Factor Income Taxation in a Horizontal Innovation Model*

Xin Long†
Asian Development Bank

Alessandra Pelloni‡
University of Rome “Tor Vergata”

February 2015

Abstract

We consider the optimal factor income taxation in a standard R&D model with technical change represented by an increase in the variety of intermediate goods. In the model there are externalities both to working and to saving, due to market power by firms, incomplete appropriability of the benefits of inventions and the cost of economic activity in terms of public expenditures. Redistributing the tax burden from labor to capital will increase employment but, in most cases, reduce the rate of growth, with conflicting effects on the deadweight loss from the externalities. For parameter values consistent with available estimates, the optimal tax rate on capital will be sizable.

Keywords: Capital Income Taxes, R&D, Growth Effect, Welfare Effect.

JEL classification: E62, H21, O41

1 Introduction

How to collect tax revenue at the lowest efficiency cost is a basic problem of public finance. The best design of a tax system is not only a fascinating topic for theorists but a hugely important one for policymakers.

*We wish to thank Barbara Annicchiarico, Costas Azariadis, Thomas Davoine, Lei Ji, Pietro Peretto, Xavier Sala-i Martin, Marika Santoro, Pedro Teles, Robert Waldmann, Joseph Zeira and participants in seminars and presentations (Asset 2014 Aix-en-Provence, Manchester 2014, Sophia Antipolis 2013, Barcelona 2013, Padua 2011, LAGV Marseille 2011, DEGIT Frankfurt 2010, SIE Catania 2010 and RCEA Rimini 2009) for comments received on previous versions of the paper. The first was titled “Welfare Improving Taxation on Saving in a Growth Model” (University of Teramo, Department of Communication, WP 71, 2011).

†South Asia Department, Asian Development Bank, 6 ADB Avenue, Mandaluyong City 1550, Metro Manila, Philippines.

‡Corresponding author: Department of Economics, Law and Institutions, University of Rome “Tor Vergata”, Via Columbia 2-0133, Rome, Italy. Email: alessandra.pelloni@uniroma2.it.
The main message from the large literature on the optimal taxation of factor incomes in dynamic settings, started with the seminal papers by Atkinson and Stiglitz (1976), Chamley (1986) and Judd (1985), is that taxing capital is a bad idea.\footnote{On the ongoing popularity of the message see Mankiw et al. (2009). The strong assumptions on which it is built are criticised in Piketty and Saez (2013). The Chamley-Judd result advocates a high tax on capital income in the initial period and a zero tax on capital income in future times. The result requires that the government can commit to zero tax in the future, so there is a problem of dynamic inconsistency (see Martin 2010). However many related quantitative investigations suggest that capital taxes should be zero or very small even in the short run (see Atkeson et al. 1999).} In accordance with this message statutory rates on capital income have fallen sharply in OECD countries since the late 1980s, both at the corporate and the personal level, even if estimates of actual tax rates calculated using national account statistics do not always confirm this declining trend.\footnote{The capital tax burden depends not only on statutory rates but on the definition of the tax base, and the extent of tax deductions, etc. See McDaniel (2007) for data on actual tax rates on capital in OECD countries.} Taxes on savings, comprising inheritance taxes, property taxes, dividend taxes and corporate or individual profit taxes, are indeed still substantial, so that their abolition would be a radical change in advanced countries: a large gap still remains between theory and policy.

In this paper, we take a new look at how the tax burden should be shared between labor and capital through the lenses of endogenous growth theory. The elements so introduced – imperfect competition, accumulation of intangibles and economies of scale – shed new light on the issue. In particular externalities to employment then arise suggesting that capital income should indeed be taxed. Our results therefore help reconcile theory and policy, in the sense that theorists need to learn from policy makers.

Capital taxation is often advocated on equity and social stability grounds, given the very marked inequality in wealth distribution we observe.\footnote{The zero capital tax prescription can hold even when individuals have different initial endowments, so its usefulness for redistribution has been put in doubt (see Judd 1985). Piketty and Saez (2013) show that no taxation of capital in the long run and redistribution of wealth are consistent, when a nonlinear capital income tax is used.} Our argument for capital taxation is instead based purely on efficiency: we show that shifting the tax burden from labor to capital may be a Pareto improvement.

To develop our analysis we adopt a workhorse of the endogenous growth literature: the so called “lab-equipment model”, where R&D leads to “horizontal-innovation” i.e. to the expansion in the range of goods available.\footnote{The model was proposed by Romer (1987) and Rivera-Batiz and Romer (1991) and is the R&D model first presented in growth textbooks from Acemoglu (2009) to Aghion and Howitt (2009) and Barro and Sala-i-Martin (2003). Given its flexibility and simplicity this model has provided a tractable framework for analyzing a wide array of issues in economic growth, synthetized in Gancia and Zilibotti (2005).} We extend it by allowing for elastic labor supply and by introducing a government.

In the model there is an infinitely lived representative agent who gets utility from consumption and leisure. The final good is produced by perfectly competitive firms with labor and differentiated intermediate products. The invention of a new intermediate has a fixed cost to be paid upfront and allows eternal
monopoly power over the good. The price of each intermediate is then higher than its marginal cost. This static inefficiency, which we label the “market power distortion”, is familiar from models of monopoly. A second, dynamic inefficiency stems from the incomplete appropriability of the social surplus from a new invention. When firms innovate they take into account only the future flows of profits the new goods will bring them, but not all of the increase in social surplus. The rate of innovation is therefore too low. We label this the “appropriability distortion”.

We assume that the only tax instruments available are linear income taxes, that the government determines the amount of fiscal revenue it wants to generate as a constant fraction of income, that a constant fraction of this revenue is transferred back to consumers and, finally, that the government budget is balanced at all times. The labor income tax rate and the capital income tax rate must adjust endogenously.\(^5\) Government consumption introduces a third distortion in the model as agents do not take into account the fact that higher income has a cost in terms of more public expenditure. We label this externality the “cost of government distortion”.

The mix of labor and capital taxation chosen to finance government affects all the distortions described above. A redistribution of the tax burden from labor to capital will increase employment. The demand of each intermediate then goes up. This increases welfare increases because, due to the “market power distortion” the production of intermediates is too low. Higher demand also means that the cost of a blueprint is distributed over more units of the good that the firm will sell: the cost of innovation for unit sold is then lower. Thus decreasing the tax on labor income could then lower the rate of growth and induce a further welfare gain through a reduction of the “appropriability distortion”. In practice the direct effect of the corresponding increase in the tax on capital income is more likely to dominate: the reduced return on savings leads to reduced incentives to innovate and grow. Nevertheless, welfare is not the same as growth and the shifting of the tax incidence from labor to capital may still be useful for welfare, while reducing growth. Finally, both higher labor and capital taxes will compress the deadweight loss from the “cost of government distortion”, but not necessarily to the same extent.

Essentially with a lower tax on labor and a correspondingly higher tax on capital households will enjoy higher consumption for a given set of intermediates, but will also suffer slower consumption growth. We investigate how the tradeoff between these two opposing effects depends on income factor shares, on the level and composition of public expenditures and on preferences. In a calibrated version of the model we explore this tradeoff in depth. To give a preview of our results, we find that capital should be taxed more the higher the capital’s income share, the tax revenue as percentage of GDP and the fraction of transfers in public outlays are. The optimal tax rate on capital is increasing in the rate of time discount, in the compensated elasticity of labor supply and in the

---

\(^5\)By a tax on capital income we mean a tax on income from savings. In the model there is no capital in the physical sense and wealth consists in patents.
inverse of the intertemporal elasticity of substitution in consumption. For reasonable parameters' values the interplay between the various channels through which the tax program works means that the optimal tax rate on capital is not only positive but quite large – given the levels of public spending observed in advanced economies. The result that capital taxation should be increasing in the capital’s income share is particularly relevant given the observed increase in this share since the 1980s (see Karabarbounis and Neiman 2014), a period during which capital taxation has not been showing any upward trend.

This paper contributes to the literature on Ramsey factor incomes taxation in growth models. Models of endogenous growth with no market failures have been used to reinforce the case against taxing capital because of the adverse effects on the rate of growth itself (see the survey in Jones and Manuelli 2005). On the other hand, a prescription of positive capital taxation has been shown to hold in models with incomplete or imperfect financial markets (e.g. Chamley 2001 and Imrohoroglu 1998), with finite horizon agents (e.g. Cremer et al. 2003, Erosa and Gervais 2002, Ho and Wang 2007, Uhlig and Yanagawa 1996 and Yakita 2003) or in which fiscal authorities cannot detect a source of income or distinguish between different ones (e.g. Correia 1996, Penalosa and Turnovsky 2005 and Reis 2011). Taxing capital may also be efficient when its marginal productivity depends positively on government spending (e.g. Barro 1990, Turnovsky 1996 and 2000 and Zhang et al. 2008), or when the alternative is to tax human capital (see Ben-Gad 2003, Chen et al 2011, Chen and Lu 2013, de Hek 2006 and Song 2002), whether a finite or infinite horizon is assumed.

The combination of two or more of these assumptions can lead to high optimal tax rates on capital (see Conesa et al. 2009 and Piketty and Saez 2013). To streamline the analysis we have deliberately avoided incorporating in our baseline model any of the assumptions listed above, but we stress that joint consideration of the various mechanisms highlighted in the literature would be necessary for formulating quantitative policy advice.

Close to our analysis are growth models in which capital should be taxed because of externalities to labor. The result can obtain through “learning by doing” à la Romer (1986), as shown by Pelloni and Waldmann (2000), through government spending à la Barro (1990), as shown by Baier and Glomm (2001), Guo and Lansing (1999) and Chen (2007), and through quality improving R&D as shown by Aghion et al. (2013) and Peretto (2007). Aghion et al. (2013)
offer a general proof that the Chamley-Judd result cannot hold when the rate of growth is a function of taxes, in which case capital income should be either subsidized or taxed.

In quality improving R&D models, the cost of taxing capital in terms of reduced growth is in general lower than in models of variety expansion like ours, because in the former each time an innovation occurs, the firm previously first in the quality ladder goes out of business: this is known as the “business stealing effect”. Negative and positive consequences of innovation coexist in these “creative destruction” models so market growth may be too low but also too high. In our model the market rate of growth is always inefficiently low. That taxing capital lowers growth but may increase welfare is therefore particularly striking. Often in papers on taxation and growth, only the growth and not the welfare effect of the tax experiments are calculated, in the belief that if in the market equilibrium growth is lower than optimal, the welfare effect of raising it will always prevail over level effects. However we show that is not necessarily the case. Indeed in calibrated examples, our counterintuitive result arises for parameters’ values well within the range of selections adopted in other settings in public finance, quantitative growth theory and business cycle analysis.

We extend the model by allowing for productive government expenditures, in the form of either R&D subsidies or of government services entering either the final sector production function or the agents’ utility function. All these extensions recommend a further shift of the tax burden from labor to capital. We also study the possibility that the aggregate production function is not linear in labor, because of a third factor of production in given supply (e.g. natural resources). In this case the result is a reduction in the optimal ratio between the tax on capital and the tax on labor.

Transitional dynamics effects may be important for taxation but we prove that our model economy always follows a unique unstable balanced growth path (BGP).

The rest of the paper is organized as follows: section 2 sets up the model and characterizes equilibrium dynamics and steady state, section 3 investigates the social planner’s solution, section 4 shows the effects on labor, growth and welfare of shifting the tax burden from labor to capital and presents a formula in...
which the optimal tax on capital is an implicit function of parameters, section 5 calibrates the model and studies its quantitative implications for the tax system design, section 6 presents four extensions and section 7 concludes. Proofs are relegated to appendices.

2 The Model

2.1 Households

We assume a continuum of length one of identical households. Each has utility $U$ given by:

$$U = \int_{t=0}^{\infty} e^{-\rho t} \left( \frac{1}{1-\sigma} C^{1-\sigma} \phi(L) \right) dt \quad (1)$$

where $C$ is consumption of the final good, whose price is the numeraire, $0 < L < 1$ is labor offered, $\rho > 0$ is the rate of time discount and $1/\sigma > 0$ is the intertemporal elasticity of substitution in consumption (IES). Conditions on $\phi$ ensuring non satiation of consumption and leisure and convexity of preferences are spelled out in Appendix A. As Zeng and Zhang (2007) note, normalizing the population to unity removes from the analysis of taxes the so called “scale effect”.

The instantaneous budget constraint consumers face is given by:

$$\dot{F} = r(1 - \tau_r)F + \pi_n(1 - \tau_r)N + w(1 - \tau_w)L - C + T. \quad (2)$$

Households derive their income by loaning entrepreneurs their financial wealth $F$ (of which all have the same initial endowment) and by supplying labor $L$ to firms, taking the interest rate $r$ and the wage rate $w$ as given. Capital income is taxed at the rate $\tau_r$ while labor income is taxed at the rate $\tau_w$. Households also receive a lump-sum transfer $T$. Optimization at an interior point implies that the marginal rate of substitution between leisure and consumption equals their relative price:

$$C = w(1 - \tau_w)(\sigma - 1)\phi'/\phi'. \quad (3)$$

Optimal consumption and leisure must also obey the intertemporal condition:

$$-\sigma \dot{C}/C + (\phi'/\phi) \dot{L} = \dot{\lambda}/\lambda = \rho - r(1 - \tau_r) \quad (4)$$

where $\lambda = \phi/C^\sigma$ is the shadow value of wealth. Given a no Ponzi game condition the transversality condition (TVC) imposes:

$$\lim_{t \to \infty} \lambda F e^{-\rho t} = 0. \quad (5)$$
2.2 Firms

In this economy there are a final good sector and an intermediate goods sector. R&D activity leads to an expanding variety of intermediate goods. The unique final good is produced competitively with the production function:

\[ Y = l^{1-\alpha} \int_0^N x(j)^\alpha dj, \quad 0 < \alpha < 1. \]  

(6)

where \( Y \) is the amount of final good produced, \( l \) is labor used and \( x(j) \) is the quantity used of the intermediate good indexed by \( j \). There is a continuum of length of such goods. \( p(j) \) is the price of the intermediate good \( j \), whose demand, by profit maximization is:

\[ x(j) = l (\alpha/p(j))^{2/\alpha}, \]  

(7)

while labor demand is:

\[ w = (1 - \alpha)Y/l. \]  

(8)

Since the final good sector is competitive and there are constant returns to scale profits are zero in equilibrium. In contrast firms which produce intermediate goods hold a patent thanks to which they earn monopoly profits for ever. The cost of production of the intermediate good \( j \), once it has been invented, is given by one unit of the final good. The present discounted value at time \( t \) of monopoly profits for firm \( j \), or in other words the value of the patent for the \( j^{th} \) intermediate good at time \( t \), \( v(j, t) \), is then:

\[ v(j, t) = \int_t^\infty (p(j) - 1)x(j)e^{-\tau(s,t)(s-t)} ds \]  

(9)

where \( \tau(s, t) \) is the average interest rate during the period of time from \( t \) to \( s \). The producer of the \( j \) intermediate good chooses \( p(j) \) to maximize \((p(j) - 1)x(j)\) where \( x(j) \) is given by (7), so for each \( j \), the equilibrium price and quantity are:

\[ p(j) = p = 1/\alpha \]  

(10)

and

\[ x(j) = x = l\alpha^{2/\alpha}. \]  

(11)

The price of each intermediate is higher than its marginal cost, so the quantity produced, \( x \), is lower than the socially optimal level, to an extent which will become apparent when comparing the market with the social planner’s equilibrium. This is a first market failure in the model, which we label the “market power distortion”. Higher labor supply implies a higher quantity produced of each intermediate and therefore a reduction in the deadweight loss from this market failure. This creates an externality to labor, because workers are atomistic and when deciding how much to work will not factor in this general equilibrium positive effect.
Combining (11) and (6) gives us:

\[ Y = xN/\alpha^2. \]  

(12)

By combining this with (8) and with (11) we see that, in the aggregate, the marginal productivity of labor

\[ MPL = \alpha^{2\alpha} N \]  

(13)

is higher than the wage:

\[ w = \alpha^{2\alpha} N(1 - \alpha). \]  

(14)

Profits are given, as a consequence of (11) and (10), by:

\[ \pi = (1/\alpha - 1) x. \]  

(15)

The cost of development of new products is \( \eta \) and there is free entry in the market for inventions. Intermediate goods producing firms will enter until the price of a patent is equal to its cost. Here we meet a second inefficiency, which is due to an appropriability problem: only the discounted value of profits, as opposed to all of social surplus originating from an invention, is taken into account when deciding whether to pay for research to develop it. The pace of innovation will be therefore too low. We will call this second market failure the “appropriability distortion”.

If we drop the \( j \) index in \( v \), (9) can be written as the Hamilton-Jacobi-Bellman equation:

\[ rv = \pi + \dot{v} \]  

(16)

which allows us to interpret it from an asset pricing perspective. The return on holding a blueprint, \( rv \), is given by dividends \( \pi \), plus the capital gains, i.e. the change in its value \( v \). In Appendix A, we show that, in a growing economy, we must have \( v = \eta \) in equilibrium at all times.\(^{13}\) But if \( v = \eta \) at all times, (16) implies that in equilibrium we will have:

\[ r = \pi/\eta \]  

(17)

which, given (15) and (11), can be written as:

\[ r = C_1 l, \]  

(18)

with

\[ C_1 = \alpha^{1+\alpha} (1 - \alpha)/\eta. \]  

(19)

The higher is labor supply the higher is the interest rate. As the sales of each intermediate good and therefore profits are increasing in labor, for their present discounted value to be equal to the given cost of an invention, the interest rate will have to increase. Looking at (18) we encounter the second positive

\(^{13}\)Our proof is an extension to the case of a variable \( L \), to the one offered in Acemoglu (2009) for the case of a fixed \( L \).
“dynamic” externality to labor in this model. Investment in an innovation is a sunk cost that is economically justified only when the anticipated revenue is sufficiently large. On the other hand, once a new idea has been discovered it can be used as many times as one wishes. Consequently the dimension of the potential market for an innovation will be a crucial determinant of whether it is profitable to pay for research to develop it. This is what Aghion et al. (2013) define the “market size effect”. In a closed economy with given population, the scale of operation of the firm depends on employment.

