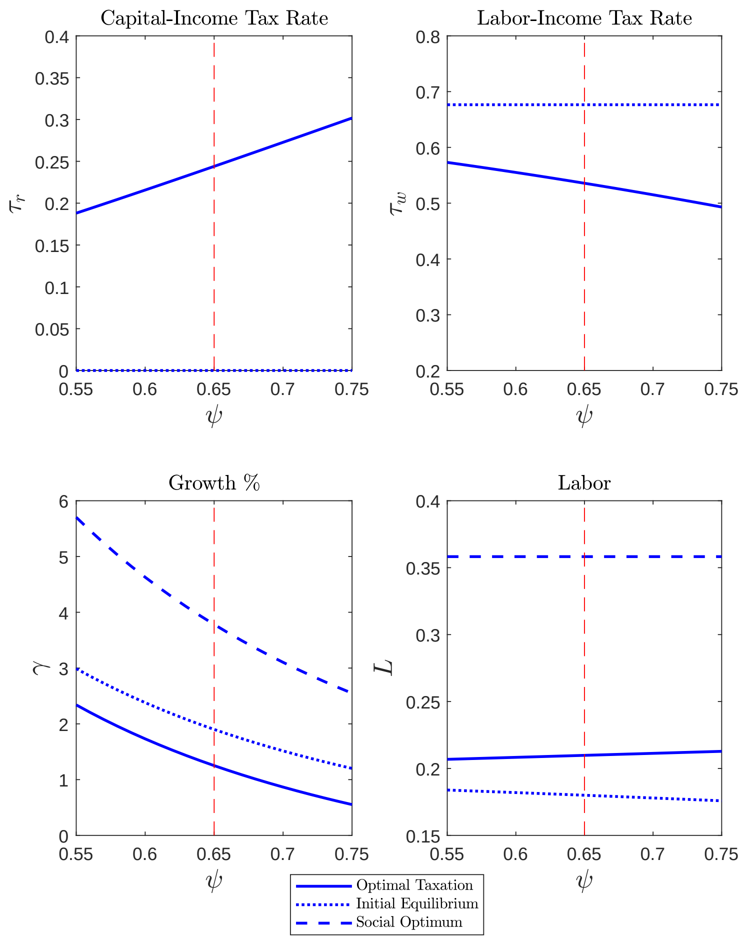
Note: continuous lines plot the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of ρ, the rate of time preference. Vertical dashed lines represent the baseline value of ρ. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with ρ, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Figure 4: Optimal Taxation and the Frisch Elasticity η



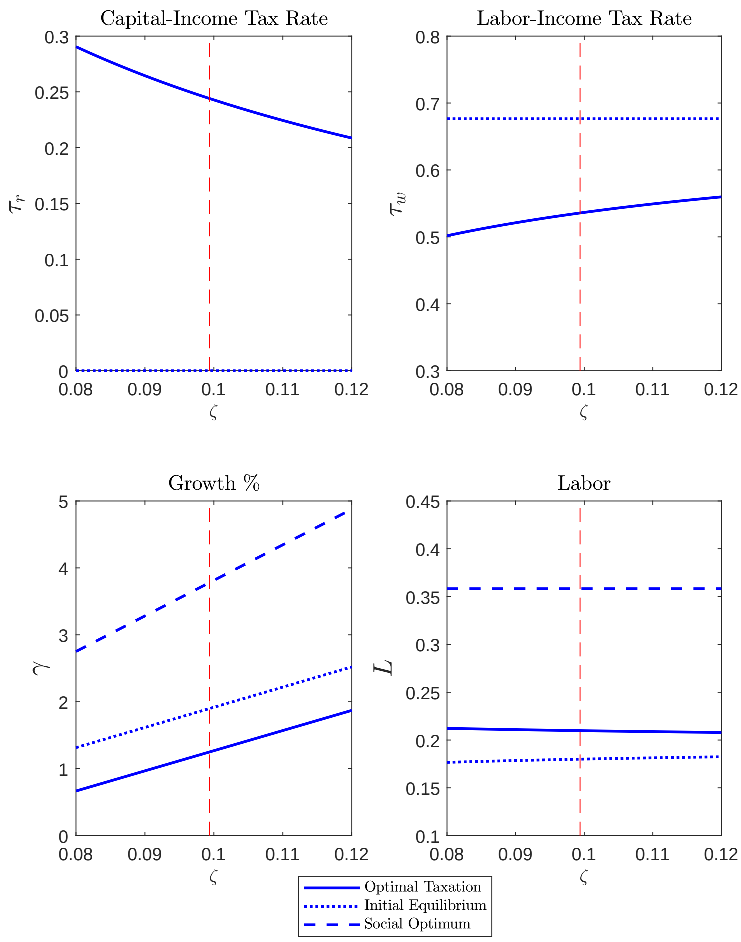
Note: continuous lines plot the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of η, the Frisch elasticity. Vertical dashed lines represent the baseline value of η. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with η, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Figure 5: Optimal Taxation and the Marginal Cost of Producing Intermediate Goods ψ



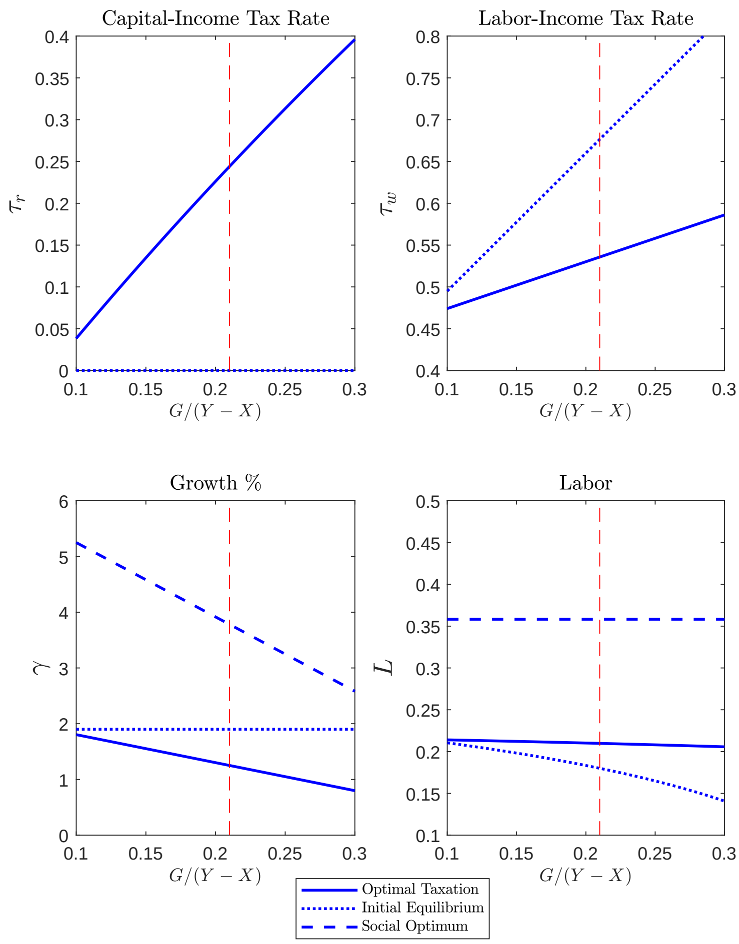
Note: continuous lines plot the optimal tax rates on capital income and the implied values for the tax rate on labor income, growth and labor for different values of ψ, the marginal cost of producing intermediate goods. Vertical dashed lines represent the baseline value of ψ. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with ψ, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Figure 6: Optimal Taxation and Productivity in the R&D Sector ζ



Note: continuous lines plot 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 ζ the productivity in the R&D sector. Vertical dashed lines represent the baseline value of ζ. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with ζ, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Figure 7: Optimal Taxation and Government Expenditure



Note: continuous lines plot 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. All the other parameters stay unchanged. Vertical dashed lines represent the baseline value of $G/(Y-X)$. Dotted lines show how the initial equilibrium values, computed at zero tax rate on capital income, change with $G/(Y-X)$, while dashed lines for growth and labor show the social planner's solution. All the other parameters stay unchanged.

