Population Ageing, Government Budgets, and Productivity Growth in Politico-Economic Equilibrium

Martín Gonzalez-Eiras and Dirk Niepelt

Working Paper 07.05

Population Ageing, Government Budgets, and Productivity Growth in Politico-Economic Equilibrium*

Martín Gonzalez-Eiras
Universidad de San Andrés†

Dirk Niepelt
Study Center Gerzensee‡
University of Bern
Institute for International Economic Studies, Stockholm University
CEPR

November 14, 2007

Abstract

We analyze the effect of changes in fertility and longevity on taxes, the composition of government spending, and productivity. To that purpose, we introduce politics in an OLG economy with endogenous growth due to human and physical capital accumulation. Population ageing shifts political power from students and workers to retirees, leading to a reallocation of resources from education spending to retirement benefits and a slowdown of productivity growth. Calibrated to U.S. data, the closed-form solutions of the model predict retirement benefits as a share of GDP to strongly increase over the next decades and the education share to fall. This effect depresses the annual productivity growth rate by 10 basis points. In spite of higher labor-income taxes, per-capita labor supply is predicted to rise, as a consequence of increased life expectancy. The equilibrium allocation is consumption and production efficient, but the political process allocates a much smaller share of resources to eduction than a Ramsey planner with balanced welfare weights.

*For useful comments, we thank Enrique Kawamura and Fabrizio Zilibotti as well as conference and seminar participants at CREI (Universidad Pompeu Fabra), EEA annual meeting (Budapest, 2007), EPRU (University of Copenhagen), IIES (Stockholm University), SED annual meeting (Vancouver, 2006), Study Center Gerzensee, and Universidad de San Andrés.

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

‡P.O. Box 21, CH-3115 Gerzensee, Switzerland. E-mail: dirk.niepelt@iies.su.se mkvA2.tex

1 Introduction

The prospect of "graying" populations in many developed economies raises concerns about the sustainability of fiscal policies. Rising old-age dependency ratios threaten to translate into growing tax burdens and thus, to depress economic activity. At the same time, generous pension and health care benefits threaten to crowd out public investment spending, for instance for education, with negative effects on productivity growth. Recent experience suggests that such concerns may be warranted. In the United States, for example, government spending on the elderly grew much faster during the second half of the twentieth century than other components of government spending (Mulligan and Sala-i-Martin, 2004); the quality of public infrastructure is deemed poor and public spending on infrastructure insufficient (American Society of Civil Engineers, 2005); and the fraction of elderly residents is negatively associated with education spending per child (Poterba, 1997).

In this paper, we develop a tractable framework to analyze the structure of government budgets in politico-economic equilibrium as well as the macroeconomic implications of this structure. We apply the framework to study the effect of the projected demographic transition—decreasing fertility and increasing longevity—on taxes, the composition of government spending, and productivity, focusing on two spending categories—transfers between workers and retirees, and public expenditures for education—that are of central importance for developed economies (accounting for nearly one half of public sector spending in the United States).

Building on a standard three-period overlapping generations model with physical and human capital, our framework endogenizes a number of political and economic choices. In their role as economic agents, households in the model take prices, taxes, education spending and retirement benefits as given when choosing consumption, savings, and labor supply. As voters, households choose among office motivated parties that offer policy platforms comprising labor and capital income taxes as well as the expenditure shares for inter-generational transfers and public education. Elections take place every period. As a consequence, the political process lacks commitment.

In this environment, the financing of government expenditures has negative effects on growth because labor and capital income taxes depress disposable income and reduce the incentive to accumulate capital. The composition of government spending also affects the growth rate, in line with empirical evidence (see, for example, Blankenau, Simpson and Tomljanovich, 2007). While transfers to retirees lower the incentive to save and further depress growth, public education expenditure fosters human capital accumulation and thus, future productivity.

Fiscal policy choices are of different concern to the different cohorts. On the one hand, the exposure of agents to capital and labor income taxes changes over the life cycle. On the other hand, retirees benefit from transfers to their group, while students and workers benefit from the effect of education spending on human capital and thus, future returns to labor and capital. When evaluating the policy platforms 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 under the presumption that voters' support for a party is subject to a small degree of randomness. This randomness induces a continuous mapping from parties' electoral platforms to vote shares, in contrast with the discontinuous mapping that arises under the more common assumption of a pivotal median voter. As a consequence, the probabilistic-voting assumption allows to capture gradual adjustments of policy in response to changes in the economic or demographic environment, even in a stylized three-period-lived overlapping-generations environment.

Tax rates and spending shares do not only affect human and physical capital accumulation, factor prices, and incomes. Absent commitment, they also affect, indirectly, future policy outcomes. In addition to the "economic" repercussions of their policy choices, voters therefore have to internalize the "political" repercussions. In particular, voters must account for the equilibrium relationship between future state variables and policy choices. We assume that only fundamental state variables enter this equilibrium relationship, excluding artificial state variables of the type sustaining trigger strategy equilibria. This restriction absolves us from having to make arbitrary assumptions about the strategies being played; it also reflects our assumption that political choices suffer from a lack of commitment, including commitment to particular enforcement strategies. While we agree that the *existence* of intergenerational transfers or public education may also owe to reputational arrangements, we focus on the Markov perfect equilibrium in order to identify the fundamental and robust forces that shape the *size* of these programs.[1]

Under standard functional form assumptions, we are able to characterize the politico-economic equilibrium in closed form. The optimal strategy of 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 while students and workers favor at least some education spending, the political process typically sustains both types of expenditures. Changes in the demographic structure affect the equilibrium allocation both directly and indirectly, by altering the balance of power and thus, policy choices. With population ageing, the politico-economic equilibrium features increasingly large budget shares flowing into intergenerational transfers and a decline of the budget share devoted to public education, in line with notions voiced in the public debate. The politico-economic equilibrium generically differs from the allocation implemented by a Ramsey planner with (arbitrary) geometric social welfare weights. Nevertheless, for a set of parameters with positive measure (including all parameterizations considered in the quantitative analysis), the allocation implemented in politico-economic equilibrium is consumption and production efficient.

To assess the quantitative implications of the model, we calibrate it to U.S. data, matching growth and interest rates as well as budget shares in selected years. Feeding historical and forecasted values for fertility and longevity into the model, we derive predictions for tax rates, government spending shares, labor supply, and productivity growth between 1975 and 2075. The model does a good job at fitting the historical trends. Out of sample, it predicts that the population ageing over the next decades will lead to an increase of the GDP share of retirement benefits by more than fifty percent, accompanied by a fall of the education share. In spite of higher tax rates, labor supply is predicted to

[1]For a discussion of Markov perfect equilibrium see, for example, Krusell, Quadrini and Ríos-Rull (1997).

rise, due to higher life expectancy. After the completion of the demographic transition, the model-implied annual productivity growth falls ten basis points short of the level that would prevail if tax rates and spending shares were at their current values.

Using the calibrated parameter values, we also compute the Ramsey policy under the assumption that welfare weights reflect population growth and households' time preference. We find that the Ramsey policy calls for a labor-income tax rate close to or even higher than the actual one, but sustains a much higher budget share for education spending and a much lower share for transfers. This holds true even if dynamic human capital externalities are weak such that growth is exogenous. With endogenous growth, productivity grows substantially faster under the Ramsey policy than in politico-economic equilibrium.

These findings have important policy implications. On the one hand, they undermine the notion among many policy makers that the political process will implement measures to boost productivity in order to "outgrow" the burden imposed by population ageing.[2] According to the model, the political process rather will do the opposite, by reallocating resources from productive use to transfers; productivity growth will only be sustained (or even strengthened) because the slowdown in population growth reduces the usual capital dilution effect. On the other hand, the analysis of the benchmark Ramsey allocation points to potentially large welfare costs (if measured by a utilitarian social welfare function), due to the comparatively low education spending in politico-economic equilibrium. Unlike in Bassetto and Sargent (2006) (who assume commitment), this "underinvestment" problem cannot be overcome by letting voters finance investment expenditures out of government debt. For the lack of commitment in our setting implies that the economic equivalence between intergenerational transfers and certain debt-plus-tax policies extends to the political sphere such that the equilibrium allocations with and without government debt coincide.[3]

We are not the first ones to incorporate politics in an overlapping-generations model to analyze the choice of productive and redistributive public spending. Bellettini and Berti Ceroni (1999) and Rangel (2003) show how societies may sustain public investment (e.g., education) even if the interests of those benefiting from the investment are not represented in the political process. In both papers, voters support public investment because a trigger strategy links investment spending to the provision of public pensions by future cohorts. Our model adopts a different perspective. Rather than emphasizing complementarities between investment and transfer payments, it focuses on the *conflict* over the size of these two spending components, and how the resolution of this conflict is shaped by fertility and life expectancy. The model also differs from previous literature in that it features various economic choices, embedded in the standard growth model, in addition to the political choices of central interest. This allows us to adopt a quantitative approach to evaluating the consequences of population ageing without having to sacrifice analytical tractability.[4]

[2]See, for example, the report in *The Economist*, October 20, 2005, or the discussions surrounding the European Union's "Lisbon Agenda."

[3]See Gonzalez-Eiras and Niepelt (2007).

