Department Discussion Paper DDP1803

ISSN 1914-2838

OPTIMAL CAPITAL TAXATION WITH INCOMPLETE MARKETS AND SCHUMPETERIAN GROWTH

Marco Cozzi
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

April, 2018

Abstract

This paper characterizes quantitatively the optimal capital income tax rate in an OLG economy with uninsurable income risk, incomplete markets and endogenous Schumpeterian growth. Contrary to the most recent literature, it is found that it is virtually never optimal to tax capital: under the optimal scheme, in a series of cases, the highest proportional tax rate on capital is found to be less than 0.2%. The reason for this result lies in the reduced GDP (and wage) growth rate stemming from a higher capital tax rate. In General Equilibrium, the interest rate rises, and the increased cost of capital reduces the endogenous rate of innovation, leading to a negative response of the growth rate. Although the equilibrium effect on the growth rate is found to be quantitatively modest (approximately half a percentage point), it still has a first order consequence on welfare. The results show that moving to the optimal income tax schedule entails large welfare gains, approximately 5% in consumption equivalent. The results are robust along a number of dimensions, including the specification of preferences. An alternative formulation of the utility function, taken from a class consistent with a Balanced Growth Path, is calibrated to obtain an empirically plausible value for the Frisch elasticity of 0.5, and confirms all the results, both qualitatively and quantitatively.

Keywords: Capital and Income Taxation, Heterogeneous Agents, Incomplete Markets, Endogenous Growth, Welfare.

JEL Classifications: D15, E21, H21, O41.

Author Contact:
Marco Cozzi, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; E-mail: mcozzi@uvic.ca; Voice: (250) 721-6535; FAX: (250) 721-6214

Optimal Capital Taxation with Incomplete Markets and Schumpeterian Growth

Marco Cozzi, University of Victoria

April 2018

Abstract

This paper characterizes quantitatively the optimal capital income tax rate in an OLG economy with uninsurable income risk, incomplete markets and endogenous Schumpeterian growth. Contrary to the most recent literature, it is found that it is virtually never optimal to tax capital: under the optimal scheme, in a series of cases, the highest proportional tax rate on capital is found to be less than 0.2%. The reason for this result lies in the reduced GDP (and wage) growth rate stemming from a higher capital tax rate. In General Equilibrium, the interest rate rises, and the increased cost of capital reduces the endogenous rate of innovation, leading to a negative response of the growth rate. Although the equilibrium effect on the growth rate is found to be quantitatively modest (approximately half a percentage point), it still has a first order consequence on welfare. The results show that moving to the optimal income tax schedule entails large welfare gains, approximately 5% in consumption equivalent. The results are robust along a number of dimensions, including the specification of preferences. An alternative formulation of the utility function, taken from a class consistent with a Balanced Growth Path, is calibrated to obtain an empirically plausible value for the Frisch elasticity of 0.5, and confirms all the results, both qualitatively and quantitatively.

JEL Classification Codes: D15, E21, H21, O41.

Keywords: Capital and Income Taxation, Heterogeneous Agents, Incomplete Markets, Endogenous Growth, Welfare.

Acknowledgments: *I am grateful to Kevin Andrew, Elisabeth Gugl, Peter Kennedy, Vincenzo Quadrini and seminar participants at the University of Victoria for useful comments.*

Contact details: *Department of Economics, University of Victoria, 3800 Finnerty Road, Victoria, BC V8P 5C2, Canada. Tel: 1-250-721-6535, Fax: 1-250-721-6214, E-mail:* `mcozzi@uvic.ca`

1 Introduction

The marked rise in income and wealth inequality observed in several economies during the last four decades has prompted a number of authors, most notably Piketty (2014), to propose taxing capital more heavily. This stand on fiscal policy is also meant to partially undo the steady decrease in capital taxation that has been observed in a number of economies, such as the U.S. and the U.K., Aoki and Nirei (2017). However, there is no consensus on whether capital income should be taxed. The classical view, presented in the seminal contributions by Judd (1985) and Chamley (1986), argued that this form of taxation should be avoided. A positive tax rate on capital income implies a compounded distortionary effect, reducing the accumulation of a productive input, leading to suboptimal allocations and welfare losses. More recently, Conesa, Kitao and Krueger (2009) challenged this perspective. In their influential paper, featuring an incomplete markets model calibrated to match some salient features of the U.S. labor income inequality, they found that in steady state comparisons welfare is highest when the capital income tax is positive, irrespective of the lower capital stock. Perhaps surprisingly, their results showed that the optimal proportional capital tax rate is large, being in the $20\% - 35\%$ range.

Peterman (2013) undertook a careful decomposition of the reasons behind a positive capital tax rate in a framework similar to the model developed in Conesa, Kitao and Krueger (2009), finding that one margin is especially important: the Frisch elasticity of labor supply. Although his results show that an alternative specification with a Frisch elasticity that is constant over the life-cycle can lower the optimal capital tax rate, one of the main findings is that the capital tax remains positive and remarkably far from zero.

The previous literature has mostly focused on the neoclassical growth model. In this paper I argue that a key channel for the positive rate on capital is that the economic growth rate is unaffected by the chosen tax schedule. Differently, I address the question of whether capital should be taxed in the context of an endogenous growth model, with rich workers' heterogeneity and technological progress that responds to different taxation schemes. I extend the Schumpeterian growth model proposed by Howitt and Aghion (1998), embedding the elements that might call for a positive tax on capital, namely a tight borrowing constraint, uninsurable income risk and taxes that are non age-dependent. A key element of the framework is that Research and Development (R&D) is capital intensive and any equilibrium effects on its cost are going to bring about endogenous responses in the rate of technological change.[1] At the core of the framework lies a trade-off between a suboptimal equilibrium distribution of economic resources (both cross-sectionally and over the life-cycle) and a lower growth rate. Quantitatively, I find that the foregone economic growth more than offsets the improved consumption smoothing (or the lack thereof, due to the absence of complete markets) that a heavier capital taxation brings about. Even with a modest capital income tax rate, the induced increase in the cost of capital reduces the rate of innovation, leading to a negative response of the growth rate. Although the equilibrium effect on the growth rate is found to be quantitatively modest, it still has a first order consequence on welfare. Contrary to the most recent literature, I find that it is virtually never optimal to tax capital: under the optimal scheme, in a series of cases, the highest proportional tax rate on capital is found to be less than 0.2%. Moreover, the results

[1] The representative agent version of this model has been used by Nuno (2011) to study the relationship between the optimal R&D and the cost of business cycles and estimated by Cozzi, Pataracchia, Pfeiffer and Ratto (2017), showing that financial crises are amplified by the endogenous innovation dynamics, partially explaining slow recoveries.

show that moving to the optimal income tax schedule with an almost zero capital income tax rate entails large welfare gains, approximately 5% in consumption equivalent (CEV).

My findings are in line with the more traditional view that capital taxation should be avoided, as argued by Atkeson, Chari and Kehoe (1999) and by the New Public Finance literature, such as Golosov, Kocherlakota, and Tsyvinski (2003). More recently, within the Schumpeterian framework, Aghion, Akcigit and Fernandez-Villaverde (2013) used an innovation-driven endogenous growth model with a representative infinitely-lived household, finding that it can be optimal to subsidize capital, depending on the size of the public expenditure and the elasticity of labor supply. My results show that even in the presence of incomplete markets, rich life-cycle dynamics, and realistic labor income risk, when capital is an input to R&D, the taxation of capital income raises the cost of producing quality enhancing intermediate goods, which slows down both the economy's growth rate and the increase of consumption over the life-cycle. These in turn lead to a negative effect on welfare.

The results are robust along a number of dimensions, including the specification of preferences. An alternative formulation of the utility function, taken from a class consistent with a Balanced Growth Path, is calibrated to obtain an empirically plausible value for the Frisch elasticity of 0.5, and confirms all the results, both qualitatively and quantitatively.

The rest of the paper is organized as follows. Section 2 presents the OLG model with endogenous growth. Section 3 briefly presents the model calibration. Section 4 discusses the main results. Section 5 concludes. Three appendices are also included: they discuss in more detail the model and the numerical methods used to solve it.

2 An OLG Model with Incomplete Markets and Schumpeterian Growth

The economy is a production economy with an endogenous asset distribution, where a government collects taxes to finance an exogenously given stream of public expenditures, denoted by G.

I work with an Overlapping Generations (OLG) structure. Agents are ex-ante identical, while they differ ex-post, due to idiosyncratic realizations of a series of shocks.

The model is an extension of the Huggett (1996) economy, appropriately modified to allow for endogenous growth, for several sources of heterogeneity in labor income and a budget-neutral reform of the income tax schedule, which introduces the possibility of taxing capital income separately from labor earnings.

Time is discrete. The economy is populated by finitely lived agents facing an age-dependent death probability π_j^d. Age is denoted with j and there are J overlapping generations, each consisting of a continuum of agents. At age J_R all agents that are still alive become retirees. The population grows at rate g_n. Beside the workers, there is a measure one of risk-neutral and infinitely lived entrepreneurs, whose consumption is denoted by C_e.

Preferences: Agents' preferences are assumed to be time-separable and represented by the utility function $U(.)$. Agents' utility is defined over stochastic consumption $\{c_j\}_{j=1}^J$ and leisure sequences $\{l_j\}_{j=1}^J$: their aim is to choose how much to consume (c_j), how much to work ($h_j = 1 - l_j$), and how much to save in an interest bearing asset (a_{j+1}) in each period of their lives, in order to maximize their objective function. The agents' problem can be defined as:

$$\max_{\{c_j,l_j,a_{j+1}\}_{j=1}^J} E_0 U(c_0, c_1, ...; l_0, l_1, ...) = \max_{\{c_j,l_j,a_{j+1}\}_{j=1}^J} E_0 \sum_{j=1}^{J} \beta^{j-1} \left[\prod_{s=1}^{j} \left(1 - \pi_s^d\right) \right] u(c_j, l_j)$$

where E_0 represents the expectation operator over the idiosyncratic sequences of shocks, and $\beta > 0$ is the subjective discount factor. In the benchmark formulation, I assume that $u(c_j, l_j) = \frac{\left(c_j^{\eta} l_j^{1-\eta}\right)^{1-\sigma} - 1}{1-\sigma}$, that is the per-period utility function is strictly increasing in both consumption and leisure, strictly concave, satisfies the Inada conditions, and has a relative risk aversion RRA$= \sigma\eta + 1 - \eta$.[2] The parameter η in the Cobb-Douglas aggregator stands for the consumption share.