Appendix A

A.1 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:

$$u = \frac{C^{1-\sigma}\left(1 + k(\sigma - 1)(1 - l)^{1+\frac{1}{\eta}}\right)^{\sigma} - 1}{1 - \sigma} \tag{A1}$$

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

Given (A1) 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 $\Omega \equiv 1 + k(\sigma - 1)(1 - l)^{1+\frac{1}{\eta}}$, so that under the assumption that $\Omega > 0$ we have:

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

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

A.2 Households Maximization Problem and Frisch Elasticity

Let $\mathcal{H}$ denote the current-value Hamiltonian, then the first-order-conditions for the consumer are the following:

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

$(ii)\ \frac{\partial \mathcal{H}_t}{\partial H_t} = 0 \Rightarrow k\sigma\frac{1+\eta}{\eta}C_t^{1-\sigma}\left(1 - k(1 - \sigma)H_t^{1+\frac{1}{\eta}}\right)^{\sigma-1}H_t^{\frac{1}{\eta}} = \mu_t w_t(1 - \tau_w);$

$(iii)\ \frac{\partial \mathcal{H}_t}{\partial F_t} = \rho\mu_t - \dot{\mu}_t \Rightarrow \mu_t r(1 - \tau_r) = \rho\mu_t - \dot{\mu}_t;$
$(i\nu)\ \lim_{t\to\infty} e^{-\rho t}\mu_t F_t = 0.$

Since $\left(1 - k(1 - \sigma)H_t^{1+\frac{1}{\eta}}\right)$ is bounded, we can write (iv) as (4) in the main text.
From (i), obtain

$$C_t^{1-\sigma} = \left[\left(1 - k(1 - \sigma)H_t^{1+\frac{1}{\eta}}\right)\mu_t^{-1/\sigma}\right]^{1-\sigma}, \tag{A2}$$

and use it into (ii),

$$k\sigma\frac{1 + \eta}{\eta}H_t^{\frac{1}{\eta}} = w_t(1 - \tau_w)\mu_t^{1/\sigma}, \tag{A3}$$

from which we can compute the Frisch elasticity:

$$\epsilon_F = \frac{dH_t}{dw_t}\frac{w_t}{H_t} = \frac{(1 - \tau_w)\mu_t^{1/\sigma}}{k\sigma\frac{1}{\eta}\frac{1+\eta}{\eta}H_t^{\frac{1}{\eta}-1}}\frac{w_t}{H_t} = \eta. \tag{A4}$$

A.3 Average Quality

In what follows we prove that the average quality $Q(t)$ in (21) is well defined.

When writing $Q(t) = \int_{\nu=0}^{1} q(t,\nu)d\nu$, since on the right hand side we have random variables, we first have to define what the integral means, as there is no universally accepted definition.

Following Uhlig (1996) we consider the partition s of the interval $[0,1] = (\nu_0, \nu_{1,...,} \nu_n)$, with $\nu_0 = 0$ and $\nu_n = 1$, while for $j = 1,2,...n-1$ ν_j $\epsilon(0,1)$, $\nu_j > \nu_{j-1}$, with mesh $m(s) = max\ (\nu_j - \nu_{j-1})$ and midpoints $\varphi_j \in [\nu_{j-1,}\nu_j])$. We then define the Riemann sum $S(s)$ for the partition s: $S(s) = \sum_{j=1}^{n} q(\varphi_j)(\nu_j - \nu_{j-1})$. $Q = \int_{\nu=0}^{1} q(t,\nu)d\nu$ can be defined as $\lim\limits_{m(s)\to 0} S(s) = Q$. Checking the existence of this limit requires preliminarily choosing a convergence criterion. Uhlig (1996) proposes using a convergence in mean square criterion (equivalent to L^2- convergence): we will then say that the limit exists if the variance of $S(s)$ converges to 0 as $m(s) \to 0$.

To prove convergence in mean square for the random variable $S(s)$ we recall that the $q(t,\nu)$'s, $\nu \in [0,1]$ are independent random variables. We calculate the common (finite) mean and variance for these variables, for t given. For simplicity and without loss of generality let us normalize $q(0,\nu) = 1$ for all ν. We then have, from (37):

$\mathrm{E}q(t,\nu) = \sum_{n=o}^{\infty} \frac{\lambda^n e^{-pt}(pt)^n}{n!} = e^{(\lambda-1)pt}$ and $\mathrm{Var}q(t,\nu) = \sum_{n=o}^{\infty} \frac{\left(\lambda^n - e^{(\lambda-1)pt}\right)^2 e^{-pt}(pt)^n}{n!} = e^{\left(\lambda^2-1\right)pt} - e^{2(\lambda-1)pt} \equiv \sigma^2(t)$.

Given the definition of $S(s)$, we then immediately see that $\mathrm{E}S(s) = e^{(\lambda-1)pt}$, while $\mathrm{E}\big(S(s) - e^{(\lambda-1)pt}\big)^2 = \sum_{j=1}^{n} \mathrm{E}\big(\big(q(\varphi_j) - e^{(\lambda-1)pt}\big)(\nu_j - \nu_{j-1})\big)^2 = \sigma^2(t)\sum_{j=1}^{n}(\nu_j - \nu_{j-1})^2 \leq \sigma^2(t)\sum_{j=1}^{n} m(s)(\nu_j - \nu_{j-1}) = \sigma^2(t)m(s)$. This goes to zero as $m(s) \to 0$, so the mean in square convergence criterion is met and Q is well defined.

Moreover, by Chebyshev's inequality, L^2- convergence implies convergence in probability (see for instance Capinski and Kopp 2004, Remark 8.14). We therefore conclude that for any real $\varepsilon > 0$, as $m(s) \to 0$, the probability that $\mid \big(S(s) - e^{(\lambda-1)pt}\big) \mid > \varepsilon$ goes to zero, so we can write $Q_t = e^{(\lambda-1)pt}$.

A.4 Social Planner's Solution

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 (55). Let μ_s denote the Lagrange multiplier associated to the constraint of the underlying current-value Hamiltonian $\mathcal{H}_s$ for the social planner's problem, the necessary conditions for a maximum are then the following:

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

$(ii)\ \frac{\partial \mathcal{H}_s}{\partial L_S} = 0 \Rightarrow C_s^{-\sigma}\left(1 - k(1-\sigma)L_s^{1+\frac{1}{\eta}}\right)^{\sigma-1} k\sigma\left(1+\frac{1}{\eta}\right)L_s^{\frac{1}{\eta}} = \frac{\mu_s \zeta(\lambda-1)}{L_s^2}$;

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

$(i\nu)\ \lim\limits_{t\to\infty} e^{-\rho t}\mu_s Q_s = 0$.