[4]In Bellettini and Berti Ceroni (1999), population ageing triggers a decrease of both public investment

Our work also relates to politico-economic models of redistribution and growth. For example, Alesina and Rodrik (1994), Persson and Tabellini (1994) and Krusell et al. (1997) argue that inequality depresses growth because anticipated redistributive taxation reduces the incentive to accumulate, or because higher inequality pushes the median voter's preferred level of public investment and taxes beyond the growth-maximizing level. Relative to these papers, we model inter- rather than intragenerational conflict and we consider a larger set of policy instruments available to policy makers. Our analysis therefore sheds light on the equilibrium size and *composition* of the government's budget, both on the financing and the spending side.[5] Glomm and Ravikumar (1992) and Perotti (1993) also analyze distributive conflict in models with human capital accumulation. They focus on the political choice of public versus private education and the effect of distortive redistribution in the presence of borrowing constraints, respectively.

The remainder of the paper is structured as follows. Section 2 describes the model and characterizes the allocation conditional on policy. Section 3 solves for the politico-economic equilibrium and analyzes its efficiency properties. Section 4 contains the quantitative analysis, and 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 production and the formation of new human capital, and save for retirement. With probability p_t, workers in period $t-1$ survive to become retirees in period t. Retirees do not work and die at the end of the period.

Each cohort consists of a continuum of homogeneous agents. The ratio of workers to retirees in period t equals ν_t/p_t which follows a deterministic process. The period-t ratio of students to workers equals ν_{t+1}. On a balanced growth path, the survival probability is constant at value p and the gross population growth rate is given by ν.

Savings of workers who die before reaching retirement age are distributed among their surviving peers (reflecting an annuities market) and among the members of the following cohort (reflecting accidental bequests). The parameter f measures the fraction of retirement savings that is annuitized while $1-f$ measures the importance of accidental bequests.[6]

and transfers. In Rangel (2003), within some limits, population ageing increases public investment.

[5] Our work shares with Krusell et al. (1997) the restriction to Markov perfect equilibrium. Methodologically, it is related to Gonzalez-Eiras and Niepelt (2008).

[6] Since bequests are distributed equally among workers (for example, because workers insure each other), there is no wealth heterogeneity within cohorts.

2.1 Technology

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

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

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 worker, k_t, therefore corresponds to the per-capita savings of workers in the previous period, s_{t-1}, normalized by ν_t. Labor is supplied by workers whose productivity is given by their human capital, H_t. We normalize the time-endowment of a worker to unity and denote workers' leisure consumption by x_t.

Production factors are paid their marginal products, due to perfect competition. The wage per unit of time, w_t, and the gross return on physical capital, R_t, therefore satisfy

$$\begin{aligned} w_t &= (1-\alpha)B_0 H_t^{1-\alpha} k_t^{\alpha}(1-x_t)^{-\alpha}, \\ R_t &= \alpha B_0 H_t^{1-\alpha} k_t^{\alpha-1}(1-x_t)^{1-\alpha} = w_t \frac{1-x_t}{k_t}\alpha' \end{aligned}$$

with $\alpha' \equiv \alpha/(1-\alpha)$. As a consequence of the (incomplete) annuitization, the gross return on savings of a worker that survives to retirement equals $\tilde{R}_t \equiv R_t(1+(1-p_t)f/p_t) = R_t(1-f+f/p_t)$. With perfect annuities markets, the gross return equals R_t/p_t.

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

$$H_{t+1} = B_{1,t} H_t^{\varepsilon(1-\delta)} I_t^{\delta}$$

with $\delta \in (0,1)$ and I_t denoting education investment per student. This human capital accumulation specification is standard in the literature.[7] Most of the analysis will be conducted under the assumption that $\varepsilon = 1$. However, to check whether the endogenous growth assumption is crucial for the model predictions, we also consider the possibility that $0 \leq \varepsilon < 1$.

As far as the specification of $B_{1,t}$ is concerned, it is useful to consider the limiting case of $\delta = 0$. The equation then postulates that H_t^{ε} and H_{t+1} are proportional to each other, with the factor of proportionality being given by $B_{1,t}$. If the acquisition of human capital is non-rival, then $B_{1,t}$ is independent of the ratio of workers to students. If the acquisition is rival, in contrast, then $B_{1,t}$ is affected by this ratio. In particular, in the rival case, it is reasonable to assume that $B_{1,t} = \tilde{B}_1 \nu_{t+1}^{-\varepsilon}$ for some fundamental constant $\tilde{B}_1$ such that $H_{t+1} = \tilde{B}_1(H_t/\nu_{t+1})^{\varepsilon}$. In the following, we allow for both cases, considering two specifications of $B_{1,t}$. In the first specification, $B_{1,t}$ is fixed or, if $\varepsilon < 1$, growing at some exogenous rate independent of demographics. In the second specification, $B_{1,t} = \tilde{B}_1 \nu_{t+1}^{-\varepsilon(1-\delta)}$. The distinction between the two specifications is irrelevant for most model predictions, but it matters for the growth implications of the demographic transition.

[7] For example, Boldrin and Montes (2005) use the above specification, subject to $\varepsilon = 1$, to compare the allocation in an economy with complete markets for private education financing to the one in an economy without these markets, but with public education and pensions.

2.2 Government

The government taxes labor income in period t at rate $\tau_t+\sigma_t$ and capital income at rate η_t. Revenues collected from workers fund transfers to retirees (the component corresponding to τ_t) as well as education investment (the component corresponding to σ_t). Revenues collected from retirees fund transfers to workers. Denoting the transfer to workers (per worker) by a_t and the transfer to retirees (per retiree) by b_t, we have

$$\begin{aligned} a_t &= s_{t-1}\tilde{R}_t\eta_t p_t/\nu_t, \\ b_t &= w_t(1-x_t)\tau_t\nu_t/p_t, \\ I_t &= w_t(1-x_t)\sigma_t/\nu_{t+1}. \end{aligned}$$

Public investment I_t as well as the transfer payments a_t and b_t must be non-negative. (We exclude lump-sum taxes.) The policy instruments therefore have to satisfy

$$\tau_t, \sigma_t, \eta_t \geq 0 \text{ for all } t. \tag{1}$$

We denote a combination of the three instruments in period t by κ_t, $\kappa_t \equiv (\tau_t, \sigma_t, \eta_t)$.

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 $\beta \in (0, 1)$. Due to the risk of death, workers' effective discount factor therefore equals βp_{t+1}. For analytical tractability, we assume that the period utility function of consumption is logarithmic. Maximizing expected utility, a worker in period t solves the following problem:

$$\begin{aligned} \max_{s_t, x_t} \quad & \ln(c_{1,t}) + v(x_t) + \beta p_{t+1}\ln(c_{2,t+1}) \\ \text{s.t.} \quad & c_{1,t} = w_t(1-x_t)(1-\tau_t-\sigma_t) + a_t - s_t + s_{t-1}R_t(1-p_t)(1-f)/\nu_t, \\ & c_{2,t+1} = s_t\tilde{R}_{t+1}(1-\eta_{t+1}) + b_{t+1}, \end{aligned}$$

where the last term in the first constraint reflects accidental bequests. 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 expected 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:

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

Substituting the expressions for a_t and b_{t+1}, the Euler equation characterizing the optimal savings choice of an *individual* household yields a closed-form solution for the *aggregate* savings function:[8]

$$s_t = z_{t+1}(\tau_{t+1}, \eta_{t+1}) w_t (1 - x_t)(1 - \tau_t - \sigma_t + \phi_t(\eta_t)),$$

where we define the aggregate savings rate

$$z_{t+1}(\tau_{t+1}, \eta_{t+1}) \equiv \frac{\alpha \beta p_{t+1}^2 (1 - f + f/p_{t+1})(1 - \eta_{t+1})}{\alpha(p_{t+1} + \beta p_{t+1}^2)(1 - f + f/p_{t+1})(1 - \eta_{t+1}) + (1-\alpha)\tau_{t+1}} \geq 0$$

and the function

$$\phi_t(\eta_t) \equiv \alpha'(\eta_t p_t (1 - f + f/p_t) + (1 - p_t)(1 - f)).$$

Note that the savings rate depends on subsequent tax rates. (If $\tau_{t+1} > 0$, retirees receive retirement benefits in addition to the return on their savings. This renders the savings rate endogenous, even with logarithmic preferences.) If these tax rates themselves depend on aggregate savings, then the above relation characterizes savings only implicitly. We will return to this point in Section 3.

