A Probabilistic Voting Model of Social Security: The Role of Myopic Agents*

Xue WEN†

DEFAP - Università Cattolica del Sacro Cuore

Oct. 2012

Abstract

This paper investigates the political incentives for the design of social security policy in competitive democracies with both far-sighted and myopic households. The social security scheme depends on both a payroll tax rate which determines the size of the pension and a Bismarckian factor that represents its redistributive component. By considering a probabilistic voting setting of electoral competition, we analyze the political game between office-seeking politicians and self-interested citizens. Politicians can win the election by targeting the voters in each group by trading off the generosity and the redistribution degree of the public pension system. In the political equilibrium, the contribution rate is U-shaped with respect to the Bismarckian factor. Moreover, the equilibrium Bismarckian factor unambiguously decreases with the proportion of myopic agents, whereas the equilibrium payroll tax rate curve is U-shaped with respect to the proportion of myopic agents.

Keywords: Myopia, Social Security, Probabilistic Voting

JEL Classifications: H55, D91, H30

---

* Preliminary versions of this paper have been presented at the RES Ph.D Meeting 2012, Jounées Louis-André Gérard-Varet Conference in Public Economics 2011, the International Workshop on Pension, Insurance and Saving in University of Dauphine 2011, Scottish Economic Society Conference 2011, The Universities of Montpellier, DEFAP doctoral Seminar and Torino (Collegio Carlo Alberto). I am grateful for helpful comments and discussions from Massimo Bordignon, Natale Piergiovanni, Michele Santoni, Philippe De Donder, Fabian Kindermann, Aniol Llorente-Saguer and the seminar participants. All remain errors are my own.

† Email: wenxue1982@gmail.com or xue.wen@unimib.it.

1. Introduction

The optimal design of social security policy remains a classic issue in public economics. Previous studies addressing the effects of social security to people's well-being mainly adopt the standard consumer behavior assumption - the life-cycle hypothesis (Ando and Modigliani, 1963), which states that consumers naturally smooth their consumption over life time. However, emerging evidence and empirical studies suggest that in reality consumer behavior does not always abide such a rule.[1] One of the key characteristics is that people often behave myopically when facing both short-term and long-term choices, even though they regret not having saved enough afterwards. Behavioral economists developed several frameworks in modeling shortsighted consumers with respect to different non-standard preferences. To new a few, Laibson (1997) proposed the quasi hyperbolic discounting framework which was previously developed by Strotz (1956) and Phelps and Pollak (1968). Gul and Pesendorfer (2001) define temptation preferences over both allocations and choices sets, which implies that under-saving is the consequence of people getting tempted by small earlier rewards. If some choices feel tempting when they are available, and if this detracts from well-being, then an individual may prefer small choice sets to large ones. Instead, Feldstein (1985) and Cremer et al. (2007, 2008) widely applied the myopic approach to the optimal design of social security. The myopic approach is a simpler framework in which agents are myopic in that they do not save at all and focuses on individuals' self-reported mistakes between their choices and their true preferences. As Pestieau and Possen (2006) put it: "[m]yopic individuals do not want to save because their immediate 'self' incites them to get instant gratification". In particular, this paper follows the assumption in myopic sophisticated agents[2] à la

[1] Many studies suggest that consumption declines sharply at retirement (See e.g Hammermesh, 1984; Banks et al., 1998). Behavioral economists suggest that individuals are not perfectly rational and sometimes have trouble behaving in their own best interests. Explanations provided for the non-standard saving behavior stress that individuals are tempted to spend and that saving requires effect and self-control. (See e.g., O'Donoghue and Rabin, 1999; Thaler and Shefrin, 1981). Moreover, the last few decades have witnessed a sharp decline in the saving rate for most developed countries. Based on the U.S data, Laibson (1997) suggests that financial innovation may have caused the ongoing decline in the savings rates, since financial innovation increases liquidity, eliminating commitment opportunities.

[2] The alternative to sophisticated myopic agents is naïve myopic agents. The notion of "naïve and sophisticated

Cremer et al. (2007, 2008). In the rest of the paper we use the term "myopic" instead of "myopic sophisticated" for simplicity.

Ever since the seminal contribution by Feldstein (1985), an increasing number of papers have incorporated the role of myopic behavior in analyzing the public pension systems.[3] Typical of this literature takes the normative view focusing on the paternalistic role of the government that uses a mandatory pension system in providing a commitment service and redistribution mechanism. Truly there are a considerable amount of studies explore the political economy of social security, however, none of them adequately addresses the presence of myopic agents.[4] Thus, this paper aims to study the political implications of incorporating the myopic consumption behavior into the optimal design of social security.

Moreover, we apply the probabilistic voting approach. Note that the existing positive literature on social security mainly adopts the majority voting approach. This approach is limited to only one characteristic of social security[5], given that the median voter theorem yields no prediction when the policy space is multi-dimensional.[6] This is rather restrictive, since the two main distinct characteristics of a pension system are

---

myopic agents" was first brought up by Strotz (1956) and Pollak (1968). Strotz (1956) suggested that people are more impatient when they make short-run trade-offs than when they make long-run trade-offs. A person could be sophisticated and know exactly what her future selves' preferences will be. Alternatively, a person could be naïve and believe her future selves' preferences will be identical to her current self's, not realizing that as she gets closer to executing decisions her tastes will have changed. He argues that in order to avoid such time inconsistency, agents could decide in a sophisticated way, making an analysis of what their actions would be in the future, as a consequence of their changing preferences. In the literature, it has assumed either one or the other behavior, without attempting to justify the choice on behavioral grounds. For instance, Akerlof (1991) assumes naive beliefs, while Laibson (1997, 1998) assume sophisticated beliefs. Harris and Laibson (2008) introduce a continuous-time model of quasi-hyperbolic time preferences to solve the intra-personal game for the sophisticated agents. Cremer et al. (2007) discuss the optimal social security for an economy consisting of both farsighted and myopic sophisticated voters through majority voting. It is worth mentioning that Hey and Lotito (2009) came up with an experimental study analyzing the dynamic inconsistent makers' attitudes towards dynamic inconsistency distinguishing subjects being naïve, resolute or sophisticated. They end up in questioning the applicability of the sophisticated myopic approach. However, given the widely application of this assumption in the theoretical literature and it seems to help generating predictions that are in line with empirical evidence on saving behavior (see e.g. Cremer et al., 2007, 2008). Therefore, one would need more experimental work before dismissing/accepting this behavioral assumption. The jury is still out.

[3] A vast literature on myopia and social security takes a normative approach, focusing on the role of myopic behavior in justifying the existence of a publicly provided pension system. More specifically, the existing literature point out that myopia is a rationale for the public pension system. See E.g., Feldstein (1985), Docquier (2002) and Andersen and Bhattacharya (2010).

[4] See Galasso and Profeta (2002) for a full review. To the author's knowledge, Cremer et al. (2007) is the only attempt studying the political economy implications of social security with myopic agents in the context of majority voting.

[5] Most of this literature focuses sorely on the payroll tax rate determining the size of the social security system.

[6] See for instance, Casamatta et al. (2000b) and Persson and Tabellini (2000).

contributions and benefits.[7] Furthermore, in a political game, the Bismarkcian factor (the redistribution degree of the social security) is no less crucial than the payroll tax rate (the size of the social security). Given that the Bismarckian factor can reduce inequality, it is essential element to target the lower income voters. The problems associated with the existence of equilibrium in high-dimensional policy spaces are well known (Plott, 1967). Nevertheless, probabilistic voting theory is considered as being an appropriate solution to these problems.[8] In addition, the probabilistic voting model allows us to analyze entrenched voter positions that are independent of personal benefits derived from economic policy. Unlike the traditional way in modeling which preferences focusing solely on the economic benefits one receives, probabilistic voting also considers ideological positions unrelated to personal economic concerns. This implies that some people's preferences are 'worth' more than others. One voter may cast his ballot in favor of a candidate that brings him lower economic benefits in exchange for a non-economic issue that for him outweighs the economic loss. This non-economic issue is referred to an "ideology" that is orthogonal to the fundamental policy dimensions of interests.[9]

To be more specific, the general setting of our model follows Feldstein (1985) in assuming that all individuals have equal earnings and identical tastes but differ in their behavior types (being either farsighted or myopic). There is no account for uncertainty in future rates of return, rates of population growth, productivity or demographics. It is assumed that individuals hold rational expectations and anticipate the effect of policy on equilibrium outcomes. We consider a simple two-period overlapping generations model, with people working in the first period and retiring in the second. Households are non-altruistic. The voting stage occurs before the market stage. Voters are characterized by their behavior types and their political ideologies. Behavior types are divided into two interest groups, myopic group and the farsighted

[7] Cremer et al. (2007) discuss both the size and the redistribution degree of the social security in a majority voting setting. The voting occurs in a sequential way. In the first stage, voters vote either for a pure Bismarckian system or a pure Beveridgean system, then, they vote for the payroll tax rate.
[8] See e.g. Lindbeck and Weibull (1987), Coughlin (1992), Dixit and Loudregan (1996) (1998), Persson and Tabellini (2002). By introducing uncertainty to the voting process, it incorporates a random element to any voter's utility calculation, which is modeled through the mechanics of probability theory.
[9] These characteristics are permanent and cannot be credibly altered in the course of electoral competition.

group.[10] Foresighted households adjust their private savings and labor supply decisions by rationally anticipating the wealth implications of the social security policy. Instead the myopic individuals focus solely on current consumption and save nothing in the market stage. Note that the problem with the myopes is not being shortsighted per se, but lacking of self-control when facing saving and labor supply decisions in the first period. However, the myopes are sophisticated in the voting stage that they vote with a long-term rationale despite the fact they suspect their behavior with nonetheless be myopic in the first period.[11]

The social security policy is determined in the probabilistic voting setting. Two office-seeking politicians enter into electoral competition. Voters vote for the Bismarckian factor and the payroll tax rate simultaneously. Voters care about both the social security policy and their own political ideologies. The farsighted and the myopic groups hold different political homogeneity representing their ability to exert political pressure in acquiring their preferred social security policy. There is no presumption on whether one group is more political homogenous than the other.[12]

The focus of the current model is on the sensitivity of the two specific groups' (myopic and farsighted) voting behavior with respect to different social security policies. Targeting the votes of the two groups can be done through two channels: the size (positively related to the contribution rate τ) and the redistribution degree (negatively related to the Bismarckian factor α). These two parameters differ in targeting the two groups and can be used as substitute tools. The unique Nash-equilibrium can be characterized analytically.[13] In the equilibrium the

[10] The standard approach is to consider two interests groups composed the young and old generations, respectively. For instance, Profeta (2002) presented a probabilistic voting model investigating a political game between the young and the old group. However, that is not the focus of this paper.

[11] The assumption of myopic sophisticated agents is crucial for generating a political equilibrium with a positive pension system. When we consider the myopic agents to be fully myopia that they are myopic not only in the market stage but as well in the voting stage, they would simply vote for no social security. On the other hand, the farsighted agents prefer private saving than social security. Thus, in the fully myopia setting, no social security will be implemented in the political equilibrium.

[12] Some may argue that the farsighted care less about social security given that they naturally smooth consumption overtime. However, this statement is fragile given that social security distorts their labor supply and, consequently, their consumption levels. Naturally, the voters inside the myopic group pay attention to social security policy as they expect the government forcing them saving and compensating them when retired. Therefore, we do consider both cases in which either the farsighted or the myopic are (exogenously) more politically homogeneous than the other group.

[13] To focus on the effects of myopia instead of the intergenerational transfers or the capital accumulation, we assume zero population growth and zero interest rate. Note that the assumption on the equality between the interest

contribution rate and the Bismarckian factor are determined endogenously. This differs from the median voter setting in Cremer et al. (2007), where voters vote sequentially first between a pure Bismarckian system and a Beveridgean system, then, the payroll tax rate for the chosen system. Furthermore, we characterize the closed-form solution with the logarithmic utility function. The results of the theoretical model are numerically illustrated in the simulation part.

To preview the main results, the payroll tax rate is a U-shaped function with respect to the Bismarckian factor in the equilibrium with a turning point at a high value of the Bismarckian factor. This implies that the tax rate will be higher under both a less Bismarckian system (in which pension benefits are less related to one's contributions) and a pure Bismarckian System (in which pension benefits fully depend on one's first period contributions) than in systems that are close to the pure Bismarckian system.[14] The underlying economic mechanism is that social security does not only create redistribution in favor of the myopic group, but also causes distortions to the labor supply for all. The distortion for the myopic agent's labor supply is higher and is only related to the payroll tax rate. Instead, the farsighted agents are less distorted as their labor supply decisions are influenced by both the payroll tax rate and the Bismarckian factor. The lower the payroll tax rate, the higher is the Bismarckian factor, the lower is the distortion. The U-shaped relationship between the payroll tax rate and the Bismarckian factor results from the trade-off between efficiency and redistribution.