2.3 Government

We assume government consumption $G$ equals a fixed fraction, $g$, of gross aggregate output: $G = gY$. Analogously $T = tY$, with $t$ given. We rule out a market for government bonds and assume that the government runs a balanced budget. As part of the revenue from income taxes is transferred back to agents in equilibrium we have:

$$\tau_r rF + \tau_w wL = (t + g)Y. \quad (20)$$

Our assumption of given $g$ and $t$ is made mainly for convenience but the public expenditure components that might be seen as exogenous in actual economies (from public wages, and 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.\footnote{In fact our modeling choice could mean that the production function in the final sector is Leontief: $Y = \min \left[ \frac{G}{g_t} \int_0^1 \int_0^1 x(j) \eta \phi(j) \frac{a_{ij}}{a_{ji}} \right]$. Of course this would be an extreme assumption as regards the role of public services in economic activity. See section 6 for a more standard representation of productive government services.} Marrero and Novales (2007) document this and show that factor income taxes may be preferable to lump-sum taxes under the assumption of a given $g$, as the former allow an internalization of the fact that higher income will lead to extra public spending. This simple effect is also at work in the baseline version of our model.

2.4 Market Equilibrium

All investment in the model is research and development of new intermediate goods $\eta \tilde{N}$. The economy-wide resource constraint is therefore given by:

$$I = Y - xN = C + \eta \tilde{N} + gY. \quad (21)$$

We are now ready for the following:

**Definition 1** In a competitive equilibrium individual and aggregate variables are the same and prices and quantities are consistent with the (private) efficiency conditions for the households (2), (3), (4) and (5), the profit maximization conditions for firms in the final good sector, (7) and (8), and for firms in the intermediate goods sector, (10) (or 11) and (18), with the government budget
constraint (20) and with the market clearing conditions for labor \((L = l)\), for capital \((F = \eta N)\), and for the final good, (21).

Given \(F = \eta N\), considering the relationship between profits and the interest rate (17), the relationship between profit and quantity sold in the intermediate sector given by (15) and the relationship between this quantity and final sector output (12), the government budget constraint (20) can be written as:

\[
\tau_w = \left( t + g \right) / (1 - \alpha) - \alpha r, \tag{22}
\]

From the definition above we can then arrive at the following:

**Proposition 2** The competitive equilibrium conditions in the model give rise to the following differential equation for labor:

\[
\dot{L} = B(L)/A(L) \tag{23}
\]

where

\[
A(L) \equiv \sigma \phi'' / \phi' + (1 - \sigma) \phi' / \phi \tag{24}
\]

and

\[
B(L) \equiv \frac{\sigma C_1 \phi (1 - \sigma) (1 - \tau_w)}{\alpha \phi'} + \rho + C_1 L \left( \tau_r - 1 + \sigma + \frac{\sigma}{\alpha} \left(1 - \frac{g}{1 - \alpha}\right) \right). \tag{25}
\]

**Proof.** See Appendix A. ■

**Proposition 3** The necessary and sufficient condition for the existence of a BGP equilibrium, in which all variables grow at the same rate, is that (23) has a fixed point \(\tilde{L}\) between 0 and 1, consistent with the TVC from (5) and with a positive growth rate \(\gamma\) for capital and consumption. \(\tilde{L}\) is implicitly defined by:

\[
B(\tilde{L}) = 0 \tag{26}
\]

while \(\gamma\) is given by:

\[
\gamma = (r(1 - \tau_r) - \rho) / \sigma. \tag{27}
\]

The TVC can be written as:

\[
\gamma (\sigma - 1) + \rho = (r(1 - \tau_r)(\sigma - 1) + \rho) / \sigma > 0 \text{ or } \gamma < r(1 - \tau_r) \tag{28}
\]

and implies

\[
\tilde{L} - (\sigma - 1) \phi' / \phi' < 0. \tag{29}
\]

**Proof.** See Appendix A. ■

(29) in light of (3) is the well-known condition that consumption must be higher than labor income for dynamic efficiency. Restrictions on parameters ensuring existence of a BGP equilibrium will be considered after introducing a specific form for the function \(\phi\). However for the general case we can establish
some interesting results on the uniqueness and stability of the BGP, assuming existence. First of all, given the definition of $B(L)$ in (25) we have:

$$\frac{B'(L)}{C_1} \equiv \frac{\sigma}{\alpha} (1 - \tau_w) (1 - \sigma) \left( 1 - \frac{\phi''}{(\phi')^2} \right) + \tau_r - 1 + \sigma + \frac{\sigma}{\alpha} \left( 1 - \frac{g}{1 - \alpha} \right). \quad (30)$$

We can now state the following:

**Proposition 4** If $\tilde{L}$ defined by $B(\tilde{L}) = 0$ exists, while $\sigma > 1$ or $\sigma < 1$ and, as is likely, $\tau_w \leq 1 - \alpha + \alpha \tau_r/(1 - \sigma) + \sigma (1 - g/(1 - \alpha)) / (1 - \sigma)$ then $B'(\tilde{L}) > 0$. This implies the BGP equilibrium is unique and locally determinate and that there is no transitional dynamics to it.

**Proof.** See Appendix A. ■

Since the necessary conditions for $B'(\tilde{L})$ negative require unrealistic parameters’ values (in particular very high $\tau_w$, $\alpha$ and $g$ and very low $\sigma$), as we will see in the calibration section, from now on we consider only the case of a determinate and unique BGP equilibrium.

## 3 The Social Planner’s Economy

We report here some key conditions characterizing the solution to the social planner’s problem and stress the differences between them and their counterparts in the market equilibrium. Variables keep the same meaning as in the decentralized economy, but the index $s$ is used to indicate that they belong to the social planner’s economy. The optimization problem of the social planner is strictly concave so it has a unique solution. There is no transitional dynamics. Derivations are in Appendix B.

The optimal quantity used of each intermediate good is:

$$x_s = (1 - g)^{1 - \alpha} \alpha^{1 - \alpha} L_s. \quad (31)$$

By comparing $x_s$ with $x$ in (11) we see that, for given labor, and ignoring $g$, $x_s$ is always larger than $x$: this is again due to the “monopoly power distortion”. Indeed the distortion remains for $g > 0$, for realistic values of $g$ and $\alpha$. $g$ has however a mitigating effect on this distortion, because the social planner takes into account that more production has a cost in terms of government consumption, while private agents do not. We also have:

$$C_s = (\sigma - 1)(1 - g)^{1 - \alpha} \alpha^{1 - \alpha} (1 - \alpha) N_s \phi(L_s) / \phi'(L_s); \quad (32)$$

in the market economy $C$ can be too low, for given $L$ and $N$. In fact by combining (14) with (3) we see that

$$C = (1 - \tau_w)(\sigma - 1)\alpha^{2 - \alpha} (1 - \alpha) N \phi / \phi'. \quad (33)$$
Setting aside $g$ and $\tau_w$, $C < C_s$. This is because, for given effort in terms of labor, the economy is inside the production efficiency frontier as not enough of this labor is used in the intermediate sector. The difference between $C$ and $C_s$ is decreasing in $g$ and increasing in $\tau_w$. Labor is implicitly given by:

\[
\frac{(1-\alpha)(1-g)^{\frac{1}{1-\alpha}}}{\eta} \alpha^{\frac{\alpha}{1-\alpha}} L_s \left(1 - \frac{(\sigma - 1)\phi}{\phi' L_s}\right) = \frac{(1-\alpha)(1-g)^{\frac{1}{1-\alpha}}}{\sigma} \alpha^{\frac{\alpha}{1-\alpha}} L_s - \rho \tag{34}
\]

In Appendix B we prove that:

**Proposition 5** In the market economy employment is always lower than its socially efficient level.

\[L_s > \bar{L}.\tag{35}\]

Finally socially optimal growth is given by:

\[\gamma_s = (C_{1s} L_s - \rho) / \sigma\tag{36}\]

where

\[C_{1s} \equiv (1-\alpha)(1-g)^{\frac{1}{1-\alpha}} \alpha^{\frac{\alpha}{1-\alpha}} / \eta.\tag{37}\]

The presence of $g$ makes for a lower $\gamma_s$, as the cost of government services, which is overlooked by savers in the market economy, increases overtime. In spite of this, growth is generally lower than socially optimal. A first reason for this suboptimality is that $\gamma_s$ is linear in $L_s$, while $\gamma$ is linear in $\bar{L}$, which is lower than $L_s$: the “monopoly power distortion” has a depressing effect on the level of activity and through the level of activity also a lowering effect on growth.

There is also a second reason for the suboptimality: the difference between $C_{1s}$ and $C_1$ themselves. Just by using (19) and (37) we can write:

\[\frac{(\gamma_s / \gamma) |_{L_s = 1}}{(\gamma_s / \gamma) |_{L_s = 1}} \equiv \frac{\alpha^{\frac{1}{1-\alpha}} (1-\alpha)(1-g)^{\frac{1}{1-\alpha}} - \rho \eta}{\alpha^{\frac{1}{1-\alpha}} (1-\alpha)(1-\tau_r) \alpha^{\frac{\alpha}{1-\alpha}} - \rho \eta}.\tag{38}\]

$(\gamma_s / \gamma) |_{L_s = 1}$ would be the ratio between first best and market growth, in a world in which leisure did not affect utility and so, given our normalization for labor, employment would be fixed at one.\(^{15}\) The ratio is always bigger than one, because, as we will see when calibrating the model, $1 - g < \alpha$ is very unrealistic for commonly used parameter values. With labor given, the “appropriability distortion” is never dominated by the “cost of government distortion” captured by $g$. Shifting the tax burden from labor to capital would always, for fixed labor, increase this difference, i.e. worsen the dynamic externality due to the "appropriability distortion".

\(^{15}\) This could be thought as the case in which $\phi(L)=1$, for any $0 < L < 1$. 

12
4 Effects of Taxes

In what follows we study the general equilibrium effects of revenue neutral shifts of the tax burden from labor to capital. For brevity we will often refer to this policy as an increase in the tax on capital. For any variable \(x\), \(dx/d\tau|_{L=L}\) indicates the derivative of \(x\) with respect to \(\tau\), when \(\tau\) simultaneously changes so as to keep the government budget balanced at the given \(g\) and \(t\).

4.1 Effect on Labor

It is relatively simple to calculate the effect of taxes on employment in this model because the wage rate does not vary with employment.\(^{16}\) We then have the following:

**Proposition 6** An increase in the tax rate on capital will increase employment, given determinacy, if and only if

\[
\left.\frac{dL}{d\tau}\right|_{L=L} = \frac{C_1\sigma}{B'(L)} \left(\frac{(\sigma - 1)\phi}{\phi} - \frac{\dot{L}}{\theta}\right) \geq 0. \tag{39}
\]

This condition always holds if \(\sigma > 1\).

The derivative can be deduced by totally differentiating (26) with respect to labor and the tax rates and noticing that by (22) \(d\tau_w/d\tau = -\alpha\). By Proposition 4, determinacy means \(B'(L) > 0\), so to understand the sign of the effect of the tax change we can look at the expression inside the parenthesis. Given (29) we see that for \(\sigma > 1\), the expression will always be positive. To explain the pivotal role of \(\sigma\) we notice that the reduction in \(\tau_w\) will increase the relative price of leisure. The substitution effect is then negative. There is no income effect on impact as the change in taxation is revenue neutral. So we expect a decrease in leisure. However if \(\sigma < 1\), consumption and leisure are complements: \(U_{CL} = \phi'/(C^\sigma) < 0\) (see Appendix A). Since the increase in \(\tau_c\) will reduce the relative price of current versus future consumption, current consumption will increase: this will make leisure more desirable. The final result could be, in theory, a reduction rather than an increase in labor. However data for developed countries tend to show that the higher the tax on labor income the lower the yearly hours worked per adult (see Ohanian et al. 2008 and references therein), while consumption and employment comove over business cycles.

4.2 Effect on Growth

By totally differentiating (27), given (39) we arrive to:

\(^{16}\)In section 6 we study the consequences of relaxing this feature of the economy and show that our qualitative results on taxation survive this change.
Proposition 7 An increase in the tax rate on capital will increase growth if and only if:

\[
\frac{d \gamma}{d \tau_r}|_{L=L} = \frac{r(1 - \tau_r)C_1}{\sigma LB'(\tilde{L})} \left( \frac{\sigma(\sigma - 1)\phi}{\phi'} - \tilde{L} - \frac{B'(\tilde{L})\tilde{L}}{C_1(1 - \tau_r)} \right) \geq 0. \quad (40)
\]

If \(d \gamma/d \tau_r|_{L=L}\) is positive then \(dL/d \tau_r|_{L=L}\) is strictly positive.

The last statement comes just by comparing (39) and (40), given \(B_0'(\tilde{L}) > 0\). Conversely the tax shift can increase employment without increasing growth. This is because from (27) we see that for growth to increase we need not only the gross of tax but the net of tax interest rate to increase. On the other hand, the condition for a shift from taxation of labor to taxation of capital to improve growth is sufficient but not necessary for the tax reform to increase welfare as we will see in the next subsection.

4.3 Ramsey Taxation

We now solve the following Ramsey Problem: given the set target for fiscal revenue choose the tax policy to maximize households’ welfare by taking into account the equilibrium reaction of private agents to the tax policy.\(^{17}\)

Since the economy will always follow a BGP, it is straightforward to express the utility attained by the representative agent in terms of \(\tau_r\) and \(\tau_w\). As \(\gamma\) and \(\tilde{L}\) are constant the integral in (1) can be simplified to yield:

\[
V = \frac{C(0)^{1-\sigma} \phi(\tilde{L})}{(1 - \sigma)(\rho - \gamma(1 - \sigma))}. \quad (41)
\]

In Appendix C we show that \(V\) can be reformulated as a function of parameters, of \(\tau_r\), of \(\tau_w\) and of equilibrium employment \(\tilde{L}\). The effect on social welfare of an increase in \(\tau_r\) and corresponding decrease in \(\tau_w\) necessary to keep the government budget constraint balanced, is then positive if \(dV/d \tau_r\) is positive. We consider the following monotonically increasing transformation of \(V\):

\[
\log \left( \frac{(1 - \sigma)V}{(1 - \sigma)} \right) \frac{d \log ((1 - \sigma)V)}{(1 - \sigma)d \tau_r} = \frac{\partial \log ((1 - \sigma)V)}{(1 - \sigma)d L} \frac{d \tilde{L}}{d \tau_r} + \frac{\partial \log ((1 - \sigma)V)}{(1 - \sigma)d \tau_r} + \frac{\partial \log ((1 - \sigma)V)}{(1 - \sigma)d \tau_w} \frac{d \tau_w}{d \tau_r}. \quad (42)
\]

In Appendix C we calculate the derivatives in the expression above and prove the following:

\(^{17}\)In the primal approach to the Ramsey problem tax rates are eliminated and quantities are chosen directly. Here we use the dual approach, i.e. we use the taxes as controls. For a simple early application of the latter approach in an endogenous growth model see Barro (1990).
Proposition 8  An increase in the tax rate on capital whose proceeds are used to reduce the tax rate on labor will be a Pareto improvement if and only if :

\[
\left(\frac{\sigma(\sigma-1)\dot{\phi}}{\phi'} - \ddot{L} - \frac{B'(\ddot{L})\dot{\phi}}{C_1(1-\tau_r)}\right)\frac{(1-\tau_r)r/\ddot{L}}{r(1-\tau_r)(\sigma-1) + \rho} + (43)
\]

We also have: 1. If the tax shift increases growth then it increases welfare. 2. If a value \(\tau_{ro}\) exists such that for \(\tau_r = \tau_{ro}\) the inequality above holds as an equality while it holds strictly for \(\tau_r < \tau_{ro}\) and does not hold for \(\tau_r > \tau_{ro}\), then \(\tau_{ro}\) is the optimal tax rate on capital. 3. At an optimum the growth effect of \(\tau_r\) is always negative, i.e. \(d\gamma(\tau_{ro})/d\tau_r < 0\).

To understand statement 1, consider that growth can be increased by raising \(\tau_r\) only if \(\ddot{L}\) increases as well (from Proposition 8). Since \(\ddot{L}\) and \(\gamma\) are always lower than socially optimal their increase will, coeteris paribus, increase welfare. Also consider that, for given \(\ddot{L}\) and \(\gamma\), increasing \(\tau_r\) and therefore lowering \(w\), has a direct positive effect on \(C(0)\) (from 3 when we recall that \((\sigma-1)\dot{\phi} > 0\) is required by leisure having positive marginal utility). The increase in \(C(0)\) also has a positive effect on welfare, as is clear from (41). The total effect on welfare of the tax change will therefore be positive.

However even if the tax change causes decreased growth agents will still benefit from higher consumption. If this benefit prevails, increasing the tax on capital will be a Pareto improvement. Statement 3 in the above proposition shows that this is exactly what happens at an optimum, i.e. when \(\tau_r = \tau_{ro}\).

This result goes against the widely held belief that when growth is suboptimal, which is always the case in this model, further decreasing it cannot possibly make agents better off: static gains could never dominate the loss in utility from slower growth.

In principle \(\tau_{ro}\) could be negative, but the uncoupling between the growth and welfare effect would still be there: pushing \(\tau_r\) below \(\tau_{ro}\) would increase growth but decrease welfare. In the next section we will show that for specifications of tastes and technology parameters commonly used in calibration exercises, \(\tau_{ro}\) will be not only positive but sizable.

