Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 13-025

"Optimal Capital Versus Labor Taxation
With Innovation-Led Growth"

by

Philippe Aghion, Ufuk Akcigit, and Jesús Fernández-Villaverde

http://ssrn.com/abstract=2272651

Optimal Capital Versus Labor Taxation with Innovation-Led Growth*

Philippe Aghion† Ufuk Akcigit‡ Jesús Fernández-Villaverde§

May 23, 2013

Abstract

Chamley (1986) and Judd (1985) showed that, in a standard neoclassical growth model with capital accumulation and infinitely lived agents, either taxing or subsidizing capital cannot be optimal in the steady state. 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. Our main result is that, for a given required trend of public expenditure, a zero tax/subsidy on capital becomes suboptimal. In particular, the higher the level of public expenditure and the income elasticity of labor supply, the less should capital income be subsidized and the more it should be taxed. Not taxing capital implies that labor must be taxed at a higher rate. This in turn has a detrimental effect on labor supply and therefore on the *market size* for innovation. At the same time, for a given labor supply, taxing capital also reduces innovation incentives, so that for low levels of public expenditure and/or labor supply elasticity it becomes optimal to subsidize capital income.

JEL Codes: H2, O3, O4.

Keywords: Capital tax, labor tax, optimal taxation, innovation, R&D, growth.

*We are particularly grateful to Raj Chetty for his continuous help and guidance throughout this project. We also thank Daron Acemoglu, Manuel Amador, Andy Atkeson, Tim Besley, Richard Blundell, Ariel Burstein, Emmanuel Farhi, Peter Howitt, Caroline Hoxby, Louis Kaplow, Huw Lloyd-Ellis, Pietro Peretto, Torsten Persson, Thomas Piketty, and John Seater, as well as seminar participants at IIES (Stockholm University), Harvard, CIFAR, Brown, Chicago Booth, UCLA, the EFJK Economic Growth Conference, the Canadian Macro Study Group, and the SKEMA Workshop on Economic Growth for very helpful comments and suggestions. Luigi Bocola provided outstanding research assistance. Finally, we acknowledge the NSF for financial support.

†Harvard University, NBER, and CIFAR 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. 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; thus, taxing the revenues from savings amounts to taxing labor income twice, and it introduces unnecessary distortions in consumption versus saving decisions. Somewhat surprisingly, growth considerations do not play much of a role in this debate. However, recent history suggests that appropriate reforms of the tax structure can spur innovation and growth in a relatively short time period.[1]

This paper tries to fill this gap by analyzing how the welfare-maximizing tax structure is affected by the introduction of endogenous technical progress. In line with new growth theory, which used the neoclassical models of Solow and Cass-Koopmans-Ramsey as its benchmark, here we use as a starting point the optimal taxation models of Chamley (1986) and Judd (1985). These two papers analyze labor versus capital taxation in the context of a standard neoclassical growth model with capital accumulation and infinitely lived agents sharing the same intertemporal utility function (not necessarily separable). A main finding is that taxing capital cannot be optimal in the steady state (i.e., in the long run). The underlying intuition is well explained by Salanié (2003): if τ 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 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.[2] By doing so, we show the usefulness of this class of models for optimal fiscal policy design and how understanding the mechanism behind innovation may change well-known policy prescriptions.

Our main result is that, for a given required trend of public expenditure, a zero tax/subsidy on capital becomes suboptimal. In particular, the higher the level of public expenditure and the income

[1] Prior to the 1991 tax reform, Sweden had a progressive tax on capital income with a maximum marginal rate of 72% and an average rate of 54%. The 1991 reform turned it into a flat 30% rate. Combined with selected reductions in public spending, this reform is believed to have contributed to a sharp increase in the savings rate (from 9% in 1990 to 12,86% in 2011), to a significant increase in labor supply, and a large reduction in fiscal evasion- both of which resulted in an increase in fiscal revenues in spite of the lower average tax rate, and to a jump in patenting (between 1990 and 2010, the annual number of patents per thousand inhabitants increased from 1 to 2,5). See Aghion and Roulet (2011) and OECD statistics.

[2] See Aghion and Howitt (1992) and Aghion *et al* (2013).

elasticity of labor supply, the less should capital income be subsidized and the more it should be taxed. Not taxing capital implies that labor must be taxed at a higher rate. This in turn has a detrimental effect on labor supply and therefore on the *market size* for innovation. At the same time, for a given labor supply, taxing capital also reduces innovation incentives (again a *market size effect*).[3] As a result, for sufficiently low levels of public expenditure and/or labor supply elasticity, it becomes optimal to subsidize capital income.

Our analysis relates to three main strands of the literature. First is the literature on optimal capital taxation. In response to the seminal contributions by Atkinson and Stiglitz (1976), Chamley (1986) and Judd (1985), all of which point to the optimality of zero capital income taxation, various attempts have been made to overturn 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 in enhancing 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.[4] However, none of these contributions[5] factor in the potential effects of taxation on growth, whereas we contribute to this whole literature by introducing (innovation-based) growth into the analysis.

A second related literature is that on taxation and growth. The first major attempt at looking at the relationship between the 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 cor-

[3] In addition, taxing capital income also discourages production by intermediate monopolists. This *market power effect* is also absent from the original Chamley-Judd model; which instead assumes full competition.

[4] 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)

[5] And the same is also true for the New Dynamic Public Finance literature (e.g., see Kocherlakota 2005, Golosov *et al.* 2006, Farhi and Werning, 2012).

porate 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.[6] 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.* (2010) find that the effective corporate tax rate has a large adverse impact on aggregate investment, FDI, and entrepreneurial activity. However, none of these papers is concerned with the choice between labor and capital income tax and more generally with the optimal design of the tax system in a dynamic framework with endogenous technical progress.[7]

More closely related to our analysis is the paper by Atkeson and Burstein (2012). This paper looks at the impact of fiscal policy on product innovation and welfare, using a growth model with expanding varieties. However, their focus is on the comparison between R&D tax credits, federal expenditures on R&D, and corporate profit tax, whereas we focus on labor versus capital income taxation and the comparison with Chamley (1986) and Judd (1985).[8]

A third strand is a recent empirical literature documenting the importance of market size as a determinant of innovation incentives.[9] In this paper, we contribute to this literature by showing how the market size effect translates into a simple expression for the aggregate growth rate as a function of the tax rates on capital and labor income, which in turn indicates very clearly how the optimal tax rate on capital income differs from the zero rate generated by Chamley and Judd.

The remainder 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

[6] More recent evidence of the effect of tax policy on growth includes Gentry and Hubbard (2004), Alesina and Ardagna (2010), Romer and Romer (2010), Barro and Redlick (2011), and Mertens and Ravn (2012).

[7] Notable exceptions consider the effect of tax structure on growth in AK models with human capital investments (see Jones, Manuelli and Rossi 1993 and Milesi-Ferretti and Roubini 1998). In these models, taxing labor rather than capital is always detrimental to growth by discouraging human capital accumulation. As we shall see below, innovation-led growth leads to more nuanced conclusions: namely, depending on the value of labor supply elasticity and on the share of government expenditure in GDP, innovation-led growth may push toward taxing capital income or toward subsidizing it. More precisely, the market size effect pushes toward subsidizing capital income more for lower values of the labor supply elasticity and of the share of government expenditure, whereas it pushes toward taxing capital more for high values of these variables.

[8] As we were completing this draft we heard about parallel work by Jaimovic and Rebelo (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 out 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 capital/profit tax can never be positive in their model.

[9] See Acemoglu and Linn (2004) and Blume-Kohout and Sood (2013), and Duggan and Morton (2010).

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. Section 5 discusses some extensions. Section 6 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 analyze an environment in which the government can fully commit to future tax rates.[10]

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 study a reduced-form endogenous growth model where productivity growth depends on the tax structure and we explain why, in general the zero capital taxation result no longer obtains in that case. Section 3 will develop a full-fledged endogenous growth model to rationalize the reduced-form specification of subsection 2.3.

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 $\{K_t\}_{t=0}^{\infty}$ 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

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

[10] To make this problem interesting, and in line with Chamley (1986), Judd (1985), and the theoretical literature on optimal taxation, we do not allow for lump-sum taxation. Full commitment is a natural first step in the investigation of optimal policy.

In this budget constraint, w_t, $r_t, \tau_{k,t}$, $\tau_{l,t}$, and δ 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 in labor such that[11]

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

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

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:

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

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

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

$$Q_{t+1} = Q_t \Phi(\tau_{k,t}, \tau_{l,t}). \tag{7}$$

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

[11] $U_n(x_1, x_2)$ denotes the partial derivative of $U(.,.)$ with respect to $x_n : U_n(x_1, x_2) \equiv \partial U(x_1, x_2)/\partial x_n$, $n \in \{1, 2\}$.

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 $\{C_t, L_t, K_{t+1}, Q_t, \tau_{k,t}, \tau_{l,t}\}_{t=0}^{\infty}$ to maximize household utility (1) subject to:

1. the economy's resource constraint

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

   which says that current final output plus capital net of depreciation provides 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;[12]

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

4. the process of growth (7).