---

rate and the population growth rate is crucial for our results as we look into the steady-state political equilibrium instead of the dynamic political equilibrium. Our model is restricted as we consider only one single generation and that voting happens only once. In a standard OLG model setting, with our assumptions on interest rates and population growth, the political equilibrium arises in a steady-state sequence where only the young vote and make the presumption that the system they adopt will as well be adopted by the next generation. On the contrary, see for example below the literature exploring the inter-generational game in a dynamic economy. Cooley and Soares (1999) and Boldrin and Rustichini (2000) stress the role of general equilibrium effects to explain why the politico-economic equilibrium is in favor of intergenerational transfers. Recently, Gonzalez-Eiras and Niepelt (2008) provide an application of the probabilistic voting model in a standard OLG model, focusing on the effect of the projected demographic transition on the politico-economic dynamic equilibrium for social security.

[14] Note that in our model, there is no room for a pure Beveridgean system (in which benefits are unrelated to a worker's contribution). This is closed related to two assumptions: first, we assume homogeneous productivity across individuals. Therefore, income inequality only depends on one's behavior type; Second, the individual preferences take an inseparable form with respect to consumption and labor supply. Therefore, when myopes exist in the economy, the farsighted always preferred policy is a purely Bismarckian system, while the myopic preferred policy is an impure contributive system, given that they are aware that farsighted individual's labor supply is less distorted if the pension system is more Bismarckian. At this stage, we assume away the productivity heterogeneity to focus on the effects of myopia and political preferences. This is because the effects of uncertainty in productivity will compound with the effects of the uncertainty in myopia and political preferences in the optimal design of social security. However, we leave to future work a more complete analysis of the interaction between these factors.

Given that the farsighted agents provide more labor supply than the myopes, the redistribution component in social security is therefore in favor of the myopic group. A society with more myopic agents therefore prefers a more redistributive pension system. Instead, the payroll tax rate curve is U-shaped with respect to the size of the myopic agents. This link again comes from the tradeoff between redistribution and efficiency.

The rest of this paper is structured as follows: Section 2 introduces the model and characterizes the equilibrium results. A numerical simulation applying the logarithmic utility function is conducted in section 3. The last section concludes.

2. The model

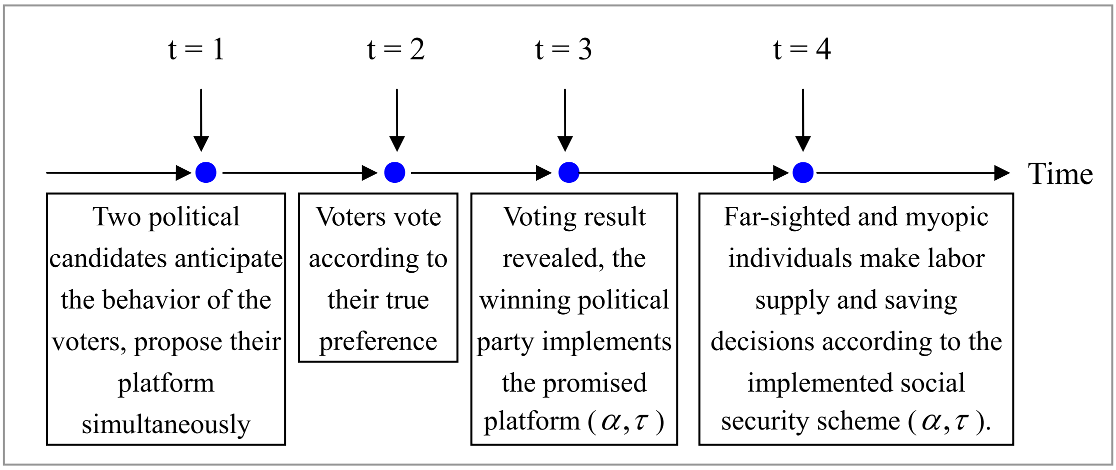
2.1 Stages of the game

Two political parties represented by two office-seeking candidates enter into electoral competition. Both candidates choose the social security policy platforms simultaneously and do not cooperate. Politicians credibly commit to their own social security policy platforms. The social security policy includes two elements: the payroll tax rate representing the size of the pension system, and the Bismarckian factor that measures the contributory parameter of the pension system. Each party holds fixed ideological positions non-related with economic policy such as positions on value issues. The party that obtains the majority of the votes wins the election and implements the promised policies.

Voters differ in two dimensions: their behavior type as farsighted or myopic agents and in their political ideology biases. Voters are motivated by both these two categories; private income and consumption on one hand, and political ideological preference on the other. Citizens may either be heterogeneously biased towards a certain party or share a common valuation of the competing parties. People vote for the party that suits his own interests, given promised economic policies as well as their own political ideological views.

Figure 1 shows the timing of the events as well as the information sets for both the political parties and the voters.

Figure 1: The Timing of the Political Game



More precisely, the political game has four sequential stages:

Stage 1 – Policy Announcements:

Taking its opponent's policy platform as given, each political candidate chooses a social security policy platform to maximize its chances of winning the election. Parties take into account the political preferences of citizens, anticipating the labor supply choice and saving pattern for each type of agents (far-sighted and myopic).

Stage 2 – Election:

During the election, the far-sighted and myopic voters vote according to their true preference given the economic policy announcements and their ideological biases. It is worth noting that when voting, the myopic sophisticated agents consider their long-time welfare anticipating that their myopic-self will emerge in the last stage when the platform is implemented. They are purely myopic when making labor supply and saving decisions in the last stage that they disregards their second period utility at this stage.

Stage 3 – Policy Implementation:

We assume the winning political candidate actually implements the policies proposed during the election campaign.

Stage 4 – Individual Decision Making:

Individuals have two roles, one as a voter and one as a participant in the labor market and benefits system. We employ a standard two-period framework with time-separable utility, wherein individuals make only two decisions in labor supply and savings. The farsighted agents naturally smooth their consumption between the two periods. Consumption decisions are assumed to exhibit a simple form of myopia for the myopes, thereby they focus merely in the present period consumption and save nothing. Each individual pays a linear income tax for social security in the first period and expect to receive a pension benefit in the second period, which consists of a flat benefit and a contributive component associated with his first period contribution.

Regarding the information sets, the political parties are informed of the size of both the myopic and farsighted groups. Parties are not informed of the party attachment of the each voter, but are aware of the distribution of relative ideological biases of each group. The voters are only aware that there are two types of agents in the economy, however, without information about either the size or the political preferences of both groups.[15] The lack of information on size is crucial for our results as it contributes to the individual's uncertainty about future pension benefits, given that it depends on the size of both groups.

We characterize the local equilibrium of the political game by backward induction.

2.1.1 Fourth Stage: Individual Decision Making

Individuals need to make decisions in two stages: the voting stage and the market stage. The voting stage happens before the market stage. It is worth mentioning that the myopic individuals are only myopic in the market stage but sophisticated in the

[15] Here we assume that the political candidates are more informed than the voters, since they are more privileged in obtaining information through organization or agencies while the normal people can not easily access such information individually. Moreover, the political parties are motivated to pay higher information costs to public organization or consulting firm so as to get elected. Instead, although a single individual might care about the social security policy, he is either too lazy (acting like a free rider) or finds it too costly to obtain such information.

voting stage, however, anticipating that they will behave myopically in the market stage.

Market Stage

We consider a simple two-period overlapping generations economy inhabited by groups of representative agents. Population consists of two types: the far-sighted rational agents and the short-sighted myopic agents. The rate of population growth and the interest rate equal zero. People produce a single good which can be consumed or saved as storage goods for future consumption. The production function is linear in labor. There are two markets in the economy: a labor market and a consumption good market. We impose market clearing conditions in each period.

Individuals live for two periods: old and young. They work in the first period and retire in the second period. Generations are unlinked, meaning that households are non-altruistic. We denote the population by $i = \{F, M\}$ where F represents the far-sighted agent and M stands for the myopic agent. The proportion for each group is λ^F and λ^M, $\lambda^F + \lambda^M = 1$. Productivity is assumed to be homogenous across population.[16] Unlike Cremer et al. (2007) where one's income is formalized of the productivity and the labor effort decision, here one's income is only the result of labor effort decision. Moreover, in Cremer et al. (2007) agents differ in myopia and productivity, in our model individuals differ in myopia and the political ideology.

The intertemporal utility function representing the "true preference" for the representative agent in group *i* is:

$$U_i = u(x_i) + \gamma_i u(d_i) = u(c_i - l_i^2/2) + \gamma_i u(d_i) \tag{1}$$

where $c, d \in \mathbb{R}$ denote the first and second period consumption, l is the labor supply in the first period. Note that quadratic disutility of labor is not crucial, what

[16] Firstly, we assume homogenous productivity as so to focus on the intragenerational transfer related to myopia but not on income inequality. In fact, if we introduced heterogeneous productivity we will have labor distortion effects not only related to being myopia but also related to wage levels, which would make it difficult to single out the role of myopia. Moreover, wage heterogeneity will complicate the political ideology biases of the voters which currently are only related to their behavior types (being farsighted or myopic). This will in turn introduce multiple effects to the optimal path of the social security policy. I doubt if it is feasible to introduce heterogeneous productivity in the probabilistic voting setting. However, trying to incorporate heterogeneous productivity into the framework would be a good direction for future research.

makes a difference is the quasi-linear specification that assumes away income effects.[17] The utility function U satisfies the following assumptions:

Assumption 1: U is twice continuously differentiable.

Assumption 2: $U_x > 0, U_d > 0, U_{xx} < 0, U_{dd} < 0$

Assumption 3: $U_{xd} > 0$

Assumption 1, 2 are rather standard, ensuring differentiability and strong monotonicity conditions on U, respectively. Assumption 3 is a sufficient condition to ensure that x and d are normal goods.

$U^{'}(x_i, d_i)$ is the marginal utility of an extra transfer of consumption for group *i*. It captures the degree of diminishing returns to private consumption. The higher is the value of marginal utility of consumption for group *i*, the easier the parties can attract the votes of group *i* by offering small economic benefits.

The parameter γ_i represents the preference for future consumption for each type, where γ_F =1 for the far-sighted and for the myopes γ_M =0 ex ante but γ_M =1 ex post. When people make decisions on labor supply and consumption/saving, the far-sighted do so with γ_F =1 but myopes do so according to their ex ante preference where γ_M =0. This implies that the far-sighted agents naturally smooth their consumption overtime, make the optimal labor supply and saving decisions considering their lifetime welfare whereas the myopic agents consume all the disposable income and save none in the first period, but regret for not saving enough when they retire in the second period.

Given that we assume homogenous productivity, wage is normalized to w =1 and we drop it henceforth. The budget constraints in each period are derived accordingly:

---

[17] In this setting, leisure is not taken into account as a normal good. See e.g. Sommacal (2006) for a detailed discussion of the role of labor supply in evaluating the redistributive impact of a pension system with different utility functions.

$$x_i + s_i = (1-\tau)l_i \tag{2}$$

$$d_i = s_i + p_i \tag{3}$$

Eq. (2) describes that an individual's first period net income after tax, $(1-\tau)l_i$, is equal to the net consumption, x_i, plus the amount of saving, s_i. Eq. (3) indicates that an individual's consumption in the second period, d_i, depends on his first period saving, s_i, plus the pension benefits after retirement, p_i. The pension, p_i, consists of a contributive component and a redistribution part:

$$p_i = \tau[\alpha l_i + (1-\alpha)El] \tag{4}$$

$$El = l_F E\lambda_F + l_M E\lambda_F \tag{5}$$

El is the average before-tax income and E is the expectation operator. The parameter $\alpha \in [0,1]$ is often referred to as the Bismarckian or the redistribution factor.[18] In a pure Beveridgean pension system ($\alpha = 0$), $p_i = p = \tau Ewl$, and each individual receives a flat social benefit. The higher α is, the more Bismarckian the social security becomes. When $\alpha = 1$, we have $p_i = \tau w_i l_i$, a pure Bismarckian system indicating that one's pension is proportional to his contribution.

