# Optimal Capital versus Labor Taxation with Innovation-Led Growth*

Philippe Aghion† Ufuk Akcigit‡ Jesús Fernández-Villaverde§

23 November 2012

**Abstract**

Chamley (1986) and Judd (1985) showed, analyzing optimal taxation in a standard neoclassical growth model with capital accumulation and infinitely-lived agents, that taxing capital cannot be optimal in the steady state (i.e. in the long run). In this paper, we introduce innovation-led growth into the Chamley-Judd framework, using a Schumpeterian growth model where productivity-enhancing innovations result from profit-motivated R&D investment, final good is produced with capital and labor, and capital accumulates over time exactly as in Chamley and Judd. Our main result is that, for a given required trend of public expenditure, a zero tax on capital becomes sub-optimal, due to a *market size effect*. In particular, for sufficiently high level of public expenditure, not taxing capital implies that labor must be taxed at a high rate. This in turn has a detrimental effect on labor supply, and therefore the market size for innovation. Simultaneously, taxing capital also reduces innovation incentives and, for some parameter values, it is optimal to subsidize it. However, the former effect dominates when the required level of public expenditure and the income elasticity of labor supply are both sufficiently high.

*We thank our discussants Pietro Peretto and Huw Lloyd-Ellis, as well as Peter Howitt, Caroline Hoxby, Torsten Persson, John Seater and seminar participants at the IIES (Stockholm University) and conference participants at Canadian Macro Study Group (2012) and SKEMA Workshop on Economic Growth (2012) for helpful comments.

†Harvard University and NBER, paghion@fas.harvard.edu

‡University of Pennsylvania and NBER, uakcigit@econ.upenn.edu

§University of Pennsylvania and NBER, jesusfv@econ.upenn.edu

# 1 Introduction

Should we tax capital income? Popular views on this issue are mixed: on the one hand, there is the view that taxing capital income is a way to restore the balance between capital owners and workers. Advocates of capital income taxation also stress that taxing labor only would encourage firms to become more capital intensive, which in turn will translate in lower wage or higher unemployment.[1] On the other hand, going back to Hobbes (1651) or John Stuart Mill (1848) is the view that individual savings already consist of income net of (labor income) tax; in other words, taxing the revenues from savings amounts to taxing labor income twice. Another argument against capital income taxation is that it discourages entrepreneurship and therefore employment and growth. Yet some observers point out that several developed countries with relatively lower unemployment rates and faster growth among OECD countries tax capital income. For example, in Sweden, capital income is taxed at a flat rate of 30%.[2]

This paper analyzes the growth and welfare effects of capital versus labor income taxation. As it turns out, growth effects do not feature prominently in existing taxation theories. For example Ordover and Phelps (1979) consider an overlapping generation model with high and low productivity individuals within each generation. They show that if consumers have separable utility functions and the government can implement the socially optimal level of capital stock, then the optimal capital income tax (which must induce truth-telling by high productivity agents) is equal to zero. More recently, Chamley (1986) and Judd (1985) analyze labor versus capital taxation in the context of a more standard neoclassical growth model with capital accumulation and infinitely-lived agents sharing the same intertemporal utility function (not necessarily separable). They show that taxing capital cannot be optimal in steady state (i.e. in the long run). The underlying intuition is well explained by Salanie (2003): if $\tau$ denotes the rate of tax or subsidy on capital, then the relative price of consumption in $T$ periods with respect to consumption today, is equal to $\left(\frac{1+r}{1+r(1-\tau)}\right)^T$ which goes to zero or infinity when $T \longrightarrow \infty$, which cannot be optimal.

In this paper we show that the Chamley-Judd result is not robust to introducing endogenous technical progress. More specifically, in Section 2, we start by reminding the reader of what happens in the Chamley-Judd model when we replace the no-growth assumption in that model by positive growth. The main finding is that the zero capital income tax results remains when growth is positive,

[1] Another argument is that not taxing capital income or taxing it differently from labor income encourages arbitrage, in particular from self-empoyed individuals who are then tempted to declare labor income as capital income (see Saez 20XX).

[2] While capital income is taxed at a relatively high rate in Sweden and other countries in Northern Europe, capital and labor income are not taxed at the same rate, which in turn is in line with our analysis in this paper.

but *exogenous*. Our second result is that the optimal tax rate on capital income ceases to be zero if growth becomes *endogenous*, that is, when the growth rate (both in the transitional dynamics and along the balanced growth path) is a function of the labor and capital tax rates. In Section 3 we open the black box of endogenous growth: specifically, we develop a Schumpeterian growth model with capital accumulation, where growth results from (profit-motivated) innovations which themselves result from R&D investments. R&D uses final good as an input, and final good is produced with capital and labor. Capital in turn accumulates over time exactly as in the Ramsey approach used by Chamley (1986) and Judd (1985). Our main result is that, for a given required trend of public expenditure, taxing capital at a zero rate may become suboptimal, due to a *market size effect*. More specifically, for sufficiently high level of public expenditure, not taxing capital implies that labor must be taxed at a high rate. This in turn has a detrimental effect on labor supply, and this in turn reduces final output which is produced with labor and intermediate inputs (machines) and profits and therefore it also reduces innovation incentives. True, taxing capital also reduces innovation incentives. However, with the Cobb-Douglas technology we assume for final good production, the former effect dominates when the required level of public expenditure and the income elasticity of labor supply are both sufficiently high. We then calibrate the model in Section 4. There our simulations show: *i*) that the growth-maximizing and welfare-maximizing tax rates on capital income in steady, are both increasing with both, the (long-run) required level of public expenditure and the income elasticity of labor supply; *ii*) that these optimal rates are negative when either of these two variables take small values and become positive otherwise; *iii*) that we are back to the Chamley-Judd zero capital tax rate result when the size of innovation goes to zero; *iv*) that the departure from the Chamley-Judd zero optimal capital tax when the size of innovation becomes positive, can be expressed as the sum of a market power effect (endogenous innovation involves monopoly rents for successful innovators) and the market size effect described earlier. One interesting implication of our analysis, is thus that the case for taxing capital income, is easier to make when the income elasticity of labor supply is high, whereas the strongest advocates of capital income taxation usually assume low labor supply elasticities.

Our analysis relates to two main strands of literature. There is first the literature on optimal capital taxation. In response to the seminal contributions by Atkinson and Stiglitz (1976), Ordover and Phelps (1979), Chamley (1986) and Judd (1985), all of which point to the optimality of zero capital income taxation, various attempts have been made at overturning this result by introducing suitable additional assumptions. Thus Chamley (2001) shows that when agents are credit-constrained, it may become optimal to also tax capital. The underlying intuition is that credit-constrained agents build up

precautionary savings to better resist the consequences of aggregate fluctuations, but credit constraints limit the extent to which insurance against aggregate risk can be achieved under laissez-faire. Taxation on accumulated savings then acts as an insurance device. Golosov *et al.* (2006) and Kocherlakota (2010) develop dynamic models with private information where the accumulation of savings leads to reduced labor supply in the future, hence the role of capital taxation to enhance future insurance possibilities. More recently, Piketty and Saez (2012) develop a dynamic model of savings and bequests with two sources of inequality: first, differences in labor incomes due to differences in ability; second, differences in inheritances due to differences in parental tastes for bequests. These two sources of inequality require two taxation instruments, not one. Hence the rationale for also taxing capital. However none of these contributions factor in the potential effects of taxation on growth, whereas we contribute to this whole literature by introducing (innovation-based) growth into the analysis.[3]

A second related literature is that on taxation and growth. A first major attempt at looking at the relationship between size of government and growth in the context of an AK model, is by Barro (1990, 1991). Using cross-country regressions, Barro finds that growth is negatively correlated with the share of public consumption in GDP, and insignificantly correlated with the share of public expenditure in GDP. More recently, Gordon and Lee (2006) perform cross country panel regressions of growth taxation over the period 1970-1997, and find a negative correlation between statutory corporate tax rates and average growth rates, both, in cross-section regressions and in panel regressions where they control for country fixed effects. Focusing more directly on entrepreneurship, Gentry and Hubbard (2004) find that both the level of the marginal tax rate and the progressivity of the tax discourage entrepreneurship (see also Gentry and Hubbard, 2003). Similarly, Petrescu (2009) shows that more tax progressivity reduces the probability of choosing self-employment and decreases the number of micro-enterprises, and that these effects are weaker in countries with higher levels of tax evasion. Similarly, Djankov *et al.* (2008) find that the effective corporate tax rate have a large adverse impact on aggregate investment, FDI, and entrepreneurial activity. However, none of these papers is concerned by the choice between labor and capital income tax and more generally by the optimal design of the tax system in a dynamic framework.

More closely related to our analysis are the papers by Jones, Manuelli and Rossi (1993) and by Atkeson and Burstein (2012). The former looks at the effects of capital and labor income tax in the context of an AK model, and they find a positive tax on capital income becomes optimal once the government optimizes on both, taxation policy and government expenditures over time. The latter

[3]There is also a literature that emphasizes that positive capital taxation is part of the best response of a government that cannot commit to future taxation paths. See, for instance Phelan and Stacchetti (2001)

looks at the impact of fiscal policy on product innovation and welfare. Their analysis depart from ours in several respects. First, they use a semi-endogenous growth model with (horizontal) product innovation whereas we develop our analysis in the context of a Schumpeterian model of innovation and endogenous growth. Second, they focus on the comparison between R&D tax credits, federal expenditures on R&D, and corporate profit tax whereas our focus is on labor versus capital income tax.[4]

The remaining part of the paper is organized as follows. Section 2 introduces growth in a reduced form into the Chamley-Judd model. Section 3 opens the growth black box using a fully-fledged Schumpeterian model of innovation and growth with capital and characterizes the optimal taxation policy under balanced growth. Section 4 calibrates our model and performs numerical simulations. And Section 5 concludes.

## 2 Introducing Growth in the Chamley-Judd Framework

Consider the standard Ramsey problem of a government that seeks to finance an exogenous stream of government expenses $\{G_t\}_{t=0}^{\infty}$ through distortionary, flat-rate, taxes on capital and labor earnings, $\{\tau_{k,t}, \tau_{l,t}\}_{t=0}^{\infty}$. The government's objective is to maximize the representative household's welfare subject to raising the required revenue. We consider an environment in which the government can fully commit to future tax rates.

In subsection 2.1, we refresh the reader's memory with the standard analysis of Chamley (1986) and Judd (1985) with no productivity growth and with their result that the optimal tax rate on capital is zero in the long-run. In subsection 2.2, we show that this result still holds when productivity grows exogenously. In subsection 2.3 we consider a reduced form endogenous growth model where productivity growth depends upon the tax structure, and explain why in general the zero capital taxation result does no longer obtain in that case. Section 3 will develop a full-fledged endogenous growth model to rationalize the reduced form specification of subsection 2.3.

[4] And as we are completing this draft we just heard about parallel work by Rebelo and Jaimovic (2012) who analyze the effect of (profit) taxation on growth in a model with endogenous innovations. A main point of their paper is that the detrimental effect of taxation on growth is non-linear: it starts being small as increasing the tax rate from zero first discourages the least talented innovators. However, the higher the tax, the more it also affects more talented entrepreneurs. Unlike in the present paper, the authors do not allow for labor taxation, and more generally they do not look at the optimal taxation structure. Also, not surprisingly, the growth-maximizing capita/profit tax can never be positive in their model.