5 Specifying the Utility Function

We consider here the following widely used class of functions for the disutility of labor:

\[
\phi(L) = (1 - L)^{1-\chi} \text{ and } \chi > 1, \sigma > 1 \text{ or } \chi < 1, \sigma < 1 < \chi + \sigma. \quad (44)
\]
\[
\hat{L} = 1 - \rho \eta (1 - \alpha)^{-1} \alpha \frac{\sigma}{\sigma - 1} \frac{(\chi - 1)}{\tau_r (1 - \tau_w)} \frac{(1 - \tau_r)}{(1 - \tau_w)} \frac{(1 - g)}{(1 - \alpha)}.
\]  

(45)

On the other hand in the social planner’s economy (given (34) we have:

\[
L_s = \frac{1 - \rho \eta (1 - \alpha)^{-1} (1 - g) \alpha \frac{\sigma}{\sigma - 1} (\chi - 1)}{1 + \frac{\tau_r - 1 + \sigma (1 + (1 - g)/(1 - \alpha))}{\alpha}}.
\]  

(46)

**Proposition 9** Necessary and sufficient conditions for the existence of a determinate equilibrium with positive growth and \( \hat{L} \) lower than one are that

\[
\frac{B'(L)}{C_1} \equiv \frac{\sigma (1 - \tau_w) (1 - \sigma)}{1 - \tau_r} + \tau_r - 1 + \sigma + \frac{\sigma}{\alpha} \left( 1 - \frac{g}{1 - \alpha} \right) > 0
\]  

(47)

and that \( \rho/C_1 \) belongs to the open interval \((a, b)\) where:

\[
a = \max \left\{ 1 - \sigma - \tau_r - \frac{\sigma (1 - \alpha - g)}{\alpha (1 - \alpha)}, \frac{- (\sigma - 1)^2 (1 - \tau_r)}{\sigma - 1 + (\alpha - 1) (1 - \tau_r)} \right\}
\]

and \(b = \frac{(\sigma - 1) (1 - \tau_w) (1 - \tau_r)}{(\sigma - 1) (1 - \tau_w) + (\chi - 1) (\alpha - 1) (1 - \tau_r)}\).

A further condition for the existence of an interior solution to the social planner’s problem is:

\[
\left( \frac{1 - g}{\alpha} \right) < \frac{C_1}{\sigma}.
\]  

(48)

We already know from section 4 that a redistribution of the tax burden from capital to labor will always cause labor to rise when \( \sigma > 1 \). However, given (45), this is also the case when \( \sigma < 1 \), as shown in Appendix D:

**Proposition 10** An increase in the tax on capital and corresponding decrease in the tax on labor will always increase labor.

\[
\frac{dL}{d\tau_r} \bigg|_{L = \hat{L}} = \frac{C_1}{B} \sigma \left( (\sigma - 1) \left( 1 - \hat{L} \right) \right) > 0.
\]  

(49)

The effect on growth is ambiguous.

However, for the parameterizations of the model we consider increasing the tax rate on capital always depresses growth.

Given (44), and using (47) Proposition 8 becomes:
Proposition 11 A positive welfare effect of a shifting of the tax burden from labor to capital requires:

\[
\frac{2 - \sigma - \chi}{(1 - \sigma)(1 - L)} - \frac{C_1(1 - \tau_r)}{C_1L(1 - \tau_r)(1 - \tau_r)(\sigma - 1) + \rho} \left( \frac{\sigma(\sigma - 1)(1 - \tilde{L})}{1 - \chi} + \tilde{L} \right) + \frac{\sigma(1 - \sigma)}{\alpha(1 - \chi)} \left( 1 + \alpha\tau_r - \frac{\rho + \rho}{1 - \alpha} \right) - 1 + \tau_r + \sigma + \frac{\sigma}{\alpha} \left( 1 - \frac{\rho}{1 - \alpha} \right) + \frac{\alpha}{1 + \alpha\tau_r - \frac{\rho + \rho}{1 - \alpha}} - \frac{C_1\tilde{L}}{C_1L(1 - \tau_r)(\sigma - 1) + \rho} \geq 0.
\]

(50)

If a value \(\tau_{ro}\) exists such that for \(\tau_r = \tau_{ro}\) the inequality above holds as an equality while it holds strictly for lower tax rates and does not hold for higher values, then \(\tau_{ro}\) is the optimal tax rate on capital.

In view of (45) and (19), given the parameters \(\{\sigma, \alpha, g, \rho, \chi, \eta, t\}\), calculating \(\tau_{ro}\) means solving the nonlinear equation in \(\tau_r\) obtained by equating to zero the LHS of (50).

5.2 Calibration

We are interested to the values of \(\tau_{ro}\) in regions in the parameters’ space that are consistent with available estimates of parameters and observed economic variables. To find these values we have to calibrate the model.\(^\text{18}\) We are completely aware that this model is not rich enough in number of variables to fit the data well. So our exercises aim more at understanding possible mechanisms of action of policy not noticed before in the literature rather than at finding precise quantitative results.

Several objects needed for the calculations have closed real-world counterparts so their calibration is relatively straightforward, while our other choices in feeding numbers to the model follow related studies (in particular Comin and Gertler 2006, Jones and Williams 2000, Strulik 2007 and Zeng and Zhang 2007).

First, we set values for the 8-tuple \(\{\sigma, \rho, \alpha, \tau_{rb}, \gamma_b, L_b, R, E\}\) where \(L_b, \gamma_b\) and \(\tau_{rb}\) indicate the baseline values for labor, growth and the capital income tax while \(R\) is total tax revenue over GDP, \(E\) is government consumption over GDP. Given \(\sigma, \rho\) and \(\gamma_b\) through (27) we calculate the rate of interest. Plugging \(L_b\) in (18) we calculate the baseline value for \(C_1\) and, given \(\alpha\), for \(\eta\) (through 19), for \(\tau_w\) (through 22) and for \(\chi\) (through 45). Finally we calculate \(g = (1 - \alpha^2)E\) and \(t = (1 - \alpha^2)R - g\). We then use (50) to obtain the optimal tax \(\tau_{ro}\), given the set \(\{\sigma, \rho, \alpha, t, g, \eta, \chi\}\).

Following Zeng and Zhang (2007) we set \(\sigma\) equal to 1.5 in our baseline economy. The former is closer to the value used in DSGE models of OECD economies than to the estimates of the parameter, which tend to be much higher than unity (see Alan and Browning 2010 for a recent study). As Zeng and Zhang (2007) and indeed most studies we set our central value for the rate of time discount \(\rho\) equal to 0.04.

\(^{18}\)The matlab code for generating results is available at www.alessandrapelloni.it.
The parameter $\alpha$ determines three different economic objects in our model, as $\alpha^2$ is the ratio of intermediate consumption to gross output, while $1/\alpha$ is the monopoly markup on intermediates and $1/(1 + \alpha)$ is the income share of labor. We choose $\alpha$ to match the US intermediates consumption which takes up around 0.45 of gross output. Hence the mark-up $1/\alpha$ is set at 1.49. This value exceeds the range $\{1.05, 1.37\}$ used by Jones and Williams (2000), but is lower than the 1.6 used by Comin and Gertler (2006). The latter note that while direct evidence is missing, given the specialized nature of the products we focus on an appropriate number for $1/\alpha$ would be at the high range of the estimates of markups for other types of goods in the US.\textsuperscript{19} The implied baseline value for the income share of labor is around 0.6, not distant from what we observe in advanced economies.

We fix $\tau_{rb}$, the baseline capital income tax rate, at 25 percent, close to the average tax rate on capital income estimated by McDaniel for the US in the period 1995-2007. Our baseline value for $E$ is 10 percent. The unweighted average of total tax revenue over GDP in OECD countries in 2005 was 35 percent. In the US the percentage was 27.1 and in France 44.1.\textsuperscript{20} We then take 35 percent as our benchmark for $R$.

Coming to labor supply, in 2005 the average US worker used 21 percent (24 percent) of her (his) time endowment to work, while the German one 13 percent.\textsuperscript{21} So we choose 0.17 as our benchmark value $L_b$ for labor. For the baseline growth rate, we use 2 percent, as the values used in related researches include 1.25 percent (Jones and Williams 2000), and 3 percent (Zeng and Zhang 2007).

Our choices and results as regards the baseline economy are summarized in Table 1:

\textsuperscript{19}Zeng and Zhang (2007) assume a benchmark value of $\alpha$ at 0.3, leading to a mark-up as big as 3.33. Cross-country comparisons show that in some other OECD countries the estimated markup value is higher than in the US. Neiss (2001) estimates for 24 OECD countries the mean of the markup to be 2.03 with standard deviation 0.78.

\textsuperscript{20}Source: OECD Tax Database.

Table 1: Baseline Economy: Parameterization and Results

<table>
<thead>
<tr>
<th>Parameters and Variables Set</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rate of time discount: $\rho$</td>
<td>0.04</td>
</tr>
<tr>
<td>initial labor: $L_b$</td>
<td>0.17</td>
</tr>
<tr>
<td>intermediate consumption to gross output ratio: $\alpha^2$</td>
<td>0.45</td>
</tr>
<tr>
<td>intertemporal elasticity of substitution (inverse): $\sigma$</td>
<td>1.5</td>
</tr>
<tr>
<td>government consumption to GDP ratio: $E$</td>
<td>0.1</td>
</tr>
<tr>
<td>tax revenue to GDP ratio: $R$</td>
<td>0.35</td>
</tr>
<tr>
<td>initial capital income tax rate: $\tau_{rb}$</td>
<td>0.25</td>
</tr>
<tr>
<td>initial GDP per capita growth: $\gamma_b$</td>
<td>0.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters and Variables Implied</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cost of inventions: $\eta$</td>
<td>0.079</td>
</tr>
<tr>
<td>parameter in the utility function: $\chi$</td>
<td>2.046</td>
</tr>
<tr>
<td>Frisch Elasticity: $F_R$</td>
<td>2.87</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Steady State Variables under Optimal Taxation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>optimal capital income tax rate: $\tau_{ro}$</td>
<td>0.288</td>
</tr>
<tr>
<td>optimal labor income tax rate: $\tau_{wo}$</td>
<td>0.392</td>
</tr>
<tr>
<td>optimal labor: $L_o$</td>
<td>0.175</td>
</tr>
<tr>
<td>optimal growth: $\gamma_o$</td>
<td>0.019</td>
</tr>
<tr>
<td>change in welfare using $\tau_{ro}$</td>
<td>0.079</td>
</tr>
</tbody>
</table>

A first point is that the capital income tax rate associated with maximum utility $\tau_{ro}$, at 28.81 percent is even higher than the rate of 25 percent suggested by the data. So the prescription arising from our simple model is more or less in line with the levels of capital income taxation observed in the real world.

A second important point is that, under our scheme, lower growth is consistent with higher welfare. This is especially interesting because as we have seen in this model the market equilibrium always generates an inefficiently low growth rate and there is a view that growth effects tend to prevail over level effects, as regards their impact on welfare. This is definitely not the case here.\textsuperscript{22} What we obtain shows clearly that as opposed to what conventional wisdom tends to maintain, even in an infinite horizon setting, the force of compounding over time does not always prevail.

The formula for the Frisch elasticity, given (44), is:

$$F_R = \frac{1/L - 1}{1 + (\chi - 1)/\sigma}. \quad (51)$$

The magnitude of the compensated aggregate labor supply elasticity is well known to be quite controversial. Researchers who use representative agent models like ours to study aggregate outcomes typically employ parameterizations

\textsuperscript{22}The technical reason why even though raising the tax on capital lowers growth, the welfare effect is positive is that for the optimizing problem to be well defined, in ours as in growth models in general use, the welfare attainable by the representative agent is assumed to be finite; this means the change in welfare due to the change in growth is finite as well and can therefore be smaller than the change in welfare due to intratemporal choices.
that imply higher aggregate labor supply elasticities than microeconometric studies would suggest. In our baseline calibration the implied value for the Frisch elasticity is 2.87. The values arising in our examples are within the range of the values used in the macro studies that typically vary from 1 to 4.\footnote{In King and Rebelo (1999) $FR=4$ allows a simple RBC model to mimic the main features of the US macro data. Smets and Wouters (2007) estimate a larger scale model and find $FR$ a mode of 1.92. Prescott (2006) suggests $FR=3$ explains differences in hours worked across OECD countries due to taxes. One explanation for the divergence between micro and macro estimates is that indivisible labor generates extensive margin responses that are not captured in micro studies of hours choices (e.g., Rogerson and Wallenius 2009). Imai and Keane (2004) find values for $FR$ as high as 4, when taking into account that the measured wage is less than the shadow wage which also reflects the value of on-the-job human capital accumulation. Domeij and Floden (2006) point out that ignoring borrowing constraints will induce a 50 percent downward bias in elasticity estimates. Rogerson (2011) argues that, if there are coordination problems that require a common workweek, idiosyncratic variations in wages typically used in micro studies may not be useful in uncovering preference parameters. Finally, Ljungqvist and Sargent (2014) focus on career length and suggest that reforms raising the officially mandated retirement age will lead to a substantial rise in aggregate labor supply elasticities.}

We now recapitulate the chain of reactions triggered by shifting the tax burden from labor to capital. On impact, lowering the fiscal wedge on labor while increasing it on interest income will cause labor supply to increase, because of the positive substitution effect, in the absence of an income effect. The increased labor supply induces a higher demand for the intermediate goods. Since the price of each intermediates is greater than its marginal cost, increased demand for an intermediate has a first order benefit for its inventor. The welfare loss from the “monopoly power distortion” is reduced. Moreover, the increase in profits induces a higher demand for investment in R&D so the interest rate will rise. The after-tax interest rate will however decrease. The BGP growth rate, as a monotonically increasing function of the after-tax interest rate, also decreases. So the welfare loss from the “appropriability distortion” is worsened. However, we find that the two changes in welfare of opposite sign associated with a redistribution of the tax incidence from labor to capital balance out for sizable levels of the tax rate on capital.

5.3 Sensitivity Analysis

We now conduct some sensitivity analysis to better understand how differences in fundamentals affects the optimal tax program and to establish how robust our results are to changes in the baseline calibration.

We first consider the following experiments:

(1) we vary the rate of time discount $\rho$ to 0.03 and to 0.05;
(2) we vary the parameter $\alpha^2$ to 0.4 and to 0.5;
(3) we vary the ratio of tax revenue to GDP, i.e. $R = (t + g)/(1 - \alpha^2)$, to 0.27 and to 0.44;
(4) we vary the ratio of government expenditures to GDP, i.e. $E = g/(1-\alpha^2)$, to 0.05 and to 0.15;
(5) we vary the inverse of the coefficient of intertemporal elasticity of substitution $\sigma$ to 1.35 and to 3.

All computations are done by changing only one parameter at a time while all the other parameters, both set and implied, are fixed at their baseline values. This means that the initial values of employment and growth, $L_b$ and $\gamma_b$, will be different from those used for the base case calibration. We calculate these different $L_b$ and $\gamma_b$ to understand in what direction the optimal tax programs push labor and growth, with respect to the initial equilibrium with $\tau_{rb}$. Results are reported in Table 2.

Table 2: Sensitivity Analysis I: one parameter changes at a time

<table>
<thead>
<tr>
<th>$\alpha^2$</th>
<th>$L_b$</th>
<th>$\gamma_b$</th>
<th>$F_R$</th>
<th>$\tau_{ro}$</th>
<th>$\tau_{wo}$</th>
<th>$L_o$</th>
<th>$\gamma_o$</th>
<th>$\frac{\Delta L_b}{L_b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.40</td>
<td>0.187</td>
<td>0.030</td>
<td>2.558</td>
<td>0.193</td>
<td>0.402</td>
<td>0.181</td>
<td>0.032</td>
<td>0.031</td>
</tr>
<tr>
<td>0.50</td>
<td>0.149</td>
<td>0.010</td>
<td>3.379</td>
<td>0.387</td>
<td>0.383</td>
<td>0.168</td>
<td>0.007</td>
<td>0.158</td>
</tr>
<tr>
<td>0.03</td>
<td>0.176</td>
<td>0.028</td>
<td>2.765</td>
<td>0.253</td>
<td>0.415</td>
<td>0.176</td>
<td>0.028</td>
<td>0.067</td>
</tr>
<tr>
<td>0.05</td>
<td>0.164</td>
<td>0.012</td>
<td>2.995</td>
<td>0.323</td>
<td>0.369</td>
<td>0.174</td>
<td>0.010</td>
<td>0.089</td>
</tr>
<tr>
<td>R=0.44</td>
<td>0.127</td>
<td>0.008</td>
<td>4.063</td>
<td>0.445</td>
<td>0.437</td>
<td>0.155</td>
<td>0.005</td>
<td>0.235</td>
</tr>
<tr>
<td>R=0.27</td>
<td>0.205</td>
<td>0.030</td>
<td>2.283</td>
<td>0.144</td>
<td>0.355</td>
<td>0.193</td>
<td>0.034</td>
<td>0.015</td>
</tr>
<tr>
<td>E=0.05</td>
<td>0.161</td>
<td>0.018</td>
<td>3.076</td>
<td>0.304</td>
<td>0.381</td>
<td>0.168</td>
<td>0.016</td>
<td>0.086</td>
</tr>
<tr>
<td>E=0.15</td>
<td>0.180</td>
<td>0.023</td>
<td>2.676</td>
<td>0.271</td>
<td>0.403</td>
<td>0.183</td>
<td>0.022</td>
<td>0.071</td>
</tr>
<tr>
<td>$\sigma=3$</td>
<td>0.448</td>
<td>0.048</td>
<td>0.913</td>
<td>0.308</td>
<td>0.378</td>
<td>0.462</td>
<td>0.045</td>
<td>0.194</td>
</tr>
<tr>
<td>$\sigma=1.35$</td>
<td>0.120</td>
<td>0.007</td>
<td>4.141</td>
<td>0.323</td>
<td>0.364</td>
<td>0.128</td>
<td>0.005</td>
<td>0.071</td>
</tr>
</tbody>
</table>