Each voter votes sincerely that he votes based on his private information. The voters are not aware of the exact amount of the far-sighted and the myopic agents in the economy. However, they do acknowledge the distribution function for each type agent. We assume λ^F and λ^M are random variables which are uniformly distributed between 0 and 1:

$$\lambda^F, \lambda^M \sim U\ [0,\ 1]$$

Therefore, in the voting stage both type agents anticipate that $E(\lambda^F) = E(\lambda^M) = \dfrac{1}{2}$. Accordingly, the pension benefits represented in Eq. (4) can be rewritten as follows:

[18] The Bismarckian factor (e.g. Cremer and Pestieau, 1998; Hassler and Lindbeck, 1997) is employed to estimate the level of intragenerational redistribution in the public pension system, Conceptually, the Bismarckian factor divides the pension benefit into a flat component and into an earnings-related component.

$$p_i = \tau[\alpha l_i + (1-\alpha)\frac{1}{2}(l_F + l_M)] \tag{6}$$

Note that in the market stage, the farsighted individuals consider the effects of pension when they are making saving and labor supply decisions. Instead, the myopes do not take into account of this effects in the market stage given they do not save anyway. However, they anticipate the effects of pension in the voting stage.

We solve the optimization problem for the far-sighted agent and myopic agent separately. By substituting Eq. (2), (3) and $\gamma_F = 1$ into Eq. (1), the problem of a far-sighted individual is thus given by:

$$\max_{s_F, l_F} U_F = u((1-\tau)l_F - l_F^2/2 - s_F) + u(s_F + p_F) \tag{7}$$

From the FOCs, yields the optimal level of labor supply and savings for the far-sighted:[19]

$$l_F = (1-\tau) + \tau\alpha \tag{8}$$

$$s_F = \frac{1}{2} - 2\tau + \frac{3}{2}\tau^2 - \tau^2\alpha^2 - \frac{1}{2}\tau^2\alpha \tag{9}$$

We assume the individuals face no borrowing constraints[20]. The farsighted individuals smooth their consumption across time. s_F is negative when the farsighted agents need to borrow their future pension from the government. The government tax the young generations to finance the pension system for the old.

The myopes do not save for retirement and choose labor supply to maximize the following utility:

$$\begin{gathered} \max_{l_i} u[(1-\tau)l_i - l_i^2/2] \\ s_M = 0 \end{gathered} \tag{10}$$

The myopic agent's labor supply is:

$$l_M = 1 - \tau \tag{11}$$

Directly derived from Eq. (8) and (11), we find that the pension system causes

---

[19] See appendix A.1 and A.2 for derivation.

[20] With borrowing constraints, the farsighted agents prefer either no social security or a pure contributive social security system with the payroll tax rate not exceeding ¼. Therefore, for a pure Bismarckian system when the payroll tax rate is higher than ¼, we simply consider the farsighted behave as the myopes. Accordingly, this assumption yields "less interesting" political equilibria.

distortions to people's labor supply. Moreover, the myopic agents are more distorted than the far-sighted agents. The distortion on the labor supply of the myopic group is merely subject to the changes in the payroll tax rate. Instead, the far-sighted see the link between the contribution component of the social security and his labor income, thus, anticipate as well the effects of the social security's redistribution degree when providing labor.

Given that $E(\lambda_F) = E(\lambda_M) = 1/2$, substituting Eq. (8) and (9) into Eq. (1), the indirect utility for a far-sighted agent with respect to the payroll tax rate and the Bismarckian factor is obtained as:

$$V^F(\tau,\alpha) = 2u[\frac{1-(1-\alpha)\tau^2}{4}] \tag{12}$$

The indirect utility function with respect to (τ,α) for a myopic agent is:[21]

$$V^M(\tau,a) = u[\frac{(1-\tau)^2}{2}] + u[\tau(1-\tau) + \frac{\alpha(1-\alpha)\tau^2}{2}] \tag{13}$$

Voting Stage

In the voting stage, an individual choose to vote for the social security policy that is closer to his own preference. We characterize the preference over social security policy for both types of agents by optimizing the indirect utility functions with respect to (τ,α) from Eq. (12) and (13)[22]. The results are summarized in the following proposition.

Proposition 1: *The group-specific preference is such that*

i) the most preferred social security policy of the farsighted group (τ_F^*, α_F^*) *is given by:*

[21] This indirect utility function is the ex post utility function for the myopic group, which reflects its true preference.

[22] See Appendix A.3 for the derivation.

$$\begin{cases} \alpha_F^* = 1, \ \tau_F^* \in [0, \ 1] \\ No \ social \ security \end{cases}$$

ii) *the most preferred social security policy of the myopic group* (τ_M^*, α_M^*) *is given by:*

$$\alpha_M^* = \frac{1}{2}, \ \tau_M^* = \frac{u'(x_M) - u'(d_M)}{u'(x_M) - \frac{7}{4}u'(d_M)}$$

iii) *for the myopic group, the relation of the most preferred payroll tax rate to the Bismarckian factor is:*
$$\begin{cases} \frac{\partial \tau_M^*}{\partial \alpha} > 0 \quad when \quad \alpha > \frac{1}{2} \\ \frac{\partial \tau_M^*}{\partial \alpha} < 0 \quad when \quad \alpha < \frac{1}{2} \end{cases}.$$

Proof: see the Appendix A.3 for derivation.

As we mentioned before, the social security affects the farsighted group in two ways: distortion effect in labor supply and redistribution effect of the labor income from the farsighted group to the myopic group. The former effect lowers the farsighted individual's private consumption in both periods. The latter effect makes private saving more favorable than social security for the farsighted. Accordingly the first point of proposition 1 shows that the farsighted either prefers no social security or votes for a contributive (Bismarckian) pension system and do not care about the size of the system or simply vote for no social security. Note that in a pure Bismarckian system, the farsighted are free from the above two effects. The government simply acts like a saving institution.

Now we turn to the second point in proposition 1 for the myopic group. Given that the far-sighted agents pay higher contribution to the pension system than myopic agents, one may intuitively presume that the myopic agents prefer a more redistributive social security system. However, the transfer to the myopic group depends both on the redistribution degree of the pension system and the labor income of the farsighted group. In deed, a more redistributive pension system entitles the myopic agents to a higher proportion of the whole pension benefits. Note that the tax base depends on the labor income of the farsighted group, which is related to the Bismarckain factor as well. A more redistributive system causes more distortion to the labor supply of the farsighted group. The tax base is reduced accordingly. In fact when the Bismarckian factor collapses to zero, the far-sighted group's labor supply is as distorted as that of the myopic agents, implying there is no transfer from the farsighted to the myopic group. Therefore, the most preferred Bismarckian factor is $\frac{1}{2}$ resulting from a trade-off between efficiency and redistribution concerns. The most preferred payroll tax for the myopic group acts more like a consumption smoothing tool and is derived accordingly with respect to the most preferred Bismarckian factor.

Lastly, for the myopic group, the most preferred payroll tax rate increases with the Bismarckian factor when it is less than $\frac{1}{2}$ but is negatively related to the Bismarckian factor when it exceeds $\frac{1}{2}$. The intuition is as follows: the optimal Bismarckian factor preferred by the myopic households is $\frac{1}{2}$, thus, the further it derives from this value, the less transfer the myopes will receive from the farsighted in the second period, thus, a higher payroll tax rate is required to smooth consumption.

2.1.2 Second stage: Voting

We apply the probabilistic voting that replaces the discrete and deterministic utility maximizing problem in the classic deterministic model with a continuously measured probabilistic calculation. In other words, voters' decisions are discrete under deterministic assumptions but continuous under probabilistic assumption. Voters are ideological, candidates are uncertain in how voters' political preference can be translated into party preferences. Voters care about the competing parties' policies as well as an ideological dimension unrelated to the policies.[23] Probabilistic assumptions

[23] See e.g., Lindbeck and Weibull (1987) or Persson and Tabellini (2000).

are reasonable when we consider that candidates are rarely certain about voters' preferences. This probabilistic element is to account for either inadvertent or deliberate uncertainty in the expected return on voting for a special candidate or proposal. This may arise if the candidates are too vague in their attitude towards key issues. Uncertainty may also arise from non-policy considerations such as personal qualities; complex issues such as the imperfect mapping of voter alternatives due to limited information; and other unpredictable factors such as technical mistakes.

In our model, uncertainty at the voting stage comes from two separate factors: voter idiosyncratic ideological bias and pension benefit. The first issue is unknown to political parties on an individual voter basis, as it is the privately held opinion of the person. The second issue unknowable to a single voter. As he is not able to fully predict the tax base for social security. Given that the tax base depends both on the labor supply and the relative size of myopic and farsighted household wealth, one individual is not necessarily informed about the size of each group. Thus, political parties must maximize their vote totals assuming that voter behavior is probabilistic, given their inaccurate information on voters' political ideologies as well as the voters' lack of knowledge on the size of the tax base.

Every individual is assigned a specific and potentially unique-probability of voting for each candidate. During the voting process, voters judge both the policy platforms and the non-policy criteria. Voters are not all alike in the eyes of the competing candidates: some voters are more important than others because they are more likely to reward policy favors with a vote. Electoral competition leads to equilibrium, with both candidates maximizing a particular social welfare function. Different voters are weighted by their "responsiveness," that is, how likely they are to reward policy favors with a vote.

The degree of political ideological bias in each group is captured by a density function. In our setting, the density functions which capture the distribution of political preferences of voters differ for the myopic and the rational groups. This density function determines the political homogeneity of each group, since more ideologically homogenous groups are more politically successful. Without loss of generality, we do not assume that the myopic individuals are more politically homogenous than the farsighted per se.[24]

We characterize the voters' response to social security policy separately for the two types of agents. As mentioned earlier, the myopic agents can be seen to have two selves, one looks for immediate gratification and one looks for long term welfare in the traditional "dual-self" models (see e.g. Cremer et al. 2007, 2008). It is their myopia side when making saving and labor supply decisions, but their rational side which emerges when they vote.

Two identical political parties are in competition. The set of candidates is denoted by $N \subset \mathbb{R}^2(A, B)$. The policy announced by candidate N is P_N. Before the election takes place, each party commits to its own policy platform P^A and P^B. Each party is seeking office and chooses the platform that maximizes its probability of winning, interpreted here as its expected number of votes. The parties make binding commitments to social security policy platform during the electoral campaign. All voters vote sincerely on the two policy alternatives. We look for the Nash equilibrium, namely a situation in which each candidate chooses a configuration of strategies such that neither party can discontinuously increase his probability of winning.

Voters differ both in their behavior type (far-sighted or myopic) and their underlying ideological affinities. We allow a whole spectrum of attitudes toward social security within each group, thus, the members of a given group are not politically identical with the same weights on candidates. σ^{ji} is the idiosyncratic ideological bias towards party A of individual j from group i. Voters with $\sigma^{ji} = 0$ care only about economic policy. We allow for these differences by letting the whole

[24] As already mentioned in footnote 32, some may argue that myopic agents are more homogenous as they care more about redistributive policies and are more dependent on social security to force themselves' saving for retirement, others declare although the rational agents do not care so much the size of pension system, however, they do care about the redistribution degree of the system as they pay the contributions. Although it is intuitively that the myopic agents are more sensitive to public pension system when they vote for two reasons: first, they are in need of a public pension system to force themselves saving; second, they desire acquiring transfers from the far-sighted individuals. On the other hand, some may argue that farsighted agents are sensitive to the social security policy as well since it influences their choices in labor supply and savings due to consumption smooth concern. Therefore, we do not want to make the assumption if one group is more politically homogenous than the other per se, instead, we consider all the possible situations.

functional form of the distribution σ^{ji} depend on *i*. It is group specific and uniformly distributed:

$$\sigma^{ji} \sim U\ [-\frac{1}{2\phi^i}, \frac{1}{2\phi^i}]$$

ϕ^i is to represent the level of "political homogeneity" of the group i. In the voting stage, considering every citizen allocating his political preferences among different issues, higher ϕ^i means that individuals in group *i* are more homogeneous in their political action, focusing on a single "issue", which in our model is the social security policy. A group with a higher ϕ^i is mostly influenced by policy as opposed to ideology, thus, will get a bigger weight. We as well allow for the distribution functions to be identical for the two groups, which would be a special case in our analysis.

Voter *j* in group *i* votes for party A if:

$$V^i(P^A) + \sigma^{ji} + \delta > V^i(P^B)$$

$V^i(P^A)$ is the indirect utility for voters in group i based on the policy vector proposed by party A, the term $\sigma^{ji} + \delta$ reflects voter *j*'s ideological preferences for party A. The distribution of political preferences is assumed to differ systematically across groups. δ is the aggregate bias towards A that affects the whole population. This can be considered as a popularity shock. This random variable is assumed to be uniformly distributed as:

$$\delta \sim U\ [-\frac{1}{2\psi}, \frac{1}{2\psi}]$$

Its expected value is zero and the density is ψ, which measures the "variability" of that population, i.e. whether it is likely to take high values. This component represents the source of electoral uncertainty, since it is assumed that δ is realized between the announcement of the party platforms and the election.