Using the "Euler Theorem" whereby

$$Y_t = K_t^{\alpha} (Q_t L_t)^{1-\alpha} = r_t K_t + w_t L_t,$$

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

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

where

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

In this section, we assume that G_t grows at the same rate as output. Government expenditures such as education, health care, or police, are basically wages paid to government employees. As the

[12] In section 5, we extend the analysis to the case where the government can raise public debt.

economy grows, so do the wages of teachers, doctors, and police officers and at the same rate as the economy's growth rate. We will come back to this point in Section 3.[13]

Note that

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

Thus, the government's problem can be summarized as

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

subject to

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

$$G_t(=\gamma Y_t)=K_t^{\alpha}\left(Q_tL_t\right)^{1-\alpha}-\widetilde{r}_tK_t-\widetilde{w}_tL_t \tag{9}$$

$$U_2\left(C_t,L_t\right)+\tilde{w}_tU_1\left(C_t,L_t\right)=0 \tag{10}$$

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

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

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: $\Phi\left(\tau_{k,t},\tau_{l,t}\right)\equiv 1$; (ii) exogenous growth: $\Phi\left(\tau_{k,t},\tau_{l,t}\right)\equiv g$; and (iii) endogenous growth: $\Phi\left(\tau_{k,t},\tau_{l,t}\right)$ 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 ($\Phi\left(\tau_{k,t},\tau_{l,t}\right)\equiv 1$ 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:

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

[13] If anything, Wagner's law suggests that the cost of many of these public services grows faster than output. It is plausible to think that education or health care are goods with income elasticities larger than one and that the political economy process delivers public spending on education and health care that outgrows output. Note also that, as in the standard Ramsey approach, we are dealing with G_t, and not with transfers (although these are also likely to be linked with output).

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:

$$\widetilde{r} = \alpha K^{\alpha-1} \left(QL \right)^{1-\alpha} = r,$$

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: $\Phi\left(\tau_{k,t}, \tau_{l,t}\right) \equiv g$. 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 (for that BGP to exist, we need to impose additional conditions on the utility function, for instance, that the $\frac{U_2(C_t,L_t)}{U_1(C_t,L_t)}$ is linear in C_t).[14] 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:

$$1 = \beta \left(\frac{\lambda_{1t+1}}{\lambda_{1t}} \left(\alpha \left(gK_0 \right)^{\alpha-1} \left(gQ_0L \right)^{1-\alpha} + 1 - \delta \right) + \frac{\lambda_{2t+1}}{\lambda_{1t}} \left(\alpha \left(gK_0 \right)^{\alpha-1} \left(gQ_0L \right)^{1-\alpha} - \widetilde{r} \right) \right). \tag{15}$$

However, note that in BGP, λ_{2t} and λ_{1t} decline at rate g with $\lambda_{1t} = \lambda_{10} g^{-t}$ and $\lambda_{2t} = \lambda_{20} g^{-t}$. Therefore, equation (15) boils down to:

$$1 = \beta \left(\frac{1}{g} \left(\alpha K_0^{\alpha-1} \left(Q_0L \right)^{1-\alpha} + 1 - \delta \right) + \frac{1}{g} \frac{\lambda_{20}}{\lambda_{10}} \left(\alpha K_0^{\alpha-1} \left(Q_0L \right)^{1-\alpha} - \widetilde{r} \right) \right). \tag{16}$$

Next, the Euler equation (11) can be re-expressed as:

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

Then, since $U_1\left(C_t, L_t\right)$ and λ_{1t} decline at the same rate g,[15] so that $U_1\left(C_t, L_t\right) = g^{-t} U_1\left(C_0, L\right)$,

[14] For the BGP to exist, we need to impose additional conditions on the utility function. For instance, if $\frac{U_2(C_t,L_t)}{U_1(C_t,L_t)}$ is linear in C_t, we can show that such a BGP exists.

[15] This follows from equation (62) in the appendix.

the Euler equation (17) is simply:

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

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

$$\widetilde{r} = \alpha K_0^{\alpha-1} \left(Q_0 L \right)^{1-\alpha}$$

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 $\Phi\left(\tau_{k,t}, \tau_{l,t}\right)$ 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 $\Phi\left(\tau_{k,t}, \tau_{l,t}\right)$. In this equilibrium, the first-order conditions for the government problem become:

$$1 = \beta \left(\begin{array}{c} \frac{1}{g}\left(\alpha K_0^{\alpha-1}\left(Q_0 L\right)^{1-\alpha} + 1 - \delta\right) + \frac{1}{g}\frac{\lambda_{20}}{\lambda_{10}}\left(\alpha K_0^{\alpha-1}\left(Q_0 L\right)^{1-\alpha} - \widetilde{r}\right) \\ + \frac{\lambda_5}{\lambda_{10}}\frac{1}{gK_0}\left(-\left(1-\alpha\right)\left(1-\tau_{k,t}\right)\Phi_1\left(\tau_{k,t},\tau_{l,t}\right) + \alpha\left(1-\tau_{l,t}\right)\Phi_2\left(\tau_{k,t},\tau_{l,t}\right)\right) \end{array} \right) \tag{19}$$

where

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

As before, from (11) we have the household's Euler equation:

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

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, we have the main result of this section:

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

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

We now open the black box of the Φ 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.

3.1 Basic Environment

3.1.1 Household

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

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

to ensure the existence of a BGP, where $\beta \in (0,1)$ is the discount factor, C_t is consumption, L_t is labor supply, η 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 of the representative household is:

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

where K_t, Π_t, X_t, r_t, and w_t stand for physical capital, gross profit, R&D expenses, the rental rate of capital, and the 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 than or equal to 1 (that is, the government cannot tax more than the total income), but we let them 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 increases 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 a higher tax on labor and a higher scale factor of disutility reduce the labor supply.[16]

3.1.2 Production

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 the numeraire, we then have, by the Euler Theorem:

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

[16] The reader should bear in mind that there are several sources of income for the representative household, not just labor income. This in turn implies that the income and substitution effects of an increase in the labor income tax rate τ_l do not exactly offset one another: here, an increase in the labor income tax, which keeps capital income unaffected, will induce the individual to save more and to supply less labor.

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:

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

3.1.3 Profits

At the beginning of any period t, in each sector i, one firm has the opportunity to innovate. A successful innovation increases productivity in the sector from $q_{i,t-1}$ to $q_{i,t} = \lambda q_{i,t-1}$, where $\lambda \geq 1$ is the size of the innovation, and in this case, the innovator becomes a local monopoly for one period in sector i. If innovation does not succeed, then $q_{i,t} = q_{i,t-1}$ and the monopoly power is randomly allocated to an individual in the economy. In both cases, there is always a fringe of firms that can potentially produce the intermediate input using the previous technology:

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

Now let us compute the equilibrium profit of a successful innovator (profit is zero if innovation does not succeed). Potential fringe producers have a marginal cost equal to $w_t \lambda / q_{i,t}$. Bertrand competition implies that this is the limit price the latest innovator in sector i can charge. Thus:

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

It follows that the equilibrium profit of a successful innovator in sector i at time t is equal to:

$$\pi_{i,t} = (p_{i,t} - MC_{i,t})\, z_{i,t}$$

or, using (29) and (27):

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

Finally, using the fact that spending on intermediate inputs $\hat{Z}_t$ accounts for a fraction $(1-\alpha)$ of total final output and that the production technology (25) is Cobb-Douglas, we get:

$$\pi_{i,t} = \frac{\lambda - 1}{\lambda}(1-\alpha)Y_t. \tag{30}$$

In this expression we can see how profit is determined by the size of the market, Y_t, and 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.

3.1.4 Factor Shares

Let us define the aggregate productivity Q_t at date t by:

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

Using (25), (26), (27), (28), and (29), and equating labor supply to labor demand (i.e., using $L_t = \int_0^1 l_{i,t} di$), one can show:

Proposition 4 *(1) The production technology* (25) *for the final good can be re-expressed in the reduced form:*

$$Y_t = K_t^{\alpha} (Q_t L_t)^{1-\alpha}; \tag{31}$$

(2) the equilibrium shares for capital, labor and profits are respectively equal to:

$$\frac{r_t K_t}{Y_t} = \alpha, \quad \frac{L_t w_t}{Y_t} = \frac{1}{\lambda}(1-\alpha), \quad \textit{and} \ \frac{\Pi_t}{Y_t} = \frac{\lambda - 1}{\lambda}(1-\alpha). \tag{32}$$

Proof. Using (28) and aggregating over all intermediate inputs, we get:

$$Q_t L_t = Z_t. \tag{33}$$

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

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

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

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

or

$$w_t L_t = \frac{1-\alpha}{\lambda} Y_t. \tag{34}$$

This, together with the Euler Theorem, which implies that capital income accounts for a fraction α of final output (which itself results from the Cobb-Douglas production technology (25)), establishes Part 2 of the Proposition.

Part 1 is then proved as follows. First, dividing (33) by (34) yields:

$$w_t = \frac{(1-\alpha)}{\lambda} Q_t \left(\frac{K_t}{Z_t}\right)^{\alpha}. \tag{35}$$

Then, combining (33), (34), and (35) yields:

$$Y_t = K_t^{\alpha} \left(Q_t L_t\right)^{1-\alpha}.$$

■

From this Proposition we learn that the share of capital, labor, and profits is determined by two parameters: α and λ. Moreover, (a) 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); (b) aggregate profits are proportional to aggregate final output. This observation implies 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.

