ESSAYS ON DYNAMIC POLITICAL ECONOMY

by

Zheng Song

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
Stockholm University

Monograph Series
No. 53

2005

THE INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
University of Stockholm

is an institute for independent scholarly research in the field of international economic relations.

DIRECTOR: Torsten Persson, Professor of Economics
DEPUTY DIRECTOR: Mats Persson, Professor of Economics

RESEARCH STAFF AND VISITING FELLOWS 2004/2005:

Professors
Lars Calmfors
Harry Flam
John Hassler
Henrik Horn
Assar Lindbeck
Mats Persson
Torsten Persson
Peter Svedberg
Fabrizio Zilibotti

Visiting Professors
Per Krusell
Kjetil Storesletten

Research Fellows
Nicola Gennaioli
Jesper Lindé (visiting)
Dirk Niepelt
David Strömberg
Jakob Svensson

Graduate Students
Martin Bech Holte
Martina Björkman
Alessandra Bonfiglioli
Thomas Eisensee
Giovanni Favara
Daria Finocchiario
Erika Färnstrand
Mathias Herzing
Anna Larsson
Martin Ljunge
Natalie Pienaar
José Mauricio Prado, Jr.
Virginia Queijo
Zheng (Michael) Song
Ulrika Stavlöt
Mirco Tonin
Gisela Waisman

Visiting Fellows
Gino Gancia
University Pompeu Fabra
Martín Gonzalez-Eiras
Universidad de San Andrés
Jean Imbs
University of Lausanne and London Business School
Paul Klein
University of Western Ontario
Jo Thori Lind
University of Oslo
Egil Matsen
Norwegian University of Science and Technology, Trondheim
Andrea Moro
University of Minnesota
Anders Olofsgård
Georgetown University
Gérard Roland
University of California, Berkeley
Betsey Stevenson
University of Pennsylvania
Lars E.O. Svensson
Princeton University
Harald Uhlig
Humboldt University
Davide Ticchi
Università degli Studi di Urbino "Carlo Bo"
Justin Wolfers
University of Pennsylvania
Joseph Zeira
Hebrew University Jerusalem

MAIL ADDRESS: S-106 91 STOCKHOLM, SWEDEN
STREET ADDRESS: Universitetsvägen 10 A, 8th floor
TELEPHONE: + 46 8 16 20 00, **TELEFAX:** + 46 8 16 14 43

ESSAYS ON DYNAMIC POLITICAL ECONOMY

by

Zheng Song

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES
Stockholm University

September 2005

Institute for International Economic Studies
Stockholm University

ISBN 91-7155-106-9

Printed by Akademitryck AB
Edsbruk, Sweden 2005

Doctoral Dissertation
Department of Economics
Stockholm University

ABSTRACT

This thesis consists of three papers in dynamic political economy.

Ideology and the Determination of Public Policy Over Time investigates how public policy responds to persistent ideological shocks in dynamic politico-economic equilibrium. To this end, we develop a tractable model to analyse the dynamic interactions among ideology, public policy and individuals' intertemporal choice. Analytical solutions are obtained to characterize the Markov perfect equilibrium. Our main finding is that the relationship between ideology and the size of government turns out to be non-monotonic. In particular, a right-leaning ideological wave may lead to higher taxation, which makes the size of government much less distinctive under different political regimes. Incorporating ideological uncertainty *per se* has its theoretical relevance. Sufficient ideological uncertainty helps pin down a unique equilibrium. This is in contrast with recent works on dynamic political economy which feature multiple equilibria and have no sharp empirical predictions.

Dynamic Inequality and Social Security analyses the dynamic politico-economic equilibrium of a model where the repeated voting on social security and the evolution of household characteristics are mutually affected over time. We incorporate within-cohort heterogeneity in a two-period Overlapping-Generation model to capture the intra-generational redistributive effect of social security transfers. Political decision-making is represented by probabilistic voting *a la* Lindbeck and Weibull (1987). We analytically characterize the unique Markov perfect equilibrium. The equilibrium social security tax rate is shown to be increasing in wealth inequality. The dynamic interaction between inequality and social security leads to growing social security programmes. The predictions of our model are broadly consistent with empirical evidence. We also perform some normative analysis, showing that the politico-economic mechanism tends to induce too large social security transfers in the long run.

A Markovian Social Contract of Social Security analyses the sustainability and evolution of the pay-as-you-go social security system in a majority voting framework with intra-cohort heterogeneity. We find that even under the temporal separation of social security contributions and benefits, there exists a Markovian social contract through which the self-interested middle-aged median voter has incentives to support the system for intra-generational redistributive reasons. This is in contrast with the approaches in the existing literature, which either resorts to the imperfect temporal separation of contributions and benefits, or builds the expectation of future social security benefits on variables that are payoff-irrelevant for future policymakers. Correspondingly, our model has a number of distinctive empirical implications. First, the social security tax rate converges along an increasing path to the steady state. Second, the growth of social security is negatively correlated with income inequality. Third, the impact of income inequality on the equilibrium social contract induces a non-monotonic relationship between income inequality and social security. These predictions are broadly consistent with the data from the OECD countries. Particularly, the Markovian social contract allows us to explain the insignificant or even negative relationship between inequality and government transfers that is hard to explain with the existing theory.

ACKNOWLEDGMENTS

I am grateful to many people making this work feasible and interesting. First of all, I wish to thank my advisor, Fabrizio Zilibotti. This work would not have been possible without his continuous support, encouragement and expert guidance. I thank Kaiji Chen, my coauthor, for his intelligence and patience, and my officemate, Giovanni Favara, for making my academic life more enjoyable. I am also indebted to Christina Lonnblad for editorial assistance and improving the language in my papers.

I thank the Institute for International Economic Studies at Stockholm University providing me the brilliant research environment to write the thesis. I also thank the Economics Department at University College London offering me the opportunity to start a very rewarding doctoral programme. Special thank is given to the Overseas Research Student Awards (ORS) in UK, the Graduate School Research Scholarship at UCL and the Jan Wallander Foundation for financial support. Finally, my warmest thanks go to Xiaoying, who has always brought me sunshine in times of frustration during this long academic journey.

TABLE OF CONTENTS

Chapter 1: Introduction 1

Chapter 2: Ideology and the Determination of Public Policy Over Time 5

Chapter 3: Dynamic Inequality and Social Security 51

Chapter 4: A Markovian Social Contract of Social Security 99

Chapter 1

Introduction

The thesis is part of a growing literature bringing dynamic politico-economic aspects into public policy analysis. The theme of the thesis is the notion that political decisions and private intertemporal choices are mutually affected over time. In particular, our work is based on a series of recent theoretical contributions including Hassler, Rodriguez Mora, Storesletten and Zilibotti (2003), Hassler, Storesletten and Zilibotti (2005) and Hassler, Krusell, Storesletten and Zilibotti (2005), which explicitly analyse the interaction among private intertemporal choices, repeated political decision-making and the evolution of political constituency. Their framework and methodology are applied to two topics: (1) the determination of public spending under a stochastic political environment with a persistent ideological shock; (2) the sustainability and the evolution of social security. We find that the dynamic politico-economic equilibrium approach may substantially change the pattern of public policy choices and thus, provide a new angle to understand the empirical evidence which is hard to explain with the existing theory. Analytical solutions are provided throughout the thesis, making the underlying politico-economic mechanism highly transparent.

Chapter 2 is motivated by the salient feature in real world democracies that voting behaviour is often driven by motives that seem hard to reconcile with mere economic factors. The empirical literature has long documented that ideology plays an important role in voting decisions, even when individuals' characteristics are accounted for. Ideological waves are also quite persistent and thus, may have nontrivial impacts on policy decision-making over time. In spite of these observations, none of the theoretical works has been devoted to analysing the role of ideology in a dynamic political economy. This chapter explores how ideology affects private intertemporal decision, public policy choice and the evolution of political constituency. To this end, we develop a tractable model, based on Hassler, Storesletten and Zilibotti (2005), to analyse the dynamic interactions among ideology, public policy and individuals' intertemporal choice. Analytical solutions are obtained to characterize the Markov perfect equilibrium. Our main finding is that the relationship between ideology and the size of government turns out to be non-monotonic. In particular, a right-leaning ideological wave may lead to high taxation, which makes the size of government much less distinctive under different political regimes. Incorporating ideological uncertainty *per se* has its theoretical relevance. Sufficient ideological uncertainty helps pin down a unique equilibrium. This is in constract with recent works on dynamic political economy which feature multiple equilibria and have no sharp empirical predictions.

Chapter 3 investigates the political decision of social security in a dynamic general equilibrium model. Most developed countries have large public pension programmes, involving not only inter-generational but also intra-generational transfers. The impact of an exogenous social security on household intertemporal choices and welfare has been extensively studied in the literature. The social security system, however, is not exogenous but endogenously determined by policy choices that reflect rich dynamic interactions between political and economic factors. Despite this, most of the existing literature has either assumed away politico-economic factors or, when considering them, it has focused on models where the size of social security is decided once-and-for-all. As a result, the effect of endogenous changes of household characteristics over time on the decision of social security transfers has been ignored altogether. In this chapter, we analyses the dynamic politico-economic equilibrium of a model where the repeated voting on social security and the evolution of household characteristics are mutually affected over time. We incorporate within-cohort heterogeneity in a two-period Overlapping-Generation model to capture the intra-generational redistributive effect of social security transfers. Political decision-making is represented by probabilistic voting *a la* Lindbeck and Weibull (1987). We analytically characterize the unique Markov perfect equilibrium. The equilibrium social security tax rate are shown to be increasing in wealth inequality. The dynamic interaction between inequality and social security leads to growing social security programmes. The predictions of our model are broadly consistent with empirical evidence. We also perform some normative analysis, showing that the politico-economic mechanism tends to induce too large social security transfers in the long run.

Chapter 4 analyses the sustainability and evolution of the social security system in a majority voting framework. The preceding chapter shows that probabilistic voting (by placing a weight on the interests of all individuals in society) is one way of explaining why a social security system can be sustained over time. However, this is arguably not the most convincing one. In the real world we know that many young people would be opposed to shutting down the system, even if this would imply for them an immediate reduction in the tax burden. The support for the pension system from the young generation may be explained by altruistic considerations vis-a-vis the current old. However, it could also occur by the logic of the social contract pointed out by Sjoblom (1985): self-interested agents perceive that future social security benefits are somehow related to the present contribution.[1] Thus, they are willing to pay the contributions in expectation that the system will not be terminated by the future generation. These considerations motivate us to study a model where the social security system is sustained over time in a median voter environment, and where the median voter is a net contributor rather than a recipient. We find that even under the temporal separation of social security contributions and benefits, there exists a Markovian social contract through which the self-interested middle-aged median voter has incentives to support the system for *intra-generational* redistributive reasons. This is in contrast with the approaches in the existing literature, which either resorts to the imperfect temporal separation of contributions and benefits, or builds the expectation of future social security benefits on variables that are payoff-irrelevant for future policymakers. Correspondingly, our model has a number of distinctive empirical implications. First, the social security tax rate converges along an increasing path to the steady state. Second, the growth of social security is negatively correlated with income inequality. Third, the impact

[1] To avoid the problem of temporal separation of contributions and benefits, some studies assume the welfare of retirees to be weighted into the preference of the policymaker, by resorting to altruism (Tabellini, 2000), probabilistic voting (Katuscak, 2002, Gonzalez-Eires and Niepelt, 2004, Song, 2005) or gerontocracy (Mulligan and Sala-i-Martin, 1999a).

of income inequality on the equilibrium social contract induces a non-monotonic relationship between income inequality and social security. These predictions are broadly consistent with the data from the OECD countries. Particularly, the Markovian social contract allows us to explain the insignificant or even negative relationship between inequality and government transfers that is hard to explain with the existing theory.

References

Hassler, John, Jose V. Rodriguez Mora, Kjetil Storesletten and Fabrizio Zilibotti, 2003, The Survival of the Welfare State, American Economic Review, 93, pp. 87-112.

Hassler, John, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti, 2005, The Dynamics of Government, forthcoming, Journal of Monetary Theory.

Hassler, John, Kjetil Storesletten and Fabrizio Zilibotti, 2005, Democratic Public Good Provision, forthcoming, Journal of Economic Theory.

Lindbeck, Assar and Jorgen W. Weibull, 1987, Balanced-Budget Redistribution as the Outcome of Political Competition, Public Choice, 52, 273-297.

Sjoblom, Kriss, 1985, Voting for Social Security, Public Choice, 45, 225-40.

Chapter 2

Ideology and the Determination of Public Policy Over Time *

1 Introduction

Modern political economy is designed to reveal the underlying mechanism of policy decision-making. The theoretical literature assumes that individuals or political groups vote for purely economical motives. A salient feature in real world democracies, however, is that voting behaviour is often driven by motives that seem hard to reconcile with mere economic factors. The empirical literature has long documented that ideology plays an important role in voting decisions, even when individuals' characteristics are accounted for.[1] For example, research has shown that about 15% to 20% electorates in the United States tend to base their policy preferences on ideology (e.g. Levitin and Miller, 1979). Ideological waves are also quite persistent. Pro-redistribution "leftist" policies, for instance, were highly popular in the 1960s, while there was a rightist mood in the late 1970s and 1980s.[2] Moreover, the influence of ideology is not necessarily limited to voting behaviour. Some recent theoretical work has shown that the evolution of political

* IIES, Stockholm University. Email: michael.song@iise.su.se. I am very grateful to Fabrizio Zilibotti for guidance and encouragement. I also thank Giovanni Favara and seminar participants at IIES for fruitful discussions.

[1] See, for example, Kau and Rubin (1979), Kalt and Zupan (1984, 1990). Similar results can also be found in the literature on political science. Levitin and Miller (1979) and Robinson and Fleishman (1984), among others, show that ideology turns out to be a significant predictor for individuals' voting choice in the U.S. presidential elections, even after taking into account the effects of party identification.

[2] See, for example, Robinson and Fleishman (1984) for a discussion of the U.S. survey data.

constituency can be endogenously driven by private intertemporal choice (e.g., Hassler, Rodriguez Mora, Storesletten and Zilibotti, 2003, Hassler, Storesletten and Zilibotti, 2004, HSRZ and HSZ henceforth). The persistency of ideological waves implies that the expectations of future policy outcomes be contingent on the current ideological state. Therefore, ideology may have additional effects on individuals' intertemporal behaviour via expectation. This, in turn, affects the current policy decision-making as well as the evolution of political constituency.

In spite of these observations, none of the theoretical works has been devoted to analysing the role of ideology in a dynamic political economy. This paper explores how ideology affects private intertemporal choice, public policy decision-making and the evolution of political constituency. We find that persistent ideological shocks substantially alter the pattern of public policy in previous research. Particularly, there is a non-monotonic U-shape correlation between ideology and the distortionary tax rate. An immediate implication is that a rightist ideology could actually increase the size of government. The mechanism behind this somewhat surprising result lies in the dynamic interaction between the ideology-contingent expectation and private intertemporal choice. When individuals condition their investment to the ideology-contingent expectation, ideology can affect the elasticity of the tax base and hence, the distortionary policy in a non-monotonic fashion. The non-monotonicity also indicates that the size of government under different political regimes tends to be much less distinctive than suggested in previous research. This may shed some lights on the weak empirical evidence for the standard partisan theory (e.g. Alseina, Roubini and Cohen, 1997).

Incorporating ideological uncertainty *per se* is theoretically relevant. There is a growing literature on the dynamics of government without commitment techniques (e.g., Besley and Coate, 1998, Hassler, Krusell, Storesletten and Zilibotti, 2005). This strand of research emphasizes the fact that in representative democracies, the incumbent government has limited abilities to commit to policies after the next election. One shortcoming, however, is that the literature often produces multiple equilibria when the identity of the incumbent is endogenous (e.g. HSRZ and HSZ). Hence, it cannot provide sharp empirical predictions. The multiplicity should not be surprising. When the government cannot commit to future policies, individuals must condition their choice on self-fulfilled expectations, which are usually not unique. The present paper shows that ideological uncertainty may help pin down a unique equilibrium, even if the government cannot commit. With sufficient uncertainty, all possible future policy outcomes can be associated with positive probability. This rules out the indeterminacy of belief in future political outcomes that can easily be self-fulfilled in the context without uncertainty. Consequently, our work can give precise empirical implications on the relationship between ideology and public policy. We may also explain the shift in beliefs that switches one political regime to the other and run comparative statics analyses, which are not fully secure under multiple equilibria.

The model economy, based on a tractable framework recently developed by HSRZ and HSZ, is populated by overlapping generations of individuals living two periods. Individuals make human capital investments when young, which can increase their probability of being rich during their life time. Two sets of policies are considered. In the benchmark case, we assume proportional income tax rates to be separately imposed on the young and old. The age-dependent taxation implies that the government has the ability to distinguish between elastic and inelastic tax bases. Then we analyse the model with age-independent taxes. Both setups give similar results. To incorporate ideology, the political decision process in our model differs from that of HSRZ and HSZ. We assume that there are two political parties running electoral competition. The right-wing and left-wing party, modeled as citizen-candidates (Osborne and Slivinski, 1996, Besley and Coate, 1997), represent the rich and the poor, respectively.[3] The election and public policy are codetermined by two fundamentals in the political economy, namely the distribution of individuals' economic situation and a persistent ideological shock.

The distinctive feature of our model is that the expectations on future political outcomes are ideology-contingent, as are private intertemporal choices. Therefore, ideology may not only directly alter the election, but also indirectly affect policy decision-making indirectly via human capital investment. For expositional reasons, we first study a static example with no private intertemporal trade-off,

[3] For simplicity, we assume zero entry cost, which shuts down the entry game in the standard citizen candidate model. However, we still regard the two-party system as a citizen-candidate model, since the party candidates cannot credibly commit to any policy platform other than their preferred policies, as in Osborne and Slivinski (1996) and Besley and Coate (1997).

in order to shut down the indirect effect of the ideology on public policy. The corresponding politico-economic equilibrium is straightforward. The left-wing party prefers a high tax rate, since the tax burden is unevenly distributed across the rich and the poor. In words, if ideology is sufficiently left-leaning, the left-wing party wins the election and then adopts higher taxes and public spending. The size of government is weakly but monotonically related to ideology. Thus, the implication of the direct effect of ideology coincides with the standard partisan model.

When ideology plays a role in intertemporal choices, the indirect effect of ideology appears and its implication turns out to be very different from the direct effect. Ideological movements lead to different expectations about future election and policy outcomes. Forward-looking individuals would thus adjust human capital investment accordingly; that is to say, the elasticity of the tax base is also ideology-contingent. To capture the indirect effect of ideology on public policy, we focus on Markov perfect equilibrium, where the dynamic interactions among ideology, private intertemporal choice and public policy can be characterized by two fixed-points: the ideology-contingent expectations on future political outcomes and the ideology-contingent public policy rule. Under quasi-linear preferences and uniformly distributed ideological shocks, the equilibrium can be solved analytically. The main finding is that there is a non-monotonic U-shape correlation between the distorting tax rate and the ideological state. Correspondingly, the size of government under different political regimes becomes much less distinctive than in the static example.

The non-monotonicity of tax rates with respect to ideology boils down to the non-monotonic elasticity of the tax base. When the ideological state is extremely left-leaning, the persistency of ideology implies that the left-wing party be reelected with certainty. Thus, the current distorting taxes do not affect the expectations on the future election outcome. This makes human capital investment rather inelastic. But when ideology becomes less left-leaning, taxes have an additional effect on human capital investment by affecting the future election and policy outcomes. This makes the tax base more elastic. Thus, the incumbent has an incentive to cut taxes. When ideology becomes more right-leaning, however, the lower expected future taxes tend to reduce the elasticity of human capital investment. This provides an incentive for the incumbent to raise the tax rate.

Our work is also related to the strategic debt literature, where policies are contingent on the reelection probability and could be inconsistent with the incumbent party's political colour.[4] The "cynical" policy in the strategic debt literature relies on the assumption that future policy outcomes can be affected by public debt. The present paper, however, suggests that the intertemporal policy instrument should not be necessary for the inconsistency between the incumbent's political colour and its policy choice. In fact, the inconsistency can naturally arise in a circumstance where future political outcomes are related to private intertemporal choice. It is also worth pointing out that the incumbent does not have any reelection concerns in our model, since political parties as citizen candidates are "short-sighted" in the sense that they only live one period. The much less distinctive difference on the size of government under different political regimes is purely driven by the non-monotonic ideology-contingent elasticity of the tax base, instead of the strategic behaviour of "long-lived" politicians in the strategic debt literature.

The rest of the paper is organized as follows. Section 2 describes the model and solves the static example. Section 3 analyses the benchmark model and gives the conditions for the existence and uniqueness of the Markov perfect equilibrium. In Section 4, we provide a closed-form solution when the ideological shock follows uniform distribution. Section 5 discusses the role of ideological uncertainty on the multiplicity of equilibria. Section 6 analyses the model with age-independent taxation and Section 7 concludes.

2 The Model

2.1 The Model Economy

The model economy is primarily based on a tractable framework recently developed by HSZ. The economy is inhabited by an infinite sequence of overlapping-generations. Each generation has a unit mass and lives two periods. There are

[4] There are a number of theoretical and empirical works on the strategic role of public debt, see for example Persson and Svensson (1989), Alesina and Tabellini (1990), Aghion and Bolton (1990), Milesi-Ferretti and Spolaore (1994), Jonsson (1997), Franzese (2001), Pettersson-Lidbom (2001) and Lambertini (2003).

two types of old individuals endowed with different productivity, referred to as the old rich and poor, respectively. The wage of the old rich equals $\delta \leq 1$ and the poor earn zero. The benefits from public good consumption g are identical across old individuals. The government imposes a proportional income tax rate τ^o on the old. Let u^{ou} and u^{os} be the utilities of the old poor and rich, respectively. These are equal to

$$u_t^{ou} = a^o g_t, \tag{2.1}$$

$$u_t^{os} = (1 - \tau_t^o)\,\delta + a^o g_t, \tag{2.2}$$

where $a^o \in (0, 2]$ is the constant marginal utility of public good for the old.

Young individuals are ex-ante homogenous. They make a human capital investment h, which increases the probability p of being rich in their life time. Without loss of generality, let $p = h$. The wage of the rich young equals unity and the poor earn zero.[5] τ^y is the proportional income tax rate for young individuals. Assuming a linear-quadratic preference over the consumption and cost of human capital investment, the utility of a young household is

$$u_t^y = h_t\,(1 - \tau_t^y) + a^y g_t - h_t^2 + \hat{\beta} E[u_{t+1}^o], \tag{2.3}$$

$$E\left[u_{t+1}^o\right] = h_t E\left[u_{t+1}^{os}\right] + (1 - h_t)\,E\left[u_{t+1}^{ou}\right], \tag{2.4}$$

where E is the expectation operator and $\hat{\beta} \in [0, 1]$ denotes the discounting rate. a^y is the constant marginal utility of public good for the young. The age-dependent taxation and marginal utility of public good have their realistic counterparts. Many public programmes and tax policies have important age-dependent elements. In addition, the young and old may evaluate public goods, such as public health care, in quite different ways. Allowing for age-dependent taxation also simplifies the analytical characterization without making any fundamental change to the results, as shall be seen below.

Through the wage structure, the old and young produce δh_{t-1} and h_t, respectively. Thus, the aggregate output y_t equals

$$y_t = \delta h_{t-1} + h_t. \tag{2.5}$$

[5] We assume that human capital depreciates over time, so that $1 - \delta$ can be regard as the depreciation rate.

Total tax incomes and public spending amount to $\tau_t^o \delta h_{t-1} + \tau_t^y h_t$ and $2g_t$, respectively. We assume that the government budget must be balanced in each period, which implies

$$g_t = \frac{\tau_t^o \delta h_{t-1} + \tau_t^y h_t}{2}. \tag{2.6}$$

2.2 The Political Decision Process

The sequence of tax rates is set through a repeated political decision process. We assume that only old individuals vote. This captures, in an extreme fashion, the phenomenon that the old are more influential in the determination of public policies.[6] It would be observationally equivalent to assume that voting occurs at the end of each period. Old individuals have no interests at stake and thus, abstain from voting. For expositional ease, we shall keep the former interpretation throughout the paper.

In HSZ, election follows the majority rule and the outcome is deterministic and solely depends on the distribution of old individuals' economic situation. However, political scientists have provided convincing evidence that, besides economic reasons, electorates' ideological label also plays a significant role in their voting choice. To incorporate ideology, we adopt a simple partisan framework with ideological uncertainty so that our model features a different political decision process from that of HSZ. The right-wing and left-wing party, modeled as citizen-candidates, represent the old rich and poor, respectively. The party candidates cannot credibly commit to any policy other than that preferred by the group they represent. For simplicity, we assume zero entry cost.[7] The election outcome is stochastic and codetermined by the distribution of old individuals' economic situation and a persistent ideological shock. The timing of events in each period is described as follows. Citizen candidates announce their policy platforms at the beginning of each period. Next, an ideological shock is realized

[6] For instance, Mulligan and Sala-i-Martin (1999) argue that the old have more influence in the political decision process because they have a lower cost of time. Empirically, the voting turnout is indeed lower for younger households (e.g. Wolfinger and Rosenstone, 1980). See HRSZ and HSZ for more detailed discussions.

[7] Zero entry cost shuts down the entry game in Osborne and Slivinski (1996) and Besley and Coate (1997). Consequently, both the candidate representing the rich and that representing the poor will participate in the electoral competition.

and then the voting occurs. The elected party implements her preferred tax rates and public spending. It follows that young individuals invest in human capital. Their being rich or poor is unfolded after the investment.

To model the impact of ideology on the election outcome, we assume that an ideological shock can switch a proportion of the poor (rich) to the right-wing (left-wing) side in terms of voting choice. The discrepancy between individuals' economic interests and their political preferences captures the impact of ideological movements. Consequently, there are two fundamentals in the political economy, i.e., the population of the old rich and the ideological state. Define the left-wingers (right-wingers) as old households voting for the left-wing (right-wing) party. The election outcome is determined by the proportion of right-wingers e_t:

$$e_t = \begin{cases} 1 & s_t \geq 1 - h_{t-1} \\ h_{t-1} + s_t & s_t \in (-h_{t-1}, 1 - h_{t-1}) \\ 0 & s_t \leq -h_{t-1} \end{cases}, \tag{2.7}$$

where s_t is the ideological shock at time t and h_{t-1} is the population of the old rich, or equivalently, the human capital investment at time $t-1$. A positive (negative) ideological shock s_t switches some of the poor (rich) to vote for the right-wing (left-wing) party. Thus, a high (low) s_t refers to a more right-leaning (left-leaning) political environment.[8] The election outcome at time t is codetermined by s_t and h_{t-1}. The right-wing party wins the election if $e_t > \frac{1}{2}$. Otherwise the left-wing is elected.[9] Note that (2.7) ensures that $e_t \in [0, 1]$ always holds. When s_t takes an extreme value (either very high or very low), the economic determinant h_{t-1} is wiped off. Outside these "ages of extremes" (Hobsbawm, 1996), economic motives may sway voters. Since ideological movements tend to be persistent (e.g. Robinson and Fleishman, 1984), s_t is assumed to follow a stationary AR(1) process, whose properties will be defined and discussed below.[10]

[8] Alternatively, s_t can be interpreted as a real shock. If the old poor are regarded as unemployed workers with zero earning, $h_{t-1} + s_t$ can be referred to as the employment rate, where $-s_t$ is a persistent unemployment shock.

[9] We assume that the left-wing comes into power if the proportion of left-wingers ties the right-wingers.

[10] The existence and uniqueness of the dynamic politico-economic equilibrium can easily be extended to an AR(n) process with $n > 1$.

2.3 Two Effects of Ideology on Public Policy

The right-wing party sets τ_t^o so as to maximize the utility of the rich u_t^{os} in (2.2), subject to the balanced-budget constraint (2.6). The assumption that $a^o \leq 2$ is sufficient for the right-wing to set $\tau_t^o = 0$.[11] The left-wing sets τ_t^o by maximizing u_t^{ou} in (2.1), which is equivalent to maximizing fiscal revenues $\tau_t^o \delta h_{t-1} + \tau_t^y h_t$. Since h_{t-1} is predetermined and τ_t^o does not distort young individuals' human capital investment, the left-wing will set $\tau_t^o = 1$. In other words, the left-wing government eliminates the income inequality of old individuals by imposing a 100% tax rate. To conclude, τ_t^o follows a binary rule

$$\tau_t^o = \begin{cases} 1 & \text{if } e_t \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}. \tag{2.8}$$

The disagreement on tax τ_t^o exhibits the feature of a two-party system. Despite the conflict of interest between left-wingers and right-wingers in terms of τ_t^o, their preferences on τ_t^y are perfectly aligned: attaining the top of the Laffer curve to maximize taxes from young individuals. This is because citizen candidates only represent the interests of the old rich and poor. None of them care about the welfare of the young.[12] So τ_t^y solves

$$\tau_t^y = \arg\max T_t, \tag{2.9}$$

where $T_t \equiv \tau_t^y h_t$. As we shall see in Section 6, the political parties would disagree on the tax rate imposed on the young if the government were not allowed to adopt age-dependent taxation. Two remarks are in order. First, the incumbent at time t would be better off if she could promise $\tau_{t+1}^o = 0$ to encourage human capital investment h_t. Without commitment techniques, however, the promise is not credible since future polices are repeatedly decided by the winners of future elections. So τ_{t+1}^o must follow the binary rule (2.8). Second, in the present political environment, both parties would like to see the right-wing being elected in the next period, since $\tau_{t+1}^o = 0$ in the right-wing regime encourages human capital investment and thus, enlarges the tax base. Note that the lack of reelection

[11] We assume that the right-wing would set $\tau_t^o = 0$ when the old rich are indifferent between private net earnings and public good, i.e., $a^o = 2$.

[12] The indifference between τ_t^y preferred by the leftist and rightist seems reasonable since the young are homogenous when candidates implement their policies.

concern is consistent with the "short-sighted" citizen candidates only living one period.

Now, we distinguish between two effects of ideology on public policy. Ideology can directly affect policy via the election. This is referred to as the direct effect of ideology, which is purely based on its impact on election outcomes. Besides the direct effect, the persistent ideological shock may also affect the private intertemporal investment via the ideology-contingent expectation on future political outcomes. This may change the elasticity of the tax base and thus, the policy decision-making. This is referred to as the indirect effect of ideology. As a warm-up exercise to facilitate the intuition, let us first study a static example with no private intertemporal trade-off. This also helps identify the direct effect of ideology on public policy.

2.3.1 A Static Example: the Direct Effect of Ideology

In this static example, we assume the probability of being rich in old age p to be exogenous. The corresponding politico-economic equilibrium is straightforward. The policy rule of τ_t^o follows (2.8). According to (2.3), young individuals' human capital investment solves

$$h = \arg\max_{\hat{h}} \left(1 - \tau^y\right) \hat{h} - \hat{h}^2, \tag{2.10}$$

which yields

$$h = \frac{1 - \tau^y}{2}. \tag{2.11}$$

We will continue to drop the time subscript when it does not create any confusion. (2.11) shows that private intertemporal choice is independent of ideology. Substituting (2.11) into (2.9), we obtain an equalized distorting tax rate and human capital investment under any ideological state.

$$\tau^y = \frac{1}{2}, \tag{2.12}$$

$$h = \frac{1}{4}. \tag{2.13}$$

(2.8) and (2.12) give the policy rules. Assuming away the intertemporal trade-off shuts down the link between ideology and current distorting policy τ_t^y. Since the political fundamental $e = p + s$, combining (2.8) and (2.12), the share of public spending can be computed as a percentage of aggregate output $\gamma \equiv 2g/y$:

$$\gamma = \begin{cases} \frac{1/2+\delta}{1+\delta} & \text{if } s \leq \frac{1}{2} - p \\ \frac{1/2}{1+\delta} & \text{otherwise} \end{cases}, \tag{2.14}$$

where γ measures the size of government. If the ideology shock $s \leq 1/2 - p$, the left-wing party, representing the interests of the poor, wins the election and adopts higher taxes and larger public spending for redistributive reasons. Otherwise, the right-wing wins and imposes lower taxes for the sake of the rich. Since the impact of ideology is limited to the election outcome, the difference in γ reflects the direct effect of ideology on public policy. Particularly, the direct effect leads to a weakly but monotonically negative correlation between ideology and the size of government. If $\delta = 0.5$, γ is three times higher in the left-wing than in the right-wing regime. The implication from the direct effect is thus in accordance with the standard prediction of partisan theory.

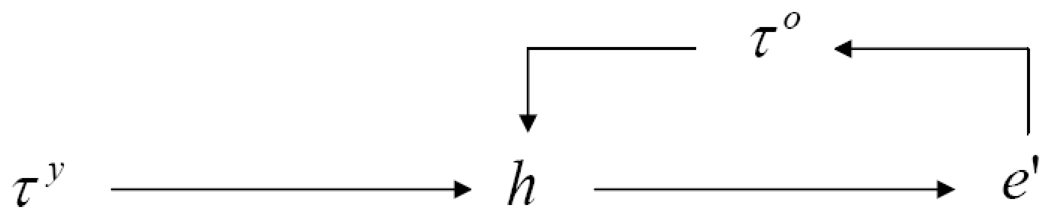

Figure 1: The Impact of Distorting Tax Rate τ^y on Human Capital Investment h.

3 The Politico-Economic Equilibrium

We turn to the benchmark setup, where the probability of being rich in old age is endogenously determined by human capital investment, i.e., $p = h$. Denote x and x' as the variable x in the current and following period, respectively. Now, the private intertemporal trade-off arises. The current distorting tax rate τ^y may change the future political fundamental e' via human capital investment

h. Moreover, changes in the future political fundamental may lead to different election outcomes and hence, different tax rates on old individuals, as indicated in (2.8). Forward-looking individuals will thus adjust human capital investment accordingly. The impact of τ^y on h is illustrated in Figure 1.

Taking into consideration the link between τ^y and e', individuals' decision problem for human capital investment turns out to be much more complicated than (2.10) in the static example. The expected utility u_t^y in (2.3) implies that h_t depends on $E[\tau_{t+1}^o]$. According to the binary tax rule (2.8), $E\tau_{t+1}^o$ is equal to $1-\pi_t$, where $\pi_t \equiv \Pr(e_{t+1} > 1/2)$ denotes the right-wing's probability of being elected at time $t+1$. Plugging (2.8) into (2.3), young individuals solve

$$h = \arg\max_{\hat{h}} \left(1-\tau^y+\beta\pi\right)\hat{h}-\hat{h}^2, \tag{2.15}$$

where $\beta \equiv \hat{\beta}\delta$. The utility from public good is irrelevant for the decision h, due to the atomic unit of individuals taking g and g' as given. (2.15) yields

$$h = \frac{1-\tau^y+\beta\pi}{2}. \tag{2.16}$$

In addition to the negative effect of τ^y on h in (2.11), (2.16) says that h increases in π, i.e. the probability for a right-wing government to be elected, since such a government will adopt the tax-free policy for old individuals. Moreover, (2.16) provides a way of pinning down the rational expectation of π. Substituting (2.16) into (2.7) and recalling that $e' = h + s'$, we obtain

$$\pi = \Pr\left(e' > \frac{1}{2}\right) = \Pr\left(\frac{1-\tau^y+\beta\pi}{2} + s' > \frac{1}{2}\right). \tag{2.17}$$

On the one hand, since π is determined by individuals' intertemporal decision, a high human capital investment h tends to increase π. On the other hand, given the expectation π, young individuals choose h according to (2.16). A high π leads to a high h. For the expectation to be self-fulfilled, π must satisfy (2.17). Hence, the rational expectation π is a fixed point of equation (2.17).

3.1 The Ideology-Contingent Expectation

First, we specify the properties of the stochastic process of s_t as follows. The density function of s_t is defined by $f : R^2 \to [0,\infty)$ with $\int f(s_t, s_{t-1})\, ds_t = 1$ for any given s_{t-1}. By (2.17), we know that π depends on τ^y and the probability of the ideological shock s', which is contingent on the current ideological state s. Hence, π can be written as a function of τ^y and s, $\pi : [0,1] \times R \to [0,1]$, which solves the following functional equation implied by (2.17):

$$\pi(\tau^y, s) = \int_{s' > \frac{\tau^y - \beta\pi(\tau^y, s)}{2}} f(s', s)\, ds'. \tag{2.18}$$

The existence of the ideology-contingent expectation $\pi(\tau^y, s)$ can easily be obtained by the following assumptions. Define $X \equiv [s_{\mathsf{min}}, s_{\mathsf{max}}]$, where $-\infty < s_{\mathsf{min}} < s_{\mathsf{max}} < \infty$. Assume

A1: s' and $s \in X$.

A2: $f(s', s)$ *is bounded and uniformly continuous.*

Lemma 1 *Assume A1 and A2. Then there exists a uniformly continuous function $\pi(\tau^y, s)$ that solves (2.18).*

Proof: See the appendix.

A2 is only a sufficient condition for the existence. $\pi(\tau^y, s)$ can exist under discontinuous distributions, as shall be seen in Section 4. The following assumption gives the sufficient condition for the uniqueness of $\pi(\tau^y, s)$.

A3: $f(s', s) < 2/\beta$ for all s' and $s \in X$.

Lemma 2 *Assume A1 and A3. Then there exists an unique $\pi(\tau^y, s)$ that solves (2.18).*

Proof: See the appendix.

Lemma 2 implies that sufficient ideological uncertainty can rule out the indeterminacy of beliefs, which features a number of recent research on dynamic politico-economic equilibrium with endogenous identity of the policymaker (e.g., HSRZ and HSZ). We will relax assumption A3 and study the multiplicity of equilibria in Section 5.

Plugging the ideology-contingent expectation $\pi(\tau^y, s)$ into (2.16), we obtain

$$h(\tau^y, s) = \frac{1 - \tau^y + \beta\pi(\tau^y, s)}{2}. \tag{2.19}$$

By (2.7), the future political fundamental e' evolves according to

$$e'\left(s',\tau^y,s\right)=\begin{cases}1 & \text{if } s'\geq 1-h\left(\tau^y,s\right)\\ h\left(\tau^y,s\right)+s' & \text{if } s'\in\left(-h\left(\tau^y,s\right),1-h\left(\tau^y,s\right)\right)\\ 0 & \text{if } s'\leq -h\left(\tau^y,s\right)\end{cases}. \tag{2.20}$$

3.2 Markov Perfect Equilibrium

Given the ideology-contingent expectation $\pi\left(\tau^y,s\right)$ solved from (2.18) and individuals' investment strategy (2.19), the incumbent sets τ^y according to (2.9). Let $T\left(\tau^y,s\right)\equiv h\left(\tau^y,s\right)\tau^y$. The problem is

$$\tau^y\left(s\right)=\arg\max_{\tau^y} T\left(\tau^y,s\right). \tag{2.21}$$

Three remarks are in order. First, the theorem of maximum implies that $\tau^y: R\rightarrow[0,1]$ be an upper hemi-continuous mapping. Second, τ^y only depends on the current ideological state s. One may guess that τ^y should also depend on the state variable $e=s+h_{-1}$, as τ^o in the Markovian tax rule (2.8). In fact, e or the identity of the incumbent has no influence on τ^y, since the objectives of two parties over τ^y are perfectly aligned: maximizing tax revenue from the young. Third, compared with the ideology-independent policy rule (2.12) in the static example, it can be found that $\tau^y\left(s\right)$ reflects the indirect effect of ideology on public policy, i.e., the impact of ideology on policy decision-making via private intertemporal choice. The indirect effect appears when individuals condition human capital investment on the ideology-contingent expectations. More specifically, ideological movements may indirectly affect policy decision-making via the ideology-contingent elasticity of tax base h:

$$\epsilon\left(\tau^y,s\right)=\frac{\tau^y-\beta\tau^y\partial\pi\left(\tau^y,s\right)/\partial\tau^y}{1-\tau^y+\beta\pi\left(\tau^y,s\right)}, \tag{2.22}$$

where ϵ denotes the absolute value of the elasticity of the tax base with respect to τ^y. An immediate observation is that, given $\partial\pi\left(\tau^y,s\right)/\partial\tau^y$, ϵ is decreasing in π. That is to say, the current tax base tends to be less elastic when ideology is more favourable for the right-wing to be elected in the next period.

In Markov perfect equilibrium, private and public choices are conditioned to payoff-relevant state variables. There are two state variables in our model: the ideological state s and the proportion of right-wingers $e=s+h_{-1}$. These two state variables are payoff-relevant since they determine the current election and thus, policy outcomes. So the Markovian political equilibrium can be defined as follows.

Definition 1 *A (Markov perfect) political equilibrium is a set of mappings* $\tau^o(e)$, $\tau^y(s)$, $\pi(\tau^y(s), s)$, *and* $h(\tau^y(s), s)$ *such that:*

(1) $\tau^o(e)$ *follows (2.8);*

(2) given $\tau^y(s)$, *the probability of election* $\pi(\tau^y(s), s)$ *solves (2.18);*

(3) given $\pi(\tau^y(s), s)$, *the human capital investment* $h(\tau^y(s), s)$ *follows (2.19);*

(4) given $h(\tau^y(s), s)$, *the incumbent solves* $\tau^y(s)$ *by (2.21).*

4 An Analytical Solution

In this section, we provide a closed-form solution of the Markov perfect equilibrium. The complete characterization of the equilibrium reveals the dynamic interactions among ideology, distortionary policy decision-making and individuals' intertemporal choice. Specifically, we assume that s' follows an AR(1) process with a symmetric uniformly distributed innovation[13]

$$s' = \rho s + \varepsilon'. \tag{2.23}$$

The ideological shock is stationary and persistent, i.e., $\rho \in (0, 1)$.[14] The density of ε equals $1/(2z)$ if $\varepsilon \in (-z, z)$ and 0 otherwise. So the conditional density function of s' is

$$f(s', s) = \begin{cases} \frac{1}{2z} & \text{if } s' \in (\rho s - z, \rho s + z) \\ 0 & \text{otherwise} \end{cases}. \tag{2.24}$$

Later we will adopt normal distribution to check the robustness of the analytical results under uniform distribution.

[13] Analytical solution is also available, though much more tedious, under more general setups. For example, s' follows an AR(n) process with the innovation that has a piecewise linear cumulative distribution function.

[14] If we interpret s as a shock on unemployment rate, ρ can be interpreted as the persistency of unemployment.

Now the functional equation (2.18) becomes

$$\pi\left(\tau^{y},s\right)=\begin{cases}1 & \text{if } \frac{\tau^{y}-\beta\pi(\tau^{y},s)}{2}\leq\rho s-z\\ \frac{1}{2z}\left(\rho s+z-\frac{\tau^{y}-\beta\pi(\tau^{y},s)}{2}\right) & \text{if } \frac{\tau^{y}-\beta\pi(\tau^{y},s)}{2}\in\left(\rho s-z,\rho s+z\right)\\ 0 & \text{if } \frac{\tau^{y}-\beta\pi(\tau^{y},s)}{2}\geq\rho s+z\end{cases}. \tag{2.25}$$

The linearity makes the analytical solution straightforward. Assumption A3 implies that $z>\beta/4$, which gives the sufficient condition for the uniqueness of $\pi\left(\tau^{y},s\right)$ under the uniform distribution (2.24). In this section, we assume that $z>\beta/4$. It can be shown that $z>\beta/4$ is also necessary. The opposite case $z<\beta/4$, which produces multiple equilibria, will be studied in Section 6.[15] Solving (2.25) yields:

$$\pi\left(\tau^{y},s\right)=\begin{cases}1 & \text{if } \tau^{y}\leq\lambda^{-}\left(s\right)\\ \frac{2(\rho s+z)-\tau^{y}}{4z-\beta} & \text{if } \tau^{y}\in\left(\lambda^{-}\left(s\right),\lambda^{+}\left(s\right)\right)\\ 0 & \text{if } \tau^{y}\geq\lambda^{+}\left(s\right)\end{cases}, \tag{2.26}$$

where $\lambda^{-}\left(s\right)\equiv2\left(\rho s-z\right)+\beta$ and $\lambda^{+}\left(s\right)\equiv2\left(\rho s+z\right)$. Note that $\lambda^{+}\left(s\right)>\lambda^{-}\left(s\right)$ as long as $z>\beta/4$. For notational convenience, we refer to $\lambda^{+}\left(s\right)\leq0$, or equivalently $s\leq-z/\rho$, as the left-dominating region, where the left-wing will be elected with probability one, irrespective of τ^{y}. Symmetrically, $\lambda^{-}\left(s\right)\geq1$, or equivalently $s\geq\left(\left(1-\beta\right)/2+z\right)/\rho$, is referred to as the right-dominating region, where the right-wing will be elected with probability one under any τ^{y}.

It immediate follows that $\partial\pi\left(\tau^{y},s\right)/\partial s\geq0$ by (2.26). A higher s leads to a higher expectation of s', which tends to increase the probability of the right-wing being elected in the next period. If $\partial\pi\left(\tau^{y},s\right)/\partial\tau^{y}\neq0$, the incumbent can choose τ^{y} "strategically" to affect the ideology-contingent expectation $\pi\left(\tau^{y},s\right)$. $\partial\pi\left(\tau^{y},s\right)/\partial\tau^{y}$ is a simple function of s:

$$\frac{\partial\pi\left(\tau^{y},s\right)}{\partial\tau^{y}}=\begin{cases}-\frac{1}{4z-\beta} & \text{if } \tau^{y}\in\left(\lambda^{-}\left(s\right),\lambda^{+}\left(s\right)\right)\\ 0 & \text{otherwise}\end{cases}. \tag{2.27}$$

Thus, $\partial\pi\left(\tau^{y},s\right)/\partial\tau^{y}\leq0$. Intuitively, a low τ^{y} encourages human capital investment and increases the share of rich individuals in the next period. This makes the right-wing more likely to win the next election. In the left-dominating

[15] $\pi\left(\tau^{y},s\right)$ does not exist if $z=\beta/4$. The non-existence of $\pi\left(\tau^{y},s\right)$ is due to the fact that the uniform distribution (2.24) is not continuous and thus, does not satisfy assumption A2.

(right-dominating) region with $\lambda^+(s) \leq 0$ ($\lambda^-(s) \geq 1$), however, τ^y cannot affect the policymaker's identity in the next period and hence, the expectation π is independent of τ^y.

