DEPARTAMENTO DE ECONOMÍA
DOCUMENTO DE TRABAJO

"Sustaining Social Security"

Martín Gonzalez Eiras y Dirk Niepelt

D.T.: N° 72   Junio 2004

Sustaining Social Security*

Martín Gonzalez-Eiras
Universidad de San Andrés[†]

Dirk Niepelt
IIES, Stockholm University[‡]

June 29, 2004

Abstract

This paper analyzes the sustainability of intergenerational transfers in politico-economic equilibrium. We argue that these transfers arise naturally in a Markov perfect equilibrium in the fundamental state variables. In contrast to earlier literature, our explanation does not resort to altruism, commitment, or trigger strategies but rests on the incentive for young households to monopolize capital accumulation, as pointed out by Kotlikoff and Rosenthal (1990). Since transfers to the old are instrumental in that respect, the vote-maximizing platform under electoral competition sustains a large social security system. Introducing fully rational voters and probabilistic voting in the standard Diamond (1965) OLG model, we find that transfers in politico-economic equilibrium are too *high* relative to the social optimum. Standard functional form assumptions yield analytical solutions for both the Ramsey and the probabilistic voting case. Under realistic parameter values, the model predicts a social security tax rate of 12 percent, as compared to a Ramsey tax rate of 3.5 percent. Other predictions of the model are also consistent with the data. Analytical solutions for the case with endogenous labor supply and tax distortions show the results of the model to be robust.

KEYWORDS: Social security; intergenerational transfers; Markov perfect equilibrium; probabilistic voting; aggregate saving; aggregate labor supply.
JEL CLASSIFICATION CODE: E62, H55.

1   Introduction

Most developed and developing countries sustain pay-as-you-go social security systems with large intergenerational transfers. These transfer schemes command strong political support, although population ageing often threatens the financial viability of the systems under the

status quo. Notwithstanding political promises, social security entitlements are not written in stone. Benefit levels, contribution rates, the retirement age and many other parameters of the social security system are politically determined and may, in principle, be altered in the regular legislative process. Why do societies nevertheless sustain large social security schemes, even if these involve net contributions by the majority of voters? And should we expect societies to continue sustaining these schemes in the light of population ageing?

Since old voters benefit from publicly funded pensions, the political viability of a social security system crucially depends on the costs imposed by the system on young voters. In this paper, we argue that these costs are likely to be smaller than suggested by simple present value calculations: Since transfers to the old allow young voters to monopolize their factor supplies, in particular aggregate savings, social security generates indirect benefits for young voters, and these indirect benefits partially compensate for the direct tax cost. We present a model based on the standard Diamond (1965) overlapping generations framework, in which social security is a political equilibrium outcome. Households are assumed to be non-altruistic. As consumers, they are price takers. As voters, they rationally take into account how policies affect prices and future political choices. Since voters are not bound by past political decisions, the politico-economic equilibrium features subgame-perfect transfer choices supporting a competitive equilibrium.

Previous literature has studied the sustainability of intergenerational transfers under the assumption that voters play trigger strategies.[1] While certainly useful, this assumption has several shortcomings: To the extent that trigger strategy equilibria are not unique, their characteristics are not robust to changes in the modeling of players' beliefs. Moreover, trigger strategy equilibria rely on extreme assumptions. They do not arise in the limit of a finite horizon game, and they are not robust to small deviations from the assumption of infinite memory.[2] With these considerations in mind, and because we view it as a useful benchmark, we instead focus on Markov perfect equilibria where policy choices are only a function of the natural state variables (in our case, the capital per worker). Also in contrast to previous literature, we do not model electoral competition under the extreme assumption of a young "median voter" (i.e., under the assumption that the policy space is unidimensional, political candidates can exactly predict citizens' voting behavior as a function of the competing policy platforms, a Condorcet winner exists, and this Condorcet winner represents the bliss point of a young voter), but instead adopt the probabilistic voting assumption. More realistically, this assumption reflects the presence of uncertainty in the electoral process, thereby inducing a continuous mapping from candidates' electoral platforms to their vote shares. In equilibrium, candidates respond to electoral uncertainty by proposing a policy that maximizes the average welfare of *all* voters, not only that of the "median voter."

Our first substantive result is that young voters have a strategic incentive to support intergenerational transfers even in a Markov perfect equilibrium without commitment, altruism, or trigger strategies. This derives from the fact that transfers to the old depress capital accumulation, thereby affecting future returns and policy choices. As savers, young households do not account for these general equilibrium and policy effects, since they take prices and aggregate choices as given. But as an interest group, they have an incentive to manipulate both general equilibrium effects (as pointed out by Kotlikoff and Rosenthal (1990)[3]) and future policy choices. We argue that, to monopolize the supply of capital, transfers from the young to the old

---

[1] See Cooley and Soares (1999), Boldrin and Rustichini (2000) or Rangel (2003).

[2] See Bhaskar (1998) who shows that informational constraints in overlapping generations games with a strictly dominant action for the old imply that the unique pure strategy equilibrium is in Markov strategies.

[3] Kotlikoff and Rosenthal (1990) assume commitment and do not model the political process.

are more effective than alternative measures, such as taxing capital accumulation. Young voters are therefore less opposed to social security than the tax cost they must bear may suggest.

We find that intergenerational transfers are sustained in a dynamically efficient economy unless old age consumption without government intervention significantly exceeds young age consumption. Furthermore, the transfers sustained in politico-economic equilibrium are *higher* than those implemented under the Ramsey policy maximizing the discounted sum of the welfare of current and future generations. The reason why the political process sustains too high transfers is that electoral competition internalizes the positive general equilibrium and policy effects of lower capital accumulation for current voters, but not the negative effects borne by later cohorts. Social security is thus sustained by a coalition of old and young voters against future generations. The government implementing the Ramsey policy, in contrast, internalizes all these effects. It pursues social security policy only to the extent that the implied intergenerational redistribution increases the weighted average of the welfare of current *and future* generations. (In our benchmark model, no sources of inefficiency are present.) The Ramsey policy implements exactly the same transfers as those sustained in a hypothetical politico-economic equilibrium with symmetric political weights, where young voters neither account for the general equilibrium benefits of depressed capital accumulation nor the effect on future political choices ("double-myopic" equilibrium).

All these results follow under the assumption of constant returns to scale and equality of households' and the government's time discount factors. Once particular functional form assumptions (logarithmic preferences and Cobb-Douglas technology) are imposed, we can derive analytical solutions. This stands in sharp contrast to much of the literature which must resort to numerical characterizations. For standard parameter values and symmetric political weights, our analytical results predict a steady-state social security tax of 12 percent, only slightly less than the actual tax rate in the U.S. If young voters did not account for the general equilibrium and policy effects of decreased capital accumulation, this equilibrium tax rate would drop to 3.5 percent, which also represents the Ramsey tax rate. More generally, the model predicts the *introduction* and extension of social security programs in response to lower population growth rates or higher labor shares, a hump-shaped relationship between the population growth rate and social security benefits per retiree, higher consumption levels of old households than young households, and stronger support for social security in closed economies. These predictions are consistent with existing empirical evidence.

Our work is part of a growing literature on dynamic politico-economic equilibrium, with voters sequentially choosing their preferred policies under rational expectations about the effects on future equilibrium outcomes (see, for example, Krusell, Quadrini and Ríos-Rull, 1997; Hassler, Rodríguez Mora, Storesletten and Zilibotti, 2003). As mentioned above, our work also relates to an extensive literature on the sources of political support for intergenerational transfers. In contrast to most models in that literature, our approach does not rely on the assumption of altruism, commitment, or trigger strategies (as imposed, respectively, by Hansson and Stuart (1989) and Tabellini (1990); Cukierman and Meltzer (1989), Conesa and Krueger (1999), and Persson and Tabellini (2002); and Cooley and Soares (1999), Boldrin and Rustichini (2000) and Rangel (2003)), nor does it restrict policy choices to be binary or population growth to be sufficiently high (to render the economy dynamically inefficient), as do some previous models. Similarly to Cooley and Soares (1999), our approach stresses the role of general equilibrium effects and voters' incentives to manipulate these.

The assumptions of probabilistic voting and Markov-perfect equilibrium only appear to have been employed in the social security context by Grossman and Helpman (1998). However,

their model does not feature any economic decisions and therefore no interaction between the political and the economic sphere which is central to the mechanism proposed here.[4] Our analysis constitutes an attempt to close this gap in the literature. We show that the natural assumption of probabilistic voting reverses some of the results derived under the traditional assumption of a median voter. In particular, in our model, the politico-economic equilibrium sustains *higher* tax rates in steady state than the Ramsey equilibrium, while the median voter equilibrium sustains *lower* (in particular, zero) tax rates. Since the Ramsey equilibrium can be interpreted as a "good" politico-economic equilibrium backed by trigger strategies, our model implies that trigger strategies may be needed to reduce rather than increase intergenerational transfers.

The rest of the paper is structured as follows: Section 2 presents the model, derives the equilibrium allocation under probabilistic voting and the Ramsey policy, and confronts testable implications of the model with existing empirical evidence. Section 3 discusses the robustness of the findings with respect to the modeling of labor supply and the available transfer policies. We show that all important results remain valid if labor supply is endogenized, taxes are distorting, and voters choose transfers both to the old and the young. Section 4 concludes.

2 The Model

We consider an overlapping generations economy inhabited by cohorts of representative agents. Households live for two periods. Young households in period t inelastically supply labor at wage w_t and pay a labor income tax levied at rate τ_t. Disposable income is allocated to consumption, $c_{1,t}$, and savings, s_t, the latter yielding a gross rate of return, R_{t+1}. The consumption of old households, $c_{2,t+1}$, equals the gross return on savings, $s_t R_{t+1}$, plus a pension benefit, b_{t+1}. The population grows at the rate $\nu - 1$, such that the ratio of young to old households is given by $\nu > 0$.

Output is produced using an aggregate production function with constant returns to scale. Per-capita output in period t positively depends on the capital labor ratio which, in turn, is proportional to the per-capita savings of the cohort born in $t-1$. Factor markets are competitive and factor prices thus correspond to marginal products. The wage and the gross interest rate are given by $w_t = w(s_{t-1})$ and $R_t = R(s_{t-1})$, strictly increasing and decreasing in s_{t-1}, respectively. Conditional on prices and policies (that is, w_t, R_{t+1}, τ_t, and b_{t+1}), the indirect utility function of a young household of cohort t is given by

$$U_t = \max_{s_t} u(c_{1,t}) + \beta u(c_{2,t+1}), \tag{1}$$

subject to the budget constraint described above. The felicity function $u(\cdot)$ is continuously differentiable, strictly increasing, and satisfies $\lim_{c \to 0} u'(c) = \infty$; $\beta \in (0,1)$.

The government sector consists of a social security administration running a pay-as-you-go system.[5] Old age pensions are financed out of the payroll taxes paid by the young: the pay-as-you-go budget constraint of the social security administration reads

$$b_t = \tau_t \nu w_t. \tag{2}$$

---

[4]After finishing the first draft of this paper, we learned about independent work by Katuscak (2002) who also adopts the probabilistic voting assumption, but does not endogenize factor returns as we do.

[5]Introducing a fully funded component of social security is inconsequential, as long as the government does not force households to save more than they would voluntarily save and investment opportunities are the same for households and the social security administration.

The sole policy instrument of the social security administration is the payroll tax rate, τ_t, imposed on the labor income of the young. This tax rate is determined in the political process (described in more detail below), subject to a non-negativity constraint, $\tau_t \geq 0$.

The timing of events is as follows: At the beginning of period t, the tax rate to be imposed in the current period is determined in the political process. When deciding which candidate to support, voters anticipate how each candidate's policy platform will affect subsequent economic and political decisions. The wage rate and the return on the predetermined savings of the old, together with the tax rate implemented by the winning candidate, determine the consumption of the old and the disposable income of the young. Young households then turn to their role as consumers and choose how much to save.

When deciding (as voters) on τ_t and (as consumers) on s_t, young households form expectations about future benefits, b_{t+1}. In a Markovian equilibrium, these benefits depend on a set of "fundamental" state variables, $\mathcal{S}_{t+1}$, whose exact elements depend on how voters' preferences are aggregated and therefore, the political institutions in place: $b_{t+1} = \nu w(s_t)\tau(\mathcal{S}_{t+1})$. It is clear, however, that s_t is included in $\mathcal{S}_{t+1}$ since s_t affects future wages and gross returns and therefore, the incomes of next period's voters. Having said this, we conjecture that s_t is sufficient for $\mathcal{S}_{t+1}$, i.e., $\tau(\mathcal{S}_{t+1}) = \tau(s_t)$. We will return later to this point when discussing the political institutions in place.[6]

To characterize the politico-economic equilibrium, we proceed by backward induction. We start by analyzing the economic choices subject to given prices and policies, and then consider the political preferences over prices and policies and their aggregation in the political process.