3.1.5 R&D, 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 clarify 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, and 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 the production of intermediate inputs, 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 the final good to innovate at rate $x_{i,t}/\bar{x}_t$, where $\bar{x}_t$ is the aggregate innovation intensity of other potential innovators.[17]

To streamline the exposition, in the remaining part of this section we will assume that firms' R&D spending is not publicly verifiable. This implies that the government cannot directly subsidize R&D costs.[18] That R&D costs are not fully verifiable is a realistic assumption. First, in practice firms often use R&D subsidies for other purposes, for example, for advertising or other devices to limit the threat of entry. Second, R&D is only the tip of a broader innovation investment iceberg, which includes the time and effort entrepreneurs spend to solve problems and to find new and better ways of doing things

[17] This linear R&D technology simplifies the analysis by delivering simpler expressions for the equilibrium R&D intensity and the equilibrium growth rate. Effectively, it leads to the same first-order condition as with a quadratic cost function (see equation 37), that is, as when solving:

$$\max_{x_{it}} \left\{ x_{it}\pi_{it} - x_{it}^2 \psi Q_t/2 \right\},$$

but our formulation has the advantage of keeping the aggregate R&D investment in the aggregate resource constraint (44) linear in Y_t. Note that the congestion externality we introduce does not have any additional implication for the aggregate dynamics due to the linear structure. In particular, both the above convex cost and the linear cost in (36) imply the same innovation rate x_{it} and exactly the same aggregate growth rate $G_t = 1 + (\lambda - 1)\, x_{it}$. Hence, our findings do not hinge on this form and the qualitative analysis and results extend to more general convex R&D technologies. An alternative way to motivate the above innovation technology is to assume that with probability x_{it}, a potential innovator produces a *radical* innovation that improves productivity by λ and becomes the monopolist, and with probability $x_{it}\left(\frac{1}{\bar{x}_t} - 1\right)$, the potential innovator produces an *incremental* innovation that improves the latest technology by $\epsilon \to 1$. Note that the incremental innovation probability increases in her own effort x_{it}, and decreases in $\bar{x}_t$, which in turn reflects a congestion externality.

[18] The calibrations in Section 5 suggest that our main findings are robust to allowing for R&D subsidies.

(and it also includes the time and effort spent by potential inventors "in their backyard," which again can hardly be identified and measured by a government). The corresponding opportunity cost of time is hard for a government to identify and therefore to subsidize.

If $\tau_{\pi,t}$ denotes the tax rate on innovation profits at date t, the resulting maximization problem for innovators is:[19]

$$\max_{x_i} (1 - \tau_{\pi,t}) \left(\frac{x_{i,t}}{\bar{x}_t} \pi_{i,t} - x_{i,t} \psi Q_t \right) \tag{36}$$

In a symmetric equilibrium, and using (30), the innovation rate is given by:

$$x_{i,t} = \bar{x}_t = \frac{\lambda - 1}{\lambda} \frac{1 - \alpha}{\psi} \frac{Y_t}{Q_t}, \tag{37}$$

which leads to the equilibrium growth rate:

$$\frac{Q_{t+1}}{Q_t} = 1 + (\lambda - 1)\bar{x}_t = 1 + \frac{(\lambda - 1)^2}{\lambda} \frac{1 - \alpha}{\psi} \frac{Y_t}{Q_t}. \tag{38}$$

This growth rate depends negatively on the share of capital α and the cost of R&D ψ but positively on markups λ, and normalized market size Y_t/Q_t. Thus, a growth-maximizing policy should aim to maximize current output. The corresponding *market size effects* lie at the heart of our analysis.

Finally, the total amount of the final good used in R&D in equilibrium is:

$$X_t = \bar{x}_t \psi Q_t = \frac{\lambda - 1}{\lambda} (1 - \alpha) Y_t = (1 - \Omega) Y_t \tag{39}$$

where $\Omega \equiv 1 - \frac{\lambda - 1}{\lambda} (1 - \alpha)$. Note that (32) and (39) imply that equilibrium aggregate profits net of R&D costs are zero:

$$\Pi_t - X_t = 0.$$

3.1.6 Government

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

$$G_t = Q_t G_0, \tag{40}$$

[19] Here we implicitly assume that *net* profits are verifiable and thus can be taxed. However, as we shall see below, in equilibrium these profits are equal to zero, so that in the end, it makes no difference to our analysis and results whether net profits are verifiable or not.

where Q_t evolves over time according to (38) and G_0 is exogenously given.[20] This assumption captures the idea that many government expenses grow with the economy. As we mentioned in Section 2, 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.

The government will choose the tax policy $\{\tau_{k,t}, \tau_{l,t}, \tau_{\pi,t}\}_{t=0}^{\infty}$ on capital, labor, and net profits to maximize intertemporal welfare[21] subject to the above dynamic equilibrium conditions, together with the aggregate resource constraint of the economy, namely

$$Y_t + (1-\delta) K_t = K_{t+1} + C_t + G_t + X_t \tag{42}$$

and the balanced budget constraint

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

where $\hat{\Pi}_t \equiv \Pi_t - X_t$. Using the fact that equilibrium aggregate net profits $\hat{\Pi}_t$ are equal to zero, the choice of $\{\tau_{\pi,t}\}_{t=0}^{\infty}$ is irrelevant and the above condition boils down to:

$$G_t = \tau_{l,t} w_t L_t + \tau_{k,t} r_t K_t. \tag{43}$$

[20] In Section 5, we will look at the case where G enters the utility function of the representative household and is thus also optimally chosen by the government.

[21] On the BGP, welfare is equal to:

$$\begin{aligned} W(\overrightarrow{\tau}) &= \max_{C_0, L_0} \sum_{t=0}^{\infty} \beta^t \left(\ln C_0 g^t - \phi \frac{L_0^{1+\eta}}{1+\eta} \right) \\ &= \max_{C_0, L_0} \left\{ \frac{\ln C_0}{1-\beta} + \frac{\beta}{(1-\beta)^2} \ln g - \frac{\phi L_0^{1+\eta}}{(1-\beta)(1+\eta)} \right\} \end{aligned} \tag{41}$$

where g is the growth rate and C_0 and L_0 are the levels of consumption and labor supply at the start of the BGP.

3.2 Equilibrium Balanced Growth: Analytical Results

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 maximizes (conditions (23) and (24)), all firms maximize (conditions (32) and (38)), the government satisfies its budget constraint (43), and markets clear (conditions (25), (39), and (42)).*

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

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

Note that the previous equations would be the same in a neoclassical growth model except that $\Omega = \lambda = 1$. Hence our economy is a minimal departure from the standard model in business cycles or optimal taxation problems.

Similarly, we can define a balanced 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.*

To simplify the analysis, in the following proposition we shall concentrate 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:*
(1) *in a 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 $L^* = \left[\frac{1-\alpha}{\phi(1-s)\lambda}\right]^{\frac{1}{1+\eta}}$ *and*

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

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

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

where $L_t^* = \left[\frac{(1-\tau_{l,t})(1-\alpha)}{\phi(1-s_t)\lambda}\right]^{\frac{1}{1+\eta}}$ *and:*

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

Proof. See the Appendix. ■

This lemma implies that equilibrium labor supply is decreasing in the tax rate on labor. More important, it also implies 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* λ *close to 1, the steady-state growth rate of the benchmark economy is approximately equal to*

$$g_t^* = 1 + (\lambda-1)^2\,\frac{1-\alpha}{\psi}\left[\beta(1-\tau_{k,t})\,\alpha\right]^{\frac{\alpha}{1-\alpha}}\left[\frac{(1-\tau_{l,t})(1-\alpha)}{\phi\left(1-(1-\alpha)\tau_{l,t}-\alpha\tau_{k,t}-\beta(1-\tau_{k,t})\,\alpha\right)}\right]^{\frac{1}{1+\eta}}$$

Proof. See the Appendix. ■

Intuitively, the higher the tax rate on labor income the lower the amount of labor supply

$$L_t^* = \left[\frac{(1-\tau_{l,t})(1-\alpha)}{\phi\left(1-(1-\alpha)\tau_{l,t}-\alpha\tau_{k,t}-\beta(1-\tau_{k,t})\,\alpha\right)}\right]^{\frac{1}{1+\eta}}$$

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* may 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 will compute the welfare-maximizing tax structure for various parameter values. For given public spending and tax policy, the equilibrium welfare is given by (21). The welfare-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. Thus, we shall decompose the departure from the Chamley-Judd welfare-maximizing tax rates into the part that comes from the market power and the part that comes from the market size effect.

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 have only 8 parameters: β, ϕ, η, α, λ, δ, ψ, 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 $\beta = 0.98$, to generate an annual interest rate of 4 percent, $\phi = 6.7$ to make hours worked to be around 1/3, and in our baseline calibration we take $\eta = 0.833$ (which implies a Frisch elasticity of 1.2) following the evidence surveyed in Chetty *et al.* (2011).[22] 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 an R&D share on GDP 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 GDP of 0.035 corresponds to the company-funded R&D to sales ratio in the U.S. (N.S.F., 2010). Moreover, the innovation step size λ corresponds to the benchmark value of Acemoglu and Akcigit (2012). A depreciation of $\delta = 0.06$ matches a capital/output ratio of around 3 (close to the one observed in the U.S. economy) and $\psi = 0.05$ delivers (conditional on the other parameters) a growth rate of 1.92 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

[22] But, below, in our first calibration exercise we allow this elasticity to vary over a whole interval from 0.5 to 2.5.

of its possible values: from 0.1 to 0.215 (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 at which 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 and 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. We 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 a steady state. Analogously, we define the log deviation of such a 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 productivity, 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, and 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 Calibration

Our main exercise below will be 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 18 to 45 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 and to make the intuition of the result more transparent, and because we do not find complicated time-dependent policies such as those implied by a pure Ramsey analysis either plausible or empirically relevant. Second, we balance the budget period by period to avoid handling an extra state variable. Also, given our choice of a constant tax rate on capital, this constraint is less important than it might seem.[23] Third, 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 in the lower range of the observed rates.[24] Finally, remember that the tax on profits is irrelevant because profits net of R&D are zero and thus we do not need to take any stand on it.