The ideology-contingent expectation gives the ideology-contingent elasticity of tax base, which plays a central role in the indirect effect of ideology on public policy. Using (2.26) and (2.27), (2.22) yields

$$\epsilon(\tau^y, s) = \begin{cases} \frac{\tau^y}{1+\beta-\tau^y} & \text{if } \tau^y \leq \lambda^-(s) \\ \kappa(\tau^y, s) & \text{if } \tau^y \in \left(\lambda^-(s), \lambda^+(s)\right) \\ \frac{\tau^y}{1-\tau^y} & \text{if } \tau^y \geq \lambda^+(s) \end{cases}, \tag{2.28}$$

where $\kappa(\tau^y, s) \equiv (1 + \beta/(4z - \beta))\,\tau^y/(1 - \tau^y + \beta(2(\rho s + z) - \tau^y)/(4z - \beta))$. (2.28) illustrates how ideology affects the elasticity of tax base h. First, note that the tax base is more elastic in the left-dominating region ($\epsilon = \tau^y/(1-\tau^y)$) than in the right-dominating region ($\epsilon = \tau^y/(1+\beta-\tau^y)$). In the left-dominating (right-dominating) region, the expectation on $\tau^o = 1$ ($\tau^o = 0$) discourages (encourages) human capital investment and hence, makes the tax base more (less) elastic. However, the impact of ideology on ϵ is not monotonic. When s becomes less friendly to the left-wing such that $\lambda^+(s) \in (0, 1)$, it is easily shown that $\kappa(\tau^y, s)$ is strictly larger than $\tau^y/(1-\tau^y)$. The emergence of the negative effect of τ^y on the ideology-contingent expectation in (2.27), $\partial\pi(\tau^y, s)/\partial\tau^y$, enhances the distortion of τ^y on h. This makes h more elastic.

Now we are well-equipped to solve for $\tau^y(s)$. By (2.19) and (2.26), tax revenues from young individuals are

$$T(\tau^y, s) = \begin{cases} \frac{1}{2}(1 - \tau^y + \beta)\,\tau^y & \text{if } \tau^y \in \left[0, \lambda^-(s)\right] \\ \frac{1}{2}\left(1 - \tau^y + \beta\frac{2(\rho s+z)-\tau^y}{4z-\beta}\right)\tau^y & \text{if } \tau^y \in \left(\lambda^-(s), \lambda^+(s)\right) \\ \frac{1}{2}(1 - \tau^y)\,\tau^y & \text{if } \tau^y \in \left[\lambda^+(s), 1\right] \end{cases}. \tag{2.29}$$

Taking s as the state variable, τ^y can be pinned down by maximizing the piecewise quadratic function $T(\tau^y, s)$ in (2.29). A full characterization of the ideology-contingent policy rule $\tau^y(s)$ is given by

Proposition 1 *Assume that (2.24) $z \geq \left(\beta + \sqrt{\beta + \beta^2}\right)/8$. Then, the Markov* *perfect equilibrium is such that*

$$\tau^y(s) = \begin{cases} \frac{1+\beta}{2} & \text{if } s \geq s^3 \\ \lambda^-(s) & \text{if } s^2 < s < s^3 \\ \phi(s) & \text{if } s^1 \leq s \leq s^2 \\ \frac{1}{2} & \text{if } s \leq s^1 \end{cases}, \tag{2.30}$$

where $\phi(s) \equiv (2\beta(\rho s + z) + 4z - \beta)/8z$, $s^1 \equiv \left(\sqrt{z(4z-\beta)} - (4z-\beta)/2 - \beta z\right)/\beta\rho$, $s^2 \equiv (16z^2 - 6\beta z + 4z - \beta)/(\rho(16z - 2\beta))$ *and* $s^3 \equiv ((1-\beta)/4 + z)/\rho$, $s^1 < s^2 < s^3$.

Proof: See the appendix.

To simplify the statement in the paper, we assume that $z \geq \left(\beta + \sqrt{\beta + \beta^2}\right)/8$.[16] The two upper panels in Figure 2 plot $\tau^y(s)$ and the ideology-contingent expectation $\pi(\tau^y(s), s)$, where we set $\beta = 0.5$ ($\hat{\beta} = 1$ and $\delta = 0.5$), $\rho = 0.5$ and $z = 0.2$. In the left-dominating region, $\partial\pi(\tau^y, s)/\partial\tau^y = 0$, distortionary taxes have no effect on the election outcome in the next period. In words, the identity of the future policymaker is fixed in the left-dominating region. (2.29) reduces to a quadratic function $(1-\tau^y)\tau^y/2$ and the incumbent sets $\epsilon = 1$, which solves $\tau^y = 1/2$.

For $\lambda^+(s) > 0$ or equivalently $s > -z/\rho$, ideology becomes less hospitable to the left-wing and τ^y plays a role in the expectation on future political outcomes. Specifically, a low τ^y can increase the human capital investment h and thus, the probability for the rightist of being elected in the next period. In this case, the future policymaker's identity becomes endogenous and dependent on τ^y. Then, the corresponding objective function T is composed of two different quadratic functions, $T = [1 - \tau^y + \beta(2(\rho s + z) - \tau^y)/(4z - \beta)]\tau^y/2$ for low τ^y and $T = (1-\tau^y)\tau^y/2$ for high τ^y. The emergence of the effect of τ^y on the expectation π makes the tax base h more elastic, as shown by (2.28). This provides an incentive for the incumbent to cut the tax rate. However, if $\lambda^+(s)$ is close to zero, the incumbent needs to cut τ^y substantially to affect expectation π and the tax base h. Proposition 1 shows that for $s < s^1$, it is still optimal to set $\tau^y = 1/2$.

[16] The other case where $z \in \left(\beta/4, \left(\beta + \sqrt{\beta + \beta^2}\right)/8\right)$ is studied in the appendix, which gives similar results.

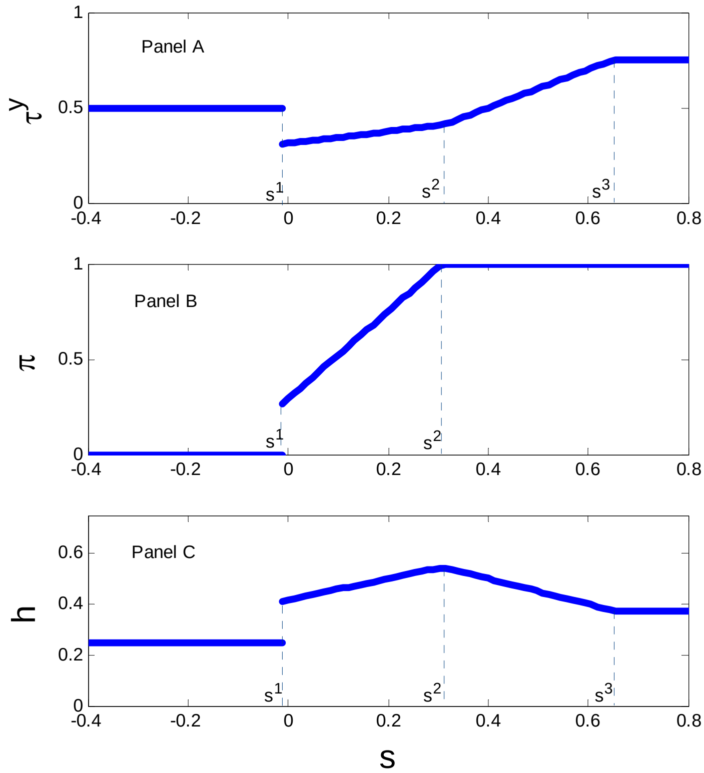

Figure 2: The Markov Perfect Equilibrium with Age-Dependent Taxation. The parameter values are $\beta = 0.5$, $\delta = 0.5$, $\rho = 0.5$, $z = 0.2$. Lower values of s correspond to a more left-wing state of ideology.

As s moves rightward, the future election outcome becomes more easily affected by tax-cutting. Particularly, when s reaches the threshold point s^1, we have an equalized maximum of the two quadratic functions in T. The incumbent becomes indifferent between $\tau^y = 1/2$ and $\tau^y = \phi(s^1)$, where

$$\phi\left(s^1\right) = \frac{\sqrt{4 - \beta/z}}{4}. \tag{2.31}$$

The indifference produces multiplicity and discontinuity of τ^y at s^1.[17] For a small increment ξ in s, the incumbent will cut τ^y from $1/2$ to $\phi\left(s^1 + \xi\right)$, to attain the top of the Laffer curve. This can be directly seen from the upper panel in Figure

[17] More specifically, $\tau^y(s)$ is not lower hemi-continuous. The theorem of maximum (e.g. Stokey and Lucas, 1989, pp. 62) only ensures that $\tau^y(s)$ is upper hemi-continuous.

2. Let $\Delta(\xi) \equiv 1/2 - \phi(s^1 + \xi)$ be the change in τ^y due to the increment ξ at s^1.

$$\lim_{\xi \to 0^+} \Delta(\xi) = \frac{1}{2} - \frac{\sqrt{4 - \beta/z}}{4} > 0. \tag{2.32}$$

(2.32) implies that the effect of τ^y on the expectation π induces a larger tax cut at s^1 under a higher β or a lower z. Intuitively, a higher β delivers a stronger feedback of future policy outcomes on private intertemporal choice. A lower z makes the expectation π more sensitive to τ^y. (2.27) and (2.28) show that both of these strengthen the effect of τ^y on π and thus, make the tax base h more elastic.

For $s > s^1$, τ^y turns out to be increasing in s. That is to say, the distortionary tax rate is higher in a more right-leaning political environment. The somewhat surprising result is due to the decreasing elasticity of h with respect to s, as shown by $\kappa(\tau^y, s)$ in (2.28). A high s increases the probability for a right-wing government of winning the next election, via its persistent impact on the future ideological state s'. This reduces the expected taxes and makes the current tax base less elastic. The lower elasticity provides an incentive for the incumbent to raise τ^y. In the right-dominating region, young individuals rationally expect the right-wing to win the next election under any τ^y. Consequently, (2.29) reduces to a quadratic function $(1 + \beta - \tau^y)\tau^y/2$, which solves $\tau^y = (1 + \beta)/2$.

From Figure 2, it can directly be seen that s has a greater impact on τ^y for $s \in (s^2, s^3)$ than for $s \in [s^1, s^2]$.[18] Note that under a modest s, τ^y has a negative effect on π, i.e. the probability for the right-wing of being elected in the next period.[19] When s is sufficiently right-leaning, however, reducing τ^y has no effect on π since π has reached its upper boundary. The lack of the effect of tax-cutting on π amplifies the impact of s on τ^y.

To conclude, the indirect effect of ideology on public policy, reflected by $\tau^y(s)$, turns out to be very different from the direct effect in the static example. Particularly, the indirect effect induces an inverted-U shaped taxation rule. This is in contrast to the conventional wisdom that the government tends to impose higher (lower) distortionary taxes in a more left-leaning (right-leaning) political environment. Our analysis shows that the non-monotonic indirect effect of ideology

[18] According to Proposition 1, $\mathrm{d}\tau^y(s)/\mathrm{d}s = \rho/4z$ for $s \in [s^1, s^2]$ and $\mathrm{d}\tau^y(s)/\mathrm{d}s = 2\rho$ for $s \in (s^2, s^3)$. The latter is greater than the former by $z > \beta/4$ and $\beta \leq 1$.

[19] For $s \in [s^1, s^2]$, $\partial\pi(\tau^y, s)/\partial\tau^y = -1/(4z - \beta) < 0$.

boils down to the non-monotonic ideology-contingent elasticity of the tax base.

Given the distortionary tax rule $\tau^y(s)$ and the ideology-contingent expectation $\pi(\tau^y(s), s)$, human capital investment h follows

Proposition 2 *Assume that (2.24) and* $z \geq \left(\beta + \sqrt{\beta + \beta^2}\right)/8$. *Then, the Markov perfect equilibrium is such that*

$$h = \begin{cases} \frac{1+\beta}{2} & \text{if } s \geq s^3 \\ \frac{1}{2} - (\rho s - z) & \text{if } s^2 < s < s^3 \\ \frac{1}{4} + \frac{\beta(\rho s + z)}{2(4z-\beta)} & \text{if } s^1 \leq s \leq s^2 \\ \frac{1}{4} & \text{if } s \leq s^1 \end{cases}. \tag{2.33}$$

The proof is straightforward and immediately follows from (2.19) and (2.30). h increases in s for $s \in [s^1, s^2]$ and peaks at $s = s^2$. Then, h decreases in s for $s \in (s^2, s^3)$. Panel C in Figure 2 plots the inverted U-shaped h. This casts doubts on the presumption that investment tends to be higher (lower) in the more right-leaning (left-leaning) political environment. The non-monotonicity of h is due to the fact that ideology has a greater effect on τ^y in (s^2, s^3). To be specific, a more right-leaning s has two opposite effects on h. First, it helps the right-wing win the next election and thus raises π, which has a positive impact on h. However, a high π makes h less elastic and thus, induces the incumbent to raise τ^y, which has a negative impact on h. For $s \in (s^1, s^2)$, the positive effect dominates the negative effect and h increases in s. For $s \in (s^2, s^3)$, however, the positive effect disappears since π has reached its upper boundary. Hence, the remaining negative effect produces a decreasing h.

Policy rules (2.8) and (2.30) capture the direct and indirect effects of ideology on public policy, respectively. Now these two effects are aggregated. Suppose that the ideological state in the previous period s_{-1} satisfies $h(\tau^y(s_{-1}), s_{-1}) + s < 1/2$. Then, the current incumbent is the left-wing. In spite of the monotonic direct effect of ideology, the inverted-U shape indirect effect can cause the size of government γ to be non-monotonically correlated with s in the left-wing regime, as can be seen from the left-hand panel of Figure 3, where $s_{-1} = -0.2$. The indirect effect is quantitatively important: it makes γ drop from 0.67 to 0.47 at $s = s^1$. The right-hand panel of Figure 3 plots γ under $s_{-1} = 0.8$. In this example, we can see that the maximum γ in the right-wing regime amounts to 0.48, greater than the minimum $\gamma = 0.47$ in the left-wing regime. That is to say, the size of government in the right-wing regime is not necessarily lower than that in the left-wing regime, even if the right-wing regime adopts the tax-free policy and its opponent imposes a 100% tax rate for old individuals.

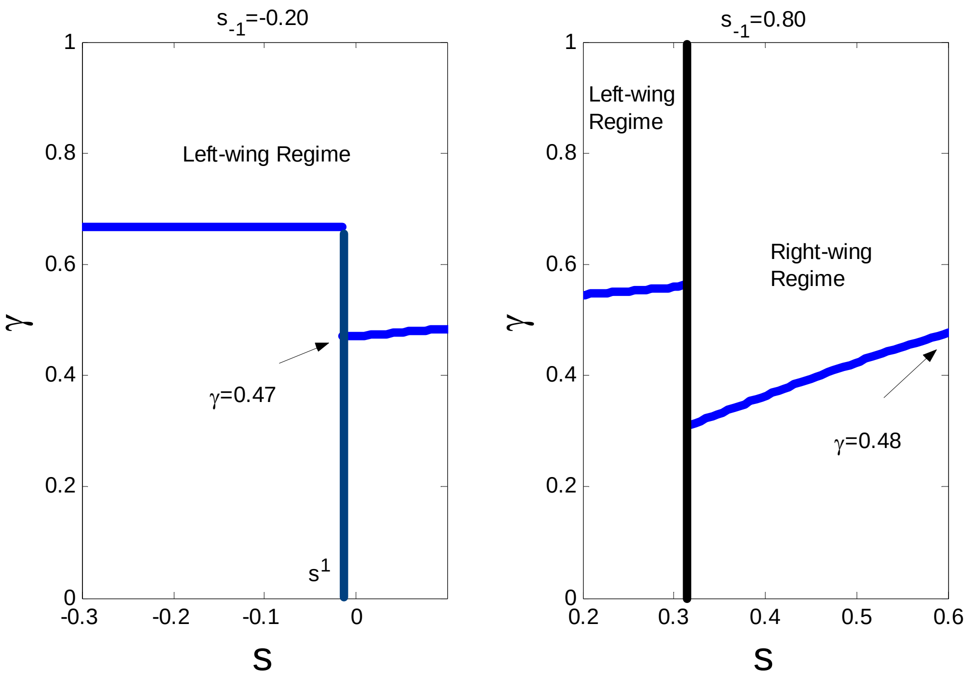

Figure 3: The Size of Government with Age-Dependent Taxation. The parameter values are $\beta = 0.5$, $\delta = 0.5$, $\rho = 0.5$, $z = 0.2$.

It is also worth noting that the size of government under different political regimes becomes substantially less distinctive due to the indirect effect of ideology. Recall that in (2.12) and (2.14), τ^y is a constant and γ is three times higher under the left-wing regime than under the right-wing regime. In that static case, human capital investment h does not increase individuals' probability of being rich. The comparison immediately indicates that the indirect effect of ideology via ideology-contingent expectation and private intertemporal choice stands in the heart of the converging sizes of government under different political regimes.

4.1 Robustness

We have analytically characterized the Markov perfect equilibrium under the uniform distribution assumption (2.24). It may be doubted whether the non-monotonicity, particularly the precipitation of τ^y at s^1, is caused by the discontinuity of $f(s', s)$. In this subsection, we adopt an alternative assumption that innovation ε is normally distributed with standard error σ, to check whether the non-monotonic relationship between τ^y and s still survives the different setup. Analytical solution is hard to obtain under normal distribution. Thus, we resort to the numerical method, which is specified in the appendix. Figure 4 plots τ^y, π and h with respect to s under the parameter values in Figure 2.[20]

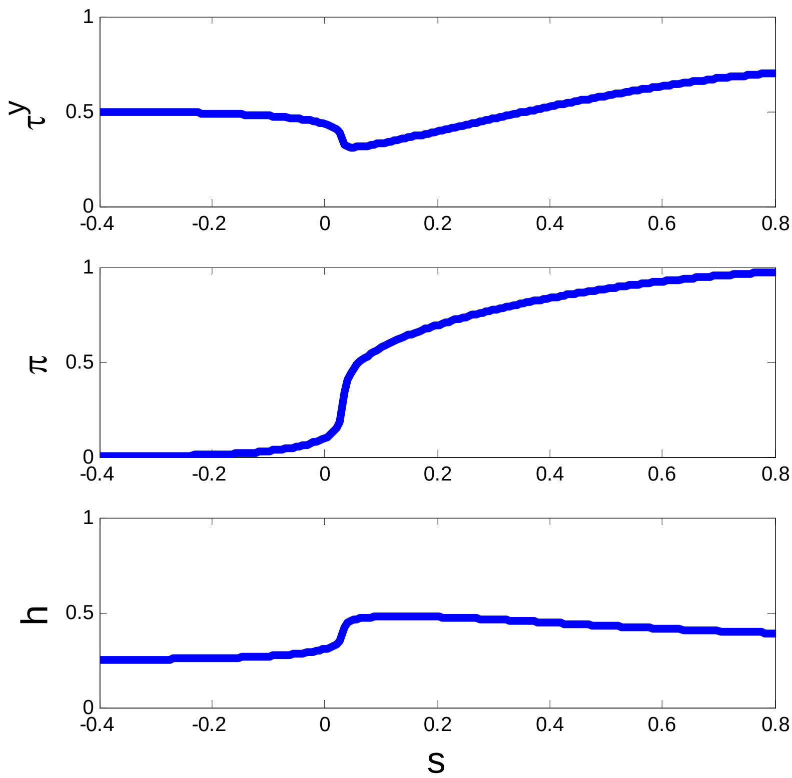

Figure 4: The Markov Perfect Equilibrium under Normal Distribution. The parameter values are $\beta = 0.5$, $\delta = 0.5$, $\rho = 0.5$, $\sigma = 0.116$.

The shapes of τ^y, π and h in Figure 4 are qualitatively the same as in Figure 2, though the discontinuity of τ^y disappears under continuous distribution. Two important properties are worth reemphasizing. First, τ^y is non-monotonically

[20] We set $\sigma = 2z/\sqrt{12}$ so that the standard error of ε under normal distribution equals that under uniform distribution.

related to s. Second, the corresponding h exhibits an inverted-U shape. The size of government γ is plotted in Figure 5. As in Figure 3, we set $s_{-1} = -0.20$ and 0.80 in the left- and right-hand panels, respectively.[21] Once more, γ under different political regimes becomes much less distinctive than that suggested by (2.14). In the left-hand panel, we can see that the left-wing incumbent may adopt a γ as low as 0.56, while the maximum of γ under $s_{-1} = 0.8$ amounts to 0.62 under the right-wing regime. Consistent with the findings in Figure 3, the size of government in the right-wing regime is not necessarily lower than that in the left-wing regime. Numerical results are suppressed under different β, δ, ρ and σ since they are robust to extensive choices of parameter values.

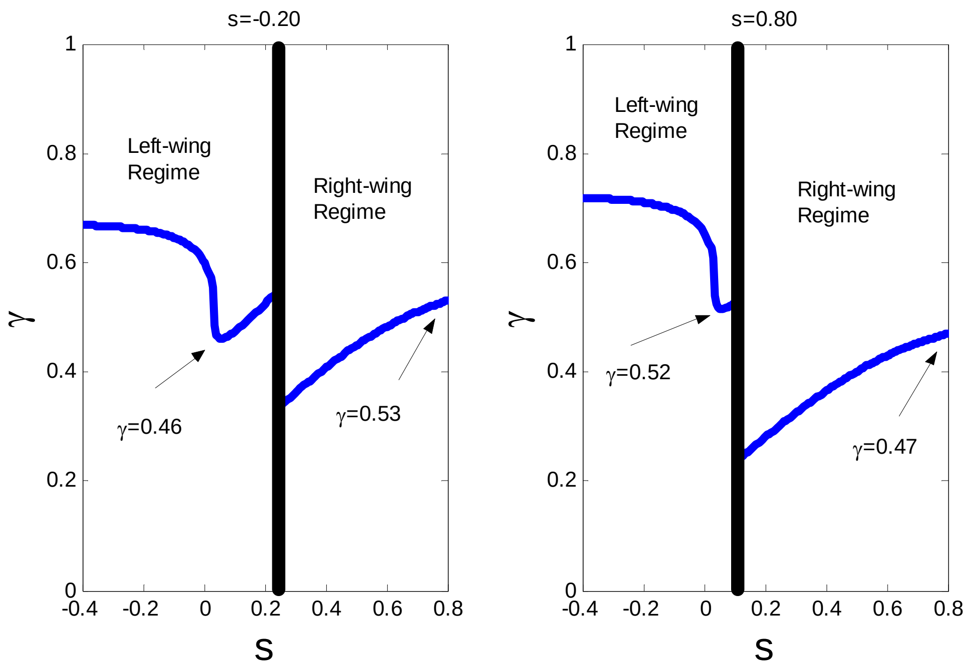

Figure 5: The Size of Government under Normal Distribution. The parameter values are $\beta = 0.5$, $\delta = 0.5$, $\rho = 0.5$, $\sigma = 0.116$.

[21] Note that the support of s is limited to $[\rho s_{-1} - z, \rho s_{-1} + z]$ under uniform distribution. There is no such limit under normal distribution.

5 Multiple Equilibria

In the previous sections, equilibria have been shown to be determinate. We regard this as an important progress, since much of the previous related literature (including HRSZ and HSZ) is plagued by multiple equilibria and indeterminacies, which undermine the ability of the theory to provide sharp empirical implications. The multiplicity should not be surprising. When the government cannot make a commitment on future policies, individuals must condition their choice on self-fulfilled expectations, which are usually not unique. Introducing ideology as a state variable has been proved decisive to eliminate multiplicity. In this section, we show that enough ideological uncertainty is indeed needed. Else, multiple Markov equilibria reemerge. We also show that the non-monotonic relationship between ideology and distortionary taxes notably survives the environment that features multiple equilibria.

If $z < \beta/4$, (2.26) no longer holds. Instead, (2.25) solves

$$\pi\left(\tau^y, s\right) = \begin{cases} 1 & \text{if } \tau^y \leq \lambda^-\left(s\right) \\ \frac{\tau^y - 2(\rho s + z)}{\beta - 4z} & \text{if } \tau^y \in \left(\lambda^+\left(s\right), \lambda^-\left(s\right)\right) \\ 0 & \text{if } \tau^y \geq \lambda^+\left(s\right) \end{cases}, \tag{2.34}$$

where $\lambda^+\left(s\right) < \lambda^-\left(s\right)$ under $z < \beta/4$. (2.34) implies that the ideology-contingent $\pi\left(\tau^y, s\right)$ is not unique for $\tau^y \in \left(\lambda^+\left(s\right), \lambda^-\left(s\right)\right)$. For $\tau^y \in \left(\lambda^+\left(s\right), \lambda^-\left(s\right)\right)$, the expectation $\left(\tau^y - 2\left(\rho s + z\right)\right) / \left(\beta - 4z\right)$ seems bizarre since it implies that $\partial\pi/\partial\tau^y > 0$ and $\partial\pi/\partial s < 0$. Both signs are counter-intuitive and hard to explain. However, the indeterminacy of $\pi\left(\tau^y, s\right)$ still remains even if we rule out the counter-intuitive self-fulfilled expectation $\left(\tau^y - 2\left(\rho s + z\right)\right) / \left(\beta - 4z\right)$ by requiring the expectation to be monotonic:

$$\pi\left(\tau^y, s\right) = \begin{cases} 1 & \text{if } \tau^y \leq \lambda^-\left(s\right) \\ 0 & \text{if } \tau^y \geq \lambda^+\left(s\right) \end{cases}. \tag{2.35}$$

Since $\lambda^+\left(s\right) < \lambda^-\left(s\right)$, π is indeterminate for $\tau^y \in \left[\lambda^+\left(s\right), \lambda^-\left(s\right)\right]$.

The indeterminacy of expectations opens the door to multiple Markov perfect equilibria. To see this, let us pick up a particular expectation rule satisfying (2.35)

$$\pi\left(\tau^y, s\right) = \begin{cases} 1 & \text{if } \tau^y < \lambda^+\left(s\right) + \psi \\ 0 & \text{if } \tau^y \geq \lambda^+\left(s\right) + \psi \end{cases},$$

where $\psi \in [0, \beta - 4z]$. Then, the human capital investment follows

$$h\left(\tau^y, s\right) = \begin{cases} \frac{1+\beta-\tau^y}{2} & \text{if } \tau^y < \lambda^+\left(s\right) + \psi \\ \frac{1-\tau^y}{2} & \text{if } \tau^y \geq \lambda^+\left(s\right) + \psi \end{cases}.$$

Correspondingly, the incumbent sets $\tau^y\left(s\right)$ by maximizing $T\left(\tau^y, s\right) = \tau^y h\left(\tau^y, s\right)$. Some algebra establishes

$$\tau^y\left(s\right) = \begin{cases} \frac{1+\beta}{2} & \text{if } s \geq \left(\left(1+\beta-2\psi\right)/4 - z\right)/\rho \\ \lambda^+\left(s\right) + \psi & \text{if } s \in \left[\left(\left(\eta-\psi\right)/2 - z\right)/\rho, \left(\left(1+\beta-2\psi\right)/4 - z\right)/\rho\right) \\ \frac{1}{2} & \text{if } s \leq \left(\left(\eta-\psi\right)/2 - z\right)/\rho \end{cases}, \tag{2.36}$$

where $\eta \equiv \left(1 + \beta - \sqrt{\beta\left(2+\beta\right)}\right)/2$. This leads to

Proposition 3 *Assume that (2.24) and $z < \beta/4$. There are multiple Markov perfect equilibria, where $\tau^y\left(s\right)$ follows (2.36) for any $\psi \in [0, \beta - 4z]$.*

Despite the indeterminacy of ψ, the non-monotonic relationship between s and τ^y still remains. (2.36) shows that the incumbent sets $\tau^y = 1/2$ when ideology is sufficiently left-leaning. When s moves rightward and reaches the threshold point $\theta \equiv \left(\left(\eta-\psi\right)/2 - z\right)/\rho$, the incumbent becomes indifferent between $\tau^y = 1/2$ and $\tau^y = \lambda^+\left(\theta\right) + \psi$. For a small increment ξ in s, the incumbent may cut τ^y from $1/2$ to $\lambda^+\left(\theta+\xi\right) + \psi$, to attain the top of the Laffer curve. τ^y increases in s for $s > \theta$, since the expectation $\pi = 1$ can be sustained by a higher τ^y in a more right-leaning political environment. Finally, τ^y reaches the maximum $\left(1+\beta\right)/2$ for $s \geq \left(\left(1+\beta-2\psi\right)/4 - z\right)/\rho$.

5.1 Discussion

The above analyses suggest that sufficient uncertainty helps pin down a unique belief on the future political fundamental. If $z < \beta/4$, $\tau^y \in \left(\lambda^+\left(s\right), \lambda^-\left(s\right)\right)$ may induce two self-fulfilled beliefs, $\pi = 0$ or 1. Given any of the self-fulfilled beliefs, the tax rate $\tau^y \in \left(\lambda^+\left(s\right), \lambda^-\left(s\right)\right)$ can be utilized to rule out the role of s on the future political fundamental. Specifically, given the belief that $\pi = 1$ (0) for $\tau^y \leq \lambda^-\left(s\right)$ ($\geq \lambda^+\left(s\right)$), any $\tau^y \in \left(\lambda^+\left(s\right), \lambda^-\left(s\right)\right)$ can lead to a rightist (leftist) government in the next period, irrespective of the current ideological state s. If $z > \beta/4$, given any $\tau^y \in \left(\lambda^-\left(s\right), \lambda^+\left(s\right)\right)$, neither the left-wing nor the right-wing can be elected with probability one for any $\tau^y \in \left(\lambda^-(s), \lambda^+(s)\right)$. Consequently, individuals must use the current ideological state s to figure out all possible future political fundamentals, $e' = h + s'$, and the corresponding policy outcomes. In words, s becomes useful information when uncertainty is sufficiently large. This highlights the role of information which dictates an unique belief on π.

Morris and Shin (1998, 2000) applied a similar methodology to pin down a unique belief in the financial market.[22] They assume that agents receive differential information on fundamentals. Noisy information destroys the common knowledge. An agent must thus consider all possible strategies of others based on its received information. The unique belief in our model does not rely on differential information. Alternatively, it has roots in the imperfect information about the future political fundamentals.

There are two advantages of introducing ideological uncertainty. First, the indeterminacy in the context without uncertainty has little to say about the shift in beliefs and the corresponding switches between different political regimes. Take HSZ as an example. If $\pi = 0$ (1), the left-wing (right-wing) will be the incumbent forever in all periods except the first election.[23] This is apparently inconsistent with the observed political cycles. Introducing electoral uncertainty provides a straightforward mechanism that switches beliefs and the corresponding politico-economic equilibrium outcomes over time. Second, sufficient ideological uncertainty gives the unique π and hence, the unique Markov perfect equilibrium. So, we may run comparative statics analyses, which are not fully secure under multiple equilibria.

6 Age-Independent Taxation

Throughout the paper, we assume that the government can condition taxes on age. Although age-dependent taxation has its realistic counterpart and substantially simplifies the analysis, this assumption is not innocuous. Since both parties are perfectly aligned with τ^y, the partisan effect only works on the non-

[22] The technique used in Herrendorf *et al.* (2000) is also related to ours. They find that sufficient heterogeneity can rule out the multiplicity of equilibria in a two-sector growth model.

[23] If the proportion of old rich exceeds $1/2$, the right-wing wins the first election. Otherwise, the left-wing wins.

distortionary tax rate τ^o. More crucially, one may wonder whether the binary taxation (2.8), which obviously overstates the effect of ideology on non-distortionary taxes, is essential for the non-monotonicity of the distortionary taxation rule $\tau^y(s)$. In this section, we would like to assess the robustness of our main findings under age-independent taxation, as in HRSZ. It will be seen that the weaker policy instrument does not lead to any qualitative change.

6.1 A Two-Period Model

For simplicity, we investigate a two-period version of HSRZ's model. Young individuals in the second period only live one period. It is worth emphasizing that under the assumption of age-dependent taxation, we can freely divide the benchmark infinite-period model into pieces of two-period models without changing any result. Hence, the results from the two-period model in this section are fully comparable to those in the preceding sections.

Let us solve the two-period model recursively. We refer to x and x' as variable x in the first and second period, respectively. The human capital investment in the second period is straightforward: $h' = (1-\tau')/2$. The left-wing incumbent maximizes tax revenues. So the age-independent tax rate τ' under the left-wing regime simply solves

$$\tau' = \underset{\hat{\tau}'}{\text{argmax}} \hat{\tau}' \left(\delta h + \frac{1-\hat{\tau}'}{2} \right). \tag{2.37}$$

τ' under the right-wing regime solves

$$\tau' = \underset{\hat{\tau}'}{\text{argmax}} \frac{a^o}{2} \hat{\tau}' \left(\delta h + \frac{1-\hat{\tau}'}{2} \right) + (1-\hat{\tau}')\,\delta. \tag{2.38}$$

The right-wing incumbent's objective is mixed, since she faces a trade-off between tax revenues and net earnings of the old rich. Two remarks are in order. First, since the incumbent is not allowed to impose separate tax rates on the elastic and inelastic tax bases h' and δh, the optimal uniform tax rate τ' must be somewhere between the age-dependent tax rates τ^y and τ^o. Second, under age-dependent taxation, the distortionary tax rate τ^y only depends on the ideological shock s, since the past human capital investment does not affect the current elastic tax base. Here, the uniform tax rate τ' becomes contingent on both state variables,

h and s'. The first order conditions of (2.37) and (2.38) yield

$$\tau'(h, s') = \begin{cases} \delta h + \frac{1}{2} & \text{if } h + s' \le \frac{1}{2} \\ \delta h + \frac{1}{2} - \frac{2\delta}{a^o} & \text{otherwise} \end{cases}. \tag{2.39}$$

For analytical convenience, we assume δ to be sufficiently small such that interior solutions always hold.[24] Due to the one-period decision of human capital investment, there is no indirect effect of ideology and policy rule (2.39) parallels the static example (2.14) in Section 2. The direct effect of ideology makes tax rate τ' weakly but monotonically decreasing in s'.

Now, we proceed to the first period. The human capital investment h solves

$$h = \frac{1 - \tau + \beta(1 - E[\tau'])}{2}. \tag{2.40}$$

Note that the probability of election π in (2.16) is replaced by $1 - E[\tau']$, where the expectation is defined as $E[\tau'] \equiv \int \tau'(h, s') f(s', s)\, ds'$. (2.40) suggests that h be determined by τ and s in equilibrium. Plugging (2.39) into (2.40) and following the similar procedures as in Section 3.1, we can establish a functional equation in terms of $h(\tau, s)$ after some algebra

$$h(\tau, s) = \frac{1 + \beta/2 - \tau + 2\beta\delta/a^o \int_{s' > \frac{1}{2} - h(\tau, s)} f(s', s)\, ds'}{2 + \beta\delta}. \tag{2.41}$$

The existence of $h(\tau, s)$ can easily be obtained under continuous distribution. The proof is analogous to Lemma 1 and left for the reader.

Lemma 3 *Assume that A1 and A2. Then there exists a uniformly continuous function $h(\tau, s)$ that solves (2.41).*

The following assumption gives the uniqueness. The proof is also analogous to Proposition 1 and left for the reader.

A4: $f(s', s) < a^o(1/2 + 1/\beta\delta)$ for all s' and $s \in X$.

Lemma 4 *Assume that A1 and A4. Then there exists a unique $h(\tau, s)$ that solves (2.41).*

---

[24] Interior solutions require $\delta h < 1/2$. Since the upper bound of h equals $\left(1 + \hat{\beta}\delta\right)/2$, we must assume $\delta < \left(\sqrt{1 + 4\hat{\beta}} - 1\right) / \left(2\hat{\beta}\right)$. Under $\hat{\beta} = 1$, $\delta < 0.6$ is sufficient for interior solutions.

Given individuals' investment strategy $h(\tau, s)$, the incumbent solves τ by

$$\tau(h_0, s) = \begin{cases} \underset{\hat{\tau}'}{\operatorname{argmax}} \hat{\tau}\left(\delta h_0 + h(\hat{\tau}, s)\right) & \text{if } h_0 + s \leq \frac{1}{2} \\ \underset{\hat{\tau}'}{\operatorname{argmax}} \frac{a^o}{2} \hat{\tau}\left(\delta h_0 + h(\hat{\tau}, s)\right) + (1 - \hat{\tau})\delta & \text{otherwise} \end{cases}. \tag{2.42}$$

As before, let us first study a static example with no private intertemporal trade-off to identify the direct effect of the ideology on public policy.

6.1.1 A Static Example: the Direct Effect of Ideology

Assume the probability of being rich in old age p to be exogenous. Young individuals' human capital investment in the first period follows $h = (1 - \tau)/2$, which is equivalent to the investment function in the second period. Therefore, the policy rule in the first period simply replicates that in the second period (2.39):

$$\tau(h_0, s) = \begin{cases} \delta h_0 + \frac{1}{2} & \text{if } h_0 + s \leq \frac{1}{2} \\ \delta h_0 + \frac{1}{2} - \frac{2\delta}{a^o} & \text{otherwise} \end{cases}. \tag{2.43}$$

Assuming away the intertemporal trade-off shuts down the link between ideology and the distorting policy τ. Under age-independent taxation, the share of public spending is $\gamma = \tau$. Hence, τ measures the size of government. If the ideology shock $s \leq 1/2 - h_0$, the left-wing party wins the election and runs a larger size of the government for redistributive reasons. Since the impact of ideology is limited to the election outcome, the difference in τ, which amounts to $2\delta/a^o$, reflects the direct effect of ideology. The direct effect leads to a weakly but monotonically negative correlation between ideology and the size of government. If $\delta = 0.5$ and $a^o = 2$, the tax rate in the left-wing regime would be 50% higher than that in the right-wing regime.

6.2 An Analytical Solution

The closed-form solution of $h(\tau, s)$ and $\tau(h_0, s)$ can be obtained under uniform distribution. Assuming (2.24), (2.41) becomes a linear functional equation. A4 implies that $z > \beta\delta/\left(2a^o\left(1 + \beta\delta/2\right)\right)$. By the linearity of (2.41), this is not only a sufficient but necessary condition for the uniqueness of $h(\tau, s)$. A full characterization of $h(\tau, s)$ and the corresponding policy rule $\tau(h_0, s)$ is given in the appendix. The main findings can be directly seen from Figure 6, where the upper, middle and lower panels plot $\tau(h_0, s)$, $\pi(\tau(h_0, s), s)$ and $h(\tau(h_0, s), s)$ with respect to s, respectively. Recall that under age-dependent taxation, both parties adopt the same taxation rule $\tau^y(s)$ for young individuals. With uniform taxation, however, the distortionary tax rate τ in the right-wing regime differs from that of the left-wing, since τ affects the net earnings of the old rich. Despite this generic difference, the shapes of the taxation rule and the human capital investment strategy in Figure 6 are qualitatively the same as those in Figure 2 and 3. The main common points are as follows. First, the incumbent would cut τ considerably when s is less left-wing, since the additional effect of τ on h via the future election and policy outcomes makes the tax base more elastic. Second, a more right-leaning ideological state s leads to a higher distortionary tax rate τ. Once more, this is because of the less elastic tax base due to the lower expected future taxes. Third, it can directly be seen from the figure that ideology has a greater impact on τ for $s \in (s^{2R}, s^{3R})$. In that region, the negative effect of τ on π disappears as π reaches its upper boundary. Consequently, the human capital investment h exhibits an inverted-U shape with respect to the ideological state. Finally, due to the non-monotonic $\tau(h_0, s)$, the size of government in the right-wing regime can be larger than that in the left-wing regime.[25]

[25] The political equilibrium is numerically solved under the assumption that innovation ε is continuously distributed. The numerical results are suppressed since they are qualitatively the same as the analytical results shown by Figure 6.

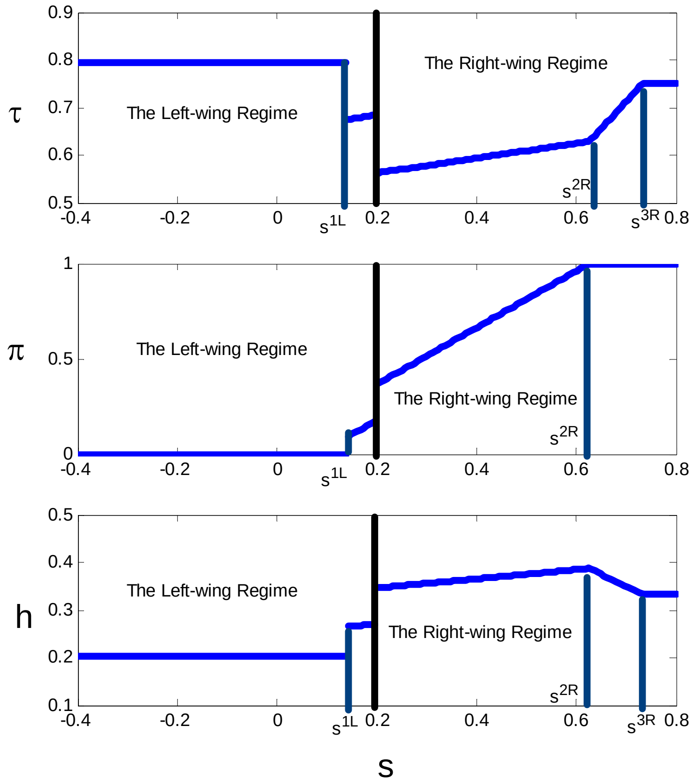

Figure 6: The Markov Perfect Equilibrium with Age-Independent Taxation. The parameter values are $\beta = 0.5$, $\delta = 0.5$, $\rho = 0.5$, $z = 0.2$ and $a^o = 2$. The initial $h_0 = 0.3$.

If $z < \beta\delta / \left(2a^o \left(1 + \beta\delta/2\right)\right)$, $h\left(\tau, s\right)$ turn out to be indeterminate. In the appendix, it is shown that there are multiple non-monotonic taxation rules $\tau\left(h_0, s\right)$ due to the indeterminacy of $h\left(\tau, s\right)$.

7 Conclusion

In spite of the growing literature on public policy decision-making in dynamic politico-economic equilibrium, most works are silent on the role of ideological waves, which tend to be persistent and have a significant impact on political outcomes. To explore the underlying mechanism of policy decision-making under stochastic ideological movements, we develop a tractable model to investigate the dynamic interactions among ideology, public policy and individuals' intertemporal choice. Our main finding is that the relationship between ideology and the size of government can be non-monotonic, which boils down to the non-monotonic ideology-contingent elasticity of the tax base. Consequently, the size of government turns out to be much less distinctive under different political regimes. Moreover, incorporating ideological uncertainty *per se* has its theoretical relevance. We prove that sufficient ideological uncertainty helps pin down a unique equilibrium. This is in contrasts to recent works on dynamic political economy which feature multiple equilibria and have no sharp empirical predictions.

The political party is modeled as citizen candidates, who only live one period of time. This approach is not innocuous, although it substantially simplifies the analysis. Particularly, it has two implications that seem questionable. First, the elected party's objective is to maximize the welfare of her constituency and thus has no interests in taking care of the young generation. Second, since citizen candidates are short-lived, they are generically "short-sighted" and have no reelection concerns. In future research, we plan to develop a model which can incorporate the political power of the young as well as the incumbent's reelection concerns.

References

Aghion, Philippe and Patrick Bolton, 1990, Government Domestic Debt and the Risk of a Default: A Political-Economic Model of a Strategic Role of Debt, 315-345, in R. Dornbusch and M. Draghi, eds., Public Debt Management: Theory and History, Cambridge: Cambridge University Press.

Alesina, Alberto and Guido Tabellini, 1990, A Positive Theory of Fiscal Deficits and Government Debt, Review of Economic Studies, 57, 403-414.

Besley, Timothy and Stephen Coate, 1997, An Economic Model of Representative Democracy, Quarterly Journal of Economics, 112, 85-114.

Besley, Timothy and Stephen Coate, 1998, Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis, American Economic Review, 88, 139-156.

Franzese, Robert J. Jr., 2001, The Positive Political Economy of Public Debt: An Empirical Examination of the OECD Postwar Experience, mimeo.

Hassler, John, Jose V. Rodriguez Mora, Kjetil Storesletten and Fabrizio Zilibotti, 2003, The Survival of the Welfare State, American Economic Review, 93, pp. 87-112.

Hassler, John, Kjetil Storesletten and Fabrizio Zilibotti, 2004, Democratic Public Good Provision, forthcoming, Journal of Economic Theory.

Hassler, John, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti, 2005, The Dynamics of Government, forthcoming, Journal of Monetary Theory.

Herrendorf, Berthold, Akos Valentinyi and Robert Waldmann, 2000, Ruling out Multiplicity and Indeterminacy: The Role of Heterogeneity, Review of Economic Studies, 67, 295-307.

Hobsbawm, Eric, 1996, The Age of Extremes: A History of the World, 1914-1991, Vintage.

Hibbs, Douglas A. Jr., 1977, Political Parties and Macroeconomic Policy, American Political Science Review, 71, 1467-1487.

Jonsson, Gunnar, 1997, Monetary Politics and Unemployment Persistence, Journal of Monetary Economics, 39, 303-325.

Kalt, Joseph H. and Mark A. Zupan, 1984, Capture and Ideology in Economic Theory of Politics, American Economic Review, 74, 279-300.

Kalt, Joseph H. and Mark A. Zupan, 1990, The Apparent Ideological Behavior of Legislators: Testing for Principal-Agent Slack in Political Institutions, Journal of Law and Economics, 33, 103-132.

Kau, James B. and Paul H. Rubin, 1979, Self-Interest, Ideology and Logrolling in Congressional Voting, Journal of Law and Economics, 22, 365-384.

Lambertini, Luisa, 2003, Are Budget Deficits Used Strategically, mimeo, UCLA.

Levitin, Teresa E. and Warren E. Miller, 1979, Ideological Interpretations of Presidential Elections, American Political Science Review, 73, 751-771.

Osborne, Martin J. and Al Slivinski, 1996, A Model of Political Competition with Citizen-Candidates, Quarterly Journal of Economics, 111, 65-96.

Pettersson-Lidbom, Per, 2001, An Empirical Investigation of the Strategic Use of Debt, Journal of Political Economy, 109, 570-583.

Milesi-Ferretti, Gian Maria and Enrico Spolaore, 1994, How Cynical Can an Incumbent Be? Strategic Policy in a Model of Government Spending, Journal of Public Economics, 55, 121-140.

Morris, Stephen and Hyun Song Shin, 1998, Unique Equilibrium in a Model of Self-fulfilling Currency Attacks, American Economic Review, 88, 587-597.

Morris, Stephen and Hyun Song Shin, 2000, Rethinking Multiple Equilibria in Macroeconomic Modelling, NBER Macroeconomics Annual, 139-161, MIT Press.

Mulligan, Casey B. and Xavier Sala-i-Martin, 1999, Gerontocracy, Retirement, and Social Security, NBER Working Paper, No. 7117.

Persson, Torsten and Lars E. O. Svensson, 1989, Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preference, Quarterly Journal of Economics, 104, 325-345.

