Transfers versus Investment: The Politics of Intergenerational Redistribution and Growth*

Martín Gonzalez-Eiras
Universidad de San Andrés†

Dirk Niepelt
Study Center Gerzensee
and
Institute for International Economic Studies, Stockholm University‡

October 13, 2006

Abstract

We analyze the tradeoff between intergenerational transfers and public education in an OLG economy where endogenous growth is driven by human and physical capital accumulation. Allowing for a wide variety of fiscal instruments available to policy makers, we contrast the Markov perfect politico-economic equilibrium with the Ramsey allocation. Calibrated to U.S. data, the closed-form solutions of the model predict shares of GDP devoted to public education and transfers, respectively, of 5 and 7 percent, in line with the data. The Ramsey policy, in contrast, calls for shares of 10 and 4 percent, respectively. Due to these discrepancies in the budget shares, the annual growth rate in politico-economic equilibrium is about 0.3 percentage points lower than under the Ramsey policy. Nevertheless, the politico-economic equilibrium allocation is consumption and production efficient.

1 Introduction

Over the last half-century, most developed countries have witnessed a secular increase in the size of the "welfare state," defined as government programs related to redistribution and social insurance. Often, this increase has not been matched by a corresponding rise in government expenditures for productive investment purposes—transfers appear to "crowd out" investment expenditures. While some observers criticize these developments by pointing to the negative long-run implications for growth and the welfare of future generations, political factors appear to render a reversal of the trend difficult.

In this paper, we develop a tractable framework that allows to analyze both the causes and consequences of the reallocation of resources in government budgets. Representing welfare state expenditures on one hand and productive investments on the other, we focus on two government

*For useful comments, we thank seminar participants at the University of Copenhagen (EPRU), the Study Center Gerzensee, and the Universidad de San Andrés.

†Vito Dumas 284, B1644BID Victoria, Pcia. Buenos Aires, Argentina. E-mail: mge@udesa.edu.ar

‡P.O. Box 21, CH-3115 Gerzensee, Switzerland. E-mail: dirk.niepelt@iies.su.se mkvA.tex

spending categories that are of central importance to developed economies: Transfers between "middle-aged" workers and "old" retirees; and public expenditures for education fostering human capital accumulation. Building on Diamond's (1965) overlapping generations model, our framework endogenizes a variety of political and economic choices. As economic agents, households take prices, taxes and pension benefits as given. As voters, households choose taxes on labor- and capital-income as well as the expenditure shares for inter-generational transfers and education investment. Realistically, we assume that the political process lacks commitment; moreover, voters are assumed to be non-altruistic.

In this setting, higher labor income taxes depress disposable income and thus, capital accumulation and growth. However, these effects caused by the financing of government spending represent only one side of the coin. As the model shows, the composition of government spending is of similar, if not greater importance. While transfers to retirees lower the incentive to save and further depress growth, public education expenditure fosters human capital accumulation. Holding taxes as a share of GDP constant, the composition of government spending therefore has strong effects on growth, intergenerational income distribution, and welfare.

The effects of policy on incomes, human and physical capital accumulation, and factor prices are of different concern to the three generations. In particular, only retirees directly benefit from transfers to their group, while only students and workers benefit from the effect of education investment on human capital and thus, future wages and returns on saving. When evaluating the policy platforms (comprising tax rates and spending shares) on offer in the political arena, the different groups of voters therefore disagree as to which platform should ideally be implemented. We model the resolution of the ensuing conflict under the assumption of probabilistic voting, representing electoral competition between parties under the presumption that voters' support for a party, conditional on the party platform, is subject to a small degree of randomness. This randomness induces a continuous mapping from parties' electoral platforms to vote shares, in contrast to the more common assumption of a pivotal median voter. As a consequence, the probabilistic-voting assumption allows us to capture gradual differences in the support for certain policies, and gradual changes in this support due to changes in the economic or demographic environment, even in a stark three-period-lived overlapping-generations environment.

Political choices of tax rates and spending shares do not only affect incomes, human and physical capital accumulation, and factor prices. Absent commitment, these choices also have implications for future policy choices. In addition to the "economic" repercussions of policy, voters therefore have to account for the "political" repercussions of their choices. To that effect, voters must form expectations about the equilibrium relationship between state variables and policy choices in the future. To characterize the subgame-perfect tax, transfer, and investment choices in politico-economic equilibrium, we have to take a stance on the set of state variables the equilibrium policy functions might possibly depend on. We restrict this set of state variables to the fundamental state variables in the economy, excluding artificial state variables of the type sustaining trigger strategy equilibria. In our view, this restriction has several advantages. First, it absolves us from having to make arbitrary assumptions about the type of trigger strategy to consider and thus, the implied equilibrium. Second, it stresses our assumption that political choices suffer from a lack of commitment, including commitment to particular trigger strategies. Most importantly, we find the restriction plausible. While we do not want to deny that the *existence* of intergenerational transfers or public education may also owe to reputational arrangements akin to a social contract, we find it plausible that the *size* of these programs depends more directly on the economic environment than through underlying trigger strategies. By focusing on the Markov perfect equilibrium, we aim at identifying these fundamental and robust forces that shape economic policy.[1]

Under standard functional form assumptions, we are able to fully characterize politico-economic equilibrium in closed form. The optimal strategy for the vote-seeking parties is to propose a policy platform maximizing a weighted average of the welfare of all voters. Since retirees favor old-age transfers and students and workers favor (some) education investment, the political process sustains both types of expenditures. Changes in the demographic structure do not only change the equilibrium allocation conditional on policy, but also the balance of political power. With population "ageing," the politico-economic equilibrium therefore features increasingly large budget shares flowing into intergenerational transfers and a decline of the share of public education investment. These findings are very intuitive and, we believe, accord well with notions voiced in the public debate.

As a benchmark, we also characterize the allocation implemented by a Ramsey planner (with arbitrary geometric welfare weights). We find that the politico-economic equilibrium generically differs from the Ramsey allocation. However, this does not imply that it is generically possible to Pareto improve upon the allocation implemented in politico-economic equilibrium. We derive criteria to check for consumption and production inefficiency and conclude that the politico-economic equilibrium is consumption and production efficient for a set of parameters with positive measure.

To assess the quantitative implications of our model, we calibrate it to U.S. data. In politico-economic equilibrium, the calibrated model predicts labor income taxes of about 17 percent. The majority of these receipts funds transfers. The latter account for approximately 7 percent of GDP while public education investment accounts for about 5 percent of GDP. In contrast, the Ramsey policy (subject to the "natural" intergenerational welfare weights, corresponding to households' time discount factor multiplied by the population growth rate) calls for slightly higher taxes whose proceeds mostly fund public education (10 percent of GDP, in contrast to 4 percent of GDP for intergenerational transfers). Switching from the balanced growth path corresponding to the politico-economic equilibrium to the Ramsey growth path would give rise to an increase in the annual growth rate of about 0.3 percentage points. Nevertheless, as it turns out, the politico-economic equilibrium is consumption and production efficient.

Our work relates to the literatures on endogenous growth due to human capital accumulation, on the one hand, and productive government spending on the other (see Lucas (1988) or Barro (1990), respectively). Relative to these literatures, our model endogenizes various policy parameters that are of crucial importance for growth and welfare, but are taken as exogenous in these papers. Glomm and Ravikumar (1992) and Perotti (1993) introduce a political choice of public versus private education or income redistribution, respectively, in endogenous growth models with human capital accumulation. These authors find that the political choice interacts with the economic sphere. We consider a richer set of political instruments available to policy makers. Our paper therefore sheds light on both the *financing* and the *allocation* of government spending, as well as on the welfare and growth effects of these political choices.

Our work also relates to a more recent literature that analyzes the link between public investment and pensions. Boldrin and Montes (2005) analyze the link between public pensions and education from a normative perspective. These authors show that, to replicate the allocation in a hypothetical economy without borrowing constraints for private financing of education, the public sector has to provide for public education *and* pensions. Our interest here lies

[1]For a discussion of Markov perfect equilibrium see, for example, Krusell, Quadrini and Ríos-Rull (1997). See also Bhaskar (1998) who shows that weak informational constraints in overlapping generations games with a strictly dominant action for the old imply that the unique pure strategy equilibrium is in Markov strategies.

in understanding the factors that shape actual tax rates and budget shares. Bellettini and Berti Ceroni (1999) and Rangel (2003) show that, under certain conditions, societies may opt for sustaining public investments (e.g., into education) even if these investments benefit yet unborn cohorts whose interests are not fully represented in the political process. In these models, voters support government spending that is of no direct benefit to themselves because a trigger strategy links their choice of public investment to the choice of public pensions by future cohorts. Rather than emphasizing the link between pensions and investments that might arise due to a trigger strategy, our model focuses on the distributive *conflict* over the *size* of these two spending components.

The remainder of the paper is structured as follows. Section 2 describes the model and derives the allocation conditional on policy. Section 3 derives the politico-economic equilibrium as well as the Ramsey allocation. Section 4 compares the two allocations both qualitatively and quantitatively and analyzes the efficiency properties of the politico-economic equilibrium. Section 5 concludes.

2 Economic Environment

We consider an economy inhabited by three overlapping generations: students, workers, and retirees. Students accumulate human capital but do not consume nor work. Workers contribute with their acquired human capital to both production and the formation of new human capital, and save for retirement. Retirees do not work. Each cohort consists of a continuum of homogeneous agents. The gross population growth rate and thus, the ratio of workers to retirees (and students to workers) equals ν.

2.1 Technology

A continuum of competitive firms transform capital and labor into output by means of a Cobb-Douglas technology. Output per retiree in period t is given by

$$B_0 s_{t-1}^{\alpha}[H_t \nu(1-x_t)]^{1-\alpha},$$

where $B_0 > 0$ and the capital share $\alpha \in (0,1)$. Capital is owned by retirees and fully depreciates after one period. The capital stock per retiree, s_{t-1}, therefore equals the per-capita savings of workers in the previous period. Labor is supplied by current workers. Normalizing their time-endowment to unity and denoting workers' leisure consumption by x_t, labor supply per retiree equals $\nu(1-x_t)$. Workers' productivity is given by their human capital, H_t.

Production factors are rewarded their marginal products, due to competition among firms. The wage per unit of time, w_t, and the gross return on private capital, R_t, therefore satisfy

$$\begin{aligned} w_t &= (1-\alpha)B_0 H_t^{1-\alpha} s_{t-1}^{\alpha}[\nu(1-x_t)]^{-\alpha}, \\ R_t &= \alpha B_0 H_t^{1-\alpha} s_{t-1}^{\alpha-1}[\nu(1-x_t)]^{1-\alpha} = w_t \frac{\nu(1-x_t)}{s_{t-1}}\alpha' \end{aligned}$$

with $\alpha' \equiv \alpha/(1-\alpha)$.

Human capital reflects education investments during previous periods. More specifically, human capital growth is a function of the ratio of education investment and human capital:

$$H_{t+1} = B_1 H_t^{1-\delta} I_t^{\delta}$$

with $B_1 > 0$, $\delta \in (0,1)$, and I_t denoting educational investment per retiree. Since this condition specifies the law of motion for human capital per worker, the constant B_1 depends on ν. In fact, $B_1 \equiv \tilde{B}_1 \nu^{-1-\delta}$ for some fundamental constant $\tilde{B}_1$.

2.2 Government

The government taxes labor income in period t at rate $\tau_t + \sigma_t + \xi_t$ and capital income at rate $\eta_t + \theta_t$. Revenues collected from workers fund transfers to retirees (the component corresponding to τ_t), education investment (σ_t), as well as a lump-sum rebate to workers (ξ_t). The only role of ξ_t therefore is to distort labor supply. Revenues collected from retirees fund transfers to workers (the component corresponding to η_t) as well as education investment (θ_t). Denoting per-capita transfers to workers and retirees by a_t and b_t, respectively, we then have

$$\begin{aligned} a_t &= w_t(1-x_t)\xi_t + s_{t-1}R_t\eta_t/\nu = w_t(1-x_t)(\xi_t + \eta_t\alpha'), \\ b_t &= \nu w_t(1-x_t)\tau_t, \\ I_t &= \nu w_t(1-x_t)\sigma_t + s_{t-1}R_t\theta_t = \nu w_t(1-x_t)(\sigma_t + \theta_t\alpha'). \end{aligned}$$

Tax rates as well as a_t, b_t, and I_t must be non-negative since we exclude lump-sum taxes. These restrictions imply that the policy instruments have to satisfy the following conditions:

$$\xi_t + \eta_t\alpha' \geq 0, \quad \tau_t \geq 0, \quad \sigma_t + \theta_t\alpha' \geq 0, \quad \tau_t + \sigma_t + \xi_t \geq 0, \quad \eta_t + \theta_t \geq 0 \text{ for all } t. \tag{1}$$

We denote a combination of the five instruments in period t as $\bar{\kappa}_t$, $\bar{\kappa}_t \equiv (\tau_t, \sigma_t, \eta_t, \theta_t, \xi_t)$.