Endowments: Agents differ in their labor endowments $\epsilon_{j,\varepsilon,f}$. There are four channels that contribute to the determination of the total efficiency units that the workers supply in the labor market. First, there is an active choice on the number of hours that the agents want to work, expressed as a share of their time endowment (normalized to 1). Second, there is a deterministic age component e_j, which is the same for all agents. Third, there is a stochastic component ε, whose log follows a stationary $AR(1)$ process: $\log \varepsilon_j = \rho_y \log \varepsilon_{j-1} + \xi_j$, with $\xi_j \sim N(0, \sigma_y^2)$. Fourth, there is a fixed effect component f, with half of the agents being born with the highest realization, and the other half with the lowest.[3] The total efficiency units a worker is endowed with are the product of the last three components, multiplied by the hours worked and the wage (w). It follows that labor earnings are $y_j^w = w h_j \epsilon_{j,\varepsilon,f} = w \times h_j \times e_j \times \varepsilon \times f$. After the common retirement age J_R, the labor endowment drops to zero, and the agents receive a pension $\overline{y}_R$ paid for with the contributions of the economically active agents. The pension is a fixed replacement rate ϕ_R of the average labor earnings, and agents pay proportional taxes (τ_R) to contribute to the balanced-budget pension scheme. They also finance the public expenditure G with their income and consumption taxes, and receive a wealth transfer TR. Agents cannot insure against their mortality risk. As a consequence, on average agents do not die with zero wealth and there are accidental bequests TR that are uniformly re-distributed across all agents in a lump-sum fashion. Apart from the accidental bequests, newborns enter the economy with a zero asset endowment and with the average realization of the stochastic component of labor earnings, which is normalized to 1.

Schumpeterian growth: The model embeds Schumpeterian growth into a model of incomplete markets. Endogenous growth is based on vertical innovations as in Howitt and Aghion (1998). Producers of final goods use labor and a continuum of intermediate goods M as inputs. The final goods sector is modeled as a constant returns to scale technology of the Cobb-Douglas form, which relies on aggregate labor L and the sum (i.e., the integral) of all the intermediate goods M_i to produce the final output Y. The intermediate goods differ in their

[2] In Cozzi (2014), a related paper focusing on whether the welfare effects stemming from a policy change can be estimated reliably, I found that heavy capital income taxation can be sub-optimal, even without an endogenous growth mechanism. However, in that paper eliminating the capital income tax was found to be welfare improving because of the exogenous labor supply assumption.

[3] These two values are chosen to match the variance of the fixed effect σ_f^2.

productivity $A_{i,t}$ and each of them is produced by a monopolistic competitive firm using capital as the only input. The amount of capital necessary to produce each intermediate good is proportional to its productivity, and more advanced products require increasingly capital-intensive techniques. In each period, there is a probability that the productivity of an intermediate good jumps to the technology frontier owing to the R&D activities of entrepreneurs. Entrepreneurs borrow resources and invest them in R&D trying to increase their probabilities of making a discovery that is going to displace the current monopolist for a specific intermediate good. If a discovery occurs, the successful entrepreneur introduces a new enhanced intermediate product in the relevant sector and becomes the new monopolist until replaced (stochastically) by another entrepreneur. The technology frontier (i.e., the productivity level of the most advanced sector) evolves endogenously as the result of positive spillovers from innovation activities.

Technology: The homogeneous final good is produced under perfect competition using labor and a continuum of intermediate products. Final output can be used interchangeably as a consumption good (C_t), a capital good (K_t), or an input to innovation (RD_t). The representative firm producing in the final good sector maximizes its profits having access to a Cobb-Douglas production function. The inputs are aggregate labor (L_t), the quantity of intermediate products of variety i ($M_{i,t}$), and the associated productivity index ($A_{i,t}$).

$$Y_t = F\left(L_t, \{A_{i,t}\}_i, \{M_{i,t}\}_i\right) = L_t^{1-\alpha} \int_0^1 A_{i,t} M_{i,t}^{\alpha} di.$$

It follows that the profits π^Y in the final sector are given by:

$$\pi^Y = L_t^{1-\alpha} \int_0^1 A_{i,t} M_{i,t}^{\alpha} di - \int_0^1 P_{i,t} M_{i,t} di - w_t L_t.$$

$M_{i,t}$ denotes the amount of intermediate product i, $P_{i,t}$ their monopolistic prices, while $A_{i,t}$ is a productivity parameter embodied in the latest version of intermediate product i. w_t is the wage rate and the price of the final output is normalized to 1. Since the final good firm is a price taker, the first order conditions for profit maximization lead to a system of demand equations $P_{i,t}(M_{i,t})$, one for each intermediate good variety, and are given by:

$$P_{i,t}(M_{i,t}) = \alpha A_{i,t} L_t^{1-\alpha} M_{i,t}^{\alpha-1}, \forall i. \tag{1}$$

Another first order condition delivers the labor demand schedule:

$$w_t = (1-\alpha) Y_t / L_t.$$

Intermediate goods: Each intermediate product i is produced by an incumbent monopolist using a capital-intensive production function (where $K_{i,t}$ is the capital used in sector i at time t):

$$M_{i,t} = K_{i,t} / A_{i,t}. \tag{2}$$

Capital is divided by the technological index $A_{i,t}$ to capture the phenomenon that successive vintages of the intermediate product are produced by increasingly capital-intensive techniques. The model at the aggregate level is deterministic, and the incumbent monopolist of each sector can correctly forecast the demand for the intermediate good they are producing. Taking the demand function for their product as given, but understanding the consequences of setting different prices, each incumbent monopolist has the following (after-tax) profit function (where τ_f is the proportional tax rate on corporate profits):

$$(1-\tau_f)\pi^f_{i,t} = (1-\tau_f)[P_{i,t}(M_{i,t})M_{i,t} - (r_t+\delta)K_{i,t}]. \tag{3}$$

Capital depreciates at the exogenous rate δ and r_t is the real rate of return to capital, so their sum $(r_t+\delta)$ is the user cost of capital. Substituting Eq. (1) and (2) into (3), the (after-tax) profit function can be rewritten as:

$$(1-\tau_f)\pi^f_{i,t} = (1-\tau_f)\left[\alpha A_{i,t}L_t^{1-\alpha}M^{\alpha}_{i,t} - (r_t+\delta)A_{i,t}M_{i,t}\right] = (1-\tau_f)A_{i,t}\left[\alpha L_t^{1-\alpha}M^{\alpha}_{i,t} - (r_t+\delta)M_{i,t}\right].$$

Since the demand functions are symmetric, and the user cost of capital is the same for every intermediate producer, each intermediate product is produced in the same amount M_t. Aggregate quantities are the integral with respect to all intermediate products, namely the average productivity across all sectors is $A_t = \int_0^1 A_{i,t}di$, and aggregate capital is obtained as the product of average productivity times the average intermediate good:

$$K_t = \int_0^1 K_{i,t}di = A_tM_t.$$

One of the convenient features of this model is that in equilibrium the final goods sector displays the familiar Cobb-Douglas aggregate production function in capital, labor and technological progress:

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

However, because of the monopolistic price distortions and the presence of profits in the intermediate good sector, the equilibrium interest rate has a different formula:

$$r = \alpha^2\frac{Y_t}{K_t} - \delta. \tag{4}$$

Notice how in Eq. (4) the term α^2 replaces α in the corresponding version for the marginal product obtained in the neoclassical growth model with a Cobb-Douglas aggregate production function.

In terms of the functional distribution of income, the share $(1-\alpha)$ of GDP belongs to labor earnings, the share α^2 to capital earnings (including depreciation), and the share $\alpha(1-\alpha)$ to profits. The flow of after-tax profits that each incumbent in the intermediate goods sector earns in every period is then:

$$(1-\tau_f)\pi^f_{i,t} = (1-\tau_f)\alpha(1-\alpha)\frac{A_{i,t}Y_t}{A_t}. \tag{5}$$

Innovation and Technological Change: Innovations result from entrepreneurship activities that advance technological knowledge. At any date there is a technology frontier that represents the most advanced technology across all sectors:

$$A_t^{max} = \underset{i}{Max}A_{i,t}. \tag{6}$$

In each period there is an endogenous probability $p_{i,t}$ that the productivity $A_{i,t}$ of an intermediate good in sector i jumps to the technology frontier. In order to achieve this goal, entrepreneurs in this sector have to undertake costly R&D activities ($RD_{i,t}$). Entrepreneurs invest resources $RD_{i,t}$ in an attempt to increase their probabilities of making a discovery and replace the current incumbents. If a discovery occurs, the successful entrepreneur introduces a new enhanced intermediate product in that sector and becomes the new monopolist, until another entrepreneur will create an even better version of that intermediate product. At the level of a single sector the evolution of the productivity index is stochastic. With probability $p_{i,t}$ the R&D activity is successful and the productivity jumps to $A_{i,t+1} = A_t^{max}$. With probability $1 - p_{i,t}$ the R&D activity is unsuccessful, the incumbent is not replaced by a new monopolist, and the productivity stays at its current value $A_{i,t+1} = A_{i,t}$.[4] Entrepreneurs incur R&D costs and the probability of a successful innovation is assumed to be:

$$p_{i,t} = 1 - e^{-\frac{RD_{i,t}}{\lambda A_t^{max}}} \tag{7}$$

where λ represents an efficiency parameter of R&D. A higher value of λ results in a less effective expenditure in R&D, because a given investment $RD_{i,t}$ corresponds to a lower probability of a successful innovation. The amount spent on R&D is also adjusted by the technology frontier variable A_t^{max}. This is meant to capture in a parsimonious way the increasing complexity of progress: as technology advances, the resource cost of further improvements increases proportionally. Notice also how this formulation guarantees that the probability of an innovation is always well behaved (i.e., $0 \leq p_{i,t} < 1, \forall (i,t)$), and that R&D investments are essential for innovations to occur (i.e., $p_{i,t} = 0$ when $RD_{i,t} = 0$). Furthermore, it is parsimonious, as it entails only one parameter (λ) that needs to be assigned a value in the quantitative implementation of the model.[5]

Entrepreneurs: The presence of monopoly power gives the incumbent in each intermediate good variety the prospect of earning some profits over the duration of their technological advantage. The value of being the incumbent in period t in sector i with current productivity level $\overline{A}$ is denoted with $V_{i,t}(\overline{A})$, and satisfies the following Bellman equation:

$$V_{i,t}(\overline{A}) = \underset{M_{i,t}}{Max} \left\{ (1 - \tau_f)\pi_{i,t}^f + \left(\frac{1 - p_{i,t}}{1 + r_t} \right) V_{i,t+1}(\overline{A}) \right\} \tag{8}$$

$V_{i,t}(\overline{A})$ is the expected discounted flow of net profits that the incumbent is expected to obtain, given that the intermediate good variety it managed to develop has an associated (and fixed) productivity equal to $\overline{A}$. It takes into consideration the process of creative destruction, as it internalizes the fact that the monopolistic position might be lost to a competitor with probability $p_{i,t}$ (hence the complement probability $1 - p_{i,t}$ stands for the incumbent's chances of survival). Following the literature, I focus on the case where there is a single entrepreneur in each period and sector that tries to replace the incumbent by developing an improved version of the related intermediate good. Since entrepreneurs are assumed to be risk neutral, and can borrow the

[4] After a successful innovation, the entrepreneur enters into Bertrand competition with the previous incumbent in that sector. Since the old intermediate good is of inferior quality, the incumbent exits. The appendix in Howitt and Aghion (1998) provides more details on how the strategic interaction unfolds.