Persson, Torsten and Guido Tabellini, 2000, Political Economics: Explaining Economic Policy, Cambridge: The MIT Press.

Robinson, John P. and John A. Fleishman, 1984, Ideological Trends in American Public Opinion, Annals of the American Academy of Political and Social Science, 742, 50-60.

Song, Zheng, 2005, Ruling Out the Multiplicity of Dynamic Politico-Economic Equilibria, mimeo, IIES.

Stokey, Nancy L. and Robert E. Lucas, Jr., 1989, Recursive Methods in Economic Dynamics, Harvard University Press.

Wolfinger, Raymond E. and Steven J. Rosenstone, 1980, Who votes?, Yale University Press, 1980.

8 Appendix

8.1 Proof of Lemma 1

Apply the Schauder fixed point theorem. Let C be a set of bounded and uniformly continuous functions mapping from $[0,1] \times X$ to $[0,1]$. Define $F = \int_{s' \geq \frac{\tau^y - \beta\alpha(\tau^y, S)}{2}} f(s', s)\, ds'$, where $\alpha(\tau^y, s) \in C$. We need to prove that the mapping F has a fixed point.

Let $\Omega = \{F(\alpha), \alpha \in C\}$. We first claim that Ω is equicontinuous, i.e., $F(\alpha)$ is bounded and uniformly continuous for any $\alpha \in C$. The boundedness is trivial since $F(\alpha)(\tau^y, s) \leq \int f(s', s)\, ds' = 1$. To prove that $F(\alpha)$ is uniformly continuous, we pick up any two vectors $x = (\tau_1^y, s_1)$ and $y = (\tau_2^y, s_2)$ from $[0,1] \times X$. It is straightforward to show that

$$
\begin{aligned}
& \left|F(\alpha)(\tau_1^y, s_1) - F(\alpha)(\tau_2^y, s_2)\right| \\
= & \left| \int_{s' \geq \frac{\tau_1^y - \beta\alpha(\tau_1^y, s_1)}{2}} f(s', s_1)\, ds' - \int_{s' \geq \frac{\tau_2^y - \beta\alpha(\tau_2^y, s_2)}{2}} f(s', s_2)\, ds' \right| \\
\leq & \left| \int_{s' \geq \frac{\tau_1^y - \beta\alpha(\tau_1^y, s_1)}{2}} f(s', s_1)\, ds' - \int_{s' \geq \frac{\tau_2^y - \beta\alpha(\tau_2^y, s_2)}{2}} f(s', s_1)\, ds' \right| \\
& + \left| \int_{s' \geq \frac{\tau_2^y - \beta\alpha(\tau_2^y, s_2)}{2}} f(s', s_1)\, ds' - \int_{s' \geq \frac{\tau_2^y - \beta\alpha(\tau_2^y, S_2)}{2}} f(s', s_2)\, ds' \right| \\
\leq & \left| \frac{\tau_1^y - \tau_2^y - \beta\left(\alpha(\tau_1^y, s_1) - \alpha(\tau_2^y, s_2)\right)}{2} \right| \left\| f(s', s_1) \right\|_{\mathsf{sup}} \\
& + \left| \sup s' - \frac{\tau_2^y - \beta\alpha(\tau_2^y, s_2)}{2} \right| \left\| f(s', s_1) - f(s', s_2) \right\|_{\mathsf{sup}}.
\end{aligned}
$$

As $\|x - y\| \to 0$, we have $\left|\left(\tau_1^y - \tau_2^y - \beta\left(\alpha(\tau_1^y, s_1) - \alpha(\tau_2^y, s_2)\right)\right)/2\right| \left\|f(s', s_1)\right\|_{\mathsf{sup}} \to 0$ by the uniform continuity of α and $\|f(s', s)\|_{\mathsf{sup}} < \infty$ (A2). Moreover, from A1 and the boundedness of α, it immediately follows that $\left|\sup s' - \frac{\tau_2^y - \beta\alpha(\tau_2^y, s_2)}{2}\right| < \infty$. By A2, $\|f(s', s_1) - f(s', s_2)\|_{\mathsf{sup}} \to 0$ as $\|x - y\| \to 0$. Therefore, we have $|F(\alpha)(x) - F(\alpha)(y)| \to 0$ as $\|x - y\| \to 0$.

Next we check the conditions of the Schauder fixed point theorem (Theorem 17.4, Stokey and Lucas, 1989). Ω has been to proved equicontinuous. And it is easily shown that C is nonempty, closed and convex and F is continuous. Thus, all conditions are satisfied. $\square$

8.2 Proof of Lemma 2

We only need to prove that, given any (τ^y, s), the following equation has a unique solution

$$x = F(x) \equiv \int_{s' \geq \frac{\tau^y - \beta x}{2}} f(s', s)\, ds'. \tag{2.44}$$

A3 implies that $\mathrm{d}F(x)/\mathrm{d}x = \beta f(s', s)/2 < 1$. The proof is complete by applying the contraction mapping theorem. $\square$

8.3 Proof of Proposition 1

For notational convenience, we define

$$\begin{aligned}
T^A(\tau^y, s) &\equiv \frac{1}{2}(1 - \tau^y + \beta)\tau^y, \\
T^B(\tau^y, s) &\equiv \frac{1}{2}\left[1 - \tau^y + \beta\frac{2(\rho s + z) - \tau^y}{4z - \beta}\right]\tau^y, \\
T^C(\tau^y, s) &\equiv \frac{1}{2}(1 - \tau^y)\tau^y.
\end{aligned}$$

$\hat{\tau}^A(s) = (1+\beta)/2$, $\hat{\tau}^B(s) = \phi(s)$ and $\hat{\tau}^C(s) = 1/2$ are the interior solutions of $\max_{\tau^y} T^A(\tau^y, s)$, $\max_{\tau^y} T^B(\tau^y, s)$ and $\max_{\tau^y} T^C(\tau^y, s)$, respectively. Moreover, let

$$\begin{aligned}
A(s) &= \max_{\tau^y \in [0, \lambda^-(s)]} T^A(\tau^y, s) & \tau^A(s) &= \operatorname*{argmax}_{\tau^y \in [0, \lambda^-(s)]} T^A(\tau^y, s) \\
B(s) &= \max_{\tau^y \in (\lambda^-(s), \lambda^+(s))} T^B(\tau^y, s) & \tau^B(s) &= \operatorname*{argmax}_{\tau^y \in (\lambda^-(s), \lambda^+(s))} T^B(\tau^y, s) \\
C(s) &= \max_{\tau^y \in (\lambda^+(s), 1]} T^C(\tau^y, s) & \tau^C(s) &= \operatorname*{argmax}_{\tau^y \in (\lambda^+(s), 1]} T^C(\tau^y, s)
\end{aligned}$$

The proof is based on the following lemma.

Lemma 5

$$\tau^y(s) = \begin{cases}
\frac{1}{2} & \textit{if } s \leq s^1 \\
\phi(s) & \textit{if } s \geq s^1,\ s < s^4,\ s < s^2 \\
\frac{1}{2} & \textit{if } s \geq s^1,\ s < s^4,\ s \geq s^2,\ s \leq s^5 \\
\lambda^-(s) & \textit{if } s \geq s^1,\ s < s^4,\ s \geq s^2,\ s \geq s^5 \\
\phi(s) & \textit{if } s \geq s^1,\ s \geq s^4,\ s \leq s^2 \\
\lambda^-(s) & \textit{if } s \geq s^1,\ s \geq s^4,\ s \in (s^2, s^3) \\
\frac{1+\beta}{2} & \textit{if } s \geq s^1,\ s \geq s^4,\ s \geq s^3
\end{cases}$$

where $s^4 \equiv (1/4 - z)/\rho$ *and* $s^5 \equiv \left(1 - \beta - \sqrt{\beta^2 + 2\beta} + 4z\right)/(4\rho)$.

When $z \geq \left(\beta + \sqrt{\beta + \beta^2}\right)/8$, it can easily be shown that $s^2 \geq s^4$. Thus, $s < s^4$ and $s \geq s^2$ cannot hold simultaneously. Moreover, since $s^4 \geq s^1$, we have $s^2 \geq s^1$. Lemma 1 leads to Proposition 1, hence. $\square$

8.3.1 Proof of Lemma 5

The solution of maximizing (2.29) is straightforward under two polarized cases, i.e., $s \geq ((1-\beta)/2 + z)/\rho$ and $s \leq -z/\rho$. Thus, we need only to focus on $s \in (-z/\rho, ((1-\beta)/2 + z)/\rho)$.

$\tau^C(s)$ equals $\hat{\tau}^C(s)$ if $\lambda^+(s) \leq 1/2$. Since $\hat{\tau}^A(s) > \hat{\tau}^B(s) > \hat{\tau}^C(s)$ for $\lambda^+(s) \leq 1/2$, It can easily be shown that $C(s) \geq B(s) \geq A(s)$ and $\tau^y(s) = \hat{\tau}^C(s)$ if $s \leq s^1$, where s^1 solves

$$T^B\left(\hat{\tau}^B\left(s^1\right), s^1\right) = T^C\left(\hat{\tau}^C\left(s^1\right), s^1\right)$$

This yields

$$s^1 = \frac{\sqrt{z(4z-\beta)} - (4z-\beta)/2 - \beta z}{\beta\rho}.$$

The other root is omitted since $s^1 > -z/\rho$.

For $s \geq s^1$, we distinguish two cases where $s \geq s^4$ and $s < s^4$, respectively. First look at $s \geq s^4$. We claim that $\tau^y(s) = \hat{\tau}^B(s)$ for $s \leq s^2$. Note that there exists a unique $s^2 = (16z^2 - 6\beta z + 4z - \beta)/(\rho(16z - 2\beta))$ such that $\hat{\tau}^B(s) \gtreqless \lambda^-(s)$ for $s \lesseqgtr s^2$. This implies that $\tau^B(s) = \hat{\tau}^B(s)$ for $s \leq s^2$. Since $\hat{\tau}^A(s) > \hat{\tau}^B(s)$,[26] we have $B(s) \geq A(s)$ and $\tau^y(s) = \hat{\tau}^B(s)$ for $s \leq s^2$. Next consider $s > s^2$. Obviously, there is a unique $s^3 = ((1-\beta)/4 + z)/\rho$ such that $\hat{\tau}^A(s) \lesseqgtr \lambda^-(s)$ for $s \gtreqless s^3$. Together with the fact that $\hat{\tau}^A(s) > \hat{\tau}^B(s)$, this implies that $\tau^y(s) = \hat{\tau}^A(s)$ for $s \geq s^3$. Moreover, since $\hat{\tau}^A(s) > \hat{\tau}^B(s)$, the intersection of $\hat{\tau}^A(s)$ and $\lambda^-(s)$ must be higher than that of $\hat{\tau}^B(s)$ and $\lambda^-(s)$. So $s^3 > s^2$. For $s \in (s^2, s^3)$, $\tau^A(s) \neq \hat{\tau}^A(s)$ and $\tau^B(s) \neq \hat{\tau}^B(s)$. Hence, $\tau^y(s)$ simply equals the boundary $\lambda^-(s)$ for $s \in (s^2, s^3)$.

Now, we proceed to the case where $s < s^4$. Two cases are investigated in order. For $s < s^2$, the above analysis establishes $\tau^y(s) = \hat{\tau}^B(s)$. For $s \geq s^2$,

[26] Since $s < ((1-\beta)/2 + z)/\rho$, we must have $s < (1/2 + z)/\rho$, the condition which satisfies $\hat{\tau}^A(s) > \hat{\tau}^B(s)$.

since $A(s) \geq B(s)$, we only need to compare $A(s)$ and $C(s)$. It immediately follows that $C(s) \geq A(s)$ and $\tau^y(s) = \hat{\tau}^C(s)$ if $s \leq s^5$ and $C(s) \leq A(s)$ and $\tau^y(s) = \lambda^-(s)$ if $s \geq s^5$, where s^5 solves

$$T^A\left(\lambda^-\left(s^5\right), s^5\right) = T^C\left(\hat{\tau}^C\left(s^5\right), s^5\right)$$

This yields

$$s^5 = \frac{1 - \beta - \sqrt{\beta^2 + 2\beta} + 4z}{4\rho}.$$

□

8.4 Normal Distribution

Assume that ε follows normal distribution. The conditional density function becomes

$$f(s', s) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(s' - \rho s)^2}{2\sigma^2}\right], \tag{2.45}$$

where σ is the standard error of innovation. If $\beta \leq 1/2$, Lemma 2 suggests that $\sigma \geq 0.10$ can ensure the uniqueness of $\pi(\tau^y, s)$. Unlike the case with uniform distribution, assumption A2 is not a necessary condition for the uniqueness $\pi(\tau^y, s)$. In fact, $\pi(\tau^y, s)$ turns out to be unique in numerical experiments as long as σ is not too small.[27]

We solve functional equation (2.18) numerically. The algorithm is straightforward. First, choose a number of grids τ_i^y and s_i in the state space $[0,1] \times [\underline{s}, \overline{s}]$. Given any τ_i^y and s_i, (2.18) reduces to a nonlinear equation (2.34). Solution x contingent on τ_i^y and s_i can easily be obtained. This leads to an approximation of function $\pi(\tau^y, s)$ and $T(\tau^y, s)$ on $[0,1] \times [\underline{s}, \overline{s}]$. The cubic spline interpolation is used to compute the value of $T(\tau^y, s)$ at points not on the grid. Then we search for the optimal tax rate τ^y that maximizes $T(\tau^y, s)$.

[27] For $f(s', s) \geq 2/\beta$, we can check the uniqueness by plotting $\int_{s^0 \geq (\tau^y - \beta x)/2} f(s', s^*)\, ds'$ with respect to x to see if it only has one cross with the 45 degree line, where s^* is chosen such that $s^* = \arg\max f(s', s)$.

8.5 Characterization of the Equilibrium under Age-Independent Taxation

(2.41) solves

$$h(\tau, s) = \begin{cases} \frac{1+\beta/2+2\beta\delta/a^o-\tau}{2(1+\beta\delta/2)} & \text{if } \tau \leq \theta^-(s) \\ \frac{1+\beta/2+\beta\delta(z-1/2)/(a^o z)-\tau+(\beta\delta\rho/(a^o z))s}{2(1+\beta\delta(1-1/(a^o z))/2)} & \text{if } \tau \in \left(\theta^-(s), \theta^+(s)\right) \\ \frac{1+\beta/2-\tau}{2(1+\beta\delta/2)} & \text{if } \tau \geq \theta^+(s) \end{cases}, \quad (2.46)$$

where $\theta^-(s) \equiv -2z + \beta/2 - \beta\delta/2 + \beta\delta(2/a^o - z) + (2+\beta\delta)\rho s$ and $\theta^+(s) \equiv 2z + \beta/2 - \beta\delta/2 + \beta\delta z + (2+\beta\delta)\rho s$, $\theta^+(s) > \theta^-(s)$ for $z > \beta\delta/(2a^o(1+\beta\delta/2))$. We refer to $\theta^+(s) \leq 0$ and $\theta^-(s) \geq 1$ as the left-dominating and right-dominating region where the left-wing and right-wing will be elected with probability one for any τ, respectively.

Correspondingly, the probability for the right-wing of being elected in the second period equals

$$\begin{aligned} \pi(\tau, s) &= \int_{s' > \frac{1}{2} - h(\tau,s)} f(s', s)\, ds \\ &= \begin{cases} 1 & \text{if } \tau \leq \theta^-(s) \\ \frac{2z+\beta/2+\beta\delta(z-1/2)-\tau+(2+\beta\delta)\rho s}{4z(1+\beta\delta(1-1/(a^o z))/2)} & \text{if } \tau \in \left(\theta^-(s), \theta^+(s)\right) \\ 0 & \text{if } \tau \geq \theta^+(s) \end{cases}. \quad (2.47) \end{aligned}$$

It immediate follows that $\partial\pi(\tau, s)/\partial s \geq 0$ and $\partial h(\tau, s)/\partial s \geq 0$. A high s leads to a high expectation of s', which tends to increase π and thus h. Like (2.27), the effect of τ on the ideology-contingent expectation, $\partial\pi(\tau, s)/\partial\tau$, is a simple function of s:

$$\frac{\partial\pi(\tau, s)}{\partial\tau} = \begin{cases} -\frac{1}{4z(1+\beta\delta(1-1/(a^o z))/2)} & \text{if } \tau \in \left(\theta^-(s), \theta^+(s)\right) \\ 0 & \text{otherwise} \end{cases}. \quad (2.48)$$

For $z > \beta\delta/(2a^o(1+\beta\delta/2))$, $\partial\pi(\tau, s)/\partial\tau \leq 0$. A low τ may intuitively lead to a high probability for the right-wing of winning the next election. In the left-dominating or right-dominating region with $\theta^+(s) \leq 0$ or $\theta^-(s) \geq 1$, however, τ cannot affect the identity of the incumbent in the next period.

(2.46) implies that the absolute value of the elasticity of the tax base $\delta h_0 + h$

is equal to

$$\epsilon(\tau, s) = \begin{cases} \frac{\tau}{2(1+\beta\delta/2)\delta h_0+1+\beta/2+2\beta\delta/a^o-\tau} & \text{if } \tau \leq \theta^-(s) \\ \frac{\tau}{2(1+\beta\delta(1-1/(a^o z))/2)\delta h_0+1+\beta/2+\beta\delta(z-1/2)/(a^o z)-\tau+(\beta\delta\rho/(a^o z))s} & \text{if } \tau \in \left(\theta^-(s), \theta^+(s)\right) \\ \frac{\tau}{2(1+\beta\delta/2)\delta h_0+1+\beta/2-\tau} & \text{if } \tau \geq \theta^+(s) \end{cases}. \tag{2.49}$$

(2.49) illustrates how ideology affects the elasticity of the tax base. Like the elasticity of the tax base under age-dependent taxation (2.28), the impact of ideology on ϵ is not monotonic. The tax base is more elastic in the left-dominating region than in the right-dominating region. But $\epsilon(\tau, s)$ can be larger for $\theta^+(s) \in (0,1)$ than for $\theta^+(s) \leq 1$. This is because the negative effect of τ on the ideology-contingent expectation in (2.48) enhances the distortion of τ on h and makes h more elastic.

Plugging (2.46) into (2.42), we see that the incumbent's objective function is piecewise quadratic. A full characterization of $\tau(h_0, s)$ is given by the following proposition.

Proposition 4 *Assume that (2.24) and*

$$z > \frac{2\beta\delta + \sqrt{4\beta^2\gamma^2 + (2+\beta)\beta\delta a^o + 2a^o\delta^2\beta(2+\beta\delta)h_0}}{8a^o(1+\beta\delta/2)} \tag{2.50}$$

Then $\tau(h_0, s)$ follows

$$\tau(h_0, s) = \begin{cases} \tau^L(h_0, s) & \text{if } h_0 + s \leq 1/2 \\ \tau^R(h_0, s) & \text{if } h_0 + s > 1/2 \end{cases}, \tag{2.51}$$

where $\tau^j(h_0, s)$, $j = L, R$, follows

$$\tau^j(h_0, s) = \begin{cases} \gamma^j(h_0) + \beta\delta/a^o & \text{if } s \geq s^{3,j} \\ \theta^-(s) & \text{if } s^{2,j} < s < s^{3,j} \\ \phi^j(h_0, s) & \text{if } s^1 \leq s \leq s^{2,j} \\ \gamma^j(h_0) & \text{if } s \leq s^1 \end{cases}. \tag{2.52}$$

Functions γ^j and ϕ^j are defined as follows:

$$
\begin{aligned}
\gamma^L(h_0) &= \delta h_0\left(1+\frac{\beta\delta}{2}\right)+\frac{1}{2}+\frac{\beta}{4},\\
\gamma^R(h_0) &= \delta h_0\left(1-\frac{2}{a^o}\right)\left(1+\frac{\beta\delta}{2}\right)+\frac{1}{2}+\frac{\beta}{4},\\
\phi^L(h_0,s) &= \delta h_0\left(1+\frac{\beta\delta}{2}\left(1-\frac{1}{a^o z}\right)\right)+\frac{1}{2}+\frac{\beta}{4}+\frac{\beta\delta}{2a^o z}\left(z-\frac{1}{2}\right)+\frac{\beta\delta\rho s}{2a^o z},\\
\phi^R(h_0,s) &= \delta h_0\left(1-\frac{2}{a^o}\right)\left(1+\frac{\beta\delta}{2}\left(1-\frac{1}{a^o z}\right)\right)+\frac{1}{2}+\frac{\beta}{4}+\frac{\beta\delta}{2a^o z}\left(z-\frac{1}{2}\right)+\frac{\beta\delta\rho s}{2a^o z}.
\end{aligned}
$$

s^1, s^2 and s^3 equal

$$
s^1 = \frac{-\Pi_2+\sqrt{\Pi_2^2-4\Pi_1\Pi_3}}{2\Pi_1}, \tag{2.53}
$$

$$
\begin{aligned}
s^{2,L} = \frac{1}{\rho\left(2+\beta\delta\left(1-1/\left(2a^o z\right)\right)\right)}&\left[\frac{1}{2}+2z-\frac{\beta}{4}+\frac{\beta\delta}{2}\right.\\
&\left.-\beta\delta\left(\frac{3}{2a^o}-z+\frac{1}{4a^o z}\right)+\delta h_0\left(1+\frac{\beta\delta}{2}\left(1-\frac{1}{a^o z}\right)\right)\right],
\end{aligned} \tag{2.54}
$$

$$
\begin{aligned}
s^{2,R} = \frac{1}{\rho\left(2+\beta\delta\left(1-1/\left(2a^o z\right)\right)\right)}&\left[\frac{1}{2}+2z-\frac{\beta}{4}+\frac{\beta\delta}{2}-\beta\delta\left(\frac{3}{2a^o}-z+\frac{1}{4a^o z}\right)\right.\\
&\left.+\delta h_0\left(1-\frac{2}{a^o}\right)\left(1+\frac{\beta\delta}{2}\left(1-\frac{1}{a^o z}\right)\right)\right],
\end{aligned} \tag{2.55}
$$

$$
s^{3,L} = \frac{1}{\rho(2+\beta\delta)}\left[\frac{1}{2}+2z-\frac{\beta}{4}+\frac{\beta\delta}{2}-\beta\delta\left(\frac{1}{a^o}-z\right)+\delta h_0\left(1+\frac{\beta\delta}{2}\right)\right], \tag{2.56}
$$

$$
\begin{aligned}
s^{3,R} = \frac{1}{\rho(2+\beta\delta)}&\left[\frac{1}{2}+2z-\frac{\beta}{4}+\frac{\beta\delta}{2}-\beta\delta\left(\frac{1}{a^o}-z\right)\right.\\
&\left.+\delta h_0\left(1-\frac{2}{a^o}\right)\left(1+\frac{\beta\delta}{2}\right)\right],
\end{aligned} \tag{2.57}
$$

with

$$
\begin{aligned}
\Pi_1 &\equiv \frac{\left(\beta\delta\rho/\left(2a^o z\right)\right)^2}{2\left(1+\beta\delta\left(1-1/\left(a^o z\right)\right)/2\right)},\\
\Pi_2 &\equiv \frac{\beta\delta\rho}{2a^o z}\left(\delta h_0+\frac{1/2+\beta/4+\beta\delta(z-1/2)/\left(2a^o z\right)}{1+\beta\delta\left(1-1/\left(a^o z\right)\right)/2}\right),\\
\Pi_3 &\equiv -\frac{\beta\delta\left(\delta h_0\right)^2}{4a^o z}+\frac{\left(1/2+\beta/4+\beta\delta(z-1/2)/\left(2a^o z\right)\right)^2}{2\left(1+\beta\delta\left(1-1/\left(a^o z\right)\right)/2\right)}+\frac{\beta\delta(z-1/2)\delta h_0}{2a^o z}-\frac{1/2+\beta/4}{2(1+\beta\delta/2)}.
\end{aligned}
$$

8.5.1 Proof of Proposition 4

We only characterize $\tau^L(h_0,s)$. The characterization of $\tau^R(h_0,s)$ is essentially the same and thus omitted. Plugging (2.46) into (2.42), the left-wing incumbent's objective function is piecewise quadratic

$$V^L(h_0,\tau,s)=\begin{cases} T^A(h_0,\tau,s) & \text{if } \tau\le\theta^-(s) \\ T^B(h_0,\tau,s) & \text{if } \tau\in\left(\theta^-(s),\theta^+(s)\right) \\ T^C(h_0,\tau,s) & \text{if } \tau\ge\theta^+(s) \end{cases}, \tag{2.58}$$

where, for notational convenience, we define

$$\begin{aligned} T^A(h_0,\tau,s) &\equiv \tau\left(\delta h_0+\frac{1+\beta/2+2\beta\delta/a^o-\tau}{2(1+\beta\delta/2)}\right), \\ T^B(h_0,\tau,s) &\equiv \tau\left(\delta h_0+\frac{1+\beta/2+\beta\delta(z-1/2)/(a^o z)-\tau+(\beta\delta\rho/(a^o z))s}{2(1+\beta\delta(1-1/(a^o z))/2)}\right), \\ T^C(h_0,\tau,s) &\equiv \tau\left(\delta h_0+\frac{1+\beta/2-\tau}{2(1+\beta\delta/2)}\right). \end{aligned}$$

$\hat{\tau}^A(h_0,s)=\gamma^L(h_0)-\beta\delta/a^o$, $\hat{\tau}^B(h_0,s)=\phi^L(s)$ and $\hat{\tau}^C(h_0,s)=\gamma^L(h_0)$ are the solutions of $\max_\tau T^A(h_0,\tau,s)$, $\max_\tau T^B(h_0,\tau,s)$ and $\max_\tau T^C(h_0,\tau,s)$, respectively. Finally, let $\omega^-\equiv -2z+\beta/2-\beta\delta/2+\beta\delta(2/a^o-z)$ and $\omega^+\equiv -2z+\beta/2-\beta\delta/2+\beta\delta z$. So $\theta^-(s)=(2+\beta\delta)\rho s+\omega^-$ and $\theta^+(s)=(2+\beta\delta)\rho s+\omega^+$.

The proof is based on the following lemma. The proof is analogous to Lemma 5 and thus omitted.

Lemma 6

$$\tau^L(h_0,s)=\begin{cases} \gamma^L(h_0) & \textit{if } s\le s^{1,L} \\ \phi^L(h_0,s) & \textit{if } s\ge s^{1,L},\ s<s^{4,L},\ s<s^{2,L} \\ \gamma^L(h_0) & \textit{if } s\ge s^{1,L},\ s<s^{4,L},\ s\ge s^{2,L},\ s\le s^{5,L} \\ \theta^-(s) & \textit{if } s\ge s^{1,L},\ s<s^{4,L},\ s\ge s^{2,L},\ s\ge s^{5,L} \\ \phi^L(h_0,s) & \textit{if } s\ge s^{1,L},\ s\ge s^{4,L},\ s\le s^{2,L} \\ \theta^-(s) & \textit{if } s\ge s^{1,L},\ s\ge s^{4,L},\ s\in\left(s^{2,L},s^{3,L}\right) \\ \gamma^L(h_0)+\beta\delta/a^o & \textit{if } s\ge s^{1,L},\ s\ge s^{4,L},\ s\ge s^{3,L} \end{cases}$$

where $s^{4,L}\equiv\delta h_0/(2\rho)+(2-\beta+8z+2\beta\delta(1-2z))/(4\rho(2+\beta\delta))$.

When (2.50) holds, it can easily be shown that $s^{2,L}\ge s^{4,L}$. Thus, $s<s^{4,L}$ and $s\ge s^{2,L}$ cannot hold simultaneously. Moreover, since $s^{4,L}\ge s^{1,L}$, we have $s^{2,L}\ge s^{1,L}$. Lemma 6 leads to Proposition 4, hence. □

8.6 Multiple Equilibria

If $z < \beta\delta / (2a^o (1 + \beta\delta/2))$, (2.46) does not hold any longer. After ruling out the counter-intuitive expectation $\pi(\tau, s)$ by requiring the expectation to be monotonic, (2.41) solves

$$h(\tau, s) = \begin{cases} \frac{1+\beta/2+2\beta\delta/a^o - \tau}{2(1+\beta\delta/2)} & \text{if } \tau \leq \theta^-(s) \\ \frac{1+\beta/2-\tau}{2(1+\beta\delta/2)} & \text{if } \tau \geq \theta^+(s) \end{cases} \tag{2.59}$$

and

$$\pi(\tau, s) = \begin{cases} 1 & \text{if } \tau \leq \theta^-(s) \\ 0 & \text{if } \tau \geq \theta^+(s) \end{cases}. \tag{2.60}$$

(2.59) and (2.60) imply that the ideology-contingent $h(\tau, s)$ and $\pi(\tau, s)$ are not unique for $\tau \in \left(\theta^+(s), \theta^-(s)\right)$. The indeterminacy of expectations creates multiple Markov perfect equilibria. Pick up a particular expectation rule satisfying (2.60)

$$h(\tau, s) = \begin{cases} \frac{1+\beta/2+2\beta\delta/a^o - \tau}{2(1+\beta\delta/2)} & \text{if } \tau < \theta^+(s) + \psi \\ \frac{1+\beta/2-\tau}{2(1+\beta\delta/2)} & \text{if } \tau \geq \theta^+(s) + \psi \end{cases}.$$

The corresponding human capital investment is

$$\pi(\tau, s) = \begin{cases} 1 & \text{if } \tau < \theta^+(s) + \psi \\ 0 & \text{if } \tau \geq \theta^+(s) + \psi \end{cases},$$

where $\psi \in [0, 2\beta\delta/a^o - 4z - 2\beta\delta z]$. The left-wing and right-wing incumbents set $\tau(h_0, s)$ by maximizing $\tau(\delta h_0 + h(\tau, s))$ and $a^o\tau(\delta h_0 + h(\tau, s))/2 + (1-\tau)\delta$, respectively. Some algebra establishes the following proposition:

Proposition 5 *Assume that (2.24) and $z < \beta\delta / (2a^o (1 + \beta\delta/2))$. There are multiple Markov perfect equilibria such that, for any $\psi \in [0, 2\beta\delta/a^o - 4z - 2\beta\delta z]$, $\tau(h_0, s)$ follows*

$$\tau(h_0, s) = \begin{cases} \tau^L(h_0, s) & \text{if } h_0 + s \leq 1/2 \\ \tau^R(h_0, s) & \text{if } h_0 + s > 1/2 \end{cases},$$

where the functions $\tau^j(h_0, s)$, $j = L, R$, are defined as follows:

$$\tau^j(h_0, s) = \begin{cases} \gamma^j(h_0) + \beta\delta/a^o & \text{if } s \geq \varphi^{2,j} \\ \theta^+(s) + \psi & \text{if } \varphi^1 \leq s < \varphi^{2,j} \\ \gamma^j(h_0) & \text{if } s \leq \varphi^1 \end{cases},$$

with

$$\begin{aligned}
\varphi^{1} &\equiv (\eta - 2z - \beta/2 + \beta\delta z - \psi) / ((2+\beta\delta)\rho), \\
\varphi^{2,j} &\equiv \left(\gamma^{j}(h_0) + \beta\delta/a^{o} - 2z - \beta/2 + \beta\delta z - \psi\right) / ((2+\beta\delta)\rho), \\
\eta &\equiv \left(1 + \beta/2 + 2\beta\delta/a^{o} + \delta h_0 (2+\beta\delta) - 2\sqrt{\beta\delta\left(1+\beta/2+2\beta\delta/a^{o}+\delta h_0(2+\beta\delta)\right)/a^{o}}\right)/2.
\end{aligned}$$

Assume that (2.24) and $z < \beta\delta/(2a^{o}(1+\beta\delta/2))$. There are multiple Markov perfect equilibria such that, for any $\psi \in [0, 2\beta\delta/a^{o} - 4z - 2\beta\delta z]$, $\tau(h_0, s)$ follows

$$\tau(h_0,s) = \begin{cases} \tau^{L}(h_0,s) & \text{if } h_0 + s \leq 1/2 \\ \tau^{R}(h_0,s) & \text{if } h_0 + s > 1/2 \end{cases},$$

where the functions $\tau^{j}(h_0, s)$, $j = L, R$, are defined as follows:

$$\tau^{j}(h_0,s) = \begin{cases} \gamma^{j}(h_0) + \beta\delta/a^{o} & \text{if } s \geq \varphi^{2,j} \\ \theta^{+}(s) + \psi & \text{if } \varphi^{1} \leq s < \varphi^{2,j} \\ \gamma^{j}(h_0) & \text{if } s \leq \varphi^{1} \end{cases},$$

with

$$\begin{aligned}
\varphi^{1} &\equiv (\eta - 2z - \beta/2 + \beta\delta z - \psi) / ((2+\beta\delta)\rho), \\
\varphi^{2,j} &\equiv \left(\gamma^{j}(h_0) + \beta\delta/a^{o} - 2z - \beta/2 + \beta\delta z - \psi\right) / ((2+\beta\delta)\rho), \\
\eta &\equiv \left(1 + \beta/2 + 2\beta\delta/a^{o} + \delta h_0 (2+\beta\delta) - 2\sqrt{\beta\delta\left(1+\beta/2+2\beta\delta/a^{o}+\delta h_0(2+\beta\delta)\right)/a^{o}}\right)/2.
\end{aligned}$$

Chapter 3

Dynamic Inequality and Social Security*

1 Introduction

Most developed countries have large public pension programmes, involving not only inter-generational but also intra-generational transfers. For instance, social security contributions are roughly proportional to income while benefits have important lump-sum components. The impact of an exogenous social security on household intertemporal choices and welfare has been extensively studied in the literature.[2] The social security system, however, is not exogenous but endogenously determined by policy choices that reflect rich dynamic interactions between political and economic factors. For instance, the evolution of the distribution of household characteristics may alter the political support for the system, since households with different characteristics tend to have different preferences over social security transfers. Despite this, most of the existing literature has either assumed away politico-economic factors or, when considering them, it has focused on models where the size of social security is decided once-and-for-all. As a result, the effect of endogenous changes of household characteristics over time on the decision of social security transfers has been ignored altogether. (e.g.

* I am very grateful to Fabrizio Zilibotti for guidance and numerous discussions. I also thank John Hassler, Dirk Niepelt and seminar participants at IIES for helpful comments.

[2] See Auerbach and Kotlikoff (1987), Imrohoroglu *et al.* (1995) and Storesletten *et al.* (1999) among many others.

Tabellini, 2000, Cooley and Soares, 1999, Conesa and Krueger, 1999).[3]

The present paper explores the positive implications and the welfare properties of a rational choice theory implying rich dynamic interactions between private intertemporal choices and political decisions on social security. To this end, we construct a dynamic general equilibrium model where agents repeatedly vote over the social security system, and confront its predictions with empirical evidence. We also analyse normative implications by comparing the political equilibrium with the Ramsey allocation chosen by a benevolent planner with a commitment technology.

In our model, the incumbent government cannot commit to future social security transfers since they are decided by future elected governments. Instead, the pension system is determined in each period by its current constituency, of which the extent of wealth inequality is a key factor. Forward-looking households adjust their private savings when rationally anticipating the equilibrium dynamics of wealth inequality and social security. A main finding is that this interaction leads to an equilibrium where social security transfers increase over time after their initial introduction. The underlying mechanism is twofold. On the one hand, the establishment of a social security system increases future wealth inequality since within-cohort transfers discourage the private savings of the poor more than that of the rich. On the other hand, the larger wealth inequality makes social security transfers more desirable in the future. This provides political support for an increasing size of social security in the following periods.

Our workhorse is a standard two-period Overlapping-Generation model. To capture the intra-generational redistributive role of social security, we incorporate within-cohort heterogeneity by assuming the young households to be born with different labour productivities. Old households are different in terms of wealth. In other words, there exists multi-dimensional heterogeneity across voters. Each group of voters has its own preferences over social security transfers. The political decision process is modelled by a repeated probabilistic voting framework.[4] In

[3] A notable exception is Boldrin and Rustichini (2000), where the interaction between private intertemporal choices and political decisions may lead to a decreasing size of social security.

[4] The probabilistic voting framework is adapted from Lindbeck and Weibull (1987). See Hassler *et al.* (2005) and Gonzalez-Eiras and Niepelt (2004) for the application of repeated probabilistic voting in a dynamic political equilibrium.

equilibrium, policymaker candidates respond to electoral uncertainty by proposing a policy platform that maximizes a weighted-average welfare of all groups of voters.

Previous literature has studied the sustainability and evolution of social security by assuming that voters play trigger strategies (e.g. Boldrin and Rustichini, 2000). Although trigger strategies may provide analytical convenience and have reasonable components, this assumption is not innocuous. Due to the multiplicity of equilibria, it is hard to provide sharp empirical predictions. More importantly, the trigger strategy equilibria are not robust to refinements such as backward induction in a finite-horizon economy when the finite-horizon tends to infinity. In the present paper, we focus on the Markov perfect equilibrium, where the size of social security is conditioned on the payoff-relevant fundamental elements, the aggregate capital stock and the distribution of assets held by old households. The Markov perfect equilibrium in our model turns is unique, which can be obtained as one takes the limit of a finite horizon environment. Thus, it does not suffer from the shortcomings of the trigger strategy equilibria discussed above. Moreover, under logarithm utility and Cobb-Douglas production technology, the unique Markov equilibrium can be characterized analytically, making the economic mechanism highly transparent. We show that the equilibrium social security tax rate is increasing in wealth inequality and independent of the aggregate capital stock.

Our theory delivers a number of empirical predictions that can be confronted with empirical evidence. First, the prediction that the size of social security programmes should grow over time after their initial introduction is consistent with stylized facts from all major OECD countries. Second, our model predicts an increasing consumption inequality over the life cycle, which is in accordance with the empirical evidence in a number of countries (Deaton and Paxson, 1994). The uniqueness of the Markov equilibrium also allows us to run comparative statics, which is problematic in models featuring multiple equilibria. We find that a high population growth rate, or a low proportion of old households, leads to a small size of social security, since it increases the weight on the welfare of young households in the political decision process. The positive relationship between the size of social security and the old-aged population is consistent with a number of cross-country studies (e.g. Tabellini, 2000, Mulligan and Sala-i-Martin, 1999a).

Since the size of social security depends on the wealth distribution, an income inequality shock does not have any immediate effects on the equilibrium level of social security transfers. Over time, however, a larger income inequality increases the wealth inequality and thus leads to a larger size of social security in the long-run. This positive relationship between income inequality and social security is also in accordance with empirical studies (e.g. Tabellini, 2000).

The tractable model allows a comparison between the politico-economic equilibrium outcomes and the efficient (Ramsey) allocation, in which a benevolent planner with a commitment technology maximizes the discounted sum of the welfare of all current and future generations. Under logarithm utility and Cobb-Douglas production technology, the Ramsey solution can be characterized analytically. There are two effects that make the *initial* social security tax rate in the political equilibrium deviate from the optimal level. First, since voters are short-lived and non-altruistic towards future generations, the political decision process does not take into account the negative impact of social security taxes on capital accumulation and the welfare of future generations. This is referred to as political effect I. Intuitively, this effect induces too large transfers. Second, the weight on the welfare of current young generation in the political decision process tends to be larger than the weight in the Ramsey problem. This is referred to as political effect II. Since social security has a negative effect on the welfare of young households, this effect yields insufficient transfers. Under reasonable parameter values, political effect II is found to dominate political effect I. Therefore, the political equilibrium provides less social security transfers in the *initial* period, compared with the efficient allocation. As for the transfers in periods other than the initial one, the Ramsey planner will impose low taxation to encourage capital accumulation since she can commit to future policies. Not surprisingly, the Ramsey tax rates turn out to be lower than the political equilibrium outcomes in the long run.

As noted above, logarithm utility helps to obtain a closed-form solution of the dynamic political equilibrium. It is worth emphasizing that in the Markov equilibrium, voters do not only hold rational expectations on future equilibrium outcomes, but may strategically affect future policies via the impact of current taxation on private intertemporal choices. Under logarithm utility, the current tax rate has symmetric effects on the private saving of the rich and poor, however. Thus, it cannot affect the future state of the economy (wealth inequality), nor future policy outcomes. In other words, strategic effects are mute in the particular case of logarithmic utility. Strategic effects are instead present under general CRRA utility when the intertemporal elasticity of substitution is different from unity. In these cases analytical results cannot be obtained, but we can numerically study the qualitative and quantitative impact of the strategic effects. To this end, it is useful to compare the Markov perfect equilibrium with an environment (referred to as "the myopic voting equilibrium"), where voters can rationally expect future policy outcomes but assume, incorrectly, that there are no strategic interaction between the current and future policies.[5] We show that if the elasticity of intertemporal substitution is smaller than unity, as suggested by most empirical literature, the strategic effect is positive: a higher current tax rate leads to a higher future wealth inequality and hence larger transfers in future. Due to the positive strategic effect, current voters have the incentive to strategically raise current social security taxes, in order to obtain larger social security benefits in the future. Under reasonable parameter values, numerical exercises indicate that the strategic effect in the Markovian equilibrium is quantitatively less important: social security transfers are about 4% higher than those in the myopic voting equilibrium.

The sustainability of the social security system has been widely discussed in the literature.[6] However, few works investigate the dynamic pattern of the system. Two exceptions are Katuscak (2002) and Forni (2005). Katuscak (2002) shows that imposing social security may discourage private savings and make the old poorer after retirement. This induces the policymaker to increase future transfers and thus, yields an increasing sequence of social security tax rates over time. The present paper differs from Katuscak's work along two dimensions: incorporating within-cohort heterogeneity and endogenizing factor prices. Our model suggests that, though the inter-generational redistribution effect is key to sustain the system, the intra-generational redistribution effect plays a central role in the evolution of the system. The growing sizes of the pension system are generated

[5] A similar notion of pseudo-equilibrium is used by Alesina and Rodrik (1994).

[6] See, for example, Boldrin and Rustichini (2000), Cooley and Soares (1999), Conesa and Krueger (1999), Mulligan and Sala-i-Martin (1999a, 1999b), Tabellini (2000), Razin *et al.* (2002), Gonzalez-Eiras and Niepelt (2004).

by the interaction between wealth inequality and social security transfers. Forni (2005) adopts the median voter framework. He shows that the social security system can be sustained by self-fulfilled expectations on the positive relationship between current and future transfers, which implies a growing system. Although Forni also focuses on the Markov equilibrium, the shortcomings in trigger strategies equilibria apply to his approach. There are multiple Markov equilibria, none of which survives a finite-horizon environment.

Our work is also part of a growing literature on dynamic politico-economic equilibrium, where current voting may change fundamentals in the future political environment and hence, affect future policy outcomes. Because of the complexity of dynamic interaction between individual intertemporal choice and voting strategy, a closed-form solution is usually implausible except in some small open economies (e.g. Hassler *et al.*, 2003, Katuscak, 2002, Hassler *et al.*, 2005). Gonzalez-Eiras and Niepelt (2004) show that an analytical solution can be obtained in a growth model with logarithm utility and Cobb-Douglas production technology. The trick is that the equilibrium voting strategy turns out to be a constant and to be independent of fundamentals in the political environment. Our work generalizes Gonzalez-Eiras and Niepelt's analytical results by incorporating within-cohort heterogeneity. More importantly, the equilibrium voting strategy becomes nontrivially dependent on fundamentals in the political environment. This is in sharp contrasts with the literature that resorts to numerical characterizations for nontrivial equilibrium voting strategies in closed-economies (e.g. Krusell *et al.*, 1997, Krusell and Rios-Rull, 1999).

The rest of the paper is organized as follows. Section 2 presents the model. In Section 3, the dynamic politico-economic equilibrium is defined and solved under logarithm utility. The predictions of the model are tested on empirical evidence. Section 4 characterizes the Ramsey solution. In Section 5, we solve the political equilibrium and the Ramsey allocation under a more general CRRA utility form. Section 6 concludes.

2 The Model

Consider an economy inhabited by an infinite sequence of overlapping-generations. Each generation lives for two periods. Households work in the first period of their life and then retire. Labour supply is inelastic and normalized to unity. Assume the gross population growth rate N^t/N^{t-1} to be a constant $n \geq 1$, where N^t denotes the population of the cohort born at time t.

Young households have probability P to be endowed with the high labour productivity γ^h and probability $1-P$ to be endowed with the low productivity γ^l ($\gamma^h > \gamma^l$). For simplicity, let $P = 1/2$.[7] Households with type $j = l$ (h) are referred to as poor (rich). Wage income is taxed at the flat rate, τ_t. The after-tax net earning for young households of type j is $(1-\tau_t)\, w_t^j$. Old households receive benefits b_t from a social security system and young households may save to finance their consumption after retirement. The corresponding intertemporal decision solves

$$\max_{k_{t+1}^j} u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right), \tag{3.1}$$

subject to

$$c_t^{y,j} = (1-\tau_t)\, w_t^j - k_{t+1}^j, \tag{3.2}$$

$$c_{t+1}^{o,j} = R_{t+1} k_{t+1}^j + b_{t+1}, \tag{3.3}$$

where $c_t^{i,j}$ and k_{t+1}^j denote the consumption and savings of households of type (i,j), $i \in \{y,o\}$ and $j \in \{l,h\}$, respectively. The discount factor is $\beta \in (0,1)$. R_{t+1} is the gross interest rate at time $t+1$. We assume that $u(c) = \log(c)$, an assumption which will be relaxed in Section 5.

Let K_t and L_t be the aggregate capital stock and effective labour used in production at time t. The clearance of factor markets requires $K_t = N^{t-1}\left(k_t^l + k_t^h\right)/2$ and $L_t = N^t\left(\gamma^l + \gamma^h\right)/2$. Without loss of generality, the average productivity $\left(\gamma^l + \gamma^h\right)/2$ is normalized to unity so that $\gamma^h = 2 - \gamma^l$ and $L_t = N^t$. Assume that production follows Cobb-Douglas technology with a constant return to scale: $AK_t^\alpha L_t^{1-\alpha}$. A denotes total factor productivity and $\alpha \in (0,1)$ is the output elasticity of capital. Factor markets are competitive and factor prices thus correspond to marginal products

$$R_t = A\alpha\left(k_t/n\right)^{\alpha-1}, \tag{3.4}$$

$$w_t = A(1-\alpha)\left(k_t/n\right)^{\alpha}, \tag{3.5}$$

where $k_t \equiv \left(k_t^h + k_t^l\right)/2$ is the average wealth holdings of old households. The individual wage rate is $w_t^j = \gamma^j w_t$. The average wage rate thus equals w_t.

[7] P has no effect on the main results below.