2.1 Choice of Individual Savings

The optimal savings decision of a young household of cohort t is characterized by the Euler equation

$$u'(c_{1,t}) = \beta R_{t+1} u'(c_{2,t+1}).$$

Since households are atomistic, they take aggregate savings and thus, next period's return on capital and, through $\tau(s_t)$ and $w(s_t)$, social security benefits as given. Households only take into account that higher individual savings increase their future financial wealth.

Conditional on $\tau(s_t)$, the Euler equation maps disposable income, $w_t(1 - \tau_t)$, and aggregate savings into an individual household's optimal savings. We denote this implicit mapping by the function

$$s(w_t(1 - \tau_t); s_t, \tau(s_t)).$$

An equilibrium aggregate savings function, $S(w_t(1 - \tau_t); \tau(\cdot))$, is defined as a fixed point of the functional equation $S(y; \tau(\cdot)) = s(y; S(y; \tau(\cdot)), \tau(S(y; \tau(\cdot)))) \ \forall y \geq 0$.

2.2 Choice of Tax Rate

To characterize society's choice of program size, we first consider the welfare implications for old and young households in general equilibrium. These welfare implications induce preferences over policies for individual households. In a second step, we consider the aggregation of these preferences over policies through the political process.

Old households prefer as high a tax rate, τ_t, as possible. This follows directly from the fact that b_t increases in τ_t, while $s_{t-1}R_t$ is independent of τ_t and the tax bill for funding the benefits

---

[6]We assume that $\tau(\cdot)$ is single-valued, as is the case in the limit of the finite horizon economy.

is solely shouldered by the young. The welfare effect for an old household of a marginal increase in the tax rate is given by

$$u'(c_{2,t})w_t\nu. \tag{3}$$

For young households, a change in the tax rate gives rise to more complex welfare implications. Differentiating U_t with respect to τ_t yields

$$-u'(c_{1,t})w_t + \beta u'(c_{2,t+1}) \left[s_t R'(s_t) + \nu \frac{\mathrm{d}(w(s_t)\tau(s_t))}{\mathrm{d}s_t} \right] \frac{\mathrm{d}S(w_t(1-\tau_t);\tau(\cdot))}{\mathrm{d}\tau_t}. \tag{4}$$

(The indirect effect through changes in the household's optimal savings cancels due to an envelope argument based on the Euler equation.) The first negative term reflects the cost of higher tax payments. This cost term is at the root of the recurring question of how intergenerational transfers can be sustained in political equilibrium. Traditionally, the social security literature has rationalized the existence of positive intergenerational transfers in the absence of altruism or commitment based on trigger strategy arguments (Cooley and Soares, 1999; Boldrin and Rustichini, 2000; Rangel, 2003). Trigger strategy equilibria rely on state variables linking the current choice of tax rate to (expectations about) future choices of tax rate: the young support positive transfers because they consider them as investment in a "social contract" promising to pay dividends in the form of subsequent contributions by later generations. In this paper, we abstract from such non-fundamental state variables and instead propose an alternative explanation for the sustainability of intergenerational transfers. In particular, we emphasize the positive welfare implications due to induced changes in fundamental, aggregate state variables as reflected in (4): By shifting disposable income from the young generation (with a positive marginal propensity to save) to the old generation (with a propensity equal to zero), an increase in the tax rate reduces aggregate savings, thereby increasing next period's return on savings and altering the social security benefits.[7] More specifically, these "general equilibrium" effects due to depressed aggregate savings take the following forms:

i. Lower aggregate savings increase next period's return on savings, which benefits young households.

ii. Lower aggregate savings increase (reduce) next period's social security benefits depending on the sign of $\mathrm{d}(w(s_t)\tau(s_t))/\mathrm{d}s_t$, which benefits (hurts) young households.

The total general equilibrium effect of depressed aggregate savings is thus positive as long as next period's political choice of benefits does not strongly increase with aggregate savings.

Two questions immediately arise: First, whether it is possible that a positive general equilibrium effect outweighs the negative tax cost effect, thereby implying that young households actually favor paying transfers to the old. Second, whether alternative policies exist that depress aggregate savings without being so costly for young households. The answer to the first question is affirmative. A very inelastic savings function combined with a technology rendering the return on savings very sensitive to the capital stock, implies that the positive effect dominates

---

[7]While the literature has pointed out the effect of social security systems on prices in general equilibrium, it has argued that general equilibrium effects *in combination* with trigger strategies (Cooley and Soares, 1999; Boldrin and Rustichini, 2000) or commitment (Cukierman and Meltzer, 1989) can sustain positive intergenerational transfers. We argue that trigger strategies or commitment are not essential, and we stress voters' incentives to *manipulate* prices.

over some range if the slope of $w(s)\tau(s)$ is not too positive.[8] Alternatively, the same result may be obtained for *some* young households by introducing intragenerational heterogeneity. If, for example, households differ with respect to their dependence on labor income on the one hand and capital income on the other, "capitalists" might favor positive payroll taxes. In the remainder of the paper, we do not adopt such extreme assumptions. Instead, we confine ourselves to the standard representative agent framework. Within this setup, we show how the general equilibrium benefits from depressed aggregate savings boost the electorate's willingness to sustain intergenerational transfers in politico-economic equilibrium.

Why should young households fund social security, if the objective is only to depress aggregate savings? The answer to this second question depends on the alternative policy considered. As a first alternative, consider a policy of using payroll taxes to fund spending that only benefits the young. Such a policy will fail to achieve the objective if government spending occurs in the form of transfers or is a close substitute for private consumption. Alternatively, consider a policy of taxing tomorrow's capital income accruing to today's young. Such a policy may not affect the savings behavior at all, or even be counter-productive if the income effect balances or even outweighs the substitution effect of lower after tax interest rates.[9] Moreover, such a policy suffers from a time-inconsistency problem. Ex post, the then old households have no interest in imposing a capital income tax, even if the receipts are reimbursed. Indeed, if the reimbursements are shared among old and young households, or if the tax-cum-subsidy policy is associated with small distortions, old households will strictly oppose such capital income taxation ex post. A tax on today's savings instead of tomorrow's capital income circumvents the time-inconsistency problem associated with the capital income tax, but it still suffers from the other problems: income effects might counteract the desired effects, and the substitution towards first-period consumption might entail significant distortions. We conclude that, with inelastic labor supply, a transfer of resources to the old dominates alternative policy instruments for depressing aggregate savings. In section 3, we analyze the case with elastic labor supply.

Up to this point, we have not been specific about the political institutions in place which aggregate old and young voters' preferences as characterized by (3) and (4). In the following, we assume that in each period, policy is implemented by the winner of a two-party electoral competition game who receives the majority of votes cast by all households currently alive. To represent such a political environment, the median voter assumption is frequently employed.[10] However, an unappealing feature of that assumption is that it completely abstracts from uncertainty in the electoral process, giving rise to extreme implications. For instance, the mapping from candidates' electoral platforms into vote shares is discontinuous, and an infinitesimal change of the population growth rate from $-\epsilon$ to $+\epsilon$ leads to a large jump in the equilibrium tax rate. Clearly, these implications are stark and unrealistic, in particular in our setup with only two groups of voters. Therefore, we instead adopt the probabilistic voting assumption.

Probabilistic voting models acknowledge the fact that voters support a party not only for its policy platform but also for party characteristics like "ideology" that are orthogonal to the

---

[8]A simple example is the case of logarithmic utility, $u(c) \equiv \ln(c)$. Disregarding future benefits, the derivative of the indirect utility function U_t with respect to the tax rate then equals $-\frac{1+\beta}{1-\tau_t} - \frac{\beta^2 w_t R'(s_t)}{(1+\beta)R(s_t)}$. This is positive for a sufficiently low tax rate and high elasticity of $R(s)$.

[9]With logarithmic utility, income and substitution effects exactly offset each other. Even if the substitution effect prevails, a given desired reduction in savings requires a stronger distortion of the consumption profile over the life cycle than in the case of intergenerational transfers, because of the need to compensate for the income effect.

[10]Since political choice is unidimensional and voters' preferences are single peaked, a Condorcet winner exists in our model.

7

fundamental policy dimension of interest. These party characteristics are permanent and cannot be credibly altered in the course of electoral competition. The valuation of party characteristics differs across voters (even if they agree on their preferred policy platform) and is subject to random aggregate shocks, realized after parties have chosen their platforms. This renders the probability of winning a voter's support as a function of the competing policy platforms continuous, in contrast to the median voter setup.

In a Nash equilibrium with two parties maximizing their expected vote share, 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 group's size and sensitivity of voting behavior to policy changes. Groups caring relatively more about policy than party characteristics have more political influence since they are more likely to shift their vote from one party to the other in response to small changes in the proposed platform. In equilibrium, these groups of "swing voters" thus tilt policy in their own favor. In contrast, if all voters are equally responsive to changes in the policy platform, electoral competition implements the utilitarian optimum with respect to voters.

In the context of our model, the probabilistic voting assumption implies that the old generally have some weight in the objective function maximized by the political process, even if the median voter is young. The policy platform proposed by the candidates solves the program

$$\max_{\tau_t \geq 0} W(s_{t-1}, \tau_t; \tau(\cdot)),$$

$$W(s_{t-1}, \tau_t; \tau(\cdot)) \equiv \omega^o u(c_{2,t}) + \omega^y \nu(u(c_{1,t}) + \beta u(c_{2,t+1}))$$

$$\text{subject to} \begin{cases} s_{t-1} \text{ given,} \\ s_t = S(w_t(1-\tau_t); \tau(\cdot)), \\ \tau_{t+1} = \tau(s_t), \\ \text{household budget constraint.} \end{cases}$$

Here, the weights ω^i reflect the sensitivity of group i's voting behavior with respect to changes in a candidate's proposed policy platform, and the household budget constraint incorporates the benefit, wage, and return functions. Next period's policy choice as a function of the state is taken as given, reflecting our assumption of Markov equilibrium. An interior optimum for the candidates is characterized by the condition that the weighted sum of (3) and (4) (where the weights are given by ω^o and ω^y, respectively) be equal to zero. Note that our earlier assumption according to which s_t is the only element of $\mathcal{S}_{t+1}$, is indeed consistent.

In a rational expectations equilibrium, the anticipated policy function coincides with the optimal one. A rational expectations equilibrium is thus given by a fixed point $\tau(\cdot)$ of the functional equation $\tau(s_{t-1}) = \arg\max_{\tau_t \geq 0} W(s_{t-1}, \tau_t; \tau(\cdot)) \ \forall s_{t-1} \geq 0$. This probabilistic voting equilibrium under rational expectations contrasts with the probabilistic voting equilibrium under "double-myopia", where naive voters expect future tax rates and factor prices to be independent of their aggregate savings choice.

It is instructive to compare the politico-economic equilibrium with the allocation implemented by a benevolent government, subject to the same set of technological and institutional constraints. If the benevolent government cannot commit, future policy choices are once more a time invariant function of the state. Conditional on an intergenerational discount factor

---

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

8

$\rho, 0 < \rho < 1$, the benevolent government solves the program

$$\max_{\tau_t \geq 0} G(s_{t-1}, \tau_t; \tilde{\tau}(\cdot)),$$

$$G(s_{t-1}, \tau_t; \tilde{\tau}(\cdot)) \equiv \rho^{t-1} \beta u(c_{2,t}) + \sum_{i=t}^{\infty} \rho^i (u(c_{1,i}) + \beta u(c_{2,i+1}))$$

$$\text{subject to} \begin{cases} s_{t-1} \text{ given,} \\ s_i = S(w_i(1 - \tau_i); \tilde{\tau}(\cdot)), i \geq t, \\ \tau_{i+1} = \tilde{\tau}(s_i), i \geq t, \\ \text{household budget constraint,} \end{cases}$$

and where $\tilde{\tau}(\cdot)$ is now a fixed point of the functional equation $\tilde{\tau}(s_{t-1}) = \arg\max_{\tau_t \geq 0} G(s_{t-1}, \tau_t; \tilde{\tau}(\cdot))$ $\forall s_{t-1} \geq 0$. In contrast to the political process, the benevolent government without access to a commitment technology values the welfare of all households, not only those currently alive (and voting). It takes into account, for example, how a change in today's tax rate affects wages and thus the consumption of tomorrow's young.

If the government does have access to a commitment technology, in contrast, the resulting Ramsey program is no longer recursive. The government is only constrained by the requirement that the chosen policy be implementable as a competitive equilibrium—it need no longer be optimal ex post. The program of the government is given by

$$\max_{\{\tau_i\}_{i=t}^{\infty} \geq 0} \rho^{t-1} \beta u(c_{2,t}) + \sum_{i=t}^{\infty} \rho^i (u(c_{1,i}) + \beta u(c_{2,i+1}))$$

$$\text{subject to} \begin{cases} s_{t-1} \text{ given,} \\ s_i = S(w_i(1 - \tau_i); \tau_{i+1}), i \geq t, \\ \text{household budget constraint.} \end{cases}$$

Finally, the allocation implemented by a social planner solves the program

$$\max_{\{c_{1,i}, c_{2,i}, s_i\}_{i=t}^{\infty}} \rho^{t-1} \beta u(c_{2,t}) + \sum_{i=t}^{\infty} \rho^i (u(c_{1,i}) + \beta u(c_{2,i+1}))$$

$$\text{subject to} \begin{cases} s_{t-1} \text{ given,} \\ f(s_{i-1}/\nu) = c_{1,i} + s_i + c_{2,i}/\nu, i \geq t, \end{cases}$$

where $f(\cdot)$ denotes the economy's production function in intensive form.