[5] Notice also that the value $A_t^{max} = 0$ can never be obtained in a growing economy, as long as $A_0^{max} > 0$.

resources needed to finance the R&D from a competitive banking sector at the interest rate r_t, they maximize the expected discounted value of becoming an incumbent in their sector of operation i in the next period:

$$\underset{RD_{i,t}}{Max}\left\{-RD_{i,t}+\left(\frac{p_{i,t}\left(RD_{i,t}\right)}{1+r_t}\right)E_tV_{i,t+1}(A_t^{max})\right\} \tag{9}$$

Notice how in Eq. (9) the optimal value function $V(.)$ representing the expected value of becoming an incumbent is evaluated at $A_t^{max} > \overline{A}$, as the entrepreneur's innovation brings the associated productivity to the technology frontier. Let's define $\rho_t \equiv \frac{RD_t}{\lambda A_t^{max}}$. The first order conditions lead to a non-linear equation in ρ_t:

$$-\rho_t = log\left[\frac{\lambda A_t^{max}}{E_tV_{t+1}(A_t^{max})}\right] = log\left[\frac{RD_t}{\rho_t E_tV_{t+1}(A_t^{max})}\right] \tag{10}$$

[Figure 1 about here]

Figure (1) plots the optimal entrepreneurs' choices regarding the (normalized) R&D investment ρ. The plot shows that the non-linear equation has only one solution, as the relationship is monotonically increasing in ρ. The left panel displays the effect of changing the parameter λ. As λ increases, R&D expenditures become less effective, and this leads to a reduction in their equilibrium value. Also, this plot shows that this parameter is uniquely identified: data on R&D expenditures (or a monotonic transformation, such as the firms' exit rate) allow to pin down its value. The right panel shows the response of ρ to changes in the interest rate, for a fixed value of λ. The lower r, the higher the discounted value of future profits as a monopolist, and the lower the costs of borrowing to finance innovation. These two effects combined expand the optimal investment in R&D when the equilibrium interest rate falls: this channel will be at play when considering a fiscal policy reform that decreases the capital income tax rate. To conclude with, notice how the optimal ρ^* is always strictly positive. This will have implications for the equilibrium growth rate.

The advancement of the technology frontier A_t^{max} is the mechanism that drives the aggregate economic growth. Innovations induce knowledge spillovers, because at any point in time the technology frontier is available to any successful innovator. This publicly available knowledge grows at a rate proportional to the aggregate rate of innovations, and each innovation moves the technology frontier by a factor $1+\gamma > 1$. At any point in time, and across different varieties of intermediates, some entrepreneurs are going to be successful at innovating, while others are going to fail. Taking this into account, together with the assumption that the likelihood of success of an innovation is independent across intermediate varieties, a law of large numbers guarantees that at the aggregate level the average productivity will evolve according to the following equation:

$$A_{t+1} = \int_0^1 \left\{p\left(\rho_t\right)\left(1+\gamma\right)A_{i,t} + \left[1-p\left(\rho_t\right)\right]A_{i,t}\right\}di = \left(1+\gamma\right)p\left(\rho_t\right)A_t + \left[1-p\left(\rho_t\right)\right]A_t = \left[1+p\left(\rho_t\right)\gamma\right]A_t \tag{11}$$

The first term in the integral represents the probability that an innovation is going to occur for a given variety, multiplied by the implied increase in the related productivity index $(1+\gamma)A_{i,t}$. The second term in the integral represents the probability that an innovation is not going to occur, hence the productivity of all these intermediates is going to stay at their current level. Thanks to the symmetry in the R&D investments across all varieties, also the probabilities of a successful innovation are the same across all varieties.

Eq. (11) can be manipulated to represent how the aggregate productivity is going to grow over time. Along a BGP $\rho_t = \rho, \forall t$, as both the R&D and the technology frontier will grow at the same rate. The (average) growth rate g is then determined as:

$$g = \frac{A_{t+1}}{A_t} - 1 = p(\rho)\gamma \tag{12}$$

This equation shows how the growth rate is equal to the spillover effect weighted by the probability of an innovation. In this framework, the spillover effect γ is an exogenous parameter. Without a strictly positive knowledge spillover the economy cannot grow over time. The second element in the formula is $p(\rho)$, which is an endogenous outcome, affected also by public policy: different taxation schemes are going to have an impact on the scaled expenditure on R&D, which in turn is going to affect the likelihood of an innovation, and ultimately the growth rate g.

Taxes and Public Expenditure: The government carries out some public expenditure G. In order to finance the cost of these purchases, consumption is taxed at rate τ_c and firms' profits at rate τ_f. Moreover, households' total income is taxed according to the functional form proposed by Gouveia and Strauss (1994) and used by Conesa, Kitao and Krueger (2009). Given taxable income y, total taxes T are given by

$$T = \kappa_0 \left[y - \left(y^{-\kappa_1} + \kappa_2 \right)^{-1/\kappa_1} \right] \tag{13}$$

The κ_0 and κ_1 parameters are set exogenously, relying on the estimates obtained by Gouveia and Strauss (1994), while κ_2 is set residually to guarantee a balanced budget.

Other market arrangements: There are no state-contingent markets to insure against the labor income risk, but workers can self-insure by saving into the risk-free asset. The households also face an exogenous borrowing limit, denoted as b, which is set to $b = 0$.[6]

3 Calibration

The economy has several parameters that need to be assigned a value. I rely on a mix of (reduced-form) estimation and calibration (in equilibrium) methods. The initial BGP is calibrated to mimic selected long-run features of the U.S. economy. Table 1 reports the full list of the calibrated parameters with their values and empirical targets.

[Table 1 about here]

[6] This is a fairly common assumption in the literature, made also by Conesa, Kitao and Krueger (2009), which I retain for comparability. It avoids capital income taxes turning into subsidies for agents in debt (without complicating the decision problem of the agents), and prevents the possibility of some households dying in debt.

Demographics: The first set of parameters are related to the demographics/life-cycle aspects: $J_R, J, g_n, \pi_j^d, e_j$. These are in line with the assumptions made in other studies in the calibration of OLG models. Agents become economically active at age 21, retire at age $J_R = 66$, and they can live up to $J = 101$ years. These are fairly conventional and innocuous assumptions. The population growth rate is set to its long-run average in the data since the early 1970's, which is equal to $g_n = 0.011$. The survival probabilities π_j^d are taken from Bell and Miller (2002), and the age profile for the efficiency units e_j from Hansen (1993).

Preferences: The discount factor β is chosen to match an equilibrium interest rate of 5%. The corresponding value is $\beta = 1.012$. The consumption share in the Cobb-Douglas aggregator is set to $\eta = 0.357$, which matches an average share of time devoted to market activities equal to 0.33. The risk aversion is set to $\sigma = 3.80$, which matches an elasticity of intertemporal substitution of 0.5.

Technology: The labor share is taken from the Penn World Tables 9.0 (PWT9.0), and implies $\alpha = 39\%$. Also the capital depreciation rate is taken from the Penn World Tables 9.0, and it is equal to $\delta = 0.0493$.

Income process and pensions: The baseline parameterization of the exogenous component in the stochastic income process relies on the values for the Panel Study of Income Dynamics reported by Guvenen (2009). His estimates are $\rho_y = 0.988$ and $\sigma_y^2 = 0.015$ for the persistent component, and $\sigma_f^2 = 0.058$ for the fixed effect. The pension replacement rate is obtained residually and its value $\phi_R = 0.4$ is close to the related statistic of 39.4% reported in OECD (2011).

Taxes and Government: In the initial (pre-reform) BGP, the income tax schedule parameters are taken from Gouveia and Strauss (1994), whose estimates are: $\kappa_0 = 0.258$, $\kappa_1 = 0.768$. κ_2 is instead set to satisfy the balanced-budget requirement.The consumption tax rate is $\tau_c = 0.05$, which is the estimate reported by Mendoza, Razin, and Tesar (1994). The payroll tax rate used to finance the pensions is set to its empirical value of $\tau_R = 0.124$. The tax rate on corporate profits is set to $\tau_f = 0.12$, which matches the average ratio of Corporate taxes/GDP of 2.7% found in the U.S. data in the 1970-2017 period. The related series are plotted in Figure (2).

[Figure 2 about here]

The calibrated value for τ_f might seem low, given that the typical corporate tax rate is in excess of 25%. However, the effective corporate tax rate is substantially lower than the statutory one, likely because of tax elusion and sophisticated accounting practices implemented to reduce the tax base. Notice also how, since the early 1980's, the Corporate taxes/GDP ratio does not exhibit a trend. Finally, The government consumption is set to match a public expenditure/GDP ratio of 17%, which leads to $G = 3.60$.

Economic Growth: The productivity of R&D is set to a value that matches a firm exit rate of 9%, which is obtained from firm dynamics data in the 1995-2007 period: $\lambda = 29.6$. The equilibrium growth rate is chosen to match its value from the Penn World Tables 9.0, and it is equal to $g = 0.018$. Given the probability of a successful innovation of about 9%, this is obtained with a spillover effect equal to $\gamma = 0.1998$.

4 Results

This section presents the main results. First, I discuss the fit of the model, which is followed by the characterization of the welfare optimizing income tax schedule.

4.1 Model Fit

The four panels in Figure (3) plot the cross-sectional profiles obtained in the benchmark model. The model is capable of capturing some salient features of the data, such as the distribution of the average wealth holdings across different age groups, the drop in consumption at retirement, and the hump-shaped profile of labor earnings. Similarly, Figure (4) shows that the model generates a substantial increase in (log-)earnings inequality, as captured by their variance, which jumps from 0.058 at age 21 to 0.999 at age 65. The endogenous growth channel does not seem to play an important role in shaping these outcomes, as similar behaviors have been documented in the literature, e.g. by Storesletten, Telmer and Yaron (2004) and Conesa, Kitao and Krueger (2009). It is worth noting that the model predicts that (on average) the agents are going to start de-cumulating assets before retirement. This is due to the fact that the labor earnings peak early in life, and the individuals rely on assets and asset income to finance a fast consumption growth. Another well-known feature of the data that is replicated by the model is the drop in consumption expenditure upon retirement. Even though individuals accumulate assets, these are not enough to smooth consumption around the retirement period, as the drop in income is substantial, and the agents find optimal to adjust their expenditure in a non-smooth fashion only once during their lives. Furthermore, the discrete increase in leisure enjoyed by the agents transitioning into retirement allows them to optimally spend less in consumption goods without experiencing a decrease in their lifetime welfare. After the quite drastic drop at model age 45, consumption expenditures in the cross-section continue to decrease monotonically, but at a lower pace. The fourth panel shows the implication of working with a Cobb-Douglas utility aggregator over consumption and leisure. The labor supply is not flat: younger individuals tend to work for longer hours, which start decreasing for older workers, and this decrease accelerates after model age 20. This behavior, coupled with the exogenous productivity profile over the life-cycle, explains the evolution of labor earnings, plotted in the third panel.

[**Figures 3 and 4 about here**]

Notice that the labor supply is stationary across generations, because it is not affected by the income growth rate. Only for this variable the cross-sectional profiles and the life-cycle profiles are identical. The other variables in Figure (3) do display different profiles, depending whether we are considering the cross-sectional ones or the life-cycle ones. Since this is going to be important for the interpretation of the welfare effects, Figure (5) reports the life-cycle profiles of the consumption expenditures. The Figure shows that consumption expenditure tends to rise as individuals age, but the drop at retirement is still present. The solid line represents the benchmark case, namely the pre-reform equilibrium. The dashed lines show how the counterfactual economies (i.e., with the optimal tax schedule), imply a faster consumption growth over the life cycle. The two panels refer to two different cases: the left panel focuses on a case where in the counterfactual economy the level of public expenditure is kept fixed at its pre-reform value. Differently, the right panel focuses on a case where the counterfactual economy has a fixed public expenditure/GDP ratio.

