Sustaining Social Security*

Martín Gonzalez-Eiras  
*Universidad de San Andrés†

Dirk Niepelt  
IIES, Stockholm University‡

March 5, 2004

Abstract  
This paper analyzes the sustainability of intergenerational transfers in politico-economic equilibrium. We argue that these transfers naturally arise 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 voters to monopoleize capital accumulation, as pointed out by Kotlikoff and Rosenthal (1990); transfers to the old are instrumental in that respect. 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 Ramsey equilibrium. Under standard functional form assumptions, we are able to analytically solve for the steady state and the complete transition dynamics in both the Ramsey and the probabilistic voting case. Under realistic parameter values, the model predicts a social security tax rate of 12 percent, compared to a Ramsey tax rate of 3.5 percent. Several qualitative predictions of the model find support in the data.

KEYWORDS: Social security; intergenerational transfers; Markov perfect equilibrium; probabilistic voting; overlapping generations.

JEL Classification Code: E62, H55.

1 Introduction

Many 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 aging often threatens the financial viability of the systems under the status quo. Political promises notwithstanding, social security entitlements are not written in stone.

---

*We thank George-Marios Angeletos, Piero Gottardi, John Hassler, Hugo Hopenhayn, Per Krusell, Torsten Persson, Kjetil Storesletten, Harald Uhlig, Jaume Ventura, Philippe Weil, Fabrizio Zilibotti, and seminar participants at CERGE-EI, CEU, ECARES, Humboldt, IIES, Maryland, MIT, San Andrés, and UTDT for comments on papers at different stages of this project; and Christina Lönnblad for editorial assistance. Large parts of this paper were written while Gonzalez-Eiras visited the IIES. He thanks the IIES for the hospitality and acknowledges financial support from Fundación Antorchas.

†Vito Dumas 284, B1644BID Victoria, Pcia. Buenos Aires, Argentina. E-mail: mge@udesa.edu.ar
‡SE-106 91 Stockholm, Sweden. E-mail: dirk.niepelt@iies.su.se
Benefit levels, contribution rates, the retirement age and many other parameters of the social security system are politically determined and can, in principle, be altered in the regular legislative process. Why do democratic societies nevertheless choose to sustain large social security schemes (or, equivalently, not to default on outstanding public debt) even if this involves contributions by the majority of voters in excess of future benefits?

To answer this question, we introduce politics into the standard Diamond (1965) overlapping generations model. 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 new households are born in each period and since these new voters are not obliged by earlier political decisions, the political process cannot commit to future transfers. Politico-economic equilibrium is therefore characterized by subgame-perfect choices of transfer policies supporting competitive equilibria.

The previous literature has studied the sustainability of intergenerational transfers under the assumption that voters play trigger strategies. While certainly useful, this assumption has several shortcomings: To the extent that trigger strategy equilibria are not unique, their characteristics are not necessarily 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. With these considerations in mind, and because we consider it to be 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 stock).

Our first substantive result is that young voters have a strategic incentive to support intergenerational transfers even in such a Markov perfect equilibrium without commitment, altruism, or trigger strategies. The incentive 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 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.

In particular circumstances, for example if the interest rate responds very elastically to the capital stock, the benefits from depressed aggregate savings outweigh the cost of higher taxes. Intergenerational transfers are then sustained in politico-economic equilibrium even under the extreme assumption of a young “median voter” (i.e., 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). We do not pursue these extreme assumptions but instead focus on the more realistic case with constant returns to scale production and probabilistic voting. The probabilistic voting assumption reflects the real-world feature of a multidimensional policy space: voters do not exclusively support political candidates for their policy platforms, but rather also take other aspects like “ideology” or “taste” for a particular candidate into consideration. With random shocks to the ideology component of voters’ preferences, candidates can no longer exactly predict citizens’ voting behavior as a function of the competing policy platforms. In equilibrium,

\(^{2}\)See Blaskar (1998).
\(^{3}\)Kotlikoff and Rosenthal (1990) assume commitment and do not model the political process.
the influence of voters is determined by the extent to which they trade off economic well being and ideology. Richer and ideologically more heterogeneous groups are less strongly represented than “swing voters” whose support for a candidate strongly responds to changes in the proposed policy platform.

Our second set of results concerns politico-economic equilibrium in this Markovian benchmark economy. We find that intergenerational transfers are sustained in a dynamically efficient economy unless old age consumption in the autarky allocation significantly exceeds young age consumption. Furthermore, we find that the transfers sustained in politico-economic equilibrium are higher than those implemented by a Ramsey government with or without commitment. The reason why the political process sustains too high transfers is that voters account for the positive general equilibrium and policy effects of lower capital accumulation, but not the negative effects borne by later cohorts: social security is sustained by a coalition of old and young voters against future generations. The Ramsey government, in contrast, accounts for all these effects. As it turns out, the Ramsey policy implements exactly the same transfers as those sustained in a hypothetical politico-economic equilibrium, where young voters account neither 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, in the Ramsey case, the equality of private and governmental discount factors. Once we also impose particular functional form assumptions (logarithmic preferences and Cobb-Douglas technology), we are able to derive analytical solutions for the steady state and the complete transition dynamics in both the Ramsey and the probabilistic voting case. This feature of our third set of results stands in sharp contrast to much of the literature which must resort to the numerical characterization of politico-economic equilibrium. For standard parameter values, our analytical results predict a steady-state social security payroll tax of 12 percent; if young voters did not account for the general equilibrium and policy effects of depressed capital accumulation, this equilibrium tax rate would drop to 3.5 percent, which also represents the Ramsey tax rate. In terms of qualitative implications, the model is able to replicate several linear and non-linear relationships found in the data.

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 Rios-Rull, 1997; Hassler, Rodriguez Morà, 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 the earlier literature, in particular Cooley and Soares (1999), our approach stresses the role of general equilibrium effects and voters’ incentives to manipulate them.

Our assumptions of probabilistic voting and Markov-perfect equilibrium appear to have been employed in the social security context only 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 that is central to our arguments.\textsuperscript{4} This gap in the literature is unfortunate, and our model constitutes an attempt to close it. We show that the more realistic probabilistic voting assumption reverses some of the results derived under the traditional assumption of a median voter and trigger strategies. In particular, in our model, the politico-economic equilibrium sustains higher tax rates in steady state than the Ramsey equilibrium. Since the latter can be interpreted as a “good” politico-economic equilibrium backed by trigger strategies, our model implies that trigger strategies may be needed to sustain low rather than high intergenerational transfers.

The rest of the paper is structured as follows: Section 2 presents the model, derives the equilibrium allocation under the assumption of probabilistic voting and in the Ramsey case, and discusses testable implications of the model in light of the data. Section 3 discusses the robustness of the findings. 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 \( \tau_t \). 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 \( t \) depends positively 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, \( w_t, R_{t+1}, \tau_t \), and \( b_{t+1} \)), the indirect utility function of a young household of cohort \( t \) is given by

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

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.\textsuperscript{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

\[
b_t = \tau_t \nu w_t. \tag{2}
\]

The sole policy instrument of the social security administration is the payroll tax rate, \( \tau_t \), 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, \( \tau_t \geq 0 \).

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

\textsuperscript{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 timing of events is as follows: At the beginning of period $t$, voters determine the tax rate to be imposed in the current period, anticipating how this choice will affect subsequent economic and political decisions. The wage rate and the return on the predetermined savings of the old, together with the chosen tax rate, 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 $\tau_t$ 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, $S_{t+1}$, whose exact elements depend on how voters’ preferences are aggregated and therefore, the political institutions in place: $b_{t+1} = \nu w(s_t) \tau(S_{t+1})$. It is clear, however, that $s_t$ is included in $S_{t+1}$ because $s_t$ affects future wages and gross returns and therefore, the endowments of next period’s voters. Having said this, we conjecture that $s_t$ is sufficient for $S_{t+1}$, i.e., $\tau(S_{t+1}) = \tau(s_t)$; we will later return to this point when discussing the political institutions in place.