Combining (i) and (ii) one obtains the expression for optimal employment (56), while the growth rate of consumption in the centralized economy (57) immediately follows from (i) and (iii).

Appendix B

This appendix presents the solution of the model under the assumption that $\psi \neq \beta$.

B.1 Intermediate Good Firms

For $\psi \neq \beta$ the optimal price condition (10) is now replaced by:

$$P_t(\nu \mid q) = \frac{\psi}{\beta} q_t(\nu). \tag{B1}$$

Using this result into (8) we get:

$$x_t(\nu \mid q) = \left(\frac{\beta}{\psi}\right)^{\frac{1}{1-\beta}} L_t \tag{B2}$$

that replaces (11). Using (B1) and (B2) into $\pi_t(\nu \mid q) = [P_t(\nu \mid q) - \psi q_t(\nu)] x_t(\nu \mid q)$ gives:

$$\pi_t(\nu \mid q) = (1-\beta) \left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} q_t(\nu) L_t, \tag{B3}$$

instead of (12).

B.2 R&D Activity

Equations (13)-(15) are still valid, while (16) and (17) are replaced by:

$$r_t = \dot{L}_t / L_t - p_t + (1-\beta) \left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} \zeta \lambda \tag{B4}$$

and

$$\pi_t(\nu \mid q) / V_t(\nu \mid q) = (1-\beta) \left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} \zeta \lambda. \tag{B5}$$

Equations (27) and (28) are still valid.

B.3 Aggregation and Market Equilibrium

By substituting the demand for intermediate goods (B2) in the production function (6), given (21), we obtain:

$$Y_t = \frac{1}{\beta} \left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} Q_t L_t \tag{B6}$$

that replaces (22). Let $X_t = \int_{\nu=0}^{1} \psi q_t(\nu) x_t(\nu \mid q) \mathrm{d}\nu$ denote the aggregate expenditure on equipments, then using (B2) and (21) we get:

$$X_t = \psi \left(\frac{\beta}{\psi}\right)^{\frac{1}{1-\beta}} Q_t L_t, \tag{B7}$$

instead of (23). Combining (B6) with (B7) we get:

$$X_t = \beta^2 Y_t \tag{B8}$$

that is equal to the expression for X (24) obtained under the normalization $\psi = \beta$.

Using (9) and (B6) we find an expression for the wage as function of the aggregate quality index Q :

$$w_t = \frac{1-\beta}{\beta}\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} Q_t, \tag{B9}$$

that replaces (25). Using (B3) and (21), aggregate profits immediately follow:

$$\Pi_t = (1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} Q_t L_t, \tag{B10}$$

instead of (26).

Using (20), given (18), (19), (28), (B6) and (B9), delivers:

$$r = \left(\frac{g+\tau}{1-\beta} - \tau_w\right)\frac{\lambda\zeta}{\tau_r \beta}(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} \tag{B11}$$

that replaces (32).

Using (5), (B6), (B7), (18) and (27), given (30) and (B9), into the resource constraint (29), simplifying and solving for p_t give:

$$p_t = \left(1 - \beta^2 - g - \frac{\eta\left(1 - k(1-\sigma)L_t^{1+\frac{1}{\eta}}\right)(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)L_t^{1+\frac{1}{\eta}}}\right)\frac{\zeta}{\beta}\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}} \tag{B12}$$

that replaces (33).

Plugging (B11) and (B12) into (B4) we obtain the following differential equation for L_t :

$$\dot{L}_t = a_\psi L_t - b_\psi L_t^{-\frac{1}{\eta}}, \tag{B13}$$

where
$a_\psi \equiv \left[1 - \beta^2 - g + \frac{\eta(1-\sigma)(1-\beta)(1-\tau_w)}{\sigma(1+\eta)} + \left(\left(\frac{g+\tau}{1-\beta} - \tau_w\right)\frac{1}{\tau_r} - \beta\right)\lambda(1-\beta)\right]\frac{\zeta}{\beta}\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}$,
$b_\psi \equiv \frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)}\frac{\zeta}{\beta}\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}$.

Proceeding as in the main text it is easy to prove the existence and uniqueness of an unstable steady state solution, $L = \left(\frac{b_\psi}{a_\psi}\right)^{\frac{\eta}{1+\eta}}$:

$$L = \left(\frac{\frac{\eta(1-\tau_w)}{\sigma k(1+\eta)}}{1 + \beta - \frac{g}{1-\beta} + \frac{\eta(1-\sigma)(1-\tau_w)}{\sigma(1+\eta)} + \left[\left(\frac{g+\tau}{1-\beta} - \tau_w\right)\frac{1}{\tau_r} - \beta\right]\lambda}\right)^{\frac{\eta}{1+\eta}}. \tag{B14}$$

B.4 Balanced Growth Path and Optimal Tax Analysis

In a BGP $\dot{L}_t/L_t = 0$, therefore from (B4), (39) and (42) we obtain:

$$r = \frac{\rho + \sigma\lambda\zeta(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}(\lambda - 1)}{1 - \tau_r + \sigma(\lambda - 1)}, \tag{B15}$$

$$\gamma = \frac{(\lambda-1)\left[\lambda\zeta(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}(1-\tau_r)-\rho\right]}{1-\tau_r+\sigma(\lambda-1)} \tag{B16}$$

and

$$p = \frac{\lambda\zeta(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}(1-\tau_r)-\rho}{1-\tau_r+\sigma(\lambda-1)}. \tag{B17}$$

The above equations replace (43)-(45). Clearly, positive growth requires the following restriction to hold: $\lambda\zeta(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}(1-\tau_r)-\rho>0$. Plugging (B15) into (B11) delivers an expression for τ_w as function of τ_r :

$$\tau_w = \frac{\tau+g}{1-\beta} - \frac{\beta\left[\rho+\sigma\lambda\zeta(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}(\lambda-1)\right]}{\lambda\zeta(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}\left[1-\tau_r+\sigma(\lambda-1)\right]}\tau_r \tag{B18}$$

that replaces (47). Hence, given the other policy parameters, τ_w goes up when $\lambda\zeta\left(\frac{1}{\psi}\right)^{\frac{\beta}{1-\beta}}$ goes up.

The optimal tax analysis is conducted as in Section 4, where now (50) is replaced by

$$C_0 = \frac{\eta Q_0(1-\beta)\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}(1-\tau_w)\left(1-k(1-\sigma)L^{1+\frac{1}{\eta}}\right)}{\beta\sigma k(1+\eta)L^{\frac{1}{\eta}}}. \tag{B19}$$

B.5 Ratios

By using (B6), (B7) and (B9), wage payments as a share of the total value-added are equal to:

$$\frac{w_tL_t}{Y_t-X_t} = \frac{\frac{1-\beta}{\beta}\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}Q_tL}{\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}\frac{1}{\beta}Q_tL-\psi\left(\frac{\beta}{\psi}\right)^{\frac{1}{1-\beta}}Q_tL_t} = \frac{1}{1+\beta}. \tag{B20}$$

Profits as a share of the total value-added can be derived by using (B6), (B7) and (B10):

$$\frac{\Pi_t}{Y_t-X_t} = \frac{\left(\frac{\psi}{\beta}-\psi\right)\left(\frac{\beta}{\psi}\right)^{\frac{1}{1-\beta}}Q_tL}{\left(\frac{\beta}{\psi}\right)^{\frac{\beta}{1-\beta}}\frac{1}{\beta}Q_tL-\psi\left(\frac{\beta}{\psi}\right)^{\frac{1}{1-\beta}}Q_tL_t} = \frac{\frac{\psi}{\beta}-\psi}{\frac{\psi}{\beta^2}-\psi} = \frac{\beta}{\beta+1}. \tag{B21}$$

Clearly the ratio between profits and labor income is β.