[Figure 5 about here]

4.2 Income Tax Schedule Reform

The equilibrium of the model is computed several times. The first time under the current policy regime, i.e. for pre-reform values of the triplet $\{\kappa_0, \kappa_1, G\}$. The other times under a counterfactual economy, i.e. after a policy change represented by the same public expenditure (either in level or as a percentage of GDP), and different income tax schedule parameters $\{\kappa_0^{new}, \kappa_1^{new}, \tau_k^{new}\}$.[7] Namely, the income tax schedule reform is such that capital income is taxed at the proportional rate τ_k^{new}, and labor earnings y^w are taxed according to the same formula above, Eq. (13), but with different parameters (and κ_2 is still set to balance the budget):

$$T^w = \kappa_0^{new} \left[y^w - \left((y^w)^{-\kappa_1^{new}} + \kappa_2 \right)^{-1/\kappa_1^{new}} \right]. \tag{14}$$

4.2.1 Optimal Capital and Labor Taxes

Starting from the baseline tax schedule, a numerical optimization routine is used to characterize the optimal tax code $\{\kappa_0^*, \kappa_1^*, \tau_k^*\}$. This procedure leads to a set of parameters for the labor income tax function and for the capital income proportional tax rate. Table 2 reports the list of tax parameters for different cases. The Status quo is characterized by $\kappa_0 = 0.258$, $\kappa_1 = 0.768$, while the optimal tax function computed by Conesa, Kitao and Krueger (2009) is $\kappa_0^* = 0.23$, $\kappa_1^* = 7.0$, with a sizable capital income tax rate $\tau_k^* = 0.36\%$

[Table 2 about here]

Regarding the capital income tax rate, the Schumpeterian endogenous growth model implies strikingly different results. With a fixed G/Y, $\kappa_0^* = 0.285$, $\kappa_1^* = 7.89$, coupled with an virtually zero capital income tax

[7] The welfare effects are going to compare two different BGP's. I consider both the percentage change in expected utility, and a consumption equivalent (CEV) welfare measure, denoted as ϖ. The former consists of either the ex-ante welfare of the households (i.e., the normalized value function $V(.)$ integrated with respect to the normalized stationary distributions, see equations (17) and (18) in Appendix A), or the Social Welfare (for its definition, see 19 in Appendix A). For the CEV welfare measure, see equations (20) and (21) in Appendix A.

rate $\tau_k^* = 0.09\%$. With a fixed G, the values of the parameters of the optimal tax function are $\kappa_0^* = 0.237$, $\kappa_1^* = 6.73$, and an infinitesimal capital income tax rate $\tau_k^* = 0.00006\%$. This represents the main result of the paper: under the optimal tax schedule, the capital income tax rate is trivially small.

[Figure 6 about here]

To confirm this finding, the panels 1 and 3 in Figure 6 show the response of two welfare measures to changes in the capital income tax rate around the optimal labor income tax function. Both the ex-ante welfare measure and the aggregate (social) welfare are negatively affected by an increasing capital income tax rate. Moreover, the welfare losses of imposing a larger capital income tax rate are substantial. Moving from the optimal zero tax rate to an empirically common value of 40% reduces the ex-ante welfare by almost 10% and the social welfare by 9.5%. Moreover, the CEV gains of moving from the Status quo to the optimal scheme are between $\varpi = 4.76\%$ and $\varpi = 4.82\%$, depending on the assumption on the public expenditure. It goes without saying that these welfare effects are sizable, especially when compared to the corresponding figure of $\varpi = 1.3\%$ obtained by Conesa, Kitao and Krueger (2009).

4.2.2 Income Tax Functions: Status quo Vs. Optimal

It is worth inspecting the mechanism behind such large welfare gains. The first reason is the increased progressivity of the labor income tax schedule, which is easily detected in Figures 7 and 8. Figure 7 is related to the case with a fixed G, while Figure 8 to the case with a fixed G/Y ratio. In the two Figures, the left panels present the tax rate as a function of income, while the right panels the total income taxes. Compared to the Status quo, the optimal scheme warrants an extended region with almost zero labor income taxes for the agents with the lowest labor income. Given that these agents are on average young and asset poor, they have a very high marginal utility of consumption. A more pronounced progressivity delivers a less unequal allocation of consumption, leading to an aggregate welfare gain. Notice also that the status quo and the optimal tax functions intersect. Interestingly, the intersection occurs at a value of labor income that is close to the average individual labor earnings. On the one hand, the lost tax revenues from taxing capital income have to be compensated with higher tax rates on the labor income rich. On the other hand, shifting away from capital taxes increases both the equilibrium profits and the aggregate consumption, leading to a lower need of taxing labor income. In other words, with a fixed public expenditure, the increased revenues from corporate profits and consumption taxes partially compensates the fall in revenues stemming from the almost absent capital taxes.

[Figures 7 and 8 about here]

When the public expenditure is not fixed, the results are quite similar, as Figure 8 reveals. The main difference is that when in the counterfactual economy the G/Y ratio is fixed, there is an increased need to collect tax revenues, as in the economy with the optimal tax scheme GDP is higher. In this case, the rise in both corporate and consumption taxes do not finance the increase in public expenditure. It follows that the income tax code becomes even more progressive and the marginal tax rate for the income rich is much larger than in the benchmark economy.

[Figure 9 about here]

Figure 9 provides six plots presenting the equilibrium response of a set of aggregate outcomes. The top two panels refer to the growth rate and the interest rate. The interest rate increases monotonically in the capital tax, as the households want to be partially compensated for the lost rate of return. For the considered range of capital income tax rates, the interest rate increases by more than a percentage point, moving from 4.4% to 5.5%. In turn, this leads to more costly R&D activities, with the end result of a reduced growth rate. A stark finding is that the fall in the growth rate is limited in size, but it is indeed sufficient to cause a decrease in welfare. For the two extreme capital tax rate values of 0% and 40% the related growth rates are 1.873% and 1.571%, respectively. Finally, from the households' point of view, the after-tax interest rate is lower with a positive capital income tax: this makes saving a less effective instrument to self-insure against the labor income risk. As far as the behavior of aggregate Output, Consumption and Capital, they all decrease monotonically in the capital tax rate, with quantitatively large responses. Differently, the labor input virtually does not have any response (as expected, given the assumed preferences).

5 Robustness Analysis

Although routinely used in quantitative work, the CRRA/Cobb-Douglas utility function assumed in the benchmark analysis fails to satisfy one of the properties outlined by King, Plosser and Rebelo (1988) and King, Plosser and Rebelo (2002) to be consistent with a Balanced Growth Path. Hence, in this section I rely on an alternative formulation of the utility function, taken from the King, Plosser and Rebelo (2002) and Jaimovich and Rebelo (2009) class of functions.[8] Another advantage of this utility function is that it does not feature the unit elasticity embedded in the Cobb-Douglas utility function. In this formulation, I assume that $u\left(c_j, l_j\right) = \frac{\left[c_j\left(1-\psi h_j^{\theta}\right)\right]^{1-\sigma}-1}{1-\sigma}$, and now the per-period utility function implies a relative risk aversion RRA$=\sigma$. Its two parameters ψ and θ can be calibrated to match the average share of time devoted to market activities, as above, but also a Frisch elasticity of 0.5 (a value that is in line with the micro-econometric evidence on this parameter).[9] This parameter

[8] Peterman (2013) undertakes a careful decomposition of the reasons behind a positive capital tax on capital, including additively separable preferences, because they feature a constant Frisch elasticity over the life-cycle. It is worth pointing out that the usual shortcut of introducing a scaling factor in additively separable preferences is not credible in the current framework. Because of the endogenous determination of the technological progress, the scaling factor would be policy dependent. In order to achieve consistency with balanced growth I would need to assume that people's preferences react to the chosen tax policy, as a different scaling factor would be needed for any equilibrium growth rate. In other words, I would need to assume that the parameters are not structural.

[9] The formula for the individual-level Frisch elasticity (ε_i^{Frisch}) is $\varepsilon_i^{Frisch} = \sigma\left(1-\psi h_i^{\theta}\right) / \left[(\theta-1)\,\sigma\left(1-\psi h_i^{\theta}\right) - \psi\theta h_i^{\theta}\right]$. In the calibration, I target the average Frisch elasticity for the working-age agents.

is important as it governs how rigid is the individual labor supply response to changes in the wage rate (in turn stemming from changes in the tax code). Table 2 lists the values of the parameters re-calibrated in equilibrium. The new discount factor is equal to $\beta = 1.008$, the RRA parameter is $\sigma = 2.0$. The disutility of work parameter is set to $\psi = 8.9$, and the convexity of the disutility of work parameter to $\theta = 3.45$. All the other parameters are kept fixed at their benchmark values.[10]

[Table 3 about here]

The panels in Figures 3 and 4 also plot the cross-sectional and life-cycle profiles implied by the version of the model with this utility function. This case is denoted with the KPR label, and plotted with the blue dashed line. The profiles are qualitatively very similar to the corresponding ones obtained in the benchmark case. Quantitatively, two important differences arise. First, the average wealth holdings have more pronounced dynamics. Moreover, as expected, the most noticeable difference pertains the behavior of the labor supply, which is substantially more rigid than in the benchmark case. In the KPR case, the hours worked are almost flat until the retirement age. Similarly, the individuals have flatter labor earnings and their average asset holdings peak later in life, reaching a value that is more than 10% higher than in the benchmark case. The combined behavior of the labor supply and increased wealth around the retirement age implies that the drop in consumption when agents leave the labor force is substantially less drastic. In terms of the optimal tax code, also in this case the optimal one displays an almost zero capital income tax rate $\tau_k^* = 0.2\%$. The optimization procedure now leads to the following parameters for the labor income tax function $\kappa_0^* = 0.287$, $\kappa_1^* = 7.93$, which are remarkably similar to the values found for the benchmark case. As far as the welfare response to capital income taxes is concerned, the panels 2 and 4 in Figure 6 show the response of two welfare measures to changes in the capital income tax rate around the new optimal labor income tax function. The drop in either measure of welfare is striking and seems to be monotonic. However, even though it cannot be visually detected, it is worth mentioning that the capital income tax rate that maximizes welfare is slightly above zero, but trivially small. To conclude with, the optimal income tax schedule with an almost zero capital income tax rate entails even larger welfare gains, which are now 5.8% in consumption equivalent.

6 Conclusions

In this paper I have embedded a Schumpeterian endogenous growth mechanism into a life-cycle model with incomplete markets and rich heterogeneity. The calibrated version of the model shows that it is virtually never optimal to set a positive tax rate on capital income. At most, the tax rate that maximizes welfare is less than 0.2%. This finding is in stark contrast with what Conesa, Kitao and Krueger (2009) found in a similar economy with an exogenous growth rate and without technological progress. The results are robust along several dimensions, such as the welfare metric, the size of the public expenditure and the behavior of the Frisch elasticity

[10] Notice that the parameters related to the R&D and the spillover effects do not need to be recalibrated, as the aggregate labor supply in the initial BGP is the same, and the discount factor is recalibrated to achieve the same interest rate. These ensure that the growth rate is the same as in the benchmark economy (1.8%), and so is the firms' exit rate (9%).

over the life-cycle. Although technically challenging, it would be interesting to extend the model to include risk averse entrepreneurs, in order to match the concentration of wealth observed in the U.S. economy. I leave these extensions and modifications for future work.