2.3 Preferences

As mentioned before, students do not work nor consume. Workers value consumption during working-age, c_1, and retirement, c_2, as well as leisure. They discount the future at factor $\beta \in (0,1)$. For analytical tractability, we assume that the period utility function of consumption is logarithmic. The indirect utility function of a worker in period t is then given by

$$\begin{aligned} \max_{s_t, x_t} \quad & \ln(c_{1,t}) + v(x_t) + \beta \ln(c_{2,t+1}) \\ \text{s.t.} \quad & c_{1,t} = w_t(1-x_t)(1-\tau_t-\sigma_t-\xi_t) + a_t - s_t, \\ & c_{2,t+1} = s_t R_{t+1}(1-\eta_{t+1}-\theta_{t+1}) + b_{t+1}. \end{aligned}$$

The felicity function of leisure is assumed to be increasing and concave.

The first-order conditions characterizing the households' savings and labor-supply decisions are standard. Conditional on factor prices, tax rates, and transfers, the marginal rate of substitution between current and future consumption is equalized with the corresponding marginal rate of transformation, the after-tax gross interest rate. Similarly, the marginal rate of substitution between consumption and leisure is equalized with the after-tax wage:

$$\begin{aligned} \frac{1}{c_{1,t}} &= \beta R_{t+1}(1-\eta_{t+1}-\theta_{t+1})\frac{1}{c_{2,t+1}}, \\ v'(x_t) &= w_t(1-\tau_t-\sigma_t-\xi_t)\frac{1}{c_{1,t}}. \end{aligned}$$

Conditional on given tax rates, the Euler equation characterizing the optimal savings choice of an *individual* household yields a closed-form solution for the *aggregate* savings function:[2]

$$s_t = z(\tau_{t+1}, \eta_{t+1}, \theta_{t+1}) w_t (1 - x_t)(1 - \tau_t - \sigma_t + \eta_t \alpha'),$$

where we define

$$z(\tau_{t+1}, \eta_{t+1}, \theta_{t+1}) \equiv \frac{\alpha\beta(1 - \eta_{t+1} - \theta_{t+1})}{\alpha(1+\beta)(1 - \eta_{t+1} - \theta_{t+1}) + (1-\alpha)\tau_{t+1}} \geq 0.$$

Note that, if future taxes are themselves functions of aggregate savings, the above relation is an implicit characterization of aggregate savings. A closed-form solution for the aggregate savings function is then in general not available. We will come back to this point in Section 3; for now, we assume future taxes to be independent of savings.

2.4 Economic Equilibrium

The fundamental state variables of the economy at time t are H_t and s_{t-1}. To simplify notation, however, we work with the state variables H_t and $q_t \equiv H_t^{1-\alpha} s_{t-1}^{\alpha}$. Substituting the expressions for wages and returns into the consumers' optimality conditions, the equilibrium allocation can recursively be expressed in terms of the following functions of policy instruments:

$$\left.\begin{array}{rcl} s_t &=& B_0(1-\alpha)\nu^{-\alpha}\; q_t\; (1-x_t)^{1-\alpha}\; (1 - \tau_t - \sigma_t + \eta_t\alpha') z(\tau_{t+1}, \eta_{t+1}, \theta_{t+1}), \\ c_{1,t} &=& B_0(1-\alpha)\nu^{-\alpha}\; q_t\; (1-x_t)^{1-\alpha}\; (1 - \tau_t - \sigma_t + \eta_t\alpha')(1 - z(\tau_{t+1}, \eta_{t+1}, \theta_{t+1})), \\ c_{2,t} &=& B_0\nu^{1-\alpha}\; q_t\; (1-x_t)^{1-\alpha}\; (\alpha(1-\eta_t-\theta_t) + (1-\alpha)\tau_t), \\ x_t &=& x(\tau_t, \sigma_t, \eta_t, \xi_t, \tau_{t+1}, \eta_{t+1}, \theta_{t+1}), \\ H_{t+1} &=& B_1\; H_t^{1-\delta}\; q_t^{\delta}\; (1-x_t)^{\delta(1-\alpha)}\; \left(B_0\nu^{1-\alpha}((1-\alpha)\sigma_t + \alpha\theta_t)\right)^{\delta}, \\ q_{t+1} &=& B_0^{\delta(1-\alpha)+\alpha} B_1^{1-\alpha} (1-\alpha)^{\alpha} \nu^{\delta(1-\alpha)^2 - \alpha^2}\; H_t^{(1-\delta)(1-\alpha)}\; q_t^{\delta(1-\alpha)+\alpha} \times \\ && (1-x_t)^{\delta(1-\alpha)^2 + \alpha(1-\alpha)} \times \\ && (1 - \tau_t - \sigma_t + \eta_t\alpha')^{\alpha} z(\tau_{t+1}, \eta_{t+1}, \theta_{t+1})^{\alpha} ((1-\alpha)\sigma_t + \alpha\theta_t)^{\delta(1-\alpha)}. \end{array}\right\} \quad (2)$$

Here, the function $x(\cdot)$ is implicitly defined by the (transformed) first-order condition characterizing labor supply,

$$v'(x_t)(1-x_t)(1 - z(\tau_{t+1}, \eta_{t+1}, \theta_{t+1})) = \frac{1 - \tau_t - \sigma_t - \xi_t}{1 - \tau_t - \sigma_t + \eta_t\alpha'}. \quad (3)$$

Note that labor supply in period t is independent of τ_t, σ_t, and θ_t if $\eta_t = \xi_t = 0$.

Conditional on initial values for the two state variables, (H_0, q_0), as well as a sequence of policy instruments, $\{\bar{\kappa}_t\}_{t=0}^{\infty}$, conditions (2) and (3) fully characterize the equilibrium allocation.

[2] To see this, note that the optimal savings choice of a worker is characterized (from the Euler equation above) by

$$s_t R_{t+1}(1 - \eta_{t+1} - \theta_{t+1}) + b_{t+1} = \beta R_{t+1}(1 - \eta_{t+1} - \theta_{t+1})[w_t(1-x_t)(1 - \tau_t - \sigma_t - \xi_t) + a_t - s_t].$$

Substituting for benefits and factor prices (and setting individual and average savings equal to each other), we arrive at

$$\begin{array}{rcl} s_t R_{t+1}(1 - \eta_{t+1} - \theta_{t+1}) + \nu w_{t+1}(1 - x_{t+1})\tau_{t+1} &=& \beta R_{t+1}(1 - \eta_{t+1} - \theta_{t+1})[w_t(1-x_t)(1 - \tau_t - \sigma_t + \eta_t\alpha') - s_t], \\ \Rightarrow (1 - \eta_{t+1} - \theta_{t+1}) + \tau_{t+1}/\alpha' &=& \beta(1 - \eta_{t+1} - \theta_{t+1})[w_t(1-x_t)(1 - \tau_t - \sigma_t + \eta_t\alpha')/s_t - 1]. \end{array}$$

The last equation yields a closed-form solution for the fixed-point problem, i.e., the aggregate savings function.

Taking logarithms of the laws of motion of the two state variables, we can express these two equations as

$$\begin{bmatrix} \ln(H_{t+1}) \\ \ln(q_{t+1}) \end{bmatrix} = \underbrace{\begin{bmatrix} 1-\delta & \delta \\ (1-\alpha)(1-\delta) & \alpha+\delta(1-\alpha) \end{bmatrix}}_{M} \begin{bmatrix} \ln(H_t) \\ \ln(q_t) \end{bmatrix} + \underbrace{\begin{bmatrix} f^H(\cdot) \\ f^q(\cdot) \end{bmatrix}}_{f_t} \tag{4}$$

where the definitions of $f^H(1-x_t(\cdot), \sigma_t, \theta_t)$ and $f^q(1-x_t(\cdot), \tau_t, \sigma_t, \eta_t, \theta_t, \tau_{t+1}, \eta_{t+1}, \theta_{t+1})$ follow from (the logarithms of) the laws of motion in (2).

In the special case of inelastic labor supply, $v'(x) = 0$, x_t equals some constant, and the equilibrium conditions (2) maintain their validity. Equation (3) and the choice of ξ_t, in contrast, become irrelevant (the latter subject to satisfying the constraint in (1)). Inspection of (2) and (3) reveals that the policy instrument θ_t is redundant:

Lemma 1. Consider a choice of policy instruments, $\bar{\kappa}_t = (\tau_t, \sigma_t, \eta_t, \theta_t, \xi_t)$, that satisfies (1). Fix the following period's policy instruments, $\bar{\kappa}_{t+1}$. Let $\mathcal{A}_t = (s_t, c_{1,t}, c_{2,t}, x_t, H_{t+1}, q_{t+1})$ be the equilibrium outcome implied by the initial condition (H_t, q_t), the policy instruments $\bar{\kappa}_t$ and $\bar{\kappa}_{t+1}$, as well as conditions (2), (3). Then, holding (H_t, q_t) and $\bar{\kappa}_{t+1}$ fixed, the same $\mathcal{A}_t$ is implied by a different choice of policy instruments, namely $\bar{\kappa}'_t = (\tau_t, \sigma_t + \theta_t\alpha', \eta_t + \theta_t, 0, \xi_t - \theta_t\alpha')$, where $\bar{\kappa}'_t$ also satisfies (1). In the special case of inelastic labor supply, without loss of generality, ξ_t can be normalized to zero as well.

Intuitively, lower retiree contributions to education can fully be replicated by higher worker contributions to education in combination with higher transfers from retirees to workers and a lower ξ_t (this last component to ensure that the choice of leisure remains unaffected). Adopting the normalization $\theta_t = 0$ for all t, the set of non-redundant policy instruments in period t, κ_t, is then given by $(\tau_t, \sigma_t, \eta_t, \xi_t)$. In the special case of inelastic labor supply, the tax rate ξ_t has no effect on the allocation and κ_t reduces to $(\tau_t, \sigma_t, \eta_t)$.

In the following, for brevity, we write $z(\tau, \eta)$ to mean $z(\tau, \eta, 0)$.

2.5 Balanced Growth Path

On a balanced growth path, all policy instruments are constant over time, implying that per-capita hours worked are time-invariant as well. From (2), the growth rates of s_t, $c_{1,t}$, and $c_{2,t}$ then are equal to the growth rate of q_t. Moreover, the laws of motion for the two state variables in (2) imply that the gross growth rate of H_t, γ_H, must equal the gross growth rate of q_t on a balanced growth path. For any time-invariant choice of instruments, the last two equations in (2) therefore pin down the ratio H_t/q_t on the corresponding balanced growth path. Given this ratio, the same two conditions pin down γ_H and thus, the balanced growth rates of $q_t, s_t, c_{1,t}$, and $c_{2,t}$. Following this logic, we find

$$\begin{aligned} \gamma_H \;=\; & \Big(B_0^\delta B_1^{1-\alpha}(1-\alpha)^{\alpha\delta}\nu^{\delta(1-2\alpha)}(1-x)^{\delta(1-\alpha)}(1-\tau-\sigma+\eta\alpha')^{\alpha\delta} \times \\ & \quad z(\tau,\eta)^{\alpha\delta}((1-\alpha)\sigma)^{\delta(1-\alpha)}\Big)^{\frac{1}{1-\alpha(1-\delta)}} \text{ s.t. (3).} \end{aligned}$$

As the equation makes clear, labor income taxes depress growth because they lower disposable income of workers (the effect captured by the expression $1-\tau-\sigma$), as do expected future pension benefits because they lower the savings rate ($z(\tau, \eta)$ is decreasing in its first argument).

At the same time, education investment fosters human capital accumulation and thus, growth (the effect captured by σ in the last term). Capital income taxes have an ambiguous effect on growth because they increase disposable income of workers but have a negative effect on their savings rate.