To check the overall robustness of our results we also redo the computations by changing each parameter/variable at a time along with the two implied ones $\chi$ and $\eta$, so as to keep the steady state values of employment and growth, $L_b$ and $\gamma_b$, at their baseline levels; this allows us to move $\sigma$ below one, and to consider the effect of assuming alternative values for $L_b$ (0.13 and 0.21) and $\gamma_b$ (0.0125 and 0.03). Results are reported in Table 3.\(^\text{24}\)

\(^{24}\)The calibration strategy behind Table 3 follows Jones et al. (2005) who adopt it to study how the elasticity of intertemporal substitution affects the relationship between growth and volatility.
### Table 3: Sensitivity Analysis II: $\eta$ and $\chi$ vary

<table>
<thead>
<tr>
<th>$\alpha^2$</th>
<th>$\eta$</th>
<th>$\chi$</th>
<th>$F_R$</th>
<th>$\tau_{ro}$</th>
<th>$\tau_{wo}$</th>
<th>$L_o$</th>
<th>$\gamma_o$</th>
<th>$\frac{\Delta Y}{Y}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.088</td>
<td>2.150</td>
<td>2.764</td>
<td>0.223</td>
<td>0.383</td>
<td>0.167</td>
<td>0.021</td>
<td>0.039</td>
</tr>
<tr>
<td>0.5</td>
<td>0.071</td>
<td>1.927</td>
<td>3.017</td>
<td>0.351</td>
<td>0.409</td>
<td>0.186</td>
<td>0.017</td>
<td>0.143</td>
</tr>
<tr>
<td>0.03</td>
<td>0.092</td>
<td>2.065</td>
<td>2.855</td>
<td>0.273</td>
<td>0.402</td>
<td>0.173</td>
<td>0.019</td>
<td>0.074</td>
</tr>
<tr>
<td>0.05</td>
<td>0.069</td>
<td>2.033</td>
<td>2.892</td>
<td>0.299</td>
<td>0.384</td>
<td>0.176</td>
<td>0.018</td>
<td>0.082</td>
</tr>
<tr>
<td>$R=0.44$</td>
<td>0.079</td>
<td>1.776</td>
<td>3.217</td>
<td>0.398</td>
<td>0.468</td>
<td>0.197</td>
<td>0.017</td>
<td>0.211</td>
</tr>
<tr>
<td>$R=0.27$</td>
<td>0.079</td>
<td>2.286</td>
<td>2.628</td>
<td>0.181</td>
<td>0.330</td>
<td>0.163</td>
<td>0.022</td>
<td>0.023</td>
</tr>
<tr>
<td>$E=0.05$</td>
<td>0.079</td>
<td>1.986</td>
<td>2.946</td>
<td>0.294</td>
<td>0.388</td>
<td>0.176</td>
<td>0.019</td>
<td>0.082</td>
</tr>
<tr>
<td>$E=0.15$</td>
<td>0.079</td>
<td>2.115</td>
<td>2.801</td>
<td>0.281</td>
<td>0.396</td>
<td>0.174</td>
<td>0.019</td>
<td>0.075</td>
</tr>
<tr>
<td>$\sigma=3$</td>
<td>0.055</td>
<td>5.057</td>
<td>2.076</td>
<td>0.441</td>
<td>0.289</td>
<td>0.196</td>
<td>0.015</td>
<td>0.367</td>
</tr>
<tr>
<td>$\sigma=1.35$</td>
<td>0.083</td>
<td>1.736</td>
<td>3.160</td>
<td>0.265</td>
<td>0.407</td>
<td>0.172</td>
<td>0.020</td>
<td>0.052</td>
</tr>
<tr>
<td>$\sigma=0.98$</td>
<td>0.093</td>
<td>0.957</td>
<td>5.104</td>
<td>0.196</td>
<td>0.453</td>
<td>0.164</td>
<td>0.022</td>
<td>0.002</td>
</tr>
<tr>
<td>$L_b=0.13$</td>
<td>0.060</td>
<td>2.434</td>
<td>3.421</td>
<td>0.303</td>
<td>0.382</td>
<td>0.136</td>
<td>0.019</td>
<td>0.088</td>
</tr>
<tr>
<td>$L_b=0.21$</td>
<td>0.098</td>
<td>1.806</td>
<td>2.447</td>
<td>0.273</td>
<td>0.402</td>
<td>0.214</td>
<td>0.019</td>
<td>0.070</td>
</tr>
<tr>
<td>$\gamma_b=0.0125$</td>
<td>0.094</td>
<td>2.019</td>
<td>2.908</td>
<td>0.310</td>
<td>0.377</td>
<td>0.178</td>
<td>0.011</td>
<td>0.085</td>
</tr>
<tr>
<td>$\gamma_b=0.03$</td>
<td>0.065</td>
<td>2.073</td>
<td>2.846</td>
<td>0.267</td>
<td>0.406</td>
<td>0.172</td>
<td>0.029</td>
<td>0.072</td>
</tr>
</tbody>
</table>

### 5.4 Discussion

We see that, with the exception of $\sigma$, qualitative results from the first set of experiments, i.e. the direction in which changes in parameters tend to push optimal taxes, are not reversed in the second set, the only difference being that variations in the tax rates are less pronounced in the second set.

We are now ready to comment on our findings. The effects of $\alpha$, $\sigma$ and $g$ offer new insights, while the effects of ratio of fiscal revenue to GDP and of $\rho$ confirm what found in Aghion et al. (2013).

We stress that in all cases a redistribution of the tax burden from labor to capital decreases market growth: this is consistent with the available empirical evidence of a negative correlation between capital tax rates and growth.

**Effect of $\alpha$** Table 2 and Figure 1 show that a higher $\alpha$ calls for a higher tax on capital. Interestingly we can prove analytically, as is done in Appendix D, that a higher $\alpha$ induces an increase in $L_s/L$ and a decrease in $(\gamma_s/\gamma)|_{L,L_o=1}$. The fact that as $\alpha$ increases market labor is pushed farther away from the optimum, while growth, for given labor, gets closer to the optimum, regardless of any other parameter in the model, suggests that for tax policy, the higher is $\alpha$ the more important increasing labor becomes, even if it has the cost of discouraging savings and growth: this is exactly what the simulations show.

To understand why $L_s/L$ goes up when $\alpha$ does we start by noticing that from (13) and (14) the ratio between the wage and the aggregate marginal productivity of labor, $1 - \alpha$, is decreasing in $\alpha$. This is because agents, being atomistic, when choosing how much to work, take all aggregate variables as given. However the increase in the demand for intermediate goods through
higher employment has a first order positive effect on profits, which, due to the “monopoly power distortion”, are increasing in the size of the market. This spillover from labor to profits is increasing in \( \alpha \) because the ratio between labor income and capital income is \( \frac{1}{\alpha} \). The tax on labor pushes labor up so the social loss from the externality on labor is reduced. And this beneficial effect will be higher the higher is \( \frac{1}{\alpha} \).

Monopolists are not able to capture the entire social gain created by their innovation. For an intuition on the effect of \( \frac{1}{\alpha} \) on \( \frac{j}{j_L L} \), consider that the contribution of a blueprint to the present discounted value of flows of future income is given by the cost of the invention \( \eta \) divided by the income share of capital, that is \( \eta \frac{(1 + \alpha)}{\alpha} \). However firms only consider that part of the contribution they can appropriate, that is \( \eta \), the PDV of profits. There is a spillover to inventing represented by \( \frac{\eta}{\alpha} \). Clearly this is decreasing in \( \alpha \): the higher the income share of capital the lower the social cost of the “appropriability distortion”.

The higher the income share of capital, the higher the tax on capital should be: this is an important policy implication of our analysis.

**Effect of \( \rho \)** Table 2 and 3 show that the ratio between \( \tau_{r_0} \) and \( \tau_{w_0} \) is increasing in \( \rho \). It is not difficult to understand intuitively this result. A lower \( \tau_{w_0} \) will lead to higher labor and to gain in utility through higher current consumption, while growth will be slower because of the higher \( \tau_{r_0} \). However this dynamic loss is discounted more heavily with higher subjective discount rate \( \rho \).

\(^{25}\)From (14), (18) and (19), and \( F = \eta N \).
Another interesting result is that a higher ratio of government tax revenue to GDP, $R$, is associated with a higher optimal ratio between the tax on capital income and the tax on labor income. This can be explained by noticing that labor is a strictly concave function of $w$. In fact from (49) and considering from (22) that $\partial \tau_w / \partial \tau_r = -\alpha$, using (47), we find:

$$
\frac{d^2 L}{d\tau_w^2} = \frac{2\sigma}{\alpha^2} \left( \frac{C_1}{B'} \right)^2 \left( \frac{\sigma(\sigma - 1)}{\chi - 1} + 1 \right) \left( \frac{L}{\sigma} - \frac{(\sigma - 1)(1 - \tilde{L})}{(\chi - 1)} \right) < 0. \quad (52)
$$

On the other hand, growth is a linearly decreasing function of $r$, for given $\tilde{L}$. Hence, as the fiscal revenue to be raised increases, increasing $r$ becomes less and less distortionary than increasing $w$. In general, this different responsiveness of the rate of growth, for given employment, and of employment itself to taxation explains why the tax rate on capital tends to be more sensitive to parameters than the tax rate on labor, as a quick look at Tables 2 and 3 shows.

The presence of $g$, i.e. of $E$, is a further source of externalities in our model economy as agents, being atomistic, when choosing how much to work and save do not consider that a higher income implies higher government consumption. In fact we can see that when $g$ is higher, $L_s$ is lower (from 46) and $\gamma_s$ is lower as well not only through the lower $L_s$ but also because the coefficient linking $\gamma_s$ to $L_s$ is lower (from 36). This is because the social planner internalizes both the externalities that $g$ induces. There is a cost of a higher level of private economic activity in terms of more public expenditures and through saving the cost increases over time.

Our numerical results suggest that the higher is $E$, for given $R$, the lower the tax on capital will be. In fact from (45) taking the sum $g + t$ as given, so that initial taxes, from (22), can be taken as given as well, we immediately see that $L$ is increasing in $g$. Intuitively a higher $g$ means that more resources are subtracted from the private sector: this causes a higher $L$, through a negative income effect, as leisure is a normal good. On the other hand $\gamma$ depends on $g$ only through $\tilde{L}$: there is no direct effect of $g$ on the marginal rate of substitution between current and future consumption.

This may help to explain why increasing $\tau_w$, which reduces $\tilde{L}$, is more efficient than increasing $\tau_r$, which reduces $\gamma$, in correcting the negative externalities from $g$.

In Table 2 the effect of $\sigma$ on the tax ratio is weak and non-monotone, while in Table 3 the higher is $\sigma$ the higher is the tax on capital. Both tables show that a higher $\sigma$ is associated with a lower $F_R$. However the negative effect of $\sigma$ on $F_R$ is much higher in Table 3: this is because, as already noticed an increase in $\sigma$ provokes an increase in $\tilde{L}$, as is clear from (45), and through this a decrease in $F_R$, even if from (51) the direct effect of $\sigma$ on $F_R$ is positive for $\sigma > 1$. In Table 3, $L_b$ is kept constant, and $\sigma$ and $\chi$ have to move in
opposite directions for this to happen (as $\partial \dot{L} / \partial \chi$ is negative) so their opposite effects on $F_R$, again from (51), tend to offset each other.

These somewhat ambiguous effects of $\sigma$ on the optimal taxes we find can be traced back to the fact that $\sigma$ has a bearing on two different characteristics of preferences: the Frisch elasticity of labor supply and the intertemporal elasticity of substitution in consumption (IES). Indeed, the real object of our interest is the impact of each of these two characteristics, rather than of $\sigma$ itself, on the optimal tax structure. To disentangle the effects of IES and $F_R$ we then proceed as follows: we note that (51) and (45) form a system of two equations in the two unknowns $L_b$ and $\chi$, if all other variables ($F_R$ and the other parameters) are taken as given.\footnote{26 Of course since $F_R$ depends on $L$ itself, the taxes have an effect on $F_R$, so $F_R$ in the final equilibrium will be different from $F_R$ in the initial equilibrium – however this effect is never very strong. In particular, for the second experiment, if a configuration of parameters gives rise in the initial equilibrium to a higher $F_R$ than another configuration, this ranking will not change with Ramsey taxes.} Indeed this system admits an analytic solution as shown in Appendix D. We then conduct two further experiments: a) we let $\sigma$ vary over the interval $[1, 3]$, while keeping the $F_R$ at the baseline value 2.87; b) we keep $\sigma$ fixed at its baseline value 1.5 and let the initial $F_R$ vary in the range $[0.87, 4.87]$.

Figure 2: Optimal tax on capital as a function of IES

Figure 2 represents the first experiment and shows that a lower IES, for given $F_R$, calls for a higher capital tax rate. Shifting the tax incidence from labor to capital lowers growth – the less the lower is the IES – and thus worsens the dynamic inefficiency in the model. However a higher elasticity of the marginal utility of consumption means consumers care relatively more about the current increase in consumption (which is lower than future consumption in a growing economy) than about the decrease in future consumption (which is higher).
So, when the current consumption is increased along with employment this increment is given more weight than the future loss.

![Figure 3](image)

Figure 3 shows that a higher $F_R$ for given IES induces a higher optimal capital tax rate. This is because the advantage of pushing up the tax on capital and down the tax on labor is bigger the bigger the increase in labor.

**Effect of $b$** We now ask ourselves how choices of $\gamma_b$ may affect results. The reassuring answer is not much. Table 2 shows that a lower $\gamma_b$ pushes up the ratio between $\tau_{ro}$ and $\tau_{wo}$. For this effect we offer the following explanation: low $\gamma_b$ implies, ceteris paribus, a low interest rate, and for given $L_b$ a low $C_1$ (high $\eta$). On its turn $\eta$ high, to be consistent with the given $L_b$, by (45) requires a low $\chi$ as $L$ is decreasing in both parameters. This pushes up the Frisch elasticity, which as we have seen pushes up the ratio between $\tau_{ro}$ and $\tau_{wo}$.

**Effect of $L_b$** We find that a lower $L_b$ pushes up the ratio between $\tau_{ro}$ and $\tau_{wo}$, but again the effect is not very large. A lower $L_b$ implies a higher $C_1$ (to be consistent with the given $\gamma_b$) i.e. a lower $\eta$. A lower $\eta$ on its turn implies (to be consistent with a lower $L_b$) a higher $\chi$. A higher $\chi$ pushes down the Frisch elasticity, which is however pushed up by the low $L_b$ itself. This last effect, which makes for higher taxation of capital in our example prevails over the direct one.

### 6 Extensions

In this section first we consider some extensions of the model. First we study the effects on Ramsey taxation of government spending being productive in three
alternative ways: the first way is that part of the spending finances a subsidy to research and development, the second is that government services enter the production function of the final sector and the third is that these services enter the utility function of agents. A further extension consist in checking what are the consequences for taxation of strictly decreasing marginal productivity of labor at the aggregate level. Derivations are in Appendix E.

6.1 A Subsidy to R&D

We now assume that R&D expenses are verifiable and that they are subsidized at the rate $s_N$. This implies a reduction in the fixed cost to pay for a blueprint. The cost (and therefore the value $v$ of a firm) is no longer $\eta$ but $\eta(1 - s_N)$. We consider two values for the subsidy rate, 0.1 and 0.4, each under two alternative assumptions 1) that the subsidies are additional public expenses; or 2) that they substitute other expenses. Results are shown in Table 4:

<table>
<thead>
<tr>
<th>$s_N$</th>
<th>$E$</th>
<th>$\tau_{wo}$</th>
<th>$\tau_{ro}$</th>
<th>$L_0, L_b$</th>
<th>$\gamma_{ro}, \gamma_b$</th>
<th>$\Delta V/V$</th>
<th>$F_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.089</td>
<td>0.396</td>
<td>0.291</td>
<td>1.78, 0.172</td>
<td>0.025, 0.026</td>
<td>0.023, 0.025</td>
<td>0.085</td>
</tr>
<tr>
<td>0.4</td>
<td>0.0124</td>
<td>0.351</td>
<td>0.327</td>
<td>0.190, 0.178</td>
<td>0.051, 0.055</td>
<td>0.039, 0.044</td>
<td>0.152</td>
</tr>
</tbody>
</table>

The second and fourth columns report results when $g$ (i.e. $E$, for given $\alpha$) is reduced so as to keep the sum of government consumption and R&D subsidies over GDP approximately constant at its baseline value 0.1.

A higher $s_N$ tends to push up the ratio $\tau_{ro}/\tau_{wo}$. It is not difficult to figure out why. The subsidy enhances growth, for given labor, thus reducing the distortion on the saving margin, so the tax system can be geared more towards pushing up labor and reducing the distortion in the work effort margin.

In the third and fifth columns $E$ is kept constant at 0.1 so the R&D subsidy pushes up total public expenditure.

The effect of the subsidy on the ratio $\tau_{ro}/\tau_{wo}$ is stronger when government consumption is not simultaneously reduced, which confirms previous findings that higher public expenses tend to increase this ratio.

6.2 G in the Production Function

We now analyse the case in which $G = gY$ enters the production function of the representative firm in the final good sector as follows:

$$Y = L^{1-\alpha} \left( \int_0^L x(j)^{\alpha/\beta} d\ell \right)^{\beta} G^{1-\beta}, \ 0 < \alpha \leq \beta < 1. \quad (53)$$
The firm is atomistic and so takes $G$ as given (i.e. not influenced by its choice of $Y$). This formulation allows us to have constant returns to scale at the firm level in the final sector, so we can keep the assumption of perfect competition. At the aggregate level $Y$ is still linear in $N$, so unbounded growth is possible. We choose for $\beta$ the lowest value we can, while keeping all other parameters at their benchmark values. With $\beta = 0.985$ equilibrium labor is still high enough to deliver low but positive growth while the labor income share ($SL = (1 - \alpha) / (1 - \alpha^2 / \beta)$), the share of intermediates in final production ($IR = \alpha^2 / \beta$) and the markup ($\beta / \alpha$) are still plausible.