### Household's Maximization Problem

We consider an infinite-horizon, discrete time, economy with an infinitely-lived representative household. The household owns the sequence of capital stocks $(K_t)_t$ and chooses consumption and labor allocations $\{C_t, L_t\}_{t=0}^{\infty}$ to maximize its lifetime utility

$$\max_{\{C_t, L_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \tag{1}$$

subject to its budget constraint

$$(1-\tau_{k,t}) r_t K_t + (1-\delta) K_t + (1-\tau_{l,t}) w_t L_t = C_t + K_{t+1}, \ \forall t.$$

In this budget constraint, $w_t$, $r_t, \tau_{k,t}$, $\tau_{l,t}$ and $\delta$ denote the wage rate, rental rate, capital income tax, labor income tax and depreciation rates, respectively. The unique final good can be consumed, used as capital or used by government as part of its spending. Household's utility is increasing and concave in consumption and decreasing labor such that

$$U(0,.) = 0, U_1(.) > 0, U_{11}(.) < 0, U_2(.) < 0, U_{22}(.) < 0.$$

The household maximization problem delivers the following optimality conditions

$$U_2(C_t, L_t) + \tilde{w}_t U_1(C_t, L_t) = 0 \tag{2}$$

$$U_1(C_t, L_t) = \beta U_1(C_{t+1}, L_{t+1})(\tilde{r}_{t+1} + 1 - \delta) \tag{3}$$

where where we have defined the after-tax input prices as

$$\widetilde{r}_t \equiv (1-\tau_{k,t}) r_t \text{ and } \widetilde{w}_t \equiv (1-\tau_{l,t}) w_t. \tag{4}$$

The first condition reflects the optimal choice between consumption and leisure, whereas the second (Euler) condition reflects the optimal choice between current and future consumption.

The final good is produced with capital and labor according to the Cobb-Douglas production function

$$Y_t = K_t^{\alpha} (A_t L_t)^{1-\alpha}.$$

Assuming that the final good sector is competitive, and choosing the final good as the numeraire,

the equilibrium prices for capital and labor are determined by:

$$r_t = \frac{\partial Y_t}{\partial K_t} = \alpha K_t^{\alpha-1} \left(A_t L_t\right)^{1-\alpha} \tag{5}$$

$$w_t = \frac{\partial Y_t}{\partial L_t} = \left(1-\alpha\right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}. \tag{6}$$

Finally, in this section we model the growth process in reduced form, as:

$$A_{t+1} = A_t \Phi\left(\tau_{k,t}, \tau_{l,t}\right). \tag{7}$$

The expression in (7) nests three alternative scenarios: $(i)$ no growth when $\Phi\left(\tau_{k,t}, \tau_{l,t}\right) = 1$ for all $\tau_{k,t}, \tau_{l,t}$, $(ii)$ exogenous growth when $\Phi\left(\tau_{k,t}, \tau_{l,t}\right) \equiv g$ for all $\tau_{k,t}, \tau_{l,t} > 0$, where $g$ is a constant and $(iii)$ endogenous growth when $\Phi\left(\tau_{k,t}, \tau_{l,t}\right)$ varies with the tax rates $\tau_{k,t}$ and $\tau_{l,t}$. We analyze the Ramsey problem under these three alternative scenarios in subsections 2.1, 2.2, and 2.3.

### Government's Maximization Problem

The government chooses a sequence $\{C_t, L_t, K_t, A_t, \tau_{k,t}, \tau_{l,t}\}_{t=0}^{\infty}$ to maximize household utility (1) subject to:

1. the economy's resource constraint

$$K_t^{\alpha} \left(A_t L_t\right)^{1-\alpha} + \left(1-\delta\right) K_t = C_t + G_t + K_{t+1},$$

   which says that current final output plus capital net of depreciation provide the resources that are being used for consumption, government spending, and investment;

2. the government's balanced budget condition

$$G_t = \tau_{k,t} r_t K_t + \tau_{l,t} w_t L_t$$

   which says that government spending at any date $t$ cannot exceed tax revenues;

3. the above optimality conditions of the representative household (2) and (3);

4. the process of growth (7).

For convenience, using the so-called "Euler equation"

$$Y_t = K_t^{\alpha} \left(A_t L_t\right)^{1-\alpha} = r_t K_t + w_t L_t$$

we can rewrite government's budget using after-tax prices (4) as

$$G_t = K_t^{\alpha} \left(A_t L_t\right)^{1-\alpha} - \widetilde{r}_t K_t - \widetilde{w}_t L_t$$

where

$$\widetilde{r}_t \equiv \left(1 - \tau_{k,t}\right) r_t \text{ and } \widetilde{w}_t \equiv \left(1 - \tau_{l,t}\right) w_t.$$

Government expenditures such as education, health, or police, are basically wages paid to government employees. As the economy grows, so do the wages of teachers, doctors, and police officers and at the same rate at the economy's growth rate. Hence, the total cost of delivering those public services will be constant as a share of output.[5]

Note that

$$\begin{aligned} \tau_{k,t} &= 1 - \frac{\widetilde{r}_t}{r_t} = 1 - \frac{\widetilde{r}_t}{\alpha K_t^{\alpha-1} \left(A_t L_t\right)^{1-\alpha}} \\ \tau_{l,t} &= 1 - \frac{\widetilde{w}_t}{w_t} = 1 - \frac{\widetilde{w}_t}{\left(1-\alpha\right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}} \end{aligned}$$

Thus the government's problem can be summarized as

$$\max_{\{C_t, L_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U\left(C_t, L_t\right)$$

subject to

$$K_t^{\alpha} \left(A_t L_t\right)^{1-\alpha} + \left(1-\delta\right) K_t = C_t + G_t + K_{t+1} \tag{8}$$

$$G_t (= \gamma Y_t) = K_t^{\alpha} \left(A_t L_t\right)^{1-\alpha} - \widetilde{r}_t K_t - \widetilde{w}_t L_t \tag{9}$$

$$U_2\left(C_t, L_t\right) + \tilde{w}_t U_1\left(C_t, L_t\right) = 0 \tag{10}$$

$$U_1\left(C_t, L_t\right) = \beta U_1\left(C_{t+1}, L_{t+1}\right)\left(\tilde{r}_{t+1} + 1 - \delta\right) \tag{11}$$

$$A_{t+1} = A_t \Phi\left(1 - \frac{\widetilde{r}_t}{\alpha K_t^{\alpha-1} \left(A_t L_t\right)^{1-\alpha}}, 1 - \frac{\widetilde{w}_t}{\left(1-\alpha\right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}}\right) \tag{12}$$

[5] If anything, Wagner's law suggests that the cost of many of these public services grow faster than output. It is plausible to think that education or health are goods with income elasticities larger than one and that the political economy process deliver public spending in education and health that outgrows output.

Let $\{\lambda_{1t}, \lambda_{2t}, \lambda_{3t}, \lambda_{4t}, \lambda_{5t},\}$ denote the Lagrangian multipliers associated with the above five sequences of constraints. In the following subsections, we will consider three cases: $(i)$ no growth: $\Phi(\tau_{k,t}, \tau_{l,t}) \equiv 1$, $(ii)$ exogenous growth: $\Phi(\tau_{k,t}, \tau_{l,t}) \equiv g$ and $(iii)$ endogenous growth: $\Phi(\tau_{k,t}, \tau_{l,t})$ varies with $\tau_{k,t}$ and $\tau_{l,t}$.

### 2.1 Steady State Without Growth

We first consider the steady-state equilibrium with no growth ($\Phi(\tau_{k,t}, \tau_{l,t}) \equiv 1$ for all $\tau_{k,t}, \tau_{l,t}$.) In this case all the Lagrange multipliers remain constant over time, hence we can drop time subscripts. In the Appendix we show that the first order conditions for the above government maximization problem imply:

$$1 = \beta\left(\alpha K^{\alpha-1}(AL)^{1-\alpha} + 1 - \delta + \frac{\lambda_2}{\lambda_1}\left(\alpha K^{\alpha-1}(AL)^{1-\alpha} - \widetilde{r}\right)\right) \tag{13}$$

However, the government must also satisfy the household's Euler condition (3), which we can rewrite in this no-growth case as:

$$1 = \beta\left(\widetilde{r} + 1 - \delta\right). \tag{14}$$

For both conditions (13) and (14) to hold simultaneously, we must have:

$$\widetilde{r} = \alpha K^{\alpha-1}(AL)^{1-\alpha} = r,$$

which in turn yields the Chamley-Judd result:

**Proposition 1** *Absent productivity growth, the optimal long-run tax rate on capital is* $\tau_{k,t} = 0$.

### 2.2 Balanced Growth Path with Exogenous Growth

We now move to the case where productivity growth is exogenous: $\Phi(\tau_{k,t}, \tau_{l,t}) \equiv g$. We focus on the balanced growth path equilibrium (BGP) with $r_t = r$, where capital and government spending also grow at rate $g$, and labor supply remains constant over time. We will identify a variable at the start of its BGP by using subindex "0".

In the Appendix we show that the first-order conditions for the government problem imply:

$$1 = \beta\left(\frac{\lambda_{1t+1}}{\lambda_{1t}}\left(\alpha(gK_0)^{\alpha-1}(gA_0L)^{1-\alpha} + 1 - \delta\right) + \frac{\lambda_{2t+1}}{\lambda_{1t}}\left(\alpha(gK_0)^{\alpha-1}(gA_0L)^{1-\alpha} - \widetilde{r}\right)\right) \tag{15}$$

However, note that in BGP, $\lambda_{2t}$ and $\lambda_{1t}$ decline at rate $g$ with $\lambda_{1t} = \lambda_{10}g^{-t}$ and $\lambda_{2t} = \lambda_{20}g^{-t}$.

Therefore equations (15) boil down to:

$$1 = \beta \left( \frac{1}{g} \left( \alpha K_0^{\alpha-1} (A_0 L)^{1-\alpha} + 1 - \delta \right) + \frac{1}{g} \frac{\lambda_{20}}{\lambda_{10}} \left( \alpha K_0^{\alpha-1} (A_0 L)^{1-\alpha} - \widetilde{r} \right) \right). \tag{16}$$

Next, the Euler equations (11) can be reexpressed as:

$$U_1 (C_t, L_t) = \beta U_1 (C_{t+1}, L_{t+1}) (\widetilde{r} + 1 - \delta) \tag{17}$$

Then, since (58) implies that $U_1 (C_t, L_t)$ and $\lambda_{1t}$ decline at the same rate $g$, so that $U_1 (C_t, L_t) = g^{-t} U_1 (C_0, L)$, the Euler equations (17) boil down to:

$$1 = \beta \frac{1}{g} (\widetilde{r} + 1 - \delta) \tag{18}$$

For conditions (16) and (18) to hold simultaneously we must have again:

$$\widetilde{r} = \alpha K_0^{\alpha-1} (A_0 L)^{1-\alpha}$$

which in turn establishes:

**Proposition 2** *When productivity growth is exogenous and constant over time, the optimal long-run tax rate on capital is* $\tau_{k,t} = 0$.