From (2), savings along the balanced growth path satisfies $s_t = B_0(1-\alpha)\nu^{-\alpha}(1-x)^{1-\alpha}(1-\tau-\sigma+\eta\alpha')\ z(\tau,\eta)H_t^{1-\alpha}s_{t-1}^{\alpha}$. Along the balanced growth path, where s_t grows at gross rate γ_H, we therefore have

$$\begin{aligned}\left(\frac{H_t}{s_{t-1}}\right)^{1-\alpha} &= \frac{\gamma_H}{B_0(1-\alpha)\nu^{-\alpha}(1-x)^{1-\alpha}(1-\tau-\sigma+\eta\alpha')\ z(\tau,\eta)} \text{ s.t. (3)},\\ R &= \frac{\alpha\gamma_H\nu^{1-\alpha}}{(1-\alpha)\nu^{-\alpha}(1-\tau-\sigma+\eta\alpha')\ z(\tau,\eta)} \text{ s.t. (3)}.\end{aligned}$$

We will use these relations for calibration purposes.

3 Equilibrium

In the previous section, we have characterized economic equilibrium, conditional on sequences for tax rates and government spending. We now turn to the question of how taxes and government spending are determined, and we simultaneously solve for policy and the equilibrium allocation. We start by considering the politico-economic equilibrium where policy choices reflect the aggregation of voters' preferences in the political process. Thereafter, we consider the allocation implemented by a Ramsey government that attaches weight to the welfare of current and future cohorts. The Ramsey allocation will serve as a benchmark to compare the politico-economic equilibrium with.

3.1 Politico-Economic Equilibrium

We assume that retirees, workers and students vote on candidates representing platforms with values for the policy instruments κ_t.[3] Voters support a candidate not only for her policy platform, but also for other characteristics like "ideology" that are orthogonal to the fundamental policy dimensions of interest. These characteristics are permanent and cannot be credibly altered in the course of electoral competition; their valuation differs across voters (even if voters agree about the preferred policy platform) and is subject to random aggregate shocks, realized after candidates have chosen their platforms. This "probabilistic-voting" setup renders the probability of winning a voter's support a continuous function of the competing policy platforms. Related, it implies that equilibrium policy platforms smoothly respond to changes in the demographic structure. This stands in contrast to the "median-voter" setup where an infinitesimal change in the population growth rate may have implausibly large effects on policy outcomes if the change alters the cohort the median voter is associated with.

In a Nash equilibrium with two candidates maximizing their expected vote share, both candidates propose the same policy platform.[4] This platform maximizes a convex combination of the objective functions of all groups of voters, where the weights reflect the groups' size and sensitivity of voting behavior to policy changes. Groups that care a lot about policy platforms

[3] By setting parameters reflecting the political influence of certain groups equal to zero, it is straightforward to analyze the case where only a subset of the three generations alive in a given period participate in the vote; see below.

[4] See Lindbeck and Weibull (1987) and Persson and Tabellini (2000) for discussions of probabilistic voting.

relative to the candidates' other characteristics have more political influence as they are more likely to shift their support from one candidate to the other in response to small changes in the proposed platform. In equilibrium, these groups of "swing voters" thus tilt policy in their own favor. If all voters are equally responsive to changes in the policy platform, electoral competition implements the utilitarian optimum with respect to voters. In the context of our model, the probabilistic-voting assumption implies that the welfare of all currently living cohorts receives some weight in the objective function maximized by the political process. In particular, the welfare of retirees receives some weight although they are outnumbered by workers and students (unless ν is very small). We consider this implication to be very realistic; after all, old voters appear to exert as strong political influence per capita as younger voters when the salient issue of intergenerational transfers is concerned (see, for example, Dixit and Londregan (1996, p. 1144) and Grossman and Helpman (1998, p. 1309)).

Owing to political competition at the beginning of each period, policy makers cannot commit to future policy platforms. Voters therefore have to form expectations about the effect of their choices on future policy outcomes. Under our assumption of Markovian equilibrium, future leisure and policy choices are functions of the fundamental state variables only, $x_{t+1} = \tilde{x}(H_{t+1}, q_{t+1})$ and $\kappa_{t+1} = \tilde{\kappa}(H_{t+1}, q_{t+1})$. If these policy functions are independent of the state variables, $\kappa_{t+s} = \tilde{\kappa},\ s \geq 1$, then (3) implies that the leisure function is independent of the state variables as well, $x_{t+s} = \tilde{x},\ s \geq 1$, and the aggregate savings function as well as the economic equilibrium conditions (2) apply (see the discussion at the end of Subsection 2.3).

In the following, we conjecture that the policy functions indeed are independent of the state variables such that (2) and (3) apply. We will derive the choice of contemporaneous policy instruments under this conjecture and verify that there exists an equilibrium with constant tax rates.

Letting ω and ψ denote the per-capita political influence of retirees and students, respectively, relative to workers, the program characterizing equilibrium policy choices in period t is given by

$$\max_{\kappa_t} W(H_t, q_t, \kappa_t; \tilde{\kappa}, \tilde{x}) \quad \text{s.t. (1)},$$

where

$$\begin{aligned} W(H_t, q_t, \kappa_t; \tilde{\kappa}, \tilde{x}) \quad &\equiv \quad \omega \ln(c_{2,t}) + \nu[\ln(c_{1,t}) + v(x_t) + \beta \ln(c_{2,t+1})] \\ &\qquad + \psi\nu^2\beta[\ln(c_{1,t+1}) + v(x_{t+1}) + \beta \ln(c_{2,t+2})] \\ \text{s.t.} \quad &\qquad (2), (3),\ H_t, q_t, \tilde{\kappa} \text{ and } \tilde{x} \text{ given.} \end{aligned}$$

Political equilibrium requires that for any combination of state variables (H_t, q_t), the κ_t solving this program is given by $\tilde{\kappa}$.

Using the equilibrium expressions for consumption from (2), the objective function can be expressed as

$$\begin{aligned} W(\cdot) \quad = \quad & \omega \ln[(1-x_t)^{1-\alpha}(\alpha(1-\eta_t) + (1-\alpha)\tau_t)] + \nu\{\ln[(1-x_t)^{1-\alpha}(1-\tau_t-\sigma_t+\eta_t\alpha')] + \\ & \quad v(x_t) + \beta \ln[(1-x_t)^{(1-\alpha)(\delta(1-\alpha)+\alpha)}(1-\tau_t-\sigma_t+\eta_t\alpha')^{\alpha}\sigma_t^{\delta(1-\alpha)}]\} \\ & + \psi\nu^2\beta\{\ln[(1-x_t)^{(1-\alpha)(\delta(1-\alpha)+\alpha)}(1-\tau_t-\sigma_t+\eta_t\alpha')^{\alpha}\sigma_t^{\delta(1-\alpha)}] + \\ & \quad \beta \ln[(1-x_t)^{(1-\alpha)(\delta(1-\delta)(1-\alpha)+(\delta(1-\alpha)+\alpha)^2)}(1-\tau_t-\sigma_t+\eta_t\alpha')^{\alpha(\delta(1-\alpha)+\alpha)} \cdot \\ & \qquad \sigma_t^{\delta(1-\alpha)(1-\delta+\delta(1-\alpha)+\alpha)}]\} \\ & + \text{t.i.p.} \quad \text{s.t. (3)}, \end{aligned}$$

where t.i.p. denotes terms that are unaffected by contemporaneous policy choices (under the conjecture). We first consider the case with inelastic labor supply.

3.1.1 Inelastic Labor Supply

If labor supply is inelastic, then $\xi_t = 0$ for all t and x_t is fixed such that

$$\begin{aligned} W(\cdot) \;\simeq\; & \omega \ln[\alpha(1-\eta_t) + (1-\alpha)\tau_t] + \\ & \nu\{\ln(1-\tau_t-\sigma_t+\eta_t\alpha') + \beta\ln[(1-\tau_t-\sigma_t+\eta_t\alpha')^{\alpha}\sigma_t^{\delta(1-\alpha)}]\} + \\ & \psi\nu^2\beta\{\ln[(1-\tau_t-\sigma_t+\eta_t\alpha')^{\alpha}\sigma_t^{\delta(1-\alpha)}] + \\ & \quad \beta\ln[(1-\tau_t-\sigma_t+\eta_t\alpha')^{\alpha(\delta(1-\alpha)+\alpha)}\sigma_t^{\delta(1-\alpha)(1-\delta+\delta(1-\alpha)+\alpha)}]\}. \end{aligned}$$

Disregarding the inequality constraints in (1), the effects of marginal policy changes are linearly dependent[5]. We focus on the equilibrium values for τ_t and σ_t, for some given choice of η_t.

Consider first the first-order condition with respect to σ_t:

$$\frac{1+\alpha\beta+\psi\nu\alpha\beta[1+\beta(\delta(1-\alpha)+\alpha)]}{1-\tau_t-\sigma_t+\eta_t\alpha'} = \beta\delta(1-\alpha)\frac{1+\psi\nu[1+\beta(1-\delta+\delta(1-\alpha)+\alpha)]}{\sigma_t}.$$

The three terms on the left-hand side of this condition reflect marginal welfare losses for workers and students. They measure, respectively, the marginal welfare loss for workers due to the tax induced reduction in disposable income and thus, first-period consumption; the marginal welfare loss for workers due to the induced reduction in second-period consumption, mediated through reduced capital accumulation (thus the influence of the capital share α) and discounted to the current period (thus the influence of β); and the marginal welfare loss for students due to reduced capital accumulation by workers (the effect on first-period consumption of students) and by themselves (the effect on second-period consumption). The two terms on the right-hand side reflect the marginal welfare gain for workers of higher disposable income in the subsequent period, due to the productivity enhancing effect of education spending, and the marginal welfare gain for students. The latter gain is comprised of two parts: Students do not only gain because education spending increases their disposable income and thus, first- and second-period consumption, but also due to the dynamic human capital externality that increases productivity and thus, capital income at the time of current students' retirement.

Conditional on η_t, the above first-order condition prescribes that σ_t and τ_t are negatively related. Intuitively, a higher value of τ_t or σ_t reduces disposable incomes; this makes it more costly to tax workers and therefore calls for a reduction of the other tax rate.

Consider next the choice of τ_t. Since the *costs* of an increase in τ_t are the same as those of an increase in σ_t (all terms in the objective function featuring $-\tau_t$ also feature $-\sigma_t$), the marginal *benefits* of increases in τ_t and σ_t must match each other in equilibrium. Disregarding the inequality constraints in (1), this implies

$$\frac{\omega}{\nu}\frac{1-\alpha}{\alpha(1-\eta_t)+(1-\alpha)\tau_t} = \beta\delta(1-\alpha)\frac{1+\psi\nu[1+\beta(1-\delta+\delta(1-\alpha)+\alpha)]}{\sigma_t},$$

where the left-hand side reflects increased consumption of retirees due to higher social security benefits. The term $(1-\alpha)$ in the numerator of the left-hand side accounts for the fact that

[5] In particular, $\partial W/\partial\eta_t = -\alpha'\partial W/\partial\tau_t$.

social security contributions comprise only part of retirees' consumption, and are proportional to wages; the ratio ω/ν reflects the weight attached to the welfare of retirees relative to workers.

Conditional on η_t, this second first-order condition defines a positive relationship between τ_t and σ_t. Intuitively, both social security benefits (higher τ_t) and public investment expenditures (higher σ_t) have decreasing marginal benefits, due to the concavity of the beneficiaries' utility function. To maintain equality of the two marginal benefits, an increase in τ_t therefore must be associated with an increase in σ_t. Note that, ceteris paribus, an increase in ω raises τ_t while an increase in ν or an increase in ψ raises σ_t. This reflects the effects of varying political influence, as captured by the probabilistic voting assumption; both higher per-capita weight in the political process and larger size of a constituency shift the political outcome in favor of that particular group.

Disregarding the inequality constraints in (1) and solving the two first-order conditions, we find

$$\begin{aligned}
\tau_t(\eta_t) &= \frac{\omega - \alpha(1-\eta_t)\left[\omega + \nu\left(1 + \beta\nu\psi[(1+\beta)\delta + \alpha^2\beta(1-\delta)^2] + \alpha\beta(1+(1+\beta\delta)\nu\psi)(1-\delta) + \beta\delta\right)\right]}{(1-\alpha)\left(\beta\psi\nu^2\left(\alpha(1-\delta)(1+\beta\delta+\alpha\beta(1-\delta)) + (1+\beta)\delta\right) + (1+\alpha\beta(1-\delta)+\beta\delta)\nu + \omega\right)}, \\
& \quad \tau_t'(\eta_t) > 0, \\
\sigma_t &= \frac{\beta\delta\nu(1+(1+\beta(1+\alpha(1-\delta)))\nu\psi)}{\beta\psi\nu^2\left(\alpha(1-\delta)(1+\beta\delta+\alpha\beta(1-\delta)) + (1+\beta)\delta\right) + (1+\alpha\beta(1-\delta)+\beta\delta)\nu + \omega} \equiv \sigma^W > 0.
\end{aligned}$$

Conditional on η_t, an increase in β, ν and ψ or a decrease in ω reduces τ_t and raises σ_t.