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

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

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 $\tau_t$ as possible. This follows directly from the fact that $b_t$ increases in $\tau_t$, while $s_{t-1}R_t$ is independent of $\tau_t$ and the tax bill from funding the benefits is solely shouldered by the young. The welfare effect for an old household of a marginal increase in the tax rate is given by

$$ u'(c_{2,t})w_t\nu. \quad (3) $$
For young households, a change in the tax rate gives rise to more complex welfare implications. Differentiating $U_t$ with respect to $\tau_t$ yields

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

(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 and Zeckhauser, 2001; 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 correctly expect that in this way they invest 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 raising the return on savings and altering the social security benefits.\(^6\) 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; this benefits young households.

ii. Lower aggregate savings raise (lower) next period’s benefits depending on the sign of $\partial(w(s_t)\tau(s_t))/\partial s_t$; this 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, implying that young households actually favor paying transfers to the old. Second, whether young households cannot adopt alternative, less costly strategies to depress aggregate savings rather than simply giving money away. 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 in place, implies that the positive effect dominates over some range if the slope of $w(s)\tau(s)$ is not too positive.\(^7\) Alternatively, the same result may be obtained for some young households by introducing intragenerational

---

\(^6\)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.

\(^7\)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 $-\frac{1 + \beta}{1 - \gamma} - \frac{\beta}{(1 + \beta)\gamma^r}$. For a sufficiently low tax rate and high elasticity of $R(s)$, this is positive.
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 example considered in detail below, we do not adopt such extreme assumptions. Instead, we restrict ourselves to a very 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 a 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 for spending that benefits the young as opposed to the old. 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. Moreover, such a policy suffers from a time inconsistency problem: Ex post, young 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 depressingaggregate 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 that aggregate old and young voters’ preferences as characterized by (3) and (4). One frequently employed modeling device, the median voter assumption, is feasible but problematic: it does not well capture electoral competition among parties and leads to extreme outcomes. Therefore, we adopt the probabilistic voting assumption. By introducing some uncertainty about the relationship between candidates’ policy platforms and their vote shares, the probabilistic voting framework smooths out the unrealistic discontinuities in the reaction function of voters present in the median voter framework. In equilibrium, the vote maximizing candidate does not only take the median voter’s preferences into account, but also the preferences of other voters.

In the context of our model, this implies that the old generally have some weight in the welfare function maximized by the political process, even if the median voter is young. The value function maximized in the political process is given by

\[
W(s_{t-1};\tau(\cdot)) = \max_{\tau_t} \left[ \omega^{old} u(c_{2,t}) + \omega^{young} \nu(u(c_{1,t}) + \beta u(c_{2,t+1})) \right]
\]

subject to

\[
\begin{aligned}
& s_{t-1} \text{ given,} \\
& s_t = S(w_t(1 - \tau_t);\tau(\cdot)), \\
& \tau_{t+1} = \tau(s_t), \\
& \text{household budget constraint.}
\end{aligned}
\]

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 lifetime than in the case of intergenerational transfers, because of the need to compensate for the income effect.

Since political choice is unidimensional and voters’ preferences are single peaked, a Condorcet winner exists.
Here, the weights $\omega^i$ reflect the sensitivity of group $i$’s voting behavior with respect to changes in a candidate’s proposed policy platform\textsuperscript{10}, 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 a 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 $\omega^{\text{old}}$ and $\omega^{\text{young}}$, respectively) be equal to zero. Note also that our earlier conjecture according to which $s_t$ is the only element of $S_{t+1}$, 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 $\tau(s_{t-1}) = \arg \max_{\tau_t} W(s_{t-1}; \tau(\cdot)) \forall s_{t-1} \geq 0$.

It is instructive to compare the politico-economic equilibrium with the allocation implemented by a Ramsey planner subject to the same set of technological and institutional constraints. If the Ramsey planner cannot commit, future policy choices are again represented by a time invariant function of the state. Conditional on a given sequence of intergenerational welfare weights $\{\omega_s\}$\textsuperscript{11}, the corresponding value function reads

$$
\mathcal{R}(s_{t-1}; \tau(\cdot)) = \max_{\tau_t} \left[ \omega_{t-1} \beta u(c_{2,t}) + \sum_{i=t}^{\infty} \omega_i (u(c_{1,i}) + \beta u(c_{2,i+1})) \right]
$$

subject to

\[
\begin{align*}
& s_{t-1} \text{ given}, \\
& s_i = S(w_i(1 - \tau_i); \tau(\cdot)) \forall i \geq t, \\
& \tau_{i+1} = \tau(s_i) \forall i \geq t, \\
& \text{ household budget constraint},
\end{align*}
\]

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

If, in contrast, the Ramsey planner does possess commitment power, her program is no longer recursive. The planner 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 constrained optimal policy then solves the program

$$
\mathcal{C}(s_{t-1}) = \max_{\{\tau_t\}_{t=0}^{\infty}} \left[ \omega_{t-1} \beta u(c_{2,t}) + \sum_{i=t}^{\infty} \omega_i (u(c_{1,i}) + \beta u(c_{2,i+1})) \right]
$$

subject to

\[
\begin{align*}
& s_{t-1} \text{ given}, \\
& s_i = S(w_i(1 - \tau_i); \tau_{i+1}) \forall i \geq t, \\
& \text{ household budget constraint}.
\end{align*}
\]

2.3 Equilibrium

We will now characterize the equilibrium allocation. We will derive all important qualitative results for the general case. Moreover, we will derive closed form solutions for the equilibrium tax rate and the savings level in steady state and during the transition, using particular functional

\textsuperscript{10}Deviations from the benchmark ($\omega^{\text{old}} = \omega^{\text{young}}$) arise, for example, due to different ideological patterns or lobbying power across groups (see Persson and Tabellini (2000)).

\textsuperscript{11}The sequence of welfare weights must decrease sufficiently quickly for the value function to be defined.
form assumptions specified below. 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. To make this point, we compare the politico-economic equilib- 
rium to two benchmark equilibria. The first benchmark is the Ramsey allocation, the second 
is an economy composed of double-myopic voters who disregard both the general equilibrium 
and the policy effects of lower savings and therefore do not appreciate the benefits of depressing 
aggregate savings. We model this feature by assuming that double-myopic voters expect next 
period’s factor prices and policy choices to depend on the inherited savings in the current as 
opposed to the next period. This particular way of modeling myopic expectations guarantees 
that the expected values of variables and their actual realizations coincide in steady state, i.e., 
expectations are self-confirming in steady state.

In the following, we assume that the Ramsey planner’s intergenerational discount factor is 
equivalent to that of the households (\( \omega_i = \omega_{i-1} \beta \nu \); we also normalize \( \omega_{i-1} \) to \( \beta^{-1} \)), and \( \beta \nu < 1 \). 
Consider first the Ramsey equilibrium under commitment. Let \( I_t \) denote the social welfare 
effect as of time \( t \) due to a marginal increase in savings in period \( t \). In the case of the Ramsey 
government with commitment, \( I_t \) equals \(^{12}\)

\[
R^C_t = \nu \beta w_{t+1}'(\tau_{t+1} - 1) (u_{t+1}' - u_{t+1}) + \nu^2 \beta^2 w_{t+1}' \frac{dS^c_{t+1}}{dw_{t+1}} w_{t+2}' (\tau_{t+2} - 1) (u_{t+2}' - u_{t+2}) + \ldots
\]