### 2.3 Balanced Growth Path with Endogenous Growth

Finally, we consider the case with endogenous growth modeled in a reduced form, namely as $\Phi (\tau_{k,t}, \tau_{l,t})$ varying with the tax rates $\tau_{k,t}$ and $\tau_{l,t}$. We again restrict attention to a BGP equilibrium with $r_t = r$ and where capital and government spending grow at the same rate $\Phi (\tau_{k,t}, \tau_{l,t})$. In this equilibrium the first order conditions for the government problem boil down to:

$$1 = \beta \begin{pmatrix} \frac{1}{g} \left( \alpha K_0^{\alpha-1} (A_0 L)^{1-\alpha} + 1 - \delta \right) + \frac{1}{g} \frac{\lambda_{20}}{\lambda_{10}} \left( \alpha K_0^{\alpha-1} (A_0 L)^{1-\alpha} - \widetilde{r} \right) \\ + \frac{\lambda_5}{\lambda_{10}} \frac{1}{gK_0} \left( - (1-\alpha) (1-\tau_{k,t}) \Phi_1 (\tau_{k,t}, \tau_{l,t}) + \alpha (1-\tau_{l,t}) \Phi_2 (\tau_{k,t}, \tau_{l,t}) \right) \end{pmatrix} \tag{19}$$

where

$$\lambda_{1t} = g^{-t} \lambda_{10},\ \lambda_{2t} = g^{-t} \lambda_{20},\ \lambda_{5t} = \lambda_5,\ U_1 (C_t, L_t) = g^{-t} U_1 (C_0, L) \text{ and } K_t = g^t K_0$$

As before, from (11) we have:

$$1 = \beta \frac{1}{g} \left( \widetilde{r} + 1 - \delta \right) \tag{20}$$

Obviously, (19) and (20) do not necessarily hold with equality when $\tau_{k,t} = 0$. For example, if

$$\Phi \left( \tau_{k,t}, \tau_{l,t} \right) = a_0 + a_1 \tau_{k,t} + a_2 \tau_{l,t},$$

then generically over the set of triplets $(a_0, a_1, a_2)$ the term $\frac{\lambda_5}{\lambda_{10}} \frac{1}{gK_0} \left( - \left( 1 - \alpha \right) \left( 1 - \tau_{k,t} \right) \Phi_1 + \alpha \left( 1 - \tau_{l,t} \right) \Phi_2 \right)$ is not equal to zero when $\tau_{k,t} = 0$. Hence:

**Proposition 3** *When growth is endogenous as in* (7)*, the optimal long-run tax rate on capital income is typically different from zero.*

***Proof.*** *See Appendix.* ■

# 3 Opening Up the Growth Black Box: The Market Size Effect

We now open the black box of the $\Phi$ function. To this end, we develop a full-fledged growth model with capital accumulation and endogenous innovation. In this model, productivity growth is generated by random sequences of quality improving innovations which themselves result from costly R&D investments. Taxes on labor and capital will affect the market values of these innovations, which in turn will then affect R&D investments and growth.

## Household

We continue with the same representative household setup of section 2. However, here we explicitly specify household utility function as:

$$\sum_{t=0}^{\infty} \beta^t \left( \ln C_t - \phi \frac{L_t^{1+\eta}}{1+\eta} \right), \tag{21}$$

where $\beta \in (0,1)$ is the discount factor, $C_t$ is consumption, $L_t$ is labor supply, $\eta$ is the inverse of the Frisch elasticity of labor supply, and $\phi > 0$ is a scale parameter for the disutility of labor. The household owns physical capital as well as all firms and collects labor income. Therefore, the budget constraint is:

$$K_{t+1} + C_t = \left( 1 - \tau_{l,t} \right) w_t L_t + \left( 1 - \tau_{k,t} \right) r_t K_t + (1 - \delta) K_t + \left( 1 - \tau_{\pi,t} \right) \left( \Pi_t - X_t \right) \tag{22}$$

where $K_t$, $\Pi_t$, $X_t$, $r_t$, and $w_t$ stand for physical capital, gross profit, R&D expenses, the rental rate of capital, and wage rate in period $t$, respectively. The tax rates imposed on profits net of R&D expenses, capital, and labor income are $\tau_{\pi,t}$, $\tau_{k,t}$, and $\tau_{l,t}$. We will assume that the tax rates are less or equal than 1 (that is, the government cannot tax more than the total income), but we let them to be negative (that is, the government can subsidize). The representative household maximizes (21) subject to the budget constraint (22).

The corresponding optimality conditions imply an Euler condition with the familiar form:

$$\frac{C_{t+1}}{C_t} = \beta \left(1 + \left(1 - \tau_{k,t+1}\right) r_{t+1} - \delta\right). \tag{23}$$

The intuition is straightforward. A higher discount factor delays consumption whereas a lower depreciation rate, a lower tax on capital, or a higher return rate increase consumption growth. Finally, the consumption-leisure arbitrage condition is simply

$$\phi C_t L_t^{\eta} = \left(1 - \tau_{l,t}\right) w_t. \tag{24}$$

For any given consumption level, a higher wage rate increases labor supply whereas higher tax on labor and higher scale of factor of disutility reduces the labor supply.

### Production and Factor Shares

We now describe the production side of the economy. We will see how, in equilibrium, the technology for the production of the final good has a reduced-form that is Cobb-Douglas in capital and labor. At the same time, the technology described in this subsection embodies the basic ingredients of an innovation-led model, namely: $(i)$ intermediate input production which is made more productive by innovation; and $(ii)$ monopoly rents rewarding successful innovators.

More specifically, final good production is such that:

$$Y_t = K_t^{\alpha} Z_t^{1-\alpha} \tag{25}$$

where $K_t$ is the capital stock at date $t$ and $Z_t$ is an intermediate goods basket produced according to the aggregator:

$$\ln Z_t = \int_0^1 \ln z_{i,t} di, \tag{26}$$

where $z_{i,t}$ is the amount of intermediate input $i$ used to produce $Z_t$ at time $t$. We assume that both

$Y_t$ and $Z_t$ are produced under perfect competition.

Taking the final good as numeraire, we then have, by Euler's identity:

$$Y_t = r_t K_t + P_{z,t} Z_t,$$

where $P_{z,t}$ is the price of the intermediate goods basket $Z_t$ at time $t$. The logarithmic structure in (26) implies that the producer of $Z_t$ will spend the same amount $\hat{Z}_t \equiv P_{z,t} Z_t$ on each intermediate input $i$. Therefore, the demand for each intermediate input $i$ at time $t$ is:

$$z_{i,t} = \frac{\hat{Z}_t}{p_{i,t}}. \tag{27}$$

Each intermediate input $i$ is produced by a one-period-lived ex-post monopolist. This firm holds the patent to the most advanced technology described by:

$$z_{i,t} = q_{i,t} l_{i,t}, \tag{28}$$

where $l_{i,t}$ is the amount of labor employed by intermediate input producer $i$ at time $t$ and $q_{i,t}$ is the corresponding labor productivity. Thus, its marginal cost of production is:

$$MC_{i,t} = \frac{w_t}{q_{i,t}}.$$

At the same time, in each sector $i$, there is a fringe of firms that can potentially produce the intermediate input using the previous technology:

$$z_{i,t} = q_{i,t-1} l_{i,t} = \frac{q_{i,t}}{\lambda} l_{i,t},$$

where $\lambda$ is the step of the innovation of the new technology. Thus, the potential producers have a marginal cost $w_t \lambda / q_{i,t}$. By Bertrand competition among all the firms in sector $i$, the latest innovator with marginal cost $w_t / q_{i,t}$ will set its price equal to the marginal cost of the previous innovator with marginal cost $w_t \lambda / q_{i,t}$, so that

$$p_{i,t} = \frac{w_t \lambda}{q_{i,t}}. \tag{29}$$

Thus, the equilibrium profit of intermediate input producer $i$ at time $t$ is equal to:

$$\pi_{i,t} = (p_{i,t} - MC_{i,t})\, y_{i,t} = \left(1 - \frac{MC_{i,t}}{p_{i,t}}\right) \hat{Z}_t = \frac{\lambda - 1}{\lambda} \hat{Z}_t.$$

In this expression we can see how the profit is determined by the innovation step (and hence, by the producing cost advantage) of the monopolist. In particular, profits go to zero when $\lambda \longrightarrow 1$. In that situation, we are back to the perfect competition/no innovation case in the traditional Chamley-Judd framework. Also, we see that the equilibrium profit is proportional to $\hat{Z}_t$. Below we will see that $\hat{Z}_t$ is itself proportional to final output $Y_t$.

From (27), (28), and (29) , the amount of labor employed by intermediate input producer $i$ is:

$$l_{i,t} = \frac{\hat{Z}_t}{w_t \lambda}. \tag{30}$$

Equating labor supply to labor demand, $L_t = \int_0^1 l_{i,t} di$, we get

$$\frac{w_t L_t}{\hat{Z}_t} = \frac{1}{\lambda}. \tag{31}$$

Note that labor share in the intermediate goods basket (26) goes to 1 as markups vanish, $\lambda \to 1$. Substituting (30) into (26), we get the wage rate as

$$w_t = \frac{(1-\alpha)}{\lambda} Q_t \left(\frac{K_t}{Z_t}\right)^{\alpha}, \tag{32}$$

where:

$$\ln Q_t \equiv \int_0^1 \ln q_{i,t} di$$

is an aggregate productivity index and the price of $Z_t$ is simply its marginal product in final good production

$$P_{z,t} = (1-\alpha) K_t^{\alpha} Z_t^{-\alpha}. \tag{33}$$

Equation (32) shows that equilibrium wages are proportional to aggregate productivity, and in a balanced growth path where $\frac{K_t}{Z_t}$ is constant, these two variables will grow at the same rate.

**Remark 4** *Combining* (31) , (32)*, and* (33)*, we can reexpress the production technology* (25) *for the final good in the reduced form:*

$$Y_t = K_t^{\alpha} (Q_t L_t)^{1-\alpha}, \tag{34}$$

*which, as claimed above, is Cobb-Douglas in capital and labor.*

Using the fact that final good production is competitive, we have:

$$r_t = \alpha K_t^{\alpha-1} Z_t^{1-\alpha}, \tag{35}$$

which implies a capital share of output equal to $\frac{r_t K_t}{Y_t} = \alpha$. Similarly from (33), the factor share for the intermediate goods basket $Z$ is equal to:

$$\frac{\hat{Z}_t}{Y_t} = (1-\alpha).$$

This factor share is itself decomposed into the labor share:

$$\frac{L_t w_t}{Y_t} = \frac{\hat{Z}_t}{Y_t \lambda} = \frac{1}{\lambda}(1-\alpha), \tag{36}$$

and the profit share:

$$\frac{\Pi_t}{Y_t} = \frac{\int \pi_{i,t} di}{Y_t} = \frac{\frac{\lambda-1}{\lambda}\hat{Z}_t}{Y_t} = \frac{\lambda-1}{\lambda}(1-\alpha), \tag{37}$$

which vindicates our claim that equilibrium profits, and therefore innovation incentives, are proportional to final output. From these last expressions, we learn that the share of capital, labor, and profits are determined by two parameters: $\alpha$ and $\lambda$. Moreover, (i) aggregate (monopoly) profits are equal to zero when $\lambda = 1$ (in other words, the case of the basic Chamley-Judd model with perfect competition and no growth in $Q_t$ analyzed in the previous section); (ii) aggregate profits are proportional to aggregate final output. This observation will imply that the equilibrium rate of innovation and the rate of growth of the economy will be proportional to final output, that is, to the *market size*. The resulting relationship between equilibrium profits and final output means that anything that enhances aggregate activity will also stimulate innovation and growth. This in turn will have consequences for the optimal taxation policy. In particular, the Chamley-Judd result that the long-run tax on capital should be zero will no longer be true in general.