For any parameter constellation, there exists a continuum of policies $(\tau_t, \sigma_t, \eta_t)$ satisfying both the inequality constraints (1) and the above relations. This can be seen by noting that (i) σ_t is always positive and (ii) raising η_t increases τ_t and thus, allows to satisfy all (non-negativity) constraints in (1). We eliminate the ensuing policy indeterminacy by focusing on the policy that minimizes the capital income tax rate, η_t. Two observations motivate this refinement. First, the economy's growth rate is maximized if the capital income tax rate is minimal. Second, the number of instruments used to implement the equilibrium allocation is minimized if the political process sets either η_t or τ_t to zero. (The former is the case when parameters are such that $\tau_t(0) \geq 0$, the latter holds otherwise.) Under our "refinement," the equilibrium tax rates are thus given by

$$\kappa^W = (\tau^W, \sigma^W, \eta^W) \text{ where } \tau^W = \max[0, \tau_t(0)] \text{ and } \eta^W = \max[0, \eta_t(0)]$$

with $\eta_t(\cdot)$ denoting the inverse of the function $\tau_t(\eta_t)$. Note that the equilibrium tax functions are indeed independent of the state variables, as conjectured. Moreover, the refined equilibrium tax rates are unique in the limit of the finite horizon economy.[6]

[6] To see this, note that the consumption of workers and retirees in the final period T is given by

$$\begin{aligned}
c_{1,T} &= w_T(1-\tau_T) + \frac{s_{T-1}R_T\eta_T}{\nu} = q_T B_0(1-\alpha)\nu^{-\alpha}(1-\tau_T+\eta_T\alpha'), \\
c_{2,T} &= q_T B_0 \nu^{1-\alpha}(\alpha(1-\eta_T) + (1-\alpha)\tau_T),
\end{aligned}$$

respectively. Tax rates are set to achieve

$$\omega = \frac{c_{2,T}}{c_{1,T}} \Rightarrow \nu[\alpha(1-\eta_T) + (1-\alpha)\tau_T] = \omega(1-\alpha)(1-\tau_T+\eta_T\alpha').$$

Under the refinement, this equation has a unique solution with either τ_T or η_T equal to zero and the other tax rate weakly positive; moreover, $\sigma_T = 0$. All three tax rates are independent of the state variables. Moving to

3.1.2 Elastic Labor Supply

With elastic labor supply, x_t generally depends on the choice of contemporaneous and future policy instruments. Recall from (3), however, that x_t is independent of τ_t and σ_t if $\xi_t = \eta_t = 0$. In this case, and under the maintained conjecture of constant policy functions, labor supply is neither directly nor indirectly affected by the choice of τ_t or σ_t, implying that the same objective function as in the case with inelastic labor supply applies. As it turns out, the equilibrium levels of ξ_t and η_t indeed equal zero under plausible conditions. In our simulations that are calibrated to match features of the U.S. economy, these conditions are always met.[7] We conclude that, with elastic labor supply, the same tax rates as in the case with inelastic labor supply, κ^W and $\xi^W = 0$, are implemented in politico-economic equilibrium. In contrast with the case of inelastic labor supply, the tax rate η_t now is in a corner and the equilibrium tax rates are no longer indeterminate. This provides another rationale for the refinement (normalizing η_t to zero) that we adopted in the case with inelastic labor supply.

3.2 Ramsey Allocation

The Ramsey program differs twofold from the program solved by political decision makers: It endows the planner with a commitment technology; and it attaches weight to the welfare of currently living *and* yet unborn generations. Denoting the intergenerational discount factor of the planner by ρ, $0 < \rho < 1$, the Ramsey policy solves the following program:[8]

$$\max_{\{\kappa_s\}_{s=t}^{\infty}} G(H_t, q_t, \{\kappa_s\}_{s=t}^{\infty}) \quad \text{s.t. (1)},$$

period $T-1$, the objective function is

$$\begin{aligned} W^{T-1}(\cdot) \quad \simeq \quad & \omega \ln[\alpha(1-\eta_{T-1}) + (1-\alpha)\tau_{T-1}] + \\ & \nu\{\ln(1-\tau_{T-1}-\sigma_{T-1}+\eta_{T-1}\alpha') + \beta\ln[(1-\tau_{T-1}-\sigma_{T-1}+\eta_{T-1}\alpha')^{\alpha}\sigma_{T-1}^{\delta(1-\alpha)}]\} + \\ & \psi\nu^2\beta\{\ln[(1-\tau_{T-1}-\sigma_{T-1}+\eta_{T-1}\alpha')^{\alpha}\sigma_{T-1}^{\delta(1-\alpha)}]. \end{aligned}$$

Constant policy functions in period T therefore imply unique, constant policy functions for σ_{T-1}, τ_{T-1}, and η_{T-1}. Similarly, constancy of the policy functions in period $T-1$ implies that the period-$(T-2)$ policy functions are independent of the state variables as well. Moreover, $W^{T-2}(\cdot) \simeq W(\cdot)$ such that $\kappa_{T-2}(H_{T-2}, q_{T-2}) = \kappa^W$. The result then follows by induction.

[7] Let $\chi \equiv (1-\alpha)(\omega + \nu + \delta(1-\delta)(1-\alpha)\nu^2\beta^2\psi + (\delta(1-\alpha)+\alpha)(\nu\beta + \nu^2\beta\psi + \nu^2\beta^2\psi(\delta(1-\alpha)+\alpha))) > 0$. Omitting time subscripts and using the equilibrium condition (3), we have

$$\begin{aligned} \frac{dW(\cdot)}{d\xi} &= \left[-\chi\frac{1}{1-x} + \nu v'(x)\right]\frac{dx}{d\xi} = \left[\left(\frac{1-\tau-\sigma-\xi}{1-\tau-\sigma+\eta\alpha'}\right)\frac{\nu}{1-z(\tau,\eta)} - \chi\right]\frac{1}{1-x}\frac{dx}{d\xi} \\ &\le \left[\frac{\nu}{1-z(\tau,\eta)} - \chi\right]\frac{1}{1-x}\frac{dx}{d\xi} \quad \text{s.t. (3)}, \end{aligned}$$

where we use the fact that $dx/d\xi \ge 0$. Similarly,

$$\frac{dW(\cdot)}{d\eta} \le \text{t.i.l.s.} + \left[\frac{\nu}{1-z(\tau,\eta)} - \chi\right]\frac{1}{1-x}\frac{dx}{d\eta} \quad \text{s.t. (3)}$$

with $dx/d\eta \ge 0$ and t.i.l.s. denoting those terms that enter the first-order condition when labor is supplied inelastically. In our simulations, the expression in square brackets is negative, implying (from (1)) that $\xi^W \le 0$. Moreover, since τ^W is strictly positive in the simulations, it follows that $\eta^W = 0$ and thus, from (1), that $\xi^W = 0$.

[8] Since the Ramsey planner can commit, "non-geometric" welfare weights would imply that the economy embarks on an unbalanced growth path. We dismiss this possibility.

where

$$G(H_t, q_t, \{\kappa_s\}_{s=t}^{\infty}) \equiv \sum_{s=t}^{\infty} \rho^{s-t}(\beta \ln(c_{2,s}) + \rho \ln(c_{1,s}) + \rho v(x_s))$$
$$\text{s.t.} \quad (2), (3) \text{ for all } s \geq t,\ H_t \text{ and } q_t \text{ given.}$$

The first-order conditions to this program are derived in Appendix A.1. We directly turn to a discussion of the findings, starting with the case of inelastic labor supply.

3.2.1 Inelastic Labor Supply

With inelastic labor supply, the Ramsey program simplifies as labor supply is unaffected by policy changes and $\xi_i = 0$ for all $i \geq t$. Similarly to the politico-economic equilibrium, the effects of policy changes *in period* t are linearly dependent.[9] As before, we therefore focus on the optimality conditions for τ_t and σ_t, conditional on some given choice of η_t; for the time being, we disregard the inequality constraints (1). The two first-order conditions (derived in Appendix A.1) can be rearranged as

$$\frac{1 + \Omega_{22}\alpha}{1 - \tau_t - \sigma_t + \eta_t \alpha'} = \frac{\delta(\Omega_{21} + \Omega_{22}(1 - \alpha))}{\sigma_t},$$
$$\frac{\beta}{\rho} \frac{1 - \alpha}{\alpha(1 - \eta_t) + (1 - \alpha)\tau_t} = \frac{\delta(\Omega_{21} + \Omega_{22}(1 - \alpha))}{\sigma_t}.$$

In parallel with the politico-economic equilibrium conditions discussed earlier, they reflect equality between the costs and benefits of a marginal increase in σ_t, and equality of the benefits of marginal increases in τ_t and σ_t, respectively. The terms $\Omega_{21} \equiv (\beta + \rho) \left[(I - \rho M)^{-1}\right]_{[2,1]}$ and $\Omega_{22} \equiv (\beta + \rho) \left[(I - \rho M)^{-1}\right]_{[2,2]}$ measure the present discounted contribution to the planner's objective of an increase in the (logarithms of the) economy's stock of human capital and q, respectively.

Consider the first equation. There are two, intuitive differences between this condition and the corresponding condition that holds in politico-economic equilibrium: First, on the left- and right-hand side, the term Ω_{22} replaces the expression $\beta(1 + \psi\nu[1 + \beta(\delta(1 - \alpha) + \alpha)])$. This is due to the fact that the planner internalizes the effect of capital accumulation on all future cohorts rather than only tomorrow's retirees and workers. Second, on the right-hand side, the term Ω_{21} replaces the expression $\psi\nu\beta^2(1 - \alpha)(1 - \delta)$, again reflecting the Ramsey planner's concern for all future cohorts as higher investment in education affects H_{t+s} and $q_{t+s},\ s \geq 1$. On the left-hand side of the second equation, the only difference to the politico-economic equilibrium is that the *political* weight of retirees relative to workers, ω/ν, is replaced by their relative weight as attached by the *planner*, β/ρ.

Holding ρ constant, changes in the population growth rate leave the Ramsey tax rates unaffected, in contrast to the situation in politico-economic equilibrium where changes in ν affect the relative weight attached to the two constituencies and thus, equilibrium tax rates.[10] Plausibly, however, the planner's intergenerational welfare weights should not be held constant if ν changes. Indeed, one might argue that $\rho = \beta\nu$ constitutes a plausible benchmark because under this assumption, the planner respects households' time preference and accounts for the size of

[9] In particular, $\partial G(\cdot)/\partial \eta_t = -\alpha' \partial G(\cdot)/\partial \tau_t$.

[10] Both Ω_{21} and Ω_{22} are independent of ν.

cohorts. If ρ is linked to ν in this way, then changes in the population growth rate do affect the Ramsey tax rates.

Disregarding the inequality constraints (1) and solving the first-order conditions conditional on some given value for η_t yields the following expressions for the optimal tax rates:

$$\tau_t(\eta_t) = \frac{(\alpha+\beta)\rho - \alpha\rho^2(1-\delta)(1+\beta) - \beta(1-\alpha) + \alpha\eta_t(\beta+\rho)(\rho(1-\delta)-1)}{(1-\alpha)(\beta+\rho)(\rho(1-\delta)-1)},\ \tau_t'(\eta_t) > 0,$$

$$\sigma_t = \frac{\delta\rho}{1-\rho(1-\delta)} \equiv \sigma^G > 0.$$

Conditional on η_t, and evaluated at $\rho \equiv \beta\nu$, an increase in β or δ raises σ_t and lowers τ_t; changes in α have no effect on σ_t, but an ambiguous effect on τ_t.