In the classic probabilistic voting Nash Equilibrium, the office seeking political candidates are only interested in power *per se*.[25] Competing parties target transfers to marginal - or "swing" – voters, i.e., those who care a lot about social security policy relative to the candidate's other characteristics have more political influence, since a small transfer to this group leads to a greater increase in political support than the same transfer to groups with more extreme ideological attachments. The preferences of all voters are taken into account in a democracy. However, as the swing voters tend to be more responsive to policies: the more a group tends to consist of swing voters, the more their preferences will be considered.

The neutral voters in each group, who are indifferent between party *A* and *B* are referred to as "swing voters". The identity of the swing voters is crucial. We define the swing voters in group *i* as ε^i, represented by the following formula:

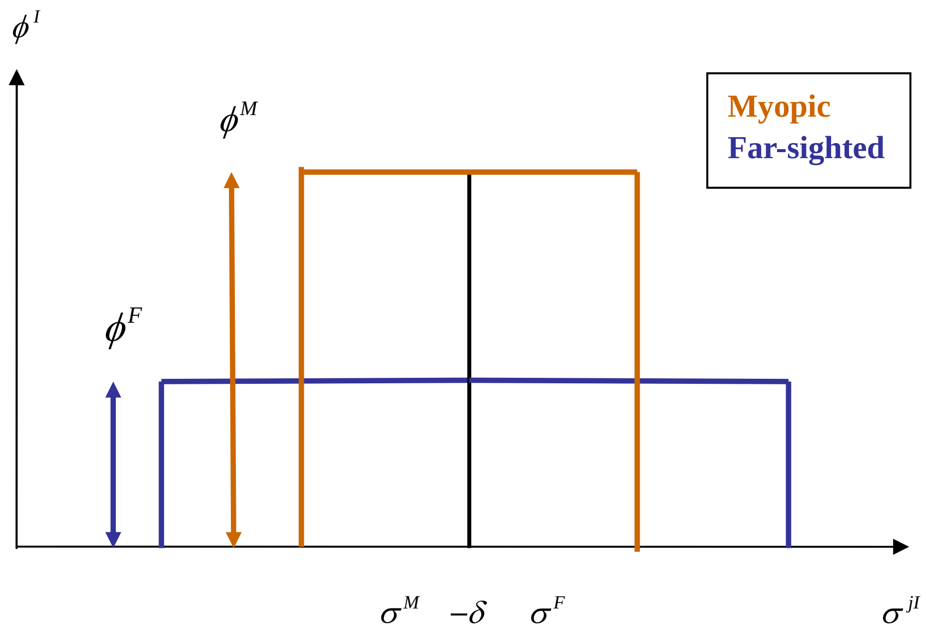
$$\varepsilon^i = V^i(P^B) - V^i(P^A) - \delta \tag{14}$$

Voters with σ^{ji} lower than ε^i vote for B and voters with σ^{ji} higher than ε^i vote for A. The equilibrium swing voters in each group are the individuals with parameter σ^{ji} equals to δ.

Figure 2 provides a graphic illustration. The height of the distribution is measured by the group density ϕ^i, which corresponds to the gain in the amount of votes in group per marginal increase in economic utility. A group with a lower density ϕ^i has a greater ideological dispersion. For illustration purpose, figure 2 assumes that the myopic group has a higher density than the far-sighted group (This does not necessarily hold for the rest of the paper). Suppose party A decides a more favorable policy to the myopes rather than the far-sighted, then party A expects a gain of votes from the myopic group equal to the number of swing voters from this group while a loss of votes equal to the number of swing voters from the far-sighted group. As the myopic group with a higher density ϕ^M is more responsive to the social security policy, party A would obtain a net gain of votes by favoring the myopic voters.

[25] See, e.g., Lindbek and Weibull (1987), Dixit and Londregan (1996), (1998), Persson and Tabellini (2002).

Figure 2: A Graphic Illustration of the Distribution of the Political Ideological Biases



Note: For illustration purpose, the political ideological bias function for the myopic group is modeled with a higher density than for the far-sighted group. However, this assumption does not hold for the rest of the paper.

2.1.3 First stage: Policy Announcements

The political candidates are aware of the size and the policy preferences of both groups. The political ideologies of single voters are unknown, but the political candidates acknowledge the political ideology distribution functions as well as the densities ϕ^i of both groups. Candidates propose their platforms P^A and P^B simultaneously to maximize the number of supporters. The vote share of party A in group *i* is the sum of the swing voters:

$$\pi_i^A = \phi^i[V^i(P^A) + \delta - V^i(P^B)] + \frac{1}{2} \tag{15}$$

Thus, the candidate A's probability of winning becomes,

$$\chi_A = \underset{\delta}{\text{Prob}}\ [\pi^A \geq \frac{1}{2}] = \frac{\psi}{\phi}\left[\sum_{i=F,M} \lambda^i \phi^i [V^i(P^A) - V^i(P^B)]\right] + \frac{1}{2} \tag{16}$$

where $\phi \equiv \sum_i \lambda^i \phi^i$ is the weighted average of ideological heterogeneity across groups.

A representative democracy with probabilistic voting behaves as if the candidates are maximizing a weighted social welfare function, where the weights represent the voter's responsiveness to marginal platform changes and group sizes. In our setting, candidates A and B maximize the expected total number of votes in the farsighted and the myopic groups. The policy vector is $p = (\alpha, \tau)$. Party A will maximize the following objective function:

$$E(\sum_{i=F,M} \lambda^i \pi_i^A) = E[\lambda^F \pi_F^A + \lambda^M \pi_M^A] \tag{17}$$

where π_i^A is given by Eq. (15). We can simplify Eq. (17) into the following form which defines party A's policy choice problem:

$$\max_{\{\tau^A, a^A\}} W^A = E \sum_{i=F,M} \lambda^i \phi^i [V^i(P^A) - V^i(P^B)] \tag{18}$$

The problem is symmetric for political party B. Therefore, the first order conditions for an interior solution for the political party $N \subset \mathbb{R}^2(A,B)$ are as follows:

$$Foc\{\tau\}: \frac{dW^N}{d\tau} = \sum_{i=F,M} \lambda^i \phi^i \frac{dV^i(P^N)}{d\tau} = 0 \tag{19}$$

$$Foc\{\alpha\}: \frac{dW^N}{d\alpha} = \sum_{i=F,M} \lambda^i \phi^i \frac{dV^i(P^N)}{d\alpha} = 0 \tag{20}$$

As shown directly from Eq. (18), as long as both groups get equal weights (the product of size, λ^i, and density, ϕ^i), the political party's task is to simply maximize the average voter's utility. Instead, the group with either a bigger size, λ^i, or a higher value in density, ϕ^i, consists of more swing voters. In other words, a group which is more sensitive to social security will receive a higher weight in the party's objective function.

2.2 Political Equilibrium

First we define a strict local equilibrium of the electoral game. The set of candidates is denoted by $N \subset \mathbb{R}^2(A,B)$, the set of voters by *J*, the compact, convex policy space by $P \subset \mathbb{R}^2$. The policy announced by candidate *N* is P_N.

Definition 1: *Let* $G = (N, J, P, W)$ *be an electoral game. A policy profile* $P^E = (\tau^E, \alpha^E)$ *is a strict local equilibrium of* G *if there exists* $\rho > 0$ *such that for all* N *and for all* P^Δ *that are within* ρ *of* P^E *(i.e.,* $\left\| \tau^\Delta - \tau^E \right\| < \rho$ *or* $\left\| \alpha^\Delta - \alpha^E \right\| < \rho$*),*

$$W^N(P^E) \geq W^N(P^\Delta)$$

with the inequality being strict whenever $P^\Delta \neq P^E$.

In the equilibrium:

1) *Both political parties announce their policy platforms that maximize their expected votes taking as given the opponent's announced social security policy, each group's budget constraints, voting and saving decisions.*
2) *The representative individual in each group votes for the party that maximizes his own well-being taking as given each party's announced platform, popularity shock and his political ideological biases.*

Now we solve the optimization problem for the political candidates. As each party faces a symmetric problem, therefore, in the political equilibrium, the proposal from party A and B will emerge $P^A = P^B = P^E$.

Proposition 2: *There exists a unique local equilibrium in the probabilistic voting setting in which both parties propose the same social security policy* P^E.

Proof: See appendix A.4 and A.5 for derivation.

The general form for the equilibrium social security policy could be derived by solving the maximization problem associated with Eq. (18). We shall be aware of the fact that the two variables are not independent as the generosity and the degree of redistribution are jointly determined. We now proceed to the voting equilibrium. We have already studied the social security preferred by the farsighted and the myopic individuals in section 2.3.1.1. Now we look into the mixed economy where the far-sighted individuals and the myopic agents coexist. The social security policy in the political equilibrium can be obtained by solving the political candidate's maximization problem given by Eq. (18). We summarize the results in the following proposition:

Proposition 3: *In the political equilibrium given by Proposition 2,*

i) the equilibrium social security policy is of the following form:

When $\lambda^F \phi^F u'(d_F) < \lambda^M \phi^M u'(d_M)$

$$(\alpha^*, \tau^*) = \begin{cases} \alpha^* = \dfrac{1}{2} + \dfrac{\lambda^F \phi^F u'(d_F)}{2\lambda^M \phi^M u'(d_M)} & 21(1) \\ \tau^* = \dfrac{[u'(x_M) - u'(d_M)]}{[u'(x_M) - 2u'(d_M)] + \dfrac{u'(d_M)}{4}[1 - \dfrac{\lambda^F \phi^F u'(d_F)}{\lambda^M \phi^M u'(d_M)}]^2} & 21(2) \end{cases}$$

When $\lambda^F \phi^F u'(d_F) \geq \lambda^M \phi^M u'(d_M)$,

$$(\alpha^*, \tau^*) = \begin{cases} \alpha^* = 1 & 22(1) \\ \tau^* = \dfrac{u'(x_M) - u'(d_M)}{u'(x_M) - 2u'(d_M)} & 22(2) \end{cases}$$

Proof: See the Appendix A.4 for derivation.

ii) In the equilibrium, the optimal value of Bismarckian factor is increasing with the relative size, the relative political homogeneity of the farsighted group and the marginal utility ratio of the farsighted group's second period consumption to that of the myopes.

Proof: See Eq. 21(1).

Proposition 1 shows that a fully farsighted economy always prefers a pure contributive (Bismarckian) social security system or no social security whereas a purely myopic society prefers the Bismarckian factor equals to one half. Accordingly, the most preferred Bismarckian factor should be at least one half or higher.[26] Therefore, no Beveridgean system exists in our model.

In the intermediate case, the equilibrium Bismarckian factor α is given by Eq. 21(1). Three effects are in concern. To start with, the relative size and the relative political ideology density of the far-sighted group to the myopic group have positive effects on the degree of redistribution. The intuition is as follows: a big size or a higher density means that the far-sighted group is politically stronger, accordingly. Thus, the political candidates would propose the social security policy that favors the far-sighted group. Lastly, the Bismarckian factor increases with the marginal utility ratio of the farsighted group's second period consumption to that of the myopes. Note that the myopic agents rely on the social security to secure their second period consumption which is dependent solely on the pension benefits. Therefore, when this ratio gets bigger, increasing the Bismarckian factor may lose some votes from the myopic group, however, gaining more votes from the farsighted group.

When $\lambda^F \phi^F u'(d_F) \geq \lambda^M \phi^M u'(d_M)$, meaning that the relative political power of the farsighted group rules over the consumption smoothing concern of the myopic group, the Bismarckian factor in the political equilibrium will become one to please the farsighted voters. However, in this case, given that the farsighted are indifferent as regards to the size of the social security, the politicians will propose the payroll tax rate according to the preference of the myopes.

The analysis of the most preferred payroll tax rate is more complex by seeing directly from the final form in Eq. 22(2). To better understand the different effects in the determination of the equilibrium payroll tax rate, we focus on the intermediate form in which we have as well the equilibrium Bismarckian factor:[27]

[26] Note that this result is closely related with two assumptions: the homogenous productivity and the inseparable preference of consumption and labor supply.