### 3.1 Innovation and Growth

We are now ready to describe how innovation and growth come about in the model. Before doing so, though, it is helpful to clarity the timing within each period $t$. Events unfold as follows: $(i)$ period $t$ starts with some initial $Q_t$ and $K_t$; $(ii)$ inventors in each intermediate sector invest in R&D; $(iii)$ successful innovators become monopolists in their intermediate sector; $(iv)$ monopolists produce, consumers consume and invest for the next period, which leads to $K_{t+1}$; $(v)$ quality improvements take place according to step $iii$ and $Q_{t+1}$ is determined accordingly; $(vi)$ period $t$ ends. In that way, technology is predetermined at the start of the period and the current period R&D only affects the technology tomorrow, a timing convention that is both empirically plausible and computationally convenient.

Growth in the model results from innovations that increase the productivity of labor in intermediate inputs production, each time by some given factor $\lambda > 1$. We assume that, at the beginning of any period $t$, there is the opportunity to innovate in order to increase labor productivity in any intermediate sector $i$. More specifically, a potential innovator in sector $i$ must spend the amount $x_{it}\psi Q_t$ of final good to innovate at rate $x_{i,t}/\bar{x}_t$ where $\bar{x}_t$ is the innovation intensity of other potential innovators. The presence of $Q_t$ in this amount just indicates that the cost of innovation grows with the already existing level of technology in a linear fashion. Since expenses on R&D can be taken out from gross profits for tax purposes, the resulting maximization problem for innovators is:

$$\max_{x_i} (1 - \tau_{\pi,t}) \left( \frac{x_{i,t}}{\bar{x}_t} \pi_{i,t} - x_{i,t}\psi Q_t \right) \tag{38}$$

We assume that, in each intermediate good sector $i$, there is always a monopolist producing. A new successful inventor gets the monopoly right to produce with the "current" technology but without the new quality jump that it creates. In a symmetric equilibrium, the innovation rate is given by:

$$x_{i,t} = \bar{x}_t = \frac{\pi_{i,t}}{\psi Q_t} = \frac{\lambda - 1}{\lambda} \frac{\hat{Z}_t}{\psi Q_t} = \frac{\lambda - 1}{\lambda} \frac{1 - \alpha}{\psi} \frac{Y_t}{Q_t},$$

which leads to the equilibrium growth rate:

$$\frac{Q_{t+1}}{Q_t} = 1 + \frac{(\lambda - 1)^2}{\lambda} \frac{1 - \alpha}{\psi} \frac{Y_t}{Q_t}. \tag{39}$$

Note that this growth rate depends negatively on the share of capital $\alpha$ and cost of R&D $\psi$; but positively on markups $\lambda$, and normalized market size $Y_t/Q_t$. Therefore, a growth-maximizing policy (which is not necessarily a welfare-maximizing one) should aim to maximize current output.

The total amount of the final good used in R&D is:

$$X_t = \bar{x}_t \psi Q_t = \frac{\lambda - 1}{\lambda} (1 - \alpha) Y_t = (1 - \Omega) Y_t \tag{40}$$

where $\Omega \equiv 1 - \frac{\lambda - 1}{\lambda} (1 - \alpha)$. Note that (37) and (40) imply that profits net of R&D is simply zero.

### Government

Government spending is assumed to grow at the same rate as the economy, namely:

$$G_t = Q_t G_0, \tag{41}$$

where $Q_t$ evolves over time according to (39) and $G_0$ is exogenously given. With this assumption we capture the idea that many government expenses grow with the economy. For example, in reality, government employees wages cannot grow more slowly than private sector wages, which in turn grow with labor productivity. Otherwise, the government would find it increasingly difficult to hire workers. Similarly, many goods and services purchased by the government have prices determined by opportunity costs that increase with economic growth and labor productivity (think, for instance, about the land required to build a public school). Technically, in a model with growth, we need to take a stand on how government spending evolves over time. If we assumed a constant spending, economic growth will make it asymptotically negligible. If we assumed that spending grows faster or slower than the economy, asymptotically we would either violate the resource constraint of the economy or be back in the case where government spending is trivially small.

In every period $t$, the government balances its budget period by period by choosing $\{\tau_{k,t}, \tau_{l,t}, \tau_{\pi,t}\}$ on capital and labor such that

$$G_t = \tau_{l,t} w_t L_t + \tau_{k,t} r_t K_t + \tau_{\pi,t} \hat{\Pi}_t. \tag{42}$$

where $\hat{\Pi}_t \equiv \Pi_t - X_t$. We impose balance budget to save on a state variable and to make the analysis more transparent.[6] Finally, the aggregate resource constraint of the economy is

$$Y_t + (1-\delta) K_t = K_{t+1} + C_t + G_t + X_t \tag{43}$$

where we added aggregate R&D spending to total expenditures compared to the aggregate resource constraint in the previous section.

### 3.2 Equilibrium and Balanced Growth Path

We can now define an equilibrium for our model.

**Definition 5** *Given tax policy $\{\tau_{l,t}, \tau_{k,t}, \tau_{\pi,t}\}_{t=0}^{\infty}$ and initial conditions $K_0$ and $Q_0$, a dynamic equilibrium for the model is a tuple $\{C_t, Y_t, K_t, L_t, X_t, \Pi_t, r_t, w_t, Q_t\}_{t=0}^{\infty}$ such that given prices, the household maximize (conditions (23) and (24), all firms maximize (conditions (35), (36), (37), and (39)), the*

[6] In our computational exercise, as we will describe below, we will not allow taxes on capital to be time-dependent. In that way, we eliminate the asymmetry present in the traditional Chamley-Judd framework at time zero, where taxation of capital is not distorsionary. In that world, the government has an incentive to tax capital as much as possible at that zero period and to use the proceedings to accumulate positive assets to reduce future taxation. Since, we do not allow for those time-dependent taxes, having the possibility to accumulate assets (or to run debt) is less important for the analysis.

*government satisfies its budget constraint (42), and markets clear (conditions (25), (40), and (43)).*

After suitable substitutions, the conditions in the above definition boil down to the following system of equations:

$$K_{t+1} + C_t + G_t = \Omega Y_t + (1-\delta) K_t \tag{44}$$

$$C_{t+1} = \beta C_t \left(1 + (1-\tau_{k,t+1}) \alpha \frac{Y_{t+1}}{K_{t+1}} - \delta\right) \tag{45}$$

$$\phi C_t L_t^{1+\eta} = (1-\tau_{l,t}) \frac{1-\alpha}{\lambda} Y_t \tag{46}$$

$$Y_t = K_t^{\alpha} (Q_t L_t)^{1-\alpha} \tag{47}$$

$$G_t = \left(\frac{1-\alpha}{\lambda}\tau_{l,t} + \alpha\tau_{k,t}\right) Y_t \tag{48}$$

and

$$\frac{Q_{t+1}}{Q_t} = 1 + \frac{(\lambda-1)^2}{\lambda} \frac{1-\alpha}{\psi} \frac{Y_t}{Q_t}. \tag{49}$$

Similarly, we can define a balance growth path (BGP):

**Definition 6** *A BGP is a dynamic equilibrium where the aggregate variables* $\{C_t, Y_t, K_t, X_t, \Pi_t, w_t, Q_t\}_{t=0}^{\infty}$ *grow at the same constant rate* $g$ *and* $\{L_t, r_t\}_{t=0}^{\infty}$ *remain constant.*

### 3.3 Analytical Solutions

We can simplify and derive further analytical results by concentrating on the case with full capital depreciation, that is, with $\delta = 1$. We then immediately obtain:

**Proposition 7** *Consider the benchmark economy with full depreciation* $\delta = 1$. *Then:*
$(i)$ *in laissez-faire economy with* $G_0 = 0$, *the equilibrium solution takes the following form*

$$C_t = (1-s) Y_t \text{ and } L_t = L^*$$

*where*

$$s = \frac{(\lambda-1)(1-\alpha)}{\lambda} + \beta\alpha \tag{50}$$

*and*

$$L^* = \left[\frac{1-\alpha}{\phi(1-s)\lambda}\right]^{\frac{1}{1+\eta}}.$$

*(ii) in an economy with $G_0 > 0$, the balanced growth equilibrium takes the following form*

$$C_t = (1-s)\, Y_t \text{ and } L_t = L^*$$

*where*

$$s = \frac{(\lambda-1)(1-\alpha)}{\lambda} + \beta\alpha + \frac{1-\alpha}{\lambda}\tau_{l,t} + \alpha(1-\beta)\,\tau_{k,t} \tag{51}$$

*and*

$$L^* = \left[\frac{(1-\tau_{l,t})(1-\alpha)}{\phi(1-s)\lambda}\right]^{\frac{1}{1+\eta}}.$$

**Proof.** See appendix. ■

This proposition implies that equilibrium labor supply is decreasing in the tax rate on labor. More importantly, it will also imply that taxing labor too much can be detrimental to growth since this could reduce the market size too much. To see this, in the following proposition we focus on steady-state growth and use a Taylor approximation around $\lambda = 1$.

**Proposition 8** *For $\lambda$ close to 1, the steady state growth rate of the benchmark economy is approximately equal to*

$$g^* = 1 + (\lambda-1)^2 \frac{1-\alpha}{\psi}\left[\beta(1-\tau_k)\alpha\right]^{\frac{\alpha}{1-\alpha}}\left[\frac{(1-\tau_l)(1-\alpha)}{\phi(1-(1-\alpha)\tau_l - \alpha\tau_k - \beta(1-\tau_k)\alpha)}\right]^{\frac{1}{1+\eta}}$$

**Proof.** See appendix. ■

Intuitively, the higher the tax rate on labor income the lower the amount of labor supply in equilibrium and therefore the lower the size of profits to successful innovators. This in turn lowers R&D incentives and thus the equilibrium innovation efforts. True, taxing capital also reduces innovation incentives. However, with the Cobb-Douglas technology we assume for final good production, the former effect dominates when labor tax is initially high (or labor supply is initially low). This *market size effect* could potentially counteract the Chamley-Judd effect pointed out in the previous section.

In section 4, we will provide a detailed quantitative investigation of these different effects and we shall compute the welfare-maximizing tax structure for various parameters values. For given public spending and tax policy, the equilibrium welfare is given by (21). The wefare-maximizing policy maximizes (21) subject to equations $(44)-(49)$. Note that when $\lambda = 1$, we are back to the Chamley-Judd model with no growth and no market power. In that section we shall decompose the departure from the Chamley-Judd welfare maximizing tax rates, into what comes from the market power part

and what comes from the market size effect part.

# 4 Quantitative Analysis

In this section, we perform a quantitative analysis of our model. First, we calibrate the structural parameters of the model for a baseline case. Second, we describe how we compute the model. Third, we characterize the optimal policy in the benchmark calibration. Fourth, we perform an extensive set of sensitivity analysis exercises.