### Table 5: $G$ in the Production Function

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>0.985</th>
<th>0.985</th>
<th>0.985</th>
<th>0.985</th>
<th>0.985</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E, R$</td>
<td>0.1, 0.35</td>
<td>0.05, 0.35</td>
<td>0.15, 0.35</td>
<td>0.1, 0.27</td>
<td>0.1, 0.38</td>
</tr>
<tr>
<td>$\tau_{eo}$</td>
<td>0.339</td>
<td>0.364</td>
<td>0.316</td>
<td>0.191</td>
<td>0.394</td>
</tr>
<tr>
<td>$\tau_{wo}$</td>
<td>0.357</td>
<td>0.341</td>
<td>0.372</td>
<td>0.321</td>
<td>0.371</td>
</tr>
<tr>
<td>$L_0$</td>
<td>0.176</td>
<td>0.169</td>
<td>0.184</td>
<td>0.193</td>
<td>0.169</td>
</tr>
<tr>
<td>$L_b$</td>
<td>0.165</td>
<td>0.155</td>
<td>0.175</td>
<td>0.200</td>
<td>0.151</td>
</tr>
<tr>
<td>$\gamma_o$</td>
<td>0.005</td>
<td>0.001</td>
<td>0.008</td>
<td>0.016</td>
<td>0.001</td>
</tr>
<tr>
<td>$\gamma_b$</td>
<td>0.007</td>
<td>0.004</td>
<td>0.010</td>
<td>0.014</td>
<td>0.004</td>
</tr>
<tr>
<td>$\Delta V /</td>
<td>V</td>
<td>$</td>
<td>0.094</td>
<td>0.104</td>
<td>0.084</td>
</tr>
<tr>
<td>$F_R$</td>
<td>2.991</td>
<td>3.222</td>
<td>2.773</td>
<td>2.358</td>
<td>3.326</td>
</tr>
</tbody>
</table>

We see from column 2, where the size and composition of the government budget are the same as in the baseline case, that when government spending is productive ($\beta$ is lower than one) the optimal tax on capital will be higher. To understand this intuitively we notice that the ratio between the wage and the marginal productivity of labor in the aggregate (equal to $\beta - \alpha$, as shown in Appendix E), and the ratio between the interest rate and the marginal productivity of $N$ (equal to $\alpha (\beta - \alpha) / (\eta \beta)$, also shown in Appendix E) are increasing in $\beta$, but the ratio of these two ratios is decreasing in $\beta$: the first misalignment is stronger with respect to the second the lower is $\beta$, i.e. the higher is the elasticity of the final product with respect to the public services. The ratio of the tax on capital to the tax on labor then has to increase. A complementary possible explanation is that a higher $\beta$ induces a lower elasticity to labor of the production function, given by $(1 - \alpha) / (\beta - \alpha)$, and therefore a lower optimal labor.

In the other columns we vary the level and composition of the budget by making $E$ and $R$ vary. In column 3(4) $E$ goes down(up), while the tax revenue $R$ is kept fixed at the baseline level. In columns 5 and 6, $E$ is kept at the baseline value while $R$ varies. We use 0.38 as the maximum value for $R$ rather than 0.44 as we did in the previous subsection, because beyond 0.38 growth would quickly become negative. As is the case with unproductive government expenditures,

27The alternative would have been to let $\chi$ and $\eta$ vary so as to let lower values of $\beta$ be consistent with given initial growth and labor. We experimented with this approach and results (available from the authors) are not markedly different.
a higher level of tax revenue $R$ makes for a higher $\tau_{ro}$ to $\tau_{wo}$ ratio, while, for given $R$ a higher $E$ has the opposite effect.

### 6.3 G in the Utility Function

Hereby we analyse the case in which $G$ enters the utility function of the representative economic agent as follows:

$$U = \int_{t=0}^{\infty} \frac{e^{-\rho t}}{1-\sigma} (C^{1-\sigma}G^\theta)^{1-\sigma} \phi(L)dt,$$

where $\theta \in [0, 1)$ is a weighting factor. We still assume $G = gY$ is taken by agents as exogenous. To calibrate the model we have to fix the new parameter $\theta$ in such a way that it is roughly consistent with observed variables. The elasticity of substitution between $C$ and $G$ equals $(1-\theta)/\theta$. The baseline value of $\theta$ is set at 0.2, following Bouakez and Rebei (2007) and we let it vary to 0.1 and 0.25. We then keep all parameters at their baseline values and let $\theta$ vary. Findings are reported in Table 6.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>0.2</th>
<th>0.1</th>
<th>0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_{ro}$</td>
<td>0.335</td>
<td>0.309</td>
<td>0.351</td>
</tr>
<tr>
<td>$\tau_{wo}$</td>
<td>0.360</td>
<td>0.377</td>
<td>0.350</td>
</tr>
<tr>
<td>$L_o, L_b$</td>
<td>0.146, 0.137</td>
<td>0.161, 0.154</td>
<td>0.139, 0.128</td>
</tr>
<tr>
<td>$\gamma_o, \gamma_b$</td>
<td>0.009, 0.011</td>
<td>0.014, 0.016</td>
<td>0.006, 0.009</td>
</tr>
<tr>
<td>$\Delta V/</td>
<td>V</td>
<td>$</td>
<td>0.100</td>
</tr>
<tr>
<td>$F_R$</td>
<td>3.612</td>
<td>3.198</td>
<td>3.867</td>
</tr>
</tbody>
</table>

As Table 6 shows, the higher the weight of public consumption in effective consumption, the higher the ratio of $\tau_{ro}$ to $\tau_{wo}$. This result can be interpreted as follows. When $G$ enters the utility function the negative externalities to working and saving it provokes are lower than when $G$ represents only a cost (or is just wasted if we prefer this interpretation). From the simulations the externality to labor is reduced less.

### 6.4 Production Function Concave in Labor

In the baseline version of the model the reduced form of the production function in the final good sector is linear in labor. If the marginal productivity of labor were decreasing at the economy-wide level, one would expect that lower employment and therefore a higher ratio between the tax on labor and the tax on capital would be optimal.

We now propose a simple example, by which the intuition above is confirmed but shown to have a limited impact on results.

Assume in each period, a final good $Y$ is produced by a representative competitive firm, using labor, a composite of intermediate inputs and a third factor,
\( P \) (this could indicate natural resources) whose supply is fixed and whose income represents rents.

\[
Y = L^\kappa P^{1-\alpha-\kappa} \sum_{j=0}^{N} x(j)^\alpha dj, \quad \alpha, \kappa \in (0, 1), \quad \alpha + \kappa < 1.
\]  

(55)

This production function is consistent with the hypotheses of constant returns to scale at the firm level (so that the replication principle is respected, the product exhaustion theorem applies and we can then keep the assumption of perfect competition) and of constant returns to the reproducible factor (N), so that we still have sustained growth. We assume that all non-labor income is taxed at the same rate. Indicating by \( \varepsilon_{YL} \) the elasticity of the final product with respect to labor we have \( \varepsilon_{YL} = \kappa/(1-\alpha) \), while \( \delta L = \kappa/(1-\alpha)^2 \), where \( \delta L \) is the income share of labor and \( \alpha^2 \) is the ratio of intermediates to final production (IR). So lower \( \kappa \) and \( \alpha \) imply lower \( \varepsilon_{YL} \), \( \delta L \) and IR. To calibrate the model we have to fix \( \alpha \) and \( \kappa \) in a way roughly consistent with the last two observables. We then consider four broadly acceptable combinations of SL and IR (cases I to IV in the table below) giving rise to decreasing values of \( \varepsilon_{YL} \). Findings are reported in Table 7.

<table>
<thead>
<tr>
<th>Case</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL,IR</td>
<td>0.59, 0.45</td>
<td>0.59, 0.35</td>
<td>0.5, 0.45</td>
<td>0.5, 0.35</td>
</tr>
<tr>
<td>( \varepsilon_{YL} )</td>
<td>0.986</td>
<td>0.939</td>
<td>0.835</td>
<td>0.796</td>
</tr>
<tr>
<td>( \tau_{ro} )</td>
<td>0.298</td>
<td>0.271</td>
<td>0.369</td>
<td>0.349</td>
</tr>
<tr>
<td>( \tau_{wo} )</td>
<td>0.386</td>
<td>0.405</td>
<td>0.331</td>
<td>0.352</td>
</tr>
<tr>
<td>( L_o, L_b )</td>
<td>0.174, 0.168</td>
<td>0.171, 0.168</td>
<td>0.163, 0.141</td>
<td>0.160, 0.141</td>
</tr>
<tr>
<td>( \gamma_o, \gamma_b )</td>
<td>0.019, 0.021</td>
<td>0.024, 0.025</td>
<td>0.024, 0.027</td>
<td>0.029, 0.031</td>
</tr>
<tr>
<td>( \Delta V/</td>
<td>V</td>
<td>)</td>
<td>0.088</td>
<td>0.074</td>
</tr>
<tr>
<td>( F_R )</td>
<td>2.925</td>
<td>2.917</td>
<td>3.583</td>
<td>3.586</td>
</tr>
</tbody>
</table>

A lower \( \varepsilon_{YL} \) has the expected effect, i.e. it leads to a lower ratio between \( \tau_{ro} \) and \( \tau_{wo} \), but its impact is not very sizable. For given SL, lower IR and \( \varepsilon_{YL} \) do indeed translate into a lower \( L_o \) and a lower ratio between \( \tau_{ro} \) and \( \tau_{wo} \) (compare column I with column II and column III with column IV). However comparing I with III and II with IV we see that when the lowering in \( \varepsilon_{YL} \) comes through a lower SL the effect on \( L_o \) and the tax ratio is the opposite. So the “income share effect” prevails over the “elasticity of labor effect” under study in this subsection.

7 Concluding Remarks

This study analyses how the tax burden should be distributed between labor and capital in a R&D model of growth. Plausible calibrations of our model imply that the optimal tax rate on capital will be positive and sizable. Our analysis helps to make sense of the fact that in advanced economy tax rates on capital
are generally well above the zero level often recommended by the literature. We find that the sign of the growth effect of a tax program is not necessarily the same as that of the welfare effect and that the two effects should be analysed separately, even in models when growth is sub-optimal.

In the model there are three externalities, one related to the market power of firms, which implies employment is too low, the second related to the incomplete appropriability of the gains from innovation activity, which implies growth is too low, and the third related to the fact that public expenditure is increasing in income, which implies both employment and growth are too high. Shifting the tax burden from labor to capital has opposite effects on the first two distortions. It will lead to an increase in employment and a reduction in the deadweight loss from monopoly power. The increase in employment would in itself make innovation more profitable, thus reducing the deadweight loss from the second distortion as well. However in our simulations the after-tax interest rate is decreasing in the capital income tax, and so the growth rate goes down. The optimal fiscal mix depends on this tradeoff between increasing employment and increasing growth, which is also influenced by the third distortion. We analyse how the tradeoff depends on the model’s primitives. A pivotal element is the functional distribution of income. A lower labor income share means a higher spillover from labor: a lower tax on labor is then in order. On the other hand, a higher capital income share means a lower spillover from innovating. So a lower labor income share pushes towards taxing capital income. In this way we isolate a new reason why the study of the decline in the income share of labor observed globally in the last decades should be on today’s economic agenda, even beyond the obvious concerns with inequality it raises. Another insight from our analysis is that the higher the ratio between the tax revenue and income, the higher the ratio between the tax on capital and the tax on labor. This is because while the marginal rate of substitution between present and future consumption is linear in the capital income tax rate, the marginal rate of substitution between consumption and leisure is a convex function of the labor income tax rate. A third important new result is that, for given fiscal revenue, a higher proportion of transfers makes for taxing capital more: this is because when resources are absorbed by the public sector rather than returned to households, due to a negative income effect employment moves up thus getting closer to its efficient level. We also show that a high time discount rate, a low elasticity of intertemporal substitution in consumption and a high Frisch elasticity of labor supply all recommend taxing capital more. The first two features imply a higher marginal rate of substitution of present for future consumption, in a growing economy. So the loss in terms of consumption growth caused by a higher capital tax will be easier to compensate with higher current consumption. The effect of the elasticity of labor supply is easy to explain since the positive welfare effect of shifting the tax burden from labor to capital depends on the policy being effective in increasing employment. Our results on capital taxation are robust to changes in the assumption of unproductive government spending: indeed we find that when government expenditures are used to subsidize R&D or to pay for services entering the final sector production function or the utility function
of agents the optimal tax on capital goes further up. Finally we show that the
linearity in labor of the production function is not necessary for the prescription
of a positive tax on capital.

In future research we plan to extend our analysis along three main direc-
tions: the first is considering whether the effects we uncover would survive in
a model incorporating vertical as well as horizontal innovation, the second is
going beyond the representative agent assumption and looking at the interaction
between the personal and the functional distribution of income, the third
is studying the implications that the relationship between growth and business
cycles may have for the optimal design of a tax system.

References

University Press, Princeton.
No. 19086.
eters Using Synthetic Residual Estimation. Review of Economic Studies,
77, 1231-1261.
Public Policies with Distortionary Taxation. Journal of Economic Dynamics
and Control, 25, 2007-2042.
MIT Press.
a Government Spending Shock? Canadian Journal of Economics, 40, 954-
979.
[14] Chamley C. 1986. Optimal Taxation of Capital Income in General Equilib-

A Appendix to Section 2

A.1 Conditions on $\phi$

For non satiation in consumption of the final good to be positive we need:

$$\phi(L) > 0$$  \hspace{1cm} (A1)

and for non satiation in leisure:

$$(1-\sigma)\phi'(L) < 0.$$  \hspace{1cm} (A2)

Strict concavity of instantaneous felicity imposes:

$$(1-\sigma)\phi''(L) < 0$$  \hspace{1cm} (A3)

and

$$\sigma\phi''\phi/(\sigma - 1) - \phi'^2 > 0.$$  \hspace{1cm} (A4)
A.2 $v = \eta$ in a Growing Economy

$v > \eta$ is never possible because of the free entry assumption in the research market. On the other hand if $v < \eta$, no research would be done so that $N = 0$, and from the economy-wide resource constraint we would have $Y - xN = C + gY$, or, using (11) and (12),

$$C = (1 - \alpha^2 - g) NL \alpha \frac{\alpha}{(1 - \alpha)(1 - \tau_w)(\sigma - 1)}. \quad (A5)$$

Plugging this, together with (8), in (3), the equilibrium level of employment would be implicitly given by:

$$\frac{\phi}{\phi'} = \frac{L(1 - \alpha^2 - g)}{(1 - \alpha)(1 - \tau_w)(\sigma - 1)}. \quad (A6)$$

So if this equation had a solution for $L$ between 0 and 1, this solution would define the equilibrium level of employment in a growthless economy, $L_{ng}$. Plugging $L_{ng}$ in (A5) and in (15), the consumption level and the profit level in this growthless economy would also be given. With labor and consumption fixed over time, the Euler equation (4) implies an interest rate equal to $\rho/(1 - \tau_r)$. Now suppose that $v = v_0 < \eta$. If $\rho/(1 - \tau_r) - L_{ng} \alpha^{2(1-\alpha)}(1/\alpha - 1)/v_0 > 0$, i.e. if $r - \pi/v_0 > 0$, then, by (16), $\dot{v}/v > 0$. So $v$ will increase and, since $\pi$ and $r$ will stay the same, $r - \pi/v$ will increase as well, i.e. $\dot{v}/v$ will be increasing. This implies that in finite time $v$ will get to $\eta$, but then $\dot{v}/v > 0$ will be no longer possible. It would then become profitable to invest in inventions and growth would start. However this would require a jump in $C$ and $L$ (no longer dictated by (A5) and (A6)) which would violate the equilibrium conditions of agents. In analogous fashion, if $\rho/(1 - \tau_r) - L_{ng} \alpha^{2(1-\alpha)}(1/\alpha - 1)/v_0 < 0$ that is if $r - \pi/v < 0$, $v$ would be decreasing at an increasing rate, reaching the value 0 in finite time. If that happened (16) could not hold any longer. So again we would have a contradiction. Finally if $\rho/(1 - \tau_r) = L_{ng} \alpha^{2(1-\alpha)}(1/\alpha - 1)/v_0$, then $v_0 < \eta$ would be the equilibrium price of existing patents and the economy would never grow. Summing up in a growing economy $v = \eta$ at all times.

A.3 Proof of Proposition 2

Combining (3), (8) and (12) with (21) we get:

$$\frac{\dot{N}}{N} = \left(1 - \alpha^2 - g + \frac{\phi(1-\sigma)(1 - \tau_w)(1 - \alpha)}{\phi'L} \right) \frac{Y}{\eta N}. \quad (A7)$$

Combining (12), (15) and (17), we find the following:

$$r = \alpha (1 - \alpha) \frac{Y}{\eta N}. \quad (A8)$$

Using (A8) and (22), (A7) becomes:

$$\frac{\dot{N}}{N} = \left(\frac{1}{\alpha} \left(1 - \frac{g}{1 - \alpha}\right) + 1\right) r + \frac{\phi(1-\sigma)(1 - \tau_w)r}{\phi'L}. \quad (A9)$$
Differentiating (3) with respect to time we obtain:

\[
\frac{\dot{C}}{C} = \frac{\dot{N}}{N} + \left(\frac{\phi'}{\phi} - \frac{\phi''}{\phi'}\right) \dot{L}.
\]  
(A10)

Plugging this expression for \(\dot{C}/C\) in (4) we obtain:

\[
\frac{\phi'}{\sigma} \dot{L} - \rho + r(1 - \tau_r) - \left(\frac{\phi'}{\phi} - \frac{\phi''}{\phi'}\right) \dot{L} = \frac{\dot{N}}{N}.
\]  
(A11)

Finally if we substitute in (A11) the expression for \(\dot{N}/N\) given by (A9) we obtain:

\[
\dot{L} = \frac{\rho - r(1 - \tau_r) + \sigma \left(\frac{1}{\alpha} \left(1 - \frac{g}{1 - \alpha}\right) + 1\right) r + \frac{\sigma(1 - \sigma)(1 - \tau_w) r}{\phi' \alpha L}}{\frac{\phi'}{\sigma} - \sigma \left(\frac{\phi'}{\phi} - \frac{\phi''}{\phi'}\right)}.
\]

Using (18) and considering that in equilibrium \(l = L\) we get (23) in the text.