As in the politico-economic equilibrium, a continuum of policies $(\tau_t, \sigma_t, \eta_t)$ satisfies both the inequality constraints (1) and the above relations. Again, we eliminate policy indeterminacy by focusing on the feasible policy with the smallest capital income tax rate, η_t. The Ramsey tax rates are thus given by

$$\kappa_t^G = (\tau^G, \sigma^G, \eta^G) \ \text{ where } \tau^G = \max[0, \tau_t(0)] \ \text{ and } \eta^G = \max[0, \eta_t(0)]$$

with $\eta_t(\cdot)$ denoting the inverse of the function $\tau_t(\eta_t)$.

Turning to the Ramsey planner's choice of κ_i^G, $i > t$, both τ_i and η_i affect the savings decision in period $i-1$ (see the program in Appendix A.1). As a consequence, no indeterminacy with respect to future policy instruments arises. In fact, as it turns out, the Ramsey tax rates for all periods $i > t$ are identical to the (refined) Ramsey tax rates in period t:

$$\kappa_i^G = \kappa_t^G \text{ for all } i > t.$$

In other words, the Ramsey policy is time consistent. Moreover, as long as $\eta_t = 0$, the implementability constraints in the Ramsey program do not bind and the Ramsey allocation therefore coincides with the social planner allocation. Appendix A.1 explains the reasons for these results.

3.2.2 Elastic Labor Supply

With elastic labor supply, changes in tax rates work through four additional channels: In period t, they (i) affect production as well as the utility from leisure in period t and (ii) have indirect effects through induced changes in the state variables in periods $t+1$ and later. In periods $i > t$, they (iii) also affect production and the utility from leisure in period $i-1$ and (iv) have indirect effects through induced changes in the state variables in periods i and later.

The resulting system of optimality conditions does not always allow for closed form solutions. Suppose that the Ramsey tax rates satisfy $\eta_i = \xi_i = 0$ for all $i \geq t$. In this case, $\partial x_t/\partial\tau_t = \partial x_t/\partial\sigma_t = 0$ (see (3)), implying that the first-order conditions for τ_t and σ_t are unchanged relative to the case with inelastic labor supply. Since $\partial z(\tau,\eta)/\partial\tau < 0$, an increase in $\tau_i, i > t$, depresses labor supply in period $i-1$ (see (3)). Relative to the case with inelastic labor supply, this introduces additional, negative terms in the first-order conditions with respect to $\tau_i, i > t$. When τ_t is strictly positive, we therefore have $\tau_t > \tau_i$. With $\tau_t \neq \tau_i$, σ_t differs from σ_i as well. Furthermore, if $\tau_i = 0$, a closed-form solution for σ_i results.[11] Similar to the situation in politico-economic equilibrium, with elastic labor supply, the optimal tax rates in the initial period are determinate, due to the presence of the marginal effects on labor supply.

[11]Under the parameter values calibrated to match features of the U.S. economy, we find that all tax rates τ and σ are interior while η and ξ are in a corner.

4 Analysis

In the previous section, we have characterized the politico-economic equilibrium and the Ramsey allocation. We now analyze in more detail how these two equilibria differ. Two sets of questions guide our analysis. The first set of questions mainly has a normative character. We are interested in identifying conditions under which the allocation in politico-economic equilibrium differs from the Ramsey allocation, and we want to understand whether the politico-economic equilibrium is consumption and production efficient. In both cases, the potential distortions arising from political preference aggregation are of central interest to us. To isolate these potential distortions, we abstract from other sources of distortions as far as possible. For that reason, we consider the case with inelastic labor supply.

The second set of questions is more positive in nature. It relates to the quantitative implications of the political choice of tax rates (relative to the Ramsey policy), and to the effect of demographic change on equilibrium outcomes. To address this second set of questions, we calibrate the model to match features of the U.S. economy. Based on this calibration and a series of robustness checks, we compute the growth differential between the politico-economic equilibrium and the Ramsey allocation. Moreover, we analyze how changes in the demographic structure affect taxes, spending patterns, and growth in politico-economic equilibrium. We consider both the case with and without tax distortions on labor supply.

4.1 Equivalence of Allocations

The equilibrium conditions in (2) imply that two allocations necessarily differ unless they are supported by the same tax rates $(\tau_t, \sigma_t, \eta_t)$ for all t. Using this fact, it is straightforward to derive parameter conditions subject to which the politico-economic equilibrium allocation coincides with the allocation implemented by the Ramsey planner. Solving the system of equations $\kappa^W = \kappa^G$ for ω and ψ yields a unique tuple

$$\hat{\omega} = \frac{\beta\nu(1-\rho)(1-\alpha\rho(1-\delta))}{\rho(1-\rho(1+\alpha(1-\delta)))}, \quad \hat{\psi} = \frac{\rho}{\beta\nu}\frac{1}{1-\rho(1+\alpha(1-\delta))}.$$

(Of course, the first-order conditions of the political decision makers on the one hand and the Ramsey planner on the other hand coincide for this tuple of political weights.) The parameter set guaranteeing equivalence of the two allocations therefore is of measure zero. Moreover, the tuple $(\hat{\omega}, \hat{\psi})$ may not be positive.

4.2 Consumption and Production Efficiency

We have found that the politico-economic equilibrium allocation and the Ramsey allocation generically differ. This finding is not very surprising, given that a priori, the Ramsey planner's welfare weights need not bear any particular relation to the per-capita political influence of voters. A more interesting question is whether the politico-economic equilibrium is consumption and production efficient, i.e., whether it does not allow for a change in consumption or investment patterns that leaves all current and future cohorts better off. We now turn to this question.

Assessing consumption efficiency of the equilibrium allocation is straightforward. The politico-economic equilibrium is consumption inefficient if reallocating consumption from workers to retirees leads to a Pareto improvement. Clearly, retirees in the initial cohort benefit from such a consumption reallocation. Members of some later cohort $i, i > t$, benefit as well if

$\beta\gamma_H \nu u'(c_{2,i+1}) \geq u'(c_{1,i})$. Using the households' consumption Euler equation, the condition for *consumption efficiency* can thus be expressed, as is standard, as

$$1 < \frac{R}{\gamma_H \nu}.$$

To assess production efficiency, we consider a sequence of changes in human and physical capital investment that leaves total investment in each period unchanged. We then check whether such a reallocation of investment spending weakly increases output in all future periods. If this is the case, the initial allocation is production inefficient (see Cass, 1972). In Appendix A.2, we derive as a criterion for *production efficiency* the requirement that

$$\frac{\delta}{\sigma} < \frac{R}{\gamma_H \nu} < \frac{1}{\sigma}.$$

4.3 Quantitative Analysis

We now turn to a quantitative assessment of the model. We calibrate the model to match stylized features of the U.S. economy, compute the implied equilibrium tax rates in politico-economic equilibrium and under the Ramsey policy, and analyze the quantitative effects of changes in key parameters. We take a period in the model to correspond to 25 years in the data.

From U.S. Census Bureau data for 1970 and 2000, we derive a gross population growth rate of $\nu = 1.3113 \approx 1.3843^{\frac{25}{30}}$. Using NIPA data, Piketty and Saez (2003) compute a time series for the capital share in post-war U.S. data. We use the average of that series for the period 1970–2003: $\alpha = 0.2815$. We assume that the per-capita political influence of students, workers and retirees is the same, $\omega = \psi = 1$, and we normalize $B_0 = 1$.[12]

In the model, the elasticity of earnings with respect to education spending equals δ (holding aggregate human capital and labor supply fixed).[13] Card and Krueger (1995) and Betts (1996) report various estimates of this elasticity, ranging from close to zero up to 0.55. We use $\delta = 0.12$, the median of the more recent estimates reported by both Card and Krueger (1995) and Betts (1996) as our baseline value, but also consider $\delta = 0.155$, the value Card and Krueger (1995) infer from their own earlier work.

To calibrate β and B_1, we use the equilibrium relations characterizing R and γ_H (see page 7), evaluated at κ^W. As target values for these two variables, we use $R^{\text{data}} = 1.0483^{25}$ and $\gamma_H^{\text{data}} = 1.0126^{25}$.[14] This implies $\beta = 0.4340 \approx 0.9672^{25}, B_1 = 2.4174$. When we consider the

[12] This normalization amounts to a choice of scaling factor. The equilibrium relations that we use to calibrate the model feature B_0 only through the expression $B_0^{\delta} B_1^{1-\alpha}$; we impose no a-priori restrictions on B_1.

[13] Letting h_t and i_{t-1} denote human capital and lagged education of an individual worker, earnings are given by

$$\frac{w_t}{H_t} h_t(1 - x_t) = (1 - \alpha) B_0 s_{t-1}^{\alpha} [H_t \nu(1 - x_t)]^{-\alpha} h_t(1 - x_t) \quad \text{s.t. } h_t = B_1 H_{t-1}^{1-\delta} (\nu^2 i_{t-1})^{\delta}.$$

[14] Campbell and Viceira (2005) report annualized gross returns for 90-day treasury-bills (1.0152), 5-year treasury-bonds (1.0289), and stocks (1.0783) for the period 1952–2002. We approximate the average return on savings by a weighted average of these returns (1.0483) where the weights are proportional to the relative size of "deposits", "credit market instruments", and "equity shares at market value, directly held plus indirectly held" in the balance sheets of households and non-profit organizations (Board of Governors of the Federal Reserve System, *Flow of Funds Accounts of the United States: Annual Flows and Outstandings*, several years; we use averages for the period 1955–2002).

According to the Bureau of Labor Statistics, multifactor productivity of private businesses grew by a factor of 1.8681 between 1952 and 2002 (`http://www.bls.gov/mfp/home.htm`, series MPU740023 (K)), or by an annual factor of 1.0126.

case with inelastic labor supply, we set $x = 2/3$. Alternatively, when considering the case with elastic labor supply, we assume $v(x) = \phi \ln(x)$ and we set $\phi = 2.6996$, implying that $x = 2/3$ when (3) is evaluated at κ^W. Finally, for the Ramsey planner's intergenerational welfare weight, we assign the value corresponding to households' discount factor multiplied by the population growth rate, $\rho = \beta\nu$. Table 1 summarizes our baseline calibration.

Table 1: Baseline Calibration

Parameter	Value
α	0.2815
ν	1.3113
ω, ψ, B_0	1.0000
δ	0.1200
β	$0.4340 \approx 0.97^{25}$
B_1	2.4174
$x,\ \phi$	0.6666, 2.6996
ρ	$= \beta\nu$

Notes: Under the assumption of inelastic labor supply, x is fixed and ϕ is redundant. Under the assumption of elastic labor supply, ϕ is fixed and x is endogenous.

Imposing these parameter assumptions on the closed-form solutions derived earlier produces the results summarized in Table 2. We highlight several central findings. First, the composition of government spending in politico-economic equilibrium starkly differs from the one under the Ramsey policy. In politico-economic equilibrium, government spending mainly flows to retirees while under the Ramsey policy, government outlays mainly fund education. This can most clearly be seen by comparing the shares of GDP devoted to transfers and education investment, respectively, defined as

$$\begin{aligned} \text{share of net transfers to retirees} &= \text{sh}_R \equiv (1-\alpha)(\tau_t - \eta_t \alpha'), \\ \text{investment share} &= \text{sh}_I \equiv (1-\alpha)\sigma_t. \end{aligned}$$

With about 7 and 5 percent in politico-economic equilibrium and about 4 and 10 percent under the Ramsey policy, respectively, these shares strongly differ across regimes. (The shares in the U.S. are approximately 7 and 5.5 percent, respectively.) Second, the different budget shares under the two policies translate into significantly different growth rates. On an annual basis, the Ramsey policy supports a 0.3 percentage points higher growth rate than the policy implemented in politico-economic equilibrium. This discrepancy arises although total tax rates under the Ramsey policy are higher than in politico-economic equilibrium. Nevertheless, third, the politico-economic equilibrium allocation is consumption and production efficient according to the criteria derived earlier. Finally, when we account for the distortions on labor supply, the politico-economic equilibrium features $\xi_t = 0$. Since $\eta_t = 0$ as well, tax distortions have no effect on the politico-economic equilibrium allocation. The Ramsey allocation, in contrast, does change once labor supply distortions are accounted for. Total tax rates fall, the composition of government spending is further tilted towards education investment, and the growth differential relative to the politico-economic equilibrium slightly increases. Moreover, labor supply rises by 2.8 percent, from 0.3333 in politico-economic equilibrium to 0.3425 under the Ramsey policy.