## 4.1 Calibration

One considerable advantage of our model is its parsimony: we only have 8 parameters: $\beta$, $\phi$, $\eta$, $\alpha$, $\lambda$, $\delta$, $\psi$, and $G_0$. In our benchmark calibration, we select values for the first 7 parameters to match certain observations of the U.S. economy at an annual rate. In that way, our choices will be quite close to standard values in the literature. With respect to the preferences, we set the discount factor $\beta = 0.979$, to generate an annual interest rate of 4 percent, $\phi = 8.5$ to make hours worked to be around 1/3, and $\eta = 0.833$ (a Frisch elasticity of 1.2) following the evidence in Chatty *et al.* (2011). With respect to technology, we set $\alpha = 0.295$ and $\lambda = 1.0522$ to give us a labor income share of 0.67, a capital income share of 0.295, and a R&D share on GPD of 0.035. A labor income share of 0.67 is a standard value in the business cycle literature and corresponds to the long run average observed in the U.S. Our choice of R&D share on GPD of 0.035 corresponds to the estimates of Acemoglu and Akcigit (2012). A depreciation $\delta = 0.06$ matches a capital/output ratio of around 2.7 (close to the one observed in the U.S. economy) and $\psi = 0.045$ delivers (conditional on the other parameters), a growth rate of 1.85 percent, roughly the per capita long-run growth rate of the U.S. economy since the late 19th century. Instead of calibrating the final parameter, $G_0$, we will explore the optimal policy for a large range of its possible values: from 0.1 to 0.18 (we will discuss below how big these are as a percentage of output). This range will be sufficiently wide as to give us quite a good view of the behavior of the optimal tax as a function of government expenditure.

## 4.2 Computation

We will solve the model by finding a third-order perturbation for the decision rules of the agents in the model given a tax policy, evaluate the welfare associated with those decision rules, and then search over the space of feasible tax policies. Third-order perturbations have become popular because they are extremely fast to compute while, at the same time, being highly accurate even far away from

the point where the perturbation is performed. This accuracy is thanks to the presence of quadratic and cubic terms and it is particularly relevant in cases where we are evaluating welfare, a non-linear function of the equilibrium variable values. Caldara *et al.* (2012) provide further background on third-order approximations, compares them numerically with alternative solution methods. Also, in appendix A2, we report the Euler equation errors of our solution. These Euler equation errors will demonstrate the more than satisfactory accuracy of our perturbation.

Before we proceed further, and since our model has long-run growth, we first need to renormalize the variables to keep an inherently local approximation such as a higher-order perturbation relevant.

Note that we can rewrite the equilibrium conditions of the model as:

$$\widetilde{K}_{t+1}\widetilde{Q}_{t+1} + \widetilde{C}_t + G_0 = \Omega\widetilde{Y}_t + (1-\delta)\,\widetilde{K}_t$$
$$\widetilde{C}_{t+1}\widetilde{Q}_{t+1} = \beta\widetilde{C}_t\left(1 + (1-\tau_k)\,\alpha\frac{\widetilde{Y}_{t+1}}{\widetilde{K}_{t+1}} - \delta\right)$$
$$\phi\widetilde{C}_t L_t^{1+\eta} = (1-\tau_l)\,\frac{1-\alpha}{\lambda}\widetilde{Y}_t$$
$$\widetilde{Y}_t = \widetilde{K}_t^{\alpha} L_t^{1-\alpha}$$
$$G_0 = \left(\frac{1-\alpha}{\lambda}\tau_{l,t} + \alpha\tau_{k,t}\right)\widetilde{Y}_t$$
$$\widetilde{Q}_{t+1} = 1 + \frac{(\lambda-1)^2}{\lambda}\frac{1-\alpha}{\psi}\widetilde{Y}_t$$

where, for an arbitrary variable $J_t$, we have defined the variable normalized by the average productivity index:

$$\widetilde{J}_t = \frac{J_t}{Q_t},$$

except for

$$\widetilde{Q}_{t+1} = \frac{Q_{t+1}}{Q_t}.$$

With that normalization, it is straightforward to find a steady state on the transformed variables and use that steady state as the approximation point of the perturbation. We will use the notation $\widetilde{J}_{ss}$ for the value of an arbitrary variable in such steady state. In the same way, we define the log deviation of such normalized variable as

$$\widetilde{J}_t = \widetilde{J}_{ss} e^{\log\frac{\widetilde{J}_t}{\widetilde{J}_{ss}}} = \widetilde{J}_{ss} e^{\widehat{\widetilde{J}}_t}$$

where $\widehat{\widetilde{J}}_t = \log\frac{\widetilde{J}_t}{\widetilde{J}_{ss}}$ and perform the perturbation in logs instead of levels. Also, for average produc-

tivity, and following the same convention as before, we get

$$Q_t = Q_{t-1}\widetilde{Q}_{ss}e^{\widehat{\widetilde{Q}}_t},$$

which will allow us to undo the normalization once we have computed the model in order to evaluate welfare and the equilibrium path of the economy.

Once we have transformed the model, we find the rescaled steady state, we substitute the unknown decision rules within the equilibrium conditions of the model, we take derivatives of those conditions and solve for the unknown coefficients in the partial derivatives of the decision rules. With these coefficients, we can find a third-order Taylor expansion of decision rules and simulate the equilibrium dynamics from any arbitrary initial condition.

### 4.3 Baseline Model

Our first exercise is to characterize optimal policy in the baseline model with a benchmark calibration. More concretely, we will implement the following experiment:

1. We will set the tax rate on capital income to be constant over time.
2. We will balance the budget period per period using labor taxes.
3. We will set as the initial level of capital the capital for a BGP when the tax rate on capital income is 0 percent.
4. Then, we will compute the tax rate on capital that maximizes the welfare of the representative household for a range of values of $G_0$ that imply a government expenditure that ranges from around 21 to 42 percent of output.

The motivation for each of these choices is as follows. First, we fix the tax on capital over time to simplify the computation of the problem, to make the intuition of the result more transparent, and because we do not find complicated time-dependent policies as those implied by a pure Ramsey analysis either plausible or empirically relevant. Second, we tax profits at the same rate than capital income because, in real life, it is extremely hard to distinguish, from an statement of profits and losses by a firm, how much of those profits correspond to payments to capital and how many to rewards to innovation. In any case, we will relax this assumption in sensitivity analysis below. Third, as explained above, we balance the budget period by period to avoid handling an extra state variable. Also, given our choice about constant tax rate on capital (and hence, the inability to run a surplus at

time zero, where taxing capital does not distort), this constraint is less important. Fourth, we set the initial capital to the one implied by a BGP when the tax rate on capital is 0 percent. This is the level of capital tax in the Chamley-Judd case. With that choice, we will have a tax rate that is the lower range of the observed rates. Finally, note that the tax on profits is irrelevant because profits net of R&D are zero.

In figure 1 we plot the optimal tax on capital (continuous, blue line) and on labor (discontinuous, red line) along the BGP. We can see how the optimal tax is increasing on the share of government expenditure in output, from -15.8 percent to 7.6 percent. The tax rate on labor goes from 38.5 percent to 59.7 percent.

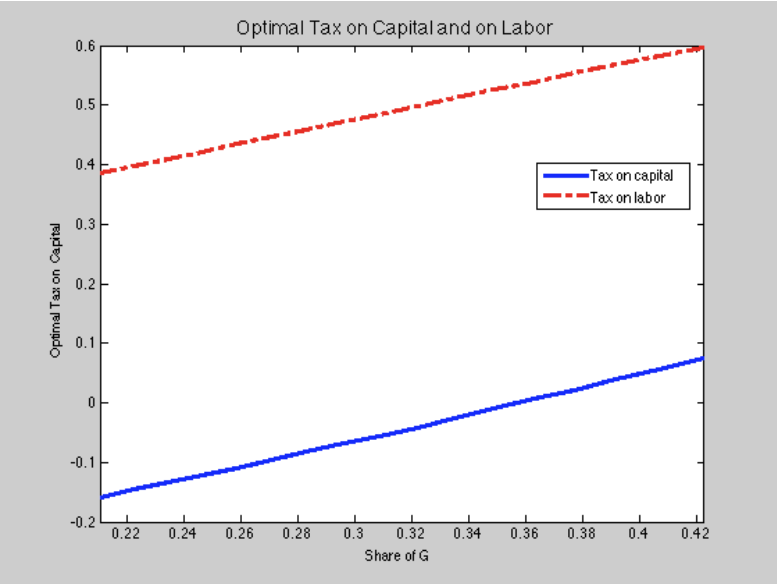



Figure 1: Optimal tax on capital and on labor as a function of the share of G on output.

The intuition of these results is as follows. When we have low levels of government expenditure, the optimal policy consists of a subsidy to capital (a negative tax rate) financed by a positive tax on labor income, which also finances the expenditure. There are two reasons for this. First, we have monopolistic competition in the intermediate inputs. Therefore, the production level is too low with respect to first best. By subsidizing capital, we induce a higher level of production. Second, by subsidizing capital, we increase the market size since more capital is accumulated and output grows. However, as we increase the size of government expenditure, the tax rate on labor must grow, increasing the distortions in labor supply, lowering hours, and with them, the market size. The only

way to minimize this impact is by reducing the subsidy to capital to the point that it eventually becomes positive and as high as 7.6 percent.

Obviously, dozens of other possible experiments are feasible. For example, none of the four assumptions in the exercise of this section are essential and we just picked them to illustrate our point more forcefully. Similarly, we can easily change any of the parameter values of our calibration. Over the next subsections, we will report on several sensitivity analysis exercises that we found particularly interesting for our main argument.

### 4.4 Small $\lambda$

A simple exercise to appreciate the effects of monopolistic competition and of the market size effect that we highlighted above is to keep all the same assumptions as in the baseline model with a benchmark calibration, but to reduce $\lambda$, the step size of the reduction of marginal costs, to 1.01. In that way, the market power nearly disappears and the market size effect is small (below we will push the argument to the limit by setting $\lambda = 1$ and hence recovering a Neoclassical environment).

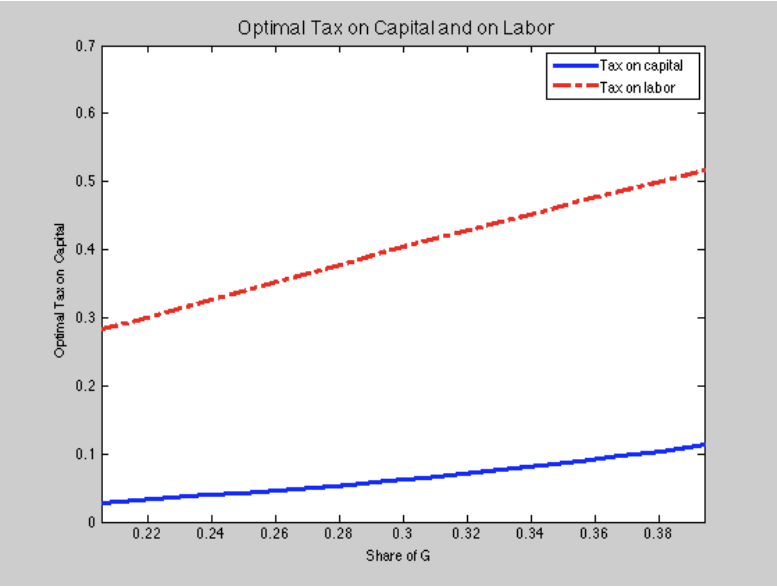



Figure 2: Optimal tax on capital and on labor as a function of the share of G on output, small llambda.

The results appear in figure 3, where we observe how now the tax rate on capital income is positive, although small. Simultaneously, the tax rate on labor income is much smaller as we do not need to finance the subsidy to capital.

The main drawback of this exercise is that we cannot neatly separate the monopoly effect from the market size effect. We will revisit this issue in subsection 4.7.

### 4.5 The Neoclassical Case

As announced above, we now reduce $\lambda$ all the way down to 1, keep all the previous assumptions unchanged, and recover a neoclassical case: both without market power and without growth.