### A.4 Proof of Proposition 3

From (3) and (14), in a BGP, i.e. when \(\dot{L} = 0\), \(C\) and \(N\) will grow at the same rate. From (4) this rate is seen to be given by (27).

In BGP condition (5) can be written as (28). Combining (28) and (A9) we get:

\[
0 > \left((1 - g/(1 - \alpha)) / \alpha + \tau_r\right) + \phi(1 - \sigma) (1 - \tau_w) / (\phi' \alpha L).
\]

From (22), \(\tau_r + (1 - g/(1 - \alpha)) / \alpha > (1 + \alpha \tau_r - (t + g)/(1 - \alpha)) / \alpha = (1 - \tau_w) / \alpha > 0\).

So the TVC implies \((\sigma - 1) \phi' / (\phi' L) > 1\).

### A.5 Proof of Proposition 4

The proof is divided into two parts. In the first part we prove that \(B'(\bar{L}) > 0\) implies uniqueness and determinacy of the BGP, with no transitional dynamics to it. In the second part we prove that if \(\sigma > 1\) or if \(\sigma < 1\) and \(\tau_w < 1 - \alpha + \alpha \tau_r/(1 - \sigma) + \sigma(1 - g/(1 - \alpha)) / (1 - \sigma)\) then \(B'(L) > 0\) for all \(L\), hence \(B'(\bar{L}) > 0\).

**First Part** Consider the phase diagram of (23) with \(B(L)\) on the vertical axis and \(L\) on the horizontal axis. We assume that at any point at which \(B(L)\) crosses the horizontal axis, i.e. when \(L\) is an element of \(\bar{L}\), where \(\bar{L}\) may be a vector of length one, then \(B'(\bar{L}) > 0\). As \(B(L)\) is a continuous function of \(L\) it cannot cross the horizontal axis from below twice in a row. This establishes the uniqueness of equilibrium given its existence and given \(B'(\bar{L}) > 0\).

To study the dynamic nature of a fixed point of (23), i.e. of BGP labor supply, we have to sign \(d\bar{L}/dL|_{L=\bar{L}}\). If this derivative is positive the fixed point \(\bar{L}\) is a repeller and the BGP is locally determinate. If the derivative is negative then \(\bar{L}\) is an attractor, i.e. there is local indeterminacy. \(A(L)\) as defined in (24), is always strictly positive for all values of \(L\), by the negative definiteness
condition of the hessian of the utility function (A3), so the differential equation (23) is defined for all values of \( L \) between 0 and 1. We have: \( \frac{dL}{dL}|_{L=\tilde{L}} = B'(\tilde{L})/A(\tilde{L}) - A'(\tilde{L})B(\tilde{L})/A^2(\tilde{L}) = B'(\tilde{L})/A(\tilde{L}) \) (since \( B(\tilde{L}) = 0 \)). So \( B'(\tilde{L}) > 0 \) implies \( \frac{dL}{dL}|_{L=\tilde{L}} > 0 \).

We have therefore established that if \( B'(\tilde{L}) > 0 \), the solution to (23) will be unique and unstable. Since for a given \( L \) the ratio between \( C \) and \( N \) is given, from (3) and (14), the economy will always be on a BGP.

**Second Part** Given the definition of \( B(L) \) in (25), taking the derivative we have:

\[
\frac{B'(L)}{C_1} = \frac{\sigma(1-\sigma)(1-\tau_w)}{\alpha} \left( 1 - \frac{\phi''}{\phi'^2} \right) + \tau_r - 1 + \frac{\sigma}{\alpha} \left( 1 - \frac{g}{1-\alpha} \right). \tag{A12}
\]

When \( \sigma > 1 \), the sign of \( 1 - \phi''/\phi'^2 \) and therefore of the first term on the right hand side (RHS) of (A12) is ambiguous. However, using (22) and rearranging we have:

\[
\frac{B'(L)}{C_1} = \frac{\sigma(1-\tau_w)}{\alpha} \left( 1 + (1-\sigma) \left( 1 - \frac{\phi''}{\phi'^2} \right) \right) + (\sigma - 1)(1-\tau_r) + \frac{\sigma t}{\alpha (1-\alpha)}. \tag{A13}
\]

We now prove the following:

\[ 1 + (1-\sigma) \left( 1 - \phi''/\phi'^2 \right) > 0. \tag{A14} \]

(A14) is self-evident if \( \sigma < 1 \), given (A1) and (A2). It is also obvious if \( \sigma > 1 \) and \( 1 - \phi''/\phi'^2 < 0 \), which is consistent with (A1) and (A2). If \( \sigma > 1 \) and \( 1 - \phi''/\phi'^2 > 0 \), which is also consistent (A1) and (A2), we use (A4) and write \( 1 - \phi''/\phi'^2 < 1 - (\sigma - 1)/\sigma = 1/\sigma \), so \( 1 + (1-\sigma) \left( 1 - \phi''/\phi'^2 \right) > 1 + (1-\sigma)/\sigma = 1/\sigma > 0 \).

In light of the just proved inequality (A14), the first term on the RHS of (A13) will be positive. So with \( \sigma > 1 \), determinacy is established.

Coming to the case of \( \sigma < 1 \), while the first term on the RHS of (A12) is certainly positive, the second term could be negative, making indeterminacy theoretically possible. However considering that by (A4) \( 1 - \phi''/\phi'^2 > 1/\sigma \), \( B'(L)/C_1 > (1-\sigma)(1-\tau_w)/\alpha + \tau_r - 1 + \sigma + \sigma (1-g/(1-\alpha))/\alpha \). From this we infer the condition for indeterminacy when \( \sigma < 1 \).

**B Appendix to Section 3**

**B.1 Solution of the Social Planner’s Problem**

The current value Hamiltonian for the social planner’s problem is:

\[
\frac{C_s^{1-\sigma} \phi(L_s)}{1-\sigma} + \frac{\mu}{\eta} \left( (1-g)L_s \int_0^{N_s} x_s(j)\alpha dj - C_s - \int_0^{N_s} x_s(j)\alpha di \right) \tag{B1}
\]
where \( \mu \) is the Lagrangian multiplier attached to the social budget constraint. The social planner decides on the optimal path of the control variables \( x_s(j) \), \( L_s \), \( C_s \), and of the state variable.

Optimality conditions with respect to the controls imply (31) and (32). Further necessary condition are:

\[
-\sigma \frac{\dot{C}_s}{C_s} + \frac{\phi'(L_s)}{\phi(L_s)} L_s = \frac{\dot{\mu}}{\mu} = \rho - \frac{1 - \alpha}{\eta} (1 - g) \frac{1}{1 - \alpha} \alpha \frac{1}{\gamma} L_s. \tag{B2}
\]

The resource constraint can then be expressed as:

\[
\frac{\dot{N}_s}{N_s} = \frac{Y_s(1 - g) - C_s - N_s x_s}{\eta N_s} = \frac{(1 - \alpha)(1 - g) \frac{1}{1 - \alpha} \alpha \frac{1}{\gamma} L_s \left(1 - \frac{(\sigma - 1)\phi(L_s)}{\sigma'(L_s)L_s}\right)}{\frac{1 - g}{\eta}} \tag{B3}
\]

where the second equality uses equations (31) and (32). Differentiating (32) we get

\[
\frac{\dot{C}_s}{C_s} = \frac{\dot{N}_s}{N_s} + L_s \left(1 - \frac{\phi''(\phi)}{\phi'^2}\right) \frac{\phi'}{\phi}. \tag{B4}
\]

Combining this with (B3) and with (B2) we get the following equation:

\[
\frac{1 - \alpha}{\sigma \eta} (1 - g) \frac{1}{1 - \alpha} \alpha \frac{1}{\gamma} L_s + \frac{\phi'(L_s)}{\sigma \phi(L_s)} L_s - \frac{\rho}{\sigma} = \frac{(1 - \alpha)(1 - g) \frac{1}{1 - \alpha} \alpha \frac{1}{\gamma} L_s \left(1 - \frac{\sigma - 1\phi(L_s)}{\sigma'(L_s)L_s}\right)}{\eta} + L_s \left(1 - \frac{\phi''(\phi)}{\phi'^2}\right) \frac{\phi'}{\phi}.
\]

From this equation, given (31), we see that neither the optimal \( L_s \) nor the optimal \( x_s \) depend on \( N_s \). The current value Hamiltonian maximised with respect to the control variables \( C_s \), \( L_s \) and \( x_s \) is therefore, in view of (32), a strictly concave function of \( N_s \). We deduce that the optimal solution is unique, so that \( L_s = 0 \) and \( L_s \) is given by (34).

From (B4) we deduce \( \dot{C}_s/C_s = \dot{N}_s/N_s = \gamma_s \) and from (B2) we get (36), which is in fact the RHS of (34).

### B.2 Proof of Proposition 5

Positive growth in the social planner’s economy, from the LHS of (34), requires:

\[
L_s - \frac{(\sigma - 1)\phi(L_s)}{\phi'(L_s)} > 0. \tag{B5}
\]

If we define the function \( f(L) = L - (\sigma - 1)\phi(L)/\phi'(L) \), \( f'(L) \) is given by the LHS of the inequality (A14) and is therefore strictly positive. We see by (29) and (B5) that \( f(L_s) > 0 > f(\bar{L}) \). Since \( f \) is a strictly increasing function of \( L \) we infer (35).
C Appendix to Section 4

C.1 Proof of Proposition 8

Given the equality in (28), (41) when we eliminate \( C(0) \) by using (3), (14) and (19) becomes:

\[
V = (\eta N(0))^{1-\sigma} \frac{\sigma-1}{\sigma} \left( \frac{C_1(1-\tau_r)}{\alpha} \right)^{1-\sigma} \phi^{2-\sigma} \frac{1}{r(1-\tau_r)(\sigma-1) + \rho}.
\]  

(C1)

Taking the logarithm of \( V \) in (C1) we get:

\[
\log \left( \frac{(1-\sigma)V}{1-\sigma} \right) = \log(\eta N(0)) + \log \left( \frac{\sigma-1}{\phi'} \right) + \log \left( \frac{C_1 (1-\tau_w)}{\alpha} \right) + \frac{2-\sigma}{1-\sigma} \log(\phi) - \frac{1}{1-\sigma} \log \left( \frac{r(1-\tau_r)(\sigma-1) + \rho}{\sigma} \right).
\]

The derivatives in (42) are then seen to be, recalling (18):

\[
\frac{\partial \log ((1-\sigma)V)}{(1-\sigma)\partial L} = -\frac{\phi''}{\phi'} + \frac{(2-\sigma)\phi'}{(1-\sigma)\phi} + \frac{r(1-\tau_r)/\dot{\bar{L}}}{r(1-\tau_r)(\sigma-1) + \rho},
\]

(C2)

\[
\frac{\partial \log ((1-\sigma)V)}{(1-\sigma)\partial \tau_r} = -\frac{r}{r(1-\tau_r)(\sigma-1) + \rho},
\]

(C3)

and

\[
\frac{\partial \log ((1-\sigma)V)}{(1-\sigma)\partial \tau_w} = -\frac{1}{1-\tau_w}.
\]

(C4)

From (22) \( d\tau_w/d\tau_r = -\alpha \). Substituting these derivatives in (42), using (18), and multiplying by \( B'/C_1 \) we arrive at (43) in the text.

To prove statement 1, consider the following. If \( d\gamma/d\tau_r > 0 \), given (27), by (40) the first term on the LHS of the inequality in (43) will be positive. We also have:

\[
\frac{\phi''}{\phi'} - \frac{(2-\sigma)\phi'}{(1-\sigma)\phi} = \frac{\phi'}{(\sigma-1)\phi} \left( 1 + (1-\sigma) \left( 1 - \frac{\phi''}{\phi'} \right) \right) > 0.
\]

(C5)

by (A14) and (A2). If \( d\gamma/d\tau_r > 0 \), then, by Proposition 7, \( dL/d\tau_r \) will be positive as well. In light of (39) and (C5), the second term on the LHS of the inequality in (43) will be positive. As the third term on the LHS of the inequality is always positive we deduce that \( d\gamma/d\tau_r > 0 \) is a sufficient condition for (43) to hold.

Statement 2 comes from basic calculus. Statement 3 is an immediate consequence of the previous two statements, by which a positive growth effect is inconsistent with a zero welfare effect.
D Appendix to Section 5

D.1 Proof of Proposition 9

(47) is just (A12) given (19) and (44). Positive growth, by (27) and (45) is equivalent to \( \rho/C_1 < b \). For \( \sigma > 1 \) just from visual inspection of (45) \( \tilde{L} \) is always lower than one while positive growth ensures the TVC is respected, by (28), so there are no related restrictions to be imposed. For \( \sigma < 1 \), from (45), \( \tilde{L} \) less than one requires \( 1 - \sigma - \tau_r - \sigma (1 - g/(1 - \alpha)) /\alpha < \rho/C_1 \) while the (28) requires:

\[
\frac{(\sigma - 1)(1 - \tau_w)(1 - \tau_r)(1 - \sigma)}{(\sigma - 1) (1 - \tau_w) + (\chi - 1) \left( \alpha \tau_r + 1 - \frac{g}{1 - \alpha} \right)} < \frac{\rho}{C_1}.
\]

We know \( \gamma_s > \gamma \), and \( L_s > \tilde{L} \) so there are no further restrictions for their positivity, while \( L_s < 1 \), from (46) and remembering from (44) that \( \tau_r > 0 \), implies: \( 1 - \sigma < \rho \eta (1 - g)^{1/(\alpha - 1)} \alpha^{\alpha/(\alpha - 1)}/(1 - \alpha), \) which is the last inequality in the proposition.

D.2 Proof of Proposition 10

We assume \( \sigma < 1 \). To prove (49) it is enough to show that

\[
\tilde{L} \geq \left( \frac{(\chi - 1)}{\sigma(\sigma - 1)} + 1 \right)^{-1}
\]

is impossible. (D1) can be rewritten, given (45), as:

\[
1 - \rho \eta (1 - \alpha)^{-1} \alpha^{\frac{2\alpha}{1 - \alpha}} \frac{(\chi - 1)}{(1 - \tau_w) \sigma(\sigma - 1)} \geq \frac{1}{\frac{(\chi - 1)}{\sigma(\sigma - 1)} + 1},
\]

or, given that from (44) and (47) the denominators of both fractions are positive, as:

\[
1 - \rho \eta (1 - \alpha)^{-1} \alpha^{\frac{2\alpha}{1 - \alpha}} \frac{(\chi - 1)}{(1 - \tau_w) \sigma(\sigma - 1)} \left( \frac{\chi - 1}{1 - \tau_r - 1 + \sigma + \frac{\alpha}{\alpha} \left( 1 - \frac{g}{1 - \sigma} \right) \alpha} \right) \geq \frac{1}{\frac{(\chi - 1)}{\sigma(\sigma - 1)} + 1}.
\]

This expression, simplifying, reordering and using (22) to eliminate \( \tau_w \) and \( \tau_r \), becomes:

\[
\sigma(\sigma - 1) \left( 1 + \frac{\alpha - g}{1 - \alpha} \right) (1 - \sigma) - \frac{t + \rho \eta \alpha^{\frac{2\alpha}{1 - \alpha}}}{1 - \alpha} \leq \rho \eta \alpha^{\frac{2\alpha}{1 - \alpha}} (\chi - 1).
\]

As \( \chi < 1 \), if \( (1 + g/(1 - \alpha)) (1 - \sigma) - (t + \rho \eta \alpha^{2\alpha/(\alpha - 1)})/(1 - \alpha) \leq 0 \), this can never be true. A necessary condition for this to hold would then be: \( (1 + g/(1 - \alpha)) (1 - \sigma) > (t + \rho \eta \alpha^{2\alpha/(\alpha - 1)})/(1 - \alpha) \). However from
we know \(1 - \sigma < \rho \eta (1 - g)^{1/(\alpha - 1)} \alpha^{\alpha/(\alpha - 1)}/(1 - \alpha)\), so a necessary condition for (D1) to hold would be:
\[
\rho \eta (1 - g)^{1/(\alpha - 1)} \alpha^{\alpha/(\alpha - 1)} (1 + \alpha - g/(1 - \alpha)) > t + \rho \eta \alpha^{2\alpha/(\alpha - 1)},
\]
and, a fortiori, \(1 + \alpha - g/(1 - \alpha) > ((1 - g)/\alpha^\alpha)^{1/(1 - \alpha)}\). This is never possible for \(g\) and \(\alpha\) belonging to the open interval \((0, 1)\).

D.3 Derivative of \((\gamma_s / \gamma) |_{L_s = 1}\) with respect to \(\alpha\).