Higher savings in period \( t \) (i) raise the wage and thus transfers in period \( t + 1 \) and (ii) reduce 
the interest rate in \( t + 1 \). The corresponding welfare effects are (i) \( \nu \beta w_{t+1}' \tau_{t+1} u_{t+1}' \) per capita 
of the old at \( t + 1 \); \( \beta w_{t+1}' (1 - \tau_{t+1}) u_{t+1}' \) per capita of the young at \( t + 1 \); and (ii) \( \beta^2 s_t R^C_{t+1} u_{t+1}' \) per capita 
of the old at \( t + 1 \). Moreover, these effects are propagated over time through higher 
wages and savings. Using the constant returns to scale property \( R^c(s_{t+1}) s_{t+1} = w'(s_{t-1}) \nu = 0 \) 
and the assumption of “matching” discount factors yields the expression for \( I_t^C \). (There is no 
direct welfare effect from induced changes in savings since households optimally choose their 
savings.) Differentiating \( C(s_{t-1}) \) with respect to \( \tau_{t+i}, i > 0 \), then yields the first-order conditions 

\[
\nu w_{t+i} \left( u'(c_{2,t+i}) - u'(c_{1,t+i}) \right) + \nu \frac{dS^C_{t+i}}{d\tau_{t+i}} R^C_{t+i} + \frac{\partial S^C_{t+i-1}}{\partial \tau_{t+i}} \nu w_{t+i} \times \left[ (\tau_{i+i} - 1) \left( u'(c_{2,t+i}) - u'(c_{1,t+i}) \right) + \frac{dS^C_{t+i}}{dw_{t+i}} \right] + \lambda_{t+i} = 0, \quad \lambda_{t+i} = 0, \quad \lambda_{t+i} = 0,
\]

where \( \lambda_{t+i} \) denotes the non-negative multiplier on the constraint that \( \tau_{t+i} \) be non-negative.\(^{13}\) 
The first term on the left-hand side represents the direct utility gain and loss for today’s old 
and young, respectively, due to higher transfers. The second term represents the welfare 
effects caused by the adjustment in savings resulting from higher current taxes (and thus, lower 
income). The third term represents the welfare effects caused by the adjustment in the 
preceding period’s savings resulting from higher current taxes (and thus transfers). The 
first-order condition with respect to the tax rate in the initial period, \( \tau_t \), does not feature the 
third effect since \( s_{t-1} \) is predetermined. Combining two successive first-order conditions yields,

\(^{12}\)We use the shorthand notation \( S_t \) to denote \( S(w, (1 - \tau_t); \tau_{t+1}) \). \( w \) and \( R \) denote wage and return as functions 
of aggregate savings. A prime denotes the first derivative.

\(^{13}\)We do not impose an upper bound on the tax rate. Since \( \lim_{\nu \to 0} u'(c) = \infty \), 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.
after some manipulations,

\[
\nu \left( u'(c_{t+1}) - u'(c_{1,t+1}) \right) \left( w_{t+1} + \lambda_{t+1} \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} \right) + \lambda_{t+1} - \lambda_{1,t+1} \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} \left( \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} - 1 \right) \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} - \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} \left( \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} - 1 \right) \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} = 0. \tag{5}
\]

Note that, in general, this commitment solution is not time consistent: Ex post, the Ramsey government does not take into account the welfare effects working through \( \partial S_{t+i}/\partial \lambda_{t+i} \). On the other hand, in a steady state with strictly positive tax rates, the terms in the second and third line on the left-hand side of (5) vanish. This implies that \( c_1 = c_2 \), such that all terms featuring the effect of taxes on savings in the preceding period—and thus the source of time inconsistency—drop out of the first-order condition. We will see below that the condition \( c_1 = c_2 \) also characterizes the solution of the Ramsey program without commitment, thereby confirming that the time inconsistency feature indeed vanishes in steady state.

The Ramsey program without commitment differs from the case with commitment in that future tax decisions are functions of the state. This has two implications: First, changes in savings do not only affect future wages and returns, but also future tax rates. The expression for \( I_t \) is therefore replaced by

\[
I_t^R = \nu \left[ \beta \left( w_{t+1}' (\tau_{t+1} - 1) + \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} \right) \left( u_{t+1}' - u_{t+1}' \right) + \right] \nonumber \nu^2 \beta^2 \left[ w_{t+2}' (\tau_{t+2} - 1) + \frac{\partial S_{t+2}}{\partial \lambda_{t+2}} \right] \left( u_{t+2}' - u_{t+2}' \right) \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} \left( \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} - 1 \right) \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} - \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} \left( \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} - 1 \right) \frac{\partial S_{t+1}}{\partial \lambda_{t+1}} = 0.
\]

Second, the effect of a change in the tax rate on the savings decision in the previous period is no longer present. Differentiating \( R(s_{t-1}; \hat{\tau}(\cdot)) \) with respect to \( \tau_t \) yields the first-order condition

\[
\nu w_t \left( u'(c_{2,t}) - u'(c_{1,t}) \right) + \nu \frac{\partial S_t}{\partial \lambda_t} \frac{\partial S_t}{\partial \lambda_t} \frac{\partial S_t}{\partial \lambda_t} + \lambda_t = 0, \quad \lambda_t \tau_t = 0, \tag{6}
\]

where \( \lambda_t \) denotes the non-negative multiplier on the constraint that \( \tau_t \) be non-negative. Combining and manipulating the first-order conditions of two successive “selves” of the government, the Ramsey policy without commitment can be characterized more compactly. In the Appendix, we follow a parallel approach to derive the generalized Euler equation (see Laibson (1997), Krusell, Kurisu and Smith (2002), and Klein, Krusell and Rios-Rull (2003)). This yields

\[
\nu w_t \left( u'(c_{2,t}) - u'(c_{1,t}) \right) + \lambda_t - \lambda_{t+1} \beta \nu \frac{\partial S_t}{\partial \lambda_t} \frac{\partial S_t}{\partial \lambda_t} \frac{\partial S_t}{\partial \lambda_t} = 0, \tag{7}
\]

implying

\[
\tau_t > 0 \Rightarrow c_{1,t} \geq c_{2,t}, \nonumber \tau_{t+1} > 0 \Rightarrow c_{1,t} \leq c_{2,t}.
\]

\footnote{We use the facts that \( \frac{\partial S_{t+i}}{\partial \lambda_{t+i}} (\tau_{t+i} - 1) = \frac{\partial S_{t+i}}{\partial w_{t+i}} w_{t+i} \) and \( \nu \beta w_{t+i}' [dS_{t+i}/dw_{t+i} I_{t+i}^c + (\tau_{t+i} - 1) (u'(c_{2,t+i}) - u'(c_{1,t+i}))] = I_{t+i-1}^c. \) }
Not surprisingly, (7) resembles (5). The difference between the two conditions is due to the
fact that a change of tax rate affects savings in the preceding period only in the case with
commitment. (A comparison of (6) and (7) shows that $\lambda_{t+1} = 0 \Rightarrow I_t^R = 0$. This result hinges
on the constant returns to scale assumption; without that assumption, equation (7) would feature
an additional term.)

The first order conditions characterizing the politico-economic equilibria under rational
expectations and double-myopia closely resemble (6). In both cases, political influence enters the
picture ($\omega^{young}$ may differ from $\omega^{old}$), and some effects internalized by the Ramsey planner
without commitment are no longer accounted for, resulting in changes of $I_t$.\footnote{To be comparable with our earlier definition, we normalize the welfare effect of a marginal increase in savings
by the weight of young voters in the political process; that is, the non-normalized welfare effect as of time $t$
due to a marginal increase in savings in period $t$ equals $I_t \omega^{young}$.} In particular, with rational
expectations, all welfare effects accruing to generations born after period $t$ remain un-
accounted for: \( I_t \) reduces to $I_t^Y \equiv \nu \beta \left[ w^t_{t+1}(\tau_{t+1} - 1) + w^t_{t+1}\tau^{'t}_{t+1} \right] u^t_{2,t+1}$. Under double-myopia, the same also holds true for the general equilibrium effects in period $t + 1$ and the policy effect
$\tau^{'t}_{t+1}$, i.e., \( I_t \) reduces to zero. We thus have\footnote{$S_t$ now serves as shorthand notation for $S(w_t(1 - \tau_t); \tau(\cdot))$.}

\begin{equation}
\nu \frac{dS_t}{d\tau_t} \omega^{young} I_t^Y + \lambda_t = 0, \quad \lambda_t \tau_t = 0, \tag{8}
\end{equation}

for the politico-economic equilibrium under rational expectations, and

\begin{equation}
\nu \frac{dS_t}{d\tau_t} \omega^{old} u^t(c^t_{2,t}) - \omega^{young} u^t(c^t_{1,t}) + \lambda_t = 0, \quad \lambda_t \tau_t = 0, \tag{9}
\end{equation}

for the politico-economic equilibrium under double-myopia.

We now characterize the different equilibria in more detail.

**Proposition 1.** Consider the Ramsey equilibrium without commitment. Suppose that strictly
positive intergenerational transfers are sustained.

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

(ii) In steady state, $\beta R = 1$ (such that the economy is dynamically efficient). The steady
state is unique.

(iii) The steady state of the Ramsey equilibrium with commitment and strictly positive
intergenerational transfers is identical to the steady state of the Ramsey equilibrium without
commitment.