The flat-rate wage income tax rate τ_t is determined through some political process that will be specified below. τ_t is imposed on the working generation to finance social security payments. In addition to the inter-generational redistribution which defines the pay-as-you-go system, pensions entail intra-generational redistributive elements. In most systems, social security contributions are proportional to income, while benefits have lump-sum or even regressive components. According to the Old Age Insurance of the U.S. social security system, for example, a 1% increase in lifetime earnings leads to a 0.90%, 0.32%, 0.15% and 0.00% increase in pension benefits from low to high income groups. Following Conesa and Krueger (1999) and many others, we assume, for analytical convenience, social security benefits to be evenly distributed within old households.[8] It is also assumed that the budget of the social security system must be balanced in each period. This implies that at any time t, social security payments $b_t N^{t-1}$ equal social security contributions $\tau_t \left(w_t^l + w_t^h\right) N^t/2$, i.e.

$$b_t = n\tau_t w_t. \tag{3.6}$$

2.1 Households' Saving Choice

Under logarithm utility, the households' saving choices can be analytically obtained by the Euler equation, $c_{t+1}^{o,j}/c_t^{y,j} = \beta R_{t+1}$, which solves (3.1). Since households are atomic, they take factor prices, aggregate savings, the current social security tax rate and future social security benefits as given. Plugging factor prices (3.4), (3.5) and the balanced budget rule (3.6) into (3.2) and (3.3), the Euler equation solves a doublet of private saving functions

$$k_{t+1}^h = S^h\left(k_t, \tau_t, \tau_{t+1}\right) \equiv \omega\left(\tau_{t+1}\right)\psi\left(\tau_{t+1}\right) A\left(1-\tau_t\right)\left(k_t/n\right)^{\alpha}, \tag{3.7}$$

$$k_{t+1}^l = S^l\left(k_t, \tau_t, \tau_{t+1}\right) \equiv \psi\left(\tau_{t+1}\right) A\left(1-\tau_t\right)\left(k_t/n\right)^{\alpha}, \tag{3.8}$$

[8] Social security wealth can be divided into two components: annuity and government transfer value. Take the U.S. as an example. Wolff (1992) shows that 85% of the social security wealth of old households in 1969 took the form of pure government transfers. Although this number has been declining over time and dropped to 66% in 1983, government transfers still constitute the major part of social security benefits. The annuity portion related to social security contributions is regarded as returns from private savings in our model. The implicit assumption is that the government runs the annuity fund in a perfect capital market such that annuity contributions play the same role as private savings.

where $\psi(\cdot)$ and $\omega(\cdot)$ are defined as:

$$\omega(\tau_{t+1}) \equiv \frac{\theta\alpha(1+\beta)+(\theta-1)(1-\alpha)\tau_{t+1}/2}{\alpha(1+\beta)-(\theta-1)(1-\alpha)\tau_{t+1}/2}, \quad (3.9)$$

$$\psi(\tau_{t+1}) \equiv \frac{\gamma^l(1-\alpha)\beta(\alpha(1+\beta)-(\theta-1)(1-\alpha)\tau_{t+1}/2)}{(1+\beta)(\alpha(1+\beta)+(1-\alpha)\tau_{t+1})}, \quad (3.10)$$

where $\theta \equiv \gamma^h/\gamma^l$ denotes the ratio of labour productivity of the rich to that of the poor. It is easy to show that $S_1^j > 0$, $S_2^j < 0$ and $S_3^j < 0$, where subscript i denotes the partial derivative with respect to the ith argument of S.[9] A higher k_t increases the wage rate and thus, private savings. The effect of a high τ_t is the opposite. Social security benefits increase the income after retirement and hence, discourage the private savings.

Note that $\theta = w_t^h/w_t^l$ and $\omega(\tau_t) = k_t^h/k_t^l$ measure young households' income inequality and old households' wealth inequality (excluding social security benefits), respectively.[10] Without social security system ($\tau_t = 0\ \forall t$), wealth inequality $\omega(0)$ coincides with income equality θ. However, the establishment of a social security system affects future wealth inequality k_{t+1}^h/k_{t+1}^l via τ_t and τ_{t+1}. First, under logarithm utility, (3.7) and (3.8) imply that τ_t has a symmetric impact on k_{t+1}^h and k_{t+1}^l and thus, does not affect k_{t+1}^h/k_{t+1}^l. Second, since $\omega(\tau_{t+1})$ increases in τ_{t+1}, a high future social security tax rate τ_{t+1} enlarges future wealth inequality. The poor receive the same amount of social security benefits as the rich after retirement, while their earnings are smaller than those of the rich. Therefore, high social security benefits discourage savings of the poor more than the rich. The results are written in Lemma 1.

Lemma 1 *Assume that $u(c) = \log(c)$. Future wealth inequality k_{t+1}^h/k_{t+1}^l increases in the future social security tax rate τ_{t+1}. Given τ_{t+1}, k_{t+1}^h/k_{t+1}^l does not depend on the current social security tax rate τ_t and aggregate capital k_t.*

Lemma 1 states an important property that will be repeatedly used in the following analysis: the choice of the current tax rate has no effect on future wealth inequality. This property is due to the assumption of logarithm utility, which

[9] We assume that $\alpha(1+\beta) > (\theta-1)(1-\alpha)/2$. Thus, the savings of the poor are always positive.

[10] To avoid confusion, wealth inequality is hereinafter referred to as inequality in terms of old households' wealth, excluding social security benefits.

cancels out the substitution and income effect and thus makes private savings proportional to labour income. As will be seen below, Lemma 1 substantially simplifies the analysis. (3.7) and (3.8) lead to the law of motion of aggregate capital

$$k_{t+1} = S(k_t, \tau_t, \tau_{t+1}) \equiv \phi(\tau_{t+1}) A (1 - \tau_t)(k_t/n)^{\alpha}, \tag{3.11}$$

where $\phi(\cdot)$ is defined as

$$\phi(\tau_{t+1}) \equiv \frac{\alpha\beta(1-\alpha)}{\alpha(1+\beta) + (1-\alpha)\tau_{t+1}}. \tag{3.12}$$

It immediately follows that $S_1 > 0$, $S_2 < 0$ and $S_3 < 0$. These aggregate results come from $S_1^j > 0$, $S_2^j < 0$ and $S_3^j < 0$ implied by the private saving functions (3.7) and (3.8).

3 Political Equilibrium

The social security tax rate τ_t is chosen by some repeated political process at the beginning of each period. In the present paper, we assume that τ_t is determined in a probabilistic voting framework (Lindbeck and Weibull, 1987). There are two policy-maker candidates running electoral competition. The winner obtains the majority of the votes of all current voters with unobservable ideological preferences towards political candidates. Since candidates only care about winning the election, they will, in equilibrium, respond to electoral uncertainty by proposing a policy platform that maximizes a weighted-average welfare of all current voters. The weights reflect the sensitivity of different groups of voters to policy changes.[11] In the context of our model, the political decision process of τ_t can be formalized as

$$\max_{\tau_t} \sum_{j=h,l} u\left(c_t^{o,j}\right) + n \sum_{j=h,l} \left(u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right)\right). \tag{3.13}$$

For notational convenience, the weights on different groups' utility are set equal.[12]

We focus on Markov perfect equilibrium, in which the state of the economy is summarized by the distribution of the assets held by old households, k_t^h and k_t^l. In the Markov equilibrium, the current political decision may affect the future asset

[11] See Persson and Tabellini (2000) for a more detailed discussion of probabilistic voting.

[12] Deviation from equal weights does not affect the main results below.

distribution and thus, the future social security tax rate. Forward-looking voters would adjust their intertemporal choice accordingly. The Markovian policy rule of τ_t can be written as

$$\tau_t = \mathbf{z}\left(k_t^h, k_t^l\right), \tag{3.14}$$

where $\mathbf{z} : R^+ \times R^+ \to [0,1]$ is assumed to be continuously differentiable. To see how the Markovian policy rule affects households' intertemporal choice, we substitute (3.14) for τ_{t+1} in (3.7) and (3.8) and solve a recursive form of private saving functions

$$k_{t+1}^j = \hat{S}^j\left(k_t, \tau_t\right). \tag{3.15}$$

Differentiating (3.7) and (3.8) with respect to τ_t pins down the partial derivatives of the saving function $\hat{S}^j$:

$$\hat{S}_2^h = \frac{S_2^h + \mathbf{z}_2\left(S_2^l S_3^h - S_2^h S_3^l\right)}{1 - \mathbf{z}_1 S_3^h - \mathbf{z}_2 S_3^l}, \tag{3.16}$$

$$\hat{S}_2^l = \frac{S_2^l + \mathbf{z}_1\left(S_2^h S_3^l - S_2^l S_3^h\right)}{1 - \mathbf{z}_1 S_3^h - \mathbf{z}_2 S_3^l}. \tag{3.17}$$

Note that $\hat{S}_2^j$ generally differs from S_2^j. Correspondingly, the law of motion of aggregate capital becomes

$$k_{t+1} = \hat{S}\left(k_t, \tau_t\right) \equiv \left(\hat{S}^h\left(k_t, \tau_t\right) + \hat{S}^l\left(k_t, \tau_t\right)\right)/2, \tag{3.18}$$

with $\hat{S}_2 = \left(\hat{S}_2^h\left(k_t, \tau_t\right) + \hat{S}_2^l\left(k_t, \tau_t\right)\right)/2$.

Given the Markovian policy rule $\mathbf{z}$, the political decision on τ_t solves (3.13), subject to budget constraints (3.2) and (3.3), factor prices (3.4) and (3.5), the balanced-budget rule (3.6), private saving functions (3.15), and the law of motion of aggregate capital (3.18). This yields an actual policy rule $\tau_t = \bar{\mathbf{z}}\left(k_t^h, k_t^l\right)$, with $\bar{\mathbf{z}} : R^+ \times R^+ \to [0,1]$. $\mathbf{z}$ is a Markovian equilibrium policy rule, if and only if $\bar{\mathbf{z}} = \mathbf{z}$. The formal definition of the Markov perfect equilibrium is given as follows.

Definition 1 *A Markov perfect political equilibrium is a triplet of functions $\tilde{S}^h$, $\tilde{S}^l$ and $\mathbf{z}$, where the private saving function $\tilde{S}^j : R^+ \times R^+ \to R^+$, $j \in \{h, l\}$, and the policy rule $\mathbf{z} : R^+ \times R^+ \to [0,1]$ are such that:*

(1) Given the policy rule $\mathbf{z}$, $\tilde{S}^j\left(k_t^h, k_t^l\right) = \hat{S}^j\left(k_t^j, \mathbf{z}\left(k_t^h, k_t^l\right)\right)$, where $\hat{S}^j$ solves the Euler equations.

(2) Given $\mathbf{z}$ *and* $\hat{S}^j$, $\bar{\mathbf{z}}$ *solves (3.13), subject to (3.2) to (3.6), (3.15) and (3.18).*

(3) $\bar{\mathbf{z}} = \mathbf{z}$.

To solve the equilibrium policy rule $\mathbf{z}$, let us look at the impact of the social security tax rate τ_t on the welfare of various groups of voters. Denote $U_t^{i,j}$ as the welfare of the households of type (i,j), $i \in \{y,o\}$ and $j \in \{l,h\}$, with $U_t^{y,j} \equiv u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right)$ and $U_t^{o,j} \equiv u\left(c_t^{o,j}\right)$. Differentiating the utility of old households with respect to τ_t yields

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = u'\left(c_t^{o,j}\right) n w_t > 0. \tag{3.19}$$

Needless to say, old households always benefit from social security transfers. Substituting for $c_t^{o,j}$ and w_t, (3.19) can be rewritten as

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{1-\alpha}{2\alpha \frac{k_t^j/k_t^i}{1+k_t^j/k_t^i} + (1-\alpha)\tau_t}, \tag{3.20}$$

where $i,j \in \{h,l\}$, $i \neq j$. Somewhat surprisingly, $\partial U_t^{o,j}/\partial \tau_t$ turns out to be independent of aggregate capital k_t and solely related to wealth distribution. This highlights the role of social security as an intra-generational redistributive policy. Specifically, the smaller is the wealth of old households, the more welfare gains can they get from social security transfers. Although the rich gain less, the aggregate welfare effect of τ_t on old households, $\partial U_t^o/\partial \tau_t = \sum_{j=h,l}\left(\partial U_t^{o,j}/\partial \tau_t\right)/2$, increases in wealth inequality due to the concavity of utility.[13]

Differentiating the utility of young households with respect to τ_t yields

$$\frac{\partial U_t^{y,j}}{\partial \tau_t} = -u'\left(c_t^{y,j}\right)\gamma^j w_t + \beta u'\left(c_{t+1}^{o,j}\right)\left(k_{t+1}^j \frac{\partial R_{t+1}}{\partial k_{t+1}} + n\tau_{t+1}\frac{\partial w_{t+1}}{\partial k_{t+1}}\right)\hat{S}_2 + \beta u'\left(c_{t+1}^{o,j}\right) n w_{t+1}\frac{\partial \tau_{t+1}}{\partial \tau_t}. \tag{3.21}$$

Note that the effect of τ_t via k_{t+1}^j cancels out due to the Euler equation. The first term in (3.21) reflects the direct cost of social security contributions. The second term captures the general equilibrium effect of τ_t via its impact on capital accumulation $\hat{S}_2$. The general equilibrium effect is twofold. On the one hand, a high τ_t reduces private savings at time t, and thus reduces the tax base of social security at time $t+1$. On the other hand, young households at time t

[13] This can be formally derived by showing $\left(\partial U_t^o/\partial \tau_t\right)/\partial\left(k_t^h/k_t^l\right) > 0$ for $k_t^h/k_t^l > 1$.

benefit from a higher interest rate R_{t+1}. As long as τ_{t+1} or wealth inequality is not very large, the interest rate effect dominates the first effect.[14] Hence, the general equilibrium effect can benefit the young households.[15] The third term is the "strategic effect", which captures the fact that voters can affect the future tax rate τ_{t+1} by their current choices of τ_t. The sign and size of the strategic effect are determined by $\partial\tau_{t+1}/\partial\tau_t$, where

$$\frac{\partial\tau_{t+1}}{\partial\tau_t} = \mathbf{z}_1\left(k_{t+1}^h, k_{t+1}^l\right)\hat{S}_2^h\left(k_t,\tau_t\right) + \mathbf{z}_2\left(k_{t+1}^h, k_{t+1}^l\right)\hat{S}_2^l\left(k_t,\tau_t\right). \tag{3.22}$$

If $\partial\tau_{t+1}/\partial\tau_t > 0$ (< 0), young households know that a higher current social security tax rate leads to higher (lower) social security benefits in the future. Thus, they may strategically increase (reduce) τ_t as compared to the case where the current political choice does not affect future policy outcomes.[16]

Then, the first-order condition of the problem (3.13) can be written as

$$\frac{\partial U_t^{o,j}}{\partial\tau_t} + n\sum_{j=h,l}\frac{\partial U_t^{y,j}}{\partial\tau_t} + \lambda_t = 0, \tag{3.23}$$

where λ_t denotes the multiplier on the non-negative constraint of τ_t, $\lambda_t = 0$ for $\tau_t > 0$ and $\lambda_t > 0$ for $\tau_t = 0$.[17] (3.23) implies a function equation for $\mathbf{z}$. Under logarithm utility, the fixed-point can be analytically obtained as the limit of finite-horizon solutions.

Proposition 1 *Assume $u(c) = \log(c)$. Define $\upsilon \equiv n(1+\alpha\beta)/(1-\alpha)$. In the Markov perfect equilibrium, the policy rule $\mathbf{z}\left(k_t^h, k_t^l\right)$ follows*

$$\mathbf{z}\left(k_t^h, k_t^l\right) = \begin{cases} H\left(k_t^h/k_t^l\right) > 0 & \text{if } \upsilon\alpha < 1 \\ 0 & \text{if } \upsilon\alpha \geq 1 \text{ and } k_t^h/k_t^l > \Theta(\upsilon) \end{cases}, \tag{3.24}$$

where

$$H\left(k_t^h/k_t^l\right) \equiv \frac{-\Phi(\upsilon) + \sqrt{\Phi(\upsilon)^2 - 4\Delta(\upsilon)\left(\frac{4\upsilon\alpha^2 k_t^h/k_t^l}{\left(1+k_t^h/k_t^l\right)^2} - \alpha\right)}}{2\Delta(\upsilon)} \tag{3.25}$$

---

[14] This can be seen by $\mathrm{sgn}\left(k_{t+1}^j\partial R_{t+1}/\partial k_{t+1} + n\tau_{t+1}\partial w_{t+1}/\partial k_{t+1}\right) = \mathrm{sgn}\left(-k_{t+1}^j/k_{t+1} + \tau_{t+1}\right)$.

[15] Gonzalez-Eiras and Niepelt (2004) show that the interest rate effect plays an important role in sustaining the social security system in the economy without within-cohort heterogeneity.

[16] In Section 5, we will study "the myopic voting equilibrium", where voters can rationally expect future policy outcomes but assume there to be no strategic interaction between the current and future policies.

[17] Recall that $U_t^{y,j} \equiv u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right)$ and $U_t^{o,j} \equiv u\left(c_t^{o,j}\right)$.

with $\Delta(\upsilon) \equiv (1-\alpha) + \upsilon(1-\alpha)^2$, $\Phi(\upsilon) \equiv -1 + 2\alpha + 2\upsilon\alpha(1-\alpha)$ *and* $\Theta(\upsilon) \equiv 2\upsilon\alpha - 1 + 2\sqrt{\upsilon\alpha(\upsilon\alpha - 1)}$. *The private saving function follows*

$$\tilde{S}^h\left(k_t^h, k_t^l\right) = \omega(\hat{\tau})\,\psi(\hat{\tau})\,A\left(1 - \mathbf{z}\left(k_t^h, k_t^l\right)\right)\left(\left(k_t^h + k_t^l\right)/(2n)\right)^\alpha, \quad (3.26)$$

$$\tilde{S}^l\left(k_t^h, k_t^l\right) = \psi(\hat{\tau})\,A\left(1 - \mathbf{z}\left(k_t^h, k_t^l\right)\right)\left(\left(k_t^h + k_t^l\right)/(2n)\right)^\alpha, \quad (3.27)$$

where $\hat{\tau}$ *is a constant solving*

$$\hat{\tau} = \mathbf{z}(\omega(\hat{\tau})) \quad (3.28)$$

Proof: See the appendix.

Four remarks about this proposition are in order. First, the political decision on the social security tax rate depends on wealth inequality. Moreover, it is easily seen that τ_t increases in k_t^h/k_t^l. That is to say, the larger the wealth inequality, the more political support the social security programme receives. Social security as an inter-generational redistribution policy has been widely studied in the literature. The within-cohort redistributive components of the pension system are often neglected, however. In the context of the present model, social security benefits the old poor more than the old rich. This is implied by (3.20), which shows that the welfare effect of τ_t, $\partial U_t^{o,j}/\partial\tau_t$, is heterogeneous and negatively related to the wealth holdings of the old. However, social security has no intra-generational redistributive effect on young households. In the appendix, we show that

$$\frac{\partial U_t^{y,j}}{\partial \tau_t} = -\frac{1+\beta\alpha}{1-\tau_t}, \quad (3.29)$$

i.e., τ_t has the same welfare effect on young households with different labour productivity. This homogenous effect is primarily due to the symmetric effect of τ_t on the private saving k_{t+1}^j, as discussed in the preceding section. It is also worth mentioning that due to the logarithmic specification, aggregate capital k_t is additively separable in the utility function. Hence, the decision of τ_t is independent of the k_t. This property does not hold under a more general utility form, which will be studied in Section 5.

Second, the conditions in Proposition 1 characterize the politico-economic environment where social security system can be sustained in the Markov equilibrium. For $\upsilon\alpha < 1$ to hold, a small n or α is needed. A small n implies a large share of old in the population and hence, a larger number of agents directly benefiting from the pension system. A low α implies that the interest rate R_{t+1} is rather elastic to aggregate capital k_{t+1}. This amplifies the general equilibrium effect and mitigates the negative welfare effect of τ_t on young households, which can be seen directly from (3.29). So a small n and α reinforce the political constituency of the social security system. When $\upsilon\alpha \geq 1$, the intra-generational redistribution becomes the key. There would be no social security system in an economy without within-cohort heterogeneity. However, a social security system can be sustained as long as there exists a sufficiently high level of wealth inequality among old households. Therefore, when $\upsilon\alpha \geq 1$, political support for a social security system largely comes from old households with low wealth which requires positive social security transfers.

Third, (3.26) and (3.27) imply that τ_t does not affect future wealth inequality in the Markov equilibrium. Once more, this is due to the symmetric effect of τ_t on the private saving under logarithm utility. Since the social security tax rate is determined by wealth inequality as shown in (3.24), the strategic effect under logarithm utility is mute, i.e., $\partial\tau_{t+1}/\partial\tau_t = 0$.[18] That is to say, although current voters can in principle influence future political outcomes via affecting future wealth inequality, they are actually unable to do this in the competitive equilibrium. The lack of any strategic effect is due to the fact that future wealth inequality is independent of the current social security tax rate, as stated in Lemma 1. This independence breaks down the dynamic link between τ_t and τ_{t+1} in the Markov equilibrium. As will be seen in Section 5, the strategic effect arises under a more general utility case, where the choice of τ_t may affect future wealth inequality and thus, future policy outcomes.

Finally, $\hat{\tau}$ which satisfies (3.28) is the rational expectation of the future tax rate. Given expectation $\hat{\tau}$, agents make intertemporal choices so that the future wealth inequality will be equal to $\omega(\hat{\tau})$. For expectation $\hat{\tau}$ to be self-fulfilled, it must equal that implied by the policy rule, i.e., $\hat{\tau} = \mathbf{z}(\omega(\hat{\tau}))$. Due to the rather complicated expression of $\mathbf{z}(\omega(\hat{\tau}))$, we are unable to characterize analytically the solution of (3.28). Extensive numerical experiments show that the self-fulfilled expectation $\hat{\tau}$ is unique. Note that the formation of the rational expectation on

[18] (3.24) implies $\mathbf{z}_1\left(k_t^h, k_t^l\right)/\mathbf{z}_2\left(k_t^h, k_t^l\right) = -k_{t+1}^l/k_{t+1}^h$. Moreover, (3.26) and (3.27) give $\hat{S}_2^h(k_t,\tau_t)/\hat{S}_2^l(k_t,\tau_t) = k_{t+1}^h/k_{t+1}^l$. Plugging these two results into (3.22) establishes $\partial\tau_{t+1}/\partial\tau_t = 0$.

the future tax rate holds for any time $t \geq 1$. Hence, all future tax rates are determined by equation (3.28) and independent of k_0^h/k_0^l and τ_0. Since the future tax rate is a constant and does not change over time, it is immediate that $\tau_t = \hat{\tau}$ and $k_t^h/k_t^l = \omega\left((\hat{\tau})\right)$ for $t \geq 1$. The constant wealth inequality after the initial period is due to the mute effect of the current tax rate on future wealth inequality and policy outcomes, as discussed above. It is worthy emphasizing that τ_t for $t \geq 1$ does follow policy rule $\mathbf{z}$ in the Markov equilibrium. The constant tax rate is due to the fact that wealth inequality becomes a constant $\omega\left(\hat{\tau}\right)$ after the initial period.

Now we can characterize the dynamics of wealth inequality and social security. Suppose that the voting for social security is unanticipatedly launched at time 0. This means that $\tau_t = 0$ for $t < 0$ and agents born at time $t < 0$ anticipate zero tax rates in future. (3.9) implies that the initial wealth inequality k_0^h/k_0^l equals income inequality θ, which gives $\tau_0 = \mathbf{z}\left(\theta\right)$ by policy rule $\mathbf{z}$. In the periods after the initial one, wealth inequality and tax rate are equal to $\omega\left(\hat{\tau}\right)$ and $\hat{\tau}$, respectively, as shown above. Therefore, k_t^h/k_t^l and τ_t converge to the steady state in two periods. Moreover, by $\omega\left(\hat{\tau}\right) \geq \omega\left(0\right) = \theta$, the positive tax rate increases future wealth inequality. This leads to a growing size of social security.[19] To conclude, we have

Corollary 1 *Assume that* $u\left(c\right) = \log\left(c\right)$. *Then in the Markov perfect equilibrium,*

(i) Wealth inequality and the social security tax rate converge to the steady state in two periods.

(ii) The subsequent wealth inequality and social security tax rates are higher than the initial one.

Note that the dynamic social security is not decided by the government with a commitment technology. Instead, the system is repeatedly determined by the political fundamental, i.e., wealth inequality. Forward-looking households, rationally perceiving the link between wealth inequality and social security, would adjust their private savings accordingly. In particular, Corollary 1 shows that

[19] Formally, the tax rate τ_t at any time $t \geq 1$ is equal to $\mathbf{z}\left(\omega\left(\hat{\tau}\right)\right)$, which is greater than the initial tax rate $\tau_0 = \mathbf{z}\left(\omega\left(0\right)\right)$.

this interaction leads to a growing size of social security in the dynamic politico-economic equilibrium. The underlying mechanism is twofold. On the one hand, the establishment of a social security system increases future wealth inequality since within-cohort transfers discourage the private savings of the poor more than the rich. On the other hand, the larger wealth inequality makes social security transfers more desirable in the future. This provides the political support for an increasing size of social security in the following periods.

3.1 Empirical Implications

This subsection investigates the empirical predictions of the model. Although the two-period OG model is very stylized, we would like to see if its predictions are broadly consistent with the facts. First, we assess the quantitative importance of the dynamic interaction between k_t^h/k_t^l and τ_t, and then compare the evolution of the size of social security with data from the OECD countries. The parameter values are set as follows. $\alpha = 0.36$, as widely adopted in the literature of macroeconomics (e.g. Prescott, 1986). Each period in the OG model is assumed to contain 30 years. The annual discounting rate equals 0.98. Then, $\beta = 0.98^{30}$ and $n = 1.384$, the latter corresponding to the gross growth rate of the U.S. population between 1970 and 2000 (Gonzalez-Eiras and Niepelt, 2004). A is calibrated such that the steady state k without a social security system is equal to unity. Set $\gamma^l = 0.5$, i.e., the labour productivity of the rich is three times higher than that of the poor.[20] Suppose that voting for a social security system is unanticipated launched at time 0. The initial wealth inequality k_t^h/k_t^l equals the income inequality $\theta = 3$. Let $k_0 = 1/2$, i.e., the economy is half-way from the steady state before the establishment of a social security system. By Proposition 1, k_0 has no effect on k_t^h/k_t^l and τ_t. Table 1 provides details on the evolution of wealth inequality, social security tax rates, factor prices and the consumption of different groups of households.

Table 1: Politico-Economic Equilibrium under Logarithm Utility

[20] A different γ^l only changes the size of the social security system, but has little impact on its evolution, which is determined by the change in wealth inequality and not by its level.

	$t=0$	$t=1$	$t=\infty$
k^h/k^l	3.000	3.637	3.637
τ	0.100	0.119	0.119
R	3.437	3.006	2.849
w	2.208	2.380	2.454
$c^{o,h}$	3.634	4.267	4.453
$c^{o,l}$	1.416	1.459	1.519
$c^{y,h}$	1.991	2.101	2.165
$c^{y,l}$	0.706	0.745	0.767

Table 1 shows that after the introduction of a social security system, wealth inequality k_t^h/k_t^l increases from 3 to 3.64 in one period and remains at that level.[21] By policy rule $\mathbf{z}$, $\tau_0 = 10\%$ and $\tau_t = 11.9\%$ for $t \geq 1$. The increase in τ_t driven by the endogenous change in k_t^h/k_t^l is quantitatively significant. The social security tax rate increases by nearly 20% within one period, which corresponds to 30 years in the present model. The significant increase in the size of the social security programme is broadly consistent with the evolution of the pension system in the OECD countries. Most OECD countries established sophisticated social security systems after World-War II. As reported in Table 2, social security transfers grew steadily during the post-war period. The picture remains almost the same if the sample is separated into two sub-groups: one with small countries and the other with G-7. From 1960 to 1985, social security transfers as a percentage of GDP nearly doubled among the OECD countries. It may be doubted whether the increase in the size of social security is caused by the aging population. According to Breyer and Craig (1997), the average public pension benefits per pensioner in the OECD countries tripled from 1960 to 1980, rising from \$1546 in 1960 to \$4653 in 1980, expressed in 1982 US dollars. The increase in social security benefits per retiree remains sizable after excluding the effect of economic growth.[22]

Table 2: Social Security Transfers as Percentage of GDP

[21] The increase in wealth inequality by 21% seems close to the result in a recent quantitative study (Fuster *et al.*, 2003). Fuster shows that introducing social security increases the Gini coefficient of the distribution of assets by 27%, from 0.51 to 0.65.

[22] The average annual and gross economic growth rates during 1960 and 1980 in the OECD countries are equal to 2.96% and 1.79, respectively.

	1960	1968	1970	1974	1980	1985
All OECD Countries	7.87 (3.69)	9.94 (4.05)	9.85 (3.71)	11.31 (3.97)	14.24 (5.47)	15.60 (5.26)
Small Countries	7.63 (3.84)	9.90 (3.98)	9.95 (3.83)	11.35 (4.32)	14.71 (6.21)	15.97 (5.83)
G-7	8.40 (3.57)	10.03 (4.50)	9.66 (3.72)	11.21 (3.42)	13.21 (3.57)	14.79 (4.05)

Notes: Standard errors are in parentheses.

Incorporating within-cohort heterogeneity, our model also provides a number of empirical predictions on the within-cohort consumption inequality over the life cycle. Though social security enlarges wealth inequality (excluding social security transfers), it does narrow the within-cohort consumption inequality. Without social security system, consumption inequality is equal to income inequality, i.e., $c_t^{o,h}/c_t^{o,l} = c_t^{y,h}/c_t^{y,l} = 3 \ \forall t < 0$. After the establishment of a social security system, there is a significant decline in the inequality in the consumption of young households, with $c_0^{y,r}/c_0^{y,p} = 2.820$, $c_1^{y,r}/c_1^{y,p} = 2.820$ and $c_\infty^{y,r}/c_\infty^{y,p} = 2.823$. Increasing future social security benefits reduces the private savings of the poor more than those of the rich, as indicated by Lemma 1. That is to say, the poor would increase their consumption more than the rich in the first period of their life. On the other hand, the inequality of the old households' consumption $c_t^{o,r}/c_t^{o,p}$ drops sharply to 2.566 at time 0, then rebounds by $c_1^{o,r}/c_1^{o,p} = 2.925$ and converges to 2.932. In the initial period, old households receive social security transfers without paying any contribution. Thus, the social security system can sharply reduce the consumption inequality of old households at $t = 0$. In subsequent periods, the net social security transfers are much smaller than the initial period, according to the constant τ_t for $t \geq 1$. The effect of social security on long-run consumption inequality $c_t^{o,r}/c_t^{o,p}$ becomes rather weak. Note that there are two opposite effects of social security transfers on the consumption inequality of the old. First, lump-sum social security transfers directly reduce the consumption inequality of the old. Second, the asymmetric effect of future transfers enlarges the wealth inequality of old households in the future. This gives the old less to consume and thus indirectly increases consumption inequality. The positive indirect effect turns out to be dominated by the negative direct effect, since the introduction of social security reduces the consumption inequality of the old, as shown in Table 1.

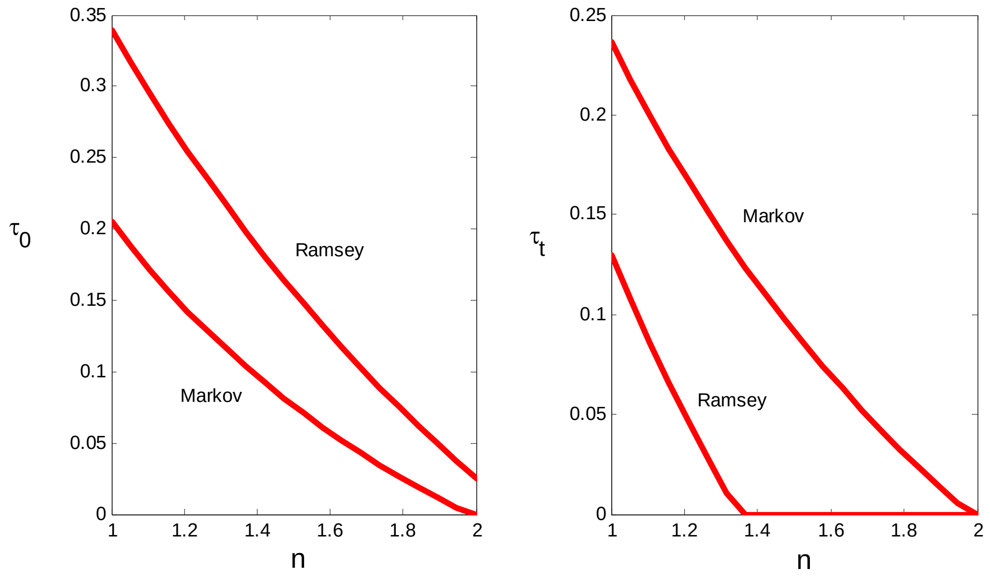

Figure 1: Comparative Statics of the Social Security Tax Rate with respect to Population Growth under Logarithm Utility.

The above results suggest that our model predicts a positive correlation between age and within-cohort consumption inequality, except in the initial period. This is consistent with empirical evidence (e.g. Deaton and Paxson, 1994). In the literature, the increasing consumption inequality with age are often explained by uninsurable earning shocks (e.g., Deaton and Paxson, 1994, Storesletten, Telmer and Yaron, 2004). The present paper, however, emphasizes the role of the social security system. To see this, first note that consumption inequality is a constant over the life cycle, when there is no social security system (recall $c_t^{o,h}/c_t^{o,l} = c_t^{y,h}/c_t^{y,l} = 3 \; \forall t < 0$). Hence, incorporating within-cohort heterogeneity *per se* does not produce the increasing consumption inequality over the life cycle. As discussed in the preceding paragraph, social security tends to narrow the consumption inequality of both the young and the old. However, the effect on the young turns out to be significantly larger than on the old, since the direct effect of the lump-sum transfers on $c_t^{o,h}/c_t^{o,l}$ is largely cancelled out by the opposite indirect effect via increased wealth inequality. The asymmetric effect of social security on $c_t^{y,h}/c_t^{y,l}$ and $c_t^{o,h}/c_t^{o,l}$ yields the increasing consumption inequality over the life cycle.

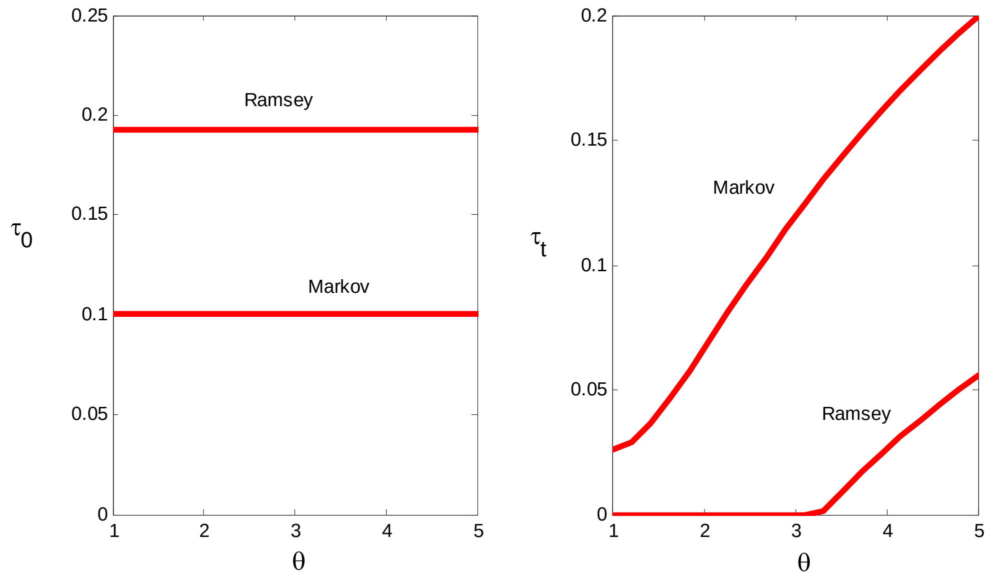

Figure 2: Comparative Statics of the Social Security Tax Rate with respect to Income Inequality under Logarithm Utility.

Our model also provides predictions on the impact of demographic structure and income inequality on the size of the social security system. The impact on τ_0 and τ_t for $t \geq 1$ is referred to as the transitory and the permanent effect, respectively. Figure 1 plots τ_0 and τ_t for $t \geq 1$ with respect to n. A lower n leads to a larger size of old households. This raises the weight on the welfare of the old in the political decision process, and thus increases social security transfers. The positive relationship between the size of social security and the old-aged population is consistent with the cross-country data (e.g. Tabellini, 2000, Mulligan and Sala-i-Martin, 1999a).

Figure 2 plots τ_0 and τ_t for $t \geq 1$ with respect to θ. Larger income inequality leads to larger wealth inequality and makes within-cohort redistribution more desirable in the future. The permanent effect produces a positive relationship between income inequality and the size of the social security system. The empirical studies based on the cross-country data give somewhat ambiguous results, however. Tabellini (2000) shows that the sizes of social security transfers are positively correlated with income inequality. But a puzzling negative correlation can also be found in the literature (e.g., Lindert, 1996 and Rodriguez, 1998). The time-series evidence is also mixed. Early tests provide supports for a positive correlation (Meltzer and Richard, 1983). But a recent study (Rodriguez, 1999) shows that there is no significant relationship between income inequality and the size of social security transfers. These ambiguous findings do not necessarily contradict our model, however. As can be seen from Figure 2, income inequality θ in fact has no transitory effect on the size of social security, which is solely determined by wealth inequality according to Proposition 1. Therefore, our model suggests that it be more appropriate to test the time-series relationship between wealth inequality and social security, instead of that between income inequality and social security.

4 Ramsey Solution

It is instructive to compare the outcomes in the political equilibrium with the Ramsey solution. To this end, we characterize the efficient allocation, where a benevolent planner with a commitment technology sets the sequence of tax rates $\{\tau_t\}_{t=0}^{\infty}$ so as to maximize the sum of the discounted utilities of all generations. The planner's constraint is that the chosen policy should be implementable as a competitive equilibrium. The corresponding programme of the Ramsey problem is

$$\max_{\{\tau_t\}_{t=0}^{\infty}} \sum_{j=h,l} u\left(c_0^{o,j}\right) + \sum_{t=0}^{\infty} \rho^{t+1} \left(\sum_{j=h,l} \left(u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right)\right)\right), \tag{3.30}$$

subject to individuals' budget constraints (3.2) and (3.3), factor prices (3.4) and (3.5), the balanced-budget rule (3.6), private saving functions (3.7) and (3.8), and the law of motion of aggregate capital (3.11). $\rho \in (0,1)$ is the intergenerational discount factor. We assume that $\rho \equiv \beta n$, i.e., the planner weighs generations by their sizes and discounts their welfare by households' discount factor. Compared with the political decision problem (3.13), the efficient allocation problem (3.30) has two distinctive features. First, the Ramsey planner cares about the welfare of all future generations, and second, she has the ability to commit to future policies.

For notational convenience, $I_{t,t+i} \equiv \partial k_{t+i}/\partial \tau_t$ is denoted as the impact of τ_t on the future capital stock k_{t+i} for $i \geq 1$, as implied by the law of motion of capital (3.11):

$$I_{t,t+i} = \begin{cases} \frac{\partial k_{t+i}}{\partial k_{t+i-1}} \frac{\partial k_{t+i-1}}{\partial k_{t+i-2}} \cdots \left(\frac{\partial k_{t+1}}{\partial \tau_t} + \frac{\partial k_{t+1}}{\partial k_t} \frac{\partial k_t}{\partial \tau_t} \right) & \text{for } t \geq 1 \\ \frac{\partial k_i}{\partial k_{i-1}} \frac{\partial k_{i-1}}{\partial k_{i-2}} \cdots \frac{\partial k_1}{\partial \tau_0} & \text{for } t = 0 \end{cases} . \tag{3.31}$$

The second line of (3.31) is due to the fact that k_0 is predetermined. τ_t also affects the capital stock at time t, since τ_t may influence the private savings in the preceding period. Its impact, denoted by $I_{t,t}$, is equal to

$$I_{t,t} = \begin{cases} \frac{\partial k_t}{\partial \tau_t} & \text{for } t \geq 1 \\ 0 & \text{for } t = 0 \end{cases} . \tag{3.32}$$

$I_{0,0} = 0$ since k_0 is predetermined. Note that τ_t directly influences the welfare of the agents born at time t and $t-1$ by affecting their after-tax net earnings and social security benefits, respectively. In addition, τ_t indirectly influences the welfare of agents born at time t and afterwards via its impact on capital accumulation $I_{t,t+i}$. τ_t has no effect on the agents born before time $t-1$.

Following the same procedure as in the preceding section, let us look at the impact of the social security tax rate τ_t on the welfare of various groups of households. Due to the envelope argument based on the Euler equation, the welfare effect of τ_t on agents born at time $t-1$, denoted by $\partial U_{t-1}^{y,j}/\partial \tau_t$, parallels its effect on old households at time t, denoted by $\partial U_t^{o,j}/\partial \tau_t$. Specifically,

$$\frac{\partial U_{t-1}^{y,j}}{\partial \tau_t} = \beta \frac{\partial U_t^{o,j}}{\partial \tau_t} = \beta \left(u'\left(c_t^{o,j}\right) n w_t + u'\left(c_t^{o,j}\right) \left(k_t^j \frac{\partial R_t}{\partial k_t} + n \tau_t \frac{\partial w_t}{\partial k_t} \right) I_{t,t} \right), \tag{3.33}$$

where $I_{t,t}$ follows (3.32). The first term on the RHS of (3.33) reflects the direct effect of τ_t, which increases social security transfers and thus benefits old households at time t. The second term captures the general equilibrium effect of τ_t through I_t. Compare (3.33) with (3.19), we see that the general equilibrium effect is absent in the political decision process, where voters take k_t as given. In the Ramsey problem, the planner has the abilities to commit to future policies. Thus, she must take into account the impact of τ_t on k_t, for $t \geq 1$. As shown in the preceding section, the general equilibrium effect is twofold. The negative $I_{t,t}$ reduces k_t and thus, the social security tax base. But a low k_t increases the interest rate. The interest rate effect dominates if τ_t or wealth inequality is not too large. The positive overall general equilibrium effect implies that the marginal benefit of τ_t to old households at time t in the Ramsey problem tends to be larger than its counterpart in the political decision process. It is worth mentioning that for $t = 0$, the welfare effect of τ_0 on old households at time 0 equals its counterpart in the political decision as shown in (3.19):

$$\frac{\partial U_0^{o,j}}{\partial \tau_0} = u'\left(c_0^{o,j}\right) n w_0, \tag{3.34}$$

since the capital in the initial period is predetermined ($I_{0,0} = 0$). The following Lemma shows that, as in the political decision process, a high social security tax rate τ_t always benefits old households at time t (or agents born at time $t-1$).

Lemma 2 *Assume that $u(c) = \log(c)$. In the Ramsey problem, the welfare effect of τ_t on old households at time t equals*

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = \begin{cases} \frac{(1+\beta)(\phi(\tau_t)+\tau_t\phi'(\tau_t))}{\alpha\gamma^j+\tau_t\phi(\tau_t)} - \frac{\beta(1-\alpha)\phi'(\tau_t)}{\phi(\tau_t)} > 0 & \textit{for } t \geq 1 \\ \frac{1-\alpha}{2\alpha\frac{k_0^j/k_0^i}{1+k_0^j/k_0^i}+(1-\alpha)\tau_0} > 0 & \textit{for } t = 0 \end{cases}. \tag{3.35}$$

The first line of (3.35) is proved in the appendix and the second line simply follows (3.20). To conclude, the welfare effect $\partial U_t^{o,j}/\partial \tau_t$ in the Ramsey problem tends to be larger than in the political decision process via the general equilibrium effect, except that they are equivalent for $t = 0$, when the general equilibrium effect is absent in both cases.

The social security tax rate τ_t also affects the welfare of all generations born at time t and afterwards. The welfare effect of τ_t on young households at time $t+i$ equals

$$\begin{aligned} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} &= -u'\left(c_{t+i}^{y,j}\right)\gamma^j w_{t+i} + u'\left(c_t^{y,j}\right)\gamma^j \frac{\partial w_{t+i}}{\partial k_{t+i}} I_{t,t+i} \\ &\quad + \beta u'\left(c_{t+i+1}^{o,j}\right)\left(k_{t+i+1}^j \frac{\partial R_{t+i+1}}{\partial k_{t+i+1}} + n\tau_{t+i+1}\frac{\partial w_{t+i+1}}{\partial k_{t+i+1}}\right) I_{t,t+i+1}. \end{aligned} \tag{3.36}$$

As in (3.21), the first term in (3.36) reflects the direct cost of social security taxes for young households. The second and third terms are the general equilibrium effects via I_{t+i} and I_{t+i+1}. Note that for $i \geq 1$, the welfare effect $\partial U_{t+i}^{y,j}/\partial \tau_t$ does not enter the political decision on τ_t, since the welfare of future generations is ignored in electoral competition. For $i = 0$, a comparison between (3.36) and (3.21) reveals that $\partial U_t^{y,j}/\partial \tau_t$ in the Ramsey problem differs from its counterpart in the political equilibrium in two respects. First, the planner takes into account the negative impact of τ_t on k_{t+i}, which reduces the social security tax base at time $t+i$. This negative general equilibrium effect is captured by the second term on the RHS of (3.36). In the political equilibrium, voters at time t would take k_t as given and hence, ignore the negative impact. Second, there is no strategic effect in the Ramsey problem, since the planner can commit to future policies. However, we have shown that the strategic effect is mute under logarithm utility. Therefore, the welfare loss of τ_t to the current young households in the Ramsey problem is greater than the political equilibrium, due to the negative general equilibrium effect.

Lemma 3 *Assume that $u(c) = \log(c)$. In the Ramsey problem, the welfare effect of τ_t on young households at time $t+i$ is equal to*

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \begin{cases} -\frac{(1+\beta\alpha)\alpha^i}{1-\tau_t} + \frac{(1+\beta\alpha)\alpha^{i+1}\phi'(\tau_t)}{\phi(\tau_t)} < 0 & \text{if } t \geq 1 \\ -\frac{(1+\beta\alpha)\alpha^i}{1-\tau_t} < 0 & \text{if } t = 0 \end{cases}. \tag{3.37}$$

The proof is given in the appendix. Four remarks are in order. First, $\partial U_{t+i}^{y,j}/\partial \tau_t < 0$ shows that τ_t incurs a net welfare loss to all generations born at time t and afterwards. Second, the welfare effect only depends on τ_t, due to the additive separability implied by logarithm utility. The irrelevance of future capital stocks and tax rates remarkably simplifies the characterization of the Ramsey allocation. Third, τ_t has the same effect on the welfare of the poor and the rich, due to the symmetric effect of τ_t on private savings k_{t+1}^j, as discussed in Section 2. Finally, from (3.37) and (3.29), it can directly be seen that for $t \geq 1$, the marginal welfare loss of τ_t to the current young households in the Ramsey problem is greater than the political equilibrium due to the negative general equilibrium effect caused by $I_{t,t}$, which reduces the social security tax base. In the initial period ($t = 0$), these are equivalent since $I_{0,0} = 0$.