However, before we engage in this main computational experiment, a preliminary exercise where we illustrate the role of market size in the optimal tax on capital, the main mechanism at work in our paper, will be most helpful at building intuition.

4.4 Tax Rate on Capital for Varying Labor Supply Elasticity

Figure 1 depicts the growth-maximizing tax rates on capital income as a function of the Frisch elasticity of labor supply, for two levels of G_0, a low level of 0.14 (discontinuous red line) and a high level of 0.18 (continuous blue line). The ratio G_0/Y implied by the figure goes from 0.15 (low G_0, low Frisch elasticity) to 0.60 (high G_0, high Frisch elasticity). That is, instead of finding the combination of tax rates on capital and labor income that maximizes welfare, we simply find the combination that maximizes the growth rate along the BGP of our economy. We could also have reported the welfare-maximizing tax rate. The lessons 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 1, initial conditions for capital stock that are sensible for some labor elasticities are not particularly good choices for other elasticities. This matters because the initial condition induces a transitional dynamics that can potentially obscure the effects we are highlighting. By focusing on the growth-maximizing tax rate,

[23] In the standard Ramsey approach, capital at time 0 is already in place and, hence, optimal government policy involves, in general, a large tax on capital income in that initial period. The tax revenue is used to buy assets of the private sector and to rely on the income generated by those assets in the future to reduce the need of raising distortionary taxation. That is, the relevant constraint in our exercise is not that the government cannot accumulate debt -something it would not like to do anyway,- but that the government cannot accumulate positive assets. Because of the arguments of empirical plausibility outlined in the main text, we are not particularly concerned about that constraint.

[24] In unreported additional sensitivity analysis we found that our results were robust to the choice of initial capital.

we avoid this problem.[25]

In Figure 1, we selected a range for the Frisch elasticity between 0.5 and 2.5 that encompasses all empirically relevant values. The lower range of this interval corresponds more to what we find in the empirical public finance literature (e.g., see Bianchi *et al.*, 2001, or Pistaferri, 2003), whereas the upper range is more in line with the macroeconomic literature on business cycles and fiscal policy (e.g., see Cho and Cooley, 1994, King and Rebelo, 1999, Prescott, 2004, and Smets and Wouters, 2007).[26]

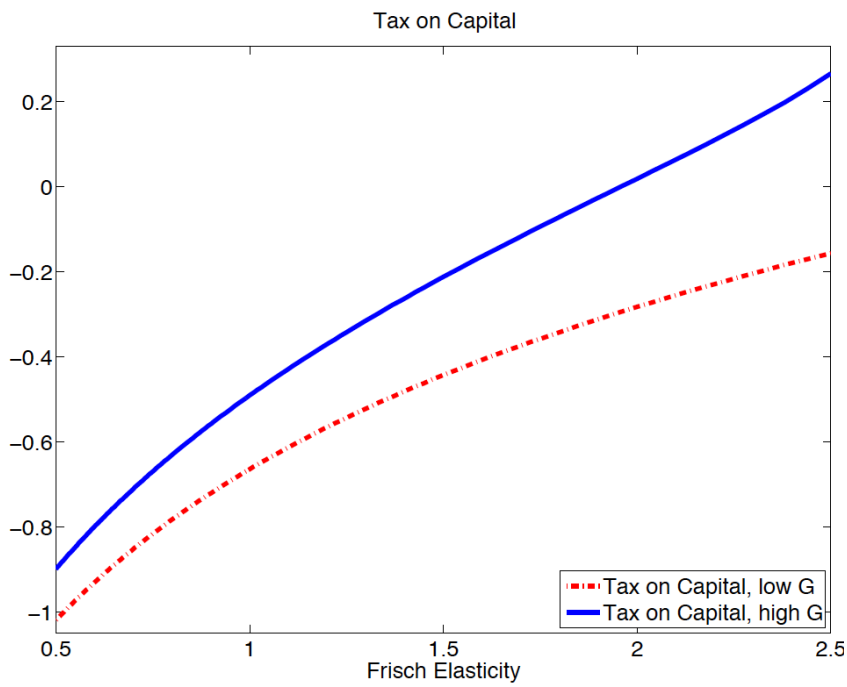

Figure 1: Growth-maximizing tax on capital as a function of the Frisch labor supply elasticity.

In Figure 1 we can see that the growth-maximizing tax on capital income may be as low as −100 percent (a subsidy as high as the rate of return) or as high as 27 percent. Figure 1 teaches us several lessons. The first lesson is that there is a positive relationship between the optimal tax rate on capital and the elasticity of labor supply. The intuition is simple: when labor is not very elastic, the optimal policy is to tax labor heavily (since it will have a small impact on hours worked) and use the proceedings to finance G and to subsidize capital. That is, market size considerations tilt fiscal policy against labor income and in favor of capital income. The opposite effect occurs when labor is

[25] Also, for very high Frisch elasticities of labor supply and high ratios G_0/Y, the model often enters into regions of indeterminacy if we keep ϕ at the value calibrated in the benchmark case. Fortunately, indeterminacy is not an issue for more empirically relevant regions of the parameter space.

[26] Note that we are keeping ϕ constant as we change the Frisch elasticity to make the numbers more easily interpretable. An alternative exercise would be to vary the Frisch elasticity and, simultaneously, change ϕ such that the average hours worked are always the same as in the data. We find this alternative experiment to be much less transparent.

very elastic: the optimal tax policy relies on taxing capital to finance G and avoid having to tax labor income. 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. Similarly, often those supporters of low taxation of capital like to assume high elasticities of labor supply in their models. The market size mechanism turns both sides of the argument around.

The second lesson is that the optimal tax on capital income is increasing in G. The reason is that, since we need to finance a larger amount of expenditures, it is relatively more costly to subsidize capital income and the distortions on labor grow bigger for all Frisch elasticities.

In summary, Figure 1 teaches us that high labor supply elasticities and high G are associated, through the market size mechanism, with high taxes on capital income. Conversely, low labor supply elasticities and low G are associated with subsidies to capital.

4.5 Welfare-Maximizing Tax Rates for Varying Government Expenditure

We now fix labor supply elasticity at its calibrated value of 1.2 and focus on how the welfare-maximizing optimal tax rates on capital and labor income vary with the share of government expenditure in output. Figure 2 shows the welfare-maximizing tax rates for capital and labor as G/Y goes from 18.16 percent to 48.60 percent. Over that range, the welfare-maximizing tax rate increases from -20.80 percent to 9.4 percent.

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 and 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 at which it eventually becomes positive.

In Figure 3, we illustrate how the differences in the BGP growth rates implied by our range of government expenditure are significant, but not too large: from 2.01 to 1.76 percent at an annual rate. We checked that a similar finding occurred when we depart from optimal policy: fiscal policy does change the rate of growth of the economy (and hence it has a significant effect on welfare), but those

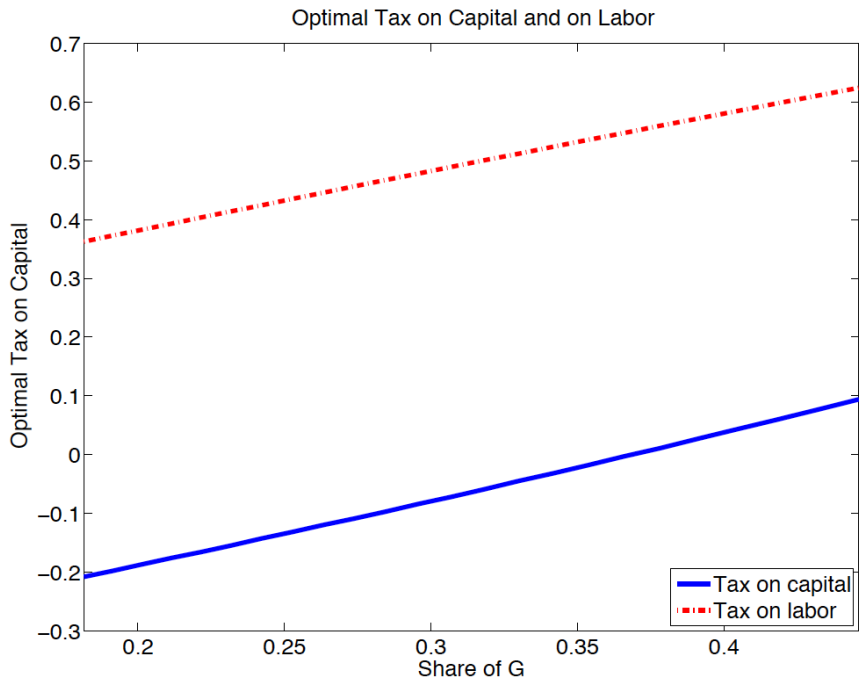

Figure 2: Optimal tax on capital and on labor as a function of the share of G on output.

effects are measured in a few dozen basis points and not in several percentage points. This reinforces our view that the model captures orders of magnitude that are empirically plausible.

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, but our framework is rich enough to accommodate many other exercises. 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. But, first, we need to take care of some unfinished business and offer a more thorough comparison with the neoclassical case.

4.6 The Neoclassical Case

Since the setup of our previous experiment is slightly different than the standard Ramsey allocation problem, we need to check how a neoclassical economy would work under the same assumptions. To illustrate that, we now reduce λ to 1, keeping all the previous assumptions unchanged, and recover the neoclassical case we are interested in: both without market power and without growth.