2.3 Equilibrium

We will now characterize the equilibrium allocation. Our objective is to demonstrate that the general equilibrium benefits of depressed aggregate savings accruing to the young strongly affect their willingness to sustain intergenerational transfers and, moreover, that in politico-economic equilibrium, these transfers are too high from a social point of view. To show the first point, we compare the politico-economic equilibrium under rational expectations to the politico-economic equilibrium under double-myopia. To show the second point, we compare the politico-economic equilibrium under rational expectations to the allocation implemented by the social planner or the Ramsey policy.

Consider first the social planner allocation, as characterized by the first-order conditions

$$\rho u'(c_{1,t+i}) = \nu\beta u'(c_{2,t+i}), i \geq 0, \qquad (5)$$
$$u'(c_{1,t+i}) = \beta f'(s_{t+i}/\nu)u'(c_{2,t+i+1}), i \geq 0.$$

According to the first condition, the planner's marginal rate of substitution between the consumption of young and old households, $\rho u'(c_{1,t+i})/(\beta u'(c_{2,t+i}))$, is equal to the corresponding marginal rate of transformation, ν. According to the first and second condition, the planner's marginal rate of substitution between consumption in two successive periods, $u'(c_{1,t+i})/(\rho u'(c_{1,t+i+1}))$, is equal to the corresponding intertemporal marginal rate of transformation, $f'(s_{t+i}/\nu)/\nu$.

In contrast to the social planner, the government implementing the Ramsey policy is not only bound by the resource constraint but must also satisfy the implementability constraint of the private sector (i.e., $s_{t+i} = S(w_{t+i}(1-\tau_{t+i}); \tau_{t+i+1})$) and the non-negativity constraint on tax rates. The former constraint is inconsequential because the social planner's dynamic first-order condition is identical to the household's Euler equation, which holds under the Ramsey policy.[12] Thus, conditional on the intergenerational wealth distribution, the savings choices induced by the Ramsey policy conform with the social planner's investment policy. Moreover, since their marginal rates of substitution between consumption of young and old households are identical, the planner and the government also aim at the same wealth distribution across cohorts. An interior Ramsey policy (i.e., a Ramsey policy where the non-negativity constraint on tax rates does not bind) therefore implements the social planner allocation. This, in turn, implies that an interior Ramsey policy is necessarily time-consistent and therefore coincides with the policy implemented by the benevolent government without commitment. (See Appendix A.1 for a more formal discussion.)

In the remainder of the paper, we assume that the government weighs generations by their size, and discounts the welfare of future generations according to the household's discount factor, i.e., we assume that $\rho \equiv \beta\nu$ (implying $\beta\nu < 1$). The social planner and—to the extent that the non-negativity constraint on tax rates does not bind—the Ramsey policy therefore equalize the per-capita consumption of old and young households at any point in time. We summarize the preceding discussion as follows:

Proposition 1. Consider the Ramsey equilibrium. Suppose that strictly positive intergenerational transfers are sustained.
    (i) $c_{1,t} = c_{2,t} \; \forall t$.
    (ii) In steady state, $\beta R = 1$ and the economy is dynamically efficient (since $\beta\nu < 1$).
    (iii) The Ramsey policy is time-consistent, implements the same allocation as the social planner, and is therefore unique.

In politico-economic equilibrium, political influence enters the picture. The political weights of households currently alive, ω^y and ω^o, replace the government's intergenerational weights. More importantly, some effects internalized by the Ramsey policy—in particular, the welfare effects of an induced change in aggregate savings on future generations—are no longer accounted for. In the politico-economic equilibrium with rational expectations, candidates only internalize the direct effect due to the social security transfer from young to old households, and the indirect welfare effects on the current young, due to the change in next period's interest rate and social security benefits. In the following, we denote these indirect welfare effects (which are

---

[12]This follows from the fact that $R_{t+i} = f'(s_{t+i-1}/\nu)$ in a competitive equilibrium.

proportional to the second term in (4)) by $\mathcal{B}_t$:

$$
\begin{aligned}
\mathcal{B}_t &\equiv \frac{\mathrm{d}S(w_t(1-\tau_t);\tau(\cdot))}{\mathrm{d}\tau_t}\nu\beta\omega^{\mathrm{y}}u'(c_{2,t+1})\left[s_t R'_{t+1} + \nu w'_{t+1}\tau_{t+1} + \nu w_{t+1}\tau'_{t+1}\right] \\
&= \frac{\mathrm{d}S(w_t(1-\tau_t);\tau(\cdot))}{\mathrm{d}\tau_t}\left[w'_{t+1}(\tau_{t+1}-1) + w_{t+1}\tau'_{t+1}\right]\nu^2\beta\omega^{\mathrm{y}}u'(c_{2,t+1}),
\end{aligned}
$$

where the second line follows from constant returns to scale. Letting $\lambda_t^{\mathcal{W}}$ denote the non-negative multiplier on the constraint that taxes be non-negative, we thus find the following first-order condition with respect to τ_t for the politico-economic equilibrium under rational expectations:

$$
w_t\left(\nu\omega^{\mathrm{o}}u'(c_{2,t}) - \nu\omega^{\mathrm{y}}u'(c_{1,t})\right) + \mathcal{B}_t + \lambda_t^{\mathcal{W}} = 0, \ \ \lambda_t^{\mathcal{W}}\tau_t = 0. \tag{6}
$$

Under double-myopia, none of the indirect welfare effects due to induced changes in aggregate savings is internalized. The first-order condition with respect to τ_t for the politico-economic equilibrium under double-myopia therefore reduces to

$$
w_t\left(\nu\omega^{\mathrm{o}}u'(c_{2,t}) - \nu\omega^{\mathrm{y}}u'(c_{1,t})\right) + \lambda_t^{\mathcal{M}} = 0, \ \ \lambda_t^{\mathcal{M}}\tau_t = 0, \tag{7}
$$

where $\lambda_t^{\mathcal{M}}$ denotes the non-negative multiplier on the constraint that taxes be non-negative.

Let $\omega \equiv \omega^{\mathrm{o}}/\omega^{\mathrm{y}}$ denote the relative weight of the old in the political process. Condition (6) implies the following result:

Proposition 2. Consider the politico-economic equilibrium under rational expectations. Suppose that strictly positive intergenerational transfers are sustained and $\mathcal{B}_t > 0$, $\omega \geq 1$.

(i) $c_{1,t} < c_{2,t} \ \forall t$.

(ii) In steady state, $\beta R > 1$ and the economy is dynamically efficient, savings are lower and the tax rate is higher than in the Ramsey equilibrium.

Results (i) and (ii) follow directly from (6). Under the balanced consumption allocation implemented by the social planner or under the Ramsey policy, the direct welfare effect of a transfer from young to old households is equal to zero (for $\omega = 1$). But if $\mathcal{B}_t > 0$, there are additional indirect benefits from higher taxes. Moreover, if $\omega > 1$, transfers to the old become more valuable. In political equilibrium, taxes and the relative consumption of old households are therefore higher than in the Ramsey equilibrium. From the household's Euler equation, $c_1 < c_2$ implies that the steady-state interest rate is higher and savings are lower than in the Ramsey equilibrium.

If lower savings indirectly benefit young voters and higher taxes depress savings, then redistribution from the young to the old beyond the extent under the Ramsey policy ($c_1 = c_2$) constitutes the vote maximizing platform in the electoral competition game, even if $\omega = 1$. Young voters accept to bear the direct cost of higher taxes because they also benefit from improved terms of trade. Correspondingly, the next generation suffers from a fall in wages. But in contrast to the government implementing the Ramsey policy, the political process does not account for this fall in wages beyond its effect on lower social security benefits, because the political process only represents the interests of voters currently alive. Social security is thus sustained by a coalition of old *and* young voters who shift part of the cost of the system to future generations. Stronger political influence by the old ($\omega > 1$) further increases the support for social security.

Adopting the median voter assumption, the social security literature has generally concluded that the politico-economic equilibrium sustains "too low" transfers, unless backed by a powerful trigger strategy.[13] The assumption of probabilistic voting turns this conclusion upside-down—

---

[13]Under the median voter assumption, our model would replicate this conventional finding.

11

the politico-economic equilibrium now sustains "too high" transfers. Trigger strategies may thus be needed to reduce intergenerational transfers.

The next proposition makes it clear that the difference between the Ramsey tax rate and the tax rate in politico-economic equilibrium under rational expectations is due to the voters' incentive to manipulate general equilibrium effects, not to the probabilistic voting assumption:[14]

Proposition 3. Consider the politico-economic equilibrium under double-myopia. Suppose that strictly positive intergenerational transfers are sustained and $\omega \geq 1$.

(i) $c_{1,t} = c_{2,t}$ $\forall t$ if $\omega = 1$, and $c_{1,t} < c_{2,t}$ $\forall t$ if $\omega > 1$.

(ii) In steady state with $\omega \geq 1$, $\beta R \geq 1$ and the economy is dynamically efficient. In steady state with $\omega = 1$, the allocation is identical to the Ramsey allocation or the allocation implemented by the social planner.

Results (i) and (ii) follow directly from (7). If political influence is balanced ($\omega = 1$), then the political process equalizes the consumption of old and young households, parallel to the Ramsey or social planner allocation. This reflects the fact that the double-myopic political process does not only disregard the negative welfare effects of social security for future generations, but also the positive general equilibrium effects and the policy repercussions of concern to young voters. Since in steady state, the interest rate is the same in the Ramsey equilibrium and the double-myopic equilibrium, the condition $c_1 = c_2$ and therefore $w(1-\tau) - s = sR + \nu\tau w$ implies that the two tax rates are also the same. As in the politico-economic equilibrium with rational expectations, stronger political influence by the old ($\omega > 1$) increases the support for social security and raises the relative consumption of old households.

To go further and derive closed-form solutions, we impose functional form assumptions:

Assumption 1. Preferences are logarithmic: $u(c) \equiv \ln(c)$. The production function is of the Cobb-Douglas type: $w(s) \equiv A\alpha(s/\nu)^{1-\alpha}$, $R(s) \equiv A(1-\alpha)(s/\nu)^{-\alpha}$, $A > 0, 0 < \alpha < 1$.

Here, α denotes the labor share, s/ν the capital-labor ratio, and A the level of productivity. Under Assumption 1, household savings are given by

$$
\begin{aligned}
s_t &= w_t(1-\tau_t) - \frac{w_t(1-\tau_t) + \nu w_{t+1}\tau_{t+1}/R_{t+1}}{1+\beta} \\
&= \frac{A\alpha\beta}{1+\beta}\left(\frac{s_{t-1}}{\nu}\right)^{1-\alpha}(1-\tau_t) - \frac{\alpha s_t \tau_{t+1}}{(1+\beta)(1-\alpha)}, \\
\Rightarrow s_t &= A\alpha\left(\frac{s_{t-1}}{\nu}\right)^{1-\alpha}(1-\tau_t)\frac{(1-\alpha)\beta}{(1-\alpha)(1+\beta)+\alpha\tau_{t+1}} \equiv s_{t-1}^{1-\alpha} \cdot z(\tau_t, \tau_{t+1}),
\end{aligned}
\tag{8}
$$

implying

$$
\begin{aligned}
c_{1,t} &= s_{t-1}^{1-\alpha}\frac{A\alpha\nu^{\alpha-1}(1-\tau_t)(1-\alpha(1-\tau_{t+1}))}{(1-\alpha)(1+\beta)+\alpha\tau_{t+1}} \equiv s_{t-1}^{1-\alpha} \cdot \gamma(\tau_t, \tau_{t+1}), \\
c_{2,t} &= s_{t-1}^{1-\alpha}A\nu^{\alpha}(1-\alpha(1-\tau_t)) \equiv s_{t-1}^{1-\alpha} \cdot \delta(\tau_t).
\end{aligned}
$$

In the above expressions, τ_{t+1} is a function of next period's state variable, s_t. The expressions therefore represent the equilibrium aggregate savings and consumption functions only in the

---

[14]We assume that double-myopic voters expect next period's factor prices and policy choices to depend on the capital-labor ratio in the current as opposed to the next period. This particular way of modeling expectations guarantees that the anticipated values of variables and their actual realizations coincide in steady state, i.e., expectations are self-confirming in steady state.

special case where $\tau'(s_t) = 0$. In that special case—which will be of particular interest—we also have

$$\mathcal{B}_t = \frac{(1-\alpha)\alpha\beta(1-\tau_{t+1})\nu\omega^y}{(1-\tau_t)(1-\alpha(1-\tau_{t+1}))} > 0 \ \forall \tau_t, \tau_{t+1} \in [0,1).$$

