On sustainable Pay As You Go contribution rules

Gabrielle DEMANGE[1]

April 22, 2009

Abstract

An unfunded social security system faces the major risk, sometimes referred to as ”political risk”, that future generations do not agree to contribute as much as expected. In order to account properly for this risk, the paper considers a political process in which the support to the system is asked from each new born generation. The analysis is conducted in an overlapping generations economy that is subject to macro-economic shocks. As a consequence, the political support varies with the evolution of the economy. The impact of various factors -intra-generational redistribution, risk aversion, financial markets, governmental debt- on the political sustainability of a pay-as-you-go system is discussed.

Keywords pay-as-you-go, social security, risk, political economy, intra-generational redistribution, overlapping generations

Classification D78, H55

[1]Paris School of Economics, 48 bd Jourdan, 75014 Paris, France e-mail demange@pse.ens.fr., and CEPR, Support of the CEPREMAP is greatfully acknowledged.

1 Introduction

In most developed countries, the pension system entails a pay-as-you-go (payg) scheme as a first pillar. A payg system is financed through compulsory contributions and unfunded. As a result, it faces the major risk, sometimes referred to as ”political risk”, that some generations do not agree to contribute as much as expected. The aim of this paper is to propose a positive analysis of this risk. I consider a political process in an overlapping generations economy in which *each* new-born generation is asked whether it supports the current social security design. Their support is determined by a comparison between the returns expected from financial markets and the social security system. Indeed, the current controversies on social security systems are partly due the fact that many young individuals expect the returns on the stock market to be much larger than those from social security.

As made clear by the changes in life expectancy and fertility rates in the last decades, crucial macro-economic and demographic variables that determine the support to a payg system are highly uncertain. So the analysis is conducted in a stochastic environment in which at each period, the economy is subject to shocks on the average labor productivity, the rate of return on capital, and the population growth rate. Shocks are assumed Markovian, and a ’state’ of the economy includes current shocks and possibly lagged ones. By considering a Markovian structure for the economy, we focus on recurrent states as opposed to transition phases, and address the question of how a system could be designed to cope with recurrent fluctuations.

Social security systems vary in two important dimensions:[2] the ”size” of the system, as measured by the contribution rate or the share of social security expenditures of GDP, and the intra-generational redistribution, ranging from the ”bismarckian” or earnings-related systems to ”beveridgean” ones with flat pension benefits. Here, the size of the system is described by a *contribution rule* that specifies the rate that applies in each state of the economy. Clearly, a pension system has to adjust to macro-economic shocks. From an economic perspective, there is no reason to exclude contingent contribution rates. From a political perspective, a contribution rule makes these adjustments more transparent. This is in line with recent reforms, such as the Riester indexation formula in Germany or the so-called *notional* system implemented in Sweden in which annuities are indexed to growth according to an explicit and agreed upon formula. The intra-generational redistribution is described through redistributive factors, which give the distortion of pension benefits with respect to a bismarckian system. We focus our attention on the political support on the size of the system. Specifically the contribution rule is subject to political approval while the redistributive factors are taken as given.

Owing to the redistribution performed by the system, contemporaries may have conflicting interests. Political support is modeled by requiring that an individual, called decisive, agrees (in a sense to be made precise) on the

[2]See for example a cross-national comparison in 44 countries in Europe at http://www.ssa.gov/policy/docs/progdesc/ssptw/2006-2007/europe/index.html.

scale of the contribution rule. In most of the analysis, individuals are two-period lived, and the decisive voter is young. Results do not qualitatively depend on who the decisive voter is, but to fix the idea he (or she) may be thought of as a median voter, or as the poorest agent.

In a stochastic setting, the support to a contribution rule -the decisive