Finally, public consumption and transfer payments as a share of the total value added are, respectively, given by:

$$\frac{G_t}{Y_t-X_t} = \frac{g}{1-\beta^2}, \tag{B22}$$

$$\frac{T_t}{Y_t-X_t} = \frac{\tau}{1-\beta^2}. \tag{B23}$$

B.6 Social Planner's Solution

The social planner problem is solved as in Section 5. (51) still holds, while (52) becomes:

$$x_s(\nu, t \mid q) = \psi^{\frac{1}{\beta-1}} L_s. \tag{B24}$$

The equations for aggregate spending on equipment and final-good production, (53) and (54), become:

$$X_s = \psi^{\frac{\beta}{\beta-1}} Q_s L_s, \tag{B25}$$

$$Y_s = \frac{1}{\beta} \psi^{\frac{\beta}{\beta-1}} Q_s L_s. \tag{B26}$$

Solving the optimization problem of the social planner and combining the necessary conditions for a maximum we find that optimal employment is still given by (56), while the growth rate of consumption in the centralized economy is easily seen to be:

$$\frac{\dot{C}_s}{C_s} = \gamma_s = \frac{\beta^{-1}\zeta(\lambda-1)\psi^{\frac{\beta}{\beta-1}}(1-g-\beta) - \rho}{\sigma}. \tag{B27}$$

Clearly, γ_s is negatively related to ψ.

Appendix C

In this appendix we characterize the decentralized equilibrium conditions and the social optimum for the model economy with dilution effect.

C.1 Decentralized Equilibrium and Optimal Tax Mix

The aggregate quality index Q_t is now given by:

$$Q_t = \int_0^N q_t(\nu)\mathrm{d}\nu, \tag{C1}$$

so that (C1) replaces (21). $N = L$ implies that in equilibrium (14) and (59) are equivalent and that, as a consequence (15) is still valid.

Recalling that F_t is the market aggregate value of firms in the intermediate sector and using (C1), we can then write $F_t = \int_0^N V(\nu \mid q)\mathrm{d}\nu$ so that (28) 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 (50) 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.

C.2 Social Planner's Solution

The social planner seeks to maximize the representative household's utility subject to the economy's resource constraint and (58). The optimality condition for each intermediate is still given by (52), which combined with (58) gives us:

$$Y_s = L_s \int_{\nu=0}^{N_s} q_s(\nu)\beta^{\frac{1}{\beta-1}}\mathrm{d}\nu. \tag{C2}$$

We now define $q_{sm} = \frac{\int_{\nu=0}^{N_s} q_s(\nu)d\nu}{N_s}$ which is the average quality of goods. We can then rewrite (C2) as:

$$Y_s = \beta^{\frac{1}{\beta-1}} L_s^2 q_{sm}, \tag{C3}$$

while the aggregate spending on equipments in the intermediate goods sectors is:

$$X_s = \beta^{\frac{\beta}{\beta-1}} L_s^2 q_{sm}, \tag{C4}$$

where the equality $L_s = N_s$ has been used.

Recalling the probability of innovation in each sector (59) and that in the sectors that innovate quality will increase by λ, we have that the law of motion of average quality is given by:

$$\dot{q}_{sm} = \frac{\zeta(\lambda-1)}{L_s^2} Z_s = \frac{\zeta(\lambda-1)}{L_s^2} \left(\beta^{\frac{1}{\beta-1}} q_{sm} L_s^2 \left(1-g-\beta\right) - C_s \right), \tag{C5}$$

where we have used the economy wide resource constraint, $Y_s - X_s - C_s - G_s = Z_s$, (C3), (C4) and $G_s = gY_s$.

Solving the optimization problem of the social planner and combining the necessary conditions for a maximum, we obtain an expression for optimal growth rate of consumption equal to (57), while employment is:

$$L_s = \left(\frac{\eta}{k\left(\frac{\sigma}{2}\left(1-\eta\right)+\eta\right)} \right)^{\frac{\eta}{\eta+1}}. \tag{C6}$$

Comparing this expression with (56) we see that now the socially optimal L_s is higher.

Appendix D

In this appendix we solve the baseline model under the assumption that R&D expenses are subsidized at the rate $S_{R\&D}$, so that at time t investing z_t in $R\&D$ has a private cost given by $z_t(1 - S_{R\&D})$.

D.1 Decentralized Equilibrium and Optimal Tax Mix

Given the subsidy on R&D, (15) is now replaced by:

$$V_t(\nu \mid q) = \frac{(1 - S_{R\&D}) L_t q_t(\nu)}{\zeta\lambda}, \tag{D1}$$

that combined with (12) delivers:

$$\pi_t(\nu \mid q)/V_t(\nu \mid q) = \frac{(1-\beta)\zeta\lambda}{1 - S_{R\&D}}, \tag{D2}$$

that in turn replaces (16). The Hamilton-Bellman-Jacobi equation (17) then becomes:

$$r_t = \dot{L}_t/L_t - p_t + \frac{(1-\beta)\zeta\lambda}{1 - S_{R\&D}}, \tag{D3}$$

where we have again dropped the arguments in $p_t(\nu \mid q)$ since we know that the arrival rate of innovation is the same across lines. The aggregate value of assets when there's a subsidy on R&D spending is then found to be:

$$F_t = \frac{(1 - S_{R\&D})L_t Q_t}{\zeta\lambda}, \tag{D4}$$

that replaces (28).

The flow budget constraint of the government is now:

$$G_t + S_{R\&D}Z_t = w_t\tau_w L_t + r_t\tau_r F_t - T_t, \tag{D5}$$

that replaces (20).

Combining (D5) with (18), (19), (22), (25), (27) and (D4) we obtain the following expression for r_t:

$$r_t = \left(g + \tau + S_{R\&D}\beta\frac{p_t}{\zeta} - \tau_w(1-\beta)\right)\frac{\lambda\zeta}{\tau_r\beta(1-S_{R\&D})}. \tag{D6}$$

From here we see that r and, therefore, growth is positively influenced by the subsidy, for given L (and therefore p).

Since (5), (9), (18), (22), (23), (27), (29) and (30) are still valid with an R&D subsidy, p_t is still given given by (33). Notice that p_t does not depend directly on the subsidy but only (positively) through L_t.

To derive the dynamic equation of labor we proceed as in Section 3, by combining (D6) with (D3) and (33):

$$\dot{L}_t = a_{S_{R\&D}}L_t - b_{S_{R\&D}}L_t^{-\frac{1}{\eta}}, \tag{D7}$$

where

$a_{S_{R\&D}} \equiv \left[1 - \beta^2 - g + \frac{\eta(1-\sigma)(1-\beta)(1-\tau_w)}{\sigma(1+\eta)}\right]\left[1 + \frac{\lambda S_{R\&D}}{\tau_r(1-S_{R\&D})}\right]\frac{\zeta}{\beta} + \left[\left(\frac{g+\tau}{1-\beta} - \tau_w\right)\frac{1}{\tau_r} - \beta\right]\frac{\lambda\zeta(1-\beta)}{(1-S_{R\&D})\beta}$,

$b_{S_{R\&D}} \equiv \frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)}\left(1 + \frac{\lambda S_{R\&D}}{\tau_r(1-S_{R\&D})}\right)\frac{\zeta}{\beta}$.