Note that, in "autarky" (i.e., if all tax rates are set to zero), the ratio of $c_{2,t}/c_{1,t}$ is given by

$$\hat{c} \equiv \frac{(1-\alpha)(1+\beta)\nu}{\alpha},$$

independent of the capital-labor ratio. Moreover, under a *time-invariant* tax rate τ, the relative consumption of old and young households equals

$$\frac{c_{2,t}}{c_{1,t}} = \frac{\hat{c} + \nu\tau}{1 - \tau}. \tag{9}$$

Equation (9) implies that a particular time-invariant tax rate,

$$\tau^{\mathcal{R}} \equiv \frac{1 - \hat{c}}{1 + \nu} = \frac{\alpha - (1-\alpha)(1+\beta)\nu}{\alpha(1+\nu)},$$

suffices to equalize the per-capita consumption of old and young households in all periods, and thus to attain the social planner allocation. If $\hat{c} \leq 1$, this tax rate does not violate the non-negativity constraint of the Ramsey program. Proposition 1 thus implies that $\tau_t = \tau^{\mathcal{R}} \ \forall t$ constitutes the unique and time-consistent Ramsey policy. If $\hat{c} > 1$, in contrast, such that old households consume more than young households in autarky, setting τ_t equal to $\tau^{\mathcal{R}}$ violates the non-negativity constraint. As shown in Appendix A.2.1, in that case, the Ramsey policy amounts to keeping all tax rates at their constrained value, $\tau_t = 0 \ \forall t$. We thus have

Proposition 4. Consider the Ramsey equilibrium. Suppose that Assumption 1 holds.
(i) There exists an equilibrium with a flat policy function, $\tau(s) = \max(\tau^{\mathcal{R}}, 0)$. In this equilibrium, the steady state is globally stable and the transition to steady state is unique.
(ii) The flat policy function is the unique equilibrium policy function.
(iii) If $\hat{c} < 1$, the Ramsey policy is time-consistent and the Ramsey allocation is identical to the allocation implemented by the social planner. If $\hat{c} > 1$ and the policy function of the benevolent government without commitment is differentiable, the Ramsey policy is time-consistent.

Result (i) follows from (8), Proposition 1, the discussion in Appendix A.2.1, and Lemma 1 in Appendix A.2.2 which proves stability. If $\hat{c} < 1$, the policy function is strictly positive and the steady-state level of savings is $s^{\mathcal{R}} \equiv (A(1-\alpha)\beta)^{1/\alpha}\nu$, which is lower than the autarky level of savings, $s^{\mathcal{A}} \equiv (A\alpha\beta\nu^{\alpha-1}/(1+\beta))^{1/\alpha}$. If $\hat{c} \geq 1$, the policy function is given by the autarky policy function, $\tau(s) = 0$, and steady-state savings are at their autarky level, $s^{\mathcal{A}}$. Given that $\{\tau_t\}$ equals $\{\max(\tau^{\mathcal{R}}, 0)\}$, the complete transition of all endogenous variables is fully characterized by the initial condition, s_{t-1}, and the coefficients $\gamma(\tau_t, \tau_{t+1})$, $\delta(\tau_t)$, and $z(\tau_t, \tau_{t+1})$. The uniqueness result under (ii) and the equivalence result under (iii) follow from Proposition 1 and the discussion in Appendix A.2.1.

Consider next the politico-economic equilibrium under rational expectations. First, we restrict the attention to the case where the candidates face a *flat* tax function (imposed by the winner of the electoral competition game in the next period), given by τ_{t+1}. Differentiating $W(\cdot)$ with respect to τ_t and defining

$$\hat{d} \equiv \frac{(1-\alpha)(1+\beta(1-\alpha))\nu}{\alpha\omega} = \hat{c}\frac{(1+\beta(1-\alpha))}{(1+\beta)\omega} < \hat{c} \text{ for } \omega \geq 1$$

13

then yields[15]

$$\frac{\mathrm{d}W(s_{t-1}, \tau_t; \tau_{t+1})}{\mathrm{d}\tau_t} = \frac{\alpha\omega(1 - \hat{d}) - \alpha\tau_t(\omega + \nu(1 + \beta(1 - \alpha)))}{(1 - \alpha(1 - \tau_t))(1 - \tau_t)}.$$

For all $\tau_t \in [0, 1)$, this derivative is strictly negative if $\hat{d} > 1$. If $\hat{d} < 1$, in contrast, this derivative is strictly positive up to the strictly positive tax rate

$$\tau^{\mathcal{W}} \equiv \frac{\alpha\omega(1 - \hat{d})}{\alpha(\omega + \nu(1 + \beta(1 - \alpha)))} = \frac{\alpha\omega - (1 - \alpha)(1 + \beta(1 - \alpha))\nu}{\alpha(\omega + \nu(1 + \beta(1 - \alpha)))}$$

rendering the numerator equal to zero, and negative thereafter. These observations lead to

Proposition 5. Consider the politico-economic equilibrium under rational expectations. Suppose that Assumption 1 holds.

(i) There exists an equilibrium with a flat policy function, $\tau(s) = \max(\tau^{\mathcal{W}}, 0)$. In this equilibrium, the steady state is globally stable and the transition to steady state is unique.

(ii) If the policy function $\tau(s)$ is continuously differentiable, the flat policy function is the unique equilibrium policy function.

Consider the politico-economic equilibrium under double-myopia. Suppose that Assumption 1 holds and the economy is dynamically efficient.

(iii) If $\omega = 1$, steady-state savings and the steady-state tax rate are those given in Proposition 4.

(iv) If $\hat{c} < \omega$, no flat policy function constitutes an equilibrium policy function. A continuously differentiable policy function is upward sloping at the steady state.

Result (i) follows from (8), the properties of $\mathrm{d}W(\cdot)/\mathrm{d}\tau_t$, and Lemma 1 in Appendix A.2.2 which proves stability. A flat policy function is consistent with equilibrium because the tax rate maximizing $W(\cdot)$ is independent of the state variable as long as next period's tax function is also independent of the capital-labor ratio. If $\hat{d} < 1$, the policy function is strictly positive, $\tau^{\mathcal{W}} > \tau^{\mathcal{R}}$ for $\omega \geq 1$, and the steady-state level of savings is $s^{\mathcal{W}} \equiv \nu(A(1 - \alpha)\beta/(\omega + \beta\nu(1 - \alpha)))^{1/\alpha} < s^{\mathcal{R}}, s^{\mathcal{A}}$ for $\omega \geq 1$. If $\hat{d} \geq 1$, the policy function is given by the autarky policy function, $\tau(s) = 0$, and steady-state savings are at their autarky level, $s^{\mathcal{A}}$. Since $\hat{c} < 1$ implies $\hat{d} < 1$, the politico-economic equilibrium always features a weakly higher tax rate than the Ramsey equilibrium if $\omega \geq 1$. Given that $\{\tau_t\}$ equals $\{\max(\tau^{\mathcal{W}}, 0)\}$, the complete transition of all endogenous variables is fully characterized. Results (ii)–(iv) are proved in Appendix A.2.3.

Similarly to $\tau^{\mathcal{R}}$, $\tau^{\mathcal{W}}$ increases in α and decreases in β and ν. To better understand this result, it is instructive to return to the first-order condition (6). For $\lambda_t^{\mathcal{W}} = 0$ and under Assumptions 1 and a constant tax rate τ, this condition reduces to

$$\frac{c_{2,t}}{c_{1,t}} - \omega = (1 - \alpha)\beta\nu,$$

where the right-hand side equals $\mathcal{B}_t c_{2,t}/(\nu w_t \omega^y)$ and measures the benefit from depressing capital accumulation (and therefore tilting the consumption profile). This incentive to raise taxes

---

[15]Under Assumption 1, $W(\cdot)$ is given by

$W(s_{t-1}, \tau_t; \tau_{t+1})$
$\quad = \quad \omega \ln[s_{t-1}^{1-\alpha}\delta(\tau_t)] + \nu\{\ln[s_{t-1}^{1-\alpha}\gamma(\tau_t, \tau_{t+1})] + \beta \ln[(s_{t-1}^{1-\alpha}z(\tau_t, \tau_{t+1}))^{1-\alpha}\delta(\tau_{t+1})]\}$
$\quad = \quad \ln[s_{t-1}][(1 - \alpha)(\omega + \nu) + (1 - \alpha)^2\nu\beta] + \omega \ln[\delta(\tau_t)] + \nu \ln[\gamma(\tau_t, \tau_{t+1})] + \ln[z(\tau_t, \tau_{t+1})](1 - \alpha)\nu\beta + \nu\beta \ln[\delta(\tau_{t+1})].$

Substituting the definitions of $\gamma(\tau_t, \tau_{t+1})$, $\delta(\tau_t)$, and $z(\tau_t, \tau_{t+1})$ and differentiating yields the result.

increases in the capital share, households' patience, and the population growth rate. On the other hand, the autarky consumption profile $\hat{c}$ (which enters in $c_{2,t}/c_{1,t}$, see equation (9)) also increases in these three parameters, thereby reducing the incentive to raise taxes. The latter effect dominates.

2.4 Empirical Evidence and Significance

In spite of its simplicity, the model entails a variety of empirical predictions for the politico-economic equilibrium under rational expectations. We now confront these predictions with available empirical evidence. We start with qualitative results, and then turn to a calibrated version of the model to assess the quantitative importance of the motive to depress aggregate savings.

In a politico-economic equilibrium with positive tax rates, pensions as a share of GDP equal

$$\frac{w\tau^{\mathcal{W}}\nu}{w\nu + Rs} = \alpha\tau^{\mathcal{W}},$$

implying that the pension share is increasing in α, and therefore in GDP per capita (since the latter increases in α), and decreasing in ν. Both these results find support in the data. Analyzing the rise of the welfare state in a sample of 30 countries during the 1880–1930 period, Lindert (1994) finds a significant positive relationship between the pension share and both (lagged) GDP per capita and the share of the elderly. A sample of OECD countries during the 1960s and 1970s produces similar findings (Lindert, 1996).

The model also replicates the apparently non-linear empirical relationship between the share of elderly in the population and public pension payments *per retiree*: For the 1880–1930 period, Lindert (1994) estimates an elasticity of the pension share in GDP with respect to the share of elderly that is larger than unity. The OECD data suggest a hump-shaped relationship between the share of the elderly and public pension payments per retiree (Lindert, 1996), and Mulligan and Sala-i-Martin (2004) conclude that there is no clear relationship.[16] Since population growth rates have declined over time, the model can account for these observations as it predicts an inverse-U shaped relationship between the share of elderly and public pension payments per retiree, corresponding to $w\tau^{\mathcal{W}}\nu$.

Since the political support for social security partly arises from the motivation to affect interest rates, the model predicts a negative relationship between a country's integration with international capital markets and its level of intergenerational transfers. This prediction is also borne out in the data, as reported by Persson and Tabellini (2003, ch. 3), who find more trade to be associated with less welfare spending in a panel of 60 countries during the 1960–1998 period.[17] Similarly, time variation in the openness of the capital account should be reflected in time varying political support for social security, consistent with the introduction or expansion of social security systems during the interwar period with its sharp reversal of capital market globalization. The old should also oppose capital account liberalization to the extent that it reduces the elasticity of the interest rate with respect to domestic savings, thus undermining the political support by the young for social security.

---

[16]Based on data for the United States and 12 European countries over the period 1965–92, Razin, Sadka and Swagel (2002) argue that the dependency ratio is negatively related to per capita transfers.

[17]On the other hand, it runs counter to Rodrik's (1998) finding of a positive relationship between public spending and trade exposure. Rodrik (1998) does not control for the share of the old as do Persson and Tabellini (2003). He suggests that the positive relationship he finds might be due to the public sector providing social insurance against external risk, an aspect not present in our deterministic framework.

Another prediction of the model (under the assumption $\omega \geq 1$) is that in societies with social security systems, old households consume more than young households. This prediction is consistent with the empirical findings discussed by Mulligan and Sala-i-Martin (2004) according to which in most developed economies, the net income of a typical elderly person is at least as high as the net income of a typical nonelderly person. Moreover, the discrepancies in consumption or living standards are even larger than those in net incomes, and they have increased during the last thirty years.[18]

To assess the quantitative importance of the motive to depress aggregate savings, we calibrate the model. In particular, we impose Assumption 1 and set α to 0.7, β to 0.98^{30}, ω to unity, and ν to the gross growth rate of the U.S. population between 1970 and 2000 (1.384).[19] Table 1 displays the implied equilibrium values of the steady-state tax rate, capital-labor ratio, factor prices (over thirty years), young and old-age consumption, and life-time welfare.