Table 2: Model Predictions

Variable	PE	Ramsey		
		inel.	el., t	el., i
τ	0.0943	0.0547	0.0547	0.0266
σ	0.0721	0.1368	0.1368	0.1409
η, ξ	0.0000	0.0000	0.0000	0.0000
sh_R	0.0677	0.0393		0.0191
sh_I	0.0518	0.0983		0.1012
R	3.2519	3.3961		3.1788
γ_H	1.3676	1.4739		1.4886
c_2/c_1	1.0320	1.0000		0.9268

Notes: The column title "PE" refers to the politico-economic equilibrium. The outcomes under the Ramsey policy are differentiated according to whether labor is supplied inelastically ("inel.") or elastically ("el."), and in the latter case, according to the period ("t" denotes the initial period of the planning horizon, "i" denotes any subsequent period).
With inelastic labor supply, ξ is normalized to zero. Otherwise, $\xi = 0$ is an equilibrium outcome.

Table 3 reports a variety of robustness checks. We analyze the sensitivity of the predicted tax rates and budget shares in politico-economic equilibrium with respect to the assumptions made about the fundamental parameters[15], and we examine the implied growth shortfall relative to the Ramsey allocation. For brevity, we only report results for the case with inelastic labor supply. The following findings emerge. First, τ^W responds more elastically to changes in the fundamental parameters than does σ^W. For example, changes in the capital share α effectively have no impact on σ^W and the growth differential while they do affect τ^W. Second, changes in ω and ψ affect τ^W and σ^W in the expected direction, and the effect on σ^W is reflected in changes in the growth shortfall. Third, the assumption of a lower growth rate in the data translates into a lower calibrated value for β and thus, a lower σ^W and higher τ^W.[16] Finally, an increase in δ leads to higher investment in education; nevertheless, the growth shortfall relative to the Ramsey allocation rises.

Returning to the baseline calibration, we analyze the effects of a change in the demographic structure. More specifically, we consider the equilibrium implications of a change in ν, holding all other fundamental parameters except $\rho \equiv \beta\nu$ and $B_1 \equiv \tilde{B}_1 \nu^{-1-\delta}$ constant. Because the effects of tax distortions on labor supply turned out to be of relatively minor importance, we consider the case with inelastic labor supply. Figure 1 displays the effect of ageing (a reduction

[15]For each change of parameter, we recalculate the implied values for β and B_1.

[16]The calibration of β hinges on the ratio of γ_H^{data} and R^{data}. To check the robustness of our findings, we therefore only need to consider changes in one of these two values.

Table 3: Robustness Checks

Variable	Baseline	α		ω		ψ		γ_H^{data}		δ
		0.2715	0.2915	0.8	1.2	0.8	1.2	1.0106^{25}	1.0146^{25}	0.155
τ^W	0.0943	0.1091	0.0790	0.0443	0.1346	0.1038	0.0855	0.0997	0.0886	0.0864
σ^W	0.0721	0.0722	0.0720	0.0634	0.0794	0.0656	0.0780	0.0683	0.0760	0.0917
sh_R	0.0677	0.0795	0.0560	0.0318	0.0967	0.0746	0.0614	0.0716	0.0637	0.0621
sh_I	0.0518	0.0526	0.0510	0.0455	0.0571	0.0471	0.0560	0.0491	0.0546	0.0659
agd	0.0030	0.0030	0.0030	0.0015	0.0045	0.0037	0.0024	0.0028	0.0033	0.0037

Notes: See explanations in the text. The row title "agd" refers to the annualized growth rate differential between the Ramsey allocation and the politico-economic equilibrium. The tax rates η^W and ξ^W always equal zero.

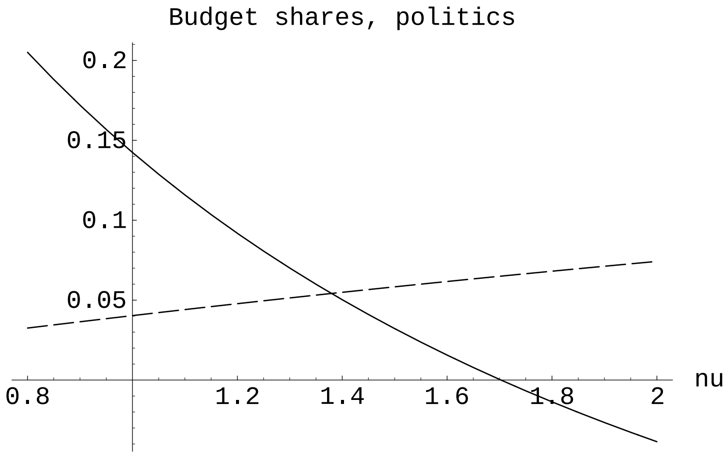

Figure 1: Effect of ν on budget shares (relative to output) in politico-economic equilibrium: sh_R^W (—), sh_I^W (- -)

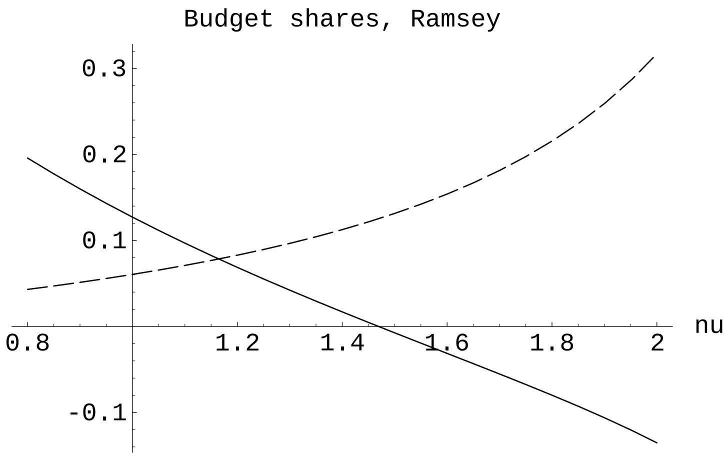

Figure 2: Effect of ν on budget shares (relative to output) in the Ramsey allocation: sh_R^G (—), sh_I^G (- -)

of ν) on transfer and investment shares. As ν decreases, the GDP-share of net transfers to retirees strongly rises. At the same time, the share of education investment falls, but only very modestly. Under the Ramsey policy, the qualitative response of the budget shares is similar to the response in politico-economic equilibrium, but much more pronounced, in particular as far as the investment share is concerned (see Figure 2). For a range of ν-values, retirees may contribute to the government's budget under the Ramsey policy while in politico-economic equilibrium they benefit from net transfers.

Figures 3 and 4 illustrate how the response of the budget shares to changes in ν corresponds with varying equilibrium tax rates. As long as both budget shares are positive, τ and σ develop similarly to the transfer and investment share, respectively. When net transfers to retirees are negative, however, τ is constant at zero and η is positive. Finally, Figure 5 displays the effect of ν on the per-capita growth rate of the economy. Faster population growth goes hand in hand with lower per-capita growth. Under the Ramsey policy, growth is higher than in politico-economic equilibrium, due to the higher share of resources flowing into growth enhancing public education

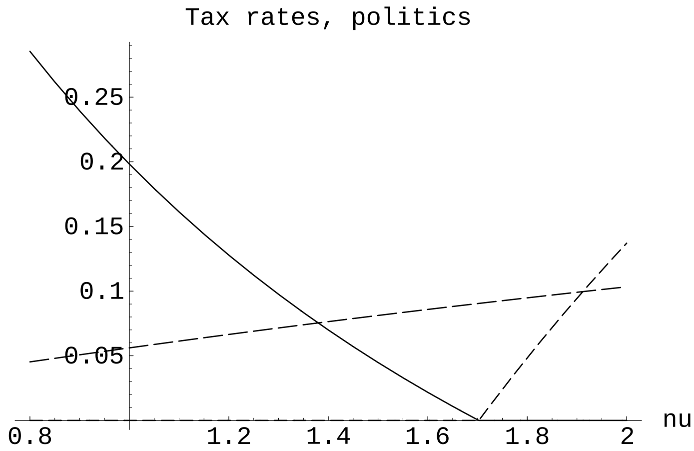

Figure 3: Effect of ν on tax rates in politico-economic equilibrium: τ^W (—), σ^W (- -), η^W (··)

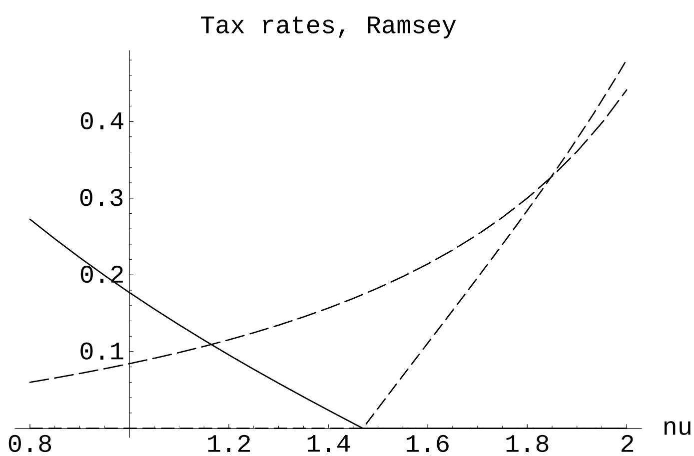

Figure 4: Effect of ν on tax rates in the Ramsey allocation: τ^G (—), σ^G (- -), η^G (··)

investment. The growth differential increases with ν because of the higher elasticity of sh_I under the Ramsey policy.

5 Conclusion

The choice between productive government investment and old-age transfers as well as the choice between capital and labor income taxes reflects the weights that political decision makers attach to two conflicting objectives: long-run growth on the one hand, and intergenerational redistribution on the other. In politico-economic equilibrium, the growth objective receives a relatively smaller weight than under the Ramsey policy because voters do not appreciate the long-term benefits of higher growth. Under our baseline parameter assumptions, the political process allocates 5 and 7 percent of GDP to public education and transfers, respectively, in line with U.S. data, but in contrast with 10 and 4 percent under the Ramsey policy. This difference in the composition of the government budget translates into a 0.3 percentage points lower per-capita growth rate in politico-economic equilibrium than under the Ramsey policy. Nevertheless,

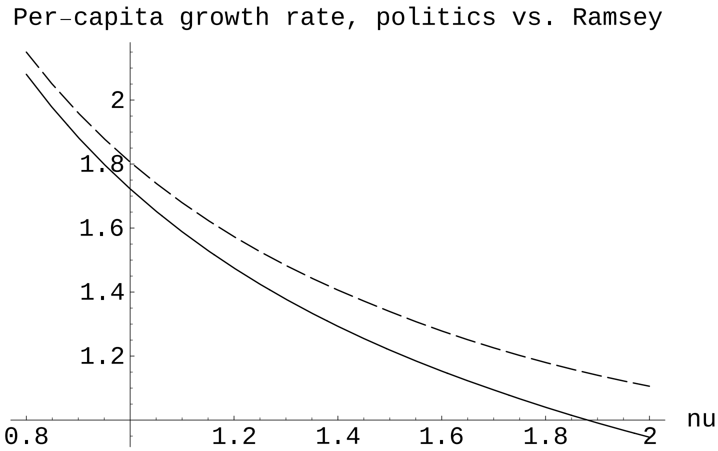

Figure 5: Effect of ν on the per-capita growth rate in politico-economic equilibrium and in the Ramsey allocation: γ_H^W (—), γ_H^G (- -)

the politico-economic equilibrium allocation is consumption and production efficient.

Our model assumes a balanced government budget, but this assumption is not restrictive. Indeed, allowing for government debt does not affect any of the results in the paper as long as the fundamental assumption of no commitment is maintained such that in equilibrium all debt is voluntarily redeemed.