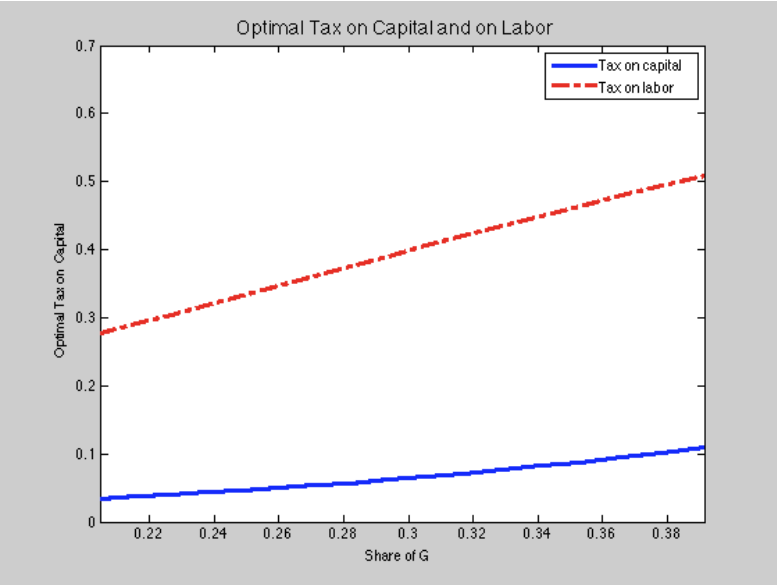



Figure 3: Optimal tax on capital and on labor as a function of the share of G on output, neoclassical.

The results are reported on figure 3, where we see that tax rate on capital income goes from roughly 3.4 percent to slightly more than 11 percent. Note that, in comparison with the standard Chamley-Judd analysis, tax on capital income is slightly positive. In Chamley-Judd, we let the government have time-varying taxes. Given that freedom, a benevolent government would like to tax capital at time 0 because it does not distort (capital is already in place). If we do not let the government to do that, it will trade-off a bit of distortion in the long-run with a with of extra tax revenue from capital in the short run (which allows to reduce taxation on labor). In any case this effect is small and the tax on capital is low. Obviously, as $\beta \to 1$, this effect disappears. As shown in figure 4, we have checked that, for $\beta = 0.999$, the optimal tax on capital is numerically zero.

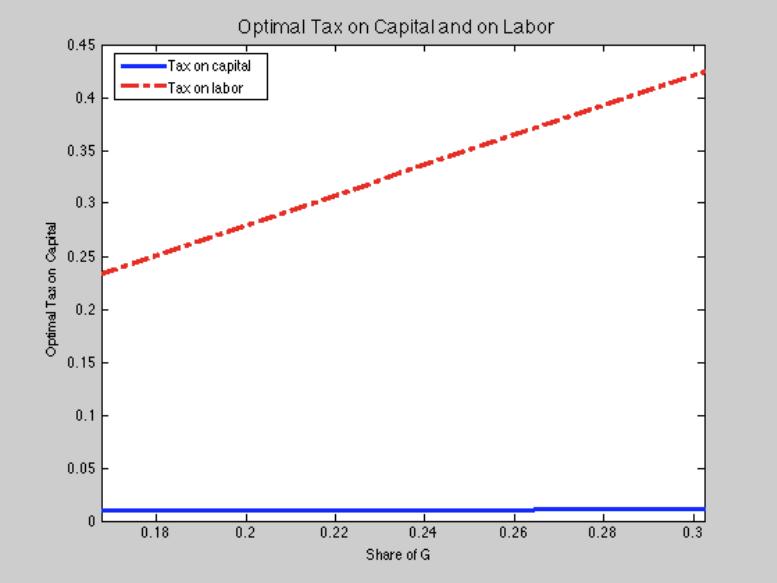



Figure 4: Optimal tax on capital and on labor as a function of the share of G on output, neoclassical, high beta

### 4.6 Exogenous Growth

We can also do the neoclassical case with exogenous growth:

$$\frac{Q_{t+1}}{Q_t} = 1 + g$$

where $g$ is calibrated to be the same than in the baseline model (around 1.85 percent at an annual level).

The results appear in figure 5 and are nearly identical to the ones in figure 4. This exercise allows us to argue that growth is not, per se, the reason for our results in the benchmark model, but the fact that growth is affected by the market size instead of being exogenous. The results are not a surprise, either, since we know that in the neoclassical case, the presence of a positive $g$ has nearly the same numerical implications than a slightly lower $\beta$.

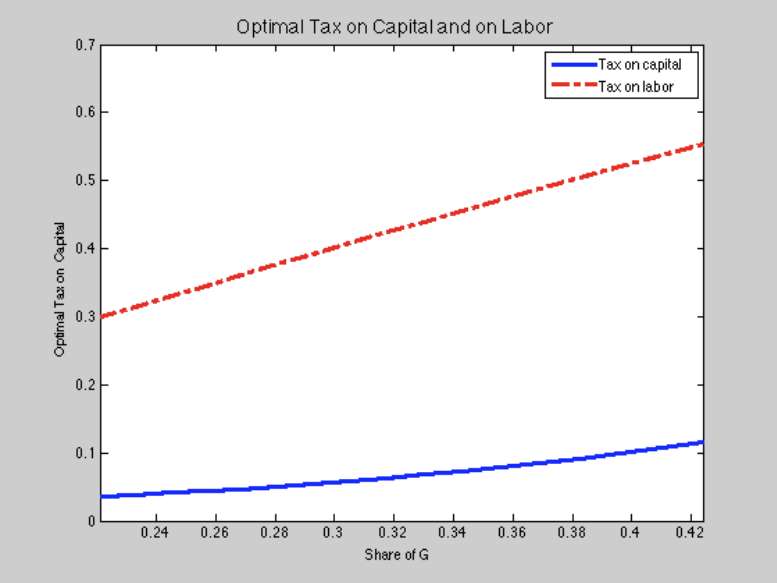



Figure 5: Optimal tax on capital and on labor as a function of the share of G on output, exogenous growth.

### 4.7 Exogenous Growth plus Market Power

In this exercise, we keep exogenous growth but we reintroduce market power. The results are shown in Figure 6. We see that market power per se (that is, abstracting from the effects of market power or the size of innovation on innovation incentives), pushes towards taxing capital income. The comparison between Figure 1 and Figure 6 is also interesting: it tells us that for the corresponding value of the labor elasticity, the market power distortion is what pushes towards taxing profits, and therefore capital income, at a higher rate for high levels of the required public expenditure. On the other hand the market size effect, i.e. the effect on innovation incentives per se, pushes towards subsidizing capital for this value of the labor market elasticity. However, for sufficiently higher labor elasticities, the market size effect also pushes towards taxing capital income for high values of the required public expenditure: in this case taxing labor income instead has a more detrimental effect on labor supply and therefore on market size which drives innovation incentives.

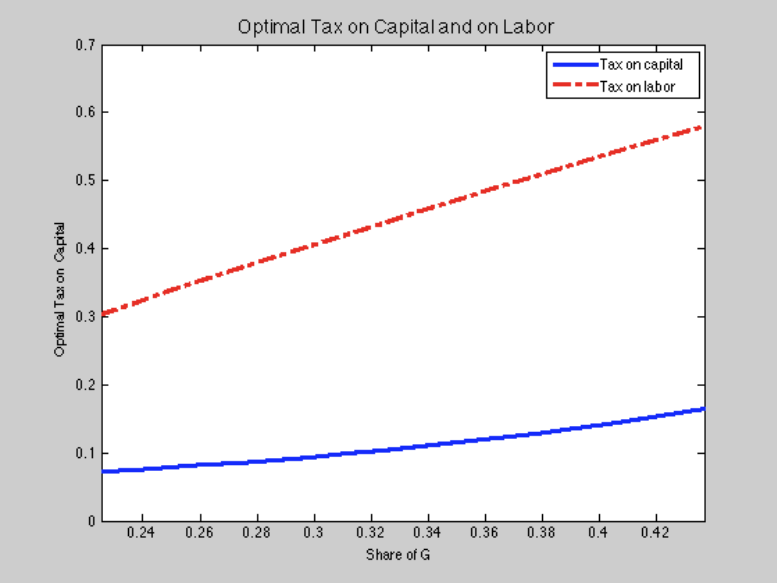



Figure 6: Optimal tax on capital and on labor as a function of the share of G on output. Exogenous growth and market power.

### 4.8 Higher Labor Supply Elasticity

Finally, Figure 7 depicts the growth-maximizing tax rates on capital income as a function of the elasticity of the Frisch labor supply, all for a given $G_0$ (where growth maximization refers to the highest growth rate along the BGP). We selected a range for the elasticity between 0.1 and 7 that encompassed all empirically relevant values. Figure 7 shows a positive relationship between the growth-maximizing rate and the elasticity of labor supply.[7] Moreover the growth-maximizing tax rate on capital income is negative for small elasticities, whereas it becomes positive for sufficiently high labor supply elasticities at this level of public expenditure (the cross is at around an elasticity of 4). The intuition is simple: when labor is not very elastic, we can tax it at high rates to generate revenue to subsidize capital with little cost in terms of distortions, increase the market size, and raise the growth rate of the economy. Interestingly, as we pointed out in the introduction, the strongest advocates of high tax rates on capital income tend to assume low values for the labor supply elasticity. Our model suggest that one

[7] We could also have reported the welfare-maximizing tax rate. The lesson for this alternative would be exactly the same. We prefer to focus on the growth-maximizing rate in this subsection because the intuition is more transparent. Also, when one is dealing with a range of Frisch elasticities as large as the one in Figure 7, what are sensible initial conditions of the model for some elasticities are not particularly good choices for elasticities very far away. This matter because the initial condition induces a transitional dynamics that can obscure the effects we are highlighting. By focusing on the growth-maximizing tax rate, we avoid this problem.

can favor one or the other, but that both positions are difficult to reconcile within the context of the theory presented in this paper.

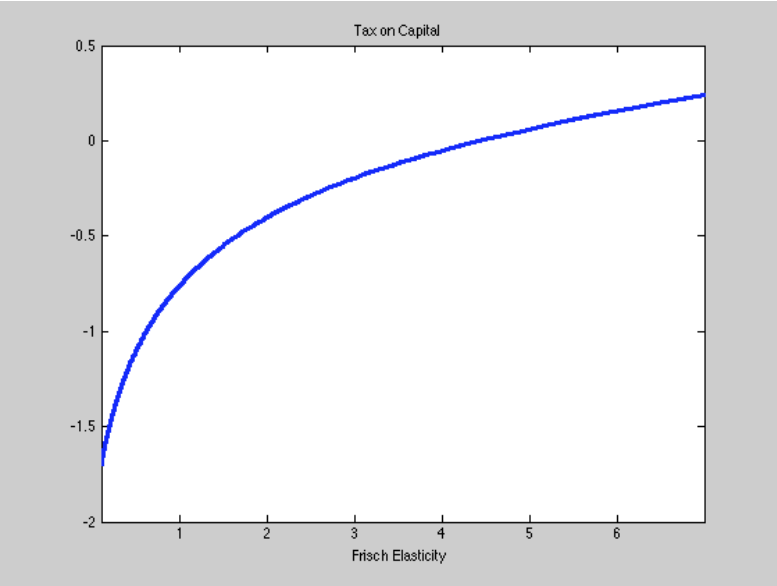



Figure 7: Growth-maximizing tax on capital as a function of the Frisch labor supply elasticity.