Results (i) and (ii) follow directly from (7). By setting $c_{1,t} = c_{2,t}$, the Ramsey government
equals its marginal rate of substitution between the consumption of young and old house-
holds with its corresponding marginal rate of transformation. The simplicity of this equilibrium
condition hinges on the constant returns to scale assumption as well as the equality of the
Ramsey government’s and households’ intertemporal discount factors. (The matching discount
factors imply that the marginal rate of substitution equals $\nu u^t(c_{1,t})/u^t(c_{2,t})$. Constant returns
to scale imply—as discussed earlier—that $I_t^R = 0$, such that the marginal rate of substitution
only reflects the direct effect of a resource transfer from young to old; it therefore equals $\nu$.)
Dynamic efficiency in steady state follows from $\beta R = 1 > \beta \nu$. Uniqueness follows from the fact
that $R$ and therefore $s$ and $w$ are pinned down by the condition $\beta R = 1$. The steady-state tax rate
is pinned down from $w(1 - \tau) - s = sR + \nu \tau w$ (since $c_1 = c_2$). Result (iii) follows from
(5). The time inconsistency property disappears in steady state because equalized consumption,
matching discount factors, and constant returns to scale imply that the social welfare effect of a marginal change in savings is zero, \( I_t^e = 0 \). Whether the government accounts for the effect of a change in tax rate on the previous period’s savings thus makes no difference to the optimal policy.

To state the next result, we define the benefit to a young household at \( t \) of increasing the tax rate as \( B_t = (dS_t/dr_t) R^t \). We also define the relative weight of the old in the political process as \( \omega = \omega_{old} / \omega_{young} \).

**Proposition 2.** Consider the politico-economic equilibrium under rational expectations. Suppose that strictly positive intergenerational transfers are sustained and \( B_t > 0, \omega \geq 1 \).

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

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

Results (i) and (ii) follow directly from (8). With \( c_1 < c_2 \) and therefore \( w(1 - \tau) - s < sR + \nu \tau \omega \), the steady-state interest rate is higher, savings are lower, and the tax rate is higher than in the Ramsey equilibrium.

If lower savings indirectly benefit young voters and higher taxes depress savings, then redistribution from young to old beyond the extent chosen by the Ramsey government (\( c_1 = c_2 \)) constitutes the vote maximizing platform in the electoral competition, even if \( \omega = 1 \). Young voters are willing 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 Ramsey government, 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.

The finding that the steady-state tax rate in politico-economic equilibrium is higher than in the Ramsey equilibrium stands in striking contrast to most of the literature where the politico-economic equilibrium sustains “too low” transfers, unless backed by a powerful trigger strategy. The present model turns this finding upside-down: under the assumption of probabilistic voting, the politico-economic equilibrium sustains “too high” transfers. Trigger strategies may thus be needed to prevent too high intergenerational transfers from being implemented.

**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 \( \omega \geq 1 \), \( \beta R > 1 \) (such that the economy is dynamically efficient); in steady state with \( \omega = 1 \), the allocation is the same as under the Ramsey policy.

Results (i) and (ii) follow directly from (9). If political influence is balanced (\( \omega = 1 \)), then the political process equates the consumption of old and young households in the same way as the Ramsey government. This surprising result 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 policy repercussions of concern to current young voters. Since in steady state, the interest rate is the same in the Ramsey and the double-myopic equilibrium, the condition \( c_1 = c_2 \) and therefore \( w(1 - \tau) - s = sR + \nu \tau \omega \) implies that the two equilibria are identical. Not surprisingly, 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 the following:

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

Under Assumption 1, household savings are given by

$$s_t = w_l(1 - \tau_l) - \frac{w_l(1 - \tau_l) + \nu w_t + \beta + \nu \omega + \alpha \omega^{t+1}/(1 - \alpha)}{1 + \beta}$$

$$= A\alpha \beta \frac{1}{1 + \beta} \left( \frac{s_{t-1}}{\nu} \right)^{1-\alpha} (1 - \tau_l) - \frac{\omega s_{t+1} \tau_l}{(1 + \beta)(1 - \alpha)},$$

$$\Rightarrow s_t = A \alpha \left( \frac{s_{t-1}}{\nu} \right)^{1-\alpha} (1 - \tau_l) - \frac{(1 - \alpha) \beta}{(1 - \alpha)(1 + \beta) + \alpha \tau_{t+1}} \equiv s_{t-1}^{1-\alpha} \cdot z(\tau_t, \tau_{t+1}),$$

(10)

implying

$$c_{1,t} = s_{t-1}^{1-\alpha} A\nu^{\alpha-1}(1 - \tau_l)(1 - \alpha(1 - \tau_{t+1})) \equiv s_{t-1}^{1-\alpha} \cdot \gamma(\tau_t, \tau_{t+1}),$$

$$c_{2,t} = s_{t-1}^{1-\alpha} A\nu^{\alpha}(1 - \alpha(1 - \tau_l)) \equiv s_{t-1}^{1-\alpha} \cdot \delta(\tau_t),$$

$$B_t = \frac{(1 - \alpha) \alpha \beta (1 - \tau_{t+1})}{(1 - \tau_l)(1 - \alpha(1 - \tau_{t+1}))}.$$

Note that $B_t > 0 \forall \tau_t, \tau_{t+1} \in [0, 1]$, and that the ratio of $c_{2,t}/c_{1,t}$ in autarky (i.e., if tax rates are set to zero) is given by

$$\hat{\epsilon} = \frac{(1 - \alpha)(1 + \beta)\nu}{\alpha},$$

which is independent of the stock of savings and therefore constant throughout the transition.

To derive our next result, it is useful to define $V(s_{t-1}, \tau_t, \{\tau_s\}_{s=t+1}^{\infty})$ to be the objective of a Ramsey government that inherits per-capita savings of the old equal to $s_{t-1}$, faces a sequence of flat, but potentially time-varying tax functions (imposed by the future selves of the government) given by $\{\tau_s\}_{s=t+1}^{\infty}$, and imposes a current tax rate of $\tau_t$. Differentiating $V(\cdot)$ with respect to $\tau_t$ yields

$$\frac{dV(s_{t-1}, \tau_t, \{\tau_s\}_{s=t+1}^{\infty})}{d\tau_t} = \frac{\alpha - (1 + \beta)\nu - \alpha \tau_t(1 + \nu) - \alpha \tau_t(1 + \nu)}{(1 - \alpha)(1 + \beta)\nu(1 - \alpha(1 - \tau_l))(1 - \tau_l)}.$$

\(^{18}\text{Under Assumption 1, } V(\cdot) \text{ is given by}

$$V(s_{t-1}, \tau_t, \{\tau_s\}_{s=t+1}^{\infty}) = \ln[s_{t-1}^{1-\alpha} \delta(\tau_t)] + \nu \ln[s_{t-1}^{1-\alpha} \gamma(\tau_t, \tau_{t+1})] + \beta \ln[(s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1})] + \beta \ln[(s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1})]$$

$$+ \nu \beta \ln((s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1}))$$

$$+ \ln \ln(\tau_l) + \ln \ln(\gamma(\tau_t, \tau_{t+1})) + \ln \ln(z(\tau_t, \tau_{t+1})) + \ln[(1 - \alpha) \nu \beta (1 + \nu) + (1 - \alpha) \nu \beta (1 + \nu) + \ldots]$$

$$+ T(s_{t-1}, \tau_t)$$

$$= \ln[s_{t-1}^{1-\alpha} \gamma(\tau_t, \tau_{t+1})] + \nu \ln[s_{t-1}^{1-\alpha} \gamma(\tau_t, \tau_{t+1})] + \beta \ln[(s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1})]$$

$$+ \nu \beta \ln((s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1}))$$

$$+ \ln \ln(\tau_l) + \ln \ln(\gamma(\tau_t, \tau_{t+1})) + \ln \ln(z(\tau_t, \tau_{t+1})) + \ln[(1 - \alpha) \nu \beta (1 + \nu) + (1 - \alpha) \nu \beta (1 + \nu) + \ldots]$$

$$+ T(s_{t-1}, \tau_t)$$

$$= \ln[s_{t-1}^{1-\alpha} \gamma(\tau_t, \tau_{t+1})] + \nu \ln[s_{t-1}^{1-\alpha} \gamma(\tau_t, \tau_{t+1})] + \beta \ln[(s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1})]$$

$$+ \nu \beta \ln((s_{t-1}^{1-\alpha} z(\tau_t, \tau_{t+1}))^{1-\alpha} \gamma(\tau_{t+1}))$$

$$+ \ln \ln(\tau_l) + \ln \ln(\gamma(\tau_t, \tau_{t+1})) + \ln \ln(z(\tau_t, \tau_{t+1})) + \ln[(1 - \alpha) \nu \beta (1 + \nu) + (1 - \alpha) \nu \beta (1 + \nu) + \ldots]$$

$$+ T(s_{t-1}, \tau_t),$$

where $T(s_{t-1}, \tau_t)$ denotes terms independent of both $s_{t-1}$ and $\tau_t$. Substituting the definitions of $\gamma(\tau_t, \tau_{t+1})$, $\delta(\tau_t)$, and $z(\tau_t, \tau_{t+1})$ and differentiating yields the desired result.
Note that, for all $\tau_t \in [0,1)$: (i) if $\hat{c} > 1$, this derivative is strictly negative; (ii) if $\hat{c} < 1$, this derivative is strictly positive up to the strictly positive tax rate rendering the numerator of the derivative equal to zero, and negative thereafter. These observations lead to