Proceeding as in the main text is easy to prove the existence and uniqueness of the unstable steady state solution

$$L = \left(\frac{\frac{\eta(1-\tau_w)}{\sigma k(1+\eta)}}{1 + \beta - \frac{g}{1-\beta} + \frac{\eta(1-\sigma)(1-\tau_w)}{\sigma(1+\eta)} + \frac{\left[\left(\frac{g+\tau}{1-\beta} - \tau_w\right) - \beta\tau_r\right]\lambda}{\tau_r(1-S_{R\&D}) + \lambda S_{R\&D}}}\right)^{\frac{\eta}{1+\eta}}. \tag{D8}$$

This shows that the subsidy, *coeteris paribus*, increases L. The closed-form expressions for γ, r and p in the presence of a subsidy on R&D spending can be derived by combining (39) with (42) and (D3) for $\dot{L}_t/L_t = 0$:

$$r = \frac{\rho + \frac{\sigma\lambda\zeta(1-\beta)}{1-S_{R\&D}}(\lambda - 1)}{1 - \tau_r + \sigma(\lambda - 1)}, \tag{D9}$$

$$\gamma = \frac{(\lambda - 1)\left[\frac{\lambda\zeta(1-\beta)(1-\tau_r)}{1-S_{R\&D}} - \rho\right]}{1 - \tau_r + \sigma(\lambda - 1)}, \tag{D10}$$

and

$$p = \frac{\frac{\lambda\zeta(1-\beta)(1-\tau_r)}{1-S_{R\&D}} - \rho}{1 - \tau_r + \sigma(\lambda - 1)}. \tag{D11}$$

These replace (43), (44) and (45). Clearly, γ, r and p are all increasing in $S_{R\&D}$.

Plugging these expressions for r and p into (D6) we obtain:

$$\tau_w = \frac{\tau + g}{1-\beta} - \frac{\beta\left[\rho\left(1-S_{R\&D}\right)+\sigma\lambda\zeta\left(1-\beta\right)\left(\lambda-1\right)\right]}{\lambda\zeta\left(1-\beta\right)\left[1-\tau_r+\sigma(\lambda-1)\right]}\tau_r + \\ + \frac{\lambda\zeta(1-\beta)(1-\tau_r)-\rho\left(1-S_{R\&D}\right)}{\lambda\zeta\left(1-\beta\right)\left[1-\tau_r+\sigma(\lambda-1)\right]}\frac{\beta\lambda S_{R\&D}}{1-S_{R\&D}} \tag{D12}$$

which replaces (47).

The optimal tax mix is found proceeding as done in main text considering that (49) and (50) are still valid.

Appendix E

This appendix describes the solution of the model under the assumption that public expenditure G enhances productivity in the economy. Specifically the production function of the representative firm in the final good sector (6) is replaced by (60).

E.1 Decentralized Equilibrium and Optimal Tax Mix

At the optimum the demand for the intermediate good, $x_t(\nu \mid q)$ is:

$$x_t(\nu \mid q) = \left[\frac{(g/\beta)^{\frac{1-\alpha}{\alpha}}q_t(\nu)L_t^{\frac{1-\beta}{\alpha}}}{P_t(\nu \mid q)}\right]^{\frac{\alpha}{\alpha-\beta}}, \tag{E1}$$

that for $\alpha = 1$ boils down into (8). The demand for labor is still given by (9).

The intermediate good producer sets the price so as to maximize profits $\pi_t(\nu \mid q) = [P_t(\nu \mid q) - \psi q_t(\nu)]x_t(\nu \mid q)$ given the demand function (E1). At the optimum the following condition must hold:

$$P_t(\nu \mid q) = \alpha q_t(\nu). \tag{E2}$$

where we have normalized ψ to β. Plugging this result into (E1) delivers:

$$x_t(\nu) = \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-\beta}}(g/\beta)^{\frac{1-\alpha}{\alpha-\beta}}L_t^{\frac{1-\beta}{\alpha-\beta}} \tag{E3}$$

that replaces (11). Substituting the above results into the profit function yields:

$$\pi_t(\nu \mid q) = \frac{1-\beta}{C_1}q_t(\nu)L_t^{\frac{1-\beta}{\alpha-\beta}}, \tag{E4}$$

where $C_1 \equiv \frac{1-\beta}{(\alpha-\beta)\left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-\beta}}(g/\beta)^{\frac{1-\alpha}{\alpha-\beta}}}$. The above expression for profits replaces (12).

In the baseline case the arrival rate of innovation in line ν with an initial quality $q\lambda^{-1}$ is given by (14) or analogously by $p_t(\nu \mid q\lambda^{-1}) = \frac{(1-\beta)\zeta z_t(\nu|q\lambda^{-1})\lambda}{\pi_t(\nu|q)}$. By analogy, in the present set-up, the arrival rate of innovation in line ν with an initial quality $q\lambda^{-1}$ is assumed to be equal to:

$$p_t(\nu \mid q\lambda^{-1}) = C_1\frac{\zeta z_t(\nu \mid q\lambda^{-1})\lambda}{L_t^{\frac{1-\beta}{\alpha-\beta}}q_t(\nu)}, \tag{E5}$$

where we have made use of (E4). We then have:

$$z_t(\nu \mid q) = \frac{p_t(\nu \mid q) q_t(\nu) L_t^{\frac{1-\beta}{\alpha-\beta}}}{\zeta C_1}. \tag{E6}$$

Proceeding as in the main text it can be shown that (15) is now replaced by:

$$V_t(\nu \mid q) = \frac{L_t^{\frac{1-\beta}{\alpha-\beta}} q_t(\nu)}{C_1 \zeta \lambda}. \tag{E7}$$

As in the baseline case the value of a firm of line ν of quality q obeys the law of motion (13), that given (E4) and (E7), can be written as:

$$r_t = \frac{1-\beta}{\alpha-\beta} \frac{\dot{L}_t}{L_t} - p_t + (1-\beta)\zeta\lambda. \tag{E8}$$

The above expression replaces (17).

Plugging (E3) and (18) in (60) and aggregating, total output in the final good sector is given by:

$$Y_t = Q_t L_t^{\frac{1-\beta}{\alpha-\beta}} g^{\frac{1-\alpha}{\alpha-\beta}} \alpha^{-\frac{\beta}{\alpha-\beta}} \beta^{\frac{\beta-1}{\alpha-\beta}}, \tag{E9}$$

that replaces (22). The aggregate expenditure on final goods used to produce intermediate goods can be obtained as in the main text and is now given:

$$X_t = \alpha^{-1}\beta^2 Y_t, \tag{E10}$$

that replaces (24). Plugging (E9) into (9) delivers an expression for the wage as function of Q and L:

$$w_t = \frac{1-\beta}{L_t} Q_t L_t^{\frac{1-\beta}{\alpha-\beta}} g^{\frac{1-\alpha}{\alpha-\beta}} \alpha^{-\frac{\beta}{\alpha-\beta}} \beta^{\frac{\beta-1}{\alpha-\beta}}. \tag{E11}$$

Aggregate profits follow from (E4):

$$\Pi_t = \frac{1-\beta}{C_1} Q_t L_t^{\frac{1-\beta}{\alpha-\beta}}. \tag{E12}$$

From (E6) aggregate R&D spending is now

$$Z_t = \frac{p_t Q_t L_t^{\frac{1-\beta}{\alpha-\beta}}}{C_1 \zeta}. \tag{E13}$$

Finally, using (E7) we obtain the aggregate value of assets:

$$F_t = \frac{Q_t L_t^{\frac{1-\beta}{\alpha-\beta}}}{C_1 \zeta \lambda}. \tag{E14}$$

The government budget constraint is given by (20), while (18) and (19) are still valid. The market clearing conditions (29) and (30) must hold, however recalling (E10), it must be that $0 < g < 1 - \alpha^{-1}\beta^2$.