[27] See Eq. A12(1) in appendix A.4.

$$\tau = \frac{\lambda^M \phi^M [u'(x_M) - u'(d_M)]}{\lambda^M \phi^M [u'(x_M) - u'(d_M)] - \underbrace{(1-\alpha)\lambda^F \phi^F u'(d_F)}_{\substack{\text{(The farsighted)} \\ \text{labor supply} \\ \text{distortion}}} - \underbrace{\lambda^M \phi^M u'(d_M)}_{\substack{\text{(The myopic)} \\ \text{labor supply} \\ \text{distortion}}} + \underbrace{(\alpha - \alpha^2)\lambda^M \phi^M u'(d_M)}_{\substack{\text{(The myopic)} \\ \text{redistribution effect}}}} \tag{23}$$

The terms in the **numerator** reflect the consumption smoothing concerns for the myopic group. Given that $d_M < x_M$ for all the myopic agents, the consumption smoothing term calls for a higher tax rate as long as the myopes' consumption gap between the two periods gets larger. This effect corresponds to the social security's role as a commitment device, forcing the myopic individuals saving for their second period. One can identify several effects regarding the denominator. The first term in the denominator serves as the standard consumption smoothing factor inside the myopic group. Therefore, in total, the consumption smoothing concerns for the myopic group is positively related with the size of the pension system.

The second and third terms represent the distortion effects on the labor supply of the farsighted and the myopic group separately. The stronger the distortion effects are, the smaller the payroll tax rate is. The effects are negative for both kinds. The last term in the denominator indicates the redistribution effects from the farsighted group to the myopic group, it is positively related to the payroll tax rate. Note that all the last three terms are weighted by the size, the density as well as the second period marginal utility of the related group. To sum up, the net effects of the presence of myopic agents on the payroll tax rate is ambiguous.

Even though the impact of myopic agents on the equilibrium payroll tax rate remains uncertain, it is worth mentioning that the equilibrium payroll tax rate when the pension system is Bismarckian is the most preferred payroll tax rate for the myopic group. (See Eq. 22(2), this result as well coincides with the desired tax rate by the myopic agents in Cremer et al., 2007). To gain intuition for this result, consider a Bismarckian system (purely contributive), the social security system generates no redistribution, thus, causes no externalities to the farsighted. There is perfect crowding out between private saving and pension for the farsighted. Thus, they are indifferent in the value of the payroll tax. In fact the social security serves merely as a commitment device for the myopic group. Accordingly, to gain the votes from the myopic group, the political candidates will set the payroll tax rate as the most preferred tax rate to the myopic group.

3. Numerical Simulation: The logarithmic utility case

We denote $k = \theta\beta$ ($\theta = \lambda^F / \lambda^M$, $\beta = \phi^F / \phi^M$) to better explain the results. θ represents the size ratio of the farsighted group to the myopic group whereas β denotes the political ideology density ratio of the farsighted group to that of the myopic group.

The payroll tax rate and the Bismarckian factor in the equilibrium are all implicit solutions, which are dependent on the concrete form of the utility function. Therefore, we conduct a numerical simulation with the following logarithmic utility function.

$$U = In(c - l^2/2) + In(d)$$

The results are displayed in the columns of Table 1 with respect to different value of β.

Table 1: The Simulation Results

The equilibrium (α^*, τ^*) w.r.t. (θ β)

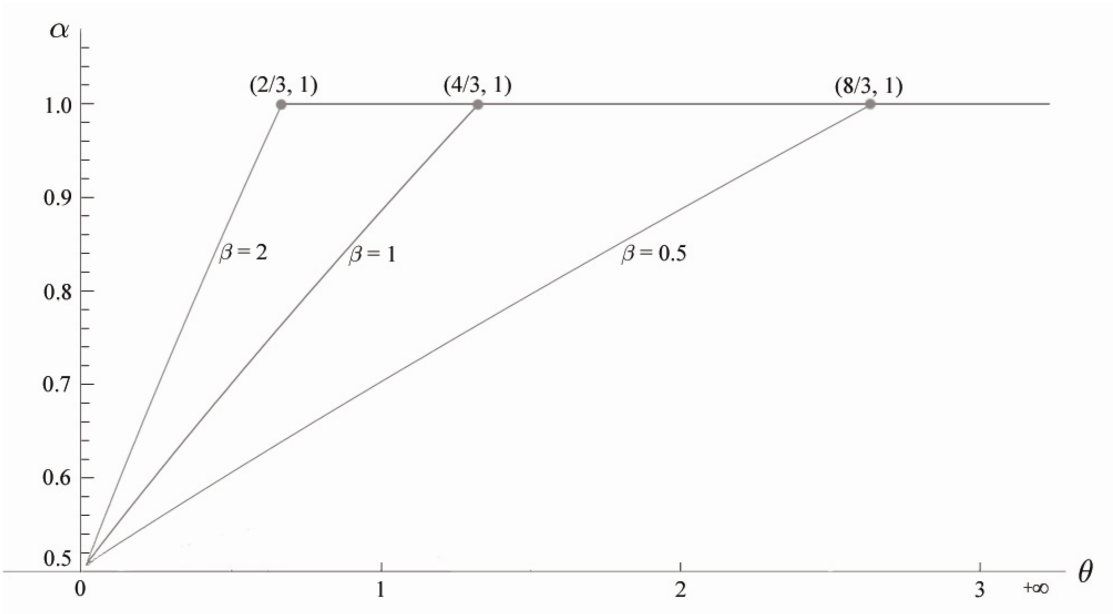
θ (λ^F)	(α, τ) with β=1	(α, τ) with β=0.5	(α, τ) with β=2
0.11 (0.10)	(0.545, 0.258)	(0.523, 0.259)	(0.589, 0.255)
0.25 (0.20)	(0.601, 0.255)	(0.551, 0.257)	(0.699, 0.251)
2/3 (0.4)			(1.000, 1/4)
1.00 (0.50)	(0.883, 0.249)	(0.699, 0.251)	——
4/3 (4/7)	(1.000, 1/4)		——
2.33 (0.70)	——	(0.941, 0.249)	——
8/3 (0.75)	——	(1.000, 1/4)	——

Note: θ is the size ratio of the farsighted group to the myopic group. When $\theta = 1$, the two groups are equally distributed. $\theta < 1$ indicates more myopic agents are in the economy whereas $\theta > 1$ means that the proportion of the farsighted group overweigh that of the myopic group. β denotes the density ratio representing the political homogeneity of the farsighted group to the myopic group. When $\beta = 1$, the farsighted and the myopic group are equally politically homogenous. $\beta < 1$ means the myopic group is politically homogenous whereas $\beta > 1$ means the farsighted group is more politically homogenous.

In particular, we study three cases. We start with the standard case with $\beta = 1$, in which the heterogeneity depends only on the size of each group but not the political homogeneity. Secondly, we assume that the myopic group is more politically homogenous, namely $\beta = \frac{1}{2}$. Lastly, the opposite case in which the farsighted individuals are more politically homogenous is taken into account: $\beta = 2$. The simulation is conducted with respect to different values of θ. The range we choose is [0.11, 0.25, 2/3, 1.00, 4/3, 2.33, 8/3], which corresponds to the proportion of the far-sighted group λ^F [0.1, 0.2, 0.4, 0.5, 4/7, 0.7, 0.75].

To understand the relation between the equilibrium social security policy and the size of myopic agents under different β, we first start showing the relation of the equilibrium Bismarkcian factor and the size ratio of the farsighted group to the myopic group, θ, with respect to different political ideological density ratio between two groups, β, in Figure 3.

Figure 3: The Relation of the Equilibrium Bismarckian Factor α^* and θ



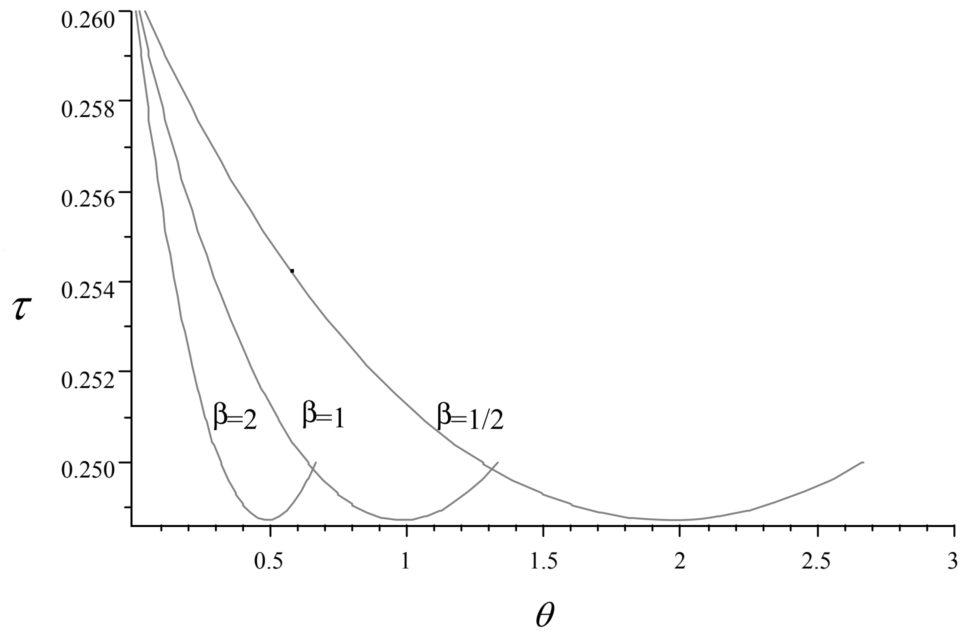
Remark: Figure 3 illustrates how the Bismarckian factor α is determined by the size ratio (θ) and the political ideology density ratio (β) of the far-sighted group to the myopic group.

The following characteristics of the Bismarckian factor are displayed: when the myopic and far-sighted group are equally politically homogenous, namely $\beta = 1$, the most preferred Bismarckian factor is equal to 1 as long as θ is higher than 2/3, in which the proportion of the rational individuals is larger than 4/7 in the society. This value is higher than 1/2 from the majority voting. The reason is that in our probabilistic voting setting, not only the size, the political homogeneity of the farsighted and the myopic groups matter, but also the sensitivity of how each group to the change in social security policy is crucial. Therefore, given that the myopic group is more sensitive to the change in the Bismarckian factor, a higher size of the farsighted group is needed to win the political game than the majority voting setting.

Moreover, the more politically homogenous the far-sighted group, for instance, β=2, the less votes are required from the far-sighted group to reach the pure Bismarckian system. When β=2, this value drops to θ=2/3. On the contrary, if the myopic agents are more politically homogenous, for instance, β=0.5, a larger proportion of the far-sighted individuals is required for a pure Bismarckian pension system. While β=0.5, this value raises up to θ=8/3.

The correlation of the equilibrium payroll tax rate and θ under different values of β is shown in the following Figure 4. As mentioned before, the total effect of the myopic agents to the equilibrium payroll tax rate is ambiguous. The U-Shaped curve depicted in Figure 4 could be explained as follows: in the left side of the graph, the payroll tax rate decreases with the size of the farsighted group, this is rather intuitive. Consider the relationship when β=1, if $0<\theta<1$ the myopic group has a bigger size than the far-sighted group. Given that the myopic group is in need of social security to secure consumption after retirement, not surprisingly, the payroll tax rate level is quite high. When θ increases indicating the increase in the size of the farsighted group, the payroll tax rate decreases as the farsighted group wants to reduce the labor distortion effect and the redistribution effect associated with the payroll tax rate. It is worth noting that the payroll tax rate falls to its lowest value until the two groups become equally politically powerful, namely θ=1. The interesting point is that when θ>1, indicating that the far-sighted group politically rules over the myopic group, the payroll tax rate starts to rise again.

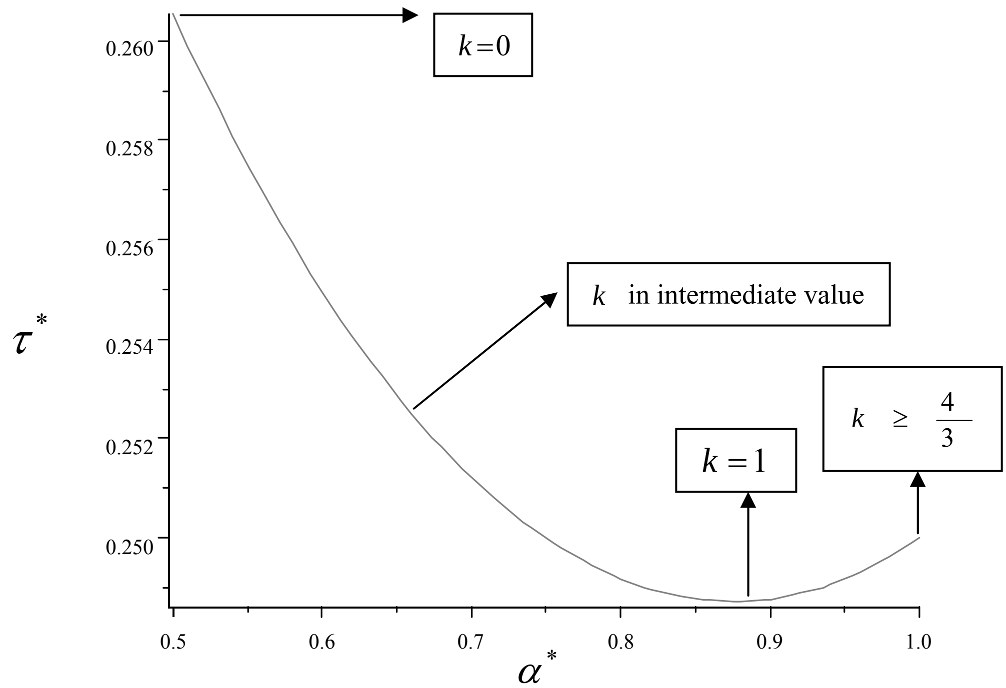
Figure 4: The Relation of the Equilibrium Tax Tate τ^* and θ



Remark: Figure 4 illustrates how the payroll tax rate τ is influenced by the size ratio ($\theta = \frac{\lambda^F}{\lambda^M}$) and the political ideology density ratio ($\beta = \frac{\phi^F}{\phi^M}$) of the far-sighted group to the myopic group.