Now, the first-order conditions of (3.30) with respect to τ_t can be written as:

$$\sum_{j=h,l} \frac{\partial U_t^{o,j}}{\partial \tau_t} + \sum_{i=0}^{\infty} \left(\rho^{i+1} \sum_{j=h,l} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} \right) + \lambda_t = 0. \tag{3.38}$$

Plug (3.34) and (3.37) into (3.38), we can solve τ_0.

Proposition 2 *Assume that $u(c) = \log(c)$. In the Ramsey solution,* *(i) the initial social security tax rate*

$$\tau_0 = \begin{cases} H\left(k_0^h/k_0^l\right) > 0 & \text{if } \upsilon\alpha < 1 \\ 0 & \text{if } \upsilon\alpha \geq 1 \text{ and } k_t^h/k_t^l > \Theta(\upsilon) \end{cases}, \tag{3.39}$$

where $H(\cdot)$ follows (3.25) with $\upsilon \equiv \rho(1+\alpha\beta)/((1-\rho\alpha)(1-\alpha))$.

(ii) $\tau_0^R \gtreqless \tau_0^M$ if and only if $\rho \lesseqgtr n/(1+\alpha n)$, where τ_0^R and τ_0^M denote the initial tax rate in the Ramsey solution and the Markov political equilibrium, respectively.

Proof is given in the appendix. The first part of Proposition 2 states that the initial tax rate τ_0 is determined by the initial wealth inequality k_0^h/k_0^l, which parallels Proposition 1 in the political equilibrium. A high k_0^h/k_0^l leads to a high τ_0, due to the within-cohort redistributive effects of τ_0. The second part of the proposition compares the initial tax rate in the Ramsey solution with its counterpart in the political equilibrium. There are two opposite effects which drive the political outcome τ_0^M to deviate from the efficient allocation τ_0^R. To see this, we rewrite the first-order condition of τ_0 (3.38) as

$$\sum_{j=h,l} \frac{\partial U_0^{o,j}}{\partial \tau_0} + n \sum_{j=h,l} \frac{\partial U_0^{y,j}}{\partial \tau_0} + \lambda_t + \sum_{i=1}^{\infty} \left(\rho^i \sum_{j=h,l} \frac{\partial U_i^{y,j}}{\partial \tau_0} \right) - (n-\rho) \sum_{j=h,l} \frac{\partial U_0^{y,j}}{\partial \tau_0} = 0. \tag{3.40}$$

The first three terms one the LHS of (3.40) coincide with the LHS of (3.23), i.e. the first-order condition in the political decision process.[23] The fourth term reflects the negative impact of τ_0 on the welfare of households born after the initial period via capital accumulation (see Lemma 3). This negative impact is ignored in the political decision process where voters are short-lived and do not care about future generations. This is referred to as political effect I, which makes τ_0^M higher than τ_0^R. The fifth term on the LHS of (3.40), referred to as political effect II, illustrates the discrepancy between the weight on the current young households in the political decision process and the Ramsey problem. Recall that $\rho \equiv \beta n < n$. The Ramsey planner would like to impose a higher τ_0 since the weight on the old is higher than in the political decision process. For $\rho > n/(1+\alpha n)$, the second effect dominates the first, and vice versa. Intuitively, a high α incurs a large

[23] Note that for $t = 0$, $\partial U_0^{o,j}/\partial \tau_0$ is the same in both of the Markov political equilibrium and the Ramsey problem, so as $\partial U_0^{y,j}/\partial \tau_0$ (see Lemma 2 and 3).

welfare loss via capital accumulation and hence, amplifies the first effect, while a high n or a low ρ weakens the second effect. As will be shown below, political effect II dominates political effect I under reasonable parameter values.

Now we proceed to τ_t for $t \geq 1$. Plug (3.35) and (3.37) into (3.38), one can find that τ_t is a constant over time. In the appendix, we prove the following proposition.

Proposition 3 *Assume that* $u(c) = \log(c)$. *In the Ramsey solution,*

(i) The social security tax rate converges to steady state in two periods.

(ii) The steady state tax rate $\bar{\tau}$ *is unique. Denote*

$$\Omega \equiv \frac{(1-\alpha)(1+\theta)^2}{2\alpha\theta} - \frac{2n\left(1+\alpha+\beta\alpha-\alpha^2\right)}{1-\beta n\alpha} + \frac{(1-\alpha)^2}{1+\alpha\beta}. \tag{3.41}$$

If $\Omega > 0$, $\bar{\tau}$ *solves the equation* $L(\bar{\tau}) = 0$ *(see the appendix for the definition of* $L(\cdot)$*).* $\bar{\tau} = 0$ *otherwise.*

The first part of the proposition parallels Corollary 1 in the political equilibrium. The distinctive feature is that the effect of the initial wealth inequality on the Ramsey social security tax rates only lasts one period. This should not be surprising. As shown in Lemma 1, future wealth inequality solely depends on the future tax rate. Thus, the planner can choose an optimal level of future wealth inequality via the future tax rate, irrespective of the initial state. (3.41) gives the condition that the social security system can be sustained or not in the Ramsey allocation. It is immediate that Ω increases in θ but decreases in n. Intuitively, a high income inequality θ increases the within-cohort redistributive benefit of social security transfers. A high population growth rate n increases the weigh on the welfare of future generations in the Ramsey problem, and thus makes social security transfers more desirable. Figure 3 plots the threshold condition of θ implied by (3.41) under $\alpha = 0.36$ and $\beta = 0.98^{30}$. $\Omega > 0$ is satisfied for any θ above the line in the figure. It can directly be seen that a high n requires a high θ to sustain the social security system in the Ramsey allocation.

Table 3: Ramsey Solution under Logarithm Utility

	$t=0$	$t=1$	$t=\infty$
k^h/k^l	3.000 (3.000)	3.000 (3.637)	3.000 (3.637)
τ	0.193 (0.100)	0.000 (0.119)	0.000 (0.119)
R	3.437 (3.437)	2.968 (3.006)	2.206 (2.849)
w	2.208 (2.208)	2.398 (2.380)	2.833 (2.454)
$c^{o,h}$	3.917 (3.634)	3.743 (4.267)	4.808 (4.453)
$c^{o,l}$	1.699 (1.416)	1.248 (1.459)	1.603 (1.519)
$c^{y,h}$	1.729 (1.991)	2.327 (2.101)	2.750 (2.165)
$c^{y,l}$	0.576 (0.706)	0.776 (0.745)	0.917 (0.767)

Note: The politico-economic equilibrium outcomes are in parentheses.

The values in Table 3 provide numerical examples on the Ramsey social security and its implications. The benchmark parameter values give $\rho = 0.76 < n/(1+\alpha n) = 0.92$. By the second part of Proposition 2, we have that $\tau_0^R = 19.3\% < \tau_0^M = 10.0\%$, i.e. the initial Ramsey social security tax rate is substantially higher than that in the political equilibrium. This suggests that the political decision process gives too much weight on the welfare of young households and thus leaves insufficient transfers to the old households in the initial period. Compared to the political equilibrium, the Ramsey allocation reduces the initial consumption of the old rich and poor by 7.16% and 19.72%, respectively, while it increases the initial consumption of the young rich and poor by 15.03% and 22.41%, respectively.

In terms of the long-run tax rate, it can easily be verified that $\Omega < 0$ under the benchmark parameter values. Thus, the second part of Proposition 3 implies that $\tau_t^R = 0$ for $t \geq 1$. That is to say, the efficient allocation does not support social security in the long run, even if it plays a role of within-cohort redistribution. This is because the zero tax rate substantially encourages capital accumulation. The gain from more capital in the future outweighs the benefits from inter- and intra-generational redistribution. Specifically, the move from the political equilibrium to the Ramsey solution increases the steady state consumption of the old rich and poor by 7.87% and 5.26%, respectively. The increase in the steady state consumption of the young rich and poor amounts to 26.73% and 19.48%, respectively.

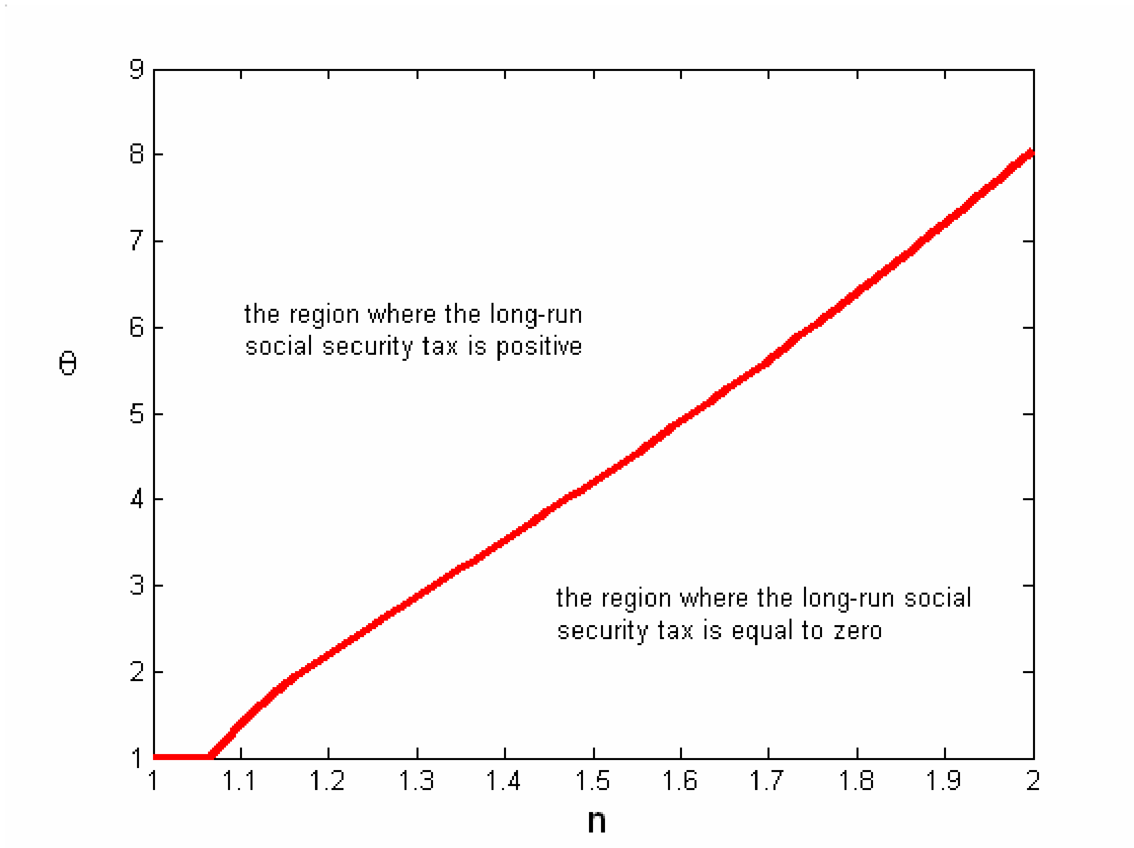

Figure 3: The Threshold Condition of θ for the Long-Run Social Security Tax Rate to be Positive.

The zero social security transfers in the Ramsey solution are somewhat surprising. If there is no within-cohort heterogeneity, the first best allocation coincides with the Ramsey solution (see Gonzalez-Eiras and Niepelt, 2004). If within-cohort heterogeneity exists, it is straightforward that the social planner would like to eliminate within-cohort consumption inequality. The first best outcome cannot be implemented as a competitive equilibrium, however, since it implies 100% tax rate, which leads to zero capital stock. Moreover, as discussed above, forward-looking households would adjust their intertemporal choices according to future social security transfers. The higher social security benefits in the future will lead to the larger wealth inequality, which considerably offsets the within-cohort redistributive effects of social security.

5 The Strategic Effect under CRRA Utility

So far, we have focused on logarithm utility. The social security tax rate being independent of the aggregate capital stock under logarithm utility substantially simplifies the analysis. However, many empirical studies suggest the elasticity of intertemporal substitution to be less than unity. It is an open question to what extent our results would be affected by the deviation from logarithm utility. Particularly, the strategic effect $\partial\tau_{t+1}/\partial\tau_t$ in the Markovian political equilibrium may arise under a less restrictive utility form. This section adopts the more general CRRA utility function, to see whether the analytical results in the preceding sections are robust with the presence of strategic effect. Specifically, we assume

$$u(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma} \tag{3.42}$$

where $\sigma > 0$ is the inverse of intertemporal substitution elasticity. Household intertemporal choices, the political decision as well as the Ramsey problem on social security tax rates are characterized in the appendix. Analytical solutions cannot be obtained for $\sigma \neq 1$ so we resort to numerical methods.

The computational strategy for the Markov perfect equilibrium adopts a standard projection method with Chebyshev collocation (Judd, 1992). The basic idea of the projection method is to approximate some unknown functions on a basis of functional space. This method turns out to be efficient for time-consistent problems in some recent research (Judd, 2003 and Ortigueira, 2004). As for the Ramsey solution, we transform the infinite-horizon problem into a finite-horizon problem by the truncated method (e.g. Jones, Manuelli and Ross, 1993). The corresponding algorithms are provided in the appendix.

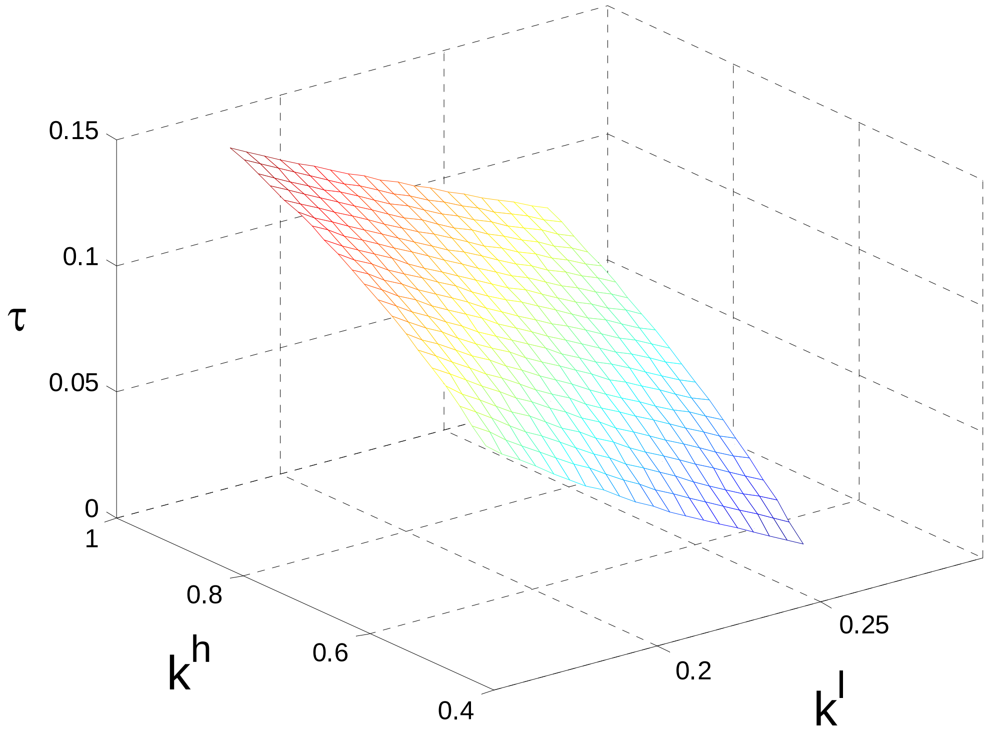

Figure 4: The Markovian Equilibrium Policy Rule under $\sigma = 3$.

Let us first look at the Markovian political equilibrium. The equilibrium policy rule $\mathbf{z}$ under $\sigma = 3$ is plotted in Figure 4. It can directly be seen that $\partial\tau_t/\partial k_t^h > 0$ and $\partial\tau_t/\partial k_t^l < 0$, which imply that social security transfers are increasing in wealth inequality. Figure 5 plots the evolution of the social security tax rate in the Markovian political equilibrium with different initial capital stocks. The sizes of social security converge to the steady state along an increasing path if the initial aggregate capital k_0 is sufficiently large. If k_0 is sufficiently low, social security tax rates may evolve along an inverted-U curve. Recall that under logarithm utility, the size of social security is independent of k_0 (Proposition 1). However, for $\sigma \neq 1$, k_0 plays a role in the choice of taxes. Specifically, the cost of τ_0 is negatively related to k_0. With a low k_0, the policymaker would like to impose a low τ_0 to encourage capital accumulation, which benefits the current young households via a larger tax base in the next period.

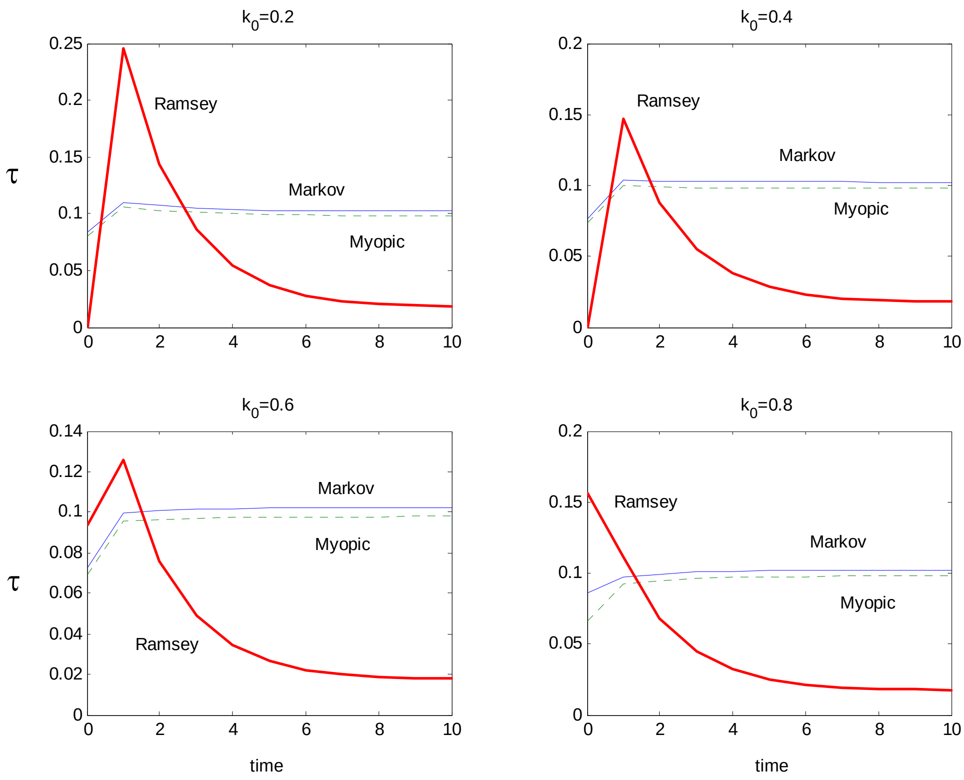

Figure 5: The Evolution of the Social Security System under $\sigma = 3$.

The strategic effect (4.21) arises for $\sigma \neq 1$. To identify the strategic effect, it is useful to study the myopic voting equilibrium where voters at time t have rational expectation on τ_{t+1} but (incorrectly) disregard the strategic effect of the current political decision τ_t on τ_{t+1}. The formal definition is given as follows.

Definition 2 *A myopic voting political equilibrium is a doublet of private saving functions* $S^j : R^+ \times [0,1] \times [0,1] \to R^+$, $j \in \{h, l\}$, *and a sequence of social security tax rates* $\{\tau_t\}_{t=0}^{\infty}$ *such that*

(1) S^j *is solved by the Euler equation.*

(2) $\{\tau_t\}_{t=0}^{\infty}$ *is solved by (3.13), subject to budget constraints (3.2), (3.3) and private saving functions* S^j.

Recall that in the log case, the strategic effect is mute and thus the myopic equilibrium and the Markov equilibrium give the same solution. The dashed line in Figure 5 plots the social security tax rates in the myopic voting equilibrium, which turns out to be slightly lower than in the Markov perfect equilibrium. The discrepancy has its roots in the strategic effect. Under log utility, the private savings of the poor and those of the rich decrease in τ_t by the same proportions. Thus, the current social security tax rate does not affect future wealth inequality. For $\sigma > 1$, the incentive for income-smoothing changes asymmetrically between the rich and poor. The asymmetric effects enlarge future wealth inequality k^h_{t+1}/k^l_{t+1} and thus raise the future social security tax rate τ_{t+1} via the equilibrium policy rule $\mathbf{Z}$. This gives rise to a positive strategic effect of τ_t on τ_{t+1}. Hence, the current young households would like to strategically vote for a higher τ_t, since it incurs higher future social security benefits. The strategic effect is quantitatively less important, however. The relative increase in the social security tax rate due to the strategic effect is less than 5%.

Finally, we turn to the Ramsey solution. Figure 5 plots the Ramsey tax rates over time. If k_0 is sufficiently large, the Ramsey tax rates converge to the steady state along an increasing path. If k_0 is sufficiently low, they may evolve along an inverted-U curve. The underlying mechanism is essentially the same as that in the political equilibrium. Under $\sigma = 3$, the long-run Ramsey tax rate is slightly above zero, but still substantially smaller than that in the Markov political equilibrium.

6 Conclusion

The redistributive transfers in the pay-as-you-go social security system create conflicts of interest among various groups of households. The evolution of households characteristics may change the political support for the system over time. Despite the extensive studies of the aggregate and distributive effects of social security, most of the existing literature is silent on how the public decision on social security responds to the time-varying political support in a dynamic environment. In this paper, we analytically characterize the Markov perfect political equilibrium in which private intertemporal choices and the repeated political decision on social security are mutually affected over time. The main finding is that the dynamic interaction between social security and wealth inequality may lead to growing sizes of social security.

Our model also gives a number of other empirical predictions. We show that consumption inequality increases in age, the size of social security is positively correlated with the old population and an income inequality shock has no immediate effect on social security transfers. The evidence for these predictions can be found in the literature. We also compared the political equilibrium with the efficient Ramsey allocation. It turns out that the political decision process induces too large social security transfers in the long run, since the short-lived voters would ignore the negative impact of taxation on the welfare of future generations via capital accumulation.

For analytical convenience, we impose a balanced budget on social security transfers. A natural extension of the model would be to relax this assumption and allow for government debt. Since the model predicts a positive correlation between wealth inequality and social security, the direct test of our theory is to investigate the dynamic comovement of wealth inequality and the size of the social security programme. The analysis of the time-series data and the extended model with public debt is left for future researches.

References

Auerbach, Alan J. and Laurence J. Koltikoff, 1987, Dynamic Fiscal Policy, Cambridge University Press.

Boldrin, Michele and Aldo Rustichini, 2000, Political Equilibria with Social Security, Review of Economic Dynamics, 3, 41-78.Breyer, Friedrich and Ben Craig, 1997, Voting on Social Security: Evidence from OECD countries, European Journal of Political Economy, 13, 705-724.

Conesa, Juan C. and Dirk Krueger, 1999, Social Security Reform with Heterogeneous Agents, Review of Economic Dynamics, 2, 757-795.

Cooley, Thomas F. and Jorge Soares, 1999, A Positive Theory of Social Security Based on Reputation, Journal of Political Economy, 107, 135-160.

De Nardi, Mariacristina, 2004, Wealth Inequality and Intergenerational Links, Review of Economic Studies, 71, 743-768.

Deaton, Angus, Pierre-Olivier Gourinchas and Christina Paxson, 2000, Social Security and Inequality over the Life Cycle, In Martin Feldstein and Jeffrey Liebman

(Eds), The Distributional Aspects of Social Security and Social Security Reform, University of Chicago Press, Chicago.

Deaton, Angus and Christina Paxson, 1994, Intertemporal Choice and Inequality, Journal of Political Economy, 102, 437-467.

Feldstein, Martin and Liebman, Jeffrey B., 2001, Social Security, NBER Working Paper, No. 8451.

Forni, Lorenzo, 2005, Social Security as Markov Equilibrium in OLG Models, Review of Economic Dynamics, 8, 178-194.

Fuster, Luisa, Ayse Imrohoroglu and Selahattin Imrohoroglu, 2003, A Welfare Analysis of Social Security in a Dynastic Framework, International Economic Review, 44, 1247-1274.

Gonzalez-Eiras, Martin and Dirk Niepelt, 2004, Sustaining Social Security, mimeo, Institute for Internation Economic Studies.

Hassler, John, Jose V. Rodriguez Mora, Kjetil Storesletten and Fabrizio Zilibotti, 2003, The Survival of the Welfare State, American Economic Review, 93, 87-112.

Hassler, John, Kjetil Storesletten and Fabrizio Zilibotti, 2003, Democratic Public Good Provision, mimeo, Institute for Internation Economic Studies.

Hassler, John, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti, 2005, The Dynamics of Government, Journal of Monetary Economics, forthcoming.

Imrohoroglu, Ayse, Imrohoroglu, Selahattin and Joines, Douglas H, 1995, A Life Cycle Analysis of Social Security, Economic Theory, 6, 83-114.

Jones, Larry E., Rodolfo E. Manuelli and Peter E. Ross, 1993, Optimal Taxation in Models of Endogenous Growth, Journal of Political Economy, 101, 485-517.

Judd, Kenneth L., 1992, Projection Method for Solving Aggregate Growth Models, Journal of Economic Theory, 58, 410-452.

Judd, Kenneth L., 2003, Existence, Uniqueness, and Computational Theory for Time Consistent Equilibria: A Hyperbolic Discounting Example, mimeo, Hoover

Institution.

Katuscak, Peter, 2002, A Positive Theory of Social Security in an Open Economy, mimeo, University of Michigan.

Krusell, Per, Vincenzo Quadrini and Jose Victor Rios-Rull, 1997, Politico-Economic Equilibrium and Economic Growth, Journal of Economic Dynamic and Controls, 21, 243-272.

Krusell, Per and Jose-Victor Rios-Rull, 1999, On the Size of US Government: Political Economy in the Neoclassical Growth Model, American Economic Review.

Lindbeck, Assar and Jorgen W. Weibull, 1987, Balanced-Budget Redistribution as the Outcome of Political Competition, Public Choice, 52, 273-297.

Lindert, Peter H, What Limits social Spending, Explorations in Economic History, 33, 1-34.

Mankiw, Gregory N., David Romer and David N. Weil, 1992, A Contribution to the Empirics of Economic Growth, Quarterly Journal of Economics, 107, 407-437.

Meltzer, Allan and Scott Richard, 1983, Tests of a Rational Theory of the Size of Government, Public Choice, 41, 403-418.

Mulligan, Casey B. and Xavier Sala-i-Martin, 1999a, Gerontocracy, Retirement, and Social Security, NBER Working Paper, No. 7117.

Mulligan, Casey B. and Xavier Sala-i-Martin, 1999b, Social Security in Theory and practice(I): Facts and Political Theories, NBER Working Paper, No. 7118.

Persson, Tortsen and Guido Tabellini, 2000, Political Economics, MIT Press, Cambridge, MA.

Prescott, Edward C., 1986, Theory ahead of Business Cycle Measurement, Federal Reserve Bank of Mineapolis Quarter Review, 10, 9-22.

OECD, 1988, Reforming Public Pensions, reported prepared by Robert Holzmann, OECD Social Policy Studies No.5.

Ortigueira, Salvador, 2004, Markovian Optimal Taxation: The Macroeconomic Implications of Two Notions of Time Consistency, mimeo, European University Institute.

Razin, Assaf, Efraim Sadka and Phillip Swagel, 2002, The Aging Population and the Size of the Welfare State, Journal of Political Economy, 110, 900-918.

Rodriguez, Francisco, 1998, Essays on Redistribution, Development and the State, Ph.D. Dissertation, Harvard University.

Rodriguez, Francisco, 1999, Does Distributional Skewness Lead to Redistribution? Evidence from the United States, Economics and Politics, 11, 171-199.

Storesletten, Kjetil, Christopher I. Telmer and Amir Yaron, 1999, The Risk Sharing Implications of Alternative Social Security Arrangements, Carnegie Rochester Conference Series on Public Policy, 50, 213-259.

Storesletten, Kjetil, Christopher I. Telmer and Amir Yaron, 2004, Consumption and Risk Sharing over the Life Cycle, Journal of Monetary Economics, 51, 609-633.

Tabellini, Guido, 2000, A Positive Theory of Social Security, Scandinavian Journal of Economics, 102, 523-545.

Wolff, Edward N., 1992, Changing Inequality of Wealth, American Economic Review (Papers and Proceedings), 82, 552-558.

7 Appendix

7.1 Proof of Proposition 1

To solve the equilibrium policy rule, we first investigate a finite-period version of the model. It will be shown that the limit of finite-horizon equilibria turns out to be equivalent to the infinite-horizon equilibrium. Suppose that the economy terminates at time T and that young households born at time T only live one period.

First consider the terminal period T. Since young households do not have any intertemporal trade-off and $c_T^{y,j}$ simply equals the net earning $(1-\tau_T)\gamma^j w_T$, the welfare effect of τ_T on young households at time T is equal to

$$\frac{\partial U_T^{y,j}}{\partial \tau_T} = -\frac{\gamma^j w_T}{c_T^{y,j}} = -\frac{1}{1-\tau_T}. \tag{3.43}$$

The welfare effect of τ_T on old households follows (3.20). Plugging (3.43) and (3.20) into the first-order condition and assuming interior solution,

$$\sum_{j=h,l} \frac{1-\alpha}{2\alpha \frac{k_T^j/k_T^i}{1+k_T^j/k_T^i} + (1-\alpha)\tau_T} - 2n\frac{1}{1-\tau_T} = 0. \tag{3.44}$$

Note that the second order condition always holds. (3.44) gives a quadratic equation of τ_T

$$\Delta(\upsilon_T)\tau_T^2 + \Phi(\upsilon_T)\tau_T + \frac{4\upsilon_T\alpha^2 k_T^h/k_T^l}{\left(1+k_T^h/k_T^l\right)^2} - \alpha = 0, \tag{3.45}$$

where $\upsilon_T \equiv n/(1-\alpha)$.

Now, we turn to corner solutions. First consider the case where $\Phi(\upsilon_T) \geq 0$. Since $\Delta(\upsilon_T) > 0$, there is a unique positive τ_T if and only if

$$\left(k_T^h/k_T^l\right)^2 + (2-4\upsilon_T\alpha)\left(k_T^h/k_T^l\right) + 1 > 0. \tag{3.46}$$

For $\upsilon_T\alpha < 1$, the condition always holds. Otherwise, we need

$$k_T^h/k_T^l > 2\upsilon_T\alpha - 1 + 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)} \text{ or } k_t^h/k_t^l < 2\upsilon_T\alpha - 1 - 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)}. \tag{3.47}$$

The first inequality in (3.47) is binding since $2\upsilon_T\alpha - 1 + 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)} > 1$ for $\upsilon_T\alpha \geq 1$. The other inequality in (3.47) cannot be satisfied since $2\upsilon_T\alpha - 1 - 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)} < 1$ for $\upsilon_T\alpha \geq 1$. Second consider the case where $\Phi(\upsilon_T) < 0$. For $\upsilon_T\alpha < 1$, (3.46) ensures a unique positive τ_T. For $\upsilon_T\alpha \geq 1$, there can be two positive roots if the LHS of (3.46) is non-positive. This implies $n < 1/(2\alpha) - 1$ and contradicts the condition that $n \geq 1/\alpha - 1$, as implied by $\upsilon_T\alpha \geq 1$.

To conclude, for $\upsilon_T\alpha < 1$, the Markovian policy rule at time T follows

$$\tau_T = \mathbf{z}^T\left(k_T^h, k_T^l\right) = \frac{-\Phi(\upsilon_T) + \sqrt{\Phi(\upsilon_T)^2 - 4\Delta(\upsilon_T)\left(\frac{4\upsilon_T\alpha^2 k_T^h/k_T^l}{\left(1+k_T^h/k_T^l\right)^2} - \alpha\right)}}{2\Delta(\upsilon_T)}. \tag{3.48}$$

For $\upsilon_T \alpha \geq 1$, $\mathbf{z}_T\left(k_T^h, k_T^l\right)$ follows (3.48) if k_T^h / k_T^l satisfies the first inequality in (3.47) and is equal to zero otherwise.

Next we consider period $T-1$. The policy rule $\mathbf{z}^T\left(k_T^h, k_T^l\right)$ at time T implies $\mathbf{z}_1^T / \mathbf{z}_2^T = -k_T^l / k_T^h$. Some algebra manipulations establish

$$\frac{\partial k_T^h}{\partial \tau_{T-1}} = \hat{S}_2^h\left(k_{T-1}, \tau_{T-1}\right) = -\omega\left(\tau_T\right) \psi\left(\tau_T\right)\left(k_{T-1}/n\right)^\alpha, \tag{3.49}$$

$$\frac{\partial k_T^l}{\partial \tau_{T-1}} = \hat{S}_2^l\left(k_{T-1}, \tau_{T-1}\right) = -\psi\left(\tau_T\right)\left(k_{T-1}/n\right)^\alpha. \tag{3.50}$$

This gives $\hat{S}_2^j\left(k_{T-1}, \tau_{T-1}\right) = S_2^j\left(k_{T-1}, \tau_{T-1}, \tau_T\right)$. Moreover, substituting (3.49) and (3.50) for $\hat{S}_2^j$ in (3.22) leads to

$$\frac{\partial \tau_T}{\partial \tau_{T-1}} = 0. \tag{3.51}$$

According to Lemma 1, for any given τ_T, τ_{T-1} does not affect the future wealth inequality k_T^h / k_T^l. Therefore, given the policy rule F^T as a function of k_T^h / k_T^l, τ_{T-1} has no impact on future policy outcome τ_T. The dynamic link between τ_{T-1} and τ_T breaks down and the strategic effect does not exist.

The welfare effect of τ_{T-1} on young households follows (3.21). Using (3.51) and the indirect utility approach discussed in the next subsection, we find

$$\frac{\partial U_{T-1}^{y,j}}{\partial \tau_{T-1}} = -\frac{1+\beta\alpha}{1-\tau_{T-1}}. \tag{3.52}$$

The welfare effect of τ_{T-1} on the old households still follows (3.20). Plugging (3.43) and (3.20) into the first-order condition and assuming interior solution, we have

$$\sum_{j=h,l} \frac{1-\alpha}{2\alpha \frac{k_{T-1}^j / k_{T-1}^i}{1 + k_{T-1}^j / k_{T-1}^i} + (1-\alpha)\tau_{T-1}} - 2n \frac{1+\beta\alpha}{1-\tau_{T-1}} = 0, \tag{3.53}$$

which gives a quadratic equation of τ_{T-1}

$$\Delta\left(\upsilon_{T-1}\right)\tau_{T-1}^2 + \Phi\left(\upsilon_{T-1}\right)\tau_{T-1} + \frac{4\upsilon_{T-1}\alpha^2 k_{T-1}^h / k_{T-1}^l}{\left(1 + k_{T-1}^h / k_{T-1}^l\right)^2} - \alpha = 0, \tag{3.54}$$

where $\upsilon_{T-1} \equiv n\left(1+\alpha\beta\right) / \left(1-\alpha\right)$. The conditions for corner solutions can easily be derived following the above procedures.

To conclude, for $\upsilon_{T-1}\alpha < 1$, the Markovian policy rule at time $T-1$ follows

$$\tau_{T-1} = \mathbf{z}^{T-1}\left(k_{T-1}^h, k_{T-1}^l\right) = \frac{-\Phi\left(\upsilon_{T-1}\right) + \sqrt{\Phi\left(\upsilon_{T-1}\right)^2 - 4\Delta\left(\upsilon_{T-1}\right)\left(\frac{4\upsilon_{T-1}\alpha^2 k_{T-1}^h / k_{T-1}^l}{\left(1 + k_{T-1}^h / k_{T-1}^l\right)^2} - \alpha\right)}}{2\Delta\left(\upsilon_{T-1}\right)}. \tag{3.55}$$

For $\upsilon_{T-1}\alpha \geq 1$, τ_{T-1} follows (3.48) if k_{T-1}^{h}/k_{T-1}^{l} satisfies

$$k_{T-1}^{h}/k_{T-1}^{l} > 2\upsilon_{T-1}\alpha - 1 + 2\sqrt{\upsilon_{T-1}\alpha\left(\upsilon_{T-1}\alpha - 1\right)} \tag{3.56}$$

and τ_{T-1} is equal to zero otherwise.

It immediately follows that the only difference in $\mathbf{z}^{T-1}$ and $\mathbf{z}^{T}$ lies in $\upsilon_{T-1} = n\left(1+\alpha\beta\right)/\left(1-\alpha\right)$ and $\upsilon_{T-1} = n/\left(1-\alpha\right)$. Young households born at time $T-1$ live for two periods and thus $\partial U_{T-1}^{y,j}/\partial\tau_{T-1}$ in (3.52) differs from $\partial U_{T}^{y,j}/\partial\tau_{T}$ in (3.43). Moreover, it can easily be seen that the political decision on τ_t for $t < T-1$ is exactly the same as in time $T-1$. The equivalence boils down to the independence of $\partial U_{t}^{y,j}/\partial\tau_{t}$ on the future tax rate and the mute strategic effect, as shown in (3.52) and (3.51), respectively. These two features transform the dynamic problem into a static one. Consequently, the key parameter is $\upsilon_t = \upsilon_{T-1}$ for $t < T-1$. The finite-horizon equilibria thus converge to the infinite-horizon Markov perfect equilibrium in two periods.

Finally, we solve the private saving function $\tilde{S}^{j}$. (3.24) implies that $\mathbf{z}_1\left(k_t^h, k_t^l\right)/\mathbf{z}_2\left(k_t^h, k_t^l\right) = -k_{t+1}^{l}/k_{t+1}^{h}$. (3.16) and (3.17) can be rewritten as

$$\hat{S}_2^h = S_2^h \frac{1+\mathbf{z}_2\left(\left(S_2^l/S_2^h\right)S_3^h - S_3^l\right)}{1+\mathbf{z}_2\left(\left(S^l/S^h\right)S_3^h - S_3^l\right)}, \tag{3.57}$$

$$\hat{S}_2^l = S_2^l \frac{1+\mathbf{z}_1\left(\left(S_2^h/S_2^l\right)S_3^l - S_2^l S_3^h\right)}{1+\mathbf{z}_1\left(\left(S^h/S^l\right)S_3^h - S_3^l\right)}. \tag{3.58}$$

Since $S_2^l/S_2^h = S^l/S^h$, (3.57) and (3.58) give $\hat{S}_2^j = S_2^j$. The same argument establishes that $\hat{S}_1^j = S_1^j$, which implies that $\hat{S}^h$ and $\hat{S}^l$ follow (3.26) and (3.27), respectively, with a constant $\hat{\tau}$ to be determined. Since future wealth inequality equals $\omega\left(\hat{\tau}\right)$, the equilibrium policy rule (3.24) implies that $\hat{\tau}$ solves (3.28). Substituting $\mathbf{z}$ for τ_t in $\hat{S}^j$ establishes $\tilde{S}^j$. □

7.2 Proof of Lemma 2

We use the indirect utility approach to simplify the derivation of the welfare effect of social security tax rates. Using individuals' budget constraints (3.2) and (3.3), factor prices (3.4) and (3.5), the balanced-budget (3.6), private saving functions (3.7) and (3.8) and the law of motion of aggregate capital (3.11), after a bit algebra, we can obtain the indirect utility of all generations born at time t in terms of k_t, τ_t and τ_{t+1}:

$$\begin{aligned} V_t^j(k_t,\tau_t,\tau_{t+1}) &= (1+\beta\alpha)\,\alpha\log k_t + (1+\beta\alpha)\log(1-\tau_t) \\ &\quad + (1+\beta)\log\left(\alpha\gamma^j + \tau_{t+1}\phi(\tau_{t+1})\right) - \beta(1-\alpha)\log\phi(\tau_{t+1}). \end{aligned} \tag{3.59}$$

The indirect utility of the old households at time 0 is

$$U_0^{o,j} = \log\left(\alpha\frac{k_0^j/k_0^i}{1+k_0^j/k_0^i} + (1-\alpha)\,\tau_0\right) + \alpha\log k_0. \tag{3.60}$$

Differentiating (3.59), the welfare effect of τ_t on the old households at time t equals

$$\frac{\partial U_t^{o,j}}{\partial\tau_t} = \frac{\partial V_{t-1}^j}{\partial\tau_t} = (1+\beta)\frac{\phi(\tau_t)+\tau_t\phi'(\tau_t)}{\alpha\gamma^j A+\tau_t\phi(\tau_t)} - \beta(1-\alpha)\frac{\phi'(\tau_t)}{\phi(\tau_t)} \tag{3.61}$$

for $t \geq 1$. Differentiating (3.60) with respect to τ_0 yields the second line of (3.35). This proves the lemma. □

7.3 Proof of Lemma 3

By (3.11), we know $\partial k_{t+i}/\partial k_{t+i-1} = \alpha k_{t+i}/k_{t+i-1}$, $\partial k_{t+1}/\partial\tau_t = -k_{t+1}/(1-\tau_t)$ and $\partial k_t/\partial\tau_t = \phi'(\tau_t)\,k_t/\phi(\tau_t)$. Thus, $I_{t,t+i}$ can be written as

$$I_{t,t+i} = \begin{cases} \alpha^{i-1}k_{t+i}\left(-\frac{1}{1-\tau_t}+\alpha\frac{\phi'(\tau_t)}{\phi(\tau_t)}\right) & \text{for } t\geq 1 \\ -\alpha^{i-1}\frac{k_i}{1-\tau_0} & \text{for } t=0 \end{cases}. \tag{3.62}$$

According to the indirect utility function (3.59), the welfare effect of τ_t on young households at time t, for $t \geq 1$, equals

$$\frac{\partial U_t^{y,j}}{\partial\tau_t} = \frac{\partial V_t^j}{\partial\tau_t} + \frac{\partial V_t^j}{\partial k_t}\frac{\partial k_t}{\partial\tau_t} = (1+\beta\alpha)\left(-\frac{1}{1-\tau_t}+\alpha\frac{\phi'(\tau_t)}{\phi(\tau_t)}\right). \tag{3.63}$$

The welfare effect of τ_t on households born after time t is

$$\frac{\partial U_{t+i}^{y,j}}{\partial\tau_t} = \frac{\partial V_{t+i}^j}{\partial k_{t+i}}I_{t,t+i} = (1+\beta\alpha)\,\alpha^i\left(-\frac{1}{1-\tau_t}+\alpha\frac{\phi'(\tau_t)}{\phi(\tau_t)}\right) \tag{3.64}$$

for $i = 1, 2, \cdots$. The second equality in (3.64) comes from the first line in (3.62). (3.63) and (3.64) give the first line of (3.37). Finally, for $t = 0$, we have

$$\frac{\partial U_0^{y,j}}{\partial\tau_0} = \frac{\partial V_0^j}{\partial\tau_0} = -\frac{1+\beta\alpha}{1-\tau_0} \tag{3.65}$$

and

$$\frac{\partial U_i^{y,j}}{\partial \tau_0} = \frac{\partial V_i^j}{\partial k_i} I_{0,i} = -\alpha^i \frac{1+\beta\alpha}{1-\tau_0} \tag{3.66}$$

for $i = 1, 2, \cdots$. The second inequality in (3.14) comes from the second line in (3.62). (3.65) and (3.66) give the second line of (3.37). □

7.4 Proof of Proposition 2

The first-order condition of (3.30) with respect to τ_0 is

$$\sum_{j=h,l} \frac{\partial U_0^{o,j}}{\partial \tau_0} + \sum_{i=0}^{\infty} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial U_i^{y,j}}{\partial \tau_0} \right) = 0, \tag{3.67}$$

where $\partial U_i^j / \partial \tau_0$ follows from (3.65) and (3.66). This leads to

$$\sum_{j=h,l} \frac{1-\alpha}{\alpha \frac{k_0^j/k_0^i}{1+k_0^j/k_0^i} + (1-\alpha)\tau_0} - 2\rho \frac{1+\beta\alpha}{1-\rho\alpha} \frac{1}{1-\tau_0} = 0, \tag{3.68}$$

which gives a quadratic equation of τ_0. Comparing (3.68) with (3.56), it is immediate that the closed-form solution of τ_0 follows (3.39) with $\upsilon \equiv \rho(1+\alpha\beta)/((1-\rho\alpha)(1-\alpha))$. A comparison of the two first-order conditions shows that $\tau_0^R \gtreqless \tau_0^M$ if and only if $\beta \gtreqless 1/(1+\alpha n)$ (recall $\rho \equiv \beta n$). □

7.5 Proof of Proposition 3

The first-order conditions of (3.30) with respect to τ_t for $t \geq 1$ are

$$\sum_{j=h,l} \frac{\partial U_{t-1}^{y,j}}{\partial \tau_t} + \sum_{i=0}^{\infty} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} \right) = 0. \tag{3.69}$$

Substituting (3.63) and (3.64) for $\partial U_{t-1}^j / \partial \tau_t$ and $\partial U_{t+i}^j / \partial \tau_t$, respectively, (3.69) leads to

$$\sum_{j=h,l} \left(\frac{(1+\beta)(\phi(\tau_t) + \tau_t \phi'(\tau_t))}{\alpha \gamma^j + \tau_t \phi(\tau_t)} \right) - 2\rho \frac{1+\beta\alpha}{1-\rho\alpha} \frac{1}{1-\tau_t} + 2\left(\rho \frac{(1+\beta\alpha)\alpha}{1-\rho\alpha} - \beta(1-\alpha) \right) \frac{\phi'(\tau_t)}{\phi(\tau_t)} = 0. \tag{3.70}$$

(3.70) solves a constant τ_t for $t \geq 1$.

Note that the second order conditions are always satisfied. It immediately follows that $\partial^2 U_{-1}^{y,j} / \partial \tau_0^2$ and $\partial^2 U_i^{y,j} / \partial \tau_0^2$ are negative by (3.60), (3.65) and (3.66).

(3.61) shows that $\partial^2 U_{t-1}^{y,j}/\partial\tau_t^2 < 0$. Differentiating (3.63) and (3.64) with respect to τ_t establishes

$$\operatorname{sgn}\left(\frac{\partial^2 U_{t+i}^{y,j}}{\partial\tau_t^2}\right) = \operatorname{sgn}\left(\frac{\alpha\left(1-\alpha\right)}{\left(\alpha\left(1+\beta\right)+\left(1-\alpha\right)\tau_t\right)^2} - \frac{1}{\left(1-\tau_t\right)^2}\right).$$

Since $\tau_t \in [0,1]$, it can easily be found that $\partial^2 U_{t+i}^{y,j}/\partial\tau_t^2 < 0$ always holds.