2.4 Economic Equilibrium

The endogenous state variables at time t are H_t and k_t. To simplify notation, we work with the state variables H_t and $q_t \equiv H_t^{1-\alpha} k_t^{\alpha}$ instead. Combining $k_t = s_{t-1}/\nu_t$ and the aggregate savings function with the dynamic budget constraint and the expressions for factor prices, the equilibrium allocation can recursively be expressed in terms of the following functions of policy instruments:

$$\left.\begin{array}{rcl}
k_{t+1} & = & \mathcal{L}_t\ (1 - \tau_t - \sigma_t + \phi_t(\eta_t))\ z_{t+1}(\tau_{t+1}, \eta_{t+1})/\nu_{t+1} = s_t/\nu_{t+1}, \\
c_{1,t} & = & \mathcal{L}_t\ (1 - \tau_t - \sigma_t + \phi_t(\eta_t))\ (1 - z_{t+1}(\tau_{t+1}, \eta_{t+1})), \\
c_{2,t} & = & \mathcal{L}_t\ \nu_t \alpha' \left((1 - f + f/p_t)(1 - \eta_t) + \frac{\tau_t}{\alpha' p_t}\right), \\
x_t & = & x_t(\tau_t, \sigma_t, \eta_t, \tau_{t+1}, \eta_{t+1}), \\
H_{t+1} & = & B_{1,t}\ H_t^{\varepsilon(1-\delta)}\ (\mathcal{L}_t \sigma_t/\nu_{t+1})^{\delta}, \\
q_{t+1} & = & \left(B_{1,t} H_t^{\varepsilon(1-\delta)} (\mathcal{L}_t \sigma_t/\nu_{t+1})^{\delta}\right)^{1-\alpha} \times \\
 & & (\mathcal{L}_t\ (1 - \tau_t - \sigma_t + \phi_t(\eta_t))\ z_{t+1}(\tau_{t+1}, \eta_{t+1})/\nu_{t+1})^{\alpha}.
\end{array}\right\} \qquad (2)$$

[8] The optimal savings choice of a worker is characterized by the condition

$$s_t \tilde{R}_{t+1}(1 - \eta_{t+1}) + b_{t+1} = \beta p_{t+1} \tilde{R}_{t+1}(1 - \eta_{t+1})[w_t(1 - x_t)(1 - \tau_t - \sigma_t) + a_t - s_t + s_{t-1} R_t (1 - p_t)(1 - f)/\nu_t],$$

implying

$$\frac{s_t \tilde{R}_{t+1}(1 - \eta_{t+1}) + b_{t+1}}{\beta p_{t+1} \tilde{R}_{t+1}(1 - \eta_{t+1})} = w_t(1 - x_t)(1 - \tau_t - \sigma_t + \alpha'[\eta_t p_t (1 - f + f/p_t) + (1 - p_t)(1 - f)]) - s_t.$$

Simplifying the left-hand side of this equation and setting individual and average savings equal to each other, we arrive at the specified expression.

Here, the variable $\mathcal{L}_t$ denotes labor income, and the function $x_t(\cdot)$ is implicitly defined by the transformed first-order condition characterizing labor supply,

$$\begin{aligned} \mathcal{L}_t &\equiv B_0(1-\alpha)q_t(1-x_t)^{1-\alpha} = w_t(1-x_t), \\ v'(x_t)(1-x_t)(1-z_{t+1}(\tau_{t+1},\eta_{t+1})) &= \frac{1-\tau_t-\sigma_t}{1-\tau_t-\sigma_t+\phi_t(\eta_t)}. \end{aligned} \tag{3}$$

Note that labor supply in period t is independent of τ_t and σ_t if $\phi_t(\eta_t) = 0$, i.e., if workers neither receive government transfers nor accidental bequests. In this case, income and substitution effects cancel.

Conditional on initial values for the two endogenous state variables, (H_0, q_0), as well as a sequence of policy instruments, $\{\kappa_t\}_{t=0}^{\infty}$, conditions (2) and (3) fully characterize the equilibrium allocation. Taking logarithms, we can express the laws of motion of the two state variables as

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

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

In the special case of inelastic labor supply, $v'(x) = 0$, $x_t = 0$, and the equilibrium conditions (2) maintain their validity while equation (3) becomes irrelevant.

2.5 Balanced Growth Path

Along a balanced growth path, all tax rates and demographic variables are constant, implying that per-capita labor supply is time-invariant as well. From (2), the growth rates of $k_t, s_t, c_{1,t}$, and $c_{2,t}$ then are equal to the growth rate of q_t.

If $\varepsilon = 1$, the economy displays endogenous growth. The laws of motion for the two state variables in (2) then imply that, along a balanced growth path, the gross growth rate of H_t, γ_H, equals the gross growth rate of q_t. For any time-invariant choice of tax rates, 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 $q_t, k_t, s_t, c_{1,t}$, and $c_{2,t}$. Following this logic, we find

$$\begin{aligned} \gamma_H &= \left(\left(B_0(1-\alpha)(1-x)^{1-\alpha}/\nu\right)^{\delta} B_1^{1-\alpha} \left((1-\tau-\sigma+\phi(\eta))z(\tau,\eta)\right)^{\alpha\delta} \sigma^{\delta(1-\alpha)} \right)^{\frac{1}{1-\alpha(1-\delta)}} \\ & \quad \text{s.t. (3).} \end{aligned} \tag{5}$$

As this equation makes clear, labor income taxes depress growth because they lower disposable income of workers (the effect captured by the expression $1-\tau-\sigma+\phi(\eta)$), as do expected future retirement benefits because they lower the savings rate ($z(\tau,\eta)$ is decreasing in its first argument).[9] At the same time, education investment fosters human

[9] Accidental bequests increase the disposable income of workers and thus, growth.

capital accumulation and thus, growth (the effect captured by σ in the last term), in line with the empirical evidence (Blankenau et al., 2007). Capital income taxes η have an ambiguous effect on growth because they increase disposable income of workers but have a negative effect on their savings rate. If $\varepsilon < 1$, the economy does not grow endogenously but converges to a steady state (unless B_1 grows exogenously).

Irrespective of the value of ε, physical capital along its long-run growth path satisfies $k_{t+1} = \mathcal{L}_t(1-\tau-\sigma+\phi(\eta))\ z(\tau,\eta)/\nu$. Since k_t grows at the gross rate γ_H, it follows that

$$\begin{aligned}\left(\frac{H_t}{k_t}\right)^{1-\alpha} &= \frac{\gamma_H \nu}{B_0(1-\alpha)(1-x)^{1-\alpha}(1-\tau-\sigma+\phi(\eta))\ z(\tau,\eta)} \text{ s.t. (3)},\\ R &= \frac{\alpha\gamma_H\nu}{(1-\alpha)(1-\tau-\sigma+\phi(\eta))\ z(\tau,\eta)} \text{ s.t. (3)}.\end{aligned}$$

We will use these relations for calibration purposes.

3 Politico-Economic Equilibrium

We assume that retirees, workers and students vote on candidates whose electoral platforms specify values for the policy instruments, κ_t.[10] Voters do not only support a candidate 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. Moreover, 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, implying that equilibrium policy platforms smoothly respond to changes in the demographic structure. This stands in sharp contrast to the "median-voter" setup where, in a model with only a few generations, an infinitesimal change in the demographic structure has implausibly large effects on policy outcomes if it alters the cohort the median voter is associated with.

In the Nash equilibrium of the game with two candidates choosing platforms to maximize their expected vote shares, both candidates propose the same policy platform.[11] 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. Those groups that care the most about policy platforms rather than other candidate characteristics are the most likely to shift their support from one candidate to the other in response to small changes in the proposed platforms. In equilibrium, such groups of "swing voters" thus gain in political influence and tilt policy in their own favor. If all voters are equally responsive to changes in the policy platforms, electoral competition implements the utilitarian optimum with respect to voters.

[10] It is straightforward to analyze the situation where only a subset of all living agents participates in the vote; see below.

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

In the context of our model, the probabilistic-voting assumption implies that the objective function maximized by the political process attaches weight to the welfare of all currently living agents. In particular, it attaches positive weight to the welfare of retirees even if these are outnumbered by workers and students. This implication is very realistic. After all, old voters appear to exert as strong a political influence per capita as younger voters when the salient issue of intergenerational transfers is at stake (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 current policy choices on future policy outcomes. Under the Markov assumption, future leisure and policy choices are functions of the fundamental state variables only, $x_{t+1} = \tilde{x}_{t+1}(H_{t+1}, q_{t+1})$ and $\kappa_{t+1} = \tilde{\kappa}_{t+1}(H_{t+1}, q_{t+1})$. (The state variables include demographic variables, thus the time indices of the policy functions.) If the policy functions are independent of (H, q), $\kappa_{t+1} = \tilde{\kappa}_{t+1}$, then (3) implies that the leisure function is independent of (H, q) as well, $x_{t+1} = \tilde{x}_{t+1}$, and both the aggregate savings function and 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 (H, q). We derive the equilibrium choice of policy instruments under this conjecture and show that this choice does not depend on (H, q), thereby verifying the conjecture.