Eq. (23) shows that the political parties care for the following effects: the distortion effects for both groups, the consumption smoothing concern for the myopic group, the redistribution effect for the myopic group. Given that these effects are related with the Bismarckian factor, we explain this point in Figure 5, in which it shows that the equilibrium payroll tax rate is U-shaped with respect to the equilibrium Bismarckian factor.

Figure 5: The Relation of the Equilibrium Tax Tate τ^* and the Equilibrium Bismarckian Factor α^*



Note: The equilibrium payroll tax rate and the Bismarckian factor depends on the value of $k = \theta\beta$, the product of the size ratio and the political ideology density ratio (β) of the far-sighted group to the myopic group.

Each point on the curve responds to a specific value of k where $k = \theta\beta$. When k is zero the Bismarckian factor is one half. In the left side of the curve, the equilibrium payroll tax rate decreases with the value of k. k equals to 1 when the curve reaches its minimum. In the right side of the curve, the equilibrium payroll tax rate increases with the value of k which corresponds to the findings in Figure 4 as well. Finally the curve stops at $(1, \frac{1}{4})$ where the farsighted group reaches its preferred Bismarckian factor. At the same time, the payroll tax rate in the equilibrium is as well the equilibrium payroll tax rate of the myopic group.[28] The equilibrium payroll tax rate in the left hand part of this U-shaped curve coincides with the left hand side of the curve in Figure 4. Obviously, in the left side part, the myopic group is more politically powerful, however, when the curve moves to the right, the myopic

[28] See the second point of the Proposition 1.

group becomes less political stronger, therefore, the political parties rise α and reduce τ to favor the far-sighted voters. This trend keeps until the payroll tax rate reaches its minimum. We can see from Figure 5, the curve reaches the bottom when the two groups are equally politically powerful. Interestingly, the payroll tax rate starts to increase when the far-sighted are more political powerful. As the myopic agents' second period consumption is less than that of the far-sighted, thus, it is easier for the politicians to turn them by rising their utility in the second period than that of the far-sighted agents. Hence, what matters now is $\frac{u'(d_F)}{u'(d_M)}$. Therefore, the right hand side of the curve could be explained that the political parties strategically increases the Bismarckian factor to keep favoring the far-sighted agents but also rises the payroll tax rate to attract the swing voters in myopic group.

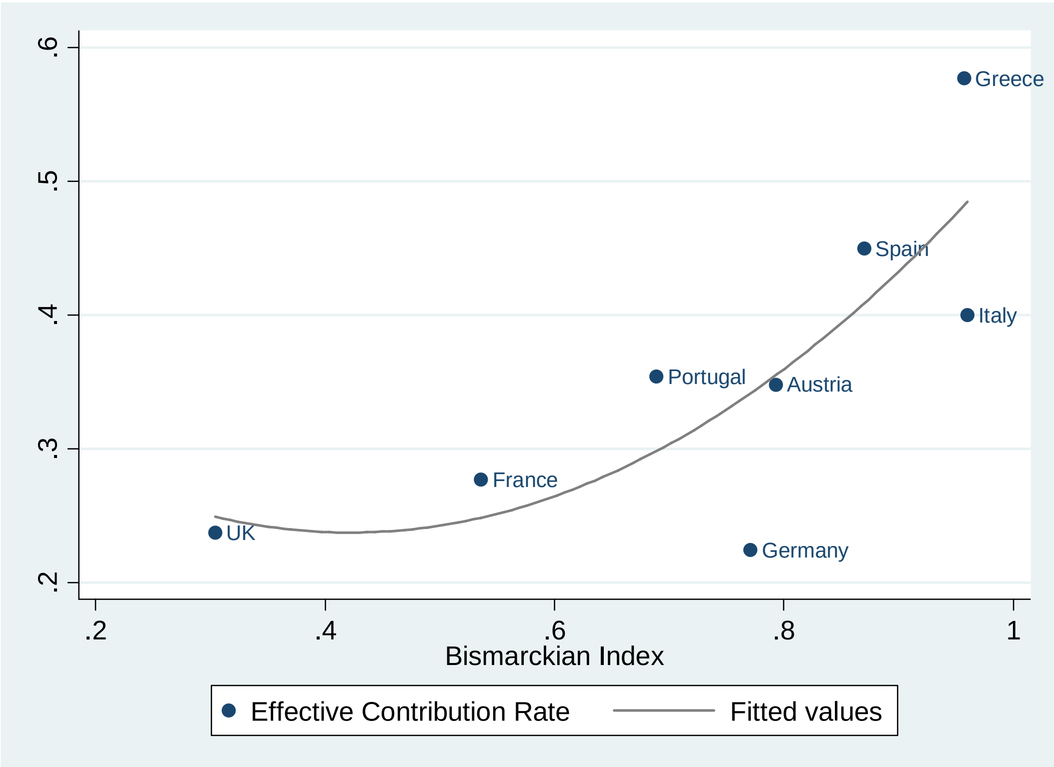
This section presented a numerical simulation of the model under the assumption of a logarithmic period utility function (namely, $u(x) = \ell n(x)$). The findings from the numerical simulations can be summarized as follows. First, the equilibrium payroll tax rate is U-shaped with respect to the equilibrium Bismarckian factor in the economy (see Figure 5). Second, the equilibrium Bismarckian factor increases with the relative size of the farsighted group in the economy (see Figure 3). The shape of the curve depends on the value of β, namely the political ideology density ratio of the far-sighted to the myopic group. The higher β is, namely the more politically homogeneous the far-sighted group becomes, the steeper the curve becomes; this implies that the far-sighted are able to obtain their preferred Bismarckian factor of unity when they have a smaller relative size (namely, when the value of θ is smaller). Finally, the optimal payroll tax rate is U-shaped with respect to the relative size of the farsighted group (see Figure 4). The shape of the curve depends on the value of β. The higher β is, the lower is the value of θ associated with a minimum of the curve and the steeper the curve becomes.

4. Link with the Stylized Facts

In this section, we relate the results from our model to some stylized facts in the characteristics of the existing pension systems. The U-shape feature between the payroll tax rate and the Bismarckian factor provided in Figure 5 may partially explain the stylized facts in Figure 6 where the payroll tax rate and the Bismarckian factor are positively related. Note that the purpose of our simulation is not to draw quantitative predictions, but instead, to provide qualitative comparison.

Figure 6 shows graphically the characteristics of social security in some OECD member countries.[29]

Figure 6: Characteristics of the Pension Systems in OECD Countries[30]



Source: The Bismarckian Index α taken from OECD (2005), The Effective Contribution Rate

---

[29] $\alpha \in [0,1]$ (For $\alpha > \frac{1}{2}$, we have Bismarckian pension system and Beveridgean system for $\alpha < \frac{1}{2}$, see i.e. Conde-Ruiz and Profeta (2007))

[30] To produce better fitted graph with the data in Table 2.B.1, we use the stata command "qfit" to illustrate two-way quadratic prediction plots.

τ taken from Disney (2004)[31]. See Table B.1 in the appendix for the detailed data.

In particular we consider the following countries: Austria, France, Germany, Greece, Portugal, Italy, Spain and UK. It is shown that when the contribution rate is sufficiently large, the Bismarckian social security system is associated with larger public pension expenditures. These facts are in contrast to the political economy theories of social security which predict that Beveridgean systems, involving intragenerational redistribution, should enjoy larger support than Bismarckian ones.[32]

Previous political economy literature[33] solves this puzzle in different frameworks rather than the probabilistic voting model. Based on European data, Koethenbuerger et al. (2008) argue that the median voter is typically middle-aged with high income and prefers a less generous system with a higher redistribution degree. Conde-Ruiz and Profeta (2007) apply the notion of issue-by-issue voting to address this puzzle focusing on income inequality, conclude a Beveridgean system may be supported by the low income voters in favor of the redistribution components while the high income supporters vote for the reduced size of the pension system. Krieger and Traub (2008) provide empirical and experimental explanations for this puzzle relying on the effects from income inequality and life expectancy. The current paper tries to explain this puzzle in an economy with both farsighted and myopic agents, where social security functions as a redistribution and commitment device for the myopic group while causes distortions to the farsighted group. The social security in the political equilibrium is to compromise the above effects.

5. Conclusion

The present paper extends the literature on the optimal design of social security systems taking into account the presence of myopic agents. Cremer et al. (2007) study

---

[31] The average effective contribution rate is derived as the average replacement rate divided by the support ratio. See Box 1 in Disney (2004) for a detailed explanation of the calculation methods.

[32] See Conde-Ruiz and Profeta (2002) for a detailed discussion.

[33] See Casamatta et al. (2000a), Cremer and Pestieau (1998) and Pestieau (1999), Cremer et al. (2007).

the optimal social security policy under majority voting when agents differ both in their productivity and in their myopic degree. People vote sequentially, first for the Bismarckian factor, then for the payroll tax rate. The Bismarckian factor takes only two extreme values: either 0 or 1. Given that the poor (and liquidity constrained) farsighted agents prefer the same policy as their myopic counterparts, their model reaches the "ends against the middle" type of equilibria where where low and high ability voters oppose the ones with intermediate ability (vote on the payroll tax for a given Bismarckian factor). Cremer et al. (2007) focus on the influence of myopic agents to the social security system but neglect the relationship of the generosity of the system and the redistribution degree taking into account in the presence of myopic agents. Instead, the present paper adopts the probabilistic voting framework, in which the optimal payroll tax rate and the Bismarckian factor are determined simultaneously in the voting equilibrium. Moreover, a more general solution is adopted where the Bismarckian factor is continuous between 0 and 1 is allowed.

The present paper is also related to the literature applying probabilistic voting in the design of social security policy. In an OLG model, Profeta (2002) focuses on the intergenerational political support from the old group's influence to the political outcome in social security policies. She emphasizes on the "single-mindedness" of the old group in the support of intergenerational transfers. The main conclusion of her work is that retirement increases the level of ideological homogeneity of the old generation as they have more leisure time. This increase in their political power allows them to win the political game and to receive a positive transfer from the young. More recently, in a probabilistic voting framework, Gonzalez-Eiras and Niepelt (2008) focus on the dynamical politico-economic equilibrium. Political competition resolves the conflict between old and young voters by shifting some of the cost of the social security system to future generations. As a consequence, intergenerational transfers are too large, relative to a system balancing the interests of all generations. Unlike the works mentioned above focusing on redistribution related to leisure, intergenerational game or income inequality, the present paper focuses specifically on the redistribution resulting from myopia.

To be more specific, we explore the social security scheme consisting of a Bismarckian factor and a payroll tax rate in a society with both myopic and farsighted agents in which the political equilibrium is determined through probabilistic voting. The individuals live for two periods: work by choosing their labor supply in the first and retire in the second period. The individuals differ in being myopic or farsighted and in their political ideologies. The farsighted agents naturally smooth their consumption between the two periods, whereas the myopic agents behave myopically in the market stage by simply consuming all of their disposable income, regretting so in the second period. The social security causes labor supply distortions to both groups. The farsighted are less distorted as they see the link between the pension payment and contribution component in the pension benefits. In the voting stage, which occurs before the market stage, young agents vote on the social security scheme offered by two competing parties. Both the farsighted agents and the myopic agents are assumed to vote rationally for the social security scheme according to their long-term preferences. This is because the myopic agents are assumed to be sophisticated. This means that they anticipate in the voting stage that their myopic behavior will emerge later in the market stage. Thus, the myopic agents use the pension system as a way to commit themselves into forced savings. In order to focus on the steady-state political equilibrium, the paper assumes that the interest rate and the population growth rate are equal and set them equal to zero.