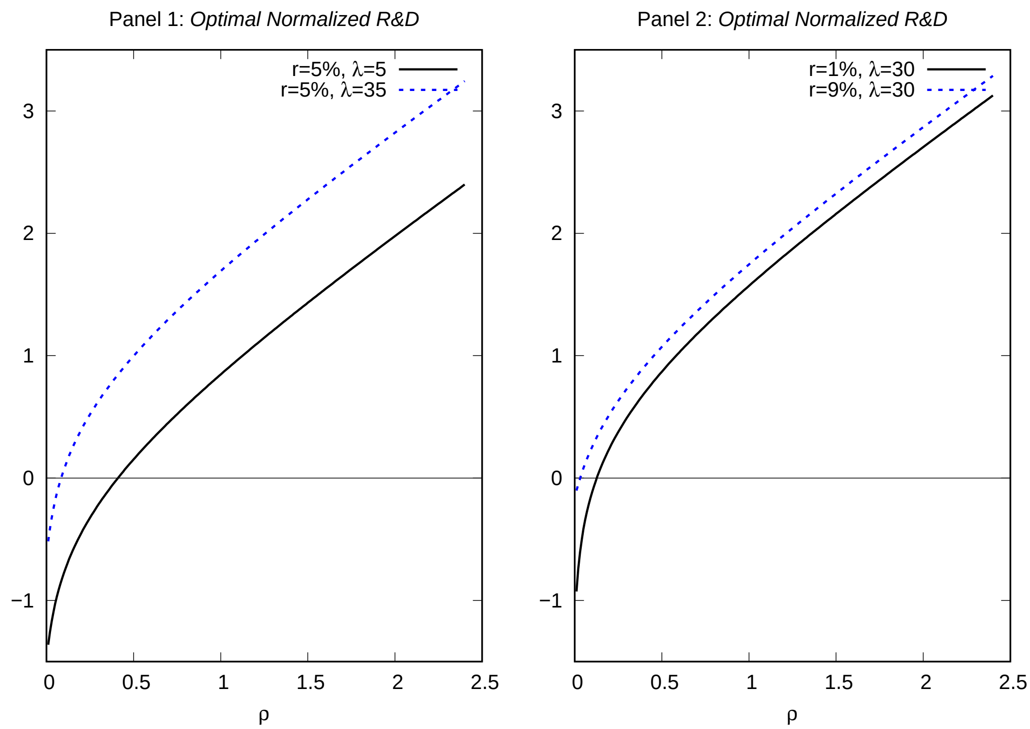

Figure 1: Optimal entrepreneurs' choices on R&D investment. The optimal ρ^* corresponds to the root of these functions.

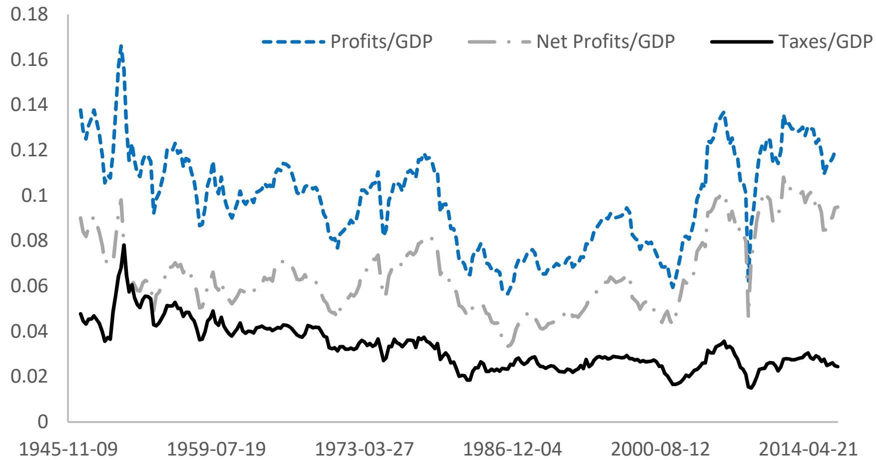

Figure 2: Corporate Profits, Corporate Net Profits (After Tax and without IVA and CCAdj), and Corporate Taxes in the U.S. Each variable is expressed as a share of GDP. Quarterly and Seasonally Adjusted series, 1947Q1-2017Q1.

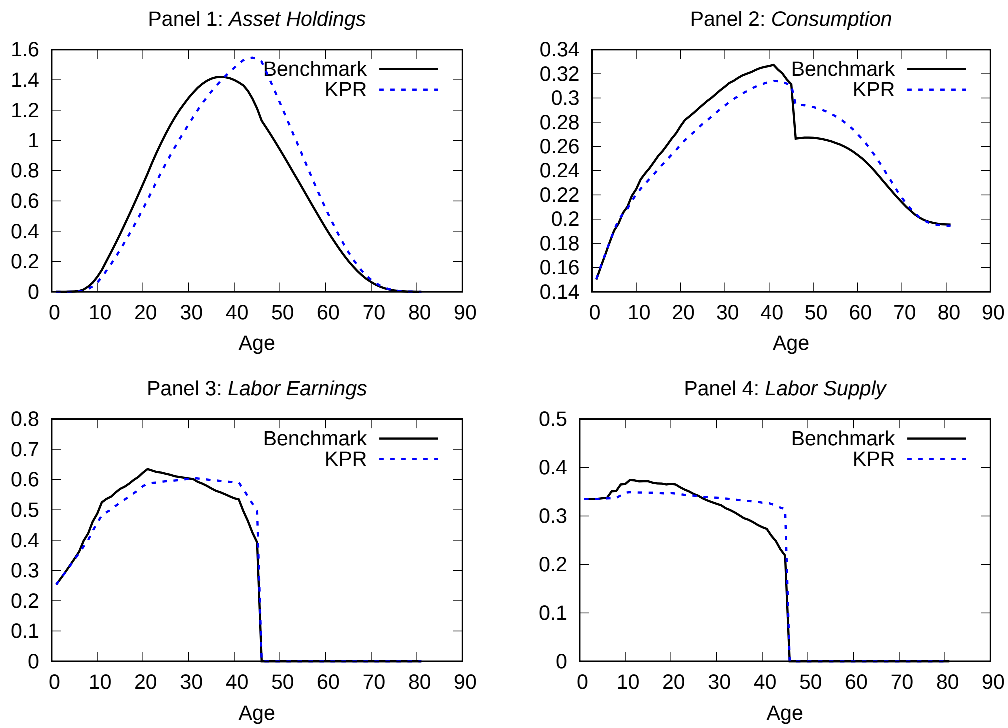

Figure 3: Average Cross Sectional profiles of Asset Holdings, Consumption, Labor Earnings, and Hours Worked. Benchmark model (solid line) and KPR preference specifications (dashed line).

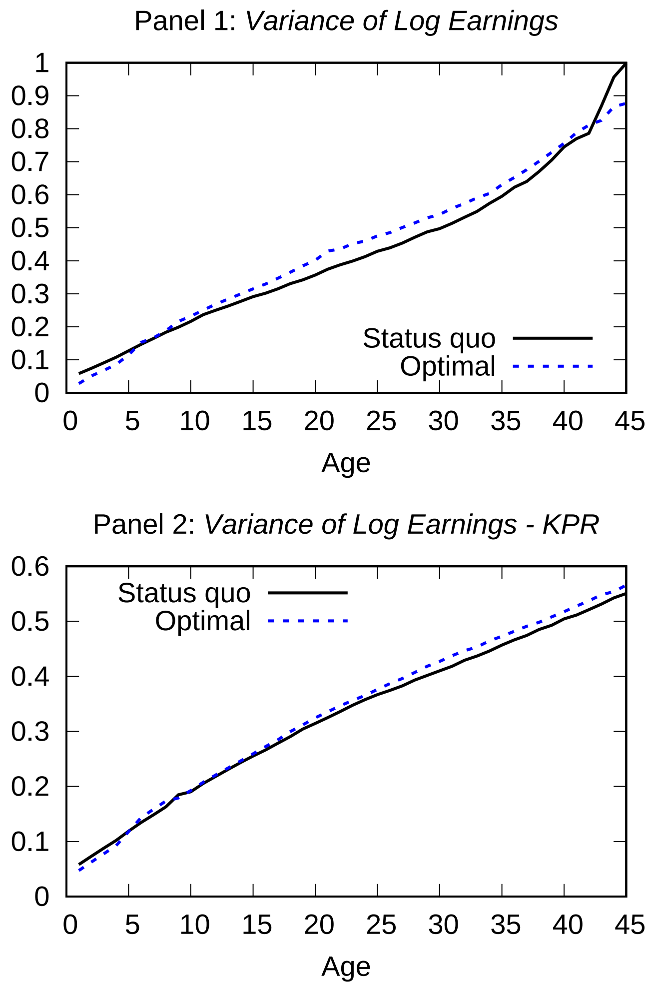

Figure 4: Pre-tax Labor Earnings Inequality. The status quo stands for the pre-reform equilibrium (solid line). The other profiles (dashed lines) are computed under the optimal taxation scheme with either assumption on preferences.

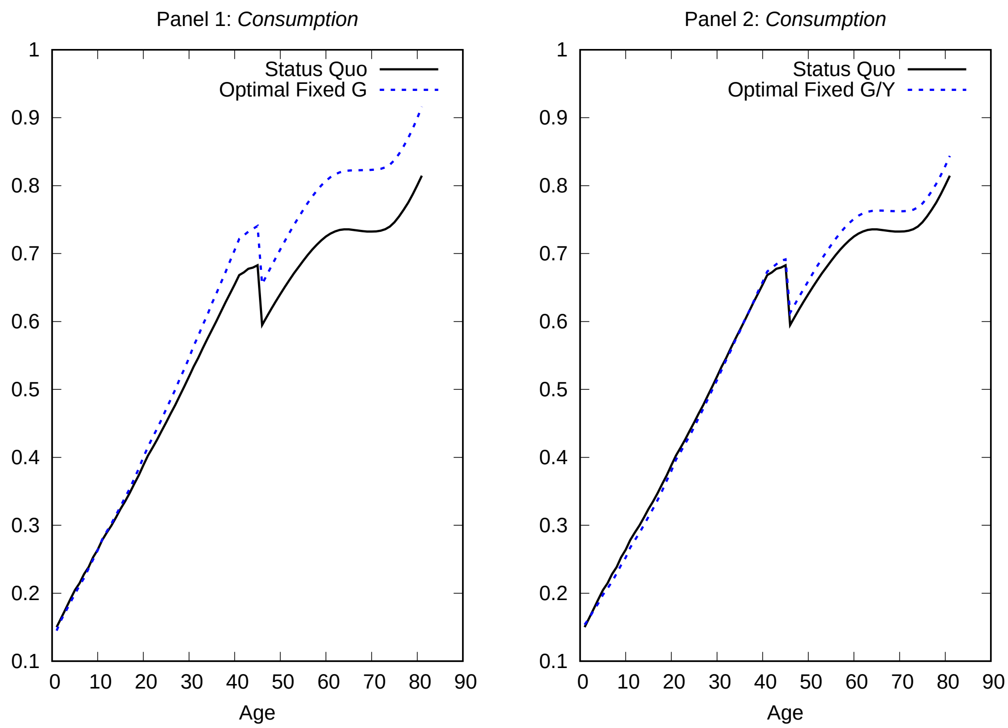

Figure 5: Average Life-cycle profiles of Consumption. The status quo stands for the pre-reform equilibrium (solid line). The other profiles (dashed lines) are computed under the optimal taxation scheme with either a fixed G or fixed G/Y.