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

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

The political objective function $W_t(\cdot)$ depends on the endogenous state variables (as well as the exogenous ones, thus the time index), the contemporaneous policy instruments, and the anticipated values of policy instruments and leisure in the following period. In particular,

$$\begin{aligned} W_t(H_t, q_t, \kappa_t; \tilde{\kappa}_{t+1}, \tilde{x}_{t+1}) \quad &\equiv \quad \omega p_t \ln(c_{2,t}) + \nu_t[\ln(c_{1,t}) + v(x_t) + \beta p_{t+1} \ln(c_{2,t+1})] \\ &\qquad + \psi \nu_t \nu_{t+1} \beta[\ln(c_{1,t+1}) + v(x_{t+1}) + \beta p_{t+2} \ln(c_{2,t+2})] \\ \text{s.t.} \qquad &\qquad (2), (3);\ H_t, q_t \text{ given};\ \kappa_{t+1} = \tilde{\kappa}_{t+1},\ x_{t+1} = \tilde{x}_{t+1}. \end{aligned}$$

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

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

$$
\begin{aligned}
W_t(\cdot) \;=\; & \omega p_t \ln[(1-x_t)^{1-\alpha}((1-f+f/p_t)(1-\eta_t)+\tau_t/(\alpha' p_t))] \\
& + \nu_t\{\ln[(1-x_t)^{1-\alpha}(1-\tau_t-\sigma_t+\phi_t(\eta_t))] + \\
& \quad v(x_t) + \beta p_{t+1} \ln[(1-x_t)^{(1-\alpha)(\delta(1-\alpha)+\alpha)}(1-\tau_t-\sigma_t+\phi_t(\eta_t))^{\alpha}\sigma_t^{\delta(1-\alpha)}]\} \\
& + \psi\nu_t\nu_{t+1}\beta\{\ln[(1-x_t)^{(1-\alpha)(\delta(1-\alpha)+\alpha)}(1-\tau_t-\sigma_t+\phi_t(\eta_t))^{\alpha}\sigma_t^{\delta(1-\alpha)}] + \\
& \quad \beta p_{t+2} \ln[(1-x_t)^{(1-\alpha)(\delta\varepsilon(1-\delta)(1-\alpha)+(\delta(1-\alpha)+\alpha)^2)}(1-\tau_t-\sigma_t+\phi_t(\eta_t))^{\alpha(\delta(1-\alpha)+\alpha)} \\
& \qquad \sigma_t^{\delta(1-\alpha)(\varepsilon(1-\delta)+\delta(1-\alpha)+\alpha)}]\} \\
& + \text{t.i.p. s.t. (3)},
\end{aligned}
$$

where t.i.p. denotes terms that are unaffected by contemporaneous policy choices (under the conjecture). Notice that the contemporaneous policy instruments do not interact with the state variables H_t or q_t. This confirms the conjecture that the equilibrium policy functions are independent of these state variables.[12] We first consider the case with inelastic labor supply.

3.1 Inelastic Labor Supply

If labor supply is inelastic then x_t is fixed such that

$$
\begin{aligned}
W_t(\cdot) \;\simeq\; & \omega p_t \ln[((1-f+f/p_t)(1-\eta_t)+\tau_t/(\alpha' p_t))] + \\
& \nu_t\{\ln[(1-\tau_t-\sigma_t+\phi_t(\eta_t))] + \beta p_{t+1}\ln[(1-\tau_t-\sigma_t+\phi_t(\eta_t))^{\alpha}\sigma_t^{\delta(1-\alpha)}]\} + \\
& \psi\nu_t\nu_{t+1}\beta\{\ln[(1-\tau_t-\sigma_t+\phi_t(\eta_t))^{\alpha}\sigma_t^{\delta(1-\alpha)}] + \\
& \quad \beta p_{t+2}\ln[(1-\tau_t-\sigma_t+\phi_t(\eta_t))^{\alpha(\delta(1-\alpha)+\alpha)}\sigma_t^{\delta(1-\alpha)(\varepsilon(1-\delta)+\delta(1-\alpha)+\alpha)}]\}.
\end{aligned}
$$

Disregarding the inequality constraints in (1), the effects of marginal policy changes are linearly dependent.[13] We address this indeterminacy issue later. For now, we focus on the equilibrium values for τ_t and σ_t conditional on a given choice of η_t.

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

$$
\frac{1+\alpha\beta p_{t+1}+\psi\nu_{t+1}\alpha\beta[1+\beta p_{t+2}(\delta(1-\alpha)+\alpha)]}{1-\tau_t-\sigma_t+\phi_t(\eta_t)} = \beta\delta(1-\alpha)\frac{p_{t+1}+\psi\nu_{t+1}[1+\beta p_{t+2}(\varepsilon(1-\delta)+\delta(1-\alpha)+\alpha)]}{\sigma_t}.
$$

The left-hand side of this equation reflects the marginal cost of an increase in σ_t. By reducing disposable income of workers and depressing capital accumulation, an increase in σ_t lowers workers' second- and third-period consumption; the associated welfare effects

[12] This result is due to the logarithmic preference assumption. However, we conjecture that the quantitative implications for equilibrium tax rates would be very similar if we generalized preferences to the CRRA class, as is the case in a setup without human capital, education, or endogenous growth (Gonzalez-Eiras and Niepelt, 2005).

[13] In particular, $\partial W/\partial\eta_t = -\alpha' p_t(1-f+f/p_t)\partial W/\partial\tau_t$.

are proportional to $1 + \alpha\beta p_{t+1}$. Lower capital accumulation also hurts students whose next-period labor income and savings falls, triggering (discounted) welfare losses that are proportional to $1 + \alpha\beta p_{t+2}$. Students additionally suffer from the fact that lower output in period $t+1$ translates into reduced education investment and thus, lower output and income in period $t + 2$; the corresponding welfare implications are proportional to $\beta p_{t+2}\delta(1 - \alpha)$.

The right-hand side of the equality reflects the marginal benefits of an increase in σ_t. These benefits work through higher productivity in the subsequent period as reflected by the term $\beta\delta(1 - \alpha)$. Workers benefit because the productivity increase affects all their sources of income during retirement. Students benefit from increased second- and third-period consumption, both directly (the welfare effect is proportional to $1 + \alpha\beta p_{t+2}$) and indirectly, through induced education investment in period $t + 1$ (proportional to $\beta p_{t+2}\delta(1 - \alpha)$) and due to the dynamic human capital externality that pays off in period $t + 2$ (proportional to $\beta p_{t+2}\varepsilon(1 - \delta)$).

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

Consider next the choice of τ_t. Since the marginal *cost* 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 coincide in equilibrium. Disregarding the inequality constraints in (1), this implies

$$\frac{\omega p_t}{\nu_t} \frac{1-\alpha}{\alpha p_t(1-f+f/p_t)(1-\eta_t)+(1-\alpha)\tau_t} = \beta\delta(1-\alpha)\frac{p_{t+1}+\psi\nu_{t+1}[1+\beta p_{t+2}(\varepsilon(1-\delta)+\delta(1-\alpha)+\alpha)]}{\sigma_t},$$

where the left-hand side reflects the welfare effect from increased consumption of retirees due to higher retirement benefits. Conditional on η_t, this second first-order condition defines a positive relationship between τ_t and σ_t. Intuitively, both retirement benefits (higher τ_t) and education investment (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.

Disregarding the inequality constraints in (1), we can solve the two first-order conditions for the equilibrium tax rates $\tau_t^W(\eta_t)$ and σ_t^W. (The resulting explicit solutions are rather complicated and we omit them here.) Only the former of the two equilibrium tax rates is conditional on the value of η_t. The equilibrium tax rate for transfers, $\tau_t^W(\eta_t)$, increases in the relative per-capita weight of retirees in the political process, ω, and decreases in the relative per-capita weight of students, ψ. It also decreases in the size of the younger cohorts, ν_t, ν_{t+1}, patience, β, the strength of the dynamic human capital externality, ε (as long as $\psi > 0$), and future longevity, p_{t+1}, p_{t+2}. The elasticity of $\tau_t^W(\eta_t)$ with respect to contemporaneous longevity, p_t, is positive for f sufficiently close to unity, i.e., if accidental bequests are small. The tax rate funding education, σ_t^W, is strictly positive and independent of the importance of accidental bequests, f. It increases in the number and the relative political influence of those that care about the future, $\nu_t/p_t, \nu_{t+1}, \psi, \omega^{-1}$, patience, β, the strength of the human capital externality, ε, and the longevity of students, p_{t+2}.

For any parameter constellation, the equilibrium tax rate σ_t^W and a continuum of policies $(\tau_t^W(\eta_t), \eta_t)$ satisfy the constraints (1). This can be seen by noting that (i) σ_t^W is always positive and (ii) $\tau_t^{W\prime}(\eta_t) > 0$ such that raising η_t allows to satisfy all non-negativity constraints in (1). The indeterminacy with respect to η_t arises because ex post (after savings have been chosen), labor and capital income taxes have exactly the opposite effect—only their difference, $\tau_t^W(\eta_t) - \eta_t\alpha' p_t(1-f+f/p_t)$, affects the allocation. Ex ante, in contrast, the two tax rates have different equilibrium implications because the capital income tax rate affects the savings rate (as long as $\tau_{t+1} \neq 0$).

We eliminate the policy indeterminacy by focusing on the policy that minimizes the capital income tax rate, η_t, conditional on satisfying (1). This refinement is motivated by the observation that the minimal feasible capital income tax rate maximizes the economy's growth rate (if $\varepsilon = 1$) or capital-labor ratio and output (if $\varepsilon < 1$).[14] The equilibrium tax rates thus are given by

$$\kappa_t^W = (\tau_t^W, \sigma_t^W, \eta_t^W), \ \tau_t^W = \max[0, \tau_t^W(0)], \ \eta_t^W = \max[0, \eta_t^W(0)]$$

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

Subject to the refinement, the equilibrium policy functions are unique in the limit of the finite horizon economy. To see this, note that the consumption of workers and retirees in the final period, T, is given by

$$\begin{aligned} c_{1,T} &= \mathcal{L}_T(1 - \tau_T + \phi_T(\eta_T)), \\ c_{2,T} &= \mathcal{L}_T \ \nu_T \alpha' \left((1-f+f/p_T)(1-\eta_T) + \frac{\tau_T}{\alpha' p_T} \right), \end{aligned}$$

respectively. Tax rates are set to achieve $\omega = \frac{c_{2,T}}{c_{1,T}}$. 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 (H_T, q_T). Moving to period $T-1$, the policy functions for σ_{T-1}, τ_{T-1}, and η_{T-1} therefore are independent of (H_{T-1}, q_{T-1}) as well. The result then follows by induction.