Unlike Cremer et al. (2007), who treat the size of the myopic and farsighted groups as common knowledge, the present paper introduces uncertainty to these variables. When an individual makes decision in voting, he needs information about the size of both groups estimate average social income, thus, to understand the pension benefits he may receive when retired.

It turns out that the optimal Bismarckian factor in the political equilibrium has the following features. With only rational individuals voting, either a purely Bismarckian system or simply no social security is preferred. The myopic group's preferred policy is instead to select an intermediate Bismarckian system, with a value of $\alpha = 0.5$ (the pure Beveridgean and Bismarckian systems being characterized by α=0 and α=1, respectively), due to the tradeoff between efficiency and redistribution concerns. The case of a mixed society is rather complicated. A pure Bismarckian system is preferred as long as the relative political power (the product of the size ratio and the political homogeneity ratio) of the farsighted group to the myopic group rules over the redistribution concern for the myopic group. However, a value of the Bismarckian parameter α between the values corresponding to the optimal systems for each of the two groups is a possibility, depending on their relative political power.

As regards the equilibrium payroll tax rate in the mixed society, several effects determine the equilibrium payroll tax rate: the consumption smoothing concern for the myopic group, the labor distortion effects for both groups, the redistribution effects from the farsighted to the myopic group. Note that all the above effects are weighted with the size, the density as well as the second period marginal utility of the relevant group. Thus, the net effect of the proportion of the myopic individuals on the equilibrium payroll tax rate is ambiguous. However, an interesting result emerges when the social security system is purely Bismarckian (see Proposition 2). In this case, the equilibrium payroll tax rate is the one preferred by the myopic group. As a pure contributive pension system generates no redistribution, the farsighted individuals are indifferent in the value of the payroll tax rate. Thus, the payroll tax serves purely as a forced saving method for the myopic individuals.

The results in the numerical simulation illustrate our theoretical findings. When a group is more politically homogenous, a lower size is required for this group to achieve its desired social security scheme. Moreover, the optimal payroll tax rate is U-shaped with respect to the Bismarckian factor which may partially explain why, in practice, Bismarckian pension systems tend to be associated with a larger pension base compared to the Beveridgean systems. A possible explanation of these features is to balance the commitment concerns of myopic agents and the efficiency concerns of the farsighted agents.

A direction for future research would be to modify or relax some of the assumptions of the model, to see whether or not the main results are sensitive to some specific assumptions being made. First, following Cremer et al. (2007), the assumption of homogeneous productivity could be relaxed. One simple way would be to introduce low productive consumers who face borrowing constraints into the model. This would allow us to introduce redistribution concerns from high productivity agents to low productivity agents into the design of the social security system. Second, the paper has assumed a unitary inter-temporal elasticity of substitution, with a log linear specification of utility in the simulation section. Given that many empirical studies suggest that the elasticity of inter-temporal substitution is less than unity, an obvious step would be to incorporate this assumption into the model. However, we know from standard OLG models that when one deviates from log linear utilities, this opens up the possibility of multiple steady-state market equilibria in the model. Thus, it remains unclear a priori to what extent our results would be affected by the deviation from the assumption of a unitary inter-temporal elasticity. Finally, the paper has assumed a form of quasi linear preferences in first period consumption (with a quadratic disutility of labor), which implies that there is no income effect on labor supply following a change in the wage rate. In turn, this means that the elasticity of labor supply with respect to the tax rate can become quite large when the system is not purely Bismarckian (in which case the Marshallian elasticity is zero for the farsighted agents) which may not be consistent with the existing empirical evidence of the tax effects on labor supply, especially for males. This for sure restricting the analysis and should be take into account in future research by the application of different utility functions.

Appendix

A.1: the optimal labor supply for each group

For the far-sighted individuals, The first order conditions to the maximization problem:

$$Foc\{l_F\}: \frac{\partial U_F}{\partial l_F} = u'(x_F)[(1-\tau)-l_F] + u'(d_F)\tau\alpha = 0$$
$$Foc\{s_F\}: \frac{\partial U_F}{\partial s_F} = u'(x_F)(-1) + u'(d_F) = 0$$

An interior solution (with s_i>0) requires the life time utility of the far-sighted differentiate with respect to saving is equal to zero. Thus, implies $x_F = d_F$. Thus, the optimal level of labor supply for the far-sighted is:

$$l_F = (1-\tau) + \tau\alpha \tag{A1}$$

As for the myopic individuals, the labor supply and saving pattern:

$$Foc\{l_M\}: \frac{\partial U_M}{\partial l_M} = u'(x_M)[(1-\tau)-l_M] = 0$$

As assumed, the myopic individuals do not save, thus, $s_M = 0$, and the optimal level of labor supply is:

$$l_M = 1-\tau \tag{A2}$$

A.2: The indirect utility function for each group

The budget constraints for both groups are:

$$x_i + s_i = (1-\tau)l_i$$
$$d_i = s_i + p_i$$

Now we need to derive the pension for each group separately,

$$p_i = \tau[\alpha l_i + (1-\alpha)El]$$
$$El = l_F E\lambda_F + l_M E(1-\lambda_F)$$

Thus, substitute l_F and l_M from A1 and A2 into Eq. (7), we obtain the pension for the far-sighted and myopic individuals:

$$\begin{aligned} Ewl &= l_F E\lambda_F + l_M E\lambda_M = (1-\tau) + \frac{1}{2}\tau\alpha \\ p_F &= \tau[\alpha l_F + (1-\alpha)El] = \tau(1-\tau) + \frac{\tau^2\alpha(1+\alpha)}{2} \\ p_M &= \tau[\alpha l_M + (1-\alpha)El] = \tau(1-\tau) + \frac{\tau^2\alpha(1-\alpha)}{2} \end{aligned} \tag{A3}$$

Thus, the indirect utility function for the far-sighted could be obtained in following steps; Firstly, from the budget constraints, the saving function for the far-sighted is:

$$s_F = \frac{l_F(1-\tau) - \frac{l_F^2}{2} - p_F}{2} \tag{A4}$$

Substitute (A2) and (A3) to (A4), the savings of far-sighted could be interpreted as a function of (τ, α):

$$s_F = \frac{1}{2} - 2\tau + \frac{3}{2}\tau^2 - \tau^2\alpha^2 - \frac{1}{2}\tau^2\alpha \tag{A5}$$

Remember, the far-sighted agents smooth their consumption over time, thus,

$$x_F = d_F = s_F + p_F = \frac{l_F(1-\tau) - \frac{l_F^2}{2} + p_F}{2}$$

Substitute p_F from Eq. (A3) into Eq. (A5),

$$x_F = d_F = \frac{1-(1-\alpha)\tau^2}{4}$$

Therefore, the indirect utility function for the far-sighted is:

$$V^F(\tau,\alpha) = u(x_F) + u(d_F) = 2u[\frac{1-(1-\alpha)\tau^2}{4}] \tag{A6}$$

As for the myopic individuals:

$$x_M = (1-\tau)l_M$$
$$d_M = p_M$$

Substitute l_M from (A2) and p_M from (A3), the indirect utility function for the myopic agents is obtained in following form:

$$V^M(\tau,\alpha) = u[\frac{(1-\tau)^2}{2}] + u[\tau(1-\tau) + \frac{\tau^2\alpha(1-\alpha)}{2}] \quad \text{(A7)}$$

A.3: Group-specific preference (Proof for Proposition 1)

The optimal social security for the far-sighted household:

The first order condition to maximize V^F w.r.t α,τ given $\alpha,\tau \in [0,1]$ is obtained accordingly,

$$\begin{cases} \dfrac{dV^F}{d\tau} = u'(x_F)\tau(\alpha-1) \\ \dfrac{dV^F}{d\alpha} = u'(x_F)\dfrac{\tau^2}{2} \end{cases} \quad \text{(A8)}$$

The critical point is $(\tau,\ \alpha) = (0,\ 1)$. According to the second derivatives test,

$$\begin{cases} \dfrac{\partial^2 V^F}{\partial\tau^2} = u''(x_F)(\alpha-1)^2\dfrac{\tau^2}{2} + u'(x_F)(\alpha-1) \\ \dfrac{\partial^2 V^F}{\partial\alpha^2} = u''(x_F)\dfrac{\tau^4}{8} \\ \dfrac{\partial^2 V^M}{\partial\alpha\,\partial\tau} = u''(x_F)\dfrac{\tau^3}{4}(\alpha-1) + u'(x_F)\tau \end{cases} \quad \text{(A9)}$$

Obviously $\dfrac{\partial^2 V^F}{\partial\tau^2}$ and $\dfrac{\partial^2 V^F}{\partial\alpha^2} < 0$ in the critical point, however, the Hessian Matrix is $H_F = 0$ implying negative semidefinite, therefore, we need to characterize the maximization problem in a more general form:

As $\dfrac{dV^F}{d\tau} \le 0$, the critical value for τ to maximize V_F is $\tau = 0$, in this case, obviously we have $\alpha_F^* \in [0,\ 1],\ \tau_F^* = 0$.

As $\dfrac{dV^F}{d\alpha} \ge 0$, the critical value for α to maximize V_F is $\alpha = 1$, although there is no specific requirement for τ, s_F needs to be non-negative. Therefore, we have $\alpha_F^* = 1,\ \tau_F^* \in [0,\ 1]$.

In sum, the optimal social security for far-sighted agents are two solution sets:

$$\begin{cases} \alpha_F^* = 1,\ \tau_F^* \in [0,\ 1] \\ No\ social\ security \end{cases}$$

The optimal social security for the myopic household:

The first order condition to maximize V^M w.r.t α, τ given $\alpha, \tau \in [0,1]$ is obtained accordingly,

$$\begin{cases} \dfrac{dV^M}{d\tau} = u'(x_M)(\tau - 1) + u'(d_M)[1 - 2\tau + \tau\alpha(1-\alpha)] \\ \dfrac{dV^M}{d\alpha} = u'(d_M)\tau^2(\dfrac{1}{2} - \alpha) \end{cases} \tag{A8}$$

The critical point we obtained is:

$$\alpha_M^* = \frac{1}{2},\ \tau_M^* = \frac{u'(x_M) - u'(d_M)}{u'(x_M) - \dfrac{7}{4}u'(d_M)}$$

Noticing that, $x_M > d_M$, thus, $u'(x_M) < u'(d_M) < \frac{7}{4}u'(d_M)$, further more, $[u'(x_M) - \frac{7}{4}u'(d_M)] > u'(x_M) - u'(d_M)$, therefore, $0 < \tau^* < 1$. Also, the more concave the utility function, the lower the value τ^* is.

Now we take the second partial derivative test:

$$\begin{cases} \dfrac{\partial^2 V^M}{\partial\tau^2} = u''(x_M)(\tau-1)^2 + u'(x_M) + u''(d_M)[1-2\tau+\tau\alpha(1-\alpha)]^2 + u'(d_M)(\alpha-\alpha^2-2) \\ \dfrac{\partial^2 V^M}{\partial\alpha^2} = u''(d_M)\tau^4(\dfrac{1}{2}-\alpha)^2 - u'(d_M)\tau^2 \\ \dfrac{\partial^2 V^M}{\partial\alpha\,\partial\tau} = u''(d_M)\tau^3(\dfrac{1}{2}-\alpha)[(1-2\tau+\tau\alpha(1-\alpha)] + u'(d_M)(1-2\alpha)\tau \end{cases} \tag{A9}$$

Obviously, $\dfrac{\partial^2 V^M}{\partial\tau^2}$ and $\dfrac{\partial^2 V^M}{\partial\alpha^2} < 0$, the determinant of the Hessian Matrix $H_M > 0$ for (α_M^*, τ_M^*). Therefore, (α_M^*, τ_M^*) is the optimal local maximum for V_M. Moreover,

$$\frac{\partial\tau_M^*}{\partial\alpha}=\frac{u''(d_M)(\alpha-\frac{1}{2})(1-\frac{7}{4}\tau)\tau^2}{\underbrace{u'(x_M)+(1-\tau)^2u''(x_M)-\frac{7}{4}u'(d_M)+(1-\frac{7}{4}\tau)u''(d_M)[1-2\tau+\alpha(1-\alpha)]}_{negative}}$$

As $u'>0,\ u''<0$, we have: $\begin{cases}\frac{\partial\tau_M^*}{\partial\alpha}>0 \quad when \quad \alpha>\frac{1}{2}\\ \frac{\partial\tau_M^*}{\partial\alpha}<0 \quad when \quad \alpha<\frac{1}{2}\end{cases}$