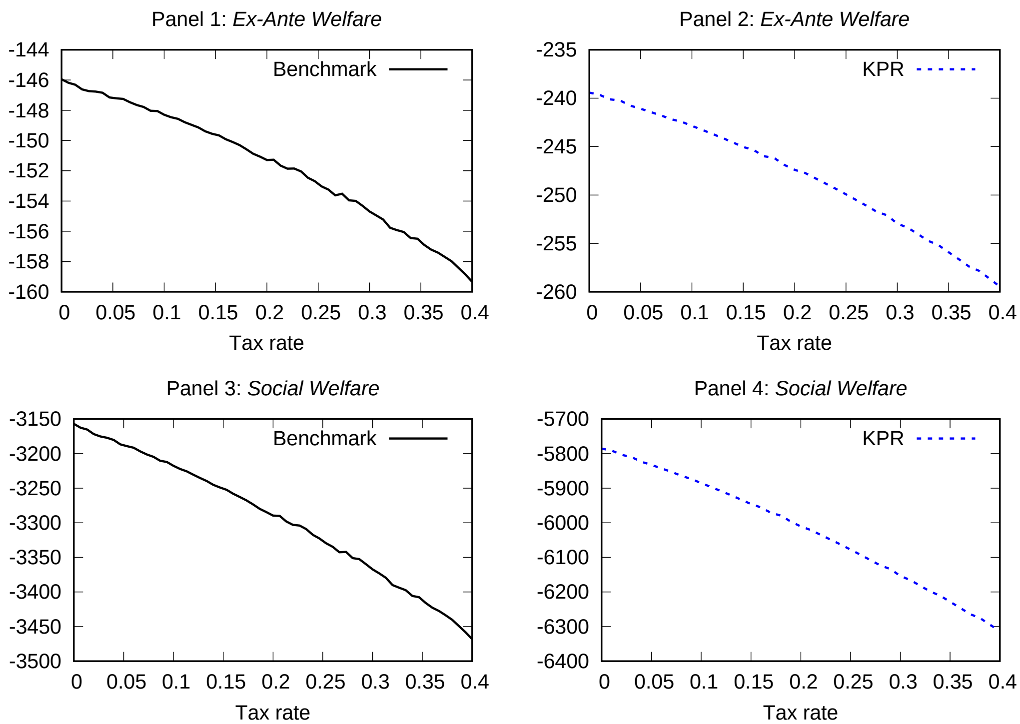

Figure 6: Ex-ante Welfare and Social Welfare responses to changes in the capital tax rate. Benchmark model (solid lines) and KPR specifications (dashed lines).

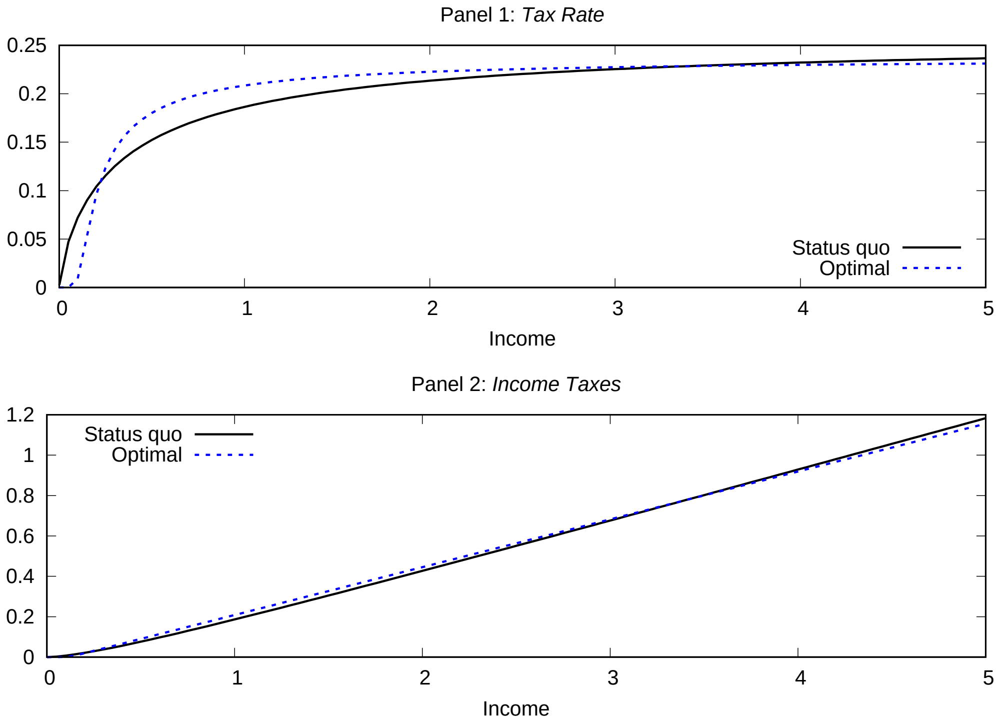

Figure 7: Tax Functions with a fixed G. Status quo Gouveia/Strauss Estimates Vs. Optimal Schedule (dashed).

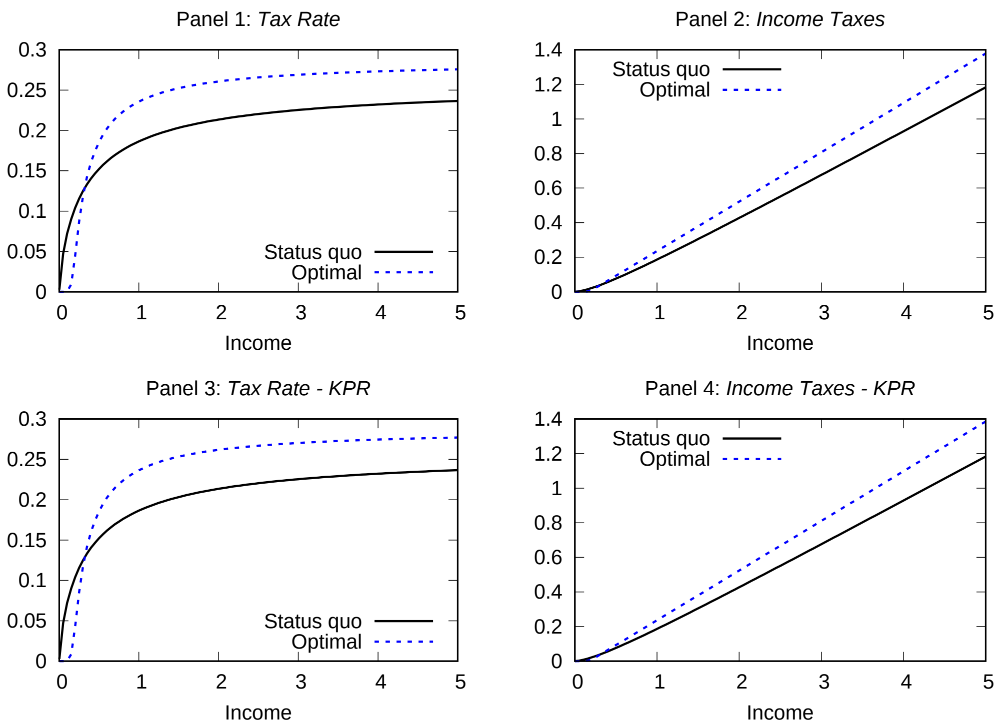

Figure 8: Income Tax Functions with a fixed G/Y ratio. Status quo Gouveia/Strauss Estimates Vs. Optimal Schedule (dashed).

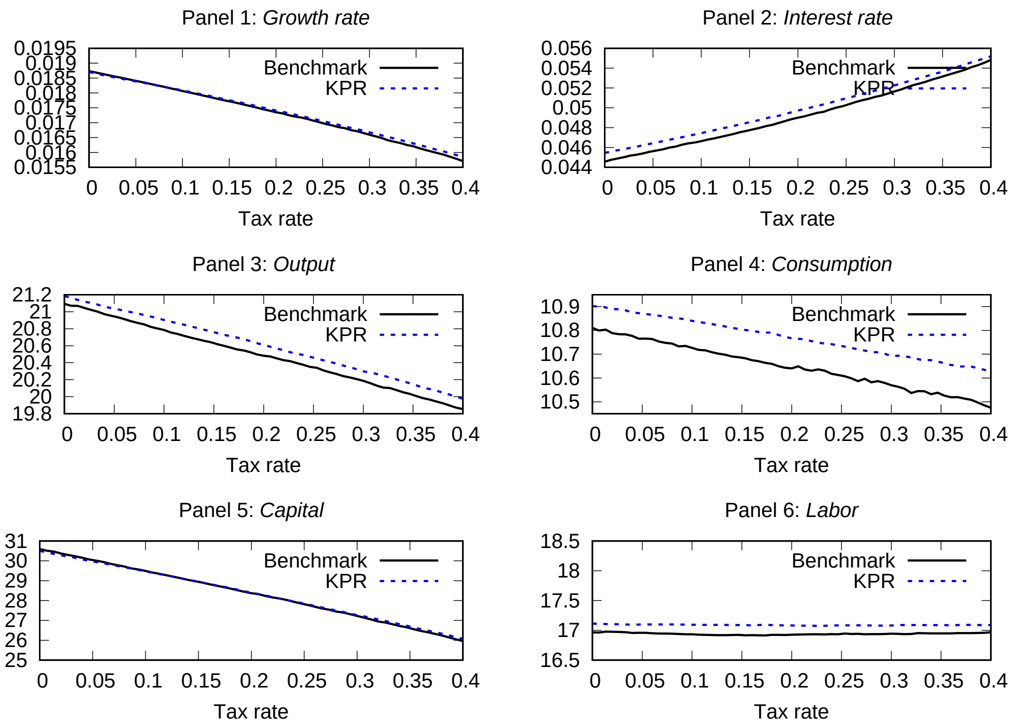

Figure 9: Growth Rate, Interest Rate, Output, Capital, Consumption and Labor equilibrium responses to changes in the capital tax rate. Benchmark model and KPR specifications (dashed).