Combining (18), (19) and (20), as well as (E9), (E11) and (E14), we obtain an expression for r:

$$r = \left(\frac{g+\tau}{1-\beta} - \tau_w \right) \frac{\lambda\zeta(1-\beta)^2 \alpha}{\tau_r \beta (\alpha - \beta)} \tag{E15}$$

that replaces (32).

Using (5), (9), (18), (E9), (E10) and (E13) into (29), recalling (30), simplifying and solving for p_t gives:

$$p_t = \left(1 - \alpha^{-1}\beta^2 - g - \frac{\eta \left(1 - k(1-\sigma) L_t^{1+\frac{1}{\eta}} \right) (1-\beta)(1-\tau_w)}{\sigma k (1+\eta) L_t^{\frac{1}{\eta}+1}} \right) \frac{\zeta}{\beta} \frac{\alpha(1-\beta)}{(\alpha-\beta)} \tag{E16}$$

that replaces (33).

Plugging (E15) and (E16) into (E8) we obtain the following differential equation for L_t :

$$\dot{L}_t = a_G L_t - b_G L_t^{-\frac{1}{\eta}}, \tag{E17}$$

where
$a_G \equiv \left[1 - \alpha^{-1}\beta^2 - g + \frac{\eta(1-\sigma)(1-\beta)(1-\tau_w)}{\sigma(1+\eta)} + \left(\frac{g+\tau}{1-\beta} - \tau_w \right) \frac{\lambda}{\tau_r}(1-\beta) - \frac{\beta}{\alpha}(\alpha-\beta)\lambda \right] \frac{\zeta\alpha}{\beta}$,
$b_G \equiv \frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)} \frac{\zeta\alpha}{\beta}$.

Proceeding as in the main text it is easy to prove the existence and uniqueness of an unstable steady state solution:

$$L = \left(\frac{\frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)}}{1 + \alpha^{-1}\beta^2(\lambda-1) - g + \frac{\eta(1-\sigma)(1-\beta)(1-\tau_w)}{\sigma(1+\eta)} + \left(\frac{g+\tau}{1-\beta} - \tau_w \right) \frac{\lambda}{\tau_r}(1-\beta) - \beta\lambda} \right)^{\frac{\eta}{1+\eta}}. \tag{E18}$$

We can see that L is increasing in α, *coeteris paribus*. Notice that (E8) in steady state implies (41), while (39) and (42) are still valid. We infer that (43), (44) and (45), which derive from (39), (41) and (42) are also still valid.

By eliminating r from (43) using (E15) we get to the following relationship between the tax rates:

$$\tau_w = \frac{\tau+g}{1-\beta} - \frac{\beta\left[\rho + \sigma\lambda\zeta(1-\beta)(\lambda-1)\right]}{\lambda\zeta(1-\beta)\alpha\left[1-\tau_r+\sigma(\lambda-1)\right]} \frac{\alpha-\beta}{1-\beta}\tau_r. \tag{E19}$$

The optimal tax mix is found as described in the main text with (49) being still valid, while (50) is replaced by

$$C_0 = \frac{\eta Q_0(1-\beta)(1-\tau_w)\left(1 - k(1-\sigma)L^{1+\frac{1}{\eta}}\right)}{\sigma k(1+\eta)L^{1+\frac{1}{\eta}}} L^{\frac{1-\beta}{\alpha-\beta}} g^{\frac{1-\alpha}{\alpha-\beta}} \alpha^{-\frac{\beta}{\alpha-\beta}} \beta^{\frac{\beta-1}{\alpha-\beta}}. \tag{E20}$$

E.2 Social Planner's Solution

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 in (E6), which in the command economy implies the following relationship between the increase over time of aggregate quality Q_s and the aggregate amount of R&D effort Z_s :

$$\dot{Q}_s = \frac{Z_s\zeta}{L_s^{\frac{1-\beta}{\alpha-\beta}}}(\lambda-1)C_1. \tag{E21}$$

The optimal quantity used of each intermediate will be:

$$x_s(\nu, t \mid q) = \left(\frac{1}{\alpha}\right)^{\frac{\alpha}{\alpha-\beta}} g^{\frac{1-\alpha}{\alpha-\beta}} \left(\frac{1}{\beta}\right)^{\frac{1}{\alpha-\beta}} L_s^{\frac{1-\beta}{\alpha-\beta}} \tag{E22}$$

which is the social planner equivalent of the decentralized quantity in (E3). The social planner internalizes the positive externality that the overall production has on productivity through G. Notice that as $\beta^{\frac{\alpha}{\alpha-\beta}} < 1$, as $\alpha > \beta$, x_s will be bigger than its private analogous.

The equations for aggregates spending on equipment and final-good production are now

$$X_s = \beta \left(\frac{L_s^{1-\beta}}{\beta} \frac{g^{1-\alpha}}{\alpha^\alpha}\right)^{\frac{1}{\alpha-\beta}} Q_s, \tag{E23}$$

$$Y_s = \beta^{\frac{1}{\beta-\alpha}} \alpha^{\frac{\beta}{\beta-\alpha}} g^{\frac{1-\alpha}{\alpha-\beta}} L_s^{\frac{1-\beta}{\alpha-\beta}} Q_s.$$

For given Q_s and labor, the level of output in the decentralized economy is lower than the optimal social value, since $\beta^{\frac{1}{\beta-\alpha}} > \beta^{\frac{1-\beta}{\beta-\alpha}}$. Solving the optimization problem of the social planner and combining the necessary conditions for a maximum as done for the baseline model, we obtain an expressions for optimal employment and the optimal growth rate of consumption:

$$L_s = \left[\frac{\eta}{k\left(\frac{\alpha-\beta}{1-\beta}\sigma + \eta + \left(\frac{\alpha-\beta}{1-\beta} - 1\right)\sigma\eta\right)}\right]^{\frac{\eta}{1+\eta}}, \tag{E24}$$

$$\frac{\dot{C}_s}{C_s} = \gamma_s = \frac{\frac{\zeta(1-\beta)\beta^{\frac{\alpha}{\beta-\alpha}}(\lambda-1)}{(\alpha-\beta)}\left[\alpha\left(1-g\right)-\beta\right]-\rho}{\sigma}. \tag{E25}$$

Notice that L_s is decreasing in α.

Appendix F

This appendix describes the solution of the model in the presence of public debt assuming a constant debt-to-GDP ratio.