3.2 Elastic Labor Supply

With elastic labor supply, x_t depends on contemporaneous and future policy instruments. As a consequence, the first-order conditions with respect to tax rates cease to be linearly dependent and the equilibrium policy is determinate. At the same time, the first-order conditions do not generally admit closed-form solutions any longer. Recall from (3), however, that x_t is independent of τ_t and σ_t if $\phi_t(\eta_t) = 0$ (such that income and substitution effects cancel). In this case, the first-order conditions with respect to τ_t and σ_t remain

[14] The equilibrium value of $\tau_t^W(\eta_t) - \eta_t\alpha' p_t(1-f+f/p_t)$ does not vary with η_t. As a consequence, the capital income tax rate affects growth or output only through the savings rate $z(\cdot)$, see equation (5) for the case $\varepsilon = 1$.

unchanged relative to the case with inelastic labor supply. As a consequence, the equilibrium tax rates in the case with inelastic labor supply, κ_t^W, also constitute the equilibrium tax rates in the case with elastic labor supply, and the endogenous nature of labor supply provides another rationale for the refinement we adopted in the case with inelastic labor supply.

Independently of $\phi_t(\eta_t)$, the equilibrium policy functions are unique in the limit of the finite horizon economy.

3.3 Efficiency

Before turning to the quantitative predictions of the model, we briefly discuss some of its normative implications. We start by comparing the politico-economic equilibrium to the allocation implemented by a Ramsey planner who—in contrast to political decision makers—can commit to future policy choices and 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:[15]

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

where

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

We discuss the solution to this program in Appendix A.

The equilibrium conditions in (2) imply that two allocations necessarily differ unless they are supported by the same tax sequences. As a consequence, the politico-economic equilibrium can only be supported by a Ramsey policy subject to intergenerational discount factor ρ if this policy specifies tax rates as time-invariant functions of the demographic structure. This is generally not the case since, as we discuss in Appendix A, the Ramsey policy generally is not time consistent. Even if the Ramsey policy *is* time consistent, the politico-economic equilibrium allocation coincides with the allocation implemented by a Ramsey planner only for a unique pair of political weights, $(\hat{\omega}, \hat{\psi})(\rho)$ say. The politico-economic equilibrium therefore generically differs from the allocation implemented by a Ramsey planner.

If the politico-economic equilibrium differs from the Ramsey equilibrium, is the former at least consumption and production efficient? Or does the politico-economic equilibrium 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

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

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 $\beta\gamma_H\nu_{i+1}u'(c_{2,i+1}) \geq u'(c_{1,i})$. Using the households' consumption Euler equation, the condition for *consumption efficiency* along a balanced growth path 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 along a balanced growth path 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 B, we derive as a criterion for *production efficiency* the requirement that

$$\frac{\delta}{\sigma} < \frac{R}{\gamma_H \nu} < \frac{\delta}{\sigma(1-\varepsilon(1-\delta))}.$$

4 Quantitative Implications

We now turn to a quantitative assessment of our framework. We calibrate the model to match stylized features of the U.S. economy, in particular the GDP shares flowing into retirement benefits on the one hand and public education on the other. We then analyze how the forecasted changes in the demographic structure affect equilibrium tax rates, budget shares and per-capita growth. We also compare the politico-economic equilibrium along a balanced growth path with the corresponding Ramsey allocation (under the assumption that $\rho = \beta\nu$) and assess consumption and production efficiency of the equilibrium. One period in the model corresponds to 25 years in the data.

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: $\alpha = 0.2815$. In the model, the elasticity of earnings with respect to education spending equals δ (holding aggregate human capital and labor supply fixed).[16] 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, as a robustness check. Without loss of generality, we normalize $B_0 = 1$. In the baseline calibration, we assume $\varepsilon = 1$ (endogenous growth).

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

$$\frac{w_t}{H_t}h_t(1-x_t) = (1-\alpha)B_0 s_{t-1}^{\alpha}[H_t\nu(1-x_t)]^{-\alpha}h_t(1-x_t) \ \text{ s.t. } h_t = B_{1,t}H_{t-1}^{\varepsilon(1-\delta)}(\nu_t\nu_{t+1}i_{t-1})^{\delta}.$$

To calibrate β, ω, ψ and f, we match four empirical moments.[17] In particular, we impose that the model-predicted GDP shares of retirement benefits and education spending,

$$\begin{aligned} \text{share of net transfers to retirees} &= \text{sh}_t^R \equiv (1-\alpha)(\tau_t - \eta_t \alpha' p_t (1 - f + f/p_t)), \\ \text{education share} &= \text{sh}_t^I \equiv (1-\alpha)\sigma_t, \end{aligned}$$

respectively, match their empirical counterparts in the years 1975 (sh_t^R) and 2005 (both sh_t^R and sh_t^I).[18] Moreover, we impose that the model's interest rate along the balanced-growth-path, R, matches the data, conditional on the observed long-run growth rate (corresponding to γ_H in the model).[19]

Finally, we associate ν_t and p_t with data about population growth and life expectancy at retirement age. In particular, we associate ν_t with the gross population growth rate of the non-retired population since, in the model,

$$\frac{\#\text{students}_t + \#\text{workers}_t}{\#\text{students}_{t-1} + \#\text{workers}_{t-1}} = \nu_t \frac{1+\nu_{t+1}}{1+\nu_t} \approx \nu_t.$$

Using data from the 2007 OASDI Trustees Report[20] (Table V.A2, estimates and projections, middle series), we compute ν_t for the years $t = 1975, 1980, \ldots, 2085$. From the same source (Table V.A4, intermediate), we compute p_t as the average Cohort Life Expectancy at Age 65, normalized by the period length of 25 years.[21] For example, for the year 2000, this procedure yields $\nu_t = \nu_{2000} = 1.2562$, $\nu_{t+1} = \nu_{2025} = 1.1438$, $p_t = p_{2000} = 0.7220$, and $p_{t+1} = p_{2025} = 0.7110$. Figure 1 displays the time series for ν_t and p_t. Based on these series and the calibrated parameters, the model generates five overlapping sequences of predictions for the endogenous variables, each with a step size of 25 years. The first such sequence covers the years 1975, 2000, 2025, ...; the second one the years 1980, 2005,

[17] There is no need to calibrate the parameter $\tilde{B}_1$ for the analysis of the transition dynamics. The parameter will be calibrated for the analysis of the balanced growth path.

[18] The empirical shares (we use 5-year centered averages, except for the year 2005) are given by $\text{sh}_{1975}^R = 0.0492$, $\text{sh}_{2005}^R = 0.0682$, and $\text{sh}_{2005}^I = 0.0533$. Retirement benefits are taken from `http://www.ssa.gov/OACT/STATS/table4a4.html` (social security and medicare), education expenditures from `http://www.bea.gov/national/nipaweb` (investment plus current expenditure).

[19] 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.

[20] See `http://www.ssa.gov/OACT/TR/TR07/V_demographic.html`.

[21] We adjust the projected life expectancy progressively by up to two years in order to account for the foreseen increase of the retirement age over the next decades, see `http://ssa.gov/pubs/retirechart.htm`. Due to data limitations, we do not correct the population growth series in the same manner.

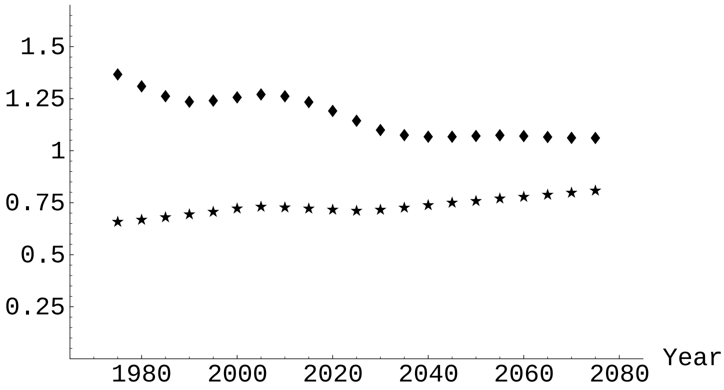

Figure 1: ν_t ($\diamond$) and p_t ($\star$)

2030, ...; etc. When reporting model predictions about transition dynamics, we merge the five sequences.[22]

We compute the predictions for the transition dynamics under the assumption of inelastic labor supply, fixing x_t at the value 2/3. Table 1 summarizes the baseline calibration. Notice that both relative political weights, ω and ψ, are calibrated to be close to unity. Our quantitative results therefore do not hinge on the presumption that some groups are much more influential in the political process than others. Notice also that f is calibrated to be close to zero, suggesting a low degree of annuitization (in line with empirical evidence) and a strong role for accidental bequests.[23] Due to these accidental bequests, an increase in life expectancy (p) has a negative income effect on workers (as they receive fewer bequests). By partly compensating for the increasing number and therefore, voice of retirees, this effect slows down the rise of τ_t.

Table 1: Baseline Calibration

Exogenous parameters:	$\alpha = 0.2815$, $\delta = 0.1200$, $B_0 = 1$, $\varepsilon = 1$
Implied parameters:	$\beta = 0.5956$, $\omega = 0.9273$, $\psi = 0.7351$, $f = 0.0638$

Notes: See explanations in the text. On an annual basis, β equals 0.9795.