Parameter	*Value*	*Target*
Model Period	*Year*	*Frequency of PSID Data*
J - Maximum Age	81	*Certain death at age* 100
J_R *- Maximum Working Age*	46	*Retirement at age* 65
π_j^d *- Death probability*	-	*Bell and Miller* (2002)
g_n *- Population growth*	0.011	*Data*
β *- Rate of time preference*	1.01213	*Pre-tax Interest rate* $\approx 5\%$
η *- Consumption share*	0.357	*Average hours* $= 0.33$
σ *- Risk Aversion*	3.80	*Elasticity of Intertemporal Substitution* $= 0.5$
δ *- Capital depreciation rate*	0.0493	*Capital depreciation estimates - PWT9.0*
λ *- Weight in prob. of successful innovation*	29.6	*Firms exit rate* $\approx 9\%$
γ *- Spillover Effect*	0.1998	*Growth rate* 1.8% *- PWT9.0*
α *- 1-Labor share*	0.39	*Labor share of output* $= 61\%$ *- PWT9.0*
σ_y^2 *- Var. of the temporary income shocks*	0.015	*Guvenen* (2009)
ρ_y *- Persistence of the temp. income shocks*	0.988	*Guvenen* (2009)
σ_f^2 *- Var. of the fixed effect*	0.058	*Guvenen* (2009)
ξ *- Government Consumption*	0.17	*G/GDP* $= 17\%$
κ_0 *- Marginal Income Tax*	0.258	*Gouveia and Strauss* (1994)
κ_1 *- Progressivity Income Tax*	0.768	*Gouveia and Strauss* (1994)
τ_c *- Consumption Tax Rate*	0.05	*Mendoza, Razin, and Tesar* (1994)
τ_f *- Profits Tax Rate*	0.12	*Corporate Taxes/GDP* $\approx 2.7\%$
τ_R *- Payroll Tax Rate*	0.124	*Data*
b - Borrowing limit	0	*No borrowing allowed*

Table 1: Calibration, Benchmark Model

Case	κ_0	κ_1	τ_k	$CEV(\%)$
Status quo	0.258	0.768	–	–
Fixed G	0.237	6.734	0.0000006	4.821
Fixed G/Y	0.285	7.891	0.0008659	4.761
KPR	0.287	7.929	0.0023872	5.795
CKK AER 2009	0.23	7.0	0.36	1.330

Table 2: Tax Schedules Parameters and Welfare changes from the Status quo to the Optimal tax schedule.

Parameter	*Value*	*Target*
β - *Rate of time preference*	1.00811	*Pre-tax Interest rate* $\approx 5\%$
ψ - *Disutility of work*	8.9	*Average hours* $= 0.33$
σ - *Risk Aversion*	2.0	*Elasticity of Intertemporal Substitution* $= 0.5$
θ - *Convexity disutility of work*	3.45	*Frisch elasticity* $= 0.5$

Table 3: Calibration in Equilibrium, KPR Model

References

Aghion, P., Akcigit, U, and Fernandez-Villaverde, J. (2013). "Optimal Capital Versus Labor Taxation with Innovation-Led Growth," *NBER Working Paper n.19086.*

Aiyagari, R. (1994). "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics,* Vol. 109, 659-684.

Aiyagari, R., and McGrattan, E. (1998). "The optimum quantity of debt," *Journal of Monetary Economics,* Vol. 42, 447-469.

Aoki, S., and Nirei, M. (1998). "Zipf's Law, Pareto's Law, and the Evolution of Top Incomes in the United States," *American Economic Journal: Macroeconomics,* Vol. 9, 36-71.

Atkeson, A., Chari, V., and Kehoe, P. (1999). "Taxing Capital Income: A Bad Idea," *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 23, 3-17.

Bell, F., and Miller, M. (2002). "Life Tables for the United States Social Security Area 1900–2100." Office of the Chief Actuary, Social Security Administration, Actuarial Study 116.

Castaneda, A., Diaz-Gimenez, J., and Rios-Rull, J.V. (2003). "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, Vol. 111, 818-857.

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

Conesa, J., Kitao, S. and Krueger, D. (2009). "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, Vol. 99, 25-48.

Cozzi, G., Pataracchia, B., Pfeiffer, P., and Ratto, M. (2017). "How much Keynes and how much Schumpeter? An Estimated Macromodel of the US Economy," Mimeo.

Cozzi, M. (2014). "Equilibrium Heterogeneous-Agent Models as Measurement Tools: some Monte Carlo Evidence," *Journal of Economic Dynamics and Control*, Vol. 39, 208-226.

Erosa, A., and Gervais, M. (2002). "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory,* Vol. 105, 338-369.

Gervais, M. (2012). "On the Optimality of Age-dependent Taxes and the Progressive U.S. Tax System," *Journal of Economic Dynamics and Control,* Vol. 36, 682-691.

Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, Vol. 70, 569-587.

Gouveia, M., and Strauss, R. (1994). "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, Vol. 47, 317-339.

Guvenen, M.F. (2009). “An Empirical Investigation of Labor Income Processes,” *Review of Economic Dynamics*, Vol. 12, 58-79.

Hansen, G. (1993). “The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked.” *Journal of Applied Econometrics*, Vol. 8, 71-80.

Heathcote, J., Perri, F., and Violante, G. (2010). “Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006,” *Review of Economic Dynamics*, Vol. 13, 15-51.

Heathcote, J., Storesletten, K., and Violante, G. (2009). “Quantitative Macroeconomics with Heterogeneous Households,” *Annual Review of Economics*, Vol. 1, 319-354.

Heathcote, J., Storesletten, K., and Violante, G. (2010). “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, Vol. 118, 681-722.

Howitt, P., and Aghion, P. (1998). “Capital accumulation and innovation as complementary factors in long-run growth,” *Journal of Economic Growth*, Vol. 3, 111-130.

Huggett, M. (1993). “The Risk-free Rate in Heterogeneous-agent Incomplete-insurance Economies,” *Journal of Economic Dynamics and Control,* Vol. 17, 953-969.

Huggett, M. (1996). “Wealth distribution in life-cycle economies,” *Journal of Monetary Economics*, Vol. 38, 469-494.

Imrohoroglu, S. (1998). “A Quantitative Analysis of Capital Income Taxation,” *International Economic Review*, Vol. 39, 307-328.

Jaimovich, N., and Rebelo, S. (2009). “Can news about the future drive the business cycle?,” *American Economic Review*, Vol. 99, 1097-1118.

Judd, K. (1985). “Redistributive Taxation in a Simple Perfect Foresight Model,” *Journal of Public Economics*, Vol. 28, 5983.

King, R., Plosser, C., and Rebelo, S. (1988). “Production, Growth and Business Cycles I. The Basic Neoclassical Model,” *Journal of Monetary Economics*, Vol. 21, 195-232.

King, R., Plosser, C., and Rebelo, S. (2002). “Production, Growth and Business Cycles: Technical Appendix,” *Computational Economics*, Vol. 20, 87-116.

Mendoza, E., Razin, A., and Tesar, L. (1994). “Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption,” *Journal of Monetary Economics*, Vol. 34, 297-323.

Nuno, G. (2011). “Optimal research and development and the cost of business cycles,” *Journal of Economic Growth*, Vol. 16, 257-283.

OECD. (2011). *Pensions at a Glance 2011: Retirement-income Systems in OECD and G20 Countries*, OECD Publishing, Paris.

Peterman, W. (2013). "Determining the Motives for a Positive Optimal Tax on Capital," *Journal of Economic Dynamics and Control,* Vol. 37, 265-295.

Piketty, T. (2014). Capital in the Twenty-First Century, Harvard University Press, 2014.

Storesletten, K., Telmer, C. and Yaron, A. (2004). "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, Vol. 51, pp. 609-633.

Ventura, G. (1999). "Flat Tax Reform: a Quantitative Exploration," *Journal of Economic Dynamics and Control*, Vol. 23, 1425-1458.

Appendix A - The OLG Model and its Recursive Representation

7 Stationary Equilibrium

In this Section, first the problem of the agents in their recursive representation is defined, then I provide a formal definition of the equilibrium concept used in this model, the recursive competitive equilibrium. The individual state variables are: age $j \in \mathcal{J} = \{1, ..., J\}$, the fixed effect $f \in \mathcal{F} = \{-\sigma_f, +\sigma_f\}$, the persistent shock component of the labor endowment $\varepsilon \in \mathcal{E} = \{\varepsilon_{\min}, ..., \overline{\varepsilon}, ..., \varepsilon_{\max}\}$ and asset holdings $a \in \mathcal{A} = [-b, \overline{a}]$. Notice that ε is discretized with the Rouwenhorst method, using a 7-state Markov chain. The transition function of the labor endowment shocks is represented by the matrix $\Pi(\varepsilon', \varepsilon) = [\pi(v, z)]$, where each element $\pi(v, z)$ is defined as $\pi(v, z) = \Pr\{\varepsilon_{j+1} = z | \varepsilon_j = v\}$, $v, z \in \mathcal{E}$. In every period the exogenous labor endowments are given by $\epsilon_{j,\varepsilon,f} = e_j \varepsilon f$. The stationary distribution of working-age agents is denoted by $\mu_j(a, \varepsilon, f)$ while that of retirees with $\mu_j^R(a)$. Φ_j denotes the share of each cohort j in the total population. These satisfy the recursion $\Phi_{j+1} = \left(\frac{1-\pi_j^d}{1+g_n}\right) \Phi_j$, and are normalized to add up to 1.

7.1 Problem of the agents

The model is solved backwards, starting from the terminal age J and with the assumption that the terminal utility value is zero, i.e. $V_{J+1} = 0$.

7.1.1 Problem of the retirees

The value function of an age-j retired agent whose current asset holdings are equal to a is denoted with $V_j^R(a)$. The problem of these agents can be represented as follows:

$$V_j^R(a) = \max_{c,a'} \left\{ u(c) + \beta \left(1 - \pi_j^d\right) V_{j+1}^R(a') \right\} \tag{15}$$

$$\begin{aligned}
&s.t. \\
&c + a' = (1 + r) a + \overline{y}_R + TR \\
&c \geq 0, \quad a' > 0
\end{aligned}$$

In the budget constraint notice the presence of the common pension payment $\overline{y}_R$. Retired agents also receive an accidental bequest TR, which is a lump-sum transfer.

7.1.2 Problem of the workers

The value function of a working-age agent whose current asset holdings are equal to a, whose current efficiency units shock is ε and whose fixed effect is f is denoted with $V_j(a, \varepsilon, f)$. The problem of these agents can be represented as follows:

$$V_j\left(a,\varepsilon,f\right)=\max_{c,l,a'}\left\{u(c,l)+\beta\left(1-\pi_j^d\right)\sum_{\varepsilon'}\pi\left(\varepsilon',\varepsilon\right)V_{j+1}\left(a',\varepsilon',f\right)\right\} \tag{16}$$

$$\begin{aligned}
&s.t.\\
&c+a'=\left(1+r\right)a+(1-\tau_R)(1-l)w\epsilon_{j,\varepsilon,f}+TR-T(y)\\
&a_0=0,\quad c\geq 0,\quad a'>-b,\quad y=ra+y^w
\end{aligned}$$

Non-retired agents have to set optimally their consumption/savings and labor supply plans. They enjoy utility from consumption and leisure, and face some uncertain events in the future. In the next period they can still be alive, and with probability $\pi\left(\varepsilon',\varepsilon\right)$ they transit from their current efficiency units ε to the value ε'. These agents pay total income taxes $T(y)$. They also pay a proportional tax τ_R on their labor earnings to finance the pension scheme. Finally, they are born with the average shock $\overline{\varepsilon}$, with no wealth, but they receive the lump-sum accidental bequest TR, and are subject to an exogenous borrowing constraint, $b\geq 0$.