F.1 Decentralized Equilibrium and Optimal Tax Mix

The representative household chooses consumption and labor to maximize (1) subject to the instantaneous budget constraint that, with public debt, is:

$$\dot{A}_t = r_t(1-\tau_r)F_t + w_t(1-\tau_w)H_t - C_t + T_t + R_t B_t, \tag{F1}$$

where $A_t = F_t + B_t$ and B_t is public debt. A_0 is given. Without loss of generality R_t, the rate of return on debt is net of tax. As both F and B are riskless financial assets, in equilibrium the following no-arbitrage condition between the two assets must hold: $r_t(1-\tau_r) = R_t$. The flow budget constraint can then be written as:

$$\dot{A}_t = r_t(1-\tau_r)A_t + w_t(1-\tau_w)H_t + T_t - C_t. \tag{F2}$$

The first-order conditions for the consumer's optimization problem (3) and (5) still hold, but the transversality condition now becomes $\lim_{t\to\infty} e^{-\rho t} C_t^{-\sigma} A_t = 0$.

Equations (6)-(19) and (21)-(30) are all still valid. The government budget constraint is now:

$$\dot{B}_t = G_t + T_t + r_t B_t - w_t \tau_w L_t - r_t \tau_r A_t, \tag{F3}$$

with

$$\lim_{t\to\infty} B(t) e^{-\int_0^t R(s)ds} \leq 0. \tag{F4}$$

We set $B_t = \kappa Y_t$, where $\kappa > 0$. The public debt-to-GDP ratio $B/(Y - X)$ is then $\kappa/(1-\beta^2)$, given (24).

Now consider debt dynamics, recalling $B_t = \kappa Y_t$ and using (9), (18), (19), (22) and (28), the flow budget constraint (F3) of the government becomes:

$$\frac{\dot{B}_t}{B_t} = \frac{\dot{Y}_t}{Y_t} = \frac{g+\tau}{\kappa} + r_t - \frac{\tau_w(1-\beta)}{\kappa} - r_t \tau_r \left(1 + \frac{\beta}{\kappa\lambda\zeta}\right), \tag{F5}$$

where $\frac{\dot{Y}_t}{Y_t} = \frac{\dot{L}_t}{L_t} + \frac{\dot{Q}_t}{Q_t}$ given (22).

Differentiating (5) with respect to time and combining the result with (3) deliver:

$$\frac{\dot{Q}_t}{Q_t} = \frac{r_t(1-\tau_r) - \rho}{\sigma} + \frac{1}{\eta}\frac{\dot{L}_t}{L_t}, \tag{F6}$$

where we have used (30). Combining (F5) with (F6) we get rid of $\frac{\dot{Q}_t}{Q_t}$ and obtain:

$$\frac{\dot{L}_t}{L_t}\left(1 + \frac{1}{\eta}\right) = \frac{g+\tau}{\kappa} - \frac{\tau_w(1-\beta)}{\kappa} + \frac{\rho}{\sigma} + r_t\left[(1-\tau_r)\left(1 - \frac{1}{\sigma}\right) - \tau_r \frac{\beta}{\kappa\lambda\zeta}\right]. \tag{F7}$$

Use (17) to get an expression for r_t and then plug the result into (F7). Then note that since (5), (9), (18), (22), (23), (27), (29) and (30) remain unchanged, p_t as a function of L_t is still given by (33). Therefore, the model dynamics is now described by the following differential equation:

$$\dot{L}_t = a_B L_t - b_B L_t^{-\frac{1}{\eta}}, \tag{F8}$$

where

$$a_B = \frac{g+\tau-\tau_w(1-\beta)+\kappa\frac{\rho}{\sigma}+\left[\kappa(1-\tau_r)\left(1-\frac{1}{\sigma}\right)-\tau_r\frac{\beta}{\lambda\zeta}\right]\left[(1-\beta)\zeta\lambda-\frac{\zeta}{\beta}\left(1-\beta^2-g+\frac{\eta(1-\sigma)(1-\beta)(1-\tau_w)}{\sigma(1+\eta)}\right)\right]}{\kappa\left[1+\frac{1}{\eta}-(1-\tau_r)\left(1-\frac{1}{\sigma}\right)\right]+\tau_r\frac{\beta}{\lambda\zeta}},$$

$$b_B \equiv \frac{\tau_r\frac{\beta}{\lambda\zeta}-\kappa(1-\tau_{r,t})\left(1-\frac{1}{\sigma}\right)}{\kappa\left[1+\frac{1}{\eta}-(1-\tau_r)\left(1-\frac{1}{\sigma}\right)\right]+\tau_r\frac{\beta}{\lambda\zeta}}\frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)}\frac{\zeta}{\beta}.$$

For a positive fixed point we need a_B, $b_B > 0$ and for determinacy we need a_B, $b_B > 0$. If $\kappa > 0$ then $\kappa\left[1 + \frac{1}{\eta} - (1-\tau_r)\left(1 - \frac{1}{\sigma}\right)\right] + \tau_r\frac{\beta}{\lambda\zeta} > 0$. Therefore, for b_B to be positive we need the following condition:

$$\frac{\tau_r\frac{\beta}{\lambda\zeta}}{(1-\tau_r)\left(1-\frac{1}{\sigma}\right)} > \kappa. \tag{F9}$$

In Table 8 we solve the model using an initial value for the capital income tax rate such that the above condition is met.

Equations (37)-(42) are all still valid. (F5), given (39), is equal to:

$$g + \tau + r\kappa - \tau_w(1-\beta) - r\tau_r\left(\kappa + \frac{\beta}{\lambda\zeta}\right) = \kappa p(\lambda - 1) \tag{F10}$$

that combined with (41) to get rid of r delivers:

$$p = \frac{(1-\beta)\zeta\lambda\left[\kappa - \tau_r\left(\kappa + \frac{\beta}{\lambda\zeta}\right)\right] + g + \tau - \tau_w(1-\beta)}{\kappa\lambda - \tau_r\left(\kappa + \frac{\beta}{\lambda\zeta}\right)}. \tag{F11}$$

Note that (43), (44), (45) are still valid. To obtain an expression of τ_w as a function of τ_r, κ and of all the other parameters of the model we use (F11) and (45). Simplifying we obtain:

$$\begin{aligned}\tau_w = &\frac{g+\tau}{1-\beta} - \frac{\beta\left[\rho + (1-\beta)\zeta\lambda\sigma(\lambda-1)\right]}{\lambda\zeta(1-\beta)\left[\sigma(\lambda-1) + (1-\tau_r)\right]}\tau_r + \\ &+ \frac{(\lambda-\tau_r)\,\rho + (1-\beta)\zeta\lambda\,(1-\tau_r)\,(\sigma-1)\,(\lambda-1)}{(1-\beta)\left[\sigma(\lambda-1) + (1-\tau_r)\right]}\kappa\end{aligned} \tag{F12}$$

that replaces (47).

The optimal tax mix is found by proceeding as done in main text, considering that (49) and (50) are still valid and using (F8) we obtain:

$$L = \left(\frac{\left(\tau_r\frac{\beta}{\lambda\zeta} - \kappa(1-\tau_{r,t})\left(1-\frac{1}{\sigma}\right)\right)\frac{\eta(1-\beta)(1-\tau_w)}{\sigma k(1+\eta)}\frac{\zeta}{\beta}}{g + \tau - \tau_w(1-\beta) + \kappa\frac{\rho}{\sigma} + \left(\kappa(1-\tau_r)\left(1-\frac{1}{\sigma}\right) - \tau_r\frac{\beta}{\lambda\zeta}\right)C_2}\right)^{\frac{\eta}{1+\eta}}, \tag{F13}$$

with $C_2 \equiv (1-\beta)\left[\zeta\lambda - \frac{\zeta}{\beta}\left(1 + \beta - \frac{g}{1-\beta} - \frac{\eta(\sigma-1)(1-\tau_w)}{\sigma(1+\eta)}\right)\right]$.