**Proposition 4.** Consider the Ramsey equilibrium without commitment. Suppose that Assumption 1 is satisfied.

(i) There exists an equilibrium with a flat policy function. In this equilibrium, the steady state is globally stable and the transition to steady state is unique. If $\hat{c} < 1$, then the policy function is strictly positive and given by

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

and the steady state level of savings is $s^R \equiv (A(1 - \alpha)\beta)^{1/\alpha}\nu$, which is lower than the autarky level of savings, $s^A \equiv (A\beta\nu^{\alpha-1}/(1 + \beta))^{1/\alpha}$. If $\hat{c} \geq 1$, then the policy function is given by the autarky policy function, $\tau(s) = 0$, and steady-state savings are at their autarky level $s^A$.

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

(iii) The steady-state results extend to the Ramsey equilibrium with commitment.

Result (i) follows from (10), the properties of $dV(\cdot)/d\tau_t$, and Lemma 1 in the Appendix which proves stability. A flat policy function is consistent with equilibrium because, under our functional form assumptions, the tax rate maximizing $V(\cdot)$ is independent of the state variable as long as future tax functions are also independent of the stock of savings. Moreover, whether this optimal tax rate is strictly positive or not, is closely linked to the autarky consumption ratio $\hat{c}$. If $\hat{c} > 1$, such that the autarky consumption of old households exceeds the autarky consumption of young households, a social security system is unable to push the consumption ratio towards unity because taxes are restricted to be non-negative. If $\hat{c} < 1$, however, such that the autarky consumption of old households falls short of the autarky consumption of young households, social security contributions and benefits can be employed to raise the consumption ratio towards unity and thereby maximize the objective function of the Ramsey government.

The tax rate achieving this goal, $\tau^R$, is independent of the stock of savings and therefore constant throughout the transition. Given that $\{\tau_t\}$ equals either $\{\tau^R\}$ or $\{0\}$, 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 $\zeta(\tau_t, \tau_{t+1})$. Result (ii) is proved in the Appendix (for the case $\hat{c} < 1$ such that the steady-state tax rate is positive). It hinges on the Cobb-Douglas property by which capital income, labor income, and (therefore) transfers are all proportional to the wage rate. Result (iii) follows from Proposition 1.

A clearer intuition about the factors determining the size of $\tau^R$ can be gained by rewriting the relative consumption of old and young households under the assumption of a constant tax rate $\tau$ as

$$\frac{c_{2,t}}{c_{1,t}} = \frac{\hat{c} + \nu\tau}{1 - \tau}. \tag{11}$$

We know that the Ramsey government aims at pushing this ratio towards unity. As long as $\hat{c} < 1$, this is possible. Moreover, in that case, the optimal tax rate increases in $\alpha$ (higher $\alpha$ reduces $Rs/(uv)$ and therefore $\hat{c}$), decreases in $\beta$ (higher $\beta$ raises the savings rate and therefore $\hat{c}$), and decreases in $\nu$ (for a given ratio of total capital to labor income, higher $\nu$ raises that ratio in per-capita terms and therefore $\hat{c}$; moreover, higher $\nu$ raises transfers per capita of the old and therefore $c_{2,t}$).
To derive our next result, we follow a parallel approach. Define $W(s_{t-1}, \tau_t, \tau_{t+1})$ to be the objective of a vote-maximizing policy maker that inherits per-capita savings of the old equal to $s_{t-1}$, faces a flat tax function (imposed by its successor) given by $\tau_{t+1}$, and imposes a current tax rate of $\tau_t$. Differentiating $W(\cdot)$ with respect to $\tau_t$ and defining

$$
\hat{d} \equiv \left(1 - \alpha\right)(1 + \beta(1 - \alpha)) \frac{\nu}{\alpha \omega} = \frac{\alpha(1 + \beta(1 - \alpha))}{(1 + \beta)\omega} \quad [< \hat{c} \text{ for } \omega \geq 1]
$$
yields\(^{19}\)

$$
\frac{dW(s_{t-1}, \tau_t, \tau_{t+1})}{d\tau_t} = \alpha \omega (1 - \hat{d}) - \alpha \tau_t (\omega + \nu(1 + \beta(1 - \alpha))) \left(1 - (1 - \tau_t)(1 - \tau_{t+1})\right).
$$

Note that, for all $\tau_t \in [0, 1]$:

(i) if $\hat{d} > 1$, this derivative is strictly negative; (ii) if $\hat{d} < 1$, this derivative is strictly positive up to the strictly positive tax rate rendering the numerator of the derivative equal to zero, and negative thereafter. These observations lead to

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

(i) There exists an equilibrium with a flat policy function. In this equilibrium, the steady state is globally stable and the transition to steady state is unique. If $\hat{d} < 1$, then the policy function is strictly positive and given by

$$
\tau(s) = \tau^W \equiv \frac{\alpha \omega - (1 - \alpha)(1 + \beta(1 - \alpha)) \nu}{\alpha(\omega + \nu(1 + \beta(1 - \alpha)))} \quad [> \tau_R \text{ for } \omega \geq 1];
$$

the steady-state level of savings is

$$
s^W \equiv \nu \left(\frac{A(1 - \alpha)\beta}{\omega + \beta \nu(1 - \alpha)}\right)^{1/\alpha} \quad [< s^R, s^A \text{ for } \omega \geq 1].
$$

If $\hat{d} \geq 1$, then the policy function is given by the autarky policy function, $\tau(s) = 0$, and steady-state savings are at their autarky level $s^A$.

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

Result (i) follows from (10), the properties of $dW(\cdot)/d\tau_t$, and Lemma 1 in the Appendix which proves stability. Again, 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 stock of savings. Note that the politico-economic equilibrium features a strictly higher tax rate than the Ramsey equilibrium (at least if $\omega \geq 1$) if $\hat{d} < 1$, and the same tax rate (equal to zero) if $\hat{d} \geq 1$ (this follows from the fact that $\hat{c} < 1 \Rightarrow \hat{d} < 1$. Given that $\{\tau_t\}$ equals either $\{\tau^W\}$ or $\{0\}$, the complete transition of all endogenous variables is fully characterized. Result (ii) is proved in the Appendix. Once more, it hinges on the Cobb-Douglas property.

\(^{19}\)Under Assumption 1, $W(\cdot)$ is given by

$$
W(s_{t-1}, \tau_t, \tau_{t+1}) = \omega \ln[s_{t-1}^{\hat{d} / \alpha} \delta(\tau_t)] + \nu[\ln[s_{t-1}^{\hat{d} / \alpha} \gamma(\tau_t, \tau_{t+1})] + \beta \ln[s_{t-1}^{\hat{d} / \alpha} x(\tau_t, \tau_{t+1})]^{1-\alpha} \delta(\tau_{t+1})]
$$

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

Substituting the definitions of $\gamma(\tau_t, \tau_{t+1}), \delta(\tau_t)$, and $x(\tau_t, \tau_{t+1})$ and differentiating yields the desired result.
Similarly to \( \tau^R \), \( \tau^W \) increases in \( \alpha \) and decreases in \( \beta \) and \( \nu \). To better understand this result, it is instructive to return to the condition characterizing \( \tau^W \) in the general case, equation (8). For \( \lambda_t = 0 \) and under Assumptions 1 and a constant tax rate \( \tau \), this condition reduces to

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

where the right-hand side equals \( B_1c_{2,t}/w_t \) and measures the benefit from depressing capital accumulation and thereby tilting the consumption profile. This incentive to raise taxes increases in the capital share, \( 1 - \alpha \), 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 (11)) also increases in these three parameters, reducing the incentive to raise taxes. The latter effect dominates.