7.2 Recursive Stationary Equilibrium

Since in equilibrium the economy is growing along a BGP, the dynamic programming problem is non-stationary. Every non stationary variable needs to be transformed into their stationary counterpart. This is achieved dividing a generic variable X_t at time t by the average technological index A_t, $\widetilde{x}_t\equiv X_t/A_t$, where the tilde denotes the transformed variable.

Definition 1 *For given public policies $\{\tau_c,\tau_f,\tau_R,\kappa_0,\kappa_1,G\}$ a recursive stationary equilibrium is a set of (transformed) decision rules, $\left\{c_j\left(\widetilde{a},\varepsilon,f\right),l_j\left(\widetilde{a},\varepsilon,f\right),a_j'\left(\widetilde{a},\varepsilon,f\right)\right\}_{j=1}^{J_R-1}$ and $\left\{c_j^R\left(\widetilde{a}\right),a_j^{R\prime}\left(\widetilde{a}\right)\right\}_{j=J_R}^{J}$, value functions, $\left\{V_j\left(\widetilde{a},\varepsilon,f\right)\right\}_{j=1}^{J_R-1}$ and $\left\{V_j^R\left(\widetilde{a}\right)\right\}_{j=J_R}^{J}$, prices $\{P,r,\widetilde{w}\}$, normalized R&D expenditure ρ, endogenous growth rate $g=(1-e^{-\rho})\gamma$, and a set of stationary distributions, $\left\{\mu_j\left(\widetilde{a},\varepsilon,f\right)\right\}_{j=1}^{J_R-1}$ and $\left\{\mu_j^R\left(\widetilde{a}\right)\right\}_{j=J_R}^{J}$, such that:*

- Given relative prices $\{r,\widetilde{w}\}$, taxes and pension benefits $\widetilde{\widetilde{y}}_R$, the individual policy functions $\left\{c_j\left(\widetilde{a},\varepsilon,f\right),h_j\left(\widetilde{a},\varepsilon,f\right),a_j'\left(\widetilde{a},\varepsilon,f\right)\right\}_{j=1}^{J_R-1},\left\{c_j^R\left(\widetilde{a}\right),a_t^{R\prime}\left(\widetilde{a}\right)\right\}_{j=J_R}^{J}$ solve the household problems (15)-(16), and $\left\{V_j\left(\widetilde{a},\varepsilon,f\right)\right\}_{j=1}^{J_R-1},\left\{V_j^R\left(\widetilde{a}\right)\right\}_{j=J_R}^{J}$ are the associated value functions.
- Given relative prices $\{P,r,\widetilde{w}\}$ and public policies, $\widetilde{K/L}$ solves the final good sector firm's problem.
- Given relative prices $\{P,r,\widetilde{w}\}$ and public policies, ρ solves the entrepreneur's problem.
- The labor market is in equilibrium, and the labor input L corresponds to the total supply of labor efficiency units

$$L=\sum_{j=1}^{J_R-1}\Phi_j\int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}}h_j\left(\widetilde{a},\varepsilon,f\right)\epsilon_{j,\varepsilon,f}d\mu_j\left(\widetilde{a},\varepsilon,f\right)$$

- The asset market clears

$$(1+g)(1+g_n)\widetilde{K} = \sum_{j=1}^{J_R-1} \Phi_j \int\limits_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} a_j'(\widetilde{a},\varepsilon,f)\,d\mu_j(\widetilde{a},\varepsilon,f) + \sum_{j=J_R}^{J} \Phi_j \int\limits_{\mathcal{A}} a_j^{R\prime}(\widetilde{a})\,d\mu_j^R(\widetilde{a})$$

- The goods market clears

$$\widetilde{Y} = \widetilde{C}+\widetilde{I}+\widetilde{G}+\widetilde{RD} = \sum_{j=1}^{J_R-1} \Phi_j \int\limits_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} c_j(\widetilde{a},\varepsilon,f)\,d\mu_j(\widetilde{a},\varepsilon,f)+\sum_{j=J_R}^{J} \Phi_j \int\limits_{\mathcal{A}} c_j^R(\widetilde{a})\,d\mu_j^R(\widetilde{a})+\widetilde{C_e}+(\delta+g_n)\,K+\widetilde{G}+\widetilde{RD}$$

- The government's budget is balanced, that is tax revenues from (capital and labor) income taxation, consumption taxes and taxes on profits are equal to the government purchases G

$$G = TAXES$$

- Total accidental bequests TR are equal to wealth holdings of the individuals that die

- The stationary distributions $\left\{\mu_j(\widetilde{a},\varepsilon,f),\mu_j^R(\widetilde{a})\right\}$ satisfy

$$\mu_{j+1}(\widetilde{a}',\varepsilon',f) = \int \nu(\widetilde{a},\varepsilon,f,j,\widetilde{a}',\varepsilon')\,d\mu_j(\widetilde{a},\varepsilon,f) \tag{17}$$

$$\mu_{j+1}^R(\widetilde{a}') = \int \nu^R(\widetilde{a},j,\widetilde{a}')\,d\mu_j^R(\widetilde{a}) \tag{18}$$

In equilibrium the measure of agents in each state is time invariant and consistent with individual decisions, as given by the above two equations (17)-(18), where $\nu(.)$ and $\nu^R(.)$ are the transition functions.

- The social welfare measure W^S is utilitarian, i.e. it weights the agents' lifetime utilities by their mass along a BGP

$$W^S = \sum_{j=1}^{J_R-1} \Phi_j \int\limits_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} V_j(a,\varepsilon,f)\,d\mu_j(a,\varepsilon,f) + \sum_{j=J_R}^{J} \Phi_j \int\limits_{\mathcal{A}} V_j^R(a)\,d\mu_j^R(a) + \left(\frac{1+r}{r}\right) C_e \tag{19}$$

- The consumption based welfare measure ϖ (ϖ_{KPR} for the KPR case) is the percentage increase in consumption in all states of the world that makes welfare in the counterfactual economy $W^1(\varpi)$ equal to welfare in the baseline one W^0

$$W^0 = W^1(\varpi)$$

$$\varpi = \left(\frac{W^1}{W^0}\right)^{\frac{1}{\eta(1-\sigma)}} - 1 \tag{20}$$

$$\varpi_{KPR} = \left(\frac{W^1_{KPR}}{W^0_{KPR}}\right)^{\frac{1}{1-\sigma}} - 1 \tag{21}$$

Appendix B - Computation

- All codes solving the model economies and simulating samples of agents were written in the FORTRAN 95 language, relying on the Intel Fortran Compiler, build 17.1 (with the IMSL library). They were compiled selecting the O2 option (maximize speed), and without automatic parallelization. They were run on different 64-bit PC platforms, all running Windows 10 Professional, either with an Intel $i7-6700k$ Quad Core processor clocked at 4.6 Ghz, or with an Intel $i7-2600k$ Quad Core processor clocked at 4.4 Ghz.

- For either version of the OLG model, the optimization with respect to the tax schedule parameters takes up to 14 hours to complete. The procedure is initialized with different guesses running in parallel, and refined starting from the local maxima. Typically from 15 to 35 iterations on the endogenous variables are needed to find each equilibrium, and the Nelder-Mead algorithm converges in 50 to 80 iterations.

- In the actual solution of the models I need to discretize the continuous state variable a. I rely on an unevenly spaced grid, with the distance between two consecutive points increasing geometrically. This is done to allow for a high precision of the policy rules at low values of a, where the change in curvature is more pronounced. I use 101 points, as increasing the number of points does not affect the results considerably. The lowest value is the borrowing constraint b and the highest one being a value $a_{\max}$ high enough not to be binding.

- In the model, ε is discretized with the Rouwenhorst method, using a 7-state Markov chain. This method has several desirable properties, especially when working with highly persistent processes.

- The model is solved with a backward recursion on the Bellman equations. I start from the terminal value $V_{J+1}^{R}=0$, and at each age, for every point in the state space, I solve the constrained maximization problem. I retrieve the policy functions, $a_j'\left(\widetilde{a},\varepsilon,f\right)$, $a_j^{R\prime}\left(\widetilde{a}\right)$, and $h_j'\left(\widetilde{a},\varepsilon,f\right)$. Notice that I do not restrict the agents' asset holdings to belong to a discrete set and I solve for the optimal decision rules relying on the Euler equations. As for the approximation method, I rely on a linear approximations of the policy functions, when evaluated at the relevant wealth value.

- The stationary distributions are computed relying on their definitions (17)-(18). I rely on these recursions and compute numerically the transition functions.

- The asset market is in equilibrium when the current guess for the interest rate r_0 achieves a capital excess demand which is less than 0.1% of the market size. In turn, this implies that the excess demand in the final good market is always less than 0.1% of the market size.

- The welfare measures W, and W^S are just the newborns value functions, and the numerical integral of the value functions (integrated with respect to the steady state distributions) plus the consumption of the entrepreneurs.

Appendix C - Solution Algorithms

This algorithm represents the computational procedure used to solve the OLG model:

1. Generate a discrete grid over the asset space $[-b, ..., a_{\max}]$.
2. Generate a discrete grid over the income shocks with the Rouwenhorst method $[\varepsilon_{\min}, ..., \varepsilon_{\max}]$.
3. Generate a discrete grid over the fixed effect f.
4. Set the values of the tax schedule parameters (κ_0, κ_1 or κ_0^{new}, κ_1^{new}, τ_k^{new}).
5. Guess the interest rate r_0.
6. Guess the accidental bequest TR_0.
7. Guess the labor supply L_0.
8. Guess the pension benefits $\overline{y}_{R,0}$.
9. Guess the income tax schedule parameter $\kappa_{2,0}$.
10. Get the capital demand K_0 and wages w_0.
11. Find the monopolists' profits π_0.
12. Find the optimal rescaled R&D expenditure ρ_0.
13. Get the growth rate g_0.
14. Transform the HH's problem into its stationary version.
15. Get the (transformed) saving functions $a_j'\left(\widetilde{a}, \varepsilon, f\right), a_j^{R\prime}\left(\widetilde{a}\right)$, the labor supply functions $h_j\left(\widetilde{a}, \varepsilon, f\right)$ and the value functions $V_j\left(\widetilde{a}, \varepsilon, f\right), V_j^R\left(\widetilde{a}\right)$.
16. Get the (transformed) stationary distributions $\mu_j\left(\widetilde{a}, \varepsilon, f\right), \mu_j^R\left(\widetilde{a}\right)$.
17. Get the equilibrium income tax schedule parameter $\kappa_{2,1}$.
18. Get the aggregate capital supply and check the asset market clearing; Get r_1.
19. Get the aggregate labor supply and check the labor market clearing; Get L_1.
20. Update r_0', TR_0', L_0', $\overline{y}_{R,0}'$, $\kappa_{2,0}'$(with a relaxation method).
21. Iterate until asset market clearing, labor market clearing, balanced budget, and aggregate consistency of the accidental bequests and pensions.
22. Get the consumption functions $c_j\left(\widetilde{a}, \varepsilon, f\right), c_j^R\left(\widetilde{a}\right)$ and check the final good market clearing.
23. Compute the ex-ante welfare W and the social welfare W^S.
24. Find the welfare maximizing tax schedule κ_0^*, κ_1^*, τ_k^* with the Nelder-Mead method, repeating steps 4-23 at each iteration.