Figures 2 and 3 display the model's predictions for the GDP shares sh_t^R and sh_t^I over the years 1975 to 2075. Conditional on the (matched) values sh_{1975}^R and sh_{2005}^R, the model

[22]To compute model predictions up to the year 2075, we need population growth predictions until 2125 and predictions for life expectancy until 2150. Since projections for population growth and life expectancy are available up to the year 2085 only, we assume that the two demographic series remain constant after that year.

[23]See Davidoff, Brown and Diamond (2003) for possible explanations of the low degree of annuitization observed in the data.

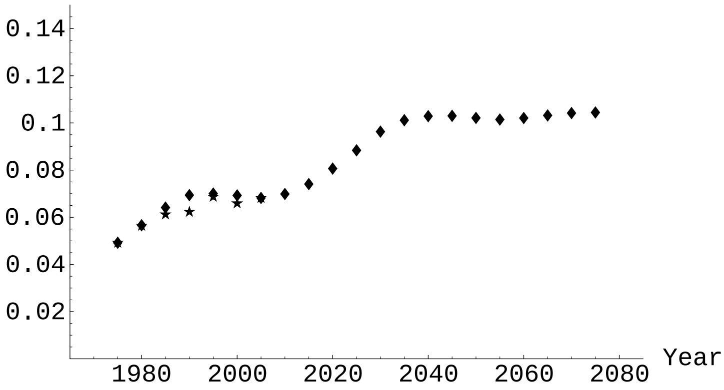

Figure 2: sh_t^R: Model predictions ($\diamond$) and data ($\star$)

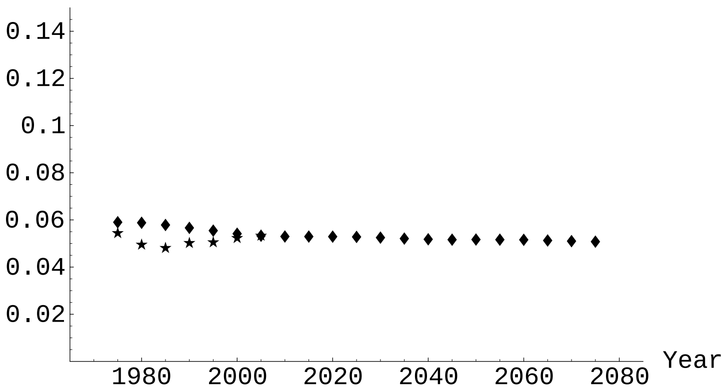

Figure 3: sh_t^I: Model predictions ($\diamond$) and data ($\star$)

does a good job at capturing the observed slowdown of the transfer share in the 1990s. As far as public eduction is concerned, conditional on the (matched) value of sh_{2005}^I, the model captures the observed trend decline in the share, but not the more short run fluctuations. Turning to the out-of-sample forecasts, the model predicts the demographic transition to have a negative effect on investment and a large, positive effect on benefits. In particular, it predicts the GDP share of retirement benefits to rise by 53 percent between the years 2005 and 2075, the share of education spending to decrease by 5 percent over the same period, and the GDP share of both government spending components to rise by 28 percent. Since the tax rate η_t remains zero throughout the sample period, the tax rates τ_t and σ_t evolve proportionally to sh_t^R and sh_t^I, respectively.[24]

These predictions are robust to changes in parameters and empirical moments used to calibrate the model, see Table 2. In particular, the predicted effect of the demographic transition on the two budget shares remains essentially unchanged if ε is reduced below

[24]The implied tax rates for the year 2005 are $\tau_{2005} = 0.0949$, $\sigma_{2005} = 0.0742$, and $\eta_{2005} = 0$.

unity such that the economy displays exogenous rather than endogenous growth. While the reduction in ε goes hand in hand with a change in ψ, to match the *levels* of the budget shares, the *elasticities* of the budget shares with respect to demographic change are essentially unaffected. This finding is intuitive. After all, the political process only internalizes the short-term repercussions of policy on output, and these short-term repercussions are not strongly affected by the dynamic human capital externality.

Table 2: Robustness Checks

	Baseline	Alternative Calibration Assumptions					
		ε	δ	$\sqrt[25]{R}$		$\sqrt[25]{\gamma_H}$	
		0.5	0.155	1.04	1.05	1.01	1.02
	Implied Parameter Values						
β	0.5956	0.5957	0.5953	0.7749	0.5656	0.5489	0.7574
ω	0.9273	0.9418	0.8960	0.9393	0.9255	0.9246	0.9380
ψ	0.7351	0.8703	0.4373	0.4270	0.8103	0.8565	0.4493
	Implied Budget Shares in the Year 2075						
sh^R_{2075}	0.1044	0.1041	0.1050	0.1039	0.1045	0.1046	0.1040
sh^I_{2075}	0.0508	0.0506	0.0516	0.0515	0.0507	0.0506	0.0514

Note: See explanations in the text.

Figure 4 shows that fertility rather than longevity is the driving force behind the rising GDP share of retirement benefits. The figure displays the evolution of sh^R_t in the baseline scenario and under the assumption that either ν_t or p_t remains constant from the year 2010 onwards. Fixing population growth (that is, keeping the ratio of students to workers higher than projected) implies a strong reduction of sh^R_t relative to the baseline, while fixing life expectancy (that is, keeping life expectancy lower than projected) has a small positive effect on the benefit share, because of higher bequests.[25] For the GDP share of public education spending, the effects are smaller and have the reverse sign, see Figure 5.

Next, we turn to the long-run growth implications of the projected demographic change, assuming that $\varepsilon = 1$. We consider the case with elastic labor supply, since

[25] Olshansky, Passaro, Hershow, Layden, Carnes, Brody, Hayflick, Butler, Allison and Ludwig (2005, p. 1138) argue that, due to obesity, "the steady rise in life expectancy during the past two centuries may soon come to an end." Before that background, the constant-p case may not only be interpreted as an analytical device, but also as a plausible scenario.

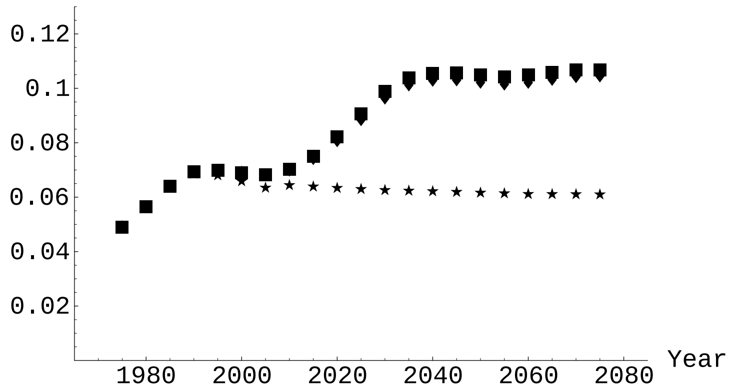

Figure 4: sh_t^R: Baseline scenario ($\diamond$), constant ν_t after 2010 ($\star$), constant p_t after 2010 ($\Box$)

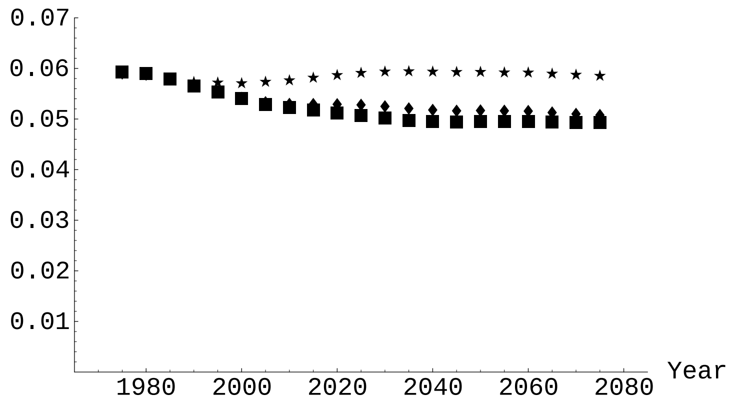

Figure 5: sh_t^I: Baseline scenario ($\diamond$), constant ν_t after 2010 ($\star$), constant p_t after 2010 ($\Box$)

even small changes in labor supply may have significant effects on the growth rate. The calibration of the model now proceeds under the assumption that the economy grows along a balanced growth path in the year 2005, the 2005-BGP say. There are five changes relative to the calibration strategy described before. First, we assume $v(x) \equiv m \ln(x)$ and calibrate m by imposing that $x_{2005} = 2/3$. Second, we calibrate $B_{1,2005}$ under the assumption that γ_H as given in (5) equals the observed long-run growth rate of $1.0126^{25} \approx 1.3676$. Third, in order to match the GDP shares of public education and social security benefits, we evaluate the expressions for the shares at the numerical solutions for tax rates under the assumption of elastic labor supply, rather than the closed-form solutions under the assumption of inelastic labor supply, κ_t. Fourth, we match the GDP shares in the year 2005 only. Rather than also matching sh^R_{1975} (as we did before), we fix f at the value calibrated before.[26] Finally, we impose that all demographic variables remain constant at their year-2005 values. Table 3 summarizes the results of the calibration. The corresponding allocation is production and consumption efficient.

Table 3: Calibration for BGP, elastic labor

Exogenous parameters:	$\alpha = 0.2815$, $\delta = 0.1200$, $B_0 = 1$, $\varepsilon = 1$, $f = 0.0638$
Implied parameters:	$\beta = 0.5265$, $\omega = 0.9474$, $\psi = 0.8326$, $m = 2.3066$, $B_{1,2005} = 2.5611$

Notes: See explanations in the text. On an annual basis, β equals 0.9747.