The results are reported in Figure 4, where we see that the tax rate on capital income is much flatter than in the previous case. Note that, in comparison with the standard Chamley-Judd analysis,

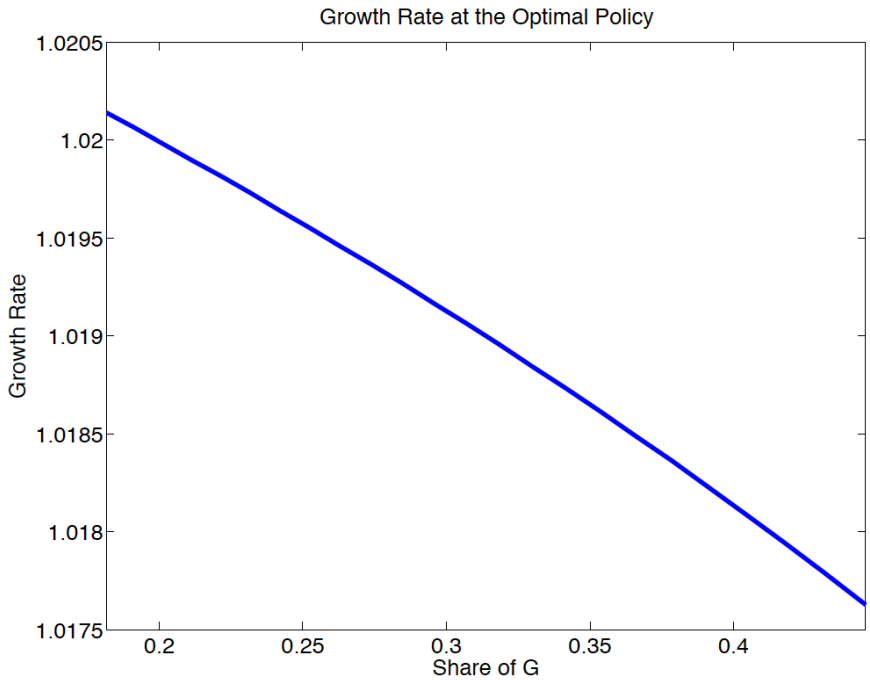

Figure 3: Growth rates in the BGP at the optimal policy

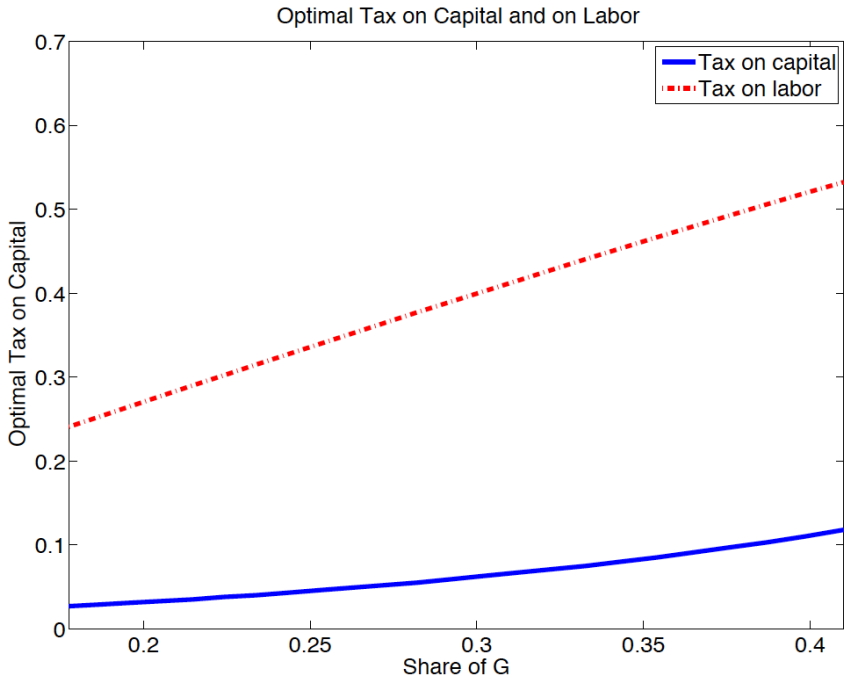

Figure 4: Optimal tax on capital and on labor as a function of the share of G on output, neoclassical.

the 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 do that, it will trade off a bit of distortion in the long run for extra tax revenue from capital in the short run (which allows a reduction in the tax 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 5, we have checked that, for $\beta = 0.999$, the optimal tax on capital is numerically zero.

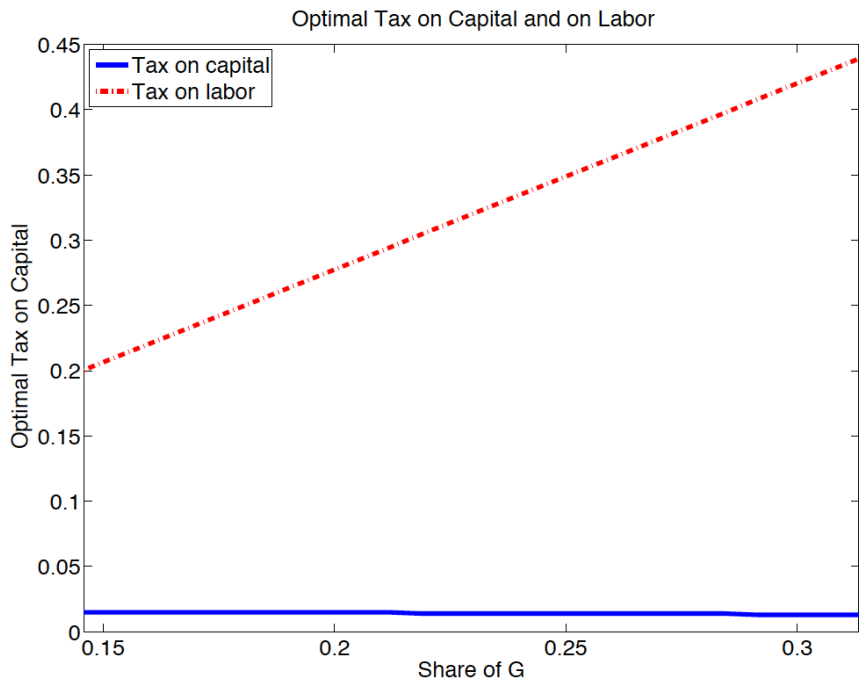

Figure 5: Optimal tax on capital and on labor as a function of the share of G on output, neoclassical, high β.

4.7 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 as in the baseline model (around 1.9 percent at an annual level).

The results appear in Figure 6 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 rather the fact that growth is affected by the market size instead of being exogenous. The results are

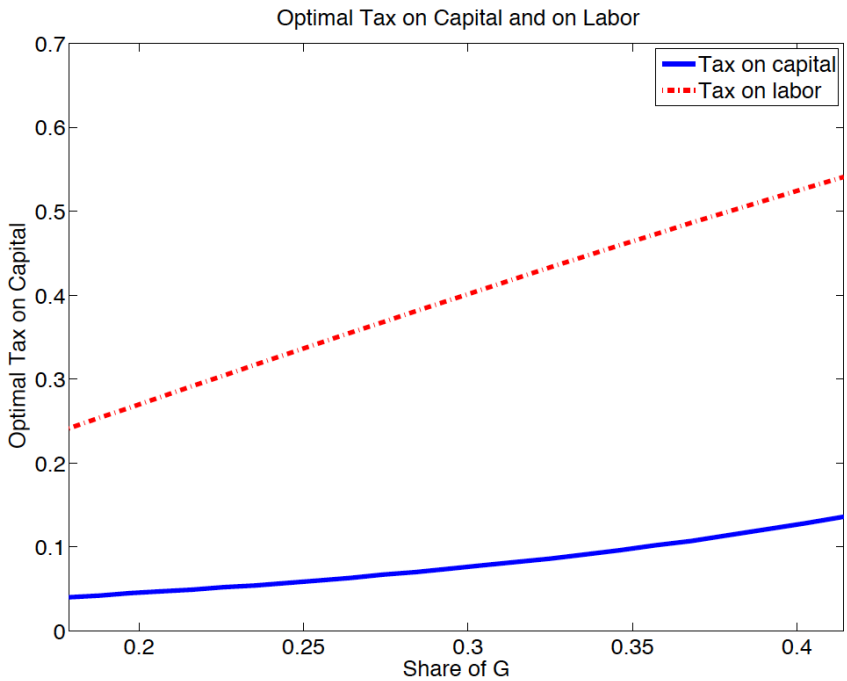

Figure 6: Optimal tax on capital and on labor as a function of the share of G on output, exogenous growth.

not a surprise, either, since we know that in the neoclassical case, the presence of a positive g has nearly the same numerical implications as a slightly lower β and, more importantly, as we reminded the reader in Section 2, that the main insight from Chamley-Judd is independent of whether or not we have positive growth in a neoclassical framework.

4.8 Exogenous Growth Plus Market Power

In this exercise, we compare the welfare-maximizing tax on capital income to the optimal capital tax schedule under exogenous growth but we reintroduce market power. In other words, we compare the optimal tax schedules on capital, respectively, with and without the market size effect (market power is at work in both cases). The results are shown in Figure 7. The continuous lines depict the welfare-maximizing schedules for the tax rate on capital income (as a function of the share of government spending in GDP) when factoring in market size, i.e., with innovation-led growth, for two values of the Frisch elasticity of labor supply, namely, the benchmark value at 1.2 and a higher value at 2.5. The dotted lines depict the corresponding optimal capital tax schedule under exogenous growth plus market power.