From (18), (19), (27), (36) and (37), \((\gamma_s / \gamma) |_{L_s = 1} = (C_{1s} - \rho \eta) / (C_1 - \rho \eta)\) and \((\gamma_s - \gamma) |_{L_s = 1} = (C_{1s} - C_1) / \sigma\). We now prove that \((\gamma_s - \gamma) |_{L_s = 1}\) is decreasing in \(\alpha\), and that, as a consequence, \((\gamma_s / \gamma) |_{L_s = 1}\) is decreasing in \(\alpha\) as well. We note:

\[
\frac{d\alpha}{d\alpha} = \frac{\alpha - \alpha \log \frac{1}{1 - \alpha}}{(1 - \alpha)} < 0 \quad \text{and} \quad \frac{d\alpha}{d\alpha} = \frac{\alpha - 1 + \log \alpha}{(1 - \alpha)^2} > 0.
\]

From (38), \((\gamma_s - \gamma) |_{L_s = 1}\), as the product of three positive functions, \(\alpha^{\alpha/(1 - \alpha)}\), \((1 - \alpha)\) and \((1 - g)^{1/(1 - \alpha)} - (1 - \tau_r) \alpha^{1/(1 - \alpha)}\), all decreasing in \(\alpha\), is therefore itself decreasing in \(\alpha\).

From (36) and (37), \((\gamma_{sa} |_{L_s = 1}\) signs as negative as \(d \left( (1 - \alpha) \alpha^{\alpha/(1 - \alpha)} \right) / d\alpha = (\alpha^{\alpha/(1 - \alpha)} \log \alpha) / (1 - \alpha)\). \(\partial (\gamma_s / \gamma) |_{L_s = 1} / \partial \alpha = ((\gamma_{sa} - \gamma_{a} \gamma_{s}) / 2) |_{L_s = 1}\) signs as \((\gamma_{sa} - \gamma_{a} \gamma_{s}) |_{L_s = 1} < (\gamma_{sa} - \gamma_{a} \gamma_{s}) |_{L_s = 1} = (\gamma_{s} (\gamma_{sa} - \gamma_{a}) |_{L_s = 1}\), with the inequality deriving from \((\gamma - \gamma_s) |_{L_s = 1} < 0\) and \(\gamma_{sa} |_{L_s = 1} < 0\). As \(\gamma_s > 0\), \(\partial (\gamma_s - \gamma) |_{L_s = 1} / \partial \alpha < 0 \to \partial (\gamma_s / \gamma) |_{L_s = 1} / \partial \alpha < 0\).

D.4 Derivative of \(L_s / \tilde{L}\) with respect to \(\alpha\).

From respectively (45) and (46) we have:

\[
\tilde{L} = \frac{L_n}{L_d}, \quad L_n \equiv 1 - \frac{\rho \eta \alpha^{-\tau_r} (\chi - 1)}{\eta (1 - \alpha) (1 - \tau_w) \sigma (\sigma - 1)} \quad \text{and} \quad (D3)
\]

\[
L_d = 1 + \frac{\alpha (\chi - 1) \tau_r - (1 + \sigma (1 + (1 - g)/(1 - \alpha)) / \alpha)}{\sigma (\sigma - 1)} \frac{1}{(1 - \tau_w)};
\]

\[
L_s = \frac{L_{sn}}{L_{sd}}, \quad L_{sn} \equiv 1 - \frac{\rho \eta (1 - g)^{\alpha - \tau_r} \alpha^{-\tau_r} (\chi - 1)}{(1 - \alpha) \sigma (\sigma - 1)} \quad \text{and} \quad L_{sd} \equiv 1 + \frac{(\chi - 1)}{\sigma}. \quad (D4)
\]

For simplicity, we omit policy variables from the calculations, as they are not essential for our purpose:

\[
\frac{L_s}{\tilde{L}} = \frac{L_d L_{sna} + L_{sn} L_d a}{L_n L_{sd}} - \frac{L_d L_{sn} L_{na}}{(L_n)^2 L_{sd}} > \frac{L_d (L_{sn} - L_n) a + L_{sn} L_d a}{L_n L_{sd}}
\]

\[
\frac{L_{sn}}{(L_n L_{sd} \sigma (\sigma - 1))} \left( \frac{d \left( \alpha^{-\tau_r} (\alpha^{-\tau_r} - 1) \right)}{da} + (\sigma - 1) L_{sn} \right),
\]

In what follows the notation \(x_u\) is sometimes used instead of \(\partial x/\partial a\).
The inequality obtains because $L_{sn}/L_n > 1$, while $-L_{na}$ from (D3) signs as the derivative of $\alpha^{2/\alpha/(\alpha-1)} / (1-\alpha)$, which is positive. In fact $\alpha^{2/\alpha/(\alpha-1)} / (1-\alpha)$ is the product of three positive functions, $\alpha^{\alpha/(\alpha-1)}$, $\alpha^{\alpha/(\alpha-1)}$, and $(1-\alpha)^{-1}$, all increasing in $\alpha$ so it is increasing in $\alpha$. In fact from (D2) $d\alpha^{\alpha/(\alpha-1)} / d\alpha = -\alpha^{2/\alpha/(\alpha-1)} (d\alpha^{\alpha/(\alpha-1)} / d\alpha) > 0$. Analogously $\alpha^{\alpha/(\alpha-1)} (\alpha^{\alpha/(\alpha-1)} - 1) / (1-\alpha)$ is the product of three positive functions increasing in $\alpha$, so if $1 < \sigma$ this is enough to prove that $(L_s^*) > 0$. If $\sigma < 1$ we write:

$$\frac{L_{d-p\eta d}}{1-\alpha} \left( \alpha^{\alpha/(\alpha-1)} (\alpha^{\alpha/(\alpha-1)} - 1) \right) + (\sigma - 1) L_{sn} > 0.$$ 

For the first inequality we have used the relation $L_{sn} < L_d$, which is shown to be always true further down. The second inequality comes from using the restriction (48), which can be written: $\rho \eta \alpha^{\alpha/(\alpha-1)} / (1-\alpha) > 1 - \sigma$, then calculating the derivative and simplifying. We are only left to show that $L_{sn} < L_d$. We have:

$$L_{sn} = 1 - \frac{\rho \eta \alpha^{\alpha/(\alpha-1)} (\chi - 1)}{(1-\alpha) \sigma (\sigma - 1)} < 1 + \frac{\chi - 1}{\sigma} < 1 + \frac{(\chi - 1) \alpha}{\sigma} + \frac{\chi - 1}{\sigma - 1} = L_d.$$ 

The first inequality comes from (48). The second inequality is immediate from (44), remembering we are considering the case when $\sigma < 1$.

### D.5 Disentangling the effects of IES and $F_R$

To see what effect a changing elasticity of intertemporal substitution has for a given Frisch elasticity we proceed as follows. For a $\sigma$ in $(1, 3]$ we write (51) as $\chi = 1 + \sigma (1 - L) / (F_{rb} L) - \sigma$, where $F_{rb} = 2.87$, the benchmark value. We then plug this expression for $\chi$ in (45), where all the parameters but $\sigma$ are given the values in Table 1. We obtain the following quadratic in $L$: $A_0 L^2 + A_1 L + A_2 = 0$, where

$$A_0 \equiv 1 - \left( \alpha + \frac{\tau_w}{\alpha - 1} \frac{1}{1-\alpha} + \frac{g}{\eta (1-\alpha)^{-1}} \frac{1}{\eta \alpha^{\alpha/(\alpha-1)}} \right) \frac{F_{rb}}{1-\tau_w},$$

$$A_1 \equiv \frac{\alpha + \frac{\tau_w}{\alpha - 1} \frac{1}{1-\alpha} - \frac{g}{\eta (1-\alpha)^{-1}} \frac{1}{\eta \alpha^{\alpha/(\alpha-1)}}}{F_{rb} (1-\tau_w)} - 1 - \frac{\rho \eta (1-\alpha)^{-1} \alpha^{\alpha/(\alpha-1)}}{\sigma - 1} \frac{1}{\eta \alpha^{\alpha/(\alpha-1)}} \frac{1}{1-\tau_w},$$

$$A_2 \equiv \frac{\rho \eta (1-\alpha)^{-1} \alpha^{\alpha/(\alpha-1)}}{F_{rb} (\sigma - 1) (1-\tau_w)}.$$ 

The positive root of the quadratic gives us the new initial equilibrium $L_b$, consistent with the chosen $\sigma$ and $F_{rb}$. Substituting this value of $L_b$ back in (51)
we find the value of $\chi$ consistent with the given $\sigma$ and with $F_{RB}$. We then solve (50), for the given $\sigma$ and the $\chi$ we have so found. The procedure is repeated for $\sigma$ varying in $[1, 3]$. To see what effect a changing Frisch elasticity has for a given elasticity of intertemporal substitution, we follow an analogous procedure but keep $\sigma$ fixed at 1.5 and let $F_R$ vary in the range $[0.87, 4.87]$.

E  Appendix to Section 6

E.1 A Subsidy to R&D

The equilibrium conditions for households and firms in the final sector stay the same so all the equations till (8) still hold, as well as their combinations (A10) and (A11). As regards the intermediate sector, all equations till (16) also hold. It is easy to show, following the same reasoning as in the baseline case, that the value of a blueprint $v$ will be time invariant. However $r = \frac{\pi}{\eta}(1 - s_N)$ replaces (17), so that

$$r = C_2 L, \quad \text{where } C_2 = C_1/(1 - s_N) \quad (E1)$$

replaces (18). The government budget constraint is no longer (20) but:

$$wL + rF = (g + t)Y + \eta s_N \dot{N}. \quad (E2)$$

Other conditions that will hold in general equilibrium are $L = l$, (21), and therefore (A7), and $F = (1 - s_N)\eta N$. (27), (28), and (41) also continue to hold. From (12), (15) and (E1) we obtain:

$$\rho \eta (1 - s_N)/ (\alpha(1 - \alpha)) = Y/N \quad (E3)$$

and, combining (8), (E2) and (E3):

$$\tau_w = -r \alpha + (g + t)/(1 - \alpha) + \eta s_N \dot{N}/ ((1 - \alpha)Y).$$

Plugging this into (A7) leads to:

$$\frac{\dot{N}}{N} = \frac{1 - g - \alpha^2 + \frac{\phi}{\sigma} (1 - \sigma) \left(1 + r \alpha - \frac{g + t}{1 - \alpha} \right) (1 - \alpha)}{N \eta \left(1 + \frac{\phi}{\sigma} (1 - \sigma) \frac{s_N}{L} \right)} Y.$$

Equating this with (A11) we get: $\dot{L} = B_1(L)/A(L)$, where $A(L)$ is defined in (24) and

$$B_1(L) \equiv \frac{1 - g - \alpha^2 + \frac{\phi}{\sigma} \left(1 + r \alpha - \frac{g + t}{1 - \alpha} \right) (1 - \alpha)}{N \eta \left(1 + \frac{\phi}{\sigma} (1 - \sigma) \frac{s_N}{L} \right)} Y - \frac{r(1 - \tau_r) - \rho}{\sigma},$$

or, given (E3) and specifying utility as in (44):

$$B_1(L) \equiv \frac{\chi(1 - s_N) \tau}{\alpha} \left(1 - g - \alpha^2 + \frac{\phi}{\sigma} \left(1 - \sigma \right) \frac{(1 - \sigma)(1 - \alpha)}{L} \left(1 + r \alpha - \frac{g + t}{1 - \alpha} \right) \right) \left(1 + \frac{s_N (1 - \sigma)(1 - L)}{(1 - \sigma) L} \right) - \frac{r(1 - \tau_r) - \rho}{\sigma}. \quad (45)$$
Since \( A(L) > 0 \), if \( B'_1(L) > 0 \) we have determinacy, where:

\[
\frac{B'_1(L)}{C_2} = \frac{(1-s_N)}{\alpha} \left( \frac{1-g-\alpha^2}{1-\alpha} + \frac{(1+\tau_r\alpha-\frac{\sigma+1}{\sigma})(\sigma-1)}{\chi-1} \right) + \frac{(r(1-\tau_r)-\rho)\sigma_N(1-\sigma)}{\sigma C_2(1-\sigma)L^2} \frac{1-\tau_r}{\sigma}.
\]

Under determinacy, for reasons analogous to those seen in the general case, the economy will always follow the unique BGP. In (4) \( L \) will be a constant so (27) will hold. Equating \( B_1(L) \) to zero gives us the quadratic:

\[
\tilde{A}_0L^2 + \tilde{A}_1L + \tilde{A}_2 = 0,
\]

with

\[
\tilde{A}_0 = \frac{(1-s_N)}{\alpha} \left( \frac{1-g-\alpha^2}{1-\alpha} + \frac{(\sigma-1)(1+\alpha\tau_r-\frac{\sigma+1}{\sigma})}{(\chi-1)} \right) + \frac{(\tau_r-1)\sigma}{(\chi-1)s_N},
\]

\[
\tilde{A}_1 = \rho \left( \frac{1+\frac{(\sigma-1)}{(\chi-1)}s_N}{\sigma C_2} \right) + \frac{(\sigma-1)}{(\chi-1)} \left( \frac{(1-\tau_r)s_N}{\sigma} - \frac{(\sigma-1)}{\alpha} \right),
\]

\[
\tilde{A}_2 = \frac{(1-\sigma)\sigma_N\rho}{(\chi-1)\sigma C_2}.
\]

The positive solution to (E4) is equilibrium \( \tilde{L} \). We now come to the calculation of the optimal taxes. By eliminating \( C(0) \) using (3), (8), (E3) and (E1), and using (27), (41) can still be formulated as (C1). To compute \( \tau_{ro} \) we still equate to zero (42) with (C3), (C2) and (C4) still valid. Differentiating (E4) with respect to \( L \) and \( \tau_r \) we get:

\[
\frac{d\tilde{L}}{d\tau_r} = -\frac{1+\frac{(\sigma-1)}{(\chi-1)}s_N(1-\sigma)}{\chi-1} \frac{\tilde{L}}{\sigma} + \frac{(\sigma-1)}{(\chi-1)} \left( \frac{1+\frac{(1-s_N)}{\sigma}}{\sigma} \right).
\]

Finally differentiating the government budget constraint, considering that there is no transitional dynamics so that \( N = N\gamma \) we obtain: \( d\tau_w/d\tau_r = -\alpha - s_N\alpha/\sigma(1-s_N) + \rho s_N\alpha/(\sigma L^2(1-s_N)) \) \( (dL/d\tau_r) \). So given (44) the welfare effect is:

\[
d\log((1-\sigma)V)/(1-\sigma)d\tau_r = \frac{\alpha}{1-\tau_w} \left( 1 + \frac{s_N}{\sigma(1-s_N)} \right) - \frac{r}{r(1-\tau_r)(\sigma-1) + \rho} \left( \frac{\sigma + \gamma - 2}{(\sigma-1)(1-\tilde{L})} - \frac{C_2(1-\tau_r)}{r(1-\tau_r)} + \frac{\rho s_N\alpha}{\sigma(1-s_N)} \right) \frac{d\tilde{L}}{d\tau_r}.
\]

As regards calibration, if we want to keep the initial ratio of government expenditure to GDP constant at 0.1 the new initial ratio between unproductive expenditures and final production, \( g \) (which is fixed at 0.1(1-\alpha^2) in the baseline case), is determined as follows: \( g = 0.1(1-\alpha^2) - \eta s_N\gamma_0\alpha^{2\alpha/(\alpha-1)}/L_0 \), from \( gY = 0.1(1-\alpha^2)Y - \eta s_N\gamma_0\tilde{N} \).
E.2 G in the Production Function

The equations pertaining to households (1-5) and therefore (A10) and (A11) still hold. As regards firms in the final sector, given (53), the optimal choice for labor is described by (8) and the demand for intermediate $j$ by:

$$p(j) = G^{\frac{1-\beta}{\alpha}} L^{\frac{1-\alpha}{\beta}} Y^{\frac{\beta-1}{\alpha}} \alpha x(j)^{\frac{1}{\alpha}}.$$  \hfill (E5)

The price chosen for intermediates is then:

$$p(j) = p = \frac{\beta}{\alpha}. \hfill (E6)$$

$\beta \geq \alpha$ is immediately seen to be a necessary condition for non-negative monopoly profits, given the marginal cost of one for intermediates. By combining (53), (E5) and (E6) we obtain:

$$x(j) = x = \left( L^{1-\alpha} \alpha^{2\beta} \beta^{-\beta} g^{1-\beta} \right)^{\frac{1}{1-\alpha}} \hfill (E7)$$

and

$$Y = \frac{\beta}{\alpha^2} x N. \hfill (E8)$$

The marginal productivity of labor (MPL) in the aggregate is obtained by taking the derivative of $Y$ with respect to $L$. Given (8), (E7) and (E8) the ratio of the wage to MPL is given by $\beta - \alpha$. So the positive externality to labor is decreasing in $\beta$. Profits are given, as a consequence of (E6), by:

$$\pi = \left( \frac{\beta}{\alpha} - 1 \right) x. \hfill (E9)$$

(9) and (16) still hold and, as the value of a new firm will be constant at $\eta$ (the proof being strictly analogous to that given for the baseline model) (17) also holds. Combining it with (E8) and (E9) we get:

$$r = \frac{\alpha}{\eta} \left( 1 - \frac{\alpha}{\beta} \right) \frac{Y}{N}. \hfill (E10)$$

The marginal productivity of $N$ from (E8) is $\beta x / \alpha^2$. The ratio of the interest rate to $\beta x / \alpha^2$ is $\alpha \left( 1 - \alpha / \beta \right) / \eta$. So the positive externality to accumulating $N$ is decreasing in $\beta$. We can also write:

$$r = C_3 L^{\frac{1-\alpha}{\beta}} \hfill (E11)$$

with $C_3 \equiv (\beta - \alpha) \left( \alpha^{\alpha+\beta} \beta^{-\beta} g^{1-\beta} \right)^{1/(\beta-\alpha)} / \eta$. The government budget constraint is still given by (20). Other equilibrium conditions are that all market clear so that $F = \eta N$, $L = l$ and (21) holds. Substituting in (20), $rL$ and $rF = r\eta N$ with their expressions from (8) and (E10) we can write:
\[
\tau_w = \frac{t + g}{1 - \alpha} - \frac{\alpha \left(1 - \frac{\beta}{\gamma}\right)}{1 - \alpha} \tau_r. \tag{E12}
\]