Conditional on this calibration consistent with the 2005-BGP, we derive predictions for the balanced growth path at the end of the century, the 2075-BGP say, under the assumption that the two demographic variables ν and p permanently change to their projected values in the year 2075. The resulting new allocation remains production and consumption efficient, but other features of the balanced growth path change. In particular, the GDP share of social security rises from 6.8 percent to over 11 percent, and the GDP share of public education falls from 5.3 to 4.7 percent. Correspondingly, the social-security tax rate τ rises and the tax rate funding public education, σ, falls (tax rate η remains zero). Note that the predicted budget shares along the 2075-BGP and the budget shares predicted to prevail at the end of the transition dynamics are close to each other. Since only the former are computed under the assumption of elastic labor supply, we conclude that the transition dynamics analysis is robust to endogenizing the labor-leisure tradeoff.

In spite of the higher labor income tax rate along the 2075-BGP, labor supply $1 - x$ is predicted to rise by about 2.7 percent, due to a negative wealth effect from fewer accidental bequests and a higher savings rate; both these effects originate in the higher

[26] The results do not hinge on this assumption. The growth implications of the demographic transition are nearly unchanged if f is fixed at different values, 0 or 0.3 say.

life expectancy, see equation (3). This prediction of increased hours per worker runs counter to the fear, expressed in the policy debate, that higher Social Security taxes would be associated with depressed labor supply. The model therefore does not only highlight the dangers of a partial equilibrium analysis, but also the importance of differentiating between changes in fertility (driving tax rates) and longevity (driving labor supply) when considering the implications of demographic ageing.

The effect of these changes on the growth rate depends on the extent to which human-capital accumulation is rival (see the discussion on page 6). If human-capital accumulation is non-rival such that $B_{1,t}$ is independent of demographics, productivity growth on the new balanced growth path increases slightly, to 1.3857. This corresponds to an annual productivity growth rate of 1.31 percent, compared with 1.26 percent along the 2005-BGP. Reduced capital dilution is at the source of this acceleration of productivity growth: Holding the economy's investment rate constant, the slowdown in population growth allows for faster capital accumulation per worker and thus, stronger growth. But this positive effect on productivity is partly undone by the negative consequences of the reallocation of government spending. If tax rates and spending shares were held constant at their 2005-BGP values, growth would accelerate even faster, to 1.4208 or 1.41 percent annually. Two thirds (10 basis points) of the potential gain in productivity growth due to reduced capital dilution (15 basis points) are therefore undone because the political process shifts resources from education to transfers.

If human-capital accumulation is rival, in contrast, $B_{1,t}$ increases as a consequence of the fall in population growth, resulting in significant physical and human capital deepening and a surge of productivity growth to around 1.9 percent annually. But the "growth cost" due to the reallocation of government spending remains roughly unchanged: If policy instruments were fixed at their 2005-BGP values, the annual growth rate would accelerate by 10 basis points more.

It is also instructive to compare the equilibrium allocation with the allocation supported by the Ramsey policy subject to $\rho = \beta\nu$, corresponding to a situation where all current and future generations vote once and for all. We consider both the case with endogenous growth and with exogenous growth. In parallel with the procedure explained earlier, we let $\varepsilon = 1.0$ or $\varepsilon = 0.5$ (endogenous or exogenous growth, respectively) and calibrate the remaining model parameters such that the politico-economic equilibrium along the 2005-BGP matches the targeted budget shares, interest rate, growth rate, and labor supply. For each of the two sets of parameters, we then compute the Ramsey policy.

With endogenous growth, the Ramsey policy features a much higher budget share for education, $\mathrm{sh}^I = 0.1536$ (compared with 0.0533 in politico-economic equilibrium or in the data), and no retirement benefits at all, $\mathrm{sh}^R = 0$ (compared with 0.0682). Although the labor income tax rate under the Ramsey policy is higher than in politico-economic equilibrium, growth is substantially stronger under the Ramsey policy (1.80 rather than 1.26 percent per year). This increase in the productivity growth rate is caused by higher labor supply (0.3478), an increased savings rate (due to reduced retirement benefits), and the much higher education spending.

With exogenous growth, the case for public investment spending becomes weaker. The Ramsey policy therefore reduces spending for public education to $\mathrm{sh}^I = 0.0936$ but maintains zero retirement benefits. These budget shares continue to be higher and lower, respectively, than the corresponding shares in politico-economic equilibrium.

5 Conclusion

What are the demographic and political factors that shape government budgets? And how does the structure of government budgets affect macroeconomic aggregates? The model presented in this paper proposes a rich, yet tractable framework to answer these questions. Building on a standard overlapping generations model, it endogenizes both economic and political choices, generating predictions for consumption, labor supply, labor- and capital-income taxes, pensions, human and physical capital accumulation as well as productivity growth.

We apply the model to study the effects of the projected trend decrease in fertility combined with the trend increase in longevity in the United States. Several robust predictions emerge. Most importantly, the projected demographic transition, in particular the decrease in fertility, triggers a reallocation of government spending from productive public education to unproductive intergenerational transfers. Labor income taxes *and* hours worked per capita rise, the latter because upward pressure on labor supply due to rising longevity more than outweighs the negative effect from tax distortions. The reallocation of government spending triggers a decrease of per-capita growth (if human capital externalities are sufficiently strong to sustain endogenous growth) or per-capita income (if they are not). In the former case, the decrease in annual productivity growth relative to a scenario without any changes to policy equals roughly ten basis points, independently of the extent to which human capital accumulation is rival.

These findings do not only undermine the notion among many policy makers that the political process will implement measures to boost productivity in order to "outgrow" the burden imposed by population ageing. A comparison between the politico-economic equilibrium and the allocation implemented by a Ramsey planner also suggests that limited political participation causes potentially severe losses if measured by a utilitarian social welfare function that accounts for all current and future generations. At the source of these losses is the fact that political decision makers optimize over a short horizon and therefore tend to invest too little, in particular if the dynamic human capital externality is strong. As noted in the introduction, this underinvestment problem cannot be overcome by letting voters finance investment expenditures out of government debt. The balanced-budget assumption maintained throughout the analysis therefore is not restrictive.

We have assumed that human capital production does not employ labor. While this assumption is useful for our purposes, one might want to relax it in order to analyze other issues, for example the reallocation of labor in response to population ageing or the effect of interest groups in the public education sector. We leave such an analysis with "teachers" for future research.

References

Alesina, A. and Rodrik, D. (1994), 'Distributive politics and economic growth', *Quarterly Journal of Economics* **109**(2), 465–490.

American Society of Civil Engineers (2005), Report card for America's infrastructure, downloadable report, http://www.asce.org/reportcard.

Bassetto, M. and Sargent, T. J. (2006), 'Politics and efficiency of separating capital and ordinary government budgets', *Quarterly Journal of Economics* **121**(4), 1167–1210.

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

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.

Blankenau, W. F., Simpson, N. B. and Tomljanovich, M. (2007), 'Public education expenditures, taxation, and growth: Linking data to theory', *American Economic Review, Papers and Proceedings* **97**(2), 393–397.

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

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

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.

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

Davidoff, T., Brown, J. R. and Diamond, P. A. (2003), Annuities and individual welfare, Working Paper 9714, NBER, Cambridge, Massachusetts.

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

Docquier, F., Paddison, O. and Pestieau, P. (2007), 'Optimal accumulation in an endogenous growth setting with human capital', *Journal of Economic Theory* **134**(1), 361–378.

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.

Gonzalez-Eiras, M. and Niepelt, D. (2005), Sustaining social security, Working Paper 1494, CESifo.

Gonzalez-Eiras, M. and Niepelt, D. (2007), Economic and politico-economic equivalence of fiscal policies. Mimeo, Study Center Gerzensee.

Gonzalez-Eiras, M. and Niepelt, D. (2008), 'The future of social security', *Journal of Monetary Economics* (forthcoming).

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

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.

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

Mulligan, C. B. and Sala-i-Martin, X. (2004), 'Internationally common features of public old-age pensions, and their implications for models of the public sector', *Advances in Economic Analysis & Policy* **4**(1).

Olshansky, S. J., Passaro, D. J., Hershow, R. C., Layden, J., Carnes, B. A., Brody, J., Hayflick, L., Butler, R. N., Allison, D. B. and Ludwig, D. S. (2005), 'A potential decline in life expectancy in the United States in the 21st century', *New England Journal of Medicine* **352**(11), 1138–1145.

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

Persson, T. and Tabellini, G. (1994), 'Is inequality harmful for growth?', *American Economic Review* **84**(3), 600–621.

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

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

Poterba, J. M. (1997), 'Demographic structure and the political economy of public education', *Journal of Policy Analysis and Management* **16**(1), 48–66.

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

A Ramsey Program

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

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

Since we focus on the Ramsey policy along a balanced growth path, we assume that $p_s = p$ for all $s \geq t$.

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 of one of the policy instruments, φ_i say, with $\varphi_i \in \{\tau_i, \sigma_i, \eta_i\},\ i \geq t$. The direct effect is given by

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

where the second and fourth lines only apply if $i > t$ as they capture effects of φ_i on household choices in the preceding period, $i-1$.[27]

[27]Using (3), the terms in the first and second line can be simplified. In particular,

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

The indirect effect is given by

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

where matrices with a negative exponent are zero by convention. The term $\partial m_{i-1}/\partial\varphi_i$ only applies if $i > t$. We first consider the case with inelastic labor supply.