A first observation is that the optimal capital tax under exogenous growth plus market power

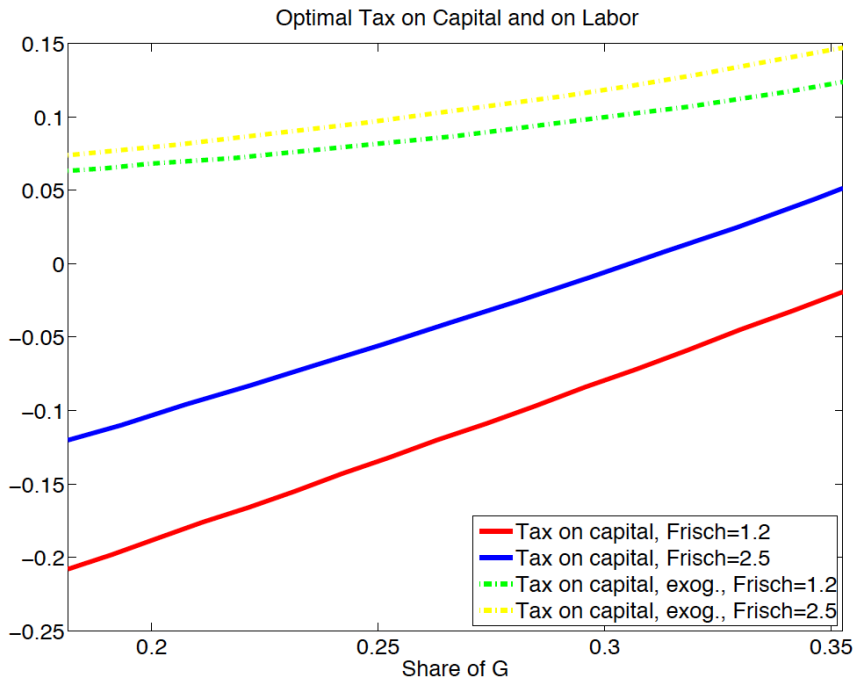

Figure 7: Optimal tax on capital under exogenous and endogenous growth with different Frisch elasticities.

varies very little with respect to both the labor elasticity or the share of government: namely, the two dotted curves corresponding to the two values of the Frisch elasticity of labor supply are very close to each other and these two curves are also almost horizontal. In contrast, the continuous curves are far apart and also strongly upward sloping. This in turn suggests that most of the effect of labor elasticity or of the share of government on the optimal capital tax is due to market size, not market power.

A second observation is that the market size effect, that is, the effect on innovation incentives per se, pushes toward subsidizing capital for these values of the labor market elasticity and of the share of government spending in GDP, although to a lower extent when the Frisch elasticity and/or the share of government spending in GDP are higher.

5 Extensions

5.1 Endogenous Choice of Government Spending

In the standard Chamley-Judd analysis, the sequence of G_t that the government has to finance is exogenously given and it does not yield any utility to the household. A natural alternative is to assume, instead, that government spending enters the utility function of the representative household

and that government spending becomes a choice variable for the government.

A simple utility function that captures that idea is Cobb-Douglas between private consumption, C_t, and public expenditure, G_t:

$$U_i = \frac{1}{1-\nu} \ln C_t^{1-\zeta} G_t^{\zeta} - \phi \frac{L_t^{1+\eta}}{1+\eta}$$

This utility function is convenient because it only introduces one new parameter: ζ. This utility function makes G_t separable from C_t and L_t, which seems a reasonable point of departure in the absence of strong evidence that G_t is a substitute or a complement of C_t and L_t. Finally, it makes the dynamics the same as in the baseline case because all the equilibrium conditions are still the same.

Since there is no generally accepted value for ζ in the literature (not a big surprise, since, as we just argued, it does not affect the equilibrium dynamics of the model and therefore cannot be identified from the macro data), we prefer to report the results for a range of values between 0.4 and 0.75. In our exercise, we compute both the optimal share of G_t over output and the optimal tax on capital income. The left-hand side of Figure 8 shows that, not surprisingly, optimal government spending increases as ζ increases. The right-hand side of Figure 8 then shows that as ζ increases, so does the optimal tax rate on capital income. Intuitively, the higher ζ the higher the optimal level of government spending and therefore the higher the optimal tax rate on capital income, given the negative market size effect of taxing labor income on innovation and growth.[27]

5.2 Allowing for R&D Subsidies

In this section, we consider R&D subsidies to test whether the market size effect can be undone through this additional policy tool. In particular, the government subsidizes private R&D at the rate χ such that the private cost of R&D is simply

$$X_{it} = (1-\chi)\, x_{it} \psi Q_t.$$

We still assume an exogenous government spending $\{G_t\}_{t=0}^{\infty}$. Then the government budget can be written as

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

Figure 9 shows how the optimal capital income tax schedule changes when we allow for positive R&D subsidies. As the figure shows, the market size effect does not disappear with R&D subsidies; if

[27] The step-wise structure of the graphs is due to the grid search we use for optimal G and the optimal tax on capital.

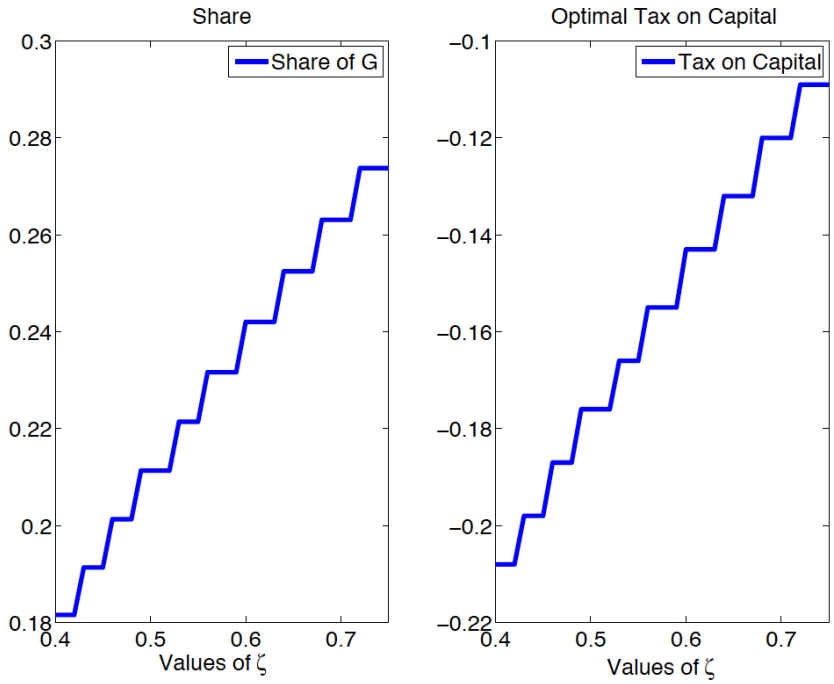

Figure 8: Welfare-maximizing G and capital tax.

anything, its impact becomes stronger. The higher the R&D subsidy rate χ, the less one should tax capital income (or the more one should subsidize it) for any value of G/Y. The intuition is simply that higher direct R&D subsidies partly offset the potentially negative market size effect of taxing labor income instead of capital income.

Note that here we are fixing the level of R&D subsidies. A natural question then is: what happens if we allow the government to fully optimize on the R&D subsidy rate as well as on the capital and labor income tax: are we back to the Chamley-Judd result in that case? The answer is no. The reason is that, with or without R&D subsidies, the market size effect is always at work: namely, not taxing capital income would still impose a higher tax rate on labor income in order to satisfy a budget balance (remember that the government does not have access to lump-sum taxes). This channel is even more important than before, since the labor tax must pay for both the exogenous government expenditure G and the R&D subsidy.

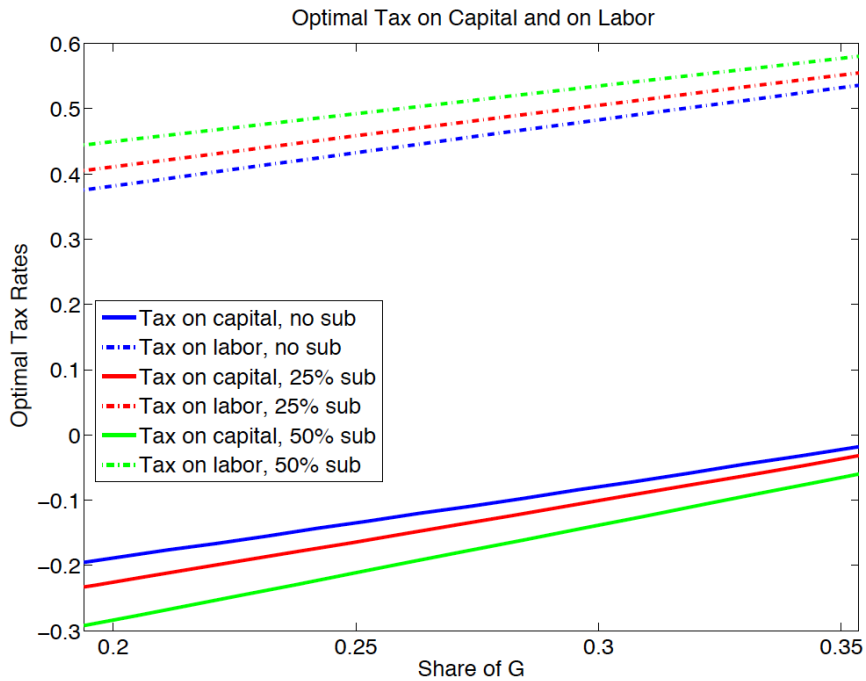

Figure 9: Optimal tax on capital and labor for several levels of subsidies to R&D expenditures.