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

\[
\frac{w\tau^W\nu}{w\nu + Rs} = \alpha\tau^W.
\]

This implies that the pension share is increasing in \( \alpha \), and therefore in GDP per capita (since this increases in \( \alpha \)), and decreasing in \( \nu \). Both of 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 explanatory power of the model extends beyond these linear relationships. In particular, the model replicates the apparently non-linear empirical relationship between the share of the elderly and public pension payments per retiree, corresponding to \( w\tau^W\nu \) in the model. For the 1880–1930 period, Lindert (1994) estimates an elasticity of the pension share in GDP with respect to the share of the elderly that is larger than unity. Razin, Sadka and Swagel (2002), on the other hand, argue based on data for the United States and 12 European countries over the period 1965–92 that the dependency ratio (a crude measure of the share of the elderly) is negatively related to per capita transfers (a crude measure of social security benefits). Finally, the OECD data suggest a hump-shaped relationship between the share of the elderly and public pension payments per retiree (Lindert, 1996). In fact, the model yields exactly such an inverse-U shaped relationship and can therefore account for the data in all three samples.

Turning to the double-myopic equilibrium, define \( M(s_{t-1}, \tau_t, \tau_{t+1}) \) to be the objective of a double-myopic policy maker that inherits per-capita savings of the old equal to \( s_{t-1} \), faces the next-period tax rate \( \tau_{t+1} \) (perceived to be unaffected by her own actions), and imposes a current tax rate of \( \tau_t \). \( M(\cdot) \) differs from \( W(\cdot) \) because next period’s interest and wage rates (which affect the second period consumption of young households) are now perceived to be identical to the current factor prices:

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

Differentiating \( M(\cdot) \) with respect to \( \tau_t \) and applying the household’s Euler equation yields

\[
\frac{dM(s_{t-1}, \tau_t, \tau_{t+1})}{d\tau_t} = \nu w_t \left( \frac{\omega}{c_{2,t} - 1} - \frac{1}{c_{1,t}} \right) = \nu w_t \left( \frac{\omega}{\nu w_t \tau_t + \nu w_t (1 - \alpha) / \alpha} - \frac{1}{w_t (1 - \tau_t) - s_t} \right)
= \left( \frac{\nu (1 + \beta)}{\alpha \tau_t + 1 - \alpha} - \frac{1}{1 - \tau_t + \nu \tau_t / R_t} \right). \]
Note that, for all $\tau_t \in [0,1)$ and $R_t > \nu$: (i) if $\hat{\tau} > \omega$, this derivative is strictly negative; (ii) if $\hat{\tau} < \omega$, this derivative is strictly positive up to the strictly positive tax rate rendering the derivative equal to zero, and negative thereafter. This implies

**Proposition 6.** Consider the politico-economic equilibrium under double-myopia. Suppose that Assumption 1 is satisfied and the economy is dynamically efficient.

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

(ii) If $\hat{\tau} < \omega$, then no flat policy function constitutes an equilibrium policy function. A continuously differentiable policy function is rather upward sloping around the steady state.

Result (i) follows from two observations: First, if $\hat{\tau} < \omega = 1$, 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 $\tau_t = \tau^R$. (If $\hat{\tau} \geq 1$, the autarky equilibrium results.) Result (ii) follows from the observation that the derivative of $M(\cdot)$ is a function of the state variable (through the interest rate) and that $\hat{\tau} < \omega$ implies strictly positive taxes.

To assess the quantitative importance of the motive to depress aggregate savings, we calibrate the model. In particular, we impose Assumption 1 and set $\alpha$ to 0.7, $\beta$ to 0.98$^{30}$, and $\nu$ to the gross growth rate of the US population between 1970 and 2000 ($1.384$)$^{20}$ Table 1 lists 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>
<thead>
<tr>
<th></th>
<th>Politico-economic equilibrium</th>
<th>Ramsey equilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau$</td>
<td>0.1187</td>
<td>0.0349</td>
</tr>
<tr>
<td>capital-labor ratio</td>
<td>0.0563</td>
<td>0.0753</td>
</tr>
<tr>
<td>$R$</td>
<td>2.2484</td>
<td>1.8332 ($= \beta^{-1}$)</td>
</tr>
<tr>
<td>$w$</td>
<td>0.2953</td>
<td>0.3223</td>
</tr>
<tr>
<td>$c_1$</td>
<td>0.1823</td>
<td>0.2067</td>
</tr>
<tr>
<td>$c_2$</td>
<td>0.2236</td>
<td>0.2067</td>
</tr>
<tr>
<td>$U$</td>
<td>-2.5190</td>
<td>-2.4362</td>
</tr>
</tbody>
</table>

For the chosen parameter values, $\hat{\tau} < 1$ (the autarky consumption of young households exceeds the autarky consumption of old households) and $\hat{\beta} < 1$. Both the Ramsey government and the political process therefore implement strictly positive tax rates.$^{21}$ Moreover, the difference between the tax rates is large: With nearly 12 percent, the tax rate implemented in politico-economic equilibrium is more than three times as high as in the Ramsey equilibrium (or, equivalently, the steady state of the double-myopic equilibrium). Ceteris paribus, the positive difference between $\tau^W$ and $\tau^R$ increases further for lower values of $\alpha$ or higher values of

$^{20}$These are standard values. See, for example, Heaton and Lucas (2000) who set $\alpha$ to 0.7 and $\beta$ to 0.95 or 0.99. The population growth rate is taken from the US census. We also set $A$ and $\omega$ to unity.

$^{21}$Ceteris paribus, tax rates are strictly positive for $\alpha > 0.6814$, $\beta < 0.9875^{30}$, and $\nu < 1.5098$ in the Ramsey case or $\alpha > 0.6251$, all $\beta \leq 1$, and $\nu < 2.0062$ in the probabilistic voting case.
$\beta$ or $\nu$. 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. Nevertheless, on an annual basis, the difference between the interest rates is modest (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 of consumption by more than 5 percent.

3 Robustness: Elastic Labor Supply and Tax Distortions

The purpose of this extension is to examine the implications of endogenous labor supply and tax distortions. Superficially, these implications seem straightforward: since tax distortions increase the cost to the young of funding social security, social security is still likely to be sustained, although possibly at a lower level. This argument is insufficient, however. It neglects the important question of whether voters still choose to sustain social security as opposed to another policy instrument in such a different environment.

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. To keep the notation simple, we assume that a young household’s felicity function is separable in consumption and leisure, $x_t$, and 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 $\theta_t$, 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 (or targeted government spending for) 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 $c_{2,t+1} = s_t R_{t+1} + b_{t+1}^*$. The tax rate $\theta$ 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 depressed 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

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

$$u'(c_{1,t}) w_t(1-\tau_t-\theta_t) = v'(x_t),$$

subject to the budget set described earlier. For convenience, we treat $T_t$ and the sum of the two tax rates, $\sigma_t \equiv \tau_t + \theta_t$, rather than $\tau_t$ and $\theta_t$, 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 inherited savings, $s_{t-1}$, aggregate labor supply, and aggregate savings as well as $\sigma_t$ 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

$$
S(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)) = s(s_{t-1}, X(\cdot), S(\cdot), \sigma_t, T_t; \sigma(S(\cdot)), T(S(\cdot)), R(S(\cdot)), w(S(\cdot)))
$$

$$
X(s_{t-1}, \sigma_t, T_t; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)) = x(s_{t-1}, X(\cdot), S(\cdot), \sigma_t, T_t; \sigma(S(\cdot)), T(S(\cdot)), R(S(\cdot)), w(S(\cdot)));
$$

$$
\forall s_{t-1}, \sigma_t, T_t \geq 0.
$$

The objective pursued by the political process is now given by

$$
W(s_{t-1}; \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot)) = \max_{\sigma_t, T_t} \left[ \omega^\text{old} u(c_{2,t}) + \omega^\text{young} \nu(u(c_{1,t}) + v(x_t) + \beta u(c_{2,t+1})) \right]
$$

subject to

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

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 $x_{t+1} = X(s_t, \sigma(s_t), T(s_t); \sigma(\cdot), T(\cdot), R(\cdot), w(\cdot))$. 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 of the support for intergenerational transfers? We can reject the second alternative under plausible conditions. To see this, consider the effect on $W$ of a marginal increase in $T_t$. This effect 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; it equals \( ^{23} \)

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

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

---

\(^{22} \)Next period’s wage and return on savings depend both on the inherited capital stock (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.