Table 1: Steady-state equilibrium values

	Politico-economic equilibrium	Ramsey equilibrium
τ	0.1187	0.0349
capital-labor ratio	0.0563	0.0753
R	2.2484	1.8332 $(= \beta^{-1})$
w	0.2953	0.3223
c_1	0.1823	0.2067
c_2	0.2236	0.2067
U	-2.5190	-2.4362

For the chosen parameter values, $\hat{c} < 1$ (the autarky consumption of young households exceeds the autarky consumption of old households) and $\hat{d} < 1$. Both the Ramsey or social planner outcome and the politico-economic equilibrium therefore feature strictly positive tax rates.[20] The predicted tax rate in politico-economic equilibrium is nearly 12 percent, very close to the actual tax rate of 12.4 percent in the United States (OASDI). Moreover, the tax rate implemented in politico-economic equilibrium is more than three times as high as in the Ramsey or double-myopic equilibrium. Ceteris paribus, the positive difference between $\tau^{\mathcal{W}}$ and $\tau^{\mathcal{R}}$ increases further for lower values of α, or higher values of β or ν, see Figure 1. As shown in Table 1, the different tax rates go hand in hand with significantly different allocations. The politico-economic equilibrium features a lower capital-labor ratio, higher interest rates, lower wages, and a steeper consumption profile than the Ramsey equilibrium. (On an annual basis, the difference between the interest rates is slightly less than 0.7 percentage points.) In terms of steady-state welfare, the move from the Ramsey equilibrium to the politico-economic equilibrium is equivalent to a permanent reduction in consumption by more than 5 percent.

[18]Alternatively, high relative consumption of old households could arise due to complementarities between consumption and leisure.

[19]Piketty and Saez (2003) find α to vary between 0.68 and 0.75 in post-war U.S. data. The population growth rate is reported by the U.S. Census Bureau. We also set A to unity.

[20]Ceteris paribus, tax rates are strictly positive for $\alpha > 0.6814$, $\beta < 0.9875^{30}$, and $\nu < 1.5098$ under the Ramsey policy or $\alpha > 0.6251$, all $\beta \leq 1$, and $\nu < 2.0052$ in the probabilistic voting case.
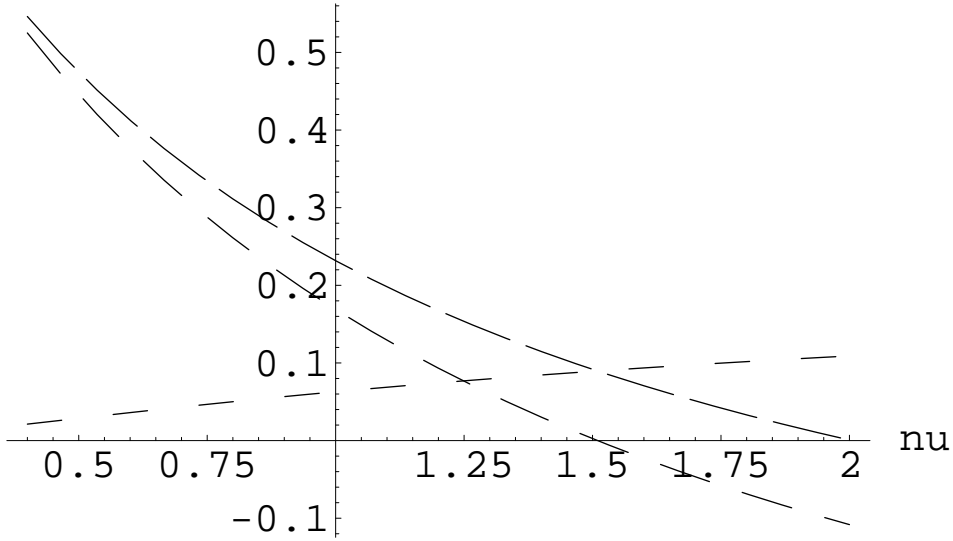
Figure 1: (Unconstrained) $\tau^{\mathcal{W}}$, $\tau^{\mathcal{R}}$, and $\tau^{\mathcal{W}} - \tau^{\mathcal{R}}$ as a function of ν (in order of decreasing length of line segments; α, β, ω at baseline values).

As a robustness check, we numerically solve for the politico-economic equilibrium under the assumption that the intertemporal elasticity of substitution, ε, differs from unity. We find relatively minor changes: For $\varepsilon = 1.5$, the capital-labor ratio increases to 0.0602, the interest rate falls to 2.1444, the tax rate increases to 0.1216, and the equilibrium tax function is upward sloping. For $\varepsilon = 0.5$, the capital-labor ratio falls to 0.0494, the interest rate increases to 2.4648, the tax rate falls to 0.1130, and the equilibrium tax function is downward sloping.

3   Elastic Labor Supply and Tax Distortions

The purpose of this extension is to examine the implications of endogenous labor supply, tax distortions, and alternative transfer schemes. If labor supply is endogenous, taxes can be used to depress aggregate savings in two ways: By reducing young households' wealth after taxes, as in the main model, and by distorting young households' allocation of disposable wealth between goods consumption (and thus savings) and leisure. We want to analyze whether the political process still sustains social security in this richer environment, or whether social security is dominated by another policy instrument that depresses labor supply and thus savings without distributing the tax receipts to the old.

To address this issue, we extend the model in two directions. First, we introduce an endogenous labor-leisure choice to model the role of tax distortions. For tractability, we assume a young household's felicity function to be separable in consumption and leisure, x_t. We also assume that old households do not work. The indirect utility function defined in (1) is thus replaced by

$$U_t = \max_{s_t, x_t} u(c_{1,t}) + v(x_t) + \beta u(c_{2,t+1}) \text{ s.t. household's budget set,}$$

where $v(\cdot)$ is continuously differentiable, strictly increasing, and satisfies $\lim_{x \to 0} v'(x) = \infty$. A

young household's time endowment is normalized to one. Second, we introduce an additional tax, levied at rate θ_t, on labor income. While the tax revenues $w_t(1 - x_t)\tau_t$ continue to fund social security, the additional tax revenues $w_t(1 - x_t)\theta_t$ fund a lump sum transfer to young households. The budget constraint of a young household thus reads

$$w_t(1 - x_t)(1 - \tau_t - \theta_t) + T_t = c_{1,t} + s_t,$$

where, in equilibrium, $T_t = w_t(1 - x_t)\theta_t$. Second-period consumption is still given by $c_{2,t+1} = s_t R_{t+1} + b_{t+1}$, where $b_{t+1} = \nu w_{t+1}\tau_{t+1}(1 - x_{t+1})$.[21] The tax rate θ_t allows labor income to be taxed without reducing the wealth of young households. With inelastic labor supply, this policy instrument had no value. With elastic labor supply, in contrast, it is valuable to the young since it allows their labor supply and therefore savings to be decreased *without* having to transfer resources to the old.

Optimal savings and labor supply decisions of a young household are characterized by the first-order conditions

$$\begin{aligned} u'(c_{1,t}) &= \beta u'(c_{2,t+1})R_{t+1}, \\ u'(c_{1,t})w_t(1 - \tau_t - \theta_t) &= v'(x_t), \end{aligned}$$

subject to the budget set earlier described. Conditional on anticipated next period policy functions $\tau(s_t)$, $\theta(s_t)$, $R(s_t)$ and $w(s_t)$, the household's first-order conditions and budget constraint map the capital-labor ratio, s_{t-1}, aggregate labor supply, and aggregate savings as well as τ_t and θ_t into a household's optimal choices of leisure and savings, $x(\cdot)$ and $s(\cdot)$, respectively.[22] Equilibrium aggregate savings and leisure functions, $S(\cdot)$ and $X(\cdot)$, respectively, are defined as fixed points of the functional equations

$$\begin{aligned} S(s_{t-1}, \tau_t, \theta_t; \tau(\cdot), \theta(\cdot), R(\cdot), w(\cdot)) &= s(s_{t-1}, X(\cdot), S(\cdot), \tau_t, \theta_t; \tau(S(\cdot)), \theta(S(\cdot)), R(S(\cdot)), w(S(\cdot))), \\ X(s_{t-1}, \tau_t, \theta_t; \tau(\cdot), \theta(\cdot), R(\cdot), w(\cdot)) &= x(s_{t-1}, X(\cdot), S(\cdot), \tau_t, \theta_t; \tau(S(\cdot)), \theta(S(\cdot)), R(S(\cdot)), w(S(\cdot))), \\ &\forall s_{t-1} \geq 0, \ 0 \leq \tau_t, \theta_t \leq 1, \ \tau_t + \theta_t \leq 1. \end{aligned}$$

The objective pursued by the political candidates is now given by

$$\max_{\tau_t, \theta_t \geq 0} \omega^o u(c_{2,t}) + \omega^y \nu(u(c_{1,t}) + v(x_t) + \beta u(c_{2,t+1}))$$

$$\text{subject to} \begin{cases} s_{t-1} \text{ given}, \\ s_t = S(s_{t-1}, \tau_t, \theta_t; \tau(\cdot), \theta(\cdot), R(\cdot), w(\cdot)), \\ x_t = X(s_{t-1}, \tau_t, \theta_t; \tau(\cdot), \theta(\cdot), R(\cdot), w(\cdot)), \\ \tau_{t+1} = \tau(s_t), \\ \theta_{t+1} = \theta(s_t), \\ \text{household budget constraint}. \end{cases}$$

In a rational expectations equilibrium, the anticipated policy functions $\tau(s_t)$ and $\theta(s_t)$ coincide with the optimal ones, and the anticipated factor price functions $R(s_t)$ and $w(s_t)$ are validated, subject to next period's aggregate leisure choice $x_{t+1} = X(s_t, \tau(s_t), \theta(s_t); \tau(\cdot), \theta(\cdot), R(\cdot), w(\cdot))$.

---

[21] To capture the dependency of benefits on individual labor supply, the benefit function can be generalized to $b_{t+1} = \nu w_{t+1}\tau_{t+1}(1 - X_{t+1})\frac{1 - x_t}{1 - X_t}$, with $1 - X$ denoting aggregate labor supply. While this affects the household's intratemporal first-order condition, the analysis still goes through without major changes.

[22] Next period's wage and return on savings both depend on the capital-labor ratio (i.e., s_t) and the endogenous labor supply. In a Markovian equilibrium, next period's labor supply is also a function of s_t. R_{t+1} and w_{t+1} are therefore functions of s_t only.

Does society still choose to sustain social security in this environment? Or does the option to depress aggregate savings without having to shift resources to the old lead to a collapse in the support for intergenerational transfers? In Appendix A.3, we show that the second alternative can generally be rejected under plausible conditions. Here, we directly focus on the special case of logarithmic preferences in consumption and Cobb-Douglas technology, since this allows us to characterize the politico-economic equilibrium in closed form.

Assumption 2. Preferences are logarithmic in consumption: $u(c) \equiv \ln(c)$. The production function is of the Cobb-Douglas type: $w(s_{t-1}, x_t) \equiv A\alpha(s_{t-1}/(\nu(1 - x_t)))^{1-\alpha}$, $R(s_{t-1}, x_t) \equiv A(1 - \alpha)(s_{t-1}/(\nu(1 - x_t)))^{-\alpha}$.

Under Assumption 2, equilibrium savings and consumption choices are given by

$$
\begin{aligned}
s_t &= s_{t-1}^{1-\alpha}(1 - x_t)^{\alpha} \cdot z(\tau_t, \tau_{t+1}), \\
c_{1,t} &= s_{t-1}^{1-\alpha}(1 - x_t)^{\alpha} \cdot \gamma(\tau_t, \tau_{t+1}), \\
c_{2,t} &= s_{t-1}^{1-\alpha}(1 - x_t)^{\alpha} \cdot \delta(\tau_t),
\end{aligned}
$$

where the functions $z(\cdot)$, $\gamma(\cdot)$, and $\delta(\cdot)$ have been defined earlier, and where τ_{t+1} is a function of s_t. Moreover, the household's intratemporal optimality condition yields

$$
v'(x_t) = \frac{1 - \tau_t - \theta_t}{(1 - x_t)(1 - \tau_t) - \frac{\beta(1-\alpha)(1-x_t)(1-\tau_t)}{(1-\alpha)(1+\beta)+\alpha\tau_{t+1}}}
$$

and thus an expression for leisure as a function of τ_t, θ_t, and τ_{t+1}, but not (directly) of s_{t-1}:

$$
x_t = x(\tau_t, \theta_t, \tau_{t+1}). \tag{10}
$$

Define $W^{\theta}(s_{t-1}, \tau_t, \theta_t; \tau_{t+1}, x_{t+1})$ to be the objective of a vote-maximizing candidate that faces per-capita savings of the old equal to s_{t-1} and *flat* future policy functions τ_{t+1} and x_{t+1}, and that imposes current tax rates of τ_t and θ_t. Under Assumption 2, $W^{\theta}(\cdot)$ is given by

$$
\begin{aligned}
&W^{\theta}(s_{t-1}, \tau_t, \theta_t; \tau_{t+1}, x_{t+1}) \\
&= W(s_{t-1}, \tau_t; \tau_{t+1}) + \ln[1 - x_t]\alpha(\omega + \nu + (1 - \alpha)\beta\nu) + \nu v(x_t) + \ln[1 - x_{t+1}]\alpha\nu\beta \\
&\qquad \text{subject to (10)} \\
&\equiv W(s_{t-1}, \tau_t; \tau_{t+1}) + g(x_t) + \ln[1 - x_{t+1}]\alpha\nu\beta \text{ subject to (10)},
\end{aligned}
$$

where the function $W(\cdot)$ has been defined earlier. An interior optimum satisfies

$$
\begin{aligned}
\frac{\partial W(\cdot)}{\partial \tau_t} + \frac{\mathrm{d}g(x_t)}{\mathrm{d}x_t}\frac{\partial x_t}{\partial \tau_t} &= 0, \\
\frac{\mathrm{d}g(x_t)}{\mathrm{d}x_t}\frac{\partial x_t}{\partial \theta_t} &= 0,
\end{aligned}
$$

and therefore features (since $\partial x_t/\partial \theta_t > 0$) the *same* optimal τ_t as in the case without tax distortions, $\tau^{\mathcal{W}}$. The optimal θ_t is then pinned down by the condition[23]

$$
\mathrm{d}g(x(\tau^{\mathcal{W}}, \theta_t, \tau^{\mathcal{W}}))/\mathrm{d}x_t = 0.
$$

---

[23]Since $\mathrm{d}g/\mathrm{d}x$ is decreasing in x, the condition is "stable"; low θ_t and thus x_t imply that $\mathrm{d}g/\mathrm{d}x > 0$, pushing x_t and thus θ_t upwards, and vice versa.