A.1 Inelastic Labor Supply

Similarly to the situation in politico-economic equilibrium, the effects of policy changes *in period* t are linearly dependent.[28] Conditional on a choice for η_t, the optimality conditions for τ_t and σ_t can be rearranged as

$$\begin{aligned}
\frac{1 + \Omega_{22}\alpha}{1 - \tau_t - \sigma_t + \phi(\eta_t)} &= \frac{\delta(\Omega_{21} + \Omega_{22}(1-\alpha))}{\sigma_t}, \\
\frac{\beta p}{\rho}\frac{1-\alpha}{\alpha p_t(1 - f + f/p_t)(1-\eta_t) + (1-\alpha)\tau_t} &= \frac{\delta(\Omega_{21} + \Omega_{22}(1-\alpha))}{\sigma_t}.
\end{aligned}$$

In parallel with the politico-economic equilibrium conditions discussed earlier, these conditions require the cost and benefits of a marginal increase of σ_t to be equal to each other, and the benefits of marginal increases of τ_t and σ_t to be equalized, respectively. The terms $\Omega_{21} \equiv (\beta p + \rho)\left[(I - \rho M)^{-1}\right]_{[2,1]}$ and $\Omega_{22} \equiv (\beta p + \rho)\left[(I - \rho M)^{-1}\right]_{[2,2]}$ measure the present discounted contribution to the planner's objective of an increase in the (logarithm 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 $\beta(p+\psi\nu_{t+1}[1+\beta p(\delta(1-\alpha)+\alpha)])$. This is due to the fact that the planner internalizes the effect of capital accumulation on all future cohorts rather than tomorrow's retirees and workers only. Second, on the right-hand side, the term Ω_{21} replaces the expression $\psi\nu_{t+1}\beta^2(1-\alpha)p\varepsilon(1-\delta)$, 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, $\omega p/\nu_t$, is replaced by their relative weight as attached by the *planner*, $\beta p/\rho$.

[28] In particular, $\partial G_t(\cdot)/\partial\eta_t = -\alpha' p_t(1 - f + f/p_t)\partial G_t(\cdot)/\partial\tau_t$.

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_t affect the relative weights attached to the constituencies and thus, equilibrium tax rates.[29] Disregarding the inequality constraints (1) and solving the first-order conditions for a given η_t yields the following tax rates:

$$\begin{aligned}
\tau_t^G(\eta_t) &= \frac{\beta p(1-\delta\rho-(1-\delta)\rho(\alpha+\varepsilon(1-\alpha\rho)))}{(1-\alpha)(\beta p+\rho)(1-(1-\delta)\varepsilon\rho)} - \alpha'(1-\eta_t)p(1-f+f/p), \\
\sigma_t^G &= \frac{\delta\rho}{1-(1-\delta)\varepsilon\rho} > 0.
\end{aligned}$$

Conditional on η_t, the tax rate τ_t^G decreases with the strength of the human capital externality, ε, while the tax rate σ_t^G increases with it.

For any parameter constellation, the Ramsey tax rate σ_t^G and a continuum of policies $(\tau_t^G(\eta_t), \eta_t)$ satisfy the constraints (1) (because σ_t^G is positive and $\tau_t^{G\prime}(\eta_t) > 0$). The source of this indeterminacy is the same as in politico-economic equilibrium: Once savings choices have been made, labor and capital income taxes have exactly opposite effects. Focusing on the feasible policy with the smallest capital income tax rate, the Ramsey tax rates are given by

$$\kappa_t^G = (\tau_t^G, \sigma_t^G, \eta_t^G),\ \tau_t^G = \max[0, \tau_t^G(0)],\ \eta_t^G = \max[0, \eta_t^G(0)]$$

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

Turning to the Ramsey planner's choice of $\kappa_i^G,\ i > t$, both τ_i and η_i affect the savings decision in period $i-1$. As a consequence, no indeterminacy with respect to future policy instruments arises. The first-order conditions with respect to τ_i and η_i, respectively, feature the following terms *in addition* to the terms that are present in the period-t first-order conditions:

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

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 $\frac{z_t(\tau_t,\eta_t)}{1-z_t(\tau_t,\eta_t)} = \Omega_{22}\alpha$ 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 $\eta_t = 0$ and $f = 1$, then it is indeed the case that $\frac{z_t(\tau_t,\eta_t)}{1-z_t(\tau_t,\eta_t)} = \Omega_{22}\alpha$ when evaluated at the optimal period-t tax rates. To understand this result, note first that $\eta_t = 0$ and $f = 1$ in combination with time-invariant tax rates (at the period-t optimal values) implies that the consumption ratio $c_{1,i}/c_{2,i}$ equals $\rho/(\beta\nu_i)$ for all $i \geq t$ (see (2)); that is, the Ramsey planner achieves the same consumption ratio as a social planner that is not constrained

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

by the implementability constraints.[30] Second, if capital income remains untaxed ($\eta_i = 0$ for all $i \geq t$) and (surviving) retirees receive the full return to capital ($f = 1$), then households' privately optimal savings choice conforms with the planner's preferred savings level (conditional on the planner's preferred level of human capital accumulation). 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.

If $\tau_t^G = 0$ and $\eta_t^G > 0$, direct and indirect feedback effects do not cancel. Nevertheless, the Ramsey policy continues to be time-consistent under certain conditions. To see this, note that $\partial z_i(0, \eta_i)/\partial \eta_i = 0$. The potential source of time-inconsistency therefore disappears as far as the choice of η_i is concerned. Moreover, the corner solution for τ_t^G implies a corner solution for τ_i^G if the following inequality (which is independent of f) holds:

$$-\frac{\partial z_i(0, \eta_i^G)}{\partial \tau_i} \left(\frac{1}{1 - z_i(0, \eta_i^G)} - \frac{\Omega_{22}\alpha}{z_i(0, \eta_i^G)} \right) < 0.$$

If this inequality is violated, then the social benefit of higher savings by workers in the preceding period falls short of the benefit of higher consumption by workers in the preceding period. The Ramsey planner therefore prefers strictly positive intergenerational transfers ex-ante, but no such transfers ex post.

A.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 optimal tax rates in the initial period t are determinate, due to the presence of the marginal effects on labor supply.

The resulting system of optimality conditions does not generally allow for closed-form solutions. Suppose that $f = 1$ and the Ramsey tax rates satisfy $\eta_i = 0$ 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. If $\tau_i = 0$, a closed-form solution for σ_i results.

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

B Production Efficiency

Consider a path with constant ν and p and let y denote output per worker. Conditional on H_t and labor supply, we have

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

Starting from the investment policy $\{k_{t+i+1}, I_{t+i}\}_{i=0}^{\infty}$, consider a sequence of small reallocations of investment spending between physical and human capital investment. This sequence involves, in each period i, a small change in human capital investment of Δ_i (per student in period i), and a corresponding change in physical capital investment of $-\Delta_i$ (per worker 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

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

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

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

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

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

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 changes in physical and human capital investment must be positive.

When $a > 0$, physical capital is over accumulated in the initial allocation. As is apparent from the above conditions, one can generate a Pareto improvement in this case by reallocating resources from physical to human capital (corresponding to $\epsilon_{t+i} > 0$). Over accumulation of physical capital is also present if $a = 0$ and $\varepsilon(1-\delta) > 0$, corresponding to the allocation in an economy without government intervention, but with markets for private education financing (see Boldrin and Montes (2005) and Docquier, Paddison and Pestieau (2007) for a characterization of this 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 $\varepsilon(1-\delta)>0$, 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 in this case 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, $\delta(1-\alpha)\sum_{j=1}^{i}(\varepsilon(1-\delta))^j\epsilon<\epsilon$. To characterize the largest $a<0$ allowing for a persistent increase in output, 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 $a\epsilon^\star_{t+i}+\delta(1-\alpha)\sum_{j=1}^{i}(\varepsilon(1-\delta))^j\epsilon^\star_{t+i-j}=0$. This implies

$$\epsilon^\star_{t+1}=\frac{\delta(1-\alpha)}{-a}\varepsilon(1-\delta)\epsilon^\star_t$$

and

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

Boundedness of the sequence and thus, production inefficiency requires $-1<\varepsilon(1-\delta)\left(1-\frac{\delta(1-\alpha)}{a}\right)<1$ which simplifies (due to $a<0$) to the condition $a<-\frac{(1-\alpha)\varepsilon\delta(1-\delta)}{1-\varepsilon(1-\delta)}$.

Since $\frac{I_{t+i}}{k_{t+i+1}}=\frac{I_{t+i}}{k_{t+i}\gamma_H}=\frac{\sigma R}{\alpha'\nu\gamma_H}$, we have $a=(1-\alpha)\left(\delta-\frac{\sigma R}{\nu\gamma_H}\right)$. The criterion for production efficiency (if $\varepsilon(1-\delta)>0$ such that $a=0$ is not efficient), $-\frac{(1-\alpha)\varepsilon\delta(1-\delta)}{1-\varepsilon(1-\delta)}<a<0$, therefore reduces to

$$\frac{\delta}{\sigma} < \frac{R}{\nu\gamma_H} < \frac{\delta}{\sigma(1-\varepsilon(1-\delta))}.$$