Denote $L\left(\tau_t\right)$ LHS of (3.70). After some algebra manipulations, $L\left(\tau_t\right)$ can be written as

$$\begin{aligned} L\left(\tau_t\right) &= \sum_{j=h,l}\left(\frac{\left(1-\alpha\right)\beta\alpha^2\left(1+\beta\right)^2}{\left(\alpha\left(1+\beta\right)+\left(1-\alpha\right)\tau_t\right)\left(\gamma^j\alpha^2\left(1+\beta\right)+\alpha\left(1-\alpha\right)\left(\gamma^j+\beta\right)\tau_t\right)}\right) \\ &\quad -2\rho\frac{1+\beta\alpha}{1-\rho\alpha}\frac{1}{1-\tau_t} - 2\left(\rho\frac{\left(1+\beta\alpha\right)\alpha}{1-\rho\alpha} - \beta\left(1-\alpha\right)\right)\frac{1-\alpha}{\alpha\left(1+\beta\right)+\left(1-\alpha\right)\tau_t}. \end{aligned}$$

The second order condition implies that $L'\left(\tau_t\right) < 0$. This implies that the solution of (3.70) is unique. It is straightforward that $\lim_{\tau_t\to 1} L\left(\tau_t\right) = -\infty$ so there is a strictly positive τ_t if and only if $L\left(0\right) > 0$. This establishes (3.41).

7.6 CRRA Utility

Given (3.42), households' problem (3.1) becomes

$$\max_{k_{t+1}^j} \frac{\left(c_t^{y,j}\right)^{1-\sigma}-1}{1-\sigma} + \beta\frac{\left(c_{t+1}^{o,j}\right)^{1-\sigma}-1}{1-\sigma}, \tag{3.71}$$

subject to (3.2) and (3.3). Households' saving choice follows the Euler equation $c_{t+1}^{o,j}/c_t^{y,j} = \left(\beta R_{t+1}\right)^{1/\sigma}$ which solves (3.71). Using budget constraints (3.2) and (3.3), factor prices (3.4) and (3.5) and balanced-budget (3.6), k_{t+1}^j follows

$$\begin{aligned} k_{t+1}^j &= G^j\left(k_t,\tau_t,\tau_{t+1},k_{t+1}\right) \qquad (3.72) \\ &\equiv \frac{\gamma^j\left(A\alpha\left(k_{t+1}/n\right)^{\alpha-1}\beta\right)^{1/\sigma}A\left(1-\alpha\right)\left(1-\tau_t\right)\left(k_t/n\right)^\alpha - A\left(1-\alpha\right)\tau_{t+1}k_{t+1}^\alpha n^{1-\alpha}}{\left(A\alpha\left(k_{t+1}/n\right)^{\alpha-1}\beta\right)^{1/\sigma} + A\alpha\left(k_{t+1}/n\right)^{\alpha-1}}. \end{aligned}$$

By $k_{t+1} = \sum_{j=h,l} k_{t+1}^j/2$, (3.72) solves the private saving functions

$$k_{t+1}^j = S^j\left(k_t,\tau_t,\tau_{t+1}\right), \tag{3.73}$$

with

$$S_i^j = \frac{G_i^j}{1-\sum_{j=h,l} G_4^j/2}, \tag{3.74}$$

for $i = 1, 2, 3$. Correspondingly, the aggregate saving function can be written as

$$k_{t+1} = S\left(k_t, \tau_t, \tau_{t+1}\right), \tag{3.75}$$

with

$$S_i = \frac{\sum_{j=h,l} G_i^j/2}{1 - \sum_{j=h,l} G_4^j/2}. \tag{3.76}$$

Given the Markovian policy rule (3.14), the recursive form of the private and aggregate saving functions can be solved.

$$k_{t+1}^j = \hat{S}^j\left(k_t, \tau_t\right), \tag{3.77}$$

with

$$\hat{S}_i^h = \frac{G_i^h - \mathbf{z}_2\left(G_i^h G_3^l - G_i^l G_3^h\right) - \left(G_i^h G_4^l - G_i^l G_4^h\right)/2}{\left(1 - G_3^l \mathbf{z}_2 - G_4^l/2\right)\left(1 - G_3^h \mathbf{z}_1 - G_4^h/2\right) - \left(G_3^h \mathbf{z}_2 + G_4^h/2\right)\left(G_3^l \mathbf{z}_1 + G_4^l/2\right)} \tag{3.78}$$

$$\hat{S}_i^l = \frac{G_i^l - \mathbf{z}_1\left(G_i^l G_3^h - G_i^h G_3^l\right) - \left(G_i^l G_4^h - G_i^h G_4^l\right)/2}{\left(1 - G_3^h \mathbf{z}_2 - G_4^h/2\right)\left(1 - G_3^l \mathbf{z}_2 - G_4^l/2\right) - \left(G_3^l \mathbf{z}_1 + G_4^l/2\right)\left(G_3^h \mathbf{z}_2 + G_4^h/2\right)} \tag{3.79}$$

for $i = 1, 2$. The welfare effect, $\partial U_t^{o,j}/\partial \tau_t$ and $\partial U_t^{y,j}/\partial \tau_t$, as well as the first-order condition still follows (3.19), (3.21) and (3.23), respectively. These derivatives will be used in the numerical solution, as will be seen in the next subsection.

Now, we turn to the Ramsey problem. The indirect utility of young households at time t can be expressed as follows.

$$W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right) \equiv \left(\gamma^j A\left(1 - \tau_t\right)\left(k_t/n\right)^{\alpha} + \tau_{t+1} k_{t+1}/\alpha\right)^{1-\sigma}\left(1 + \beta^{1/\sigma}\left(A\alpha\left(k_{t+1}/n\right)^{\alpha-1}\right)^{1/\sigma-1}\right)^{\sigma}. \tag{3.80}$$

(3.73) and (3.80) solve the indirect utility function $V_t^j\left(k_t, \tau_t, \tau_{t+1}\right)$ with

$$\begin{aligned}
\frac{\partial V_t^j}{\partial k_t} &= \frac{\partial W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right)}{\partial k_t} + \frac{\partial W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right)}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial k_t}, \\
\frac{\partial V_t^j}{\partial \tau_t} &= \frac{\partial W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right)}{\partial \tau_t} + \frac{\partial W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right)}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial \tau_t}, \\
\frac{\partial V_t^j}{\partial \tau_{t+1}} &= \frac{\partial W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right)}{\partial \tau_{t+1}} + \frac{\partial W^j\left(k_t, \tau_t, \tau_{t+1}, k_{t+1}\right)}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial \tau_{t+1}}.
\end{aligned}$$

The welfare effect can be written as follows. For $t \geq 1$,

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{\partial V_{t-1}^j}{\partial \tau_t} \tag{3.81}$$

and

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \begin{cases} \frac{\partial V_{t+i}^{j}}{\partial k_{t+i}} I_{t,t+i} & \text{if } i \geq 1 \\ \frac{\partial V_t^j}{\partial \tau_t} + \frac{\partial V_t^j}{\partial k_t} I_{t,t} & \text{if } i = 0 \end{cases}. \tag{3.82}$$

For $t = 0$, $k_t = 0$ is predetermined and hence, $I_{0,0} = 0$. The first-order condition of the Ramsey problem still follows (3.38).

7.7 Numerical Method for the Markovian Political Equilibrium

A direct application of the projection method for the present problem with heterogeneous agents is to approximate $\mathbf{z}$, $\hat{S}^h$ and $\hat{S}^l$ by three two-dimensional n-order Chebyshev polynomials with tensor products. Consequently, we need to pin down $3 \times n^2$ coefficients of the polynomials that satisfy the Euler equation and the first-order condition (3.23). That is to say, the computation will be involved in solving $3 \times n^2$ nonlinear equations.

However, the analysis in the preceding subsection suggests that computing functions $\hat{S}^j$ is not necessary. In fact, only the derivatives $\hat{S}_i^j$, rather than the function $\hat{S}^j$, are of importance for the equilibrium policy rule $\mathbf{z}$. The following strategy substantially reduces the computational cost. Specifically, the number of nonlinear equations drops from $3 \times n^2$ to n^2. First, we approximate $\mathbf{z}$ by

$$\mathbf{z}\left(k^h, k^l\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \phi_{ij}\left(k^h, k^l\right), \tag{3.83}$$

where $\phi_{ij}\left(k^h, k^l\right)$ are the tensor products of one-dimensional Chebyshev polynomials. The second step is to pin down the partial derivatives appearing in the first-order condition (3.23). S_i^j is easy to compute. Plugging $\mathbf{z}_1$, $\mathbf{z}_2$ and S_i^j into (3.78) and (3.79), $\hat{S}_i^j$ can be solved. Finally, choose n points in the state space $\left[k^{h,\min}, k^{h,\max}\right]$ and $\left[k^{l,\min}, k^{l,\max}\right]$, respectively, by Chebyshev collocation. The first-order condition (3.23) has to be satisfied for each point. Thus, the functional equation is transformed into n^2 nonlinear equations, which solve n^2 unknown coefficients a_{ij} in (3.83).

Following Judd (1992), the accuracy of the approximation can be indirectly assessed by the Euler equation error. Let $\tilde{\mathbf{z}}$ be the approximated $\mathbf{z}$. The Euler equation error on any given pair $\left(k^h, k^l\right)$ is measured by the percentage deviation from τ_t implied by the approximated equilibrium policy rule $\mathbf{z}\left(k^h, k^l\right)$ to the "true" optimal τ_t that solves (3.23) as if $\mathbf{z} = \tilde{\mathbf{z}}$. The accuracy increases with the order of Chebyshev polynomial. However, the improvement tends to be less significant with higher degrees, which increase the computation cost exponentially. In our case, the polynomial of 8-order turns out to be sufficient. The Euler equation errors over 900 points that are uniformly collected in the state space are computed. The maximum errors in all numerical experiments are below 10^{-3}.

A common problem associated with the projection method is that the convergence of the solution for unknown coefficients highly depends on the initial guess. In a standard growth model, a good initial guess can be obtained by linearizing the policy function around the steady state. This problem turns out to be much more serious in the present environment since we essentially have no idea about the steady state. Fortunately, we know the closed-form solution $\mathbf{z}$ under logarithm utility.. So we adopt a simple continuation method, i.e., use the analytical solution $\mathbf{z}$ as an initial guess for $\sigma = 1 + \varepsilon$. Some perturbations on the initial guess are used to check the local convergence of the solution. The equilibrium policy rule $\mathbf{z}$ turns out to be unique in the numerical experiments so far.

7.8 Numerical Method for the Ramsey Solution

Given the indirect utility V_t^j, the Ramsey problem (3.30) can be rewritten as

$$\max_{\{\tau_t\}_{t=0}^{\infty} \text{ and } \{k_t\}_{t=1}^{\infty}} \sum_{j=h,l} U_0^{o,j}\left(k_0^j, k_0, \tau_0\right) + \sum_{t=0}^{\infty} \rho^{t+1} \left(\sum_{j=h,l} V_t^j\left(k_t, \tau_t, \tau_{t+1}\right) \right), \tag{3.84}$$

subject to the law of motion of aggregate capital (3.75). The first-order conditions with respect to τ_t and k_t for $t \geq 1$ are

$$\sum_{j=h,l} \frac{\partial V_{t-1}^j}{\partial \tau_t} + \rho \sum_{j=h,l} \frac{\partial V_t^j}{\partial \tau_t} = \mu_{t-1} \frac{\partial k_t}{\partial \tau_t} + \rho \mu_t \frac{\partial k_{t+1}}{\partial \tau_t}, \tag{3.85}$$

$$\rho \sum_{j=h,l} \frac{\partial V_t^j}{\partial k_t} = -\mu_{t-1} + \rho \mu_t \frac{\partial k_{t+1}}{\partial k_t}, \tag{3.86}$$

where μ_t is the Lagrangian multiplier associated with k_{t+1}. Let $\bar{x}$ be the steady state value of variable x. For notational convenience, we denote $\bar{V}_1^j$, $\bar{V}_2^j$ and $\bar{V}_3^j$ as the steady states of $\partial V_t^j / \partial k_t$, $\partial V_t^j / \partial \tau_t$ and $\partial V_t^j / \partial \tau_{t+1}$, respectively. Similarly, $\bar{S}_1$,

$\bar{S}_2$ and $\bar{S}_3$ are referred to as the steady states of $\partial S_t/\partial k_t$, $\partial S_t/\partial \tau_t$ and $\partial S_t/\partial \tau_{t+1}$, respectively. Then (3.86) leads to

$$\bar{\mu} = -\frac{\rho \sum_{j=h,l} \bar{V}_1^j}{1-\rho \bar{S}_1}. \tag{3.87}$$

Using (3.87), (3.85) implies

$$\sum_{j=h,l} \bar{V}_3^j + \rho \sum_{j=h,l} \bar{V}_2^j + \frac{\rho\left(\rho \bar{S}_2 + \bar{S}_3\right)}{1-\rho \bar{S}_1} \sum_{j=h,l} \bar{V}_1^j = 0. \tag{3.88}$$

Moreover, (3.75) gives

$$\bar{k} = S\left(\bar{k}, \bar{\tau}, \bar{\tau}\right). \tag{3.89}$$

(3.88) and (3.89) solve the steady state capital stock $\bar{k}$ and the steady state social security tax rate $\bar{\tau}$.

Following Jones, Manuelli and Ross (1993), we adopt the truncated method to solve the Ramsey allocation. Assume that the economy reaches the steady state after period T. Then the infinite-horizon problem (3.84) can be approximated by a finite-horizon one

$$\max_{\{\tau_t\}_{t=0}^{T-1} \text{ and } \{k_t\}_{t=1}^{T}} \sum_{j=h,l} U_0^{o,j}\left(k_0^j, k_0, \tau_0\right) + \sum_{t=0}^{T-1} \rho^{t+1} \left(\sum_{j=h,l} V_t^j\left(k_t, \tau_t, \tau_{t+1}\right) \right) + \Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right), \tag{3.90}$$

subject to the law of motion of aggregate capital (3.75). The value of continuation $\Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right)$ is equal to

$$\Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right) = \sum_{t=T}^{\infty} \rho^{t+1} \left(\sum_{j=h,l} V_t^j\left(k_T, \bar{\tau}, \bar{\tau}\right) \right), \tag{3.91}$$

which corrects the error caused by "end effects". Therefore, standard nonlinear programming techniques can be applied to solve (3.90). For interior solutions, $\{\tau_t\}_{t=0}^{T-1}$ may be directly solved by the first-order conditions. The effect of τ_t on $\Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right)$ is

$$\frac{\partial \Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right)}{\partial \tau_t} = \sum_{i=T}^{\infty} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i} \right) = \frac{\rho^{T+1} \sum_{j=h,l} \bar{V}_1^j}{1-\rho \bar{S}_1} I_{t,t+T}. \tag{3.92}$$

Using (3.82) and (3.92), the first-order conditions of (3.84) with respect to τ_t for $t \geq 1$ can be written as

$$\sum_{j=h,l} \frac{\partial V_{t-1}^j}{\partial \tau_t} + \rho \sum_{j=h,l} \frac{\partial V_t^j}{\partial \tau_t} + \sum_{i=0}^{T-1} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i} \right) + \frac{\rho^{T+1} \sum_{j=h,l} \bar{V}_1^j}{1-\rho \bar{S}_1} I_{t,t+T} = 0. \tag{3.93}$$

Similarly, we have the first-order condition of (3.84) with respect to τ_0

$$\sum_{j=h,l}\frac{\partial U_0^{o,j}}{\partial \tau_0}+\rho\sum_{j=h,l}\frac{\partial V_0^j}{\partial \tau_0}+\sum_{i=0}^{T-1}\rho^{i+1}\left(\sum_{j=h,l}\frac{\partial V_i^j}{\partial k_i}I_{0,i}\right)+\frac{\rho^{T+1}\sum_{j=h,l}\bar{V}_1^j}{1-\rho\bar{S}_1}I_{0,T}=0. \tag{3.94}$$

(3.93) and (3.94) constitute a nonlinear equation system which solves $\{\tau_t\}_{t=0}^{T-1}$.

Chapter 4

A Markovian Social Contract of Social Security*

1 Introduction

In the preceding chapter, we have shown that probabilistic voting (by placing a weight on the interests of all individuals in society) is one way of explaining why a social security system can be sustained over time. However, this is arguably not the most convincing one. In the real world we know that many young people would be opposed to shutting down the system, even if this would imply for them an immediate reduction in the tax burden. The support for the pension system from the young generation may be explained by altruistic considerations vis-a-vis the current old. However, it could also occur by the logic of the social contract pointed out by Sjoblom (1985): self-interested agents perceive that future social security benefits are somehow related to the present contribution.[2] Thus, they are willing to pay the contributions in expectation that the system will not be terminated by the future generation. These considerations motivate us to study a model where the social security system is sustained over time in a median voter environment, and where the median voter is a net contributor rather than a recipient.

* We thank Caroline Betts, Ayse Imrohoroglu, Michael Magill, Kjetil Storesletten, and especially Fabrizio Zilibotti, for very helpful discussions.

[2] To avoid the problem of temporal separation of contributions and benefits, some studies assume the welfare of retirees to be weighted into the preference of the policymaker, by resorting to altruism (Tabellini, 2000), probabilistic voting (Katuscak, 2002, Gonzalez-Eires and Niepelt, 2004, Song, 2005) or gerontocracy (Mulligan and Sala-i-Martin, 1999a).

To formulate such a dynamic linkage between current contributions and future benefits, the literature has so far relied on two approaches. The first approach simply assumes away the temporal separation problem by allowing "once-and-for-all-voting" in the sense that the initial median voter has full commitment to future social security benefits.[3] The second approach adopts the "trigger strategy equilibrium", where the expectation of future benefits can be self-fulfilled in an infinite-horizon dynamic game with a system of constructed punishments.[4] While certainly useful, trigger strategy equilibria are generically indeterminate and hence, unable to provide sharp empirical predictions. Moreover, the collective coordination on some specific form of punishment required by the trigger strategy equilibrium seems hard to achieve in the real world economy.

This paper analyses the Markov perfect equilibrium which sustains the pay-as-you-go social security system in a majority voting framework with intra-cohort heterogeneity. In the class of equilibria investigated, we show that the decisive voter is always middle-aged and associated with lower income.[5] At first sight, sustaining the pension system through Markov equilibria may seem desperate. Under Markov strategies that only depend on the payoff-relevant state variables, the current policy choice has no direct effect on future pension benefits. The middle-aged contributors might thus optimally choose to overturn the system by voting for zero social security taxes. However, our result shows that even if contributions and benefits are temporally separated, there exists a Markovian social contract through which the self-interested median voter has incentives to sustain the system. The sustainability in this setup is twofold. First, the current policy choice indirectly affects the future policy outcomes via private intertemporal decision, which forms a self-fulfilled expectation of future social security benefits being positively related to current contributions. This solves the temporal separation problem of benefits and contributions. Second, given the positive linkage, the median voter has an incentive to support the system for *intra-generational*

[3] See, among others, Browning (1975), Conesa and Krueger (1999), Persson and Tabellini (2000, chapter 6).

[4] See, among others, Cooley and Soares (1999) and Boldrin and Rustichini (1999).

[5] According to the United Nations (2000), the four eldest countries in the world are Monaco, Italy, Greece and Sweden with respectively 29, 22, 22 and 22 percent of their population above age 60. Therefore, the voter of median age is not a retiree but a taxpaying worker, as pointed out by Mulligan and Sala-i-Martin (1999b).

redistribution. We prove the uniqueness of the differentiable Markov perfect equilibrium in the space of linear functions.[6] Numerical exercises show that the equilibrium strategy in the space of differentiable nonlinear functions indeed converges to the linear one. Our model gives a variety of empirical predictions that are broadly consistent with the cross-country data. Particularly, the Markovian social contract allows us to explain the empirical observation that there is an insignificant or even negative relationship between inequality and social security (or government) transfers. This relationship is hard to explain with the existing theory, which implies that a higher level of inequality provides more political support for redistributive legislation.[7]

Our model economy is a small open economy. There are three living generations in each period of time, the young, the middle-aged and the old. Agents work in the first two periods of their life and retire in the third. Young agents can make human capital investment to increase their labour productivity and thus, their wage income. The state of the economy is characterized by the human capital of the middle-aged. All agents have linear utility on consumption and human capital investment involves a quadratic loss for young agents. The linear-quadratic preference helps obtain closed-form solution. Labour income of the young and the middle-aged is taxed to finance the social security benefits of the old. To capture the intra-generational redistributive effect of social security, we introduce intra-cohort heterogeneity by assuming each agent to be born with either high or low ability. Social security benefits are uniformly distributed across different types of old agents.

We solve for the differentiable Markov perfect equilibrium where the policy choice of the decisive voter is a differentiable function of the payoff-relevant state variable, i.e., her human capital stock. In this class of equilibria, the middle-aged agent with low ability, referred to as "the middle poor", is always decisive under repeated majority voting. Moreover, it turns out that future social security benefits are positively related to the current contributions. On the one hand,

[6] The dynamic strategic interaction may yield multiple equilibria, even if the strategies are restricted to being linear. See, for example, Phelps and Pollak (1968), Lockwood (1996), and Azariadis and Galasso (2002).

[7] See, for example, Meltzer and Richard (1981), Alesina and Rodrik (1994), Persson and Tabellini (1994).

the current payroll tax rate distorts investment and reduces the human capital stock of the median voter in the next period. On the other hand, a lower human capital stock implies a lower labour income, which reduces the cost of the social security tax rate paid by the next period's median voter. This induces her to choose a higher tax rate. Due to the interaction between the economic and political choices, individuals rationally expect that increasing current social security contributions has a positive effect on future benefits. The political choice of the current contributions takes this link into consideration. Therefore, the median voter would like to choose a tax rate equalizing the marginal cost of current contributions and future marginal benefits. When income inequality is very low, however, the marginal cost outweighs the benefits and the median voter will choose zero tax rate. Positive social security taxes are sustained in equilibrium, only if inequality is sufficiently large. This highlights the role of intra-generational redistribution on the sustainability of the social security system.

The social security system entails both inter- and intra-generational redistributive aspects. Therefore, the political decision on social security is determined by demographic factors as well as within-cohort inequality. In our model, there are two effects of dependency ratio and income inequality on the size of social security. First, there is a direct effect on the median voter's policy choices. Taking the equilibrium return of contributions implied by the social contract as a parameter,[8] a higher dependency ratio makes the inter-generational transfers less desirable. The direct effect thus implies that the dependency ratio reduces the social security transfers, as predicted by the existing theory. Second, dependency ratio inequality affects the current median voter's policy choices via its impact on the equilibrium return of contributions. We refer to this channel, which has so far largely been neglected in the literature, as *the social contract effect*. A higher dependency ratio implies a lower quantity of labour supply and makes the future tax base less elastic. This induces the middle-aged median voter to raise social security taxes, which implies a high return of benefits to the past contributions. In equilibrium, we obtain a positive relationship between dependency ratio and the return to contributions. Consequently, a high dependency ratio tends to increase transfers. Like the dependency ratio, income inequality has two opposite

[8] We refer to the marginal return of future benefits to current contributions implied by the social contract as the equilibrium return of social security contributions.

effects on the size of social security. The sign of the overall impact depends on which is the dominating effect.

The Markovian social contract gives rise to a variety of distinctive empirical predictions. First of all, the social security tax rate converges to the steady state along an increasing path. The intuition is the following. Suppose there to be no social security system before the initial period. Since social security taxes distort human capital investment, the median voter in the initial period has more human capital than median voters in subsequent periods. As a result, the initial median voter tends to vote for a lower tax rate because the cost of taxing human capital is larger. Similarly, since the second period median voter has less human capital than the initial period median voter but more than median voters in the following periods, she would impose a tax rate higher than the initial tax rate but lower than the future ones. The social security tax rate keeps growing until it converges to the steady state.

Second, the growth of social security is positively correlated with the dependency ratio and negatively correlated with income inequality. This result boils down to the nature of the social contract effect. Social security growth reflects the responsiveness of future transfers to current ones, a dynamic linkage embedded in the equilibrium social contract. For the median voter, such an equilibrium relationship provides a positive return of social security benefits to the contributions. As mentioned above, the social contract effect implies that the dependency ratio (income inequality) has a positive (negative) impact on the equilibrium return. A high equilibrium return makes future transfers more responsive to current ones and hence, leads to a high growth in social security.

Third, the steady state size of social security is negatively correlated to the dependency ratio but non-monotonically correlated to income inequality. Due to the opposite impact of the direct and social contract effect, the sign of the overall impact depends on which effect eventually dominates. Our numerical results show that in the steady state, the social security tax rate decreases in the dependency ratio, but turns out to be an inverted-U function of income inequality. This indicates that the direct effect of the dependency ratio always dominates the social contract effect. The direct effect of income inequality dominates the social contract effect when inequality is sufficiently low, while the social contract effect starts overshadowing the direct effect when inequality increases to some critical level.

We use data from OECD countries to examine the consistency between the prediction of our model and the facts. The first prediction is clearly consistent with the history of social security systems in the OECD, which features growing sizes of social security. To test the second prediction, we regress the average growth rate of social security benefits on the dependency ratio and income inequality. The results show that income inequality correlates negatively with social security growth, as predicted by our theory. This provides empirical evidence for the existence of the social contract effect. The demographic impact on the growth of social security is not significant, however. Coherent with the third prediction, the size of social security is positvely correlated with the dependency ratio but negatively correlated with income inequality. Particularly, the latter result suggests that the social contract effect of income inequality on the size of social security outweighs the direct effect.

This paper contributes to the literature on the political sustainability of the social security system in several respects. We show that under the temporal separation of contributions and benefits, social security can be sustained by a middle-aged median voter who plays Markov strategy. In previous efforts to construct the linkage between future benefits and current contribution, future policies are often conditioned on variables that are payoff-irrelevant for future policymakers. The assumption of full commitment in once-and-for-all voting is obviously far from realistic, though it substantially simplifies the analysis. A trigger strategy approach makes a major theoretical advance by allowing repeated voting. However, the dependence of policy choices on the payoff-irrelevant voting history leads to the indeterminacy of political equilibria. The multiplicity of equilibria may also exist in a non-trigger strategy approach. Azariadis and Galasso (2002), for example, find multiple equilibria where the median voter conditions her voting choices on capital stock. However, in their two-period OG model with exogenous factor prices, capital held by the old is payoff-irrelevant for the young median voter. A related but somewhat different problem can be found in Forni (2005), who uses capital as the state in a general equilibrium framework. While being payoff-relevant, capital *per se* does not affect the median voter's choices of social security. Capital affects the political choice if and only if the median voter expects future policies to be contingent on future capital. In our model, the linkage between policy choice and the state variable does not depend on expectations. Therefore, our work indicates that the fundamental linkage between policy choices and the payoff-relevant state variable may help overcome the indeterminacy of equilibria.[9]

Another advantage of our approach over the once-and-for-all voting and trigger strategy equilibrium is that the Markovian social contract generates nontrivial dynamics and thus, provides an explanation for the evolution of social security during the postwar period, while the size of social security is usually constant over time in the two other approaches.[10] The evolution of social security in our model is induced by the positive linkage between future social security benefits and current contributions, which serves as a necessary condition for the political sustainability of social security. Therefore, our model shows that the growing sizes of social security can be an intrinsic feature in the politico-economic equilibrium. This complements some recent studies on the dynamic patterns of social security, which resort to interactions between social security and private savings (Katuscak, 2002) or wealth inequality (Song, 2005a).[11]

The role of intra-generational redistribution of social security is often neglected in the literature. Our work shows that income inequality plays an essential role in the sustainability of the system. Particularly, social security can be sustained in a dynamic efficient economy if income inequality is not too low. Else, the system cannot compensate the contributions paid by the median voter with lower income and hence, will be overturned. This is in contrast with the existing Markovian

[9] McCallum (1983) finds that in a wide-class of linear rational expectation models, non-uniqueness of solutions occurs because unnecessary or 'extraneous' components are permitted to influence expected (and therefore actual) values of endogenous variables. Maskin and Tirole (2001) argue that by preventing non-payoff-relevant variables from affecting strategic behavior, Markov perfect equilibrium is often successful in eliminating or reducing a large multiplicity of equilibria in dynamic game.

[10] A notable exception is Boldrin and Rustichini (1999). However, their model predicts a decreasing sequence of social security tax rates that is inconsistent with the data.

[11] Social security would survive the backward refinement in Katuscak (2002) and Song (2005a), which adopt a probabilistic voting framework. Our model as well as the previous studies using trigger strategies, do not have the property. That is to say, social security can be sustained only if the horizon is infinite. This is a natural result implied by the temporal separation of costs and benefits, since the contributor in the last period will never vote for a positive social security tax rate.

approach, where the general equilibrium effect stands in the heart of the political sustainability of the social security system in a dynamic efficient economy (Forni, 2005, Gonzalez-Eiras and Dirk Niepelt, 2004).

Within the context of empirical work, evidence on the relationship between social security and income inequality is far from onclusive. For example, Tabellini (2000) finds that when applying cross-country regression (more than 40 countries), the size of social security is positively correlated with income inequality.[12] However, an opposite negative relationship across the OECD has been repeatedly reported (e.g., Lindert, 1996 and Rodriguez, 1998) and further confirmed in the present paper. Moreover, since social security constitutes the largest part of government transfers, it would not be surprising to see the lack of positive correlation between inequality and government transfers in the literature (see, e.g., Perotti, 1996 and Benabou, 1996). These empirical findings are in contrast to the conventional wisdom that government transfers should be increasing in income inequality. Taking into account the strategic interaction among median voters over time, our model shows that the puzzling ambiguous correlation between inequality and social security (or government transfers) in the OECD can be well explained by the social contract effect ignored in the existing theory.

The paper is organized as follows. Section 2 describes the economic environment and the voting procedure. We characterize the political equilibrium in Section 3. Section 4 examines the impact of income inequality and the dependence ratio on the size and the growth rate of social security, respectively. In section 5, we check the consistency of theory and empirical evidence. Section 6 concludes.

2 The Model Economy

Consider a small open economy inhabited by an infinite sequence of overlapping-generations. Each generation lives three periods. An agent works in the first two periods of her life and retires in the last. Labour supply in each of the first two periods is inelastic and normalized to unity. Young agents can make human capital investment to increase labour productivity.

[12] Persson and Tabellini (2000, Chapter 6) note that the measure of inequality is bounded to be imperfect for such a large sample of countries.

There is heterogeneity within each cohort. Agents can either have high ability (w^s) or low ability (w^u), and will be referred to as the rich or the poor, respectively. Let h_t^j be the human capital investment of a young agent born at time t of type j, $j = u, s$. Human capital and wage income in her working ages equal h_t^j and $w^j h_t^j$, respectively.[13] In addition to the inter-generational transfer component by which it is defined, the pay-as-you-go system also has important intra-generational redistributive elements since social security contributions are usually proportional to income, while benefits are more or less regressive.[14] For analytical convenience, following Conesa and Krueger (1999) and many others, we assume that social security benefits are evenly distributed within old households. Then, the life-time wealth A_t^j follows

$$A_t^j = (1 - \tau_t) w^j h_t^j + \frac{(1 - \tau_{t+1}) w^j h_t^j}{R} + \frac{p_{t+2}}{R^2}, \tag{4.1}$$

where τ_t is the proportional wage income tax rate levied on working generations and p_{t+2} is the social security benefits to retirees born at time t.

To obtain a closed-form solution, we assume agents to have a linear-quadratic preference over life-time wealth and costs of human capital investment (see Hassler *et al.*, 2003):

$$\max_{h_t^j} A_t^j - \frac{1}{2} \left(h_t^j\right)^2, \tag{4.2}$$

subject to (4.1). Solving (4.2) yields

$$h_t^j = \left(1 - \tau_t + \frac{1 - \tau_{t+1}}{R}\right) w^j. \tag{4.3}$$

Human capital investment increases in ability and decreases in tax rates.

The proportion of the poor is a constant λ in each cohort. We assume $\lambda \geq 1/2$ so that the poor are the majority of the population. Average productivity is normalized to unity so that $w^h = (1 - \lambda w^u) / (1 - \lambda)$. The weighted average wage incomes for agents born at time t, denoted by $\overline{w}^t$, are equal to

$$\overline{w}_t^t = \overline{w}_{t+1}^t = \lambda w^u h_t^u + (1 - \lambda) w^s h_t^s. \tag{4.4}$$

[13] We could then assume that human capital depreciates over time. Then, the wage income for the middle-aged would be equal to $\delta w^j h_t^j$, where $1 - \delta$ is the depreciation rate. The main result would not change under this extension.

[14] See Song (2005a) for further discussion. Like Song (2005a), this paper takes the composition of inter- and intra-generational redistribution aspects as exogenous. We will leave it to future research to endogenize the composition of the two-way redistributive transfers.

The first equality is due to the fact that middle-aged workers have the same productivity as when they are young. The gross population growth rate is a constant $n > 1$. Plugging (4.3) into (4.4), we obtain the output per retiree:

$$\begin{aligned} y_t &= n\overline{w}_t^{t-1} + n^2\overline{w}_t^t \\ &= \Phi\left(\frac{h_{t-1}^u}{w^u} + n\left(1 - \tau_t + \frac{1-\tau_{t+1}}{R}\right)\right), \end{aligned} \tag{4.5}$$

where $\Phi \equiv n\left(\lambda\left(w^u\right)^2 + (1-\lambda)\left(w^h\right)^2\right)$. We use the fact that $h_{t-1}^s/h_{t-1}^u = w^s/w^u$ implied by (4.3). The output per retiree y_t is the current tax base for social security transfers. The future tax base y_{t+1} is determined by h_t^j and h_{t+1}^j, and independent of the current human capital stock h_{t-1}^j.

$$y_{t+1} = \Pi - \Phi\left(\tau_t + \left(n + \frac{1}{R}\right)\tau_{t+1} + \frac{n}{R}\tau_{t+2}\right), \tag{4.6}$$

where $\Pi \equiv \Phi\left(1+n\right)\left(1+\frac{1}{R}\right)$.

We assume that the budget of the social security system must balance in each period. This implies that in each period, the benefits paid to the old generation equal the contributions collected from the working generations:

$$p_t = \tau_t y_t. \tag{4.7}$$

Substituting (4.6) and (4.7) into (4.1), the indirect utility functions of the middle-aged and old of type j, denoted by $v^{m,j}$ and $v^{o,j}$, can be written as

$$v^{m,j}\left(h_{t-1}^j, \tau_t, \tau_{t+1}, \tau_{t+2}\right) = \left(1-\tau_t\right)w^j h_{t-1}^j + \frac{\tau_{t+1}\left(\Pi - \Phi\left(\tau_t + \left(n+\frac{1}{R}\right)\tau_{t+1} + \frac{n}{R}\tau_{t+2}\right)\right)}{R} \tag{4.8}$$

$$v^{o,j}\left(h_{t-1}^u, \tau_t, \tau_{t+1}\right) = \Phi\tau_t\left(\frac{h_{t-1}^u}{w^u} + n\left(1-\tau_t + \frac{1-\tau_{t+1}}{R}\right)\right). \tag{4.9}$$

Note that $v^{o,s} = v^{o,u}$ since social security benefits p_{t+1} are evenly distributed across retirees.

3 Political Equilibrium

The sequence of social security tax rates $\{\tau_t\}_{t=0}^{\infty}$, or equivalently, the sequence of social security benefits $\{p_t\}_{t=0}^{\infty}$, is determined through repeated political decisions at the beginning of each period. More specifically, τ_t is chosen by the decisive voter (e.g. the median voter) by maximizing her indirect utility. We assume that there exists a cap on τ_t, denoted by $\overline{\tau}$. $\overline{\tau} \leq 1$ can be considered as the upper boundary of politically acceptable social security tax rates in the legislative process.

We assume that young agents do not vote. This reflects, though somewhat excessively, the phenomenon that the older are more influential in the determination of public policies.[15] In the Markov perfect equilibrium, τ_t evolves according to a policy rule T that conditions policy choices to the payoff-relevant state variables. There are two state variables at time t, human capital stock h_{t-1}^s and h_{t-1}^u. Since $h_{t-1}^s / h_{t-1}^u = w^s / w^u$ we can, without loss of generality, write the policy rule T as a function of h_{t-1}^u only:

$$\tau_t = T\left(h_{t-1}^u\right), \tag{4.10}$$

where $T : \left[\underline{h}, \overline{h}\right] \to [0, \overline{\tau}]$, $\underline{h} \equiv w^u (1 + 1/R)(1 - \overline{\tau})$ and $\overline{h} \equiv w^u (1 + 1/R)$ are the lower and upper bound of h_{t-1}^u, respectively. Social security benefits at time $t+1$ are not predetermined for the decisive voter at time t. But given the policy rule T, she can strategically affect τ_{t+1} by the distortionary effect of τ_t on h_t^u. Plugging (4.10) into (4.3), we have

$$h_t^u = w^u \left(1 - \tau_t + \frac{1 - T\left(h_t^u\right)}{R}\right). \tag{4.11}$$

Equation (4.11) defines the human capital investment strategy of the young $H : [0, \overline{\tau}] \to \left[\underline{h}, \overline{h}\right]$, which solves

$$H\left(\tau_t\right) = w^u \left(1 - \tau_t + \frac{1 - T \circ H\left(\tau_t\right)}{R}\right). \tag{4.12}$$

Combining (4.10) and (4.12) yields

$$\tau_{t+1} = T \circ H\left(\tau_t\right) \equiv B\left(\tau_t\right), \tag{4.13}$$

with $B : [0, \overline{\tau}] \to [0, \overline{\tau}]$. In the following text, B will be referred to as a social contract characterizing the evolution of the social security system. Then, the political choice τ_t solves

$$\tau_t = \arg \max_{\tau_t \in [0, \overline{\tau}]} v^{dec}, \tag{4.14}$$

[15] For instance, Mulligan and Sala-i-Martin (1999) argue that the old have more influence in the political decision process because they have a lower cost of time. Empirically, voting turnout is indeed lower for younger households (e.g. Wolfinger and Rosenstone, 1980). See Hassler *et al.* (2003) and Song (2005a) for further discussion.

subject to $\tau_{t+1} = B(\tau_t)$ and $\tau_{t+2} = B(B(\tau_t))$. v^{dec} is the indirect utility function of the decisive voter who determines τ_t. (4.8) and (4.9) show that the choice of τ_t only depends on h^u_{t-1}, irrespective of which group is decisive. Define $\tau_t = \tilde{T}\left(h^u_{t-1}\right)$ as the solution of (4.14). T is an equilibrium policy rule if and only if $T = \tilde{T}$.

We assume T and H to be continuous and differentiable. The corresponding equilibrium is referred to as the differentiable Markov perfect equilibrium. For analytical convenience, the assumption of differentiability has often been adopted in recent studies on social security (e.g. Azariadis and Galasso, 2002, Katuscak, 2002, Forni, 2005) and fiscal policy (e.g. Klein *et al.*, 2003).[16] Later we will see that the differentiability also helps pin down the identity of the median voter in dynamic politico-economic equilibrium. The definition of the equilibrium is given by

Definition 1 *A (differentiable) Markov perfect political equilibrium is a pair of differentiable functions $\langle T, H\rangle$, where $T : \left[\underline{h}, \overline{h}\right] \to [0, \overline{\tau}]$ is the policy rule of the social security tax rate and $H : [0, \overline{\tau}] \to \left[\underline{h}, \overline{h}\right]$ is a private decision rule of human capital investment. T and H solve the following functional equations:*

(1) $T\left(h^u_{t-1}\right) = \arg\max_{\tau_t \in [0,\overline{\tau}]} v^{dev}$, subject to $\tau_{t+1} = T \circ H(\tau_t)$ and $\tau_{t+2} = T \circ H(T \circ H(\tau_t))$.

(2) $H(\tau_t) = w^u \left(1 - \tau_t + \frac{1 - T\circ H(\tau_t)}{R}\right)$.

We will show that in the class of differentiable Markov perfect equilibria, the middle poor are always decisive for τ_t under majority voting. This is in contrast to the approach that the old play at least some roles on the political decision of social security benefits.[17] For expositional reasons, we first solve a dictatorship equilibrium where the political power rests in the hands of the middle poor. Later, we show that the equilibrium under dictatorship in fact coincides with the equilibrium under majority voting.

[16] Krusell and Smith (2003) show that there could be in principle an infinitely large number of nondifferentiable Markov equilibria. Markov Equilibria with non-differentiable strategies may exist in our model. We will however not attempt to characterize such equilibria.

[17] See for example the probabilistic voting framework adopted by Katuscak (2002), Gonzalez-Eiras and Dirk Niepelt (2004) and Song (2005a).

3.1 Dictatorship

We define "dictatorship of the middle poor" (DMP) as the regime where the social security tax rate is chosen at the beginning of each period by the middle poor. It immediately follows that the policy rule $T(h_t^u) = 0$ for all $h_t^u \in [\underline{h}, \overline{h}]$ trivially satisfies the conditions in Definition 1. This is referred to as the trivial equilibrium. We say that an equilibrium is nontrivial if $T(h_t^u) \neq 0$ at least for some $h_t^u \in [\underline{h}, \overline{h}]$.

Given a nontrivial equilibrium social contract B, the future tax base (4.6) can be written in a recursive fashion

$$y_{t+1} = Y(\tau_t) = \Pi - \Phi\left(\tau_t + \left(n + \frac{1}{R}\right) B(\tau_t) + \frac{n}{R} B(B(\tau_t))\right). \tag{4.15}$$

Differentiating $Y(\tau_t)$ yields $Y'(\tau_t) = -\Phi\left(1 + (n + 1/R) B'(\tau_t) + n (B'(\tau_t))^2 / R\right) < 0$. The inequality holds under any B. That is to say, increasing τ_t always reduces the future tax base y_{t+1}. Maximizing the indirect utility of the middle poor, (4.14) solves

$$w^u h_{t-1}^u = \frac{B'(\tau_t) Y(\tau_t) + B(\tau_t) Y'(\tau_t)}{R}, \tag{4.16}$$

where the LHS of (4.16) captures the marginal cost of the social security tax rate, while the marginal benefit is given by the RHS. We suppress two multipliers associated with the constraints $\tau_t \geq 0$ and $\tau_t \leq \bar{\tau}$, since corner solutions can easily be ruled out by the differentiability of $T(\cdot)$.[18] The following lemma establishes the monotonicity in the nontrivial equilibrium.

Lemma 1 *In the nontrivial equilibrium under DMP, we have* $T'(\cdot) < 0$, $H'(\cdot) < 0$ *and* $B'(\cdot) > 0$.

Proof: Suppose that $B'(\tau_t) \leq 0$ for some $\tau_t \in (0, \bar{\tau})$. Since $Y'(\cdot) < 0$, (4.16) cannot hold. Contradiction. $H'(\cdot) < 0$ and $T'(\cdot) < 0$ immediately follow from equations (4.12) and (4.13). □

The negative sign of $H'(\cdot)$ is due to the distortionary effect of taxes on human capital investment. The negative sign of $T'(\cdot)$ stems from the fundamental impact

[18] If the constraints are binding for some $h_{t-1}^u \in (\underline{h}, \overline{h})$, then τ_t must be equal to 0 or $\overline{\tau}$ for all $h_{t-1}^u \in [\underline{h}, \overline{h}]$ by the differentiability of T. $T(\cdot) = 0$ is a trivial equilibrium. If instead $T(\cdot) = \overline{\tau}$, $B'(\cdot) = 0$. Then (4.16) implies that $\tau_t = 0$. Contradiction.

of the decisive voter's human capital stock on his political choice. Intuitively, a low human capital h_{t-1}^{u} is associated with a low marginal cost of τ_t, as captured by the LHS of (4.16). This induces the decisive voter to raise the tax rate. Combining with $H'(\cdot) < 0$, we obtain a positive relationship between τ_{t+1} and τ_t. That is to say, the more the working generations contribute to the social security system, the more they can expect to receive from the system in the future. The corresponding positive $B'(\cdot)$ relates current contributions to future benefits. This solves the temporal separation problem of benefits and contributions and thus, provides a necessary condition for the decisive voter to support the social security system since otherwise, she would rationally choose zero contribution.

Under the quasi-linear preference, the equilibrium can be solved analytically.

Proposition 1 *(1) A nontrivial DMP equilibrium exists if*

$$\begin{aligned} \phi_0 + \phi_1 \overline{h} &\geq 0, & (4.17)\\ \phi_0 + \phi_1 \underline{h} &\leq \bar{\tau}, & (4.18) \end{aligned}$$

where $\phi_0 \equiv -(b_1\pi_0 + b_0\pi_1)/(2b_1\pi_1) > 0$ *and* $\phi_1 \equiv Rw^u/(2b_1\pi_1) < 0$ *(see the appendix for the definition of* b_0*,* b_1*,* π_0 *and* π_1*).*

(2) Assume (4.17) and (4.18). There exists a unique linear nontrivial DMP equilibrium such that

$$\begin{aligned} T\left(h_{t-1}^{u}\right) &= \phi_0 + \phi_1 h_{t-1}^{u}, & (4.19)\\ H(\tau_t) &= \frac{w^u}{1 + w^u\phi_1/R}\left(1 + \frac{1-\phi_0}{R} - \tau_t\right), & (4.20)\\ B(\tau_t) &= b_0 + b_1\tau_t, & (4.21)\\ Y(\tau_t) &= \pi_0 + \pi_1\tau_t, & (4.22) \end{aligned}$$

where $b_0 > 0$*,* $b_1 > 0$*,* $\pi_0 > 0$ *and* $\pi_1 < 0$*.*

(3) Assume (4.17) and (4.18). In the nontrivial DMP equilibrium, we have $b_0 > 0$ *and* $b_0 + b_1 < \overline{\tau}$*. The social security tax rates monotonically converge to the steady state* $b_0/(1-b_1) \in [0, \overline{\tau}]$*.*

Proof: See the appendix.