A.4: The policy in the political equilibrium in a mixed economy

In order to obtain the most preferred payroll tax rates τ and the level of α, we have to substitute $V^M(P^N)$ and $V^F(P^N)$ which are defined respectively in Eq. (A4) and (A5). Parties act simultaneously, taking the choice of the other party candidate as given, and do not cooperate. Therefore, for instance, politician A would maximize a weighted welfare function obtained from Eq. (18). Given the problem is symmetric for politician B, we write the optimization problem for politician *N* as follows:

$$W^N=E\{\sum_{i=F,M}\lambda^i\phi^i[V^i(P^N)-V^i(P^{-N})]\} \tag{A10}$$

Thus, the FOCs are obtained accordingly:

$$\begin{aligned}&Foc\{\tau\}:\frac{dW^N}{d\tau}=\sum_{i=F,M}\lambda^i\phi^i\frac{dV^i(P^N)}{d\tau}=0\\&Foc\{\alpha\}:\frac{dW^N}{d\alpha}=\sum_{i=F,M}\lambda^i\phi^i\frac{dV^i(P^N)}{d\alpha}=0\end{aligned} \tag{A11}$$

Substitute Eq. (A7) and (A9) into first order Eq. (A11), we obtain the general solution for the payroll tax rates and the Bismarckian factor. Setting the first order conditions equal to zero yields a system of two equations that jointly determine the most preferred value for τ and α.

$$\tau=\frac{\lambda^M\phi^M[u'(x_M)-u'(d_M)]}{\lambda^M\phi^M[u'(x_M)-u'(d_M)]-(1-\alpha)\lambda^F\phi^Fu'(d_F)-\lambda^M\phi^Mu'(d_M)+\alpha(1-\alpha)\lambda^M\phi^Mu'(d_M)} \quad \text{A12(1)}$$

$$\alpha^* = \frac{1}{2} + \frac{\lambda^F \phi^F u'(d_F)}{2\lambda^M \phi^M u'(d_M)} \qquad \text{A12(2)}$$

Note, $\alpha \in (0,1)$, therefore, the Bismarckian factor α in the local equilibrium is derived accordingly,

$$\alpha^* = \begin{cases} \dfrac{1}{2} + \dfrac{\lambda^F \phi^F u'(d_F)}{2\lambda^M \phi^M u'(d_M)} & \text{for } \lambda^F \phi^F u'(d_F) < \lambda^M \phi^M u'(d_M) \qquad \text{A13(1)} \\ 1 & \text{for } \lambda^F \phi^F u'(d_F) \geq \lambda^M \phi^M u'(d_M) \qquad \text{A13(2)} \end{cases}$$

When $\lambda^F \phi^F u'(d_F) < \lambda^M \phi^M u'(d_M)$, $\alpha^* = \frac{1}{2} + \frac{\lambda^F \phi^F u'(d_F)}{2\lambda^M \phi^M u'(d_M)}$, we obtain a more compact form of the equilibrium tax rate τ by substituting the equilibrium Bismarckian factor α from Eq. A13(1) into the Eq. (A11).

$$\tau^* = \frac{[u'(x_M) - u'(d_M)]}{[u'(x_M) - 2u'(d_M)] + \frac{u'(d_M)}{4}[1 - \frac{\lambda^F \phi^F u'(d_F)}{\lambda^M \phi^M u'(d_M)}]^2} \qquad \text{(A14)}$$

See from A13(2), when $\lambda^F \phi^F u'(d_F) \geq \lambda^M \phi^M u'(d_M)$, we can not derive the equilibrium social security policy from the Eq. (A12) directly. Thus, it is considered as a special case. Therefore, we have to substitute the equilibrium Bismarckian factor in which $\alpha^* = 1$ directly to the Lagrange problem A(10) so as to derive the equilibrium payroll tax rate, and the expression of the equilibrium payroll tax rate is as follows:

$$\tau = \frac{u'(x_M) - u'(d_M)}{u'(x_M) - 2u'(d_M)}$$

To sum up, in the mixed economy, the equilibrium social security policy is in the following form:

When $\lambda^F \phi^F u'(d_F) < \lambda^M \phi^M u'(d_M)$,

$$(\alpha^*, \tau^*) = \begin{cases} \alpha^* = \dfrac{1}{2} + \dfrac{\lambda^F \phi^F u'(d_F)}{2\lambda^M \phi^M u'(d_M)} \\ \tau^* = \dfrac{[u'(x_M) - u'(d_M)]}{[u'(x_M) - 2u'(d_M)] + \frac{u'(d_M)}{4}[1 - \frac{\lambda^F \phi^F u'(d_F)}{\lambda^M \phi^M u'(d_M)}]^2} \end{cases} \qquad \text{A(15)}$$

When $\lambda^F \phi^F u'(d_F) \geq \lambda^M \phi^M u'(d_M)$,

$$(\alpha^*, \tau^*) = \begin{cases} \alpha^* = 1 \\ \tau^* = \dfrac{u'(x_M) - u'(d_M)}{u'(x_M) - 2u'(d_M)} \end{cases} \qquad \text{A(16)}$$

A.5: Proof for Proposition 2

To show the second order condition, we first conduct the Hessian Matrix for the optimization problem associated with Eq.(18):

$$H(\alpha^*, \tau^*) = \begin{bmatrix} \dfrac{\partial^2 W^N}{\partial \tau^2} & \dfrac{\partial^2 W^N}{\partial \alpha\, \partial \tau} \\ \dfrac{\partial^2 W^N}{\partial \tau\, \partial \alpha} & \dfrac{\partial^2 W^N}{\partial \alpha^2} \end{bmatrix}$$

Therefore, there is a unique local equilibrium as long as the Hessian Matrix is negative definite with respect to the equilibrium solution (τ^*, α^*). In our case when $\lambda^F \phi^F u'(x_F) < \lambda^M \phi^M u'(d_M)$, this condition is satisfied. When $\lambda^F \phi^F u'(x_F) \geq \lambda^M \phi^M u'(d_M)$, the equilibrium social security (τ^*, α^*) is $(\dfrac{u'(x_M) - u'(d_M)}{u'(x_M) - 2u'(d_M)}$, 1). The second order conditions are satisfied as well.

Table B.1: Characteristics of the pension systems

Country	Bismarckian Index	Effective Contribution Rate
Austria	**0.793**	**34.8**
France	**0.536**	**27.7**
Germany	**0.771**	**22.4**
Greece	**0.957**	**57.7**
Portugal	**0.689**	**35.4**
Italy	**0.96**	**40**
Spain	**0.87**	**45**
UK	**0.304**	**23.7**

Source: The Bismarckian Index α is taken from OECD(2005), The Effective Contribution Rate τ is taken from Disney (2004).

References

Andersen, T. M. and J. Bhattacharya (2010). "On Myopia as Rationale for Social Security," *Economic Theory*, 47(1), 135-158.

Burden, B. (1997). “Deterministic and Probabilistic Voting Models,” *American Journal of Political Science*, 41, 1150–1169.

Casamatta, G., H. Cremer and P. Pestieau (2000). The Political Economy of Social Security, *Scandinavian Journal of Economics*, 102(3), 503-22.

Cigno, A. (2008). "Is there a social security tax wedge," *Labour Economics*, 15(1), 68-77.

Conde-Ruiz, I. and P. Profeta (2007), "The redistributive design of social security systems," *Economic Journal*, 117, 686-712.

Coughlin, P. (1992). *Probabilistic Voting Theory*, Cambridge: Cambridge University Press.

Cremer, H., P. De Donder, D. Maldonado, and P. Pestieau (2007). "Voting over type and generosity of a pension system when some individuals are myopic," *Journal of Public Economics*, 91(10), 2041-2061.

Cremer, H., P. De Donder, D. Maldonado, and P. Pestieau (2008). "Designing a linear pension scheme with forced savings and wage heterogeneity," *International Tax and Public Finance*, 15(5), 547-562.

Cremer, H., P. De Donder, D. Maldonado, and P. Pestieau (2009). “Forced Saving, Redistribution and Nonlinear Social Security Schemes”, *Southern Economic Journal*, 76(1), 86-98.

Cremer, H. and P. Pestieau (1998). "Social insurance, majority voting and labor mobility," *Journal of Public Economics*, 68(3), 397-420.

Cremer, H. and P. Pestieau (2011). "Myopia, redistribution and pensions," *European Economic Review*, 55(2), 165-175.

Diamond, P. A. and B. Koszegi (2003). "Quasi-Hyperbolic Discounting and Retirement," *Journal of Public Economics*, 87, 1839–1872.

Disney, R. (2004). "Are contributions to public pension programmes a tax on employment," *Economic Policy*, 39, 267–311.

Dixit, A., and J. Londregan (1996). "The Determinants of Success of Special Interests in Redistributive Politics," *Journal of Politics*, 58(4), 1132- 1155

Dixit, A., and J. Londregan (1998). "Ideology, Tactics, And Efficiency In Redistributive Politics," *The Quarterly Journal of Economics*, 113(2), 497-529.

Feldstein, M. S. (1985). "The Optimal Level of Social Security Benefits," *The Quarterly Journal of Economics*, 100(2), 303-20.

Feldstein, M. And J. B. Liebman (2002). "Social security," pp. 2245-2324 in: A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, vol.4, North-Holland, Amsterdam.

Gonzalez-Eiras, M. and D. Niepelt (2008). "The future of social security," *Journal of Monetary Economics*, 55(2), 197-218.

Hassler, J. and A. Lindbeck (1997). "Optimal actuarial fairness in pension systems: A note. Economics Letters," 55, 251-255.

Imrohoroglu, A., S. Imrohoroglu and D. H. Joines (2003)."Time-Inconsistent Preferences and Social Security," *The Quarterly Journal of Economics*, 118(2), 745-784.

Kaplow, L. (2006)."Myopia and the Effects of Social Security and Capital Taxation on Labor Supply," NBER Working Papers 12452.

Krieger, K. and S. Traub (2008). "Back to Bismarck? Shifting Preferences for Intragenerational Redistribution in OECD Pension Systems," Working Papers 13, University of Paderborn, CIE Center for International Economics.

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

Loewenstein, G. (1996). "Out of Control: Visceral Influences on Behavior," *Organizational Behavior and Human Decision Processes*, 65(3), 272-292.

Koethenbuerger, M., P. Poutvaara, and P. Profeta (2008). "Why are more redistributive social security systems smaller? A median voter approach" *Oxford Economic Papers*, 60(2), 275-292

Meltzer, A. and S. Richard (1981). "A rational theory of the size of government," *Journal of Political Economy*, 89, 914–927.

Mueller, D. C. (1989). *Probabilistic Majority Rule*, Kyklos, 42(2), 151-70.

Munnell, A. H. (1977). *The Future of Social Security*. Brookings, Washington, D.C.

OECD (2005). *Pensions at a Glance. Public policies across OECD countries*, OECD, Paris.

O'Donoghue, T. and M. Rabin (1999). "Doing It Now or Later," *American Economic Review*, 89, 103–124.

Patty, J. W. (2005). "Local equilibrium equivalence in probabilistic voting models," *Games and Economic Behavior*, 51(2), 523-536.

Pestieau, P. (1999). "The Political Economy of Redistributive Social Security," IMF Working Papers 99/180, International Monetary Fund.

Persson, T. and G. Tabellini (2002). *Political Economics: Explaining Economic Policy*, MIT Press Books, The MIT Press, Edition 1.

Phelps, E. S., and R. A. Pollak (1968). "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35(2), 185-199.

Pigou, A. C. (1920). *The Economics of Welfare*, London: Macmillan.

Plott, C. (1967). "A Notion of Equilibrium and its Possibility under Majority Rule," *American Economic Review*, 57, 787-806.

Profeta, P. (2002). "Retirement and Social Security in a Probabilistic Voting Model," *International Tax and Public Finance*, 9(4), 331-348.

Samuelson, P. A. (1975). "Optimum Social Security in a Life-Cycle Growth Model," *International Economic Review*, 16(3), 539-544.

Strotz, R. H. (1956). "Myopia and Inconsistency in Dynamic Utility Maximization," *The Review of Economic Studies*, 23(3), 165-180.

Tabellini, G. (1991). "The politics of intergenerational redistribution," *Journal of Political Economy*, 99, 335– 357.

Tenhunen, S. and M. Tuomala (2010). "On optimal lifetime redistribution policy", *Journal of Public Economic Theory*, 12(1), 171-198.

Venti, S. (2006). "Choice, Behavior and Retirement Saving". In *Oxford Handbook of Pensions and Retirement Income,* (Ed.) G. Clark, A. Munnell and M. Orzsag, Oxford University Press, Oxford.

Weiss, D. M. (1991). "Paternalistic Pension Policy: Psychological Evidence and Economic Theory, " *University of Chicago Law Review*, 58, 1275-1319.