Appendix G

This appendix presents the solution of the two-classes model.

G.1 Decentralized Equilibrium and Optimal Tax Mix

Given the lifetime utility function (61) and the individual budget constraint (62) of wealth owners, their optimal choices are analogous to (3) and (5) and (4), i.e. must respect these conditions:

$$\frac{\dot{C}_{1,t}}{C_{1,t}} + \frac{k_1(1-\sigma)(1+\eta)H_{1,t}^{\frac{1}{\eta}}}{\eta\left(1 - k_1(1-\sigma)H_{1,t}^{1+\frac{1}{\eta}}\right)}\dot{H}_{1,t} = \frac{r_t(1-\tau_r) - \rho}{\sigma}, \tag{G1}$$

$$\frac{w_{1,t}(1-\tau_{w,1})}{\sigma k_1 C_{1,t}} = \frac{(1+\eta)H_{1,t}^{\frac{1}{\eta}}}{\eta\left(1 - k_1(1-\sigma)H_{1,t}^{1+\frac{1}{\eta}}\right)} \tag{G2}$$

and $\lim\limits_{t\to\infty} C_{1,t}^{-\sigma}A_t e^{-\rho t} = 0$.

Given the preferences of the lower skilled workers in (63), and their individual budget constraint (64), at the optimum their labor supply is described by the following equation:

$$\left(1 - H_{2,t}^{1+\frac{1}{\eta}}k_2\left(1 + \sigma\frac{1}{\eta}\right)\right)w_{2,t}(1-\tau_{w,2}) = T_{2,t}\sigma\left(1 + \frac{1}{\eta}\right)k_2 H_{2,t}^{\frac{1}{\eta}}. \tag{G3}$$

Given (6) and (65) we have:

$$w_{1,t} = \alpha\frac{1-\beta}{L_{1,t}}Y_t, \quad \text{(G4)}$$

$$w_{2,t} = (1-\alpha)\frac{1-\beta}{L_{2,t}}Y_t. \quad \text{(G5)}$$

Combining these two equations, if we define $w_t \equiv \frac{w_{1,t}L_{1,t}+w_{2,t}L_{2,t}}{L_t}$, (9) is still valid. Indeed all the equations in Subsections 3.2, 3.3, 3.4 and 3.5 still hold till (29), with the exception of (20) which must be replaced by:

$$G_t = w_{1,t}\tau_{w,1}L_{1,t} + w_{2,t}\tau_{w,2}L_{2,t} + r_t\tau_r F_t - T_t, \quad \text{(G6)}$$

where $T_t = \chi T_{1,t} + (1-\chi)T_{2,t} = \chi\tau_1 Y_t + (1-\chi)\tau_2 Y_t$.

Equation (29) still holds if we define:

$$C_t = \chi C_{1,t} + (1-\chi)C_{2,t}, \quad \text{(G7)}$$

while (30) is replaced by

$$L_{1,t} = \chi H_{1,t} \quad \text{(G8)}$$

and

$$L_{2,t} = (1-\chi)H_{2,t}. \quad \text{(G9)}$$

Recalling that A_t denotes individual wealth, capital market clearing requires $F_t = \chi A_t$.

We can now proceed as in the main text and see that r_t is indeed constant over time

$$r = \left(\frac{g+\tau}{1-\beta} - \tau_{w,1}\alpha - \tau_{w,2}(1-\alpha)\right)\frac{\lambda\zeta}{\tau_r\beta}(1-\beta), \quad \text{(G10)}$$

Plugging (G5), $T_{2,t} = \tau_2 Y_t$, (G9) into (G3) delivers the fraction of time spent at work by each individual unskilled worker:

$$H_2 = \left(\frac{(1-\tau_{w,2})(1-\alpha)(1-\beta)}{\tau_2(1-\chi)\sigma\left(1+\frac{1}{\eta}\right)k_2 + k_2\left(1+\sigma\frac{1}{\eta}\right)(1-\tau_{w,2})(1-\alpha)(1-\beta)}\right)^{\frac{\eta}{1+\eta}}. \quad \text{(G11)}$$

Clearly, H_2 is constant over time. Consumption of unskilled workers immediately follows from (64), using (G5), $T_{2,t} = \tau_2 Y_t$ and (G9):

$$C_{2,t} = \frac{(1-\beta)(1-\alpha)}{(1-\chi)}(1-\tau_{w,2})Y_t + \tau_2 Y_t. \quad \text{(G12)}$$

Plugging (18), (22), (23), (27), (G2), given (G4), (G8), (G9) and (G12), into the resource constraint (29) and then simplifying we obtain an expression of p_t as a function of $H_{1,t}$:

$$p_t = \left(\begin{array}{c} 1-\beta^2-g-\frac{(1-\beta)\alpha(1-\tau_{w,1})\eta\left(1-k(1-\sigma)H_{1,t}^{1+\frac{1}{\eta}}\right)}{\sigma k(1+\eta)H_{1,t}^{\frac{1}{\eta}+1}}+ \\ -(1-\beta)(1-\alpha)(1-\tau_{w,2})-\tau_2(1-\chi) \end{array}\right)\frac{\zeta}{\beta}. \quad \text{(G13)}$$

Plugging (G10) and (G13) into (17) and using the fact that $\dot{L}_t/L_t = \alpha\dot{H}_{1,t}/H_{1,t}$ we obtain the following non linear differential equation for $H_{1,t}$:

$$\dot{H}_{1,t} = a_{two} H_{\ 1,t} - b_{two} H_{2,t}^{-\frac{1}{\eta}}, \tag{G14}$$

where

$a_{two} \equiv \frac{1}{\alpha}\left[\left(1-\beta^2 - g - (1-\beta)(1-\alpha)(1-\tau_{w,2}) - \tau_2\left(1-\chi\right) + \frac{\eta(1-\sigma)\alpha(1-\beta)(1-\tau_{w,1})}{\sigma(1+\eta)}\right) + (1-\beta)D_2\right]\frac{\zeta}{\beta}$,

$b_{two} \equiv \frac{1}{\alpha}\frac{\eta(1-\beta)\alpha(1-\tau_{w,1})}{\sigma k(1+\eta)}\frac{\zeta}{\beta} > 0$

with $D_2 \equiv \left(\frac{g+\tau}{1-\beta} - \tau_{w,2}(1-\alpha) - \tau_{w,1}\alpha\right)\frac{\lambda}{\tau_r} - \beta\lambda$.

Reasoning as in the baseline case, it is easy to prove that H_1 will be at all times given by:

$$H_1 = \left[\frac{\frac{\alpha(1-\tau_{w,1})\eta}{\sigma k_1(1+\eta)}}{1+\beta-\frac{g}{1-\beta}+\frac{\alpha(1-\tau_{w,1})\eta(1-\sigma)}{\sigma(1+\eta)} - (1-\alpha)(1-\tau_{w,2}) - \frac{\tau_2(1-\chi)}{1-\beta} + D_2}\right]^{\frac{1}{1+\eta}}. \tag{G15}$$

The tax mix maximizing the social welfare function defined in (66) is then found by proceeding as in all the other cases, with the search being now conducted over $\tau_{w,1}, \tau_{w,2}$ and τ_r. In our last exercise ζ is set as in (67).