The first part of the proposition gives conditions for the sustainability of social security under DMP. Given (4.17), the middle poor would choose a non-negative social security tax rate for any $h_{t-1} \in [\underline{h}, \overline{h}]$. (4.18) ensures that $\tau_t \leq 1$ is not binding.[19] It turns out that (4.17) and (4.18) can be satisfied under a wide variety of parameter values. An example is plotted in Figure 1, where we set $R = 1.04^{30}$ and $\lambda = 0.6$. Condition (4.18) fails to be satisfied only for very large values of n (e.g. $n > 20$ outside the range in the figure). (4.17) holds for all w^u and n in the diagram of Figure 1, except the triangle with high w^u and low n. An immediate observation is that the social security system under DMP can exist in a dynamic efficient economy ($R > n$ in our model), as long as income inequality is not too low, i.e., w^u is not too high. The classical solution to the existence of the social security system is that if the economy is dynamically inefficient, the introduction of the system will be Pareto-improving (e.g. Aaron, 1966) since it provides a higher return than the financial market. The sustainability of social security in a dynamic efficient economy in the current environment is twofold. First, the social contract conditions future benefits to the current contributions. As shown above, this gives the necessary condition for the sustainability of the pension system. More importantly, a higher income inequality makes the poor more willing to support the system for intra-generational redistributive reasons, since it implies a higher return to the poor from the social security system. That is why the system can be sustained in a dynamic efficient economy with sufficiently high income inequality.[20]

[19] Since $T'(\cdot) < 0$, the minimum (maximum) value of τ_t is located at $h_{t-1} = \overline{h}$ ($\underline{h}$).

[20] As mentioned in the introduction, Forni (2005), Gonzalez-Eiras and Dirk Niepelt (2004) study the political sustainability of the pension system in a general equilibrium framework without intra-cohort heterogeneity. In their models, the social security system can also be sustained in a dynamic efficient economy. The sustainability comes from the general equilibrium effect, however. The working generation is willing to accept a positive payroll tax rate since it can discourage capital accumulation and increase the future interest rate.

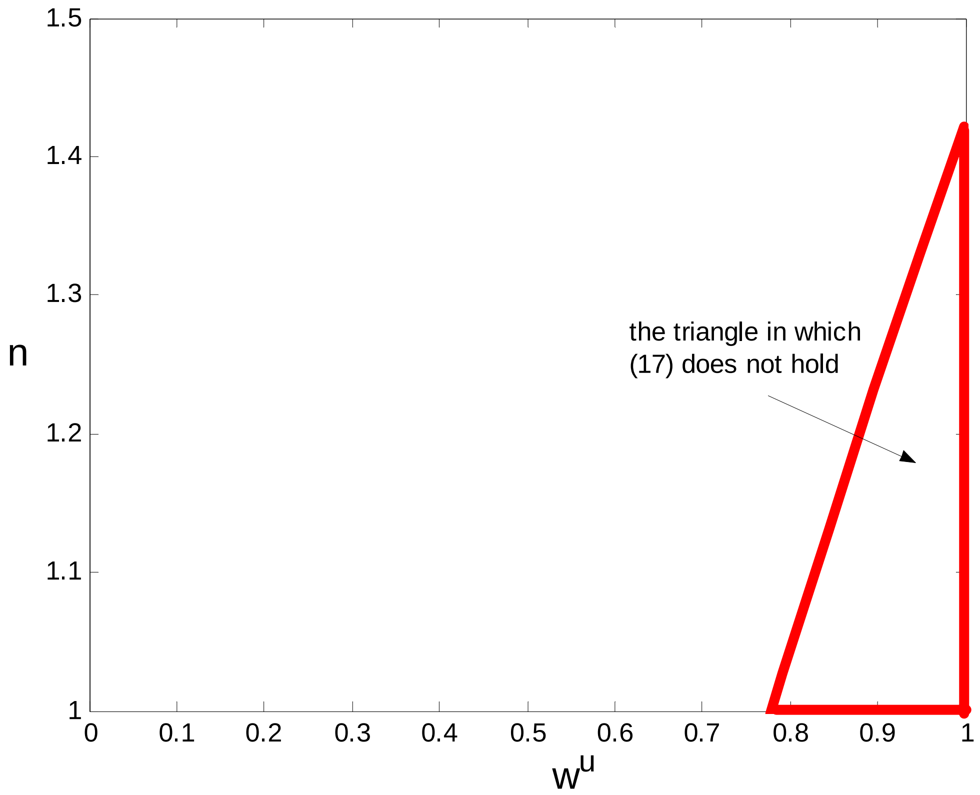

Figure 1: The Existence Condition of the Social Security System under DMP. We set $R = 1.04^{20}$ and $\lambda = 0.60$. (17) and (18) hold for all n and w^u in the diagram except in the triangle in which (17) does not hold.

The second part of the proposition gives the uniqueness of the linear Markov perfect equilibrium. Although we cannot analytically rule out nonlinear solutions, numerical simulations show that the equilibrium strategy in the space of nonlinear functions indeed converges to the linear one (4.19).[21] The uniqueness sharply contrasts the multiplicity of linear equilibria in Azariadis and Galasso (2002), though their model has a number of features in common with ours (e.g. the linear technology and the temporal separation of contributions and benefits). Azariadis and Galasso construct a policy rule contingent on the capital held by old households. However, due to the linear technology, old households' capital is

[21] The computational strategy adopts a standard projection method with Chebyshev collocation (Judd, 1992), to approximate the equilibrium strategy on the basis of a high-order polynomial functional space.

not payoff-relevant for the decisive voter, i.e. the young households in their model. Therefore, the sustainability of social security solely relies on the self-fulfilled expectation of future social security benefits that are essentially indeterminate. The policy rule T in our model is instead a rigorous Markovian strategy in the sense that human capital is payoff-relevant for the middle poor.

The third part of the proposition characterizes the evolution of a social security system under DMP. Suppose that there is no social security taxes before time 0 and DMP arises unanticipatedly at time 0. Then, we have $h_0^u = \overline{h}$ and $T\left(\overline{h}\right) < b_0/\left(1 - b_1\right)$. This implies that $\{\tau_t\}_{t=0}^{\infty}$ be an increasing sequence. The underlying mechanism is in the following. The initial middle poor have more human capital than the middle poor in subsequent periods. Since taxes are increasing in human capital, the initial median voter will choose the lowest tax rate. Similarly, since the second period median voter has less human capital than the initial period median voter but more human capital than the median voters in the following periods, she will impose a tax rate that is higher than the initial tax rate but lower than the future tax rates. The increasing size of the social security system is implied by the positive $B'(\cdot)$, which is an intrinsic feature in a dynamic political equilibrium under DMP. Without the positive link b_1 between current contributions and future benefits, the coordination failure among generations that has its roots in the temporal separation problem would destroy the welfare state. Therefore, the growing social security can be regarded as a natural outcome of the incentive mechanism which props the political support of the middle-aged for inter-generational transfers.

3.1.1 Expectational Stability

In the Markovian social contract, $b_1 = \mathrm{d}\tau_t/\mathrm{d}\tau_{t-1}$ can be referred to as the equilibrium return of social security benefits to the contributions. We have shown the positive b_1 to be crucial for the sustainability and evolution of the social security system. The rational expectation of b_1 commands full knowledge of the structure of the model. However, it appears more natural for agents to have limited knowledge about the economy. Therefore, we are interested in knowing whether the equilibrium outcome b_1 can be achieved through some learning process. To this end, we construct a temporary equilibrium framework with "artificial time" $\hat{t}$ to conduct the expectational stability (E-stability) analysis (Evans and Honkapohja, 2001). If b_1 is expectationally unstable, it would be hard to believe that the Markovian social contract in Proposition 1 can be realized and have any practical relevance.

Denote $\tilde{b}_1$ as the return of contributions expected by the agents. The first-order condition (4.16), after some algebra, implies that the middle poor set the tax rate as

$$\tau_t = C + F\left(\tilde{b}_1\right)\tau_{t-1}, \tag{4.23}$$

where

$$F\left(\tilde{b}_1\right) \equiv \frac{R}{2\left(\lambda + (1-\lambda)\left(1/w^u - \lambda\right)^2\right)\left(1 + (n + 1/R)\,\tilde{b}_1 + n\tilde{b}_1^2/R\right)n\tilde{b}_1 - 1} \tag{4.24}$$

and C is an unimportant constant.[22] (4.23) shows that, given the past social security contribution τ_{t-1}, the current policymaker would like to provide social security benefits $C + F\left(\tilde{b}_1\right)\tau_{t-1}$. So the actual return for the past contributions, denoted by $\hat{b}_1$, is equal to

$$\hat{b}_1 = F\left(\tilde{b}_1\right). \tag{4.25}$$

$F : [0,1] \to [0,1]$ is the mapping from the expected return to the actual return. The equilibrium return b_1 corresponds to the fixed point of F. It is easily shown that $F'(\cdot) < 0$. As $\tilde{b}_1$ increases, (4.15) implies that the future tax base become more elastic with respect to τ_t. This is because under a high $\tilde{b}_1$, future tax rates are more responsive to τ_t. The increased elasticity induces the policymaker to cut τ_t and hence, leads to a low actual return for the past social security contribution, τ_{t-1}. According to Evans and Honkapohja (2001), the equilibrium return b_1 is said to be expectationally stable (E-stable) if the differential equation

$$\frac{db_1}{d\hat{t}} = F(b_1) - b_1$$

is asymptotically stable. In the present model, $F'(\cdot) < 0$ gives the E-stability of b_1. Since the other coefficients in (4.19), (4.21) and (4.22) are determined by b_1 (see the appendix), we have

Proposition 2 *The nontrivial DMP Equilibrium is E-stable.*

[22] We use the definition of ϕ_1 and π_1 provided in the appendix and substitute (4.20) for h_{t-1}^u.

Proposition 2 says that even if individuals cannot "compute" b_1, they can adjust their expected return $\tilde{b}_1$ according to the actual return provided to the current retirees $\hat{b}_1$. The adjustment leads to an expected return that converges to the equilibrium. As will be seen in Section 4, the interaction between the expected and the actual return in (4.23) helps reveal the mechanism on how the equilibrium return b_1 is affected by demographic structure and income inequality.

3.2 Majority Voting

Now we turn to majority voting, where the median voter is decisive for τ_t. Before proceeding, it is worth pointing out that the strategic interaction between private intertemporal choice and the policy decision can in principle switch the identity of the median voter over time. The dynamic politico-economic equilibrium is hence very hard to characterize (the exceptions are Hassler *et al.*, 2003 and Hassler *et al.*, 2004, Song, 2005b). In this paper, we focus on the differentiable Markovian strategies. The differentiability helps rule out the time-varying identity of the median voter and makes the analysis substantially simpler.

Lemma 2 *The identity of the median voter is constant over time in the differentiable Markov perfect equilibrium.*

Proof: See the appendix.

The intuition is the following. When the identity of the median voter switches over time, we will observe different political regimes in the equilibrium. This leads to a discontinuous or nondifferentiable policy rule T (see, for example, Hassler *et al.*, 2003). Therefore, the technical restriction that the policy function must be differentiable actually amounts to imposing a constant identity of the decisive voters over time, irrespective of the current policy choice. Next, we show that only the middle poor can be the median voter. The argument consists of two parts. First, we check if the middle poor can be the median voter. Then, we rule out the possibility for the middle rich or the old of being decisive in the majority voting.

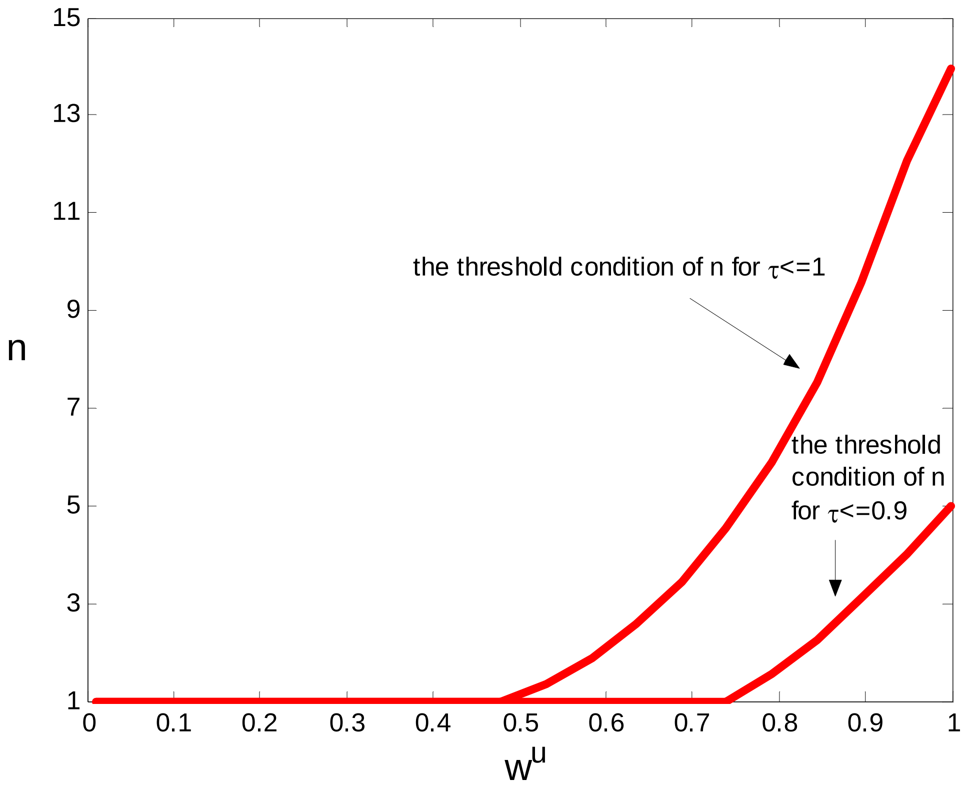

Figure 2: The Threshold Condition of n Implied by (27) with Different Upper Bound of Social Security Tax Rate. We set $R = 1.04^{20}$ and $\lambda = 0.60$.

Let $\tau_t^{m,s}$, $\tau_t^{m,u}$ and τ_t^{o} be the preferred tax rate of the middle rich, the middle poor and the old, respectively. The equilibrium under DMP survives the majority voting if and only if the median voter is always the middle poor, i.e.,

$$\tau_t^{m,s} \leq \tau_t^{m,u} \leq \tau_t^{o} \tag{4.26}$$

for any $h_{t-1}^{u} \in \left[\underline{h}, \overline{h}\right]$. The first inequality is straightforward. Since the middle rich obtain the same social security benefits as the middle poor, but must pay higher taxes $w^s h_{t-1}^s \tau_t$, the middle rich always prefer a lower tax rate than the middle poor. So we only need to check the second inequality in (4.26). Consider the following inequality:

$$\phi_0 + \phi_1 \underline{h} \leq \frac{1 + (1 - b_0)/R + \underline{h}/(n w^u)}{2(1 + b_1/R)}. \tag{4.27}$$

The LHS of (4.27) is the highest tax rate for which the middle poor would vote. In the appendix, we show that the lowest tax rate for which the old would vote is equal to the RHS of (4.27). Condition (4.27) implies that $\tau_t^{m,u} \leq \tau_t^o$ for any $h_{t-1}^u \in \left[\underline{h}, \overline{h}\right]$ and thus ensures that the political equilibrium under repeated majority voting replicates that under DMP. Note that (4.27) is more likely to be satisfied by a higher $\underline{h}$ or a lower $\overline{\tau}$. This is because $\tau_t^{m,u}$ deceases in h_{t-1}, while the old tend to impose a higher tax rate with a larger inelastic human capital stock h_{t-1}. Figure 2 plots the threshold condition of n implied by (4.27) for $\overline{\tau} = 1$ and 0.9, respectively. (4.27) is satisfied for any n above the line in the figure. It can be seen that the region satisfying (4.27) is substantially enlarged as $\overline{\tau}$ decreasing from 1 to 0.9. For $\overline{\tau} < 0.86$, (4.27) holds for all w^u and n in the diagram. The reason for $\tau_t^{m,u} \leq \tau_t^o$ is twofold. Increasing τ_t raises the tax burden of the middle-aged, while the tax burden of the old is always zero. Moreover, the elasticities of the current tax base y_t for the old and the future tax base y_{t+1} for the middle-aged are different with respect to τ_t. Increasing τ_t does not only discourage h_t^u, but also h_{t+1}^u. So y_{t+1} tends to be more elastic than y_t, as long as the current human capital stock h_{t-1}^j is not too low.

If condition (4.27) is violated, the middle poor would not be the median voter. The natural question is thus which group can otherwise be decisive in the majority voting. In the appendix, we prove that neither the old nor the middle rich can be the median voter, under some additional restrictions on the parameter values. The intuition is as follows. Suppose that the old are decisive. They would vote for a higher τ_t as h_{t-1}^u is increasing. This leads to a positive ϕ_1 and a negative b_1, which aggravates the temporal separation problem of benefits and contributions and induces the middle-aged (the majority of the voters) to vote for zero taxes. Hence, the old cannot be the median voter. Then, we consider the dictatorship of the middle rich (DMR), i.e., the middle rich are always decisive. Since the first inequality of (4.26) must hold, the middle rich would be the median voter in the majority voting if and only if $\tau_t^o \leq \tau_t^{m,s}$ for any $h_{t-1}^u \in \left[\underline{h}, \overline{h}\right]$. This gives the condition

$$\frac{1 + 1/R - b_0^{m,s}/R + \overline{h}/nw^u}{2\left(1 + b_1^{m,s}/R\right)} > \phi_0^{m,s} + \phi_1^{m,s}\overline{h} \tag{4.28}$$

which rejects the middle rich to be the median voter. The derivation of (4.28) and the definitions of $b_0^{m,s}$, $b_1^{m,s}$, $\phi_0^{m,s}$ and $\phi_1^{m,s}$ are provided in the appendix. The

LHS and the RHS of (4.28) are the highest and the lowest tax rate preferred by the old and the middle rich under DMR, respectively. Condition (4.28) holds in all numerical experiments we have done so far, since the old tend to impose the highest social security tax rate while the middle rich impose the lowest.[23] The above analysis establishes

Proposition 3 *Assume (4.17), (4.18), (4.27) and (4.28). Then the median voter in any (differentiable) Markov perfect equilibrium must be the middle poor. The equilibrium outcomes are equivalent to those under the DMP equilibrium, as stated in proposition 1.*

4 Comparative Static Analysis

Since the social security system works as a redistributive policy among and within cohorts, we are particularly interested in the impact of demographic structure and income inequality on the size of the system. Before proceeding, let us first investigate the impact of w^u and n on two parameters b_0 and b_1 in the social contract B that characterizes the size of social security in the dynamics.

Proposition 4 *Assume (4.17), (4.18), (4.27) and (4.28). Then $\partial b_1/\partial n < 0$ and $\partial b_1/\partial w^u > 0$.*

Proof: See the appendix.

Previous research suggests the size of social security to be increasing in population growth and income inequality. Since the return b_1 tends to increase the size of social security, the negative correlation between b_1 and population growth or income inequality may seem counter-intuitive. However, we argue that it arises as a natural outcome in the politico-economic equilibrium. To see this, let us look at the decision-making of τ_t in (4.23). Given the expected return $\tilde{b}_1$, the actual return of the past contributions decided by the middle poor is equal to $F\left(\tilde{b}_1\right)$. It immediately follows that $F\left(\tilde{b}_1\right)$ is negatively and positively correlated with n and w^u, respectively. Since the future tax base y_{t+1} increases in the future labour supply, a higher population growth rate n makes y_{t+1} more elastic with respect

[23] For example, (4.28) holds for any n and w^u in the diagram of Figure 1 or 2.

to τ_t, as implied by (4.15). The median voter thus will set a lower τ_t, i.e., a lower actual return for the past contribution τ_{t-1}. The dotted line in Figure 3 plots the new $F\left(\tilde{b}_1\right)$ under a higher n. It can directly be seen directly that this moves $F\left(\tilde{b}_1\right)$ downwards and results in a lower equilibrium return b_1. The effect of w^u on b_1 is analogous. Since the average wage rate is normalized to unit, a high w^u leads to a low w^h. Since a low w^h discourages the human capital investment of the rich more than the increased investment of the poor by a high w^h, the future tax base y_{t+1} and the elasticity of y_{t+1} are decreasing in w^u. This induces the median voter to increase τ_t, i.e., a high actual return for τ_{t-1}. So there is a positive relationship between $F\left(\tilde{b}_1\right)$ and w^u, which leads to $\partial b_1/\partial w^u > 0$. This can be seen from the dashed line in Figure 3, which plots the new $F\left(\tilde{b}_1\right)$ under a higher w^u. Moreover, when w^u is very low, increasing w^u (or equivalently reducing income inequality) sharply reduces the future tax base y_{t+1}. This implies that as w^u is increasing, the elasticity of y_{t+1} tends to decline more rapidly for low w^u than for high w^u. Hence, the effect of w^u on b_1 turns out to be highly nonlinear. When w^u is low, increasing w^u can substantially raise the equilibrium return b_1. The effect wipes off as w^u approaches one.[24]

[24] Numerical simulation shows that $\partial b_1/\partial n$ is much smoother than $\partial b_1/\partial w^u$.

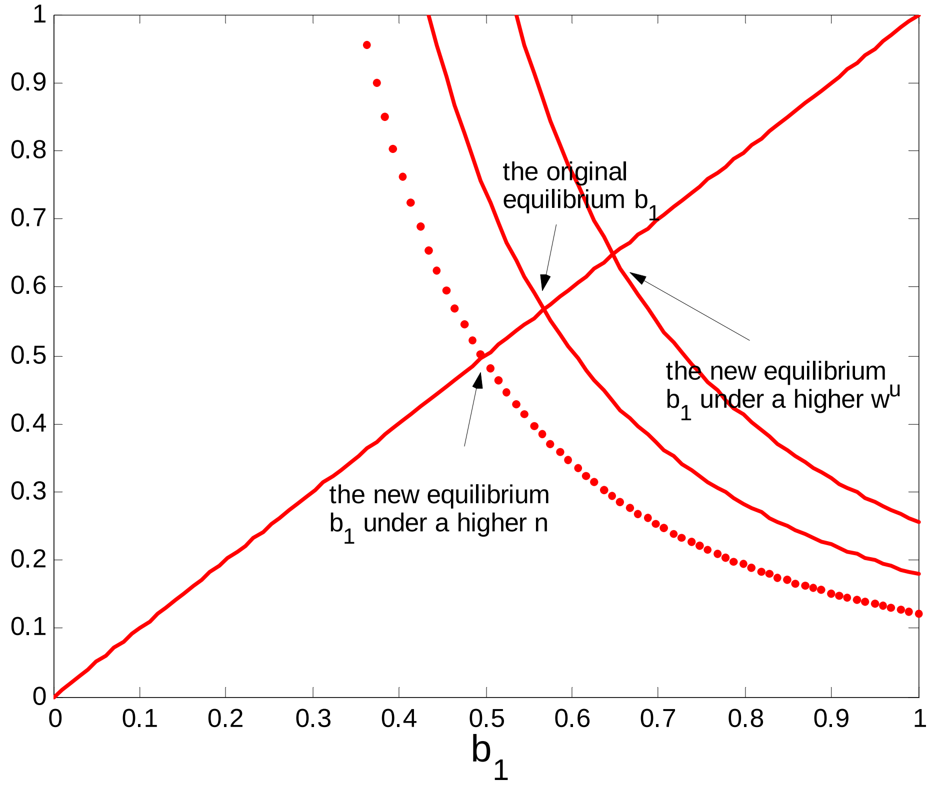

Figure 3: The Determination of the Equilibrium Return of Social Security Contribution b_1. The solid line, dotted line and dashed line refer to $F(b_1)$ under the benchmark parameter values, a higher n and a high w^u, respectively. We set $R = 1.04^{20}$, $\lambda = 0.60$, $n = 1.384$ and $w^u = 0.5$ in the benchmark and let n and w^u increase by 25% each in the two other cases.

The impact of n and w^u on the constant term b_0 in the social contract is more complex. There are two effects:

$$\frac{\partial b_0}{\partial x} = \left.\frac{\partial b_0}{\partial x}\right|_{b_1} + \frac{\partial b_0}{\partial b_1}\frac{\partial b_1}{\partial x} \tag{4.29}$$

where $x = w^u$ or n. The first term on the RHS of (4.29) is the direct effect of x, taking the equilibrium return b_1 as a parameter. The second term captures the indirect effect through b_1. Due to the complexity of b_0 (see the appendix), it is hard to get analytical details on $\partial b_0/\partial x$. So we resort to numerical experiments. It turns out that the following properties hold under all experiments we have done so far.

$$\left.\frac{\partial b_0}{\partial n}\right|_{b_1} > 0 \tag{4.30}$$

$$\left.\frac{\partial b_0}{\partial w^u}\right|_{b_1} < 0 \tag{4.31}$$

$$\frac{\partial b_0}{\partial b_1} > 0 \tag{4.32}$$

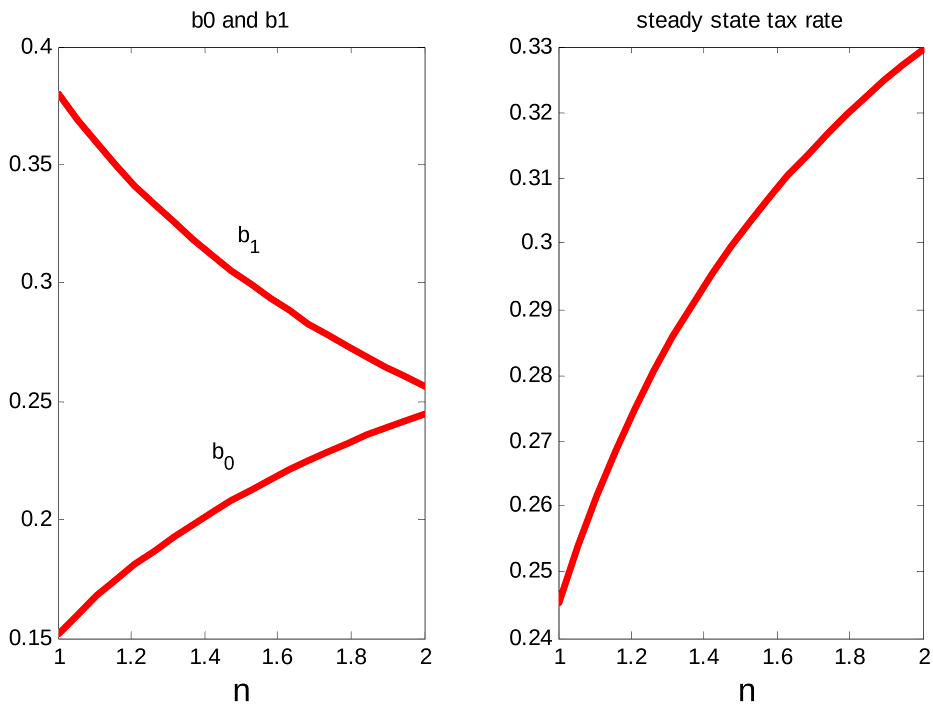

Figure 4: The Effects of n on the Markov Perfect Political Equilibrium. We set $w^u = 0.5$, $R = 1.04^{20}$ and $\lambda = 0.60$.

Intuitively, taking b_1 as a parameter, a low n or a high w^u leads to a low level of inter- or intra-generational transfers, respectively. This produces a low constant term b_0 in the social contract as shown by (4.30) and (4.31). On the other hand, b_0 is increasing in b_1. A high return provides incentives for the taxpayers to increase τ_t. This results in a positive $\partial b_0/\partial b_1$. Since $\partial b_1/\partial w^u > 0$ and $\partial b_1/\partial n < 0$, the indirect effect is opposite to the direct effect. Numerical simulations show that the direct effect of n always dominates the indirect effect. Correspondingly, b_0 is monotonically increasing in n. Figure 4 gives an example. From Figure 5, we see that the indirect effect of w^u dominates the direct effect for low w^u while the direct effect dominates for high w^u. So b_0 turns out to be an inverted-U function of w^u. Compared with the dominated indirect effect of n, the relatively stronger indirect effect of w^u for low w^u should not be very surprising. As shown above, $\partial b_1/\partial w^u$ is highly nonlinear. For a low w^u, the effect of w^u on b_1 is sufficiently large to outweigh the direct effect.

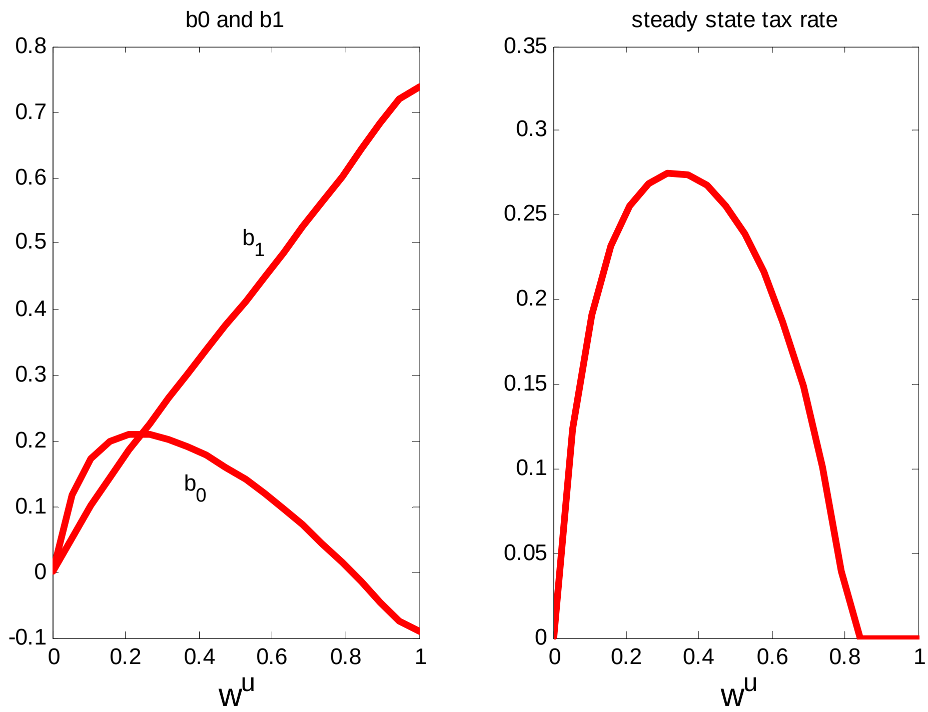

Figure 5: The Effects of w^u on the Markov Perfect Political Equilibrium.

4.1 The Static and Social Contract Effects on the Size of Social Security

In this subsection, we distinguish two channels through which n and w^u may affect the size of social security. These are referred to as the static effect and the social contract effect, respectively. The first-order condition (4.16) implies that the optimal tax rate for the middle poor at time t equals

$$\tau_t = \frac{-\pi_0\left(b_0, b_1\right) b_1 - \pi_1\left(b_1\right) b_0 + R w^u h_{t-1}^u}{2\pi_1\left(b_1\right) b_1} \tag{4.33}$$

Recall that b_0 and b_1 are the constant term and the return of contributions in the social contract (4.21), respectively. π_0 and π_1, determined by b_0 and b_1 (see the appendix), are the constant term and the marginal impact of τ_t on the future tax base in (4.22), respectively. The impact of n and w^u on τ_t can be written as

$$\frac{\partial \tau_t}{\partial x} = \underbrace{\frac{\partial \tau_t}{\partial x}|_{b_0,b_1}}_{\text{the static effect}} + \underbrace{\frac{\partial \tau_t}{\partial b_0}\frac{\partial b_0}{\partial x} + \frac{\partial \tau_t}{\partial b_1}\frac{\partial b_1}{\partial x}}_{\text{the social contract effect}} \tag{4.34}$$

We define the impact of x under fixed b_0 and b_1, $\partial \tau_t / \partial x\,|_{b_0,b_1}$, as the *static effect* in the sense that it abstracts the impact of x on future policy outcomes via the endogenous social contract B. The abstracted impact is captured by the last two terms on the RHS of (4.34), which are referred to as the social contract effect. Let us investigate these two effects in order.

Taking b_0 and b_1 as parameters, the decision of τ_t reduces to a static problem and the size of social security turns out to be positively correlated with n and negatively correlated with w^u.[25] Hence, the static effect implies that the size of social security be larger in the economy associated with higher population growth rate or income inequality. This coincides with the conventional wisdom on the determination of social security.

Table 1: The Effects of n and w^u on the Size of Social Security

	n	w^u
the static effect	$+$	$-$
the social contract effect	$-$	$+$

The static effect is just one side of the story. The social contract *per se* is an endogenous equilibrium outcome. The social contract effect is twofold, through b_0 and b_1. The last term on the RHS of (4.34), $(\partial \tau_t / \partial b_1) \, / \, (\partial b_1 / \partial x)$, reflects the social contract effect of x through b_1. Intuitively, a high return b_1 induces more social security contributions.[26] Together with Proposition 4, $(\partial \tau_t / \partial b_1)\,(\partial b_1 / \partial x)$ should be negative and positive for $x = n$ and w^u, respectively. That is to say, a larger proportion of retirees or higher income inequality tends to reduce the size of social security according to the social contract effect through b_1.

[25] Numerical experiments confirm $\partial \tau_t / \partial n\,|_{b_0,b_1} > 0$ and $\partial \tau_t / \partial w^u\,|_{b_0,b_1} > 0$.

[26] Numerical experiments confirm $\partial \tau_t / \partial b_1 > 0$.

The social contract effect through b_0, i.e. $(\partial\tau_t/\partial b_0)(\partial b_0/\partial x)$, is more complicated. A higher constant term in the social contract b_0 implies higher future tax rates, which increase the elasticity of the future tax base.[27] The higher elasticity induces the policymaker to cut taxes, which gives $\partial\tau_t/\partial b_0 < 0$. The sign of the social contract effect through b_0 is thus the opposite to the sign of $\partial b_0/\partial x$. We have shown that $\mathrm{sgn}(\partial b_0/\partial x)$ is determined by two opposite effects. For $x = n$, the direct effect always dominates and hence, $(\partial\tau_t/\partial b_0)(\partial b_0/\partial n) < 0$. Therefore, the overall social contract effect of n via b_0 and b_1 is negative. For $x = w^u$, $\mathrm{sgn}(\partial b_0/\partial w^u)$ depends on w^u: $(\partial\tau_t/\partial b_0)(\partial b_0/\partial w^u) < 0$ for small w^u and $(\partial\tau_t/\partial b_0)(\partial b_0/\partial w^u) > 0$ for large w^u. Numerical simulations show that the negative effect through b_0 for small w^u is dominated by the positive effect via b_1. The overall social contract effect of w^u is therefore positive. Table 1 summarizes the above analyses. To conclude, the social contract effect that has largely been neglected in the literature turns out to be opposite to the static effect, which reflects conventional wisdom in determining social security.

The aggregate impact of n or w^u on the size of social security is not obvious due to the conflicting effects stated in Table 1. Particularly, we find numerically that the steady state social security tax rate $\tau^* \equiv b_0/(1-b_1)$ is increasing in n and non-monotonically related to w^u, see Figures 4 and 5 for an example. Figure 4 shows that the positive static effect of n dominates the negative social contract effect in the steady state. While Figure 5 shows that there is an inverted-U shaped correlation between w^u and τ^*. That is to say, the positive social contract effect of w^u dominates the negative static effect under sufficiently high income inequality. The static effect starts to outweigh the social contract effect when income inequality is below some critical level.

We have distinguished the static and social contract effect. The social contract effect does not only have a quantitative impact, but can qualitatively change the relationship between the size of social security and income inequality. The relatively stronger social contract effect of w^u for low w^u is not surprising. As discussed before, the effect of w^u on the equilibrium return b_1 is substantial when income inequality is sufficiently large.

[27] Formally, we have $\partial\pi_0/\partial b_0 < 0$ (see the definition of π_0 in the appendix).

4.2 The Social Contract Effect on the Growth of Social Security

The growth of social security g_t is defined as

$$g_t = \ln \tau_t - \ln \tau_{t-1} \tag{4.35}$$

Solving the first-order difference equation $\tau_t = b_0 + b_1 \tau_{t-1}$, we have $\tau_t = b_0 \left(1 - b_1^t\right) / \left(1 - b_1\right) + b_1^t \tau_0$. If $b_1^t \tau_0$ is close to zero, the growth rate of τ_t can be approximated by

$$g_t \approx \ln \left(\frac{1 - b_1^t}{1 - b_1^{t-1}} \right) \tag{4.36}$$

It can be shown that the RHS of (4.36) increases in b_1 for any positive integer t.[28] Together with Proposition 4, we find g_t to be negatively correlated with n, but positively correlated with w^u. This relationship is a natural implication of the equilibrium social contract discussed above. It also implies that the diversified growth patterns of social security systems across countries can be explained by the various levels of dependency ratio and income inequality. The distinctive prediction on the correlation between g_t and n or w^u helps distinguish between the present model and many others. Below, we will see that the prediction of our model is broadly consistent with the data from OECD countries.

5 Empirical Evidence

In this section, we would like to examine whether the predictions of our model are consistent with empirical evidence. Recall that there are three main predictions of the model.

(1) The size of social security increases over time.

(2) The growth of social security is positively correlated with the dependency ratio but negatively correlated with income inequality.

(3) The size of social security in the long run is decreasing in the dependency ratio but non-monotonically related to income inequality.

The first prediction is in line with the increasing sizes of social security during the postwar period, which is believed to be a stylized fact of the evolution of

[28] Numerical experiments show that the approximation (4.36) is reliable, i.e., the growth rate of τ_t is positively correlated with b_1.

social security (see Mulligan and Sala-i-Martin, 1999a for more details). From Table 2, we see that both the absolute and the relative size of social security grow steadily in the OECD countries over time. In particular, real benefits per social security beneficiary in 1990 amount to seven thousand USD, three times more than in 1960. Public social security per social security beneficiary has also approximately triple from $1,546 in 1960 to $4,653 in 1980.[29]

Table 2: Social Security Program in the OECD (Averages)

	1960	1970	1980	1990
Benefits Per GNP	0.0633	0.0708	0.110	0.119
Public Pensions Per GNP	0.0462	0.0588	0.0924	-
Real Benefits Per Capita	0.293	0.530	0.976	1.362
Public Pension Per Capita	0.228	0.433	0.810	-
Real Benefit Per Pensioner	1.962	3.186	5.713	7.157
Public Pension Per Pensioner	1.546	2.583	4.653	-

Source: Breyer and Craig (1997, table 2). Absolute amount in 1000 of 1982 US Dollars.

We use data from OECD countries over the period 1960-1985 to test the second prediction. We run a cross-country regression of the average growth rate of real social security benefits per beneficiary during 1960-1985 on both dependency ratio and income inequality. The sample countries are divided into two groups: G7 and small OECD countries (the definition of small OECD countries follows the OECD Analytical Database). The measure of income inequality is the average Gini coefficient. The dependency ratio is defined as the ratio of the population aged above 65 to the population aged 15-64. The description of the data source is provided in the appendix. However, the remarkable demographic changes in many OECD countries over the past four decades are not consistent with the prerequisite of our cross-section regression, where the dependency ratio should be rather stable during the relevant period. So we collect the data of the dependency ratio in 1960 and 1980 and run regressions on the dependency ratio of these two

[29] Note that the growth of GDP per capita in OECD countries is far below the growth of pension benefits and hence, cannot explain the rapid expansion of the social security system. See Song (2005a) for further discussions.

years separately, to see if there is any significant discrepancy in the estimated coefficients.

Table 3: The Effects on the Growth of Social Security

	All Sample Countries		Small Open Economies	
	(1)	(2)	(3)	(4)
Gini	-0.138^{*}	-0.139^{*}	-0.325^{**}	-0.313^{**}
	(0.098)	(0.101)	(0.130)	(0.138)
dependency ratio (1960)	-0.091	-	0.091	-
	(0.141)		(0.169)	
dependency ratio (1980)	-	-0.027	-	0.065
		(0.114)		(0.141)
adjusted R^2	0.118	0.10	0.412	0.396

Note: The dependent variable is the average growth rate of social security. Columns (1) and (2) report OLS regression for all 20 sample countries; columns (3) and (4) report OLS regression for 13 small OECD countries. ***, ** and * refer to significance at the 1%, 5% and 10%, respectively. Standard errors are in parenthesis. All specifications include a constant intercept.

Columns (1) and (2) in Table 3 give the results in the group of all sample countries, with the dependency ratio of 1960 and 1980 included separately. It can directly be seen from Figure 6 that income inequality and the average growth of social security are negatively correlated across all sample countries. The estimated coefficients on Gini are significantly different from zero at the 10 percent level, though not at the 5% level. Conditioning on the dependency ratio in different years only slightly changes the estimated coefficients on Gini. These findings support the prediction of our theory that social security growth should be negatively correlated with income inequality. However, when we proceed the dependency ratio, the coefficients are insignificant and the signs are opposite to the prediction. As mentioned above, this might be due to the rather unstable demographic structure.

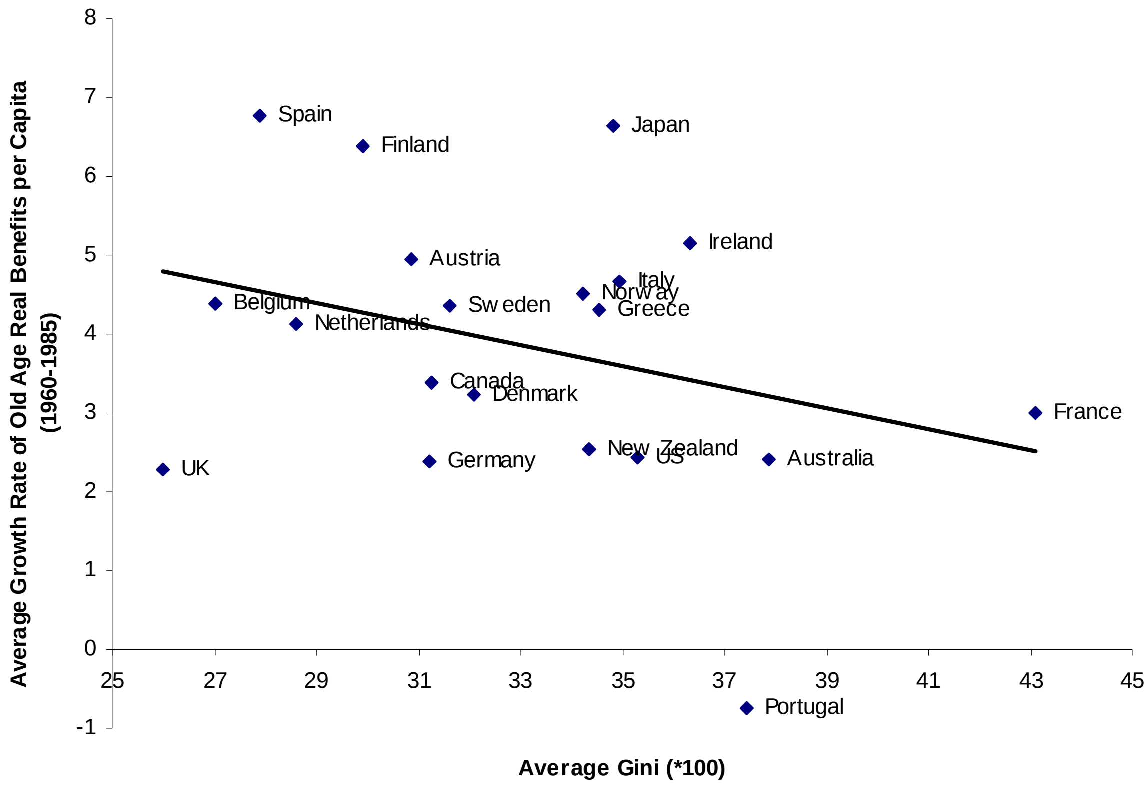

Figure 6: Cross-country Relationship between Income Inequality and the Average Growth Rate of Social Security (All Sample Countries).

Since our model is developed for a small open economy where factor prices are exogenous, we would like to run a regression for the subgroup of small OECD countries. Columns (3) and (4) report the results. Consistent with the prediction of the model, the estimated coefficients on Gini are significantly positive. Moreover, the significance level reaches 5% and R^2 jumps from 10% to 40%. Figure 7 reveals a rather close relationship between social security growth and income inequality in small OECD countries. This indicates that the social contract effect of income inequality can well explain the diversified growth of social security among small OECD countries. The much improved fitness also suggests that it might be important to include general equilibrium consideration for a better explanation of social security growth in large countries. We leave this work to future research. The social contract effect of the dependency ratio is insignificant, though the signs become positive and consistent with our prediction.

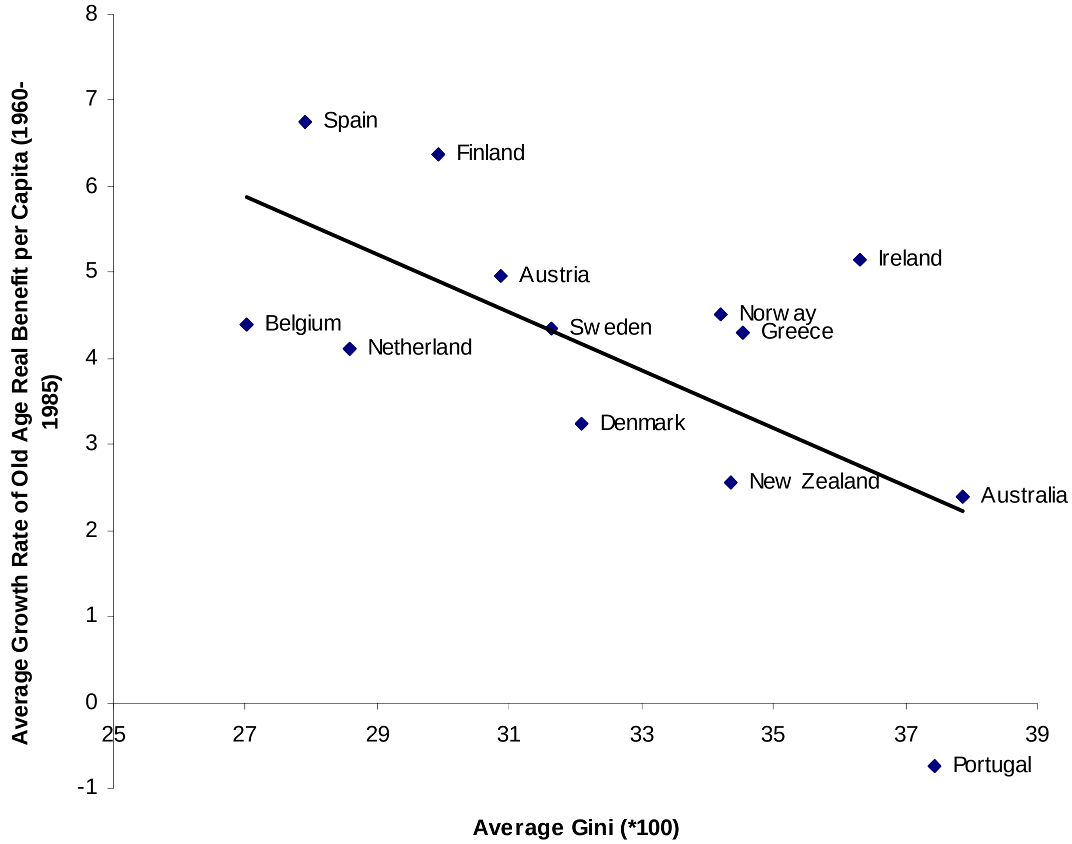

Figure 7: Cross-country Relationship between Income Inequality and the Average Growth Rate of Social Security (Small Open Economies).