## 5 Conclusion

In this paper we have extended the Chamley-Judd framework by introducing innovation based growth. We showed that the long term optimal tax rate on capital ceases to be zero when we introduce innovation-led growth. This departure reflects two effects: first, a market power effect driven by the monopoly distortion associated with endogenous innovation; second, a market size effect which relates the equilibrium rate of innovation to the equilibrium level of aggregate final output. For low levels of public investment or for low elasticity of labor supply, the market size effect pushes towards taxing labor income while subsidizing capital income. For high levels of public investment *and* high elasticity of labor supply, the market size effect pushes towards taxing capital income in order to spare labor income and thereby preserve labor supply.

The analysis can be extended in several directions. A first avenue is to allow for different kinds of labor incomes, i.e. to introduce skilled versus unskilled labor subject to different supply elasticities. In the same vein, one could analyze the case where (skilled) labor is a direct input into R&D. A second avenue is to introduce bequests and dynastic wealth accumulation considerations into our model of

taxation with endogenous innovation. A third avenue is to use cross-country or cross-region panel data to test the relationship between growth and the tax structure interacted with variables such as the elasticity of labor supply or the size of government spending. All these extensions await further research.

## References


[1] Acemoglu, D, and U. Akcigit (2012), "Intellectual Property Rights Policy, Competition and Innovation", *Journal of the European Economic Association,* 10, 1-42.

[2] Aghion, P, and P. Howitt (1998), *Endogenous Growth Theory,* MIT Press.

[3] Aghion, P, and P. Howitt (2009), *The Economics of Growth,* MIT Press.

[4] Atkeson, A, and A. Burstein (2011). "Aggregate Implications of Innovation Policy." Mimeo.

[5] Atkinson, A and J. Stiglitz (1976). "The Design of Tax Structure: Direct Versus Indirect Taxation." *Journal of Public Economics,* 6, 55-75.

[6] Barro, R (1990), "Government Spending in a Simple Model of Endogenous Growth", *Journal of Political Economy,* 98, 103-125.

[7] Barro, R (1991), "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics,* 106, 407-443.

[8] Chamley, C (1986). "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica,* 54, 607-622.

[9] Chamley, C (2001). "Capital Income Taxation, Wealth Redistribution and Borrowing Constraints." *Journal of Public Economics,* 79, 55-69

[10] Chetty, R., Guren, A, Manoli, D, and A. Weber (2011). "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins." *American Economic Review Papers and Proceedings* 101, 1-6.

[11] Cho, J., and T. Cooley. (1994) "Employment and Hours Over the Business Cycle." *Journal of Economic Dynamics and Control* 18, 411-32.

[12] Djankov, S, Ganser, T, McLiesh, C, Ramalho, R, and A. Shleifer (2010), "The Effect of Corporate Taxes on Investment and Entrepreneurship", *American Economic Journal: Macroeconomics,* 2, 31-64.

[13] Farhi, E, and I. Werning (2011), "Capital Taxation: Quantitative Exploration of the Inverse Euler Equation", mimeo Harvard-MIT.

[14] Gentry, W, and G. Hubbard (2004), "Success Taxes, Entrepreneurial Entry and Innovation", NBER Working Paper 10551.

[15] Golosov, M, Tsyvinski, A, and I. Werning (2006), "New Dynamic Public Finance: A User's Guide", NBER Macroeconomic Annual.

[16] Goolsbee, A (1998), "Does Government R&D Policy Mainly Benefit Scientists and Engineers", mimeo Chicago GSB.

[17] Gordon, R (1986). "Taxation of Investment and Savings in a Wolrd Economy." *American Economic Review,* 76, 1086-1102.

[18] Gordon, R, and Y. Lee (2006), "Interest Rates, Taxes and Corporate Financial Policies", Working Paper, University of California at San Diego.

[19] Jones, L, Manuelli, R, and P. Rossi (1993). "Optimal Taxation in Models of Endogenous Growth." *Journal of Political Economy,* 3, 485-517.

[20] Judd, K (1985). "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics, 28*, 59-83.

[21] King, R., and S. Rebelo (1999), "Resuscitating Real Business Cycles." In *Handbook of Macroeconomics,* ed. John B. Taylor and Michael Woodford, 927-1007. Amsterdam: North-Holland.

[22] Kocherlakota, N (2010), *The New Dynamic Public Finance,* Princeton University Press.

[23] Ordover, J, and E. Phelps (1979), "The Concept of optimal Taxation in the Overlapping Generations Model of Capital and Wealth." *Journal of Public Economics,* 12, 1-26.

[24] Petrescu, I (2009), "Income Taxation and Self-Employment: The Impact of Progressivity in Countries with Tax Evasion", Working Paper, Harvard University.

[25] Phelan, C. and E. Stacchetti (2001). "Sequential Equilibria in a Ramsey Tax Model," Econometrica 69, 1491-1518.

[26] Piketty, T, and E. Saez (2012). "A Theory of Optimal Capital Taxation." NBER Working Paper #17989.

[27] Prescott, E. (2004) "Why Do Americans Work So Much More Than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 2–13.

[28] Smets, F., and R. Wouters. (2007) "Shocks and Frictions in US Business Cycle: A Bayesian DSGE Approach." *American Economic Review* 97, 586-606.

# A Appendix

## A.1 Support material for Section 2

Let $\{\lambda_{1t}, \lambda_{2t}, \lambda_{3t}, \lambda_{4t}, \lambda_{5t},\}$ denote the Lagrangian multipliers associated with the maximization program:

$$\max_{\{C_t, L_t, K_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U\left(C_t, L_t\right)$$

subject to

$$K_t^{\alpha}\left(A_t L_t\right)^{1-\alpha} + (1-\delta) K_t = C_t + G_t + K_{t+1} \tag{52}$$

$$G_t (= \gamma Y_t) = K_t^{\alpha}\left(A_t L_t\right)^{1-\alpha} - \widetilde{r}_t K_t - \widetilde{w}_t L_t \tag{53}$$

$$U_2\left(C_t, L_t\right) + \widetilde{w}_t U_1\left(C_t, L_t\right) = 0 \tag{54}$$

$$U_1\left(C_t, L_t\right) = \beta U_1\left(C_{t+1}, L_{t+1}\right)\left(\tilde{r}_{t+1} + 1 - \delta\right) \tag{55}$$

$$A_{t+1} = A_t \Phi\left(1 - \frac{\widetilde{r}_t}{\alpha K_t^{\alpha-1}\left(A_t L_t\right)^{1-\alpha}}, 1 - \frac{\widetilde{w}_t}{(1-\alpha) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}}\right) \tag{56}$$

The Lagrangian for this program is expressed as:

$$\sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{c} U\left(C_t, L_t\right) \\ +\lambda_{1t}\left(K_t^{\alpha}\left(A_t L_t\right)^{1-\alpha} + (1-\delta) K_t - C_t - G_t - K_{t+1}\right) \\ +\lambda_{2t}\left(K_t^{\alpha}\left(A_t L_t\right)^{1-\alpha} - \widetilde{r}_t K_t - \widetilde{w}_t L_t - G_t\right) \\ +\lambda_{3t}\left(U_1\left(C_t, L_t\right)\widetilde{w}_t + U_2\left(C_t, L_t\right)\right) \\ +\lambda_{4t}\left(U_1\left(C_t, L_t\right) - \beta U_1\left(C_{t+1}, L_{t+1}\right)\left(\widetilde{r}_{t+1} + 1 - \delta\right)\right) \\ +\lambda_{5t}\left(\Phi\left(1 - \frac{\widetilde{r}_t}{\alpha K_t^{\alpha-1}\left(A_t L_t\right)^{1-\alpha}}, 1 - \frac{\widetilde{w}_t}{(1-\alpha) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}}\right) - \frac{A_{t+1}}{A_t}\right) \end{array} \right\}$$

The solution to the government's maximization problem must satisfy the following first-order

conditions:

$$K_{t+1} \quad : \quad \lambda_{1t} = \beta \left( \begin{array}{c} \lambda_{1t+1} \left( \alpha K_{t+1}^{\alpha-1} \left( A_{t+1} L_{t+1} \right)^{1-\alpha} + 1 - \delta \right) \\ +\lambda_{2t+1} \left( \alpha K_{t+1}^{\alpha-1} \left( A_{t+1} L_{t+1} \right)^{1-\alpha} - \widetilde{r}_{t+1} \right) \\ +\lambda_{5t+1} \left( \begin{array}{c} -\Phi_1 \left( \tau_{k,t+1}, \tau_{l,t+1} \right) \frac{1-\alpha}{\alpha} \frac{\widetilde{r}_{t+1}}{K_{t+1}^{\alpha} (A_{t+1} L_{t+1})^{1-\alpha}} \\ +\Phi_2 \left( \tau_{k,t+1}, \tau_{l,t+1} \right) \frac{\alpha}{1-\alpha} \frac{\widetilde{w}_{t+1}}{K_{t+1}^{\alpha+1} A_{t+1}^{1-\alpha} L_{t+1}^{-\alpha}} \end{array} \right) \end{array} \right) \tag{57}$$

$$C_t \quad : \quad \left\{ \begin{array}{c} U_1 \left( C_t, L_t \right) - \lambda_{1t} + \lambda_{3t} \left( U_{11} \left( C_t, L_t \right) \widetilde{w}_t + U_{21} \left( C_t, L_t \right) \right) \\ +\lambda_{4t} U_{11} \left( C_t, L_t \right) - \lambda_{4t-1} U_{11} \left( C_t, L_t \right) \left( \widetilde{r}_t + 1 - \delta \right) \end{array} \right\} = 0 \tag{58}$$

$$L_t \quad : \quad \left\{ \begin{array}{c} U_2 \left( C_t, L_t \right) + \lambda_{1t} \left( 1 - \alpha \right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha} \\ +\lambda_{2t} \left( \left( 1 - \alpha \right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha} - \widetilde{w}_t \right) \\ +\lambda_{3t} \left( U_{12} \left( C_t, L_t \right) \widetilde{w}_t + U_{22} \left( C_t, L_t \right) \right) + \lambda_{4t} U_{12} \left( C_t, L_t \right) \\ -\lambda_{4t-1} U_{12} \left( C_t, L_t \right) \left( \widetilde{r}_t + 1 - \delta \right) \\ -\lambda_{5t} \left( \begin{array}{c} -\Phi_1 \left( \tau_{k,t}, \tau_{l,t} \right) \frac{1-\alpha}{\alpha} \frac{\widetilde{r}_t}{K_t^{\alpha-1} A_t^{1-\alpha} L_t^{2-\alpha}} \\ +\Phi_2 \left( \tau_{k,t}, \tau_{l,t} \right) \frac{\alpha}{1-\alpha} \frac{\widetilde{w}_t}{K_t^{\alpha} (A_t L_t)^{1-\alpha}} \end{array} \right) \end{array} \right\} = 0 \tag{59}$$

$$A_{t+1} \quad : \quad \left\{ \begin{array}{c} \left( \lambda_{1t+1} + \lambda_{2t+1} \right) K_{t+1}^{\alpha} \left( A_{t+1} L_{t+1} \right)^{1-\alpha} - \frac{\lambda_{5t}}{\beta} \frac{1}{A_t} + \lambda_{5t+1} \left( \frac{A_{t+2}}{A_{t+1}^2} \right) \\ +\lambda_{5t+1} \Phi_1 \left( \tau_{k,t+1}, \tau_{l,t+1} \right) \left( 1 - \alpha \right) \frac{\widetilde{r}_{t+1} A_{t+1}^{\alpha-2}}{\alpha K_{t+1}^{\alpha-1} L_{t+1}^{1-\alpha}} \\ +\lambda_{5t+1} \Phi_2 \left( \tau_{k,t+1}, \tau_{l,t+1} \right) \left( 1 - \alpha \right) \frac{\widetilde{w}_{t+1} A_{t+1}^{\alpha-2}}{(1-\alpha) K_{t+1}^{\alpha} L_{t+1}^{-\alpha}} \end{array} \right\} = 0 \tag{60}$$

$$\widetilde{w}_t \quad : \quad -\lambda_{2t} L_t + \lambda_{3t} U_1 \left( C_t, L_t \right) - \lambda_{5t} \Phi_2 \left( \tau_{k,t}, \tau_{l,t} \right) \frac{1}{\left( 1 - \alpha \right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}} = 0 \tag{61}$$

$$\widetilde{r}_t \quad : \quad \lambda_{2t} K_t + \lambda_{4t-1} U_1 \left( C_t, L_t \right) - \lambda_{5t} \Phi_1 \left( \tau_{k,t}, \tau_{l,t} \right) \frac{1}{\alpha K_t^{\alpha-1} \left( A_t L_t \right)^{1-\alpha}} = 0 \tag{62}$$