Plugging in (21) the expression for \(C\) obtained using (3) and (8) and eliminating \(x\) using (E8) we get:

\[
\frac{\dot{N}}{N} = \frac{Y}{\eta N} \left(1 - g - \frac{\alpha^2}{\beta} + \frac{(1 - \sigma) (1 - \tau_w) (1 - \alpha) \phi}{\phi'L}\right).
\]

Equating this with the expression for \(\dot{N}/N\) in (A11) we get: \(\dot{L} = B_2(L)/A(L)\) where \(A(L)\) is defined in (24) and

\[
B_2(L) \equiv \rho - r(1 - \tau_r) + \frac{\sigma Y}{\eta N} \left(1 - g - \frac{\alpha^2}{\beta} + \frac{(1 - \sigma) (1 - \tau_w) (1 - \alpha) \phi}{\phi'L}\right).
\]

Using the specification (44) as well as (E10) we have:

\[
B_2(L) = \rho + \frac{r \sigma \left(1 - g - \frac{\alpha^2}{\beta} - \frac{(1 - \sigma)(1 - \tau_w)(1 - \alpha)}{(\chi - 1)L} \left(1 - \frac{1}{\beta - \alpha}\right)\right)}{\alpha - \frac{\alpha^2}{\beta}} + r (\tau_r - 1). \tag{E13}
\]

If the equilibrium is determinate it is unique, i.e. if \(B_2'\) \(\dot{L}\) > 0, \(B_2 = 0\) at all times, as in the baseline case. We have:

\[
\frac{\dot{L} B_2'(\dot{L})}{r} = \frac{\sigma \left(t + (1 - \alpha)(1 - \tau_w) \left(\frac{\sigma + \chi - 2}{\chi - 1} - \frac{1 - \gamma}{\beta - \alpha} \left(\frac{(\sigma - 1)(1 - \dot{L})}{(\chi - 1)L} - 1\right)\right)\right)}{\alpha - \frac{\alpha^2}{\beta}}
+ \left(\sigma - \frac{1 - \alpha}{\beta - \alpha}\right) (1 - \tau_r).
\]

Given (E11), \(B_2(L) = 0\) has to be solved numerically to find \(\dot{L}\). The welfare calculations are strictly analogous to those for the baseline case. As the economy is always in steady state (27) and (28) hold. To express (41) in terms of labor and the taxes we combine (3), (8), (E10) and (E11) to eliminate \(C(0)\) and use the equality in (28). We arrive to:

\[
V = \sigma(\eta N(0))^{1 - \sigma} \sigma \left(\frac{\sigma - 1 - \beta (1 - \alpha)(1 - \tau_w) C_2 L^{1 - \beta / \alpha}}{(\chi - 1) \alpha (\beta - \alpha)}\right)^{1 - \sigma} \frac{1 - \sigma}{r(1 - \tau_r)(\sigma - 1) + \rho} \left(1 - \dot{L}\right)^{2 - \sigma - \chi}.
\]

The optimal \(\tau_r\) corresponds to the value equating \(d \log ((1 - \sigma)V) / ((1 - \sigma)d \tau_r)\) as defined in (42) to zero. The derivatives in the expression given (44) are as follows:

\[
\frac{\partial \log ((1 - \sigma)V)}{(1 - \sigma)\partial L} = \frac{1 - \beta}{(\beta - \alpha) L} - \frac{2 - \sigma - \chi}{(1 - \sigma)(1 - \dot{L})} + \frac{(1 - \alpha)(1 - \tau_r) C_2 L^{1 - \beta / \alpha}}{(\beta - \alpha)(r(1 - \tau_r)(\sigma - 1) + \rho)}.
\]

Totally differentiating (E13), since from (E12) \(d \tau_w/d \tau_r = -\alpha (1 - \alpha/\beta) / (1 - \alpha)\), we get:
By combining these expressions, (C4) and (C3), \( \tau_{ro} \) will solve (after expressing \( \tau_w \) in terms of \( \tau_r \) using E12 and considering that \( B_2(\tilde{L}) = 0 \)):

\[
\frac{d\tilde{L}}{d\tau_r} = \frac{(\beta-a)\left(\frac{\sigma(\sigma-1)(1-\tilde{L})}{\chi-1} - \tilde{L}\right)}{\alpha - \frac{\sigma}{\alpha^2} \left(1 - \frac{a^2}{\beta} - g + \frac{(1-\tau_w)(\sigma-1)}{\chi-1} \left(1 - \alpha - \frac{1-a}{L}\right)\right) + \tau_r - 1}.
\]

E.3 \( G \) in the Utility Function

Given (54), (A1) and (A2) are still needed, while strict concavity of instantaneous felicity imposes (A3) and

\[
\left(1 - \frac{1}{(1-\sigma)(1-\theta)}\right)\phi'' - \phi'^2 > 0. \tag{E14}
\]

Consumers face the instantaneous budget constraint (2). Marginal rate of substitution between leisure and consumption under optimization equals their relative price:

\[
\frac{\phi'}{\phi} = \frac{w(1-\tau_w)(\sigma-1)(1-\theta)}{C}. \tag{E15}
\]

The intertemporal condition for optimal consumption and leisure is now:

\[
((1-\theta)(1-\sigma) - 1) \frac{\dot{C}}{C} + \theta(1-\sigma) \frac{\dot{G}}{G} + \frac{\phi'}{\phi} \dot{L} = \frac{\dot{\lambda}}{\lambda} = \rho - r(1-\tau_r) \tag{E16}
\]

where \( \lambda = (1-\theta)C(1-\theta)(1-\sigma) - 1G^{\theta(1-\sigma)} \phi \) is the shadow value of wealth. (5) is still necessary. The assumptions on the productive sector and the government stay the same so equations from (6) to (22) still hold along with the market clearing conditions for labor (\( L = l \)), for wealth (\( F = \eta N \)), and for the final good (21).

To get the equilibrium value of \( L \), we plug in (21) the expression for \( C \) obtained using (E15) and (8) and the expression for \( x \) in (12) thus arriving to:

\[
\frac{N}{\dot{N}} = \frac{r}{\alpha} \left(1 + \alpha - \frac{g}{1-\alpha} - (1-\theta)(1-\tau_w)\frac{(\sigma-1)\phi}{\phi'} L\right). \tag{E17}
\]

Differentiating (E15) with respect to time we still obtain (A10). Plugging (A10) for \( \frac{\dot{C}}{C} \) in (E16) we get:

\[
\frac{N}{\dot{N}} = \frac{\left(\frac{\phi'}{\sigma} + ((1-\theta)(1-\sigma) - 1)\left(\frac{\phi'}{\sigma} - \phi''\right)\right)}{\sigma} \tilde{L} - \rho + r(1-\tau_r) \tag{E18}
\]
Substituting in (E18) the expression for \( \dot{N}/N \) given by (E17) we obtain
\[
\dot{L} = \frac{B_3(L)}{A_3(L)}, \quad \text{with} \quad A_3(L) \equiv \frac{(1 - \theta)(1 - \sigma)}{\phi} + \frac{(1 - (1 - \theta)(1 - \sigma))}{\phi'} \quad (E19)
\]
and 
\[
B_3(L) \equiv \rho + r \left( \frac{\sigma}{\alpha} \left( 1 + \alpha - \frac{g}{1 - \alpha} - \frac{(\sigma - 1)\phi(1 - \tau_w)(1 - \theta)}{\phi'L} \right) + \tau_r - 1 \right). \quad (E20)
\]

\( A_3(L) \) is positive by (A2) and (E14). Using (44), (18) and (19) we derive from (E20) the condition for determinacy (which implies uniqueness of the solution):
\[
B_3(L) \equiv C_1 \left( \frac{\sigma}{\alpha} \left( 1 + \alpha - \frac{g}{1 - \alpha} + \frac{(\sigma - 1)(1 - \tau_w)}{(\chi - 1)} + \tau_r - 1 \right) \right) > 0.
\]

Notice if \( \sigma > 1 \), the case we focus on, indeterminacy is impossible if \( g/(1 - \alpha) > 0 \) as always assumed. Equating \( B_3(L) = 0 \) gives us equilibrium \( \bar{L} \).

We also calculate the Frisch elasticity given (54).
\[
\bar{L} = \frac{\frac{\sigma}{\alpha} \left( \frac{(\sigma - 1)}{\alpha (\chi - 1)} (1 - \tau_w)(1 - \theta) - \frac{\rho}{C_1} \right)}{\frac{\sigma}{\alpha} \left( 1 + \alpha - \frac{g}{1 - \alpha} + \frac{(\sigma - 1)}{(\chi - 1)} (1 - \tau_w)(1 - \theta) \right) + \tau_r - 1}.
\]

\( \bar{L} \) is decreasing in \( \theta \), so the effect of \( \theta \) on \( F_R \) ambiguous (negative directly but positive through a lower \( \bar{L} \)). As \( \bar{L} = 0 \) at all times and \( G = gY \), we know from (21) and (12) that \( C, N \) and \( G \) grow at the same rate, which by (E16), is as in (27).

The welfare calculations are strictly analogous to those for the baseline case. First we integrate (54), given (44), then eliminate \( C(0) \) using (E15) and (14), and \( G(0) \) using (12) and (11), finally we use the equality in (28). We arrive to:
\[
V = \frac{(\frac{\eta N(0)C_1}{\alpha})^{1-\sigma} \left( \frac{(1 - \theta)(\sigma - 1)(1 - \tau_w)}{(\chi - 1)} \right)^{(1 - \theta)(1 - \sigma)} \left( \frac{a\bar{L}}{1 - \alpha} \right)^{\theta(1 - \sigma)}}{(1 - \sigma) \left( \frac{\tau(1 - \tau_r)(\sigma - 1) + \rho}{\sigma} \right) \left( 1 - \bar{L} \right)^{\sigma + \chi - 2 + \theta(1 - \sigma)}}. \quad (E22)
\]

The optimal \( \tau_r \) corresponds to the value equating \( d \log ((1 - \sigma)V)/(1 - \sigma)d\tau_r \) as defined in (42) to zero. (C3) and \( d\tau_r/d\tau_r = -\alpha \) are still valid, while, using (18), we have:
\[
\frac{\partial \log ((1 - \sigma)V)}{(1 - \sigma)\partial \bar{L}} = \frac{\sigma + \chi - 2 + \theta(1 - \sigma)}{(1 - \sigma) \left( 1 - \bar{L} \right)} + \frac{\theta}{\bar{L}} + \frac{C_1(1 - \tau_r)}{C_1 \bar{L}(1 - \tau_r)(\sigma - 1) + \rho}.
\]

Also, \( \partial \log ((1 - \sigma)V)/(1 - \sigma) \partial \tau_w = -(1 - \theta)/(1 - \tau_w) \). Finally, from (E21), by using (22) we get:
\[
\frac{d \bar{L}}{d\tau_r} = \frac{(1 - \theta) \frac{(\sigma - 1)(1 - \bar{L}) - \bar{L}}{\chi - 1} - \bar{L}}{1 + \frac{1}{\alpha} \left( \frac{1 - \sigma}{\chi - 1} \right) + \frac{(1 - \theta)}{\alpha} \left( 1 + \alpha \tau_r - \frac{\eta N(0)}{C_1} \frac{\sigma - 1}{\chi - 1} \right) \frac{(\sigma - 1)}{(\chi - 1)} + \frac{\tau_r - 1}{\sigma}}.
\]
Putting all this together the optimal tax on capital solves:

\[
\frac{\alpha (1 - \theta)}{1 - \tau_w} - \frac{r}{\sigma (1 - \tau_r) (\sigma - 1) + \rho} + \left( \frac{\sigma + \chi - 2}{(1 - \sigma) (1 - \bar{L})} + \frac{C_1 (1 - \tau_r)}{C_1 \bar{L} (1 - \tau_r) (\sigma - 1) + \rho} + \frac{\theta}{(1 - \bar{L}) \bar{L}} \right) \frac{d\bar{L}}{d\tau_r} = 0.
\]

E.4 Production Function Concave in Labor

Given (55), the profit maximizing conditions in the final good sector are

\[ w = \frac{Y}{L}, \quad p_P = \frac{(P^{1 - \alpha} - \kappa L^\kappa / p(j))^{1/(1 - \alpha)}}{1 - \alpha - \kappa} \]

Without loss of generality we fix \( P = 1 \). Given the demand curve for the intermediate sector firm \( j \), the optimal price chosen is still \( \frac{1}{2} \), while the optimal quantity is \( x(j) = \frac{1}{2} L \). We then have:

\[ Y = \left( \frac{1}{1 - \alpha - \kappa} \right) N. \]

Profits in the intermediate sector are \( \pi N = \frac{(1 - \alpha - \kappa) Y}{\alpha} \). In equilibrium \( \pi N = \eta \gamma N \), and \( r = C_1 L^\kappa (1 - \alpha - \kappa) \), where \( C_1 \) is the same as in (19). Equations describing households behavior (1-5) and their implications (A10) and (A11) still hold. Using (21), (3) and the expressions derived for \( w \) and \( Y \), while assuming the functional form for \( \phi \) in (44), we find:

\[ \frac{\dot{N}}{N} = \frac{C_1 L \tau_w}{\alpha (1 - \alpha)} \left( 1 - g - \alpha^2 - \kappa \frac{(1 - \tau_w) (1 - \sigma)}{(1 - \chi) L} \right). \]

Equating this to (A11), given (44) we get: \( (\chi + \sigma - 1) / (1 - L) \sigma \dot{L} = B_4(L) \), with

\[ B_4(L) = \frac{C_1 L \tau_w}{\alpha (1 - \alpha)} \left( 1 - g - \alpha^2 - \frac{\kappa (1 - \tau_w) (1 - \sigma)}{(1 - \chi) L} - \frac{\alpha (1 - \alpha - \kappa)}{\sigma} \right) + \frac{\rho}{\sigma}. \]

As the coefficient of \( \dot{L} \) is always positive, determinacy requires \( B_4'(L) > 0 \), or:

\[ 1 - \frac{g}{1 - \alpha} + \alpha \left( 1 - \frac{(1 - \tau_r)}{\sigma} \right) + \frac{(1 - \tau_w) (1 - \sigma)}{(1 - \chi) (1 - \alpha)} \left( \kappa + \frac{1 - \alpha - \kappa}{L} \right) > 0, \]

which is always the case when \( \sigma > 1 \), and indeed for values not too distant from one. Given determinacy we have uniqueness and no transitional dynamics. \( \dot{L} \) is the solution to \( B_4(L) = 0 \). As a consequence (27) and (28) still hold. Assume that all non-labor income is taxed at the same rate: using the ratios between \( wL, rF \) and \( p_P P \) we can then rewrite the government budget constraint \( wL \tau_w + rF \tau_r + p_P P \tau_r = (t + g) Y \) as:

\[ \tau_w = \frac{g + t}{\kappa} - \frac{(1 - \alpha^2 - \kappa)}{\kappa} \tau_r. \]  

(E23)

As to the welfare calculations, proceeding as in the previous cases, (41) can be written as:
\[ V = \frac{(\eta N(0))^{1-\sigma}}{1-\sigma} \frac{\sigma r \kappa (1-\tau_w)}{(1-\alpha) \alpha L} \frac{(1-\kappa) L^{1-\sigma}}{r(1-\tau_r)(\sigma-1)+\rho}. \]

(42) still applies, as well as (C3) and (C4). We have:

\[ \frac{\partial \log ((1-\sigma)V)}{(1-\sigma)\partial \tilde{L}} = \frac{\frac{(1-\sigma)(L-1)}{(1-\alpha)\alpha L} \frac{d\tau_w}{d\tau_r} \frac{(1-\alpha)\tilde{L}}{\kappa \sigma} - \frac{\tau_w}{C_1 \sigma}}{1-\tau_w} \frac{1}{(1-\alpha) \alpha L} \frac{1}{(1-\tau_r)(\sigma-1)+\rho}. \]

From \( B_4(\tilde{L}) = 0 \) we find:

\[ \frac{d\tilde{L}}{d\tau_r} = \frac{(1-\sigma)(L-1)}{(1-\alpha)\alpha L} \frac{d\tau_w}{d\tau_r} \frac{(1-\alpha)\tilde{L}}{\kappa \sigma} \frac{C_1 \sigma}{\rho L^{\frac{\alpha}{\chi-1}}}. \]

Finally from (E23) we have: \( d\tau_w/d\tau_r = -(1-\alpha^2-\kappa)/\kappa \). \( \tau_r \) can then be calculated, after eliminating \( \tau_w \) by using (E23), by solving:

\[ 0 = \frac{1-\alpha^2}{\kappa(1-\tau_w)} - \frac{r}{\kappa(1-\tau_r)(\sigma-1)+\rho} + \frac{(\alpha+\kappa-1)}{L} \frac{\kappa(1-\tau_r)C_1 L^{\frac{\kappa}{\chi-1}}}{r(1-\tau_r)(\sigma-1)+\rho} + \frac{\sigma + \chi - 2}{(1-\sigma) \left(1 - \frac{\tilde{L}}{\tilde{L}}\right)} \frac{(\sigma-1)(1-L)(1-\alpha^2-\kappa)}{(1-\alpha)\alpha L} \frac{1}{\kappa} \left(\frac{(1-\tau_w)}{(1-\tau_r)(\sigma-1)+\rho} \frac{C_1 \sigma}{\rho L^{\frac{\alpha}{\chi-1}}} - \frac{(1-\alpha)\tilde{L}}{(1-\alpha)\alpha L} \frac{1}{\kappa} \left(\frac{(1-\tau_w)}{(1-\tau_r)(\sigma-1)+\rho} \frac{C_1 \sigma}{\rho L^{\frac{\alpha}{\chi-1}}} \right) \right). \]