The third prediction states that the long-run social security size is positively correlated with the dependency ratio, but non-monotonically correlated with income inequality. We regress the share of social security expenditure as a percentage of GDP in 1980 on the dependency ratio and Gini. The reason for choosing the year of 1980 is that the share of social security starts to level off in the 1980s after several decades of growth. To assess the sensitivity of demographic changes, we use the dependency ratio in 1960 and 1980 separately as an independent variable.

Columns (1) and (2) in Table 4 report the results for all sample countries. The estimated coefficients on the dependency ratio are significantly positive at 5%. The coefficients on Gini are significantly negative at 10% and 5%, respectively.

The significant correlations, which can directly be seen from Figures 8 and 9, are in accordance with the prediction. The results for the subgroup of small OECD countries are reported in Column (3) and (4). The negative correlation between social security and income inequality turns out to be robust and the significance increases to 1%. The correlation between the size of social security and the dependency ratio in 1980 is also significantly negative at 5%. But the coefficient on the dependency ratio in 1960 turns out to be insignificant.

Table 4: The Effects on the Size of Social Security

	All Sample Countries		Small Open Economies	
	(1)	(2)	(3)	(4)
Gini	-0.397^{*}	-0.462^{**}	-1.12^{***}	-1.19^{***}
	(0.258)	(0.250)	(0.330)	(0.288)
dependency ratio (1960)	-	0.743^{**}	-	0.785^{**}
		(0.366)		(0.384)
dependency ratio (1980)	0.525^{**}	-	0.347	-
	(0.298)		(0.351)	
adjusted R^2	0.263	0.296	0.567	0.657

Note: The dependent variable is the size of social security. Columns (1) and (2) report OLS regressions for all 20 sample countries; columns (3) and (4) report OLS regressions for 13 small OECD countries. ***, ** and * refer to significance at the 1%, 5% and 10%, respectively. Standard errors are in parenthesis. All specifications include constant intercept.

The robust negative correlation between inequality and social security transfers is closely related to the insignificant or even negative correlation between inequality and welfare spendings found in recent empirical studies (see, e.g., Lindert, 1996 and Rodriguez, 1998), since social security occupies the largest part of government transfers. The puzzling feature, however, can be reconciled with the non-monotonic relationship predicted by our model.[30] Specifically, it implies that the social contract effect of income inequality dominates the direct effect on the size of social security. This is consistent with the significant social contract

[30] The inverted-U relationship can be tested by adding a quadratic term of income inequality to the regression equation. The coefficients on Gini square turn out to be insignificant, however.

effect of income inequality on the growth of social security reported in Table 3. Thus, our model sheds light on the puzzling non-positive correlation between inequality and government transfers that are hard to explain with the existing theories.

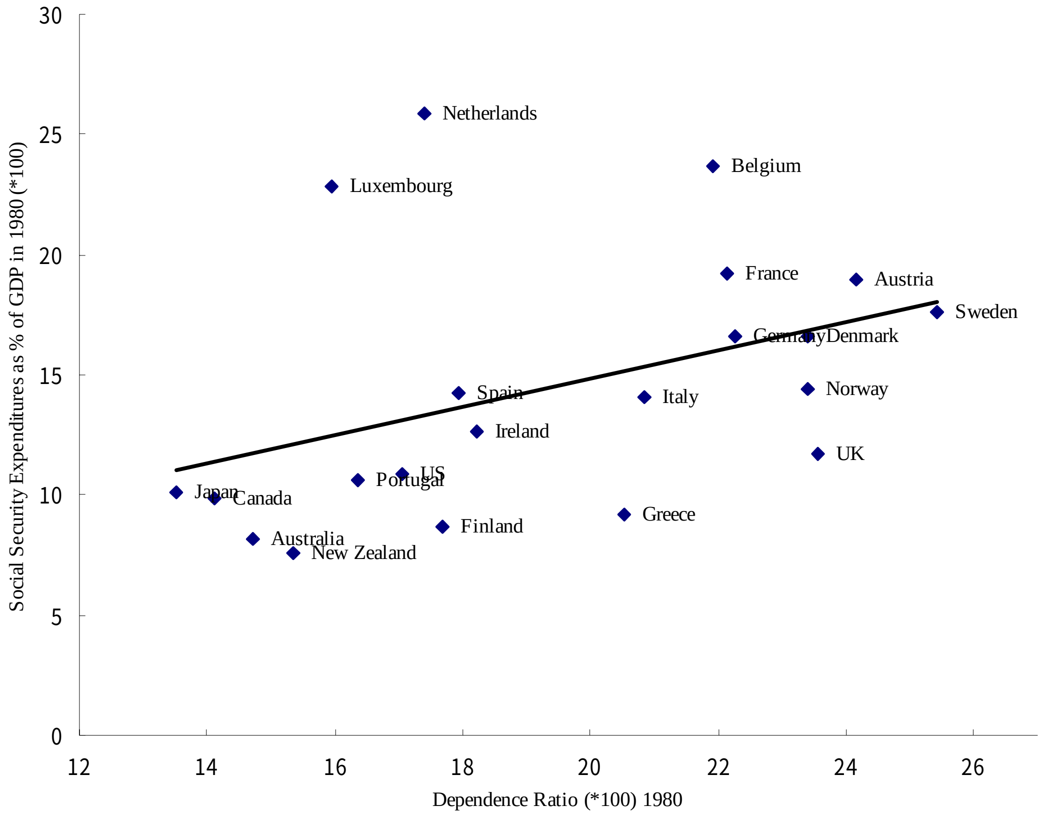

Figure 8: Cross-country Relationship between the Dependency Ratio and the Size of Social Security (All Sample Countries).

Finally, it is worth mentioning that Razin *et al.* (2002) find a negative effect of the dependency ratio on the size of social security from a panel of OECD data. At a first sight, this result seems inconsistent with our model. Nevertheless, our model implies that the growing sizes of social security may be an intrinsic feature of the dynamic politico-economic equilibrium. To give an unbiased estimation of demographic effects in the panel data, the potential growth trend of social security must be controlled for. To see this more clearly, we look at the demographic change in the data used by Razin *et al.* (2002). During the data period (from

1970 to 1991), the average dependency ratio fell from 58% to 54% in the OECD countries. Together with the growing sizes of social security in the same period, it would not be surprising to obtain a negative relationship between social security and the dependency ratio, even if the dependency ratio had no effects on the size of social security.

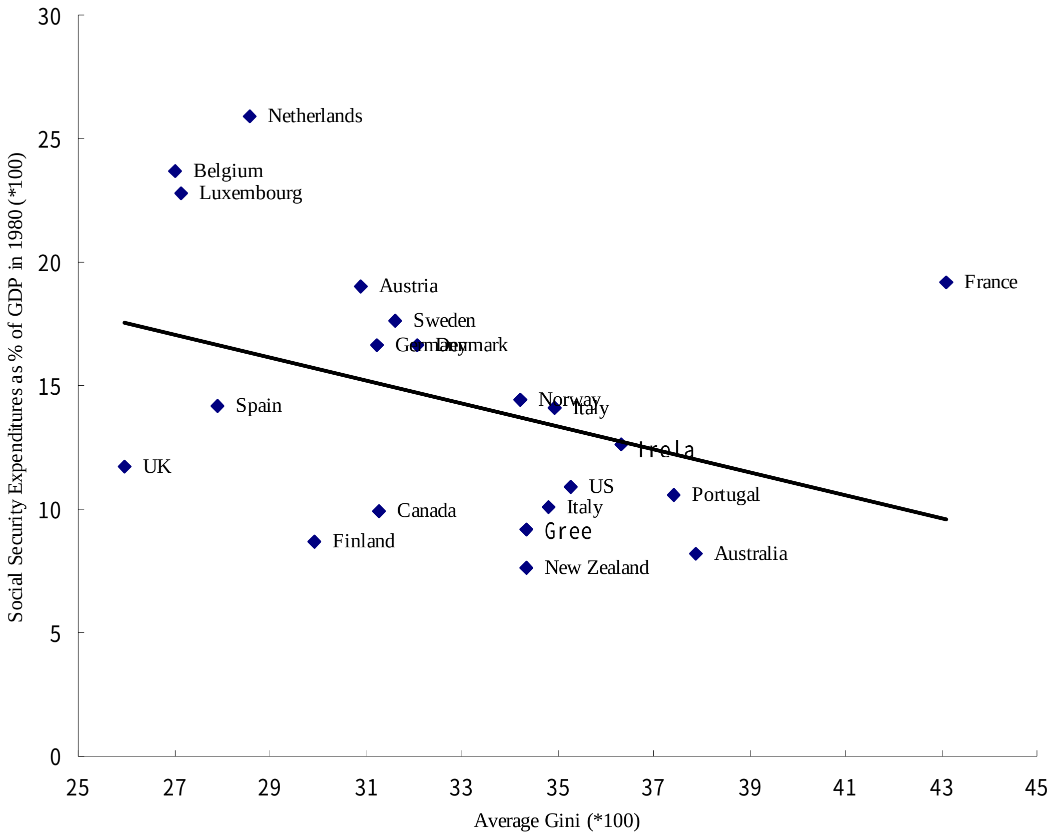

Figure 9: Cross-country Relationship between Income Inequality and the Size of Social Security (All Sample Countries).

6 Conclusion

This paper develops a positive theory of social security in a majority voting framework with a middle aged median voter. We find that there exists a Markovian social contract that sustains the system. The social contract positively relates future social security benefits to current contributions and thus, solves the temporal separation problem of benefits and contributions. This relationship is achieved through two fundamental linkages. One is a negative reaction of the next period median voter's human capital stock to the current payroll tax rate due to its distortionary effects on human capital investment. The other is a negative reaction of the payroll tax rate chosen by the next period's median voter to her human capital stock, due to the proportional income taxation. Given this positive relationship, the median voter with less income is more willing to support the social security system for intra-generational redistributive reasons. Therefore, the system can be sustained in a dynamic efficient economy with sufficiently high income inequality. The linear-quadratic preference allows us to obtain the close-form solution of the Markov perfect equilibrium.

The major theoretical contribution of this paper is twofold. First, we show that even under the temporal separation of contributions and benefits, the self-interested median voter may choose to sustain social security in Markov perfect equilibrium. Therefore, our model contrasts with the previous studies on political sustainability of social security which either resort to the imperfect temporal separation of contributions and benefits or non-Markovian expectation on future policy outcomes. Second, we incorporate within-cohort heterogeneity to capture the intra-generational redistributive effects of social security transfers. It turns out that income inequality plays an essential role in the sustainability of the social security system.

Our model provides a variety of empirical predictions that can be confronted with the data. First, the size of social security converges monotonically to the steady state. This implies increasing sizes of social security over time, after the introduction of the system. Second, the growth of social security is positively correlated with the dependency ratio but negatively correlated with income inequality. Third, the size of social security in the long run is increasing in the dependency ratio but non-monotonically related to income inequality. These predictions are based on the Markovian social contract and differ fundamentally from the Ramsey allocation. In the appendix, we show that the Ramsey allocation implies an oscillatory convergence of social security tax rates. And income inequality does not affect the Ramsey allocation due to the linear preferences over lifetime wealth. Our empirical evidence shows that the predictions of the Markovian social contract are broadly consistent with the data in the OECD countries. Two novel results are worth reemphasizing. First, the growth of social security is positively related to income inequality and second, the puzzling negative correlation between inequality and government transfers in the literature can be reconciled with our theory.

There are several interesting extensions that can be made in future research. First, we would like to study the effect of continuously changing population growth rates. This can help us analyse the dynamics of social security with anticipated demographic changes. Second, the general equilibrium effect is ignored in the model. This leads to the difficulty in accounting for the evolution of the social security system in a large economy such as the United States. For analytical convenience, we impose a balanced budget on social security transfers. A natural extension of the model would be to relax this assumption and allow for government debt.

Reference

Aaron, H. J., 1966, The Social Insurance Paradox, Canadian Journal of Economics and Political Science, 32, 371-376.

Alesina, A. and D. Rodrik, 1994, Distributive Politics and Economic Growth, Quarterly Journal of Economics, 109, 465-490.

Azariadis, Costas and Vincenzo Galasso, 2002, Fiscal Constitutions, Journal of Economic Theory, 103, 255-281.

Benabou, Roland, 1996, Inequality and Growth, in Ben S. Bernanke and Julio J. Rotemberg, eds., NBER Macro Annual, Vol. 11, 11-74, MIT Press, Cambridge.

Benabou, Roland, 2001, Unequal Societies: Income Distribution and the Social Contract, The American Economic Review, Vol. 90, 96-129.

Boldrin, Michele and Aldo Rustichini, 2000, Political Equilibria with Social Security, Review of Economic Dynamics, 3, 41-78.

Breyer, Friedrich and Ben Craig, 1997, Voting on Social Security: Evidence from

OECD countries, European Journal of Political Economy, 13, 705-724.

Conesa, Juan C. and Dirk Krueger, 1999, Social Security Reform with Heterogeneous Agents, Review of Economic Dynamics, 2, 757-795.

Cooley, Thomas F. and Jorge Soares, 1999, A Positive Theory of Social Security Based on Reputation, Journal of Political Economy, 107, 135-160.

Cutler, David M. and Richard Johnson, 2002, The Birth and Growth of the Social Insurance States: Explaining Old Age and Medical Insurance Across Countries, mimeo.

Deininger, Klaus and Lyn Squire, 1996, A New Data Set Measuring Income Inequality, The World Bank Economic Review, Vol. 10, 565-91. Updated dataset available on http://www.worldbank.org.

Diamond, Peter A., 1965, National Debt in a Neoclassical Growth Model, American Economic Review, 55, 1126-50.

Disney, Richard and Paul Johnson (Editors), 2001, social security systems and Retirement Incomes across OECD Countries, Edward Elgar, Cheltenham.

Evans, George W. and Seppo Honkapohja, 2001, Learning and Expectations in Macroeconomics, Princeton University Press.

Feldstein, Martin and Liebman, Jeffrey B., 2001, Social Security, NBER Working Paper, No. 8451.

Forni, Lorenzo, 2005, Social Security as Markov Equilibrium in OLG Models, Review of Economic Dynamics, 8, 178-194.

Gonzalez-Eiras, Martin and Dirk Nieplet, 2004, Sustaining Social Society, mimeo, IIES.

Gordon, Margaret S., 1988, Social Security Policies in Industrial Countries: A Comparative Analysis, Cambridge University Press, New York.

Hassler, John, Jose V. Rodriguez Mora, Kjetil Storesletten and Fabrizio Zilibotti, 2003, The Survival of the Welfare State, American Economic Review, 93, 87-112.

Hassler, John, Kjetil Storesletten and Fabrizio Zilibotti, 2004, Democratic Public Good Provision, Journal of Economic Theory, forthcoming.

Judd, Kenneth L., 1992, Projection Method for Solving Aggregate Growth Models, Journal of Economic Theory, 58, 410-452.

Katuscak, Peter, 2002, A Positive Theory of Social Security in an Open Economy, mimeo.

Klein, Paul, Per Krusell and Jose-Victor Rios-Rull, 2003, Time-Consistent Public Expenditure, mimeo.

Kohlberg, Elon, 1976, A Model of Economic Growth with Altruism between Generation, Journal of Economic Theory, 13, 1-13.

Krusell, Per and Jose-Victor Rios-Rull, 1999, On the Size of US Government: Political Economy in the Neoclassical Growth Model, American Economic Review.

Krusell, Per and Anthony A. Smith, Jr, 2003, Consumption-Savings Decisions with Quasi-Geometric Discounting, Econometrica, 71,

Lindert, Peter H, What Limits social Spending, Explorations in Economic History, 33, 1-34.

Lockwood, Ben, 1996, Uniqueness of Markov-Perfect Equilibrium in Infinite-Time Affine-Quadratic Differential Games, Journal of Economic Dynamics and Control, 20, 751-765.

Maskin, Eric and Jean Tirole, 2003, Markov Perfect Equilibrium, I. Observable Actions, Journal of Economic Theory, 100, 191-219.

Meltzer, Allan and Scott Richard, 1983, A Rational Theory of the Size of Government, Journal of Political Economy, 89, 914-927.

McCallum, Bennet, 1983, On Non-uniqueness in Rational Expectional Models, Journal of Monetary Economy, 11, 139-168.

Mulligan, Casey B. and Xavier Sala-i-Martin, 1999a, Gerontocracy, Retirement, and Social Security, NBER Working Paper, No. 7117.

Mulligan, Casey B. and Xavier Sala-i-Martin, 1999b, Social Security in Theory and practice(I): Facts and Political Theories, NBER Working Paper, No. 7118.

OECD, 1988, Reforming Public Pensions, reported prepared by Robert Holzmann, OECD Social Policy Studies No.5.

Perotti, Roberto, 1996, Growth, Income Distribution, and Democracy: What the Data Say, Journal of Economic Growth, 1, 149-87.

Persson, Torsten and Guido Tabellini, 1994, Is Inequality Harmful for Growth, American Economic Review, 84, 600-621.

Persson, Torsten and Guido Tabellini, 2000, Political Economics: Explaining Economic Policy, MIT Press, Cambridge, MA.

Phelps, E. S. and R. A. Pollak, 1968, On Second Best National Savings and Game Equilibrium Growth, Review of Economic Studies, 35, 185-199.

Razin, Assaf, Efraim Sadka and Phillip Swagel, 2002, The Aging Population and the Size of the Welfare State, Journal of Political Economy, 110, 900-918.

Rodriguez, Francisco, 1998, Essays on Redistribution, Development and the State, Ph.D. Dissertation, Harvard University.

Sjoblom, Kriss, 1985, Voting for Social Security, Public Choice, 45, 225-40.

Song, Zheng, 2005a, Dynamic Inequality and Social Security, mimeo, IIES.

Song, Zheng, 2005b, Ideology and the Determination of Public Policy Over Time, mimeo, IIES.

Storesletten, Kjetil, Christopher I. Telmer and Amir Yaron, 2004, Consumption and Risk Sharing over the Life Cycle, Journal of Monetary Economics, 51, 609-633.

Tabellini, Guido, 2000, A Positive Theory of Social Security, Scandinavian Journal of Economics, 102, 523-545.

United Nations, 2000, Demographic Year Book, Historical Supplement, 1948-1997.

Wolfinger, Raymond E. and Steven J. Rosenstone, 1980, Who votes?, Yale University Press, 1980.

7 Appendix

7.1 The Definition of Parameters in Proposition 1

Let $q \equiv \left(2 - 27n^2 R \left(w^u\right)^2 / 2\Phi\right) / 27n^3$. b_0, b_1, π_0, π_1 are defined as follows

$$b_1 \equiv \sqrt[3]{-q + \sqrt{q^2 - \frac{1}{729n^6}}} + \sqrt[3]{-q - \sqrt{q^2 - \frac{1}{729n^6}}} - \frac{1}{3n}, \tag{4.37}$$

$$\pi_1 \equiv -\Phi \left(1 + \left(n + \frac{1}{R}\right) b_1 + \frac{n}{R} b_1^2\right), \tag{4.38}$$

$$b_0 \equiv \frac{\Pi b_1 - \left(w^u\right)^2 (1 + R)}{\Phi b_1 \left(n + 1/R + n/R \left(1 + b_1\right)\right) - \pi_1 \left(1 + 2b_1\right) - \left(w^u\right)^2}, \tag{4.39}$$

$$\pi_0 \equiv \Pi - \Phi \left(\left(n + \frac{1}{R}\right) b_0 + \frac{n}{R} b_0 \left(1 + b_1\right)\right), \tag{4.40}$$

7.2 Proof of Proposition 1

Due to the linear-quadratic preference, it would be natural to guess that the policy rule T is linear

$$T\left(h_{t-1}^u\right) = \phi_0 + \phi_1 h_{t-1}^u, \tag{4.41}$$

where ϕ_0 and ϕ_1 are two undetermined coefficients. Substitute into (4.11)

$$H\left(\tau_t\right) = \left(1 + \frac{1 - \phi_0}{R} - \tau_t\right) \frac{w^u}{1 + w^u \phi_1 / R}. \tag{4.42}$$

Combining (4.41) and (4.42), we obtain a linear social contract B

$$B\left(\tau_t\right) = b_0 + b_1 \tau_t, \tag{4.43}$$

where $b_0 \equiv \phi_0 + w^u \phi_1 \left(1 + R - \phi_0\right) / \left(R + w^u \phi_1\right)$ and $b_1 \equiv -R w^u \phi_1 / \left(R + w^u \phi_1\right)$. Plugging (4.43) into (4.15), the function tax base is

$$Y\left(\tau_t\right) = \pi_0 + \pi_1 \tau_t, \tag{4.44}$$

where π_0 and π_1 are defined by (4.40) and (4.38), respectively.

The first-order condition (4.16) yields

$$\tau_t = -\frac{b_1\pi_0 + b_0\pi_1}{2b_1\pi_1} + \frac{Rw^u}{2b_1\pi_1} h^u_{t-1}. \tag{4.45}$$

(4.45) pins down ϕ_0 and ϕ_1 in (4.41)

$$\phi_0 = -\frac{b_1\pi_0 + b_0\pi_1}{2b_1\pi_1}, \tag{4.46}$$

$$\phi_1 = \frac{Rw^u}{2b_1\pi_1}. \tag{4.47}$$

(4.47) implies a four-order polynomial of b_1:

$$2\Phi\frac{n}{R}b_1^4 + 2\Phi\left(n + \frac{1}{R}\right)b_1^3 + 2\Phi b_1^2 - (w^u)^2 b_1 - R(w^u)^2 = 0. \tag{4.48}$$

Factorizing (4.48), one root of b_1 equals $-R$, which should be omitted by Lemma 1. The other three roots solve

$$nb_1^3 + b_1^2 - \Psi = 0, \tag{4.49}$$

where $\Psi \equiv R(w^u)^2/2\Phi$. Rearrange (4.49):

$$b_1^2 = \frac{\Psi}{1+nb_1}. \tag{4.50}$$

It is straightforward to see that the LHS and the RHS of (4.50) have a unique cross with $b_1 > 0$, which gives the only real root of b_1, i.e., (4.37). The other undetermined coefficient b_0 can easily be solved.

Then we need to check whether $T\left(h^u_{t-1}\right) \in [0, \overline{\tau}]$ for all $h^u_{t-1} \in \left[\underline{h}, \overline{h}\right]$. This gives the existence conditions (4.17) and (4.18).

Finally, $T\left(\overline{h}\right) \geq 0$ requires $\phi_0 + \phi_1\overline{h} \geq 0$, which implies $\phi_0 \leq -\phi_1 w^u(1+R)/R$. By the definition of b_0, it is easily shown that $b_0 \geq 0$. On the other hand, $T(0) \leq \overline{\tau}$ implies that $b_0 + b_1 \leq \overline{\tau}$. Together with $b_1 > 0$ implied by Lemma 1, we have $b_1 \in (0, \overline{\tau}]$ and $b_0/(1-b_1) \in [0, \overline{\tau}]$. □

7.3 Proof of Lemma 2

Without loss of generality, let us assume that there are three sets, $\Theta^{m,s}$, $\Theta^{m,u}$ and Θ^o, with $\Theta^{m,s} \cap \Theta^o = \oslash$, $\Theta^{m,u} = \left[\underline{h}, \overline{h}\right] \setminus (\Theta^{m,s} \cup \Theta^o)$. The median voter at time t is the middle rich, the middle poor and the old, for $h^u_{t-1} \in \Theta^{m,s}$, $\Theta^{m,u}$ and Θ^o, respectively. Denote $T^{m,s}\left(h_{t-1}^{u}\right)$, $T^{m,u}\left(h_{t-1}^{u}\right)$ and $T^{o}\left(h_{t-1}^{u}\right)$ the preferred tax rate of the middle rich, middle poor and the old, respectively. Then, the equilibrium policy rule can be written as

$$T\left(h_{t-1}^{u}\right)=\begin{cases} T^{m,s}\left(h_{t-1}^{u}\right) & \text{if } h_{t-1}^{u}\in\Theta^{m,s} \\ T^{m,u}\left(h_{t-1}^{u}\right) & \text{if } h_{t-1}^{u}\in\Theta^{m,u} \\ T^{o}\left(h_{t-1}^{u}\right) & \text{if } h_{t-1}^{u}\in\Theta^{o} \end{cases}$$

It is straightforward from (4.16) that $T^{m,s}\left(h_{t-1}^{u}\right)<T^{m,u}\left(h_{t-1}^{u}\right)\ \forall h_{t-1}^{u}$. So $T\left(h_{t-1}^{u}\right)$ would be discontinuous if the median voter is the middle rich, except for $\Theta^{m,u}\cup\Theta^{o}=\oslash$. The contradiction establishes $\Theta^{m,s}=\oslash$ or $\Theta^{m,u}\cup\Theta^{o}=\oslash$. That is to say, if the middle rich becomes the median voter, she must be the only median voter in the dynamic political equilibrium.

For the old to be the median voter, we need $T^{m,s}\left(h_{t-1}^{u}\right)<T^{o}\left(h_{t-1}^{u}\right)<T^{m,u}\left(h_{t-1}^{u}\right)$. Otherwise the the middle-aged would be the median voter. Suppose $\Theta^{o}\neq\oslash$ and $\Theta^{m,u}\neq\oslash$. By the above analysis, this implies $\Theta^{m,s}=\oslash$. Then, there exists some $\hat{h}$ such that $T^{o}\left(\hat{h}+\varepsilon\right)\lesseqgtr T^{m,u}\left(\hat{h}+\varepsilon\right)$ and $T^{o}\left(\hat{h}-\varepsilon\right)\gtreqless T^{m,u}\left(\hat{h}-\varepsilon\right)$. This implies that $\mathrm{d}T^{o\prime}\left(\hat{h}\right)/\mathrm{d}\hat{h}\neq\mathrm{d}T^{m,u\prime}\left(\hat{h}\right)/\mathrm{d}\hat{h}$, while the differentiability requires $\mathrm{d}T^{o\prime}\left(\hat{h}\right)/\mathrm{d}\hat{h}=\mathrm{d}T^{m,u\prime}\left(\hat{h}\right)/\mathrm{d}\hat{h}$. The contradiction establishes $\Theta^{o}=\oslash$ or $\Theta^{m,u}=\oslash$. $\square$

7.4 Proof of Proposition 2

The proof consists of two steps. First, we check if the middle poor can be the median voter. Then we rule out the possibility for the middle rich or the old of being decisive in the majority voting.

Given the equilibrium social contract B, the current tax base y_t can be written as

$$y_t=Y_c\left(h_{t-1}^{u},\tau_t\right)=\Phi\left(\frac{h_{t-1}^{u}}{w^{u}}+n\left(1-\tau_t+\frac{1-B\left(\tau_t\right)}{R}\right)\right). \tag{4.51}$$

Maximizing the indirect utility of the old, (4.14) solves

$$\tau_t^{o}=\min\left\{1,\frac{-Y_c\left(h_{t-1}^{u},\tau_t\right)}{\partial Y_c\left(h_{t-1}^{u},\tau_t\right)/\partial\tau_t}\right\}$$

. In words, the old choose τ_t^{o} to attain the top of the Laffer curve. Substituting the social contract B (4.21) into (4.51), we have

$$\tau_t^o = \min\left\{1, \frac{1 + 1/R - b_0/R + h_{t-1}^u/nw^u}{2\left(1 + b_1/R\right)}\right\}, \tag{4.52}$$

where b_0 and b_1 are defined in (4.39) and (4.37), respectively. Since τ_t^o increases in h_{t-1}^u, the minimum τ_t^o locates at $h_{t-1}^u = \underline{h}$, which gives the RHS of (4.27).

Now we move to the second step. First, consider the gerontocracy where the old are always decisive. Guess that the policy rule under the gerontocracy, denoted by T^o, follows $\tau_t = \phi_0^o + \phi_1^o h_{t-1}^u$. The differentiability rules out the corner solution and hence, (4.52) establishes:

$$\tau_t = \frac{1 + R - b_0^o}{2\left(R + b_1^o\right)} + \frac{R/nw^u}{2\left(R + b_1^o\right)} h_{t-1}^u, \tag{4.53}$$

where $b_0^o \equiv \phi_0^o + w^u \phi_1^o \left(1 + R - \phi_0^o\right) / \left(R + w^u \phi_1^o\right)$ and $b_1^o \equiv -Rw^u \phi_1^o / \left(R + w^u \phi_1^o\right)$ are the coefficients in the social contract B^o under gerontocracy. (4.53) pins down ϕ_0^o and ϕ_1^o:

$$\phi_0^o = \frac{1 + R - b_0^o}{2\left(R + b_1^o\right)}, \tag{4.54}$$

$$\phi_1^o = \frac{R/nw^u}{2\left(R + b_1^o\right)}. \tag{4.55}$$

(4.55) implies a linear equation of ϕ_1^o, which solves

$$\begin{aligned} \phi_1^o &= \frac{R}{\left(2nR - 1\right) w^u} > 0, \\ b_1^o &= -\frac{1}{2n} < 0. \end{aligned}$$

It can further be obtained that $\phi_0^o = \left(1 + R\right) / \left(2 + R\right)$ and $b_0^o = \left(n + 1\right) \phi_0^o / n$. Given $b_1^o < 0$, the middle-aged would vote for zero tax rate and thus, T^o or B^o cannot be sustained under majority voting.

Next consider the dictatorship of the middle rich (DMR). It immediately follows that the policy rule and the social contract under DMR, denoted by $T^{m,s}$ and $B^{m,s}$, follow $\tau_t = \phi_0^{m,s} + \phi_1^{m,s} h_{t-1}^u$ and $\tau_t = b_0^{m,s} + b_1^{m,s} \tau_{t-1}$, respectively. The definitions of $\phi_0^{m,s}$, $\phi_1^{m,s}$, $b_0^{m,s}$ and $b_1^{m,s}$ follow the definitions of ϕ_0, ϕ_1, b_0 and b_1, where w^u is replaced with w^s. Then we need to show that DMR cannot survive the the majority voting. Since the first inequality of (4.26) must hold, the middle rich would be the median voter if and only if

$$\tau_t^o \leq \tau_t^{m,s} \tag{4.56}$$

for any $h^u_{t-1} \in \left[0, \overline{h}\right]$. Paralleled with (4.52), we have

$$\tau^o_t = \min\left\{1, \frac{1 + 1/R - b^{m,s}_0/R + h^u_{t-1}/nw^u}{2\left(1 + b^{m,s}_1/R\right)}\right\}, \tag{4.57}$$

given that $b^{m,s}_0$ and $b^{m,s}_1$. (4.56) will not be satisfied if (4.28) holds.

7.5 Proof of Proposition 4

Differentiating (4.49) with respect to x, we have

$$\frac{\partial b_1}{\partial x} = \frac{\partial \Psi/\partial x}{3nb_1^2 + 2b_1}$$

where x refers to w^u and n, respectively. Since $b_1 > 0$, sgn$(\partial b_1/\partial x)$ =sgn$(\partial \Psi/\partial x)$. It immediate follows that $\partial \Psi/\partial w^u > 0$ and $\partial \Psi/\partial n < 0$. □

7.6 Ramsey Allocation

In this section, we character the Ramsey allocation where a benevolent planner with a commitment technology sets the sequence of tax rates $\{\tau_t\}^\infty_{t=0}$ so as to maximize the sum of the discounted utilities of all generations. The constraint of the planner is that the chosen policy should be implementable as a competitive equilibrium. The evolution of social security in the Ramsey allocation will be compared with the political equilibrium outcome. Substituting (4.3) into (4.2), the Ramsey problem can be written as

$$\begin{aligned}
&\max_{\{\tau_t\}^\infty_{t=0}} \tau_0 \left[n\left(\lambda w^u h^u_{-1} + (1-\lambda) w^s h^s_{-1}\right) + n^2\left(\lambda w^u h^u_0(\tau_0, \tau_1) + (1-\lambda) w^s h^s_0 h^u_0(\tau_0, \tau_1)\right)\right] \\
&+\lambda\left((1-\tau_0) w^u h^u_{-1} + \frac{y_1(\tau_0, \tau_1, \tau_2)\tau_1}{R}\right) + (1-\lambda)\left((1-\tau_0) w^s h^s_{-1} + \frac{y_1(\tau_0, \tau_1, \tau_2)\tau_1}{R}\right) \\
&+\sum_{t=0}^{\infty} \beta^t \left(\lambda V^u_t(\tau_t, \tau_{t+1}, \tau_{t+2}, \tau_{t+3}) + (1-\lambda) V^s_t(\tau_t, \tau_{t+1}, \tau_{t+2}, \tau_{t+3})\right)
\end{aligned} \tag{4.58}$$

where $\beta \in (0, 1)$ is the discount factor of the planner on the welfare of all generations born after time 0. h^j_{-1} is the predetermined human capital before the initial period. V^j_t follows

$$\begin{aligned}
V^j_t(\tau_t, \tau_{t+1}, \tau_{t+2}, \tau_{t+3}) &= (1-\tau_t) w^j h^j_t(\tau_t, \tau_{t+1}) + \frac{(1-\tau_{t+1}) w^j h^j_t(\tau_t, \tau_{t+1})}{R} \\
&+\frac{y_{t+2}(\tau_{t+1}, \tau_{t+2}, \tau_{t+3})\tau_1}{R^2} - \frac{\left(h^j_t(\tau_t, \tau_{t+1})\right)^2}{2}
\end{aligned} \tag{4.59}$$

We assume that τ_t cannot exceed one, but can be negative. The negative τ_t refers to the intergenerational redistribution from the old to working generations. For analytical convenience, we focus on interior solutions. (4.58) is a standard sequential problem. Note that the derivatives $\partial h_t^j(\tau_t,\tau_{t+1})/\partial\tau_t = -w^j$, $\partial h_t^j(\tau_t,\tau_{t+1})/\partial\tau_{t+1} = -w^j/R$, $\partial y_{t+1}(\tau_t,\tau_{t+1},\tau_{t+2})/\partial\tau_t = -\Phi$, $\partial y_{t+1}(\tau_t,\tau_{t+1},\tau_{t+2})/\partial\tau_{t+1} = -\Phi(n+1/R)$ and $\partial y_{t+1}(\tau_t,\tau_{t+1},\tau_{t+2})/\partial\tau_{t+2} = -\Phi n/R$ are time-invariant. The first-order condition of (4.58) with respect to τ_t for any $t \geq 2$ is

$$\begin{aligned}&\left(\frac{-\Phi\frac{n}{R}\tau_{t-1}}{R^2}\right) + \beta\left(\frac{-\Phi\left(n+\frac{1}{R}\right)\tau_t + y_t(\tau_{t-1},\tau_t,\tau_{t+1})}{R^2}\right)\\
&+\beta^2\sum_{j=u,s}\mu^j\left(-(1-\tau_{t-1})\left(w^j\right)^2/R - \frac{(1-\tau_t)\left(w^j\right)^2/R}{R} + \frac{-\Phi\tau_{t+1}}{R^2}\right)\\
&+\beta^3\sum_{j=u,s}\mu^j\left(-(1-\tau_t)\left(w^j\right)^2 - \frac{(1-\tau_{t+1})\left(w^j\right)^2}{R}\right) = 0 \qquad (4.60)\end{aligned}$$

(4.60) yields a second-order linear difference equation after some algebra

$$\gamma_1\tau_{t+1} + \gamma_2\tau_t + \gamma_3\tau_{t-1} = \Delta \qquad (4.61)$$

where $\gamma_1 \equiv \beta\left(\beta^2 - \beta n/R - n^2/R^2\right)/R$, $\gamma_2 \equiv \beta\left(\beta^2 + \beta/R^2 - 2n(n+1/R)/R^2\right)$, $\gamma_3 \equiv \gamma_1/\beta$ and $\Delta \equiv \beta(\beta(1+1/R)(\beta+1/R) - n(1+n)(1+R)/R)$. The first-order condition of (4.58) with respect to τ_0 is

$$\begin{aligned}&\sum_{j=u,s}\mu^j\left(\tau_0 n^2\left(-\lambda\left(w^j\right)^2\right) + y_0\right) + \sum_{j=u,s}\mu^j\left(-w^j h_{-1}^j + \frac{-\Phi\tau_1}{R}\right)\\
&+\sum_{j=u,s}\mu^j\left(-(1-\tau_0)\left(w^j\right)^2 - \frac{(1-\tau_1)\left(w^j\right)^2}{R}\right) = 0 \qquad (4.62)\end{aligned}$$

where $y_0 = n\left(\lambda w^u h_{-1}^u + (1-\lambda)w^s h_{-1}^s\right) + n^2\left(\lambda w^u h_0^u(\tau_0,\tau_1) + (1-\lambda)w^s h_0^s h_0^u(\tau_0,\tau_1)\right)$. The first-order condition of (4.58) with respect to τ_1 is

$$\begin{aligned}&\frac{\tau_0 n^2}{R}\left(-\lambda\left(w^u\right)^2 - (1-\lambda)\left(w^s\right)^2\right) + \frac{-\Phi(n+1/R)\tau_1 + y_1(\tau_0,\tau_1,\tau_2)}{R}\\
&+\sum_{j=u,s}\mu^j\left(-(1-\tau_0)\left(w^j\right)^2/R - \frac{(1-\tau_1)\left(w^j\right)^2/R}{R} + \frac{-\Phi\tau_2}{R^2}\right)\\
&+\beta\sum_{j=u,s}\mu^j\left(-(1-\tau_1)\left(w^j\right)^2 - \frac{(1-\tau_2)\left(w^j\right)^2}{R}\right) = 0 \qquad (4.63)\end{aligned}$$

We consider two cases of the discount factor: $\beta = 1/R$ and $\beta = n/R$. The second case implies that the planner weighs generations by their relative sizes and discounts their welfare by the interest rate. Assume the economy to be dynamic efficient so that $\beta < 1$ holds.

Proposition 5 *In the Ramsey allocation, we have*

(1) If $\beta = 1/R$, the social security tax rate converges to a positive $\tau^R \in (0,1)$ in an oscillatory way, where

$$\tau^R = \frac{\Delta}{\gamma_1 + \gamma_2 + \gamma_3} \tag{4.64}$$

(2) If $\beta = n/R$, the social security tax rate converges to zero in an oscillatory way.

If $\beta = (1 + z(n-1))/R$, then the above two cases refer to $z = 0$ and $z = 1$. For $z \in (0,1)$, numerical simulations show that Proposition 2 is robust. Two remarks are in order. First, the oscillatory convergence of social security taxes in the Ramsey allocation is fundamentally different from the monotonic convergence in the Markovian political equilibrium. Second, due to the linear preferences over lifetime wealth, income inequality does not affect the Ramsey allocation.

7.6.1 Proof of Proposition 5

(1) The steady state payroll tax rate in the Ramsey allocation τ^R is solved by substituting τ^R for τ_t in (4.60). Substituting $\beta = 1/R$ into Δ, it is immediate that $\Delta < 0$, $\gamma_1 < 0$, $\gamma_2 < 0$ and $\gamma_3 < 0$ by $n > 1$. Moreover, $|\Delta| < |\gamma_1 + \gamma_2 + \gamma_3|$. So $\tau^R \in (0,1)$. The eigenvalues of (4.61) are $\eta_1 = \left(-\gamma_2 + \sqrt{\gamma_2^2 - 4\gamma_1\gamma_3}\right) / (2\gamma_1)$ and $\eta_2 = \left(-\gamma_2 - \sqrt{\gamma_2^2 - 4\gamma_1\gamma_3}\right) / (2\gamma_1)$. Both of these have a negative real part. $n > 1$ implies that $\gamma_2 < \gamma_1 + \gamma_3$. So η_1 and η_2 are real numbers, $\eta_1 \in (0,1)$ and $\eta_2 < -1$.

(2) Substituting $\beta = n/R$ into Δ, it immediately follows that $\Delta = 0$. The other proofs are the same as above. $\square$

7.7 Data Source

The cross-country data on the average growth rate of real social security benefits per beneficiary are taken from OECD (1988). The data on the social security benefit as a percentage of GDP are taken from Economic Outlook of OECD. As we lack the data for the average growth rate of social security benefit per beneficiary for France and Greece, we approximate the average growth rate of social security tax rates by the average growth rate of the share of social security expenditure in GDP for these two countries. For most countries, the data cover the period 1960-1985.[31]

The data for the dependency ratio in 1960 and 1980 are from authors' calculations based on the demographic data in United Nations (2000). For those countries missing data for the above specific years, we use the data for the following year.

The data of the average Gini coefficients are from the updated data set of Deininger and Squire (1996, Table 1).[32] The data show that inequality does not vary to any considerable extent during the sample period, though there is a trend of increasing inequality. In addition, the coverage period for the inequality statistics is broadly consistent with the coverage period for the average growth rate of real social security benefits per beneficiary.

[31] The starting year for Australia is 1961 and the ending year for Portugal and Sweden is 1984. The average growth rates for Belgium and Spain are for periods 1971-1984 and 1974-1985, respectively.

[32] For Austria, the calculation of the statistics excludes the fraction of population who are self-employed.

MONOGRAPH SERIES

1. **Michaely, Michael *The Theory of Commercial Policy: Trade and Protection,* 1973**
2. **Söderström, Hans Tson *Studies in the Microdynamics of Production and Productivity Change*, 1974**
3. **Hamilton, Carl B. *Project Analysis in the Rural Sector with Special Reference to the Evaluation of Labour Cost,* 1974**
4. **Nyberg, Lars and Staffan Viotti *A Control Systems Approach to Macroeconomic Theory and Policy in the Open Economy*, 1975**
5. **Myhrman, Johan *Monetary Policy in Open Economies*, 1975**
6. **Krauss, Melvyn *International Trade and Economic Welfare*, 1975**
7. **Wihlborg, Clas *Capital Market Integration and Monetary Policy under Different Exchange Rate Regimes*, 1976**
8. **Svensson, Lars E.O. *On Competitive Markets and Intertemporal Resources Allocation*, 1976**
9. **Yeats, Alexander J. *Trade Barriers Facing Developing Countries*, 1978**
10. **Calmfors, Lars *Prices, Wages and Employment in the Open Economy*, 1978**
11. **Kornai, János *Economics of Shortage*, Vols I and II, 1979**
12. **Flam, Harry *Growth, Allocation and Trade in Sweden. An Empirical Application of the Heckscher-Ohlin Theory*, 1981**
13. **Persson, Torsten *Studies of Alternative Exchange Rate Systems. An Intertemporal General Equilibrium Approach*, 1982**
14. **Erzan, Refik *Turkey's Comparative Advantage, Production and Trade Patterns in Manufactures. An Application of the Factor Proportions Hypothesis with Some Qualifications*, 1983**
15. **Horn af Rantzien, Henrik *Imperfect Competition in Models of Wage Formation and International Trade*, 1983**
16. **Nandakumar, Parameswar *Macroeconomic Effects of Supply Side Policies and Disturbances in Open Economies*, 1985**
17. **Sellin, Peter *Asset Pricing and Portfolio Choice with International Investment Barriers*, 1990**
18. **Werner, Ingrid *International Capital Markets: Controls, Taxes and Resources Allocation*, 1990**
19. **Svedberg, Peter *Poverty and Undernutrition in Sub-Saharan Africa: Theory, Evidence, Policy*, 1991**
20. **Nordström, Håkan *Studies in Trade Policy and Economic Growth*, 1992**
21. **Hassler, John, Lundvik, Petter, Persson, Torsten and Söderlind, Paul *The Swedish Business Cycle: Stylized facts over 130 years*, 1992**
22. **Lundvik, Petter *Business Cycles and Growth*, 1992**
23. **Söderlind, Paul *Essays in Exchange Rates, Business Cycles and Growth*, 1993**
24. **Hassler, John A.A. *Effects of Variations in Risk on Demand and Measures of Business Cycle Comovements*, 1994**
25. **Daltung, Sonja *Risk, Efficiency, and Regulation of Banks*, 1994**
26. **Lindberg, Hans *Exchange Rates: Target Zones, Interventions and Regime Collapses*, 1994**
27. **Stennek, Johan *Essays on Information-Processing and Competition*, 1994**
28. **Jonsson, Gunnar *Institutions and Incentives in Monetary and Fiscal Policy*, 1995**
29. **Dahlquist, Magnus *Essays on the Term Structure of Interest Rates and Monetary Policy*, 1995**
30. **Svensson, Jakob *Political Economy and Macroeconomics: On Foreign Aid and Development,* 1996**
31. **Blix, Mårten *Rational Expectations and Regime Shifts in Macroeconometrics,* 1997**
32. **Lagerlöf, Nils-Petter *Intergenerational Transfers and Altruism,* 1997**
33. **Klein, Paul *Papers on the Macroeconomics of Fiscal Policy,* 1997**
34. **Jonsson, Magnus *Studies in Business Cycles,* 1997**
35. **Persson, Lars *Asset Ownership in Imperfectly Competitive Markets,* 1998**
36. **Persson, Joakim *Essays on Economic Growth*, 1998**
37. **Domeij, David *Essays on Optimal Taxation and Indeterminacy,* 1998**
38. **Flodén, Martin *Essays on Dynamic Macroeconomics,* 1999**
39. **Tangerås, Thomas *Essays in Economics and Politics: Regulation, Elections and International Conflict*, 2000**
40. **Lidbom, Per Pettersson *Elections, Party Politics and Economic Policy*, 2000**

41. **Vestin, David** ***Essays on Monetary Policy,*** **2001**
42. **Olofsgård, Anders** ***Essays on Interregional and International Political Economics,*** **2001**
43. **Johansson, Åsa** ***Essays on Macroeconomic Fluctuations and Nominal Wage Rigidity,*** **2002**
44. **Groth, Charlotta** ***Topics on Monetary Policy,*** **2002**
45. **Gancia, Gino A.** ***Essays on Growth, Trade and Inequality,*** **2003**
46. **Harstad, Bård** ***Organizing Cooperation: Bargaining, Voting and Control,*** **2003**
47. **Kohlscheen, Emanuel** ***Essays on Debts and Constitutions,*** **2004**
48. **Olovsson, Conny** ***Essays on Dynamic Macroeconomics,*** **2004**
49. **Stavlöt, Ulrika** ***Essays on Culture and Trade,*** **2005**
50. **Herzing, Mathias** ***Essays on Uncertainty and Escape in Trade Agreements,*** **2005**
51. **Bonfiglioli, Alessandra** ***Essays on Financial Markets and Macroeconomics,*** **2005**
52. **Pienaar, Natalie** ***Economic Applications of Product Quality Regulation in WTO Trade Agreements,*** **2005**
53. **Song, Zheng** ***Essays on Dynamic Political Economy*****, 2005**