\(^{23} \) $S_t$ 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.
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 which, due to constant returns to scale, are proportional to the change in wages.\(^{24}\)

(The effects on \(W\) of a marginal increase in \(\sigma_t\) closely resemble the terms above: The term in i. is multiplied by \(-w_t(1 - x_t)\), and the expressions in ii. and iii. have the derivatives with respect to \(T_t\) replaced by the derivatives with respect to \(\sigma_t\).)

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 to the young \((I_t^W < 0\) in our earlier notation). 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 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, as long as \(\omega u'(c_{2,t}) > u'(c_{1,t})\) (a sufficient condition for which is \(c_{2,t} \leq c_{1,t}\) and \(\omega \geq 1\)), the political process does not sustain 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 candidates prefer to fund social security. The intuition for this result is straightforward: refunding the tax revenue to the young would take resources from the needy old, raise savings (which runs counter to the strategic interest of young voters), and alter factor prices in the current period. While benefiting young voters (through higher wages), the last effect would hurt old voters (through lower income from savings and pensions due to factor price effects, and lower pensions due to lower labor supply).

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 noisy relationship between candidates’ platforms and citizens’ voting behavior, or probabilistic voting—strikes us as a much more natural and realistic assumption than the supposition of a deterministic link between policy platforms and votes cast that underlies the median voter framework. Not only is the probabilistic voting assumption realistic; introducing it in the standard Diamond (1965) model preserves that model’s tractability and delivers intuitive and novel results in a strikingly transparent fashion.

Throughout the paper, we have imposed the Markov assumption: political and economic choices are functions of the single state variable, the capital-labor ratio, only. This contrasts with the more common assumption in the social security literature, that support for intergenerational transfers depends on intricate trigger strategies. To introduce trigger strategies in our setup is possible; as we have argued, however, it reverses the traditional finding according to which trigger strategies are needed to improve on suboptimally low intergenerational transfers.

Since old voters benefit from publicly funded pensions, the political sustainability of a social security system, and its size, crucially depend on the cost the system imposes on (the majority of) young voters. These cost are likely to be smaller than suggested by simple present value

\(^{24}\)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.
calculations: Social security creates benefits for young voters by enabling them to monopolize their factor supplies, in particular aggregate savings, and these benefits partially compensate for the direct tax cost. How large the benefits are depends on the elasticity of factor prices and thus, among other factors, the size and openness of the economy. One prediction of the model is therefore that small, open economies have smaller social security systems than large or closed economies (holding other factors as for instance the exposure to risk constant), and that time variation in the openness of the capital account is reflected in time varying political support for social security. The expansion of social security systems during the interwar period with its sharp reversal of capital market globalization may be interpreted as supportive evidence in that respect. Another prediction of the model concerns the role of population growth: An increase in the old-age dependency ratio should be accompanied by the introduction or expansion of social security programmes, while pensions per retiree as a function of the old-age dependency ratio should be hump-shaped. Both of these model 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.

While the political process in our model is "inclusive" in the sense of representing the interests of all voters, it is not inclusive enough from a 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 partially shifting the cost of the social security system to future generations. As a consequence, the social security system is too large (rather than too small as frequently suggested) relative to a system balancing the interests of current and future generations. Both of 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.
A  Generalized Euler Equation

We derive the generalized Euler equation (GEE) from the conditions

\[
\tau(s) = \arg \max_{\theta} u(sR(s) + \nu w(s)\theta) + \nu u(w(s)(1 - \theta) - K(s, \theta)) + \beta \nu V(K(s, \theta)) + \lambda(s)\theta, \\
V(s) = u(sR(s) + \nu w(s)\tau(s)) + \nu u(w(s)(1 - \tau(s)) - K(s, \tau(s))) + \beta \nu V(K(s, \tau(s))),
\]

where we use the shorthand notation \(K(s, \theta)\) for \(S(w(s)(1 - \theta); \tau(\cdot))\) and \(\lambda(s)\) denotes the multiplier on the non-negativity constraint of \(\theta\). From the first equation, the optimal policy function satisfies

\[
w(s) \left( u'(c_2(s)) - u'(c_1(s)) \right) + \frac{\partial K(s, \tau(s))}{\partial \theta} \left( -u'(c_1(s)) + \beta V'(K(s, \tau(s))) \right) + \lambda(s)/\nu = 0,
\]

where \(c_1(s) \equiv w(s)(1 - \tau(s)) - K(s, \tau(s))\) and \(c_2(s) \equiv sR(s) + \nu w(s)\tau(s)\). Combining this first-order condition with the envelope condition from the second equation yields

\[
V'(s) = \nu w'(s)(\tau(s) - 1) \left( u'(c_2(s)) - u'(c_1(s)) \right) + u'(c_2(s)) \left( sR'(s) + \nu w'(s) \right)
+ R(s)u(c_2(s)) + \nu \frac{\partial K(s, \tau(s))}{\partial s} \left( \beta V'(K(s, \tau(s))) - u'(c_1(s)) \right).
\]

The same envelope condition, evaluated at \(\tau(s) = 0\), results in the case where \(\lambda(s) > 0\). Substitution of the first-order condition into the envelope condition and application of the household’s Euler equation yields the GEE\footnote{Primes on variables (as opposed to functions) denote values in the next period. We use the fact that}

\[

\nu w(s) \left( u'(c_2(s)) - u'(c_1(s)) \right) = -\lambda(s) + \beta \nu \frac{\partial K(s, \tau(s))}{\partial \theta} \left\{ \lambda(s') \left[ \frac{w'(s')(\tau(s') - 1)}{w(s')} \right] - u'(c_2(s)) \left( sR'(s) + \nu w'(s) \right) \right\}.
\]

Under constant returns to scale, the right-most term in the curly brackets vanishes.

B  Proofs

B.1 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) \(\tau'(s_l) = 0 \forall s_l\) or (ii) \(\tau'(s_l) \geq 0 \forall s_l\) and \((1 - \alpha)(1 - \tau(s)) - s\tau'(s) \geq 0.\)
Proof. Using the functional form assumptions, we have\(^{26}\)

\[
\frac{ds_t}{ds_{t-1}} = \frac{\beta \ A \ s_{t-1}^{-\alpha} (1 - \alpha)(1 - \tau(s_{t-1})) - s_{t-1} \tau'(s_{t-1}))}{1 + \frac{\alpha \ \tau(s_{t-1})}{1 - \alpha \ 1+\beta} + s_{t-1} \ \frac{s_{t-1}^{\alpha}}{1 - \alpha \ 1+\beta}},
\]

\[
\frac{ds_t}{ds_{t-1}} \bigg|_{\text{steady state}} = \frac{\beta \ A \ (1 - \alpha) \nu^{1-\alpha} \left(1 + \frac{\alpha \ \tau(s)}{1 - \alpha \ 1+\beta} - s^{1-\alpha} \tau'(s)\right)}{1 + \frac{\alpha \ \tau(s)}{1 - \alpha \ 1+\beta} + s^{\alpha} \ \frac{s^{\alpha}}{1 - \alpha \ 1+\beta}}.
\]

If \(\tau'(s) = 0\), the second equation implies \(\frac{ds_t}{ds_{t-1}} \bigg|_{\text{steady state}} = 1 - \alpha\) (and therefore local stability). Moreover, if \(\tau'(s_t) = 0 \ \forall s_t\), then the first equation implies \(\frac{ds_t}{ds_{t-1}}\) 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 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 \((1 - \alpha)(1 - \tau(s)) - s \tau'(s) \geq 0\).