5.3 Allowing for Public Debt

In this section, we let the government save (or borrow) intertemporally by issuing government bonds B_t. Therefore, the budget constraint of the household is:

$$K_{t+1} + \frac{B_{t+1}}{R_t} + C_t = (1 - \tau_{l,t})\, w_t L_t + (1 - \tau_{k,t})\, r_t K_t + (1 - \delta) K_t + \Pi_t - X_t + B_t$$

where R_t is the gross rate of return on one-period bonds. As in Ljungqvist and Sargent (2004), we assume that bonds are tax exempt (this assumption is just for notational convenience: a tax on the bond yields will be immediately reflected in the pretax rate of return to ensure that the bonds are held by the public).

In every period t, the government's budget constraint becomes:

$$G_t = \tau_{l,t} w_t L_t + \tau_{k,t} r_t K_t + \frac{B_{t+1}}{R_t} - B_t.$$

Moreover, the present value of the government's future resources should exceed its initial debt position

$$B_0 \leq \left(\frac{1-\alpha}{\lambda} \tau_{l,0} + \alpha \tau_{k,0} \right) Y_0 - G_0 + \sum_{t \geq 0} \frac{1}{R_t} [\left(\frac{1-\alpha}{\lambda} \tau_{l,t+1} + \alpha \tau_{k,t+1} \right) Y_{t+1} - G_{t+1}]. \tag{52}$$

Finally, we have two transversality conditions with respect to both assets:

$$\lim_{T\to\infty}\left(\prod_{s=1}^{T-1}\frac{1}{R_s}\right)K_{T+1}=0 \text{ and } \lim_{T\to\infty}\left(\prod_{s=1}^{T-1}\frac{1}{R_s}\right)\frac{B_{T+1}}{R_T}=0. \tag{53}$$

This time, in addition to the optimality conditions with respect to consumption and leisure, the household optimality conditions include a no-arbitrage condition between two assets (capital and bonds):

$$R_t=\left(1-\tau_{k,t+1}\right)r_{t+1}+1-\delta.$$

Then we can write the government's maximization as:

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

subject to

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

$$G_t=K_t^{\alpha}\left(Q_tL_t\right)^{1-\alpha}-\widetilde{r}_tK_t-\widetilde{w}_tL_t+\frac{B_{t+1}}{R_t}-B_t$$

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

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

$$Q_{t+1}=Q_t\left(1+\frac{\left(\lambda-1\right)^2}{\lambda}\frac{\left(1-\alpha\right)}{\psi}\frac{K_t^{\alpha}\left(Q_tL_t\right)^{1-\alpha}}{Q_t}\right)$$

$$R_t=\left(1-\tau_{k,t+1}\right)r_{t+1}+1-\delta$$

$$(52) \text{ and } (53).$$

Then the FOC with respect to K_{t+1} in the steady state is:

$$1=\frac{\beta}{g}\left(\begin{array}{c}\left(\alpha K_0^{\alpha-1}\left(Q_0L_0\right)^{1-\alpha}+1-\delta\right)+\frac{\lambda_{20}}{\lambda_{10}}\left(\alpha K_0^{\alpha-1}\left(Q_0L_0\right)^{1-\alpha}-\widetilde{r}\right)\\+\frac{\lambda_5}{\lambda_{10}}\frac{(\lambda-1)^2}{\lambda}\frac{(1-\alpha)}{\psi}\frac{\alpha K_0^{\alpha-1}(Q_0L_0)^{1-\alpha}}{Q}\end{array}\right) \tag{54}$$

In the steady state, the household's Euler equation remains unchanged

$$1=\beta\frac{1}{g}\left(\widetilde{r}+1-\delta\right) \tag{55}$$

Again, in the steady state, $\tau_k=0$ is not sufficient to equate (54) to (55) due to the *endogenous*

growth term:

$$\frac{\lambda_5}{\lambda_{10}} \frac{(\lambda-1)^2}{\lambda} \frac{(1-\alpha)}{\psi} \frac{\alpha K_t^{\alpha-1} (Q_t L_t)^{1-\alpha}}{Q} \neq 0.$$

This result shows that the government's ability to borrow or lend does not change our long-run conclusion.

6 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 that 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 toward taxing labor income while subsidizing capital income. For high levels of public investment *and* high elasticity of labor supply, the market size effect pushes toward 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 analyze the implication of endogenous innovation and the resulting market size effect in alternative optimal taxation models, in particular in the context of a model with finitely lived individuals and bequests. This would allow us to blend dynastic wealth accumulation considerations with endogenous growth.[28] A third avenue would be to extend the analysis to the case of an open economy. There we conjecture that the market size effect would still operate not only on non-traded goods but also on the traded goods for which the domestic economy is currently a technological leader. A fourth 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 of these extensions await further research.

[28] Also, one should explore the implications of introducing endogenous innovation in New Dynamic Public Finance models.

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] Acemoglu, D., and J. Linn (2004). "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 119, 1049-1090.

[3] Aghion, P., and P. Howitt (1992). "A Model of Growth Through Creative Destruction," *Econometrica*, 60, 323-351.

[4] Aghion, P., Akcigit, U., and P. Howitt (2013). "What Do We Learn from Schumpeterian Growth Theory?" NBER Working Paper 18824.

[5] Aghion, P., and A. Roulet (2011). *Repenser l'Etat,* Editions du Seuil.

[6] Alesina, A., and S. Ardagna (2010). "Large Changes in Fiscal Policy Taxes versus Spending" in J.R. Brown (ed) *Tax Policy and the Economy*, NBER and University of Chicago Press.

[7] Atkeson, A., and A. Burstein (2012). "Aggregate Implications of Innovation Policy," UCLA Working Paper.

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

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

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

[11] Barro, R., and C. J. Redlick (2011). "Macroeconomic Effects of Government Purchases and Taxes," *Quarterly Journal of Economics*, 126: 51-102.

[12] Bianchi, M., Gudmundsson, B. R., and G. Zoega (2001). "Iceland's Natural Experiment in Supply-Side Economics," *American Economic Review*, 91, 1564-79.

[13] Blume-Kohout, M. E., and N., Sood (2013). "Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development," *Journal of Public Economics*, 97, 327-336.

[14] Caldara, D., Fernández-Villaverde, J., Rubio-Ramírez, J. F., and W. Yao (2012). "Computing DSGE Models with Recursive Preferences and Stochastic Volatility," *Review of Economic Dynamics*, 15, 188–206.

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

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

[17] 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.

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

[19] 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.

[20] Duggan, M., and F. S. Morton (2010). "The Effect of Medicare Part D on Pharmaceutical Prices and Utilization," *American Economic Review*, 100, 590-607.

[21] Farhi, E., and I. Werning (2012). "Capital Taxation: Quantitative Exploration of the Inverse Euler Equation," *Journal of Political Economy*, 120, 398-445.

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

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

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

[25] Hobbes, T. (1651). *Leviathan,* reprinted in 1965 from the edition of 1651, Clarendon Press.

[26] Jaimovich, N., and S. Rebelo (2012). "Non-linear Effects of Taxation on Growth," Duke University Working Paper.

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

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

[29] 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.

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

[31] Kocherlakota, N. (2005). "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica,* 73, 1587-1621.

[32] Ljungqvist, L., and T. Sargent (2004). *Recursive Macroeconomic Theory,* MIT Press.

[33] Mertens, K., and M. Ravn (2012). "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," *American Economic Review*, forthcoming.

[34] Milesi-Ferretti, G. M., and N. Roubini (1998). "On the Taxation of Human and Physical Capital in Models of Endogenous Growth," *Journal of Public Economics*, 237-254.

[35] Mill, J. S. (1848). *Principles of Political Economy with some of their Applications to Social Philosophy*, London: West Strand.

[36] National Science Board (2010). *Science and Engineering Indicators 2010*, Arlington, VA: National Science Foundation.

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

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

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

[40] Pistaferri, L. (2003). "Anticipated and Unanticipated Wage Changes, Wage Risk, and Intertemporal Labor Supply," *Journal of Labor Economics*, 21, 729-54.

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

[42] Romer, C., and D. Romer (2010). "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100, 763-801.

[43] Salanié, B. (2003). *The Economics of Taxation*, MIT Press.

[44] Smets, F., and R. Wouters (2007). "Shocks and Frictions in US Business Cycles: 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:

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

subject to

$$K_t^{\alpha} (A_t L_t)^{1-\alpha} + (1-\delta) K_t = C_t + G_t + K_{t+1} \tag{56}$$

$$G_t (= \gamma Y_t) = K_t^{\alpha} (A_t L_t)^{1-\alpha} - \widetilde{r}_t K_t - \widetilde{w}_t L_t \tag{57}$$

$$U_2(C_t, L_t) + \tilde{w}_t U_1(C_t, L_t) = 0 \tag{58}$$

$$U_1(C_t, L_t) = \beta U_1(C_{t+1}, L_{t+1})(\tilde{r}_{t+1} + 1 - \delta) \tag{59}$$

$$A_{t+1} = A_t \Phi\left(1 - \frac{\widetilde{r}_t}{\alpha K_t^{\alpha-1} (A_t L_t)^{1-\alpha}}, 1 - \frac{\widetilde{w}_t}{(1-\alpha) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}}\right). \tag{60}$$

The Lagrangian for this program is expressed as:

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

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

conditions:

$$K_{t+1} \quad : \quad \lambda_{1t} = \beta \left(\begin{array}{c} \lambda_{1t+1} \left(\alpha K_{t+1}^{\alpha-1} \left(A_{t+1} L_{t+1} \right)^{1-\alpha} + 1 - \delta \right) \\ +\lambda_{2t+1} \left(\alpha K_{t+1}^{\alpha-1} \left(A_{t+1} L_{t+1} \right)^{1-\alpha} - \widetilde{r}_{t+1} \right) \\ +\lambda_{5t+1} \left(\begin{array}{c} -\Phi_1 \left(\tau_{k,t+1}, \tau_{l,t+1} \right) \frac{1-\alpha}{\alpha} \frac{\widetilde{r}_{t+1}}{K_{t+1}^{\alpha} (A_{t+1} L_{t+1})^{1-\alpha}} \\ +\Phi_2 \left(\tau_{k,t+1}, \tau_{l,t+1} \right) \frac{\alpha}{1-\alpha} \frac{\widetilde{w}_{t+1}}{K_{t+1}^{\alpha+1} A_{t+1}^{1-\alpha} L_{t+1}^{-\alpha}} \end{array} \right) \end{array} \right) \tag{61}$$

$$C_t \quad : \quad \left\{ \begin{array}{c} U_1 \left(C_t, L_t \right) - \lambda_{1t} + \lambda_{3t} \left(U_{11} \left(C_t, L_t \right) \widetilde{w}_t + U_{21} \left(C_t, L_t \right) \right) \\ +\lambda_{4t} U_{11} \left(C_t, L_t \right) - \lambda_{4t-1} U_{11} \left(C_t, L_t \right) \left(\widetilde{r}_t + 1 - \delta \right) \end{array} \right\} = 0 \tag{62}$$

$$L_t \quad : \quad \left\{ \begin{array}{c} U_2 \left(C_t, L_t \right) + \lambda_{1t} \left(1 - \alpha \right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha} \\ +\lambda_{2t} \left(\left(1 - \alpha \right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha} - \widetilde{w}_t \right) \\ +\lambda_{3t} \left(U_{12} \left(C_t, L_t \right) \widetilde{w}_t + U_{22} \left(C_t, L_t \right) \right) + \lambda_{4t} U_{12} \left(C_t, L_t \right) \\ -\lambda_{4t-1} U_{12} \left(C_t, L_t \right) \left(\widetilde{r}_t + 1 - \delta \right) \\ -\lambda_{5t} \left(\begin{array}{c} -\Phi_1 \left(\tau_{k,t}, \tau_{l,t} \right) \frac{1-\alpha}{\alpha} \frac{\widetilde{r}_t}{K_t^{\alpha-1} A_t^{1-\alpha} L_t^{2-\alpha}} \\ +\Phi_2 \left(\tau_{k,t}, \tau_{l,t} \right) \frac{\alpha}{1-\alpha} \frac{\widetilde{w}_t}{K_t^{\alpha} (A_t L_t)^{1-\alpha}} \end{array} \right) \end{array} \right\} = 0 \tag{63}$$

$$A_{t+1} \quad : \quad \left\{ \begin{array}{c} \left(\lambda_{1t+1} + \lambda_{2t+1} \right) K_{t+1}^{\alpha} \left(A_{t+1} L_{t+1} \right)^{1-\alpha} - \frac{\lambda_{5t}}{\beta} \frac{1}{A_t} + \lambda_{5t+1} \left(\frac{A_{t+2}}{A_{t+1}^2} \right) \\ +\lambda_{5t+1} \Phi_1 \left(\tau_{k,t+1}, \tau_{l,t+1} \right) \left(1 - \alpha \right) \frac{\widetilde{r}_{t+1} A_{t+1}^{\alpha-2}}{\alpha K_{t+1}^{\alpha-1} L_{t+1}^{1-\alpha}} \\ +\lambda_{5t+1} \Phi_2 \left(\tau_{k,t+1}, \tau_{l,t+1} \right) \left(1 - \alpha \right) \frac{\widetilde{w}_{t+1} A_{t+1}^{\alpha-2}}{(1-\alpha) K_{t+1}^{\alpha} L_{t+1}^{-\alpha}} \end{array} \right\} = 0 \tag{64}$$

$$\widetilde{w}_t \quad : \quad -\lambda_{2t} L_t + \lambda_{3t} U_1 \left(C_t, L_t \right) - \lambda_{5t} \Phi_2 \left(\tau_{k,t}, \tau_{l,t} \right) \frac{1}{\left(1 - \alpha \right) K_t^{\alpha} A_t^{1-\alpha} L_t^{-\alpha}} = 0 \tag{65}$$

$$\widetilde{r}_t \quad : \quad \lambda_{2t} K_t + \lambda_{4t-1} U_1 \left(C_t, L_t \right) - \lambda_{5t} \Phi_1 \left(\tau_{k,t}, \tau_{l,t} \right) \frac{1}{\alpha K_t^{\alpha-1} \left(A_t L_t \right)^{1-\alpha}} = 0. \tag{66}$$

When productivity growth is exogenous: $\Phi \left(\tau_{k,t}, \tau_{l,t} \right) \equiv g$, (61) 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, λ_{2t} and λ_{1t} decline at rate g with $\lambda_{1t} = \lambda_{10} g^{-t}$ and $\lambda_{2t} = \lambda_{20} g^{-t}$. Thus, this latter equation simplifies 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 $\Phi \left(\tau_{k,t}, \tau_{l,t} \right)$,

equation (61) 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 re-express the above system as:

$$\tilde{k}_{t+1}\frac{L_{t+1}Q_{t+1}}{L_tQ_t}+\tilde{c}_t+\frac{G_0}{L_t}=\Omega\tilde{y}_t+\left(1-\delta\right)\tilde{k}_t \tag{67}$$

$$\frac{\tilde{c}_{t+1}}{\tilde{c}_t}\frac{L_{t+1}Q_{t+1}}{L_tQ_t}=\beta\left(1-\tau_{k,t+1}\right)\alpha\frac{\tilde{y}_{t+1}}{\tilde{k}_{t+1}} \tag{68}$$

$$\phi\tilde{c}_tL_t^{1+\eta}=\left(1-\tau_{l,t}\right)\frac{1-\alpha}{\lambda}\tilde{y}_t \tag{69}$$

$$\tilde{y}_t=\tilde{k}_t^{\alpha} \tag{70}$$

$$G_0=\left(\frac{1-\alpha}{\lambda}\tau_{l,t}+\alpha\tau_{k,t}\right)L_t\tilde{y}_t \tag{71}$$

$$\frac{Q_{t+1}}{Q_t}=1+\frac{\left(\lambda-1\right)^2}{\lambda}\frac{1-\alpha}{\psi}\tilde{y}_tL_t \tag{72}$$

Solve (71) for G_0/L_t and substitute into (67). Similarly, use conjecture $\tilde{c}_t=\left(1-s\right)\tilde{y}_t$ in (68) and

solve for $\tilde{k}_{t+1}\frac{Q_{t+1}}{Q_t}$ and substitute into (67). 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 (69) 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 (68) 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 second-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

A standard measurement of the accuracy of a solution to a dynamic equilibrium model is the Euler equation error function (Judd, 1992). In figure A1, we plot the Euler equation errors of the model solved with our third-order perturbation for the benchmark calibration as a function of the log of (normalized) capital and for $G_0=0.1$. 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 -7 (one can interpret this number as equivalent to making a mistake of \$1 for each \$10 million spent). Hence, the performance of our solution method in terms of accuracy is most satisfactory. Even if we move 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.

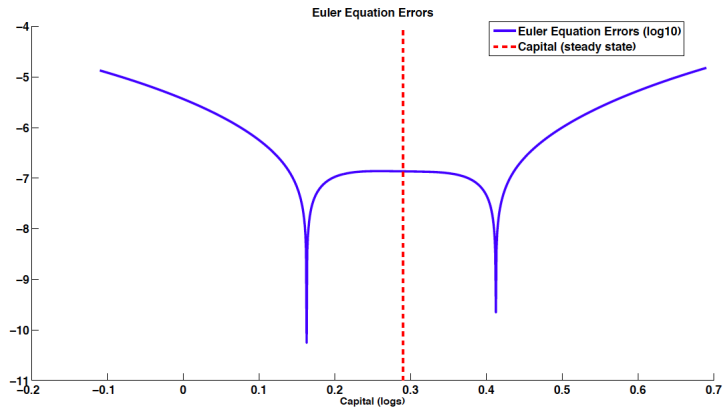

Figure A1: Euler equation errors, benchmark calibration.

In Figure A2, we plot the Euler equation errors as a function of both capital and different values of G_0. Again, we see a very good performance of our third-order perturbation in all relevant ranges of the variables.

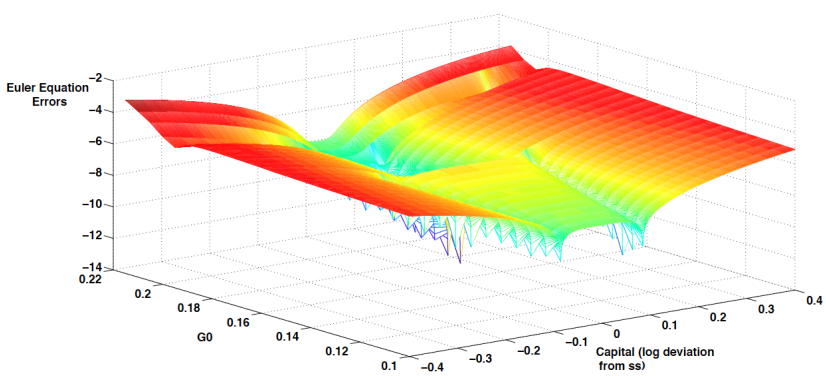

Figure A2: Euler equation errors, optimal policies.

References

[1] Judd, K.L. (1992). "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory* 58, 410-452.