Summing up, for an interior choice of θ_t, the same policy function $\tau(\cdot)$ as in the main model arises in politico-economic equilibrium. The policy function $\theta(\cdot)$ is flat and—confirming our initial guess—the anticipated next-period policy functions for τ_{t+1} and x_{t+1} are also flat.[24] In particular, for $v(x) \equiv \phi \ln(x)$, our benchmark parameter values imply an equilibrium tax rate θ_t equal to 0.066.[25]

Returning to the question motivating this extension, we conclude that society sustains social security even in an environment with endogenous labor supply, distorting taxes, and alternative transfer schemes. From the perspective of a young voter, an increase in θ_t dominates an increase in τ_t, since the former decreases aggregate savings and increases wages. From the viewpoint of "society" (i.e., the collective of current voters), however, social security constitutes a more efficient instrument than a pure distortion of labor supply. While both policy instruments depress savings, the latter shifts resources from the needy old to the less needy young by lowering the return on savings and the labor supply funding pensions.

4 Conclusion

We have argued that the political support for intergenerational transfers reflects a variety of interests. All voters rather than a young median voter alone directly influence the size of the social security system. The micropolitical foundation for that view—a non-deterministic relationship between candidates' platforms and citizens' voting behavior, or probabilistic voting—is more natural and has more realistic implications than the alternative of a deterministic link between policy platforms and votes cast that underlies the median voter framework. Not only is the probabilistic voting assumption more appealing, introducing it in the standard Diamond (1965) model preserves that model's tractability and delivers intuitive and novel results in a strikingly transparent fashion.

Since old voters benefit from publicly funded pensions, the political viability of a social security system, and its size, crucially depend on the costs imposed on young voters by the system. We have shown that these costs are likely to be smaller than suggested by simple present value calculations, since social security generates indirect benefits by enabling young voters to monopolize their factor supplies, in particular aggregate savings. Among other factors, the magnitude of this effect depends on the elasticity of factor prices and thus, the size and openness of the economy. Small open economies should therefore be expected to have smaller social security systems than relatively closed economies. Moreover, time variation in the openness of the capital account should be reflected in time-varying political support for social security, and the old should be expected to oppose policies (as for instance the liberalization of the capital account) reducing the elasticity of the interest rate with respect to domestic savings.

Population ageing plays a central, but ambivalent, role in the model. An increase in the old-age dependency ratio is accompanied by the introduction or expansion of social security programs, although pensions per retiree as a function of the old-age dependency ratio are hump-shaped. Both these predictions, as well as the prediction of a positive relationship between per-capita GDP and the share of pensions in GDP, are consistent with the data. Moreover, they have important implications for the policy discussion where it is often argued that cuts in social

---

[24]We have not established uniqueness of this equilibrium. However, in numerical simulations, we have not found other equilibria.

[25]Under this functional form assumption, $x(\tau_t, \theta_t, \tau_{t+1}) = \frac{\phi(1-\tau_t)(1-\alpha(1-\tau_{t+1}))}{(1-\tau_t)((1-\alpha)(1+\beta+\phi)+\alpha(1+\phi)\tau_{t+1})-((1-\alpha)(1+\beta)+\alpha\tau_{t+1})\theta_t}$. The equilibrium tax rate is independent of ϕ.

security benefits herald the dismantling of pay-as-you-go social security systems. According to the model, this argument is misguided since population ageing will further increase the size of social security systems, even if the benefits per retiree might have surpassed their maximum.

While the political process in our model is "inclusive" in the sense of representing the interests of all voters, it is not sufficiently inclusive from a social welfare point of view. The interests of cohorts yet unborn are not represented. In effect, political competition establishes a coalition of old and young voters that partially shifts the costs of the social security system to future generations. As a consequence, the social security system is too large relative to a system balancing the interests of current and future generations. Both these implications—support for social security by a coalition of current voters at the expense of future generations, and suboptimally high intergenerational transfers—accord well with frequently expressed notions in the social security debate.

Throughout the paper, we have imposed the Markov assumption that political and economic choices are only functions of the single state variable, i.e., the capital-labor ratio. This contrasts with the more common assumption in the social security literature that support for intergenerational transfers depends on trigger strategies. Introducing trigger strategies in our setup is possible but, as we have argued, such strategies would be used to reduce rather than increase intergenerational transfers. Finally, we have imposed a balanced government budget throughout the analysis. A natural extension of the model would be to relax this assumption and allow for government debt, thereby introducing an alternative instrument to depress capital accumulation and redistribute across generations. We leave the analysis of this extended model with both capital and debt as state variables, and with endogenous tax, deficit, and default choices to future research.

A Appendix

A.1 Relationship Between Social Planner and Ramsey Allocation

Let $n_{t+i} \equiv \rho^{t+i-1}(\nu\beta u'(c_{2,t+i}) - \rho u'(c_{1,t+i}))$ denote the net benefit to the social planner, the government implementing the Ramsey policy, or the benevolent government of transferring one unit of resources in period $t + i$ from young to old households. From (5), the planner's optimal policy is to set $n_{t+i} = 0, i \geq 0$.

Let I_{t+i} denote the effect on the objective function of a marginal increase in savings in period $t + i$. Under the Ramsey policy, I_{t+i} equals[26]

$$I_{t+i}^{\mathcal{R}} = w'_{t+i+1}(\tau_{t+i+1} - 1)n_{t+i+1} + w'_{t+i+1}\frac{\mathrm{d}S_{t+i+1}}{\mathrm{d}w_{t+i+1}}w'_{t+i+2}(\tau_{t+i+2} - 1)n_{t+i+2} + \ldots.$$

Higher savings (i) raise the wage and thus transfers in the next period and (ii) reduce the interest rate. The corresponding welfare effects are (i) $\beta\rho^{t+i}\nu w'_{t+i+1}\tau_{t+i+1}u'(c_{2,t+i+1})$ on account of the old, $\rho^{t+i+1}w'_{t+i+1}(1-\tau_{t+i+1})u'(c_{1,t+i+1})$ on account of the young, and (ii) $\beta\rho^{t+i}s_{t+i}R'_{t+i+1}u'(c_{2,t+i+1})$ on account of the old.[27] The initial increase in savings is also propagated over time through higher wages and savings, and thus causes parallel welfare effects in subsequent periods. The expression for $I_{t+i}^{\mathcal{R}}$ then results, since the constant returns to scale assumption implies that $R'(s)s + w'(s)\nu = 0$. With $I_{t+i}^{\mathcal{R}}$ thus representing the shadow value of a marginal increase in savings, the Ramsey policy satisfies the following first-order condition with respect to $\tau_{t+i}, i \geq 0$:

$$w_{t+i}n_{t+i} + \frac{\mathrm{d}S_{t+i}}{\mathrm{d}\tau_{t+i}}I_{t+i}^{\mathcal{R}} + \frac{\partial S_{t+i-1}}{\partial\tau_{t+i}}w'_{t+i}\left[(\tau_{t+i} - 1)n_{t+i} + \frac{\mathrm{d}S_{t+i}}{\mathrm{d}w_{t+i}}I_{t+i}^{\mathcal{R}}\right] + \lambda_{t+i}^{\mathcal{R}} = 0, \quad (11)$$

$$\lambda_{t+i}^{\mathcal{R}}\tau_{t+i} = 0,$$

where $\lambda_{t+i}^{\mathcal{R}}$ denotes the non-negative multiplier on the constraint that τ_{t+i} be non-negative.[28] The first term on the left-hand side represents the direct welfare gain and loss for old and young households, respectively, due to higher transfers. The second term represents the welfare effects caused by the adjustment in savings resulting from higher taxes (and thus, lower disposable income). The third term represents the welfare effects caused by the adjustment in the savings in the *preceding* period, resulting from higher current taxes (and thus transfers).[29] Since $n_{t+i+s} = 0, s \geq 1$, implies that $I_{t+i}^{\mathcal{R}} = 0$, the *distribution* of consumption implemented by the social planner, $n_{t+i+s} = 0, s \geq 0$, satisfies the Ramsey first-order condition as long as the non-negativity constraint on tax rates does not bind.[30] Moreover, since the savings choices induced by the Ramsey policy conform with the social planner's investment policy, the consumption *levels* of old and young households in the social planner allocation and an interior Ramsey equilibrium also coincide.

---

[26]We use the short-hand notation S_t to denote $S(w_t(1-\tau_t); \tau_{t+1})$. w and R denote wage and return as functions of aggregate savings. A prime denotes the first derivative.

[27]There is no direct welfare effect from induced changes in savings, since savings choices are privately optimal.

[28]We do not impose an upper bound on the tax rate. Since $\lim_{c\to 0} u'(c) = \infty$, however, and since, in equilibrium, households cannot borrow against future benefits (the capital stock must be non-negative), tax rates will always be lower than unity in equilibrium.

[29]The first-order condition with respect to the tax rate in the initial period, τ_t, does not feature the third term in (11), since s_{t-1} is predetermined. The implications for time-consistency are discussed below.

[30]This would even be the case if the savings choices under the Ramsey policy conflicted with the social planner's investment policy, since the shadow value of a marginal increase in savings in an interior Ramsey allocation equals zero.

An interior Ramsey policy therefore implements the social planner allocation, and it is thus necessarily time-consistent, i.e., implements the same allocation as the benevolent government without commitment. This can also be seen from (11): If $n_{t+i} = 0, i \geq 0$, and therefore $I_{t+i}^{\mathcal{R}} = 0$, the terms in square brackets in (11) add up to zero, and the potential source of time inconsistency—the effect of a change in tax rate on savings in the preceding period—disappears.

The first-order condition of the benevolent government without commitment differs from (11) in two respects: First, since the government takes future tax choices to be functions of the state, it only chooses the contemporaneous tax rate, τ_t. Second, since changes in savings do not only affect future wages and returns, but also future tax rates, the expression for I_t is replaced by[31]

$$
\begin{aligned}
I_t^{\mathcal{G}} &= \left[w'_{t+1}(\tau_{t+1} - 1) + w_{t+1}\tau'_{t+1} \right] n_{t+1} + \\
&\quad \left[w'_{t+2}(\tau_{t+2} - 1) + w_{t+2}\tau'_{t+2} \right] \left(w'_{t+1}\frac{\mathrm{d}S_{t+1}}{\mathrm{d}w_{t+1}} + \tau'_{t+1}\frac{\mathrm{d}S_{t+1}}{\mathrm{d}\tau_{t+1}} \right) n_{t+2} + \ldots \\
&= \left[w'_{t+1}(\tau_{t+1} - 1) + w_{t+1}\tau'_{t+1} \right] n_{t+1} + \\
&\quad \left[w'_{t+2}(\tau_{t+2} - 1) + w_{t+2}\tau'_{t+2} \right] \frac{\mathrm{d}S_{t+1}}{\mathrm{d}\tau_{t+1}} \left(\tau'_{t+1} + w'_{t+1}\frac{\tau'_{t+1} - 1}{w_{t+1}} \right) n_{t+2} + \ldots.
\end{aligned}
$$

The first-order condition with respect to τ_t reads

$$
w_t n_t + \frac{\mathrm{d}S_t}{\mathrm{d}\tau_t} I_t^{\mathcal{G}} + \lambda_t^{\mathcal{G}} = 0, \quad \lambda_t^{\mathcal{G}} \tau_t = 0, \tag{12}
$$

where $\lambda_t^{\mathcal{G}}$ denotes the non-negative multiplier on the constraint that τ_t be non-negative. Since $n_{t+i} = 0, i \geq 1$, implies that $I_t^{\mathcal{G}} = 0$, and since the implementability constraints under the Ramsey policy and the policy without commitment are identical (conditional on tax rates), (12) confirms that an interior Ramsey policy is time-consistent.