References

Barro, R. J. (1990), 'Government spending in a simple model of endogenous growth', *Journal of Political Economy* **98**(5), S103–S125. 3

Bellettini, G. and Berti Ceroni, C. (1999), 'Is social security really bad for growth?', *Review of Economic Dynamics* **2**, 796–819. 4

Betts, J. R. (1996), Is there a link between school inputs and earnings? fresh scrutiny of an old literature, *in* G. Burtless, ed., 'Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success', Brookings Institution, Washington, chapter 6, pp. 141–191. 16

Bhaskar, V. (1998), 'Informational constraints and the overlapping generations model: Folk and anti-folk theorems', *Review of Economic Studies* **65**(1), 135–149. 3

Boldrin, M. and Montes, A. (2005), 'The intergenerational state: Education and pensions', *Review of Economic Studies* **72**, 651–664. 3, 28, 29

Campbell, J. Y. and Viceira, L. M. (2005), The term structure of the risk-return tradeoff, Working Paper 11119, NBER, Cambridge, Massachusetts. 16

Card, D. and Krueger, A. B. (1995), The economic return to school quality: A partial survey, *in* W. E. Becker and W. J. Baumol, eds, 'Assessing Educational Practices: The Contribution of Economics', MIT Press, Cambridge, Massachusetts, chapter 6. 16

Cass, D. (1972), 'On capital overaccumulation in the aggregative, neoclassical model of economic growth: A complete characterization', *Journal of Economic Theory* **4**, 200–223. 16

Diamond, P. A. (1965), 'National debt in a neoclassical growth model', *American Economic Review* **55**(5), 1126–1150. 2

Dixit, A. and Londregan, J. (1996), 'The determinants of success of special interests in redistributive politics', *The Journal of Politics* **58**(4), 1132–1155. 9

Glomm, G. and Ravikumar, B. (1992), 'Public versus private investment in human capital: Endogenous growth and income inequality', *Journal of Political Economy* **100**(4), 818–834. 3

Grossman, G. M. and Helpman, E. (1998), 'Intergenerational redistribution with short-lived governments', *Economic Journal* **108**(450), 1299–1329. 9

Krusell, P., Quadrini, V. and Ríos-Rull, J.-V. (1997), 'Politico-economic equilibrium and economic growth', *Journal of Economic Dynamics and Control* **21**(1), 243–272. 3

Lindbeck, A. and Weibull, J. W. (1987), 'Balanced-budget redistribution as the outcome of political competition', *Public Choice* **52**, 273–297. 8

Lucas, R. E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics* **22**, 3–42. 3

Perotti, R. (1993), 'Political equilibrium, income distribution, and growth', *Review of Economic Studies* **60**, 755–776. 3

Persson, T. and Tabellini, G. (2000), *Political Economics*, MIT Press, Cambridge, Massachusetts. 8

Piketty, T. and Saez, E. (2003), 'Income inequality in the United States, 1913–1998', *Quarterly Journal of Economics* **118**(1), 1–39. 16

Rangel, A. (2003), 'Forward and backward intergenerational goods: Why is social security good for the environment?', *American Economic Review* **93**(3), 813–834. 4

A Appendix

A.1 Ramsey Program

Denoting a typical term in the objective function by π_s, we have

$$\begin{aligned}\pi_s &\equiv \beta\ln(c_{2,s}) + \rho\ln(c_{1,s}) + \rho v(x_s) \;\text{ s.t. } (2), (3) \\ &= \beta\ln[q_s(1-x_s)^{1-\alpha}(\alpha(1-\eta_s)+(1-\alpha)\tau_s)] + \\ &\quad \rho\ln[q_s(1-x_s)^{1-\alpha}(1-\tau_s-\sigma_s+\eta_s\alpha')(1-z(\tau_{s+1},\eta_{s+1}))] + \\ &\quad \rho v(x_s) + \text{constant terms} \;\text{ s.t. } (3) \\ &= \ln(q_s)(\beta+\rho) + \ln(1-x_s)(1-\alpha)(\beta+\rho) + \rho v(x_s) + \\ &\quad \beta\ln(\alpha(1-\eta_s)+(1-\alpha)\tau_s) + \rho\ln(1-\tau_s-\sigma_s+\eta_s\alpha') + \\ &\quad \rho\ln(1-z(\tau_{s+1},\eta_{s+1})) + \text{constant terms} \;\text{ s.t. } (3).\end{aligned}$$

Consider the direct and indirect effects (the latter working through induced changes in q_s) on the objective function that are triggered by a marginal change in one of the policy instruments, ϕ_i say with $\phi_i \in \{\tau_i, \sigma_i, \eta_i, \xi_i\},\ i \geq t$. The direct effect is given by

$$\begin{aligned}\frac{dG(H_t, q_t, \{\kappa_s\}_{s=t}^{\infty})}{d\phi_i}\Big|_{\text{dir}} &= \rho^{i-t}(\rho v'(x_i) - (1-\alpha)(\beta+\rho)/(1-x_i))\frac{\partial x_i}{\partial \phi_i} + \\ &\quad \rho^{i-1-t}(\rho v'(x_{i-1}) - (1-\alpha)(\beta+\rho)/(1-x_{i-1}))\frac{\partial x_{i-1}}{\partial \phi_i} + \\ &\quad \rho^{i-t}\frac{\partial \beta\ln(\alpha(1-\eta_i)+(1-\alpha)\tau_i) + \rho\ln(1-\tau_i-\sigma_i+\eta_i\alpha')}{\partial\phi_i} + \\ &\quad \rho^{i-1-t}\frac{\partial\rho\ln(1-z(\tau_i,\eta_i))}{\partial\phi_i},\end{aligned}$$

where the second and fourth lines only apply if $i > t$ as they capture effects of ϕ_i on choices in the preceding period, $i-1$.[17]

The indirect effect is given by

$$\begin{aligned}\frac{dG(H_t, q_t, \{\kappa_s\}_{s=t}^{\infty})}{d\phi_i}\Big|_{\text{ind}} &= \sum_{s=t}^{\infty}\rho^{s-t}(\beta+\rho)\frac{d\ln(q_s)}{d\phi_i} \;\text{ s.t. } (4) \\ &= (\beta+\rho)\sum_{s=t}^{\infty}\rho^{s-t}\left[M^{s-1-i}\frac{df_i}{d\phi_i} + M^{s-i}\frac{df_{i-1}}{d\phi_i}\right]_{[2]} \\ &= (\beta+\rho)\left[\rho^{i+1-t}\frac{df_i}{d\phi_i}\sum_{k=0}^{\infty}(\rho M)^k + \rho^{i-t}\frac{df_{i-1}}{d\phi_i}\sum_{k=0}^{\infty}(\rho M)^k\right]_{[2]} \\ &= (\beta+\rho)\rho^{i-t}\left[(I-\rho M)^{-1}\right]_{[2,\cdot]}\left(\rho\frac{df_i}{d\phi_i} + \frac{df_{i-1}}{d\phi_i}\right),\end{aligned}$$

[17] Using (3), the terms in the first and second line can be simplified to

$$\begin{aligned}\left(\rho v'(x_i) - \frac{(1-\alpha)(\beta+\rho)}{1-x_i}\right)x_i' &= \left(\rho v'(x_i)(1-x_i) - (1-\alpha)(\beta+\rho)\right)\frac{x_i'}{1-x_i} = \\ &= \left(\frac{1-\tau_i-\sigma_i-\xi_i}{1-\tau_i-\sigma_i+\eta_i\alpha'}\frac{\rho}{1-z(\tau_{i+1},\eta_{i+1})} - (1-\alpha)(\beta+\rho)\right)\frac{x_i'}{1-x_i}.\end{aligned}$$

where matrices with a negative exponent are zero by convention. For the same reason as above, the term $\partial f_{i-1}/\partial \phi_i$ only applies if $i > t$.

We now focus on the case with inelastic labor supply. Disregarding the inequality constraints (1), the first-order conditions with respect to τ_t, σ_t, and η_t read

$$\begin{aligned}
\frac{(1-\alpha)\beta}{\alpha(1-\eta_t)+(1-\alpha)\tau_t} - \frac{\rho(1+\Omega_{22}\alpha)}{1-\tau_t-\sigma_t+\eta_t\alpha'} &= 0, \\
-\frac{\rho(1+\Omega_{22}\alpha)}{1-\tau_t-\sigma_t+\eta_t\alpha'} + \frac{\rho\delta(\Omega_{21}+\Omega_{22}(1-\alpha))}{\sigma_t} &= 0, \\
\frac{-\alpha\beta}{\alpha(1-\eta_t)+(1-\alpha)\tau_t} + \frac{\rho\alpha'(1+\Omega_{22}\alpha)}{1-\tau_t-\sigma_t+\eta_t\alpha'} &= 0,
\end{aligned}$$

respectively. Here, $\Omega_{21} \equiv (\beta+\rho)\left[(I-\rho M)^{-1}\right]_{[2,1]}$ and $\Omega_{22} \equiv (\beta+\rho)\left[(I-\rho M)^{-1}\right]_{[2,2]}$. These two terms represent the present discounted contribution to the planners objective of an increase in the (logarithms of the) economy's stock of human capital and q, respectively. Clearly, the first first-order condition is proportional to the last one; the tax rate η_t can therefore be set to some exogenous value, subject to satisfying (1). Rearranging and solving the above conditions yields the first-order conditions and equilibrium tax rates that are reported in the text.

Turning to the first-order conditions with respect to τ_i and η_i, $i > t$, additional feedback effects come into play as taxes affect households' savings decisions in the preceding period. (In contrast, the choice of σ_i does not affect preceding savings decisions, see (2).) Formally, the first-order conditions with respect to τ_i and η_i, respectively, feature the following terms *in addition* to the terms listed above:

$$\begin{aligned}
&-\frac{\partial z(\tau_i,\eta_i)}{\partial \tau_i}\left(\frac{1}{1-z(\tau_i,\eta_i)} - \frac{\Omega_{22}\alpha}{z(\tau_i,\eta_i)}\right), \\
&-\frac{\partial z(\tau_i,\eta_i)}{\partial \eta_i}\left(\frac{1}{1-z(\tau_i,\eta_i)} - \frac{\Omega_{22}\alpha}{z(\tau_i,\eta_i)}\right).
\end{aligned}$$

These terms reflect the fact that induced changes in the *preceding* period's savings rate affect the planner's objective both directly (altered consumption of workers, see the left-hand side term in the brackets) and indirectly (implications of altered capital accumulation, see the right-hand side term in the brackets). If $\frac{z(\tau_t,\eta_t)}{1-z(\tau_t,\eta_t)} = \Omega_{22}\alpha$ when evaluated at the optimal period-t tax rates, then the direct and indirect feedback effects cancel, implying that the optimal tax rates in period t are also optimal in all subsequent periods.

If $\eta_t = 0$, then it is indeed the case that $\frac{z(\tau_t,\eta_t)}{1-z(\tau_t,\eta_t)} = \Omega_{22}\alpha$ when evaluated at the optimal period-t tax rates. To understand this result note first that $\eta_t = 0$ in combination with time-invariant tax rates (at the period-t optimal values) implies that the consumption ratio $c_{1,s}/c_{2,s}$ equals $\rho/(\beta\nu)$ for all $s \geq t$ (see (2)); that is, the Ramsey planner achieves the same consumption ratio as a social planner that is not constrained by the implementability constraints.[18] Second, if capital income remains untaxed ($\eta_i = 0$ for all $i \geq t$) then households' privately optimal capital accumulation conforms with the planner's preferred savings level (conditional on the planner's preferred level of human capital accumulation). As long as the optimal capital income tax rate equals zero, the implementability constraints therefore are non-binding and the Ramsey planner implements the same allocation as a social planner that is only constrained by the resource constraint. But such a social-planner allocation is necessarily time-consistent.

[18]The social planner equalizes the marginal rate of substitution between consumption of retirees and workers, $(\beta/c_{2,s})/(\rho/c_{1,s})$, and the corresponding marginal rate of transformation, $1/\nu$.

If the optimal $\eta_t > 0$, direct and indirect feedback effects do not cancel. Nevertheless, the Ramsey policy continues to be time-consistent. To see this, note that a strictly positive η_t implies $\tau_t = 0$. Moreover, $\partial z(0, \eta_t)/\partial \eta_t = 0$. The potential source of time-inconsistency therefore disappears as far as the choice of η_i is concerned. Moreover, one can show that parameter constellations for which $\tau_t = 0$ and $\eta_t > 0$ necessarily imply the inequality

$$-\frac{\partial z(0,\eta_t)}{\partial \tau_t}\left(\frac{1}{1-z(0,\eta_t)} - \frac{\Omega_{22}\alpha}{z(0,\eta_t)}\right) < 0.$$

As a consequence, a corner solution for κ_t^G goes hand in hand with a corner solution for κ_i^G for all $i > t$ and the potential source of time-inconsistency also disappears as far as the choice of τ_i is concerned.

In conclusion, the Ramsey government sets time invariant tax rates, and this policy is time-consistent.