\[\square\]

B.2 Proof of Proposition 4

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

Along the transition to the steady state, \(c_{1,t} = c_{2,t},\ i.e.,\)

\[w_t(1 - \tau_t) \left(1 - \frac{\beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau(s_t)}\right) = \nu w_t \tau_t + s_{t-1} R_t.\]

Since \(s_{t-1} R_t = \nu w_t \frac{1 - \alpha}{\alpha}\), this implies the following relation between the current and the expected future tax rate:

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

(12)

In a first step, we now prove that convergent, non-oscillatory dynamics towards the steady state imply a flat \(\tau(s_t),\ i.e.,\ \tau(s_t) = \tau\) with \(\tau\) and \(s\) denoting steady-state values. Consider the first-order approximation

\[\tau(s_t) = \tau - \epsilon \tau',\]

\[\tau(s_{t-1}) = \tau - (\epsilon + \delta) \tau',\]

with \(\tau'\) 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 \(\epsilon \delta \geq 0\). A first-order approximation of (12) around the steady state yields

\[\tau'(\delta + \epsilon) \left(1 - \frac{\beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau}\right) - \tau'(1 - \tau) \frac{\beta(1 - \alpha) \epsilon}{(1 - \alpha)(1 + \beta) + \alpha \tau} = -\nu (\delta + \epsilon) \tau',\]

\[^{26}\text{The second equation follows from the steady-state relationship}\]

\[1 + \frac{s}{1 - \alpha \ 1+\beta} = \frac{\beta \ A \ s}{1 - \alpha \ 1+\beta} \frac{\alpha \ \tau(s)}{1 - \alpha \ 1+\beta} \nu \left(1 - \tau(s)\right).\]
or
\[
\tau' \delta \left( \nu + 1 - \frac{\beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau} \right) + \\
\tau' \epsilon \left( \frac{(1 - \tau)\beta(1 - \alpha)\alpha}{((1 - \alpha)(1 + \beta) + \alpha \tau)^2} - \left( 1 - \frac{\beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau} \right) - \nu \right).
\]

The term multiplying \( \delta \tau' \) on the left-hand side is positive. Using the steady-state expression for the tax rate and \( \beta R = 1 \), the term multiplying \( \epsilon \tau' \) on the right-hand side simplifies to
\[
\frac{(1 + \nu)((1 - \alpha)\nu - R)}{1 - \alpha + R},
\]
which is negative since \( R = \beta^{-1} > \nu \). A stable, non-oscillatory steady state therefore requires \( \tau' \) to be (locally) equal to zero. The argument can be extended in a parallel fashion beyond the steady state; the policy function must thus be globally flat.

It remains to be shown that the (non-degenerate) steady state indeed features \( 0 \leq \frac{dS_t}{ds_{t-1}} \text{steady state} \leq 1 \) or, if not, that still \( \tau' = 0 \). Note first that a slope greater than one is ruled out by the fact that there is only one steady state and \( \lim_{t \to \infty} s_t < \infty \). Now consider the case of a negative slope: From (12), we have (a) \( \frac{dS_{t+1}}{dS_t} > 1 \). At the same time, under our assumption, (b) \( \frac{dS_{t}}{dS_{t-1}} \text{steady state} < 0 \). Assume first that (c) \( \tau' > 0 \). Take some \( s_{t-1} \) slightly lower than \( s \). From (b), \( s_{t} > s \). From (c), \( \tau(s_{t}) > \tau(s_{t-1}) \). Now take a slightly lower starting value, \( \delta_{t-1} < s_{t-1} \). From (b), \( \delta_{t} > s_{t} \). From (c), \( \tau(s_{t}) > \tau(s_{t-1}) > \tau(\delta_{t-1}) \), contradicting (a). Alternatively, assume that (d) \( \tau' < 0 \). Take some \( s_{t-1} \) slightly lower than \( s \). From (b), \( s_{t} > s \). From (d), \( \tau(s_{t}) < \tau(s_{t-1}) \). Now take a slightly higher starting value, \( s_{t-1} < \delta_{t-1} < s \). From (b), \( \delta_{t} < s_{t} \). From (d), \( \tau(s_{t-1}) > \tau(\delta_{t-1}) > \tau(\delta_{t}) > \tau(s_{t}) \), contradicting (a). We conclude that \( \tau' = 0 \). 

**B.3 Proof of Proposition 5**

**Proof.** We prove that, if \( \tau(s) \) is continuously differentiable and strictly positive, the equilibrium policy function must be flat. We assume that \( \omega = 1 \).

From (8), using the functional form assumptions, we find
\[
\frac{c_{1,t}}{c_{2,t}} = 1 + \frac{dS_t}{d\tau_t} \alpha \left( \frac{1 - \tau(s_t) - \frac{s_{t+1}}{s_{t}} \tau'(s_t)}{w_t} \right).
\]

From (10),
\[
\frac{dS_t}{d\tau_t} = \frac{w_t \beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau(s_t) + \alpha s_t \tau'(s_t)}.
\]

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

Together, these conditions imply
\[
\frac{(1 - \tau_t) \left( 1 - \frac{\beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau(s_t)} \right)}{\nu \left( \frac{1 - \alpha}{\alpha} + \tau_t \right)} = \\
1 - \frac{\beta(1 - \alpha)\alpha}{(1 - \alpha)(1 + \beta) + \alpha \tau(s_t) + \alpha s_t \tau'(s_t)} \left( 1 - \tau(s_t) - \frac{s_{t+1}}{s_{t}} \tau'(s_t) \right).
\]

(13)
Note that, since $B_t > 0$ implies $c_{1,t} < c_{2,t}$, we must have

\[(1 - \alpha)(1 + \beta) + \alpha \tau(s_t) + \alpha s_t \tau'(s_t) > 0. \quad (14)\]

(Suppose not. Then $dS_t/d\tau_t > 0$ and $1 - \tau(s_t) - \frac{\delta}{1 - \alpha} \tau'(s_t)$ must be negative, implying $\tau'(s_t) > 0$ and thus giving rise to a contradiction.) In a first step, we now prove that convergent, non-oscillatory dynamics towards the steady state imply a flat $\tau(s_t)$, i.e., $\tau(s_t) = \tau$ with $\tau$ and $s$ denoting steady-state values. Consider the first-order approximation

\[
\begin{align*}
\tau(s_t) &= \tau - \epsilon \tau', \\
\tau(s_{t-1}) &= \tau - (\epsilon + \delta) \tau',
\end{align*}
\]

with $\tau'$ 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

\[
\begin{align*}
a &\equiv \frac{\beta(1 - \alpha)}{(1 - \alpha)(1 + \beta) + \alpha \tau}, \\
b &\equiv \frac{1 - \alpha}{\alpha}, \\
c &\equiv \frac{\beta(1 - \alpha)\alpha}{(1 - \alpha)(1 + \beta) + \alpha \tau + \alpha s \tau'}, \\
c' &\equiv \frac{\beta(1 - \alpha)\alpha}{((1 - \alpha)(1 + \beta) + \alpha \tau + \alpha s \tau')^2}.
\end{align*}
\]

A first-order approximation of (13) around the steady state yields

\[
(\epsilon + \delta)\tau' \frac{1 - a}{v b} - \epsilon \tau' \frac{1 - \tau}{v b} \alpha a' + (\epsilon + \delta)\tau' \frac{1 - \tau}{v b^2} (1 - a) = -\epsilon \tau c \frac{2 - \alpha}{1 - \alpha} - \epsilon 2 \alpha \tau' c' \left(1 - \tau - \frac{s \tau'}{1 - \alpha}\right)
\]

or

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

for some positive term $y$. The term multiplying $\tau' \epsilon$ on the right-hand side is negative because of (14), and because

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

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, is satisfied 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 $\tau'$ to be (locally) equal to zero. The argument can be extended in a parallel fashion beyond the steady state; the policy function must thus be globally flat.

It remains to be shown that the (non-degenerate) steady state indeed features $0 \leq \frac{dS_t}{ds_{t-1}}|_{\text{steady state}} \leq 1$ or, if not, that still $\tau' = 0$. An argument parallel to that in the proof for the Ramsey case (replacing (12) by (13)) establishes this result. We conclude that $\tau' = 0$. \qed
References