A.2 Proofs

A.2.1 Proposition 4

Since $\hat{c} > 1$, the autarky allocation satisfies $c_{2,t+i} > c_{1,t+i}$, $i \geq 0$, and therefore $n_{t+i} < 0$, $i \geq 0$, and $I_{t+i}^{\mathcal{R}} > 0$, $i \geq 0$. Consider the effect of a marginal increase in τ_{t+i}, $i \geq 0$:

i. Direct effect: $c_{2,t+i}$ rises and $c_{1,t+i}$ falls, negatively contributing to the government's objective.

ii. General equilibrium effects via an induced fall in s_{t+i}: $c_{2,t+i+1}$ rises and $c_{1,t+i+1}$ falls, negatively contributing to the government's objective. Parallel (negative) effects arise in subsequent periods. (The change in savings has no direct welfare effects, since savings are optimally chosen by households.)

iii. General equilibrium effects via an induced fall in s_{t+i-1} (if $i > 0$): $c_{2,t+i}$ rises and $c_{1,t+i}$ falls, negatively contributing to the government's objective. Parallel (negative) effects arise in subsequent periods. (The change in savings has no direct welfare effects, since savings are optimally chosen by households.)

---

[31]S_t now serves as short-hand notation for $S(w_t(1 - \tau_t); \tilde{\tau}(\cdot))$.

Starting from the autarky allocation, a marginal increase in τ_{t+i}, $i \geq 0$, thus decreases the government's objective. At the same time, all terms n_{t+i+j}, $j \geq 0$, become more negative and $I^{\mathcal{R}}_{t+i}$ remains strictly positive. A strictly positive tax rate τ_{t+i}, $i \geq 0$, (associated with $\lambda^{\mathcal{R}}_{t+i} = 0$) can therefore not satisfy the first-order condition (11), which can be rewritten as

$$
n_{t+i} \underbrace{\left[w_{t+i} + \frac{\partial S_{t+i-1}}{\partial \tau_{t+i}} w'_{t+i}(\tau_{t+i} - 1) \right]}_{>0} + I^{\mathcal{R}}_{t+i} \underbrace{\left[\frac{\mathrm{d}S_{t+i}}{\mathrm{d}\tau_{t+i}} + \frac{\partial S_{t+i-1}}{\partial \tau_{t+i}} w'_{t+i} \frac{\mathrm{d}S_{t+i}}{\mathrm{d}w_{t+i}} \right]}_{<0} + \lambda^{\mathcal{R}}_{t+i} = 0.
$$

We conclude that keeping tax rates at zero in all periods constitutes the unique Ramsey policy.

To see that the Ramsey policy is time-consistent even if the non-negativity constraint on tax rates is binding, consider the program of the benevolent government without commitment in the case $\hat{c} > 1$. Combining conditions (12) in t and $t+1$ and exploiting the relationship between $I^{\mathcal{G}}_t$ and $I^{\mathcal{G}}_{t+1}$ yields, after some manipulations,

$$
w_t n_t + \lambda^{\mathcal{G}}_t - \frac{\mathrm{d}S_t}{\mathrm{d}\tau_t} \lambda^{\mathcal{G}}_{t+1} \left[\tau'_{t+1} + \frac{w'_{t+1}(\tau_{t+1} - 1)}{w_{t+1}} \right] = 0, \ \ \lambda^{\mathcal{G}}_t \tau_t = 0, \ \ \lambda^{\mathcal{G}}_{t+1}\tau_{t+1} = 0.
$$

If the policy function is differentiable, then $\lambda^{\mathcal{G}}_{t+1}\tau'_{t+1} = 0$ and the third term in the left-hand equation is weakly negative. The equation can therefore be satisfied under the autarky allocation $n_t < 0$ and a binding non-negativity constraint ($\tau_t = 0$, $\lambda^{\mathcal{G}}_t > 0$). A strictly positive τ_t, in contrast, necessarily violates the condition, because an increase in τ_t from zero renders n_t further negative but implies that $\lambda^{\mathcal{G}}_t = 0$. We conclude that the benevolent government without commitment chooses $\tau_t = 0 \ \forall t$, and therefore that the Ramsey policy is time-consistent.

A.2.2 Lemma

Lemma 1. Under Assumption 1, the steady state is globally stable and unique (except possibly for the degenerate steady state $s = 0$) if (i) $\tau'(s_t) = 0 \ \forall s_t$ or (ii) $\tau'(s_t) \geq 0 \ \forall s_t$ and $(1 - \alpha)(1 - \tau(s)) - s\tau'(s) \geq 0$.

Proof. Using the functional form assumptions, we have[32]

$$
\frac{\mathrm{d}s_t}{\mathrm{d}s_{t-1}} = \frac{\beta}{1 + \beta} \frac{A\alpha}{\nu^{1-\alpha}} \frac{s_{t-1}^{-\alpha} \left((1 - \alpha)(1 - \tau(s_{t-1})) - s_{t-1}\tau'(s_{t-1}) \right)}{1 + \frac{\alpha}{1-\alpha} \frac{\tau(s_t)}{1+\beta} + s_t \frac{\alpha}{1-\alpha} \frac{\tau'(s_t)}{1+\beta}},
$$

$$
\frac{\mathrm{d}s_t}{\mathrm{d}s_{t-1}} \Big|_{\text{steady state}} = \frac{\beta}{1 + \beta} \frac{A\alpha}{\nu^{1-\alpha}} \frac{(1 - \alpha)\nu^{1-\alpha} \frac{1+\beta}{A\alpha\beta} \left(1 + \frac{\alpha}{1-\alpha} \frac{\tau(s)}{1+\beta} \right) - s^{1-\alpha}\tau'(s)}{1 + \frac{\alpha}{1-\alpha} \frac{\tau(s)}{1+\beta} + s \frac{\alpha}{1-\alpha} \frac{\tau'(s)}{1+\beta}}.
$$

If $\tau'(s) = 0$, the second equation implies $\frac{\mathrm{d}s_t}{\mathrm{d}s_{t-1}}\big|_{\text{steady state}} = 1 - \alpha$ (and therefore local stability). If $\tau'(s_t) = 0 \ \forall s_t$, the first equation implies that s_t as a function of s_{t-1} is strictly increasing and concave and therefore, global stability and uniqueness (except possibly for the degenerate steady state $s = 0$). If $\tau'(s_t) \geq 0 \ \forall s_t$, the first equation implies that s_t as a function of s_{t-1} remains strictly increasing (but less so than in the case with a flat $\tau(s_t)$ function) and concave

---

[32]The second equation follows from the steady-state relationship

$$
1 + \frac{\alpha}{1 - \alpha} \frac{\tau(s)}{1 + \beta} = \frac{\beta}{1 + \beta} \frac{A\alpha}{\nu} \left(\frac{s}{\nu} \right)^{-\alpha} (1 - \tau(s)).
$$

24

up to the point where the numerator of the derivative becomes zero. Global stability and uniqueness (except possibly for the degenerate steady state $s = 0$) is thus guaranteed as long as $(1 - \alpha)(1 - \tau(s)) - s\tau'(s) \geq 0$. □

A.2.3 Proof of Proposition 5

(ii) (We assume that $\omega = 1$.)

Part A: We prove that, if $\tau(s)$ is continuously differentiable and strictly positive, the equilibrium policy function must be flat.

Step 1: We prove that $\tau'(s) = 0$ at the non-degenerate steady state.

From (6), using the functional form assumptions, we find

$$\frac{c_{1,t}}{c_{2,t}} = 1 + \frac{\mathrm{d}S_t}{\mathrm{d}\tau_t}\alpha\left(\frac{1 - \tau(s_t) - \frac{s_t}{1-\alpha}\tau'(s_t)}{w_t}\right). \tag{13}$$

From (8),

$$\frac{\mathrm{d}S_t}{\mathrm{d}\tau_t} = -\frac{w_t\beta(1-\alpha)}{(1-\alpha)(1+\beta) + \alpha\tau(s_t) + \alpha s_t\tau'(s_t)}.$$

Moreover,

$$\frac{c_{1,t}}{c_{2,t}} = \frac{1 - \tau_t}{\nu\left(\tau_t + \frac{1-\alpha}{\alpha}\right)}\left(1 - \frac{(1-\alpha)\beta}{(1-\alpha)(1+\beta) + \alpha\tau(s_t)}\right).$$

Together, these conditions imply that

$$\frac{(1 - \tau_t)\left(1 - \frac{\beta(1-\alpha)}{(1-\alpha)(1+\beta)+\alpha\tau(s_t)}\right)}{\nu\left(\frac{1-\alpha}{\alpha} + \tau_t\right)} =$$

$$1 - \frac{\beta(1-\alpha)\alpha}{(1-\alpha)(1+\beta) + \alpha\tau(s_t) + \alpha s_t\tau'(s_t)}\left(1 - \tau(s_t) - \frac{s_t\tau'(s_t)}{1-\alpha}\right). \tag{14}$$

Note that, since $\mathcal{B}_t > 0$ implies that $c_{1,t} < c_{2,t}$, we must have

$$(1 - \alpha)(1 + \beta) + \alpha\tau(s_t) + \alpha s_t\tau'(s_t) > 0. \tag{15}$$

(Suppose not. Then $\mathrm{d}S_t/\mathrm{d}\tau_t > 0$ and $1 - \tau(s_t) - \frac{s_t}{1-\alpha}\tau'(s_t)$ must be negative, implying that $\tau'(s_t) > 0$ and thus giving rise to a contradiction.) We first prove that convergent, non-oscillatory dynamics towards the steady state imply $\tau'(s) = 0$ around the non-degenerate steady state. Consider the first-order approximation around the non-degenerate steady state

$$\begin{aligned}\tau(s_t) &= \tau - \epsilon\tau', \\ \tau(s_{t-1}) &= \tau - (\epsilon + \delta)\tau',\end{aligned}$$

with τ' denoting the slope of the policy function at the steady state, $\epsilon = s - s_t$, and $\delta = s_t - s_{t-1}$. If the steady state is stable, then $\epsilon\delta \geq 0$. Let

$$\begin{aligned}a &\equiv \frac{\beta(1-\alpha)}{(1-\alpha)(1+\beta) + \alpha\tau}, \quad \hat{a} \equiv \frac{\beta(1-\alpha)}{((1-\alpha)(1+\beta) + \alpha\tau)^2}, \\ b &\equiv \tau + \frac{1-\alpha}{\alpha}, \\ c &\equiv \frac{\beta(1-\alpha)\alpha}{(1-\alpha)(1+\beta) + \alpha\tau + \alpha s\tau'}, \quad \hat{c} \equiv \frac{\beta(1-\alpha)\alpha}{((1-\alpha)(1+\beta) + \alpha\tau + \alpha s\tau')^2}.\end{aligned}$$

25

A first-order approximation of (14) around the steady state yields

$$(\epsilon + \delta)\tau'\frac{1-a}{\nu b} - \epsilon\tau'\frac{1-\tau}{\nu b}\alpha\hat{a} + (\epsilon + \delta)\tau'\frac{1-\tau}{\nu b^2}(1-a) = -\epsilon\tau'c\frac{2-\alpha}{1-\alpha} - \epsilon 2\alpha\tau'\hat{c}\left(1 - \tau - \frac{s\tau'}{1-\alpha}\right)$$

or

$$y\tau'\delta = \epsilon\tau'\left(-c\frac{2-\alpha}{1-\alpha} - 2\alpha\hat{c}\left(1 - \tau - \frac{s\tau'}{1-\alpha}\right) - \frac{1-a}{\nu b} - \frac{1-\tau}{\nu b^2}(1-a) + \frac{1-\tau}{\nu b}\alpha\hat{a}\right)$$

for some positive term y. The term multiplying $\tau'\epsilon$ on the right-hand side is negative because of (15), and because

$$\left(1 - \frac{\beta(1-\alpha)}{(1-\alpha)(1+\beta) + \alpha\tau}\right)\left(1 + \frac{1-\tau}{\frac{1-\alpha}{\alpha} + \tau}\right) > (1-\tau)\frac{\beta(1-\alpha)}{((1-\alpha)(1+\beta) + \alpha\tau)^2}\alpha$$

is implied by

$$1 - \frac{\beta(1-\alpha)}{(1-\alpha)(1+\beta) + \alpha\tau} > (1-\tau)\frac{\beta(1-\alpha)}{((1-\alpha)(1+\beta) + \alpha\tau)^2}\alpha,$$

which, in turn, holds due to

$$R = \frac{((1-\alpha)(1+\beta) + \alpha\tau)\nu}{\alpha\beta(1-\tau)}$$

(derived from the steady-state expression for c_1/c_2) and $R > \beta^{-1} > \nu$. A stable, non-oscillatory steady state therefore requires τ' to be (locally) equal to zero.

Note that $\frac{ds_t}{ds_{t-1}}|_{\text{steady state}} > 1$ is ruled out by the fact that there is only one steady state and $\lim_{t\to\infty} s_t < \infty$. It therefore remains to be shown that even if the (non-degenerate) steady state features $\frac{ds_t}{ds_{t-1}}|_{\text{steady state}} < 0$, it will be the case that $\tau' = 0$. From (14), we have (a) $\frac{d\tau_{t+1}}{d\tau_t} > 0$. At the same time, under our assumption, (b) $\frac{ds_t}{ds_{t-1}}|_{\text{steady state}} < 0$. First assume that (c) $\tau' > 0$. Take some s_{t-1} slightly lower than s. From (b), $s_t > s$, and from (c), $\tau(s_t) > \tau(s_{t-1})$. Now, take a slightly lower starting value, $\tilde{s}_{t-1} < s_{t-1}$. From (b), $\tilde{s}_t > s_t$, and from (c), $\tau(\tilde{s}_t) > \tau(s_t) > \tau(s_{t-1}) > \tau(\tilde{s}_{t-1})$, contradicting (a). Alternatively, assume that (d) $\tau' < 0$. Take some s_{t-1} slightly lower than s. From (b), $s_t > s$, and from (d), $\tau(s_t) < \tau(s_{t-1})$. Now, take a slightly higher starting value, $s_{t-1} < \tilde{s}_{t-1} < s$. From (b), $\tilde{s}_t < s_t$, and from (d), $\tau(s_{t-1}) > \tau(\tilde{s}_{t-1}) > \tau(\tilde{s}_t) > \tau(s_t)$, contradicting (a). We conclude that $\tau' = 0$ at the non-degenerate steady state.

Step 2: We prove that $\tau(t) = \tau(s)$, $0 \le t \le s$, where s denotes the non-degenerate steady-state savings level.

Consider condition (14). Since all $\tau'(s_t)$-terms in this equation are multiplied by s_t, it follows that $\tau(0) = \tau(s)$. Differentiability implies that there exists at least one value $\hat{s}, 0 < \hat{s} < s$, with $\tau'(\hat{s}) = 0$. It follows that $\tau(0) = \tau(s) = \tau(\hat{s})$. The same argument can now be extended to the whole interval of interest.

Step 3: We prove that $\tau(t) = \tau(s)$, $t > s$.

In the limit, $\tau'(s_t)$ has to approach zero at a rate faster than $1/s_t$. (Otherwise, $\tau(s_t)$ would not be bounded.) This implies that $\lim_{s_t\to\infty} s_t\tau'(s_t) = 0$. A parallel argument as under step 2 therefore implies that $\lim_{s_t\to\infty} \tau(s_t) = \tau(s)$. Differentiability implies that there exists at least one value $\hat{s}, s < \hat{s} < \infty$, with $\tau'(\hat{s}) = 0$. It follows that $\tau(s) = \tau(\hat{s}) = \lim_{s_t\to\infty} \tau(s_t)$. The same argument can now be extended to the whole interval of interest.

Part B: We extend the proof to the case where taxes are not necessarily strictly positive.

If the non-negativity constraint on the tax rate is binding, (13) is replaced by an inequality, with the left-hand side of (13) being smaller than the right-hand side. Note that (13) depends on s_t only through $\tau(s_t)$. If tax rates are constrained Differentiability implies $\tau'(s) = 0$ and $\tau'(0) = 0$.

This concludes the proof of (ii).

(iii)–(iv) Turning to the double-myopic equilibrium, define $M(s_{t-1}, \tau_t; \tau_{t+1})$ to be the objective of a candidate that faces per-capita savings of the old equal to s_{t-1} and next-period's tax rate equal to τ_{t+1} (perceived to be unaffected by current policy choices), and that imposes a current tax rate of τ_t. $M(\cdot)$ differs from $W(\cdot)$ because next period's interest and wage rates (which affect the second period consumption of young households) are perceived to be unaffected by the choice of τ_t:

$$M(s_{t-1}, \tau_t; \tau_{t+1}) = \omega \ln[s_{t-1}R_t + \nu w_t \tau_t] + \nu\{\ln[w_t(1 - \tau_t) - s_t] + \beta \ln[s_t R_t + \nu w_t \tau_{t+1}]\}.$$

Differentiating $M(\cdot)$ with respect to τ_t and applying the household's Euler equation yields

$$
\begin{aligned}
\frac{\mathrm{d}M(s_{t-1}, \tau_t; \tau_{t+1})}{\mathrm{d}\tau_t} &= \nu w_t \left(\frac{\omega}{c_{2,t}} - \frac{1}{c_{1,t}}\right) = \nu w_t \left(\frac{\omega}{\nu w_t \tau_t + \nu w_t (1-\alpha)/\alpha} - \frac{1}{w_t(1-\tau_t) - s_t}\right) \\
&= \left(\frac{\alpha\omega}{\alpha\tau_t + 1 - \alpha} - \frac{\nu(1+\beta)}{1 - \tau_t + \nu\tau_t/R_t}\right).
\end{aligned}
$$

For all $\tau_t \in [0, 1)$ and $R_t > \nu$, this derivative is strictly negative if $\hat{c} > \omega$. If $\hat{c} < \omega$, in contrast, then this derivative is strictly positive up to the strictly positive tax rate rendering the derivative equal to zero, and negative thereafter. Result (iii) follows from two observations: First, if $\hat{c} < \omega = 1$, such that taxes are strictly positive, then $c_{1,t} = c_{2,t}$ (see Proposition 3). Second, in steady state, this implies $R = \beta^{-1}$ such that the derivative of $M(\cdot)$ equals zero at $\tau_t = \tau^{\mathcal{R}}$. (If $\hat{c} \geq 1$, the autarky equilibrium results.) Result (iv) follows from the observation that the derivative of $M(\cdot)$ is an increasing function of the state variable (through the interest rate) and that $\hat{c} < \omega$ implies strictly positive taxes.

A.3 Elastic Labor Supply and Tax Distortions

For convenience, we treat T_t and the sum of the two tax rates, $\sigma_t \equiv \tau_t + \theta_t$, rather than τ_t and θ_t as the independent political choice variables. Old age benefits and the composition of tax rates are implicitly defined. Conditional on anticipated next period policy functions $\sigma(s_t)$, $T(s_t)$, $R(s_t)$ and $w(s_t)$, the household's first-order conditions and budget constraint map the capital-labor ratio, s_{t-1}, aggregate labor supply, and aggregate savings as well as σ_t and T_t into a household's optimal choices of leisure and savings, $x(\cdot)$ and $s(\cdot)$, respectively. Equilibrium aggregate savings and leisure functions, $S(\cdot)$ and $X(\cdot)$, respectively, are defined as fixed points of the functional equations

$$
\begin{aligned}
S(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)) &= s(s_{t-1}, X(\cdot), S(\cdot), \sigma_t, T_t; \sigma(S(\cdot)), T(S(\cdot)), R(S(\cdot)), w(S(\cdot))), \\
X(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)) &= x(s_{t-1}, X(\cdot), S(\cdot), \sigma_t, T_t; \sigma(S(\cdot)), T(S(\cdot)), R(S(\cdot)), w(S(\cdot))), \\
\forall s_{t-1}, \sigma_t, T_t &\geq 0.
\end{aligned}
$$

The objective pursued by the political parties is now given by

$$\max_{\sigma_t, T_t \geq 0} \omega^{\mathrm{o}} u(c_{2,t}) + \omega^{\mathrm{y}} \nu(u(c_{1,t}) + v(x_t) + \beta u(c_{2,t+1}))$$

$$\text{subject to} \begin{cases} s_{t-1} \text{ given,} \\ s_t = S(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)), \\ x_t = X(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)), \\ \sigma_{t+1} = \sigma(s_t), \\ T_{t+1} = T(s_t), \\ \text{household budget constraint.} \end{cases}$$

In a rational expectations equilibrium, the anticipated policy functions $\sigma(s_t)$ and $T(s_t)$ coincide with the optimal ones, and the anticipated factor price functions $R(s_t)$ and $w(s_t)$ are validated, subject to next period's aggregate leisure choice $x_{t+1} = X(s_t, \sigma(s_t), T(s_t); \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot))$.

Consider the effect of a marginal increase in T_t, which consists of three parts:

i. The direct welfare effect due to lower transfers from young to old households,

$$-\nu(\omega u'(c_{2,t}) - u'(c_{1,t})).$$

ii. The welfare effect on young voters due to lower aggregate savings. This effect parallels the general equilibrium and policy effects in the main model and equals[33]

$$\frac{\partial S_t}{\partial T_t} \nu \beta u'(c_{2,t+1}) \left\{ s_t R'(s_t) + \frac{\nu \mathrm{d}[w(s_t)(1 - x(s_t))\sigma(s_t) - T(s_t)]}{\mathrm{d}s_t} \right\}.$$

iii. The welfare effects on young and old voters due to changes in the labor supply. These effects, which did not arise in the main model, equal

$$\frac{\partial X_t}{\partial T_t} \nu \left\{ -\omega u'(c_{2,t}) w_t \sigma_t - \frac{\partial w(s_{t-1}, x_t)}{\partial x_t} (1 - x_t)(1 - \sigma_t)(\omega u'(c_{2,t}) - u'(c_{1,t})) \right\},$$

where we use the household's intratemporal optimality condition as well as the constant returns to scale property. The terms in curly brackets represent the loss for old households from lower social security benefits (due to lower labor supply), and the general equilibrium welfare effects on young and old households. Higher wages (due to lower labor supply) benefit young households and also old households (due to the effect on pensions) but, at the same time, old households suffer from lower returns on their savings. Due to constant returns to scale, this latter effect is proportional to the change in wages.[34]

(The effects of a marginal increase in σ_t closely resemble the terms above. The term in i. is multiplied by $-w_t(1 - x_t)$, and in the expressions in ii. and iii., the derivatives with respect to T_t are replaced by the derivatives with respect to σ_t.)

The second of the above three effects is negative if conditions parallel to those relevant for Proposition 2 are satisfied. In particular, the effect is negative if aggregate savings increase in the lump-sum subsidy T_t, and if depressing aggregate savings is beneficial for the young. The first of the three effects is negative if $\omega u'(c_{2,t}) - u'(c_{1,t})$ is positive, i.e., if old households consume

---

[33]S_t and X_t denote $S(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot))$ and $X(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot))$, respectively.

[34]The welfare loss for young households from lower consumption (due to lower labor supply) and the welfare gain from higher leisure consumption exactly offset each other.

relatively little. Finally, the third effect is negative if $\omega u'(c_{2,t}) - u'(c_{1,t})$ is positive and leisure is a normal good on the aggregate level, i.e., if an increase in T_t reduces young households' labor supply. In sum, this suggests that, starting from an autarky allocation satisfying $\omega u'(c_{2,t}) > u'(c_{1,t})$ (a sufficient condition for which is $c_{2,t} \leq c_{1,t}$ and $\omega \geq 1$), the political process does *not* introduce transfers to the young but *does* sustain (for $\sigma_t > 0$) transfers to the old. Even if the young might prefer the distorting labor income tax to fund a lump-sum transfer to themselves, the vote maximizing policy is to fund social security. If, due to social security payments to the old, the politico-economic equilibrium features relatively high old-age consumption ($\omega u'(c_{2,t}) < u'(c_{1,t})$), then optimal T_t might differ from zero.

References

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

Boldrin, M. and Rustichini, A. (2000), 'Political equilibria with social security', *Review of Economic Dynamics* **3**, 41–78.

Conesa, J. C. and Krueger, D. (1999), 'Social security reform with heterogeneous agents', *Review of Economic Dynamics* **2**(4), 757–795.

Cooley, T. F. and Soares, J. (1999), 'A positive theory of social security based on reputation', *Journal of Political Economy* **107**(1), 135–160.

Cukierman, A. and Meltzer, A. H. (1989), 'A political theory of government debt and deficits in a Neo-Ricardian framework', *American Economic Review* **79**(4), 713–732.

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

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

Hansson, I. and Stuart, C. (1989), 'Social security as trade among living generations', *American Economic Review* **79**(5), 1182–1195.

Hassler, J., Rodríguez Mora, J., Storesletten, K. and Zilibotti, F. (2003), 'The survival of the welfare state', *American Economic Review* **93**(1), 87–112.

Katuscak, P. (2002), A positive theory of social security in an open economy. Mimeo, University of Michigan.

Kotlikoff, L. J. and Rosenthal, R. W. (1990), Some inefficiency implications of generational politics and exchange, Working Paper 3354, NBER, Cambridge, Massachusetts.

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.

Lindert, P. H. (1994), 'The rise of social spending, 1880–1930', *Explorations in Economic History* **31**(1), 1–37.

Lindert, P. H. (1996), 'What limits social spending?', *Explorations in Economic History* **33**(1), 1–34.

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).

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

Persson, T. and Tabellini, G. (2002), Political economics and public finance, *in* A. J. Auerbach and M. Feldstein, eds, 'Handbook of Public Economics', Vol. 3, Elsevier Science, Amsterdam, chapter 24.

Persson, T. and Tabellini, G. (2003), *The Economic Effects of Constitutions*, 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.

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

Razin, A., Sadka, E. and Swagel, P. (2002), 'The aging population and the size of the welfare state', *Journal of Political Economy* **110**(4), 900–918.

Rodrik, D. (1998), 'Why do more open economies have bigger governments?', *Journal of Political Economy* **106**(5), 997–1032.

Tabellini, G. (1990), A positive theory of social security, Working Paper 3272, NBER, Cambridge, Massachusetts.