A.2 Production Efficiency

Let y denote output per retiree; conditional on the state variables in period t, we have

$$\ln(y_{t+i+1}) \simeq \alpha \ln(s_{t+i}) + \delta(1-\alpha)\sum_{j=0}^{i}(1-\delta)^j \ln(I_{t+i-j}),\ i \geq 0.$$

Conditional on the state variables in period t and some initial sequence of investment spending, $\{s_{t+i}, I_{t+i}\}_{i=0}^{\infty}$, consider a sequence of small changes in the investment policy that leaves total investment in each period unchanged. This latter sequence involves, in each period i, a small change in human capital investment of Δ_i (per retiree in period i), and a corresponding change in physical capital investment of $-\Delta_i/\nu$ (per retiree in period $i+1$). If this policy change weakly increases output in all subsequent periods, then it amounts to a Pareto improvement and the initial allocation is production inefficient. Formally, the conditions for production inefficiency are given by

$$\begin{aligned} d\ln(y_{t+i+1}) &= -\frac{\alpha}{\nu}\frac{\Delta_{t+i}}{s_{t+i}} + \delta(1-\alpha)\sum_{j=0}^{i}(1-\delta)^j \frac{\Delta_{t+i-j}}{I_{t+i-j}} = \\ &= -\frac{\alpha}{\nu}\frac{I_{t+i}}{s_{t+i}}\epsilon_{t+i} + \delta(1-\alpha)\sum_{j=0}^{i}(1-\delta)^j \epsilon_{t+i-j} \geq 0 \ \text{ for all } i \geq 0, \end{aligned}$$

where we define $\epsilon_{t+i} \equiv \Delta_{t+i}/I_{t+i}$, and where at least one inequality must hold strictly. Since the initial allocation corresponds to a balanced growth path, the recurrent term

$$a \equiv -\frac{\alpha}{\nu}\frac{I_{t+i}}{s_{t+i}} + \delta(1-\alpha)$$

is time-invariant. The conditions for production inefficiency can therefore be summarized as

$$\begin{aligned} &a\epsilon_t \geq 0, \\ &a\epsilon_{t+i} + \delta(1-\alpha)\sum_{j=1}^{i}(1-\delta)^j \epsilon_{t+i-j} \geq 0 \ \text{ for all } i \geq 1, \end{aligned}$$

where at least one inequality must hold strictly.

Intuitively, the term a (multiplied by the amount of physical capital investment) represents the effect of an infinitesimal reallocation from physical to human capital investment on output in the subsequent period. To increase output in period $t+1$, ϵ_t must have the same sign as a. To increase output in periods later than period $t+1$, the combined effect of the lagged change in physical capital investment and the cumulative change of human capital investment must be positive. This latter, cumulative change is reflected in the summation that appears in the conditions for periods $i \geq 1$.

When $a \geq 0$, physical capital is over accumulated in the initial allocation. As is apparent from the above conditions, in this case one can generate a Pareto improvement by reallocating resources from physical to human capital (corresponding to $\epsilon_{t+i} > 0$). Specifically, over accumulation of physical capital is present if $a = 0$, corresponding to the allocation in an economy without government intervention, but with markets for private education financing (see Boldrin and Montes (2005) and Appendix A.3 for a characterization of this hypothetical economy). In such a complete markets setting, savings is allocated across human and physical capital investment in such a way that output in the *subsequent* period cannot be increased. However, if $1-\delta > 0$, the stock of human capital contributes to future human capital accumulation, and a slight reallocation from physical to human capital investment therefore increases output in all *later* periods, as is apparent from the above conditions. The complete markets allocation ($a = 0$) is not Pareto optimal because it does not properly account for the dynamic human capital externality.

When a is negative and large in absolute value, the allocation again is production inefficient. In this case, a reallocation of resources from human to physical capital accumulation (corresponding to $\epsilon_{t+i} < 0$) generates a Pareto improvement. For example, if $a = -1$, a sequence of $\epsilon_{t+i} = \epsilon < 0$ for all $i \geq 0$ increases production in all future periods because the positive effect from additional physical capital investment, $a\epsilon = -\epsilon > 0$, dominates the cumulative negative effect from reduced human capital accumulation, $\delta(1-\alpha)\sum_{j=1}^{i}(1-\delta)^j\epsilon < \epsilon$. To characterize the largest $a < 0$ allowing for an improvement of production efficiency, we consider a sequence $\{\epsilon^\star_{t+i}\}_{i=0}^{\infty}$ with $\epsilon^\star_t < 0$ where $\{\epsilon^\star_{t+i}\}_{i=1}^{\infty}$ is recursively defined by the requirement that $d\ln(y_{t+i}) = 0$ for all $i \geq 2$. If such a sequence is bounded then production is inefficient. For $i \geq 1$, the terms of such a sequence satisfy $a\epsilon^\star_{t+i} + \delta(1-\alpha)\sum_{j=1}^{i}(1-\delta)^j\epsilon^\star_{t+i-j} = 0$ and thus,

$$\begin{aligned}
\epsilon^\star_{t+i} &= \frac{\delta(1-\alpha)}{-a}\sum_{j=1}^{i}(1-\delta)^j\epsilon^\star_{t+i-j} \\
&= \frac{\delta(1-\alpha)}{-a}(1-\delta)\epsilon^\star_{t+i-1} + \frac{\delta(1-\alpha)}{-a}\sum_{j=2}^{i}(1-\delta)^j\epsilon^\star_{t+i-j} \\
&= \frac{\delta(1-\alpha)}{-a}(1-\delta)\epsilon^\star_{t+i-1} + \frac{\delta(1-\alpha)}{-a}(1-\delta)\sum_{j=1}^{i-1}(1-\delta)^j\epsilon^\star_{t+i-j-1} \\
&= \frac{\delta(1-\alpha)}{-a}(1-\delta)\epsilon^\star_{t+i-1} + (1-\delta)\epsilon^\star_{t+i-1} \\
&= (1-\delta)\left(1-\frac{\delta(1-\alpha)}{a}\right)\epsilon^\star_{t+i-1}.
\end{aligned}$$

Boundedness of the sequence and thus, production inefficiency, requires $-1 < (1-\delta)\left(1-\frac{\delta(1-\alpha)}{a}\right) < 1$ which simplifies to the condition $a < -(1-\alpha)(1-\delta)$. Under some time-invariant policy (τ, σ, η), we have $a = (1-\alpha)\left(\delta - \frac{\sigma R}{\nu \gamma_H}\right)$ because $\frac{I_{t+i}}{s_{t+i}} = \frac{I_{t+i}}{s_{t+i-1}\gamma_H} = \frac{\sigma R}{\alpha' \gamma_H}$. The criterion for production efficiency, $-(1-\alpha)(1-\delta) < a < 0$, therefore reduces to

$$\frac{\delta}{\sigma} < \frac{R}{\nu \gamma_H} < \frac{1}{\sigma}.$$

A.3 Laissez Faire and Markets for Private Education Financing

We consider the situation without government intervention, but with complete markets for private education financing. Students individually decide how much to invest in education and thus, how much to borrow (from workers). This is the setup considered by Boldrin and Montes (2005), with the difference that we allow for population growth.

Due to the absence of taxes and benefits, disposable labor income equals $w_t(1-x_t)$ and old-age income equals $d_t R_{t+1}$ with d_t denoting per-capita savings of a worker. These savings differ from capital accumulation s_t due to the presence of student loans. Let i_t denote the private education choice of a typical student and define the market wage per average stock of human capital, w_t^N, as

$$w_t^N \equiv w_t/H_t = (1-\alpha)B_0 s_{t-1}^{\alpha}[H_t \nu (1-x_t)]^{-\alpha}.$$

Since we normalized aggregate education investment I_t with respect to the number of retirees, we have $i_t \nu^2 = I_t$.

The private education choice i_{t-1} affects the income of a worker twofold: On one hand, it generates private human capital h_t for the worker and thus, labor income; on the other hand, it triggers repayment of the student loan used to finance the education. Taking w_t^N as given, students in period $t-1$ then solve

$$\max_{i_{t-1}} \; -R_t i_{t-1} + w_t^N h_t (1-x_t) \;\text{ s.t. } h_t = B_1 H_{t-1}^{1-\delta}(\nu^2 i_{t-1})^{\delta}.$$

This program yields the first-order condition $R_t = \delta w_t^N (1-x_t) h_t / i_{t-1}$ and thus,

$$R_t i_{t-1} = \delta w_t (1-x_t) \;\Rightarrow\; \delta s_{t-1} = \alpha' \nu i_{t-1}.$$

The first-order conditions characterizing the consumption-savings and labor-leisure tradeoff of the household remain unchanged, except for the fact that all tax rates and subsidies equal zero.

In general equilibrium, per-capita savings of a worker are split between physical capital accumulation and student loans extended to the following cohort:

$$d_t = \frac{\beta}{1+\beta}[w_t(1-x_t) - R_t i_{t-1}] = s_t + \nu i_t.$$

Combining these results, we have

$$\begin{aligned} d_t &= s_t + \nu i_t = s_t\left(1 + \frac{\delta}{\alpha'}\right), \\ d_t &= \frac{\beta}{1+\beta}[w_t(1-x_t) - R_t i_{t-1}] = \frac{\beta}{1+\beta}(1-\delta) w_t (1-x_t). \end{aligned}$$

Eliminating d_t and substituting, we arrive at the following characterization of the complete-markets allocation:

$$\begin{aligned}
s_t &= \frac{\beta}{1+\beta}\frac{\alpha'}{\alpha'+\delta}(1-\delta)(1-\alpha)B_0\nu^{-\alpha}(1-x_t)^{1-\alpha}q_t, \\
c_{1,t} &= \frac{1}{1+\beta}(1-\delta)(1-\alpha)B_0\nu^{-\alpha}(1-x_t)^{1-\alpha}q_t, \\
c_{2,t} &= (\alpha'+\delta)\nu(1-\alpha)B_0\nu^{-\alpha}(1-x_t)^{1-\alpha}q_t, \\
x_t &= x:\ v'(x)(1-x)=\frac{1+\beta}{1-\delta}\ \text{(with elastic labor supply)}, \\
H_{t+1} &= B_1\left(\frac{\delta\nu}{\alpha'}\right)^{\delta}H_t^{1-\delta}\left[\frac{\beta}{1+\beta}\frac{\alpha'}{\alpha'+\delta}(1-\delta)(1-\alpha)B_0\nu^{-\alpha}(1-x_t)^{1-\alpha}q_t\right]^{\delta}, \\
q_{t+1} &= \left[B_1\left(\frac{\delta\nu}{\alpha'}\right)^{\delta}H_t^{1-\delta}\right]^{1-\alpha}\left[\frac{\beta}{1+\beta}\frac{\alpha'}{\alpha'+\delta}(1-\delta)(1-\alpha)B_0\nu^{-\alpha}(1-x_t)^{1-\alpha}q_t\right]^{\delta(1-\alpha)+\alpha}.
\end{aligned}$$

On a balanced growth path, s_t, H_t, and q_t all grow at the same rate. Imposing constancy of s_t/H_{t+1}[19] and solving, we find the balanced growth rate γ_H to satisfy

$$\gamma_H = \left[B_1\left(\frac{\delta\nu}{\alpha'}\right)^{\delta}\right]^{\frac{1-\alpha}{1-\alpha(1-\delta)}}\left[\frac{\beta}{1+\beta}\frac{\alpha'}{\alpha'+\delta}(1-\delta)(1-\alpha)B_0\nu^{-\alpha}(1-x)^{1-\alpha}\right]^{\frac{\delta}{1-\alpha(1-\delta)}}.$$

The fractions of GDP flowing into education and repayments of student loans, respectively, are given by

$$\begin{aligned}
\text{investment share} &= \beta\delta\frac{(1-\alpha)(1-\delta)}{(1+\beta)(\alpha'+\delta)}, \\
\text{student loan repayment share} &= \delta(1-\alpha).
\end{aligned}$$

To understand the latter share, note that a worker's repayment of student loans, relative to her labor income, equals $\frac{i_{t-1}R_t}{w_t(1-x_t)}$. Using the optimality condition characterizing education, $i_{t-1}R_t = \delta w_t(1-x_t)$, it follows that this ratio equals δ. Normalized by $1/(1-\alpha)$, this translates into the stated share of GDP.

[19] On a balanced growth path,

$$\frac{s_t}{H_{t+1}} = \left(\frac{\left[\frac{\beta}{1+\beta}\frac{\alpha'}{\alpha'+\delta}(1-\delta)(1-\alpha)B_0\nu^{-\alpha}(1-x)^{1-\alpha}\right]^{1-\delta}}{B_1\left(\frac{\delta\nu}{\alpha'}\right)^{\delta}}\right)^{\frac{1}{1-\alpha(1-\delta)}}.$$