When productivity growth is exogenous: $\Phi \left( \tau_{k,t}, \tau_{l,t} \right) \equiv g$, (57) can be rewritten as:

$$1 = \beta \left( \frac{\lambda_{1t+1}}{\lambda_{1t}} \left( \alpha \left( g K_0 \right)^{\alpha-1} \left( g A_0 L \right)^{1-\alpha} + 1 - \delta \right) + \frac{\lambda_{2t+1}}{\lambda_{1t}} \left( \alpha \left( g K_0 \right)^{\alpha-1} \left( g A_0 L \right)^{1-\alpha} - \widetilde{r} \right) \right)$$

However, in BGP, $\lambda_{2t}$ and $\lambda_{1t}$ decline at rate $g$ with $\lambda_{1t} = \lambda_{10} g^{-t}$ and $\lambda_{2t} = \lambda_{20} g^{-t}$. Therefore this latter equation boils down to:

$$1 = \beta \left( \frac{1}{g} \left( \alpha K_0^{\alpha-1} \left( A_0 L \right)^{1-\alpha} + 1 - \delta \right) + \frac{1}{g} \frac{\lambda_{20}}{\lambda_{10}} \left( \alpha K_0^{\alpha-1} \left( A_0 L \right)^{1-\alpha} - \widetilde{r} \right) \right).$$

When $\Phi \left( \tau_{k,t}, \tau_{l,t} \right)$ varies with the tax rates $\tau_{k,t}$ and $\tau_{l,t}$, and again restricting attention to a BGP equilibrium with $r_t = r$ and where capital and government spending grow at the same rate $\Phi \left( \tau_{k,t}, \tau_{l,t} \right)$,

the (57) equations can be rewritten as:

$$1=\beta\left(\begin{array}{c}\frac{1}{g}\left(\alpha K_0^{\alpha-1}\left(A_0L\right)^{1-\alpha}+1-\delta\right)+\frac{1}{g}\frac{\lambda_{20}}{\lambda_{10}}\left(\alpha K_0^{\alpha-1}\left(A_0L\right)^{1-\alpha}-\widetilde{r}\right)\\+\frac{\lambda_5}{\lambda_{10}}\frac{1}{gK_0}\left(-\left(1-\alpha\right)\left(1-\tau_{k,t}\right)\Phi_1\left(\tau_{k,t},\tau_{l,t}\right)+\alpha\left(1-\tau_{l,t}\right)\Phi_2\left(\tau_{k,t},\tau_{l,t}\right)\right)\end{array}\right)$$

where

$$\lambda_{1t}=g^{-t}\lambda_{10},\ \lambda_{2t}=g^{-t}\lambda_{20},\ \lambda_{5t}=\lambda_5,\ U_1\left(C_t,L_t\right)=g^{-t}U_1\left(C_0,L\right)\ \text{and}\ K_t=g^tK_0.$$

**Proof of Proposition 3.** First note that $\tau_{l,t}<1$ (otherwise, we do not have an equilibrium with positive labor supply). Then:

$$\left(1-\tau_{l,t}\right)\frac{\lambda_5}{\lambda_{10}}a_1\alpha\frac{1}{gk_0}>0$$

Thus,

$$R+1-\delta+\frac{\lambda_{20}}{\lambda_{10}}\left(R-\widetilde{r}\right)>\widetilde{r}+1-\delta\Rightarrow$$
$$R>\widetilde{r}$$

and the result follows. ■

**Proof of Proposition 7.** Using the notation "lower-case tilde" to denote variables per effective worker (for example $\tilde{x}_t\equiv\frac{X_t}{Q_tL_t}$), we can reexpress the above system as:

$$\tilde{k}_{t+1}\frac{L_{t+1}Q_{t+1}}{L_tQ_t}+\tilde{c}_t+\frac{G_0}{L_t}=\Omega\tilde{y}_t+\left(1-\delta\right)\tilde{k}_t \tag{63}$$

$$\frac{\tilde{c}_{t+1}}{\tilde{c}_t}\frac{L_{t+1}Q_{t+1}}{L_tQ_t}=\beta\left(1-\tau_{k,t+1}\right)\alpha\frac{\tilde{y}_{t+1}}{\tilde{k}_{t+1}} \tag{64}$$

$$\phi\tilde{c}_tL_t^{1+\eta}=\left(1-\tau_{l,t}\right)\frac{1-\alpha}{\lambda}\tilde{y}_t \tag{65}$$

$$\tilde{y}_t=\tilde{k}_t^{\alpha} \tag{66}$$

$$G_0=\left(\frac{1-\alpha}{\lambda}\tau_{l,t}+\alpha\tau_{k,t}\right)L_t\tilde{y}_t \tag{67}$$

$$\frac{Q_{t+1}}{Q_t}=1+\frac{\left(\lambda-1\right)^2}{\lambda}\frac{1-\alpha}{\psi}\tilde{y}_tL_t \tag{68}$$

Solve (67) for $G_0/L_t$ and substitute into (63). Similarly, use conjecture $\tilde{c}_t=\left(1-s\right)\tilde{y}_t$ in (64) and

solve for $\tilde{k}_{t+1}\frac{Q_{t+1}}{Q_t}$ and substitute into (63). This gives us

$$\beta\left(1-\tau_{k,t}\right)\alpha+\left(1-s\right)+\frac{1-\alpha}{\lambda}\tau_{l,t}+\alpha\tau_{k,t}=\Omega$$

which then solves for $s$ as in (50).

To find $L^*$ just use (65) and the conjecture. ■

**Proof of Proposition 8.** Note that the growth rate is

$$g_\lambda^*=1+\frac{\left(\lambda-1\right)^2}{\lambda}\frac{1-\alpha}{\psi}\tilde{y}^*L^*$$

Note that from (64) we get

$$\tilde{k}_{\lambda=1}^*=\left[\beta\left(1-\tau_{k,t}\right)\alpha\right]^{\frac{1}{1-\alpha}}$$

since $Q_{t+1}=Q_t$ when $\lambda=1$. Then the $2^{nd}$ order Taylor approximation is simply

$$\begin{aligned}
g_\lambda^* &\approx g_{\lambda=1}^*+\left(\lambda-1\right)\left.\frac{dg}{d\lambda}\right|_{\lambda=1}+\frac{\left(\lambda-1\right)^2}{2}\left.\frac{d^2g}{d\lambda^2}\right|_{\lambda=1}\\
&= 1+\frac{\left(\lambda-1\right)^2}{2}2\frac{1-\alpha}{\psi}\tilde{y}_{\lambda=1}^*L_{\lambda=1}^*\\
&= 1+\left(\lambda-1\right)^2\frac{1-\alpha}{\psi}\left[\beta\left(1-\tau_{k,t}\right)\alpha\right]^{\frac{\alpha}{1-\alpha}}\left[\frac{\left(1-\tau_{l,t}\right)\left(1-\alpha\right)}{\phi\left(1-\left(1-\alpha\right)\tau_{l,t}-\alpha\tau_{k,t}-\beta\left(1-\tau_{k,t}\right)\alpha\right)}\right]^{\frac{1}{1+\eta}}
\end{aligned}$$

■

## A.2 Euler Equation Errors

In figure A1, we plot the Euler equation errors (Judd, 1992) of the model solved with our third-order perturbation for the benchmark calibration as a function of the log of (normalized) capital. Following standard practice, we plot the decimal log of the absolute value of the Euler equation error. For values close to the normalized level at the rescaled steady state, the Euler equation errors are around -10 (one can interpret this number as equivalent to making a mistake of \$1 for each \$10 billion spent). Hence, the perfomence of the model in terms of accuracy is most satisfactory. Even if we move quite away from the rescaled steady state (and remember that we are normalizing all the variables by the level of technology and that, consequently, the approximation stays relevant along the BGP), the Euler equation errors are still below -5.

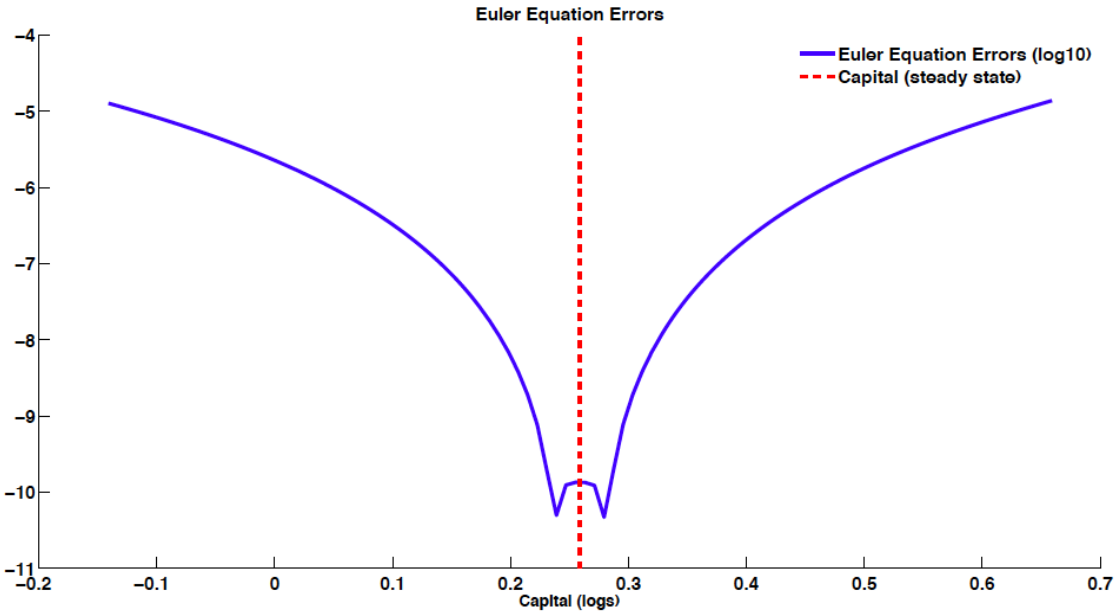



Figure A1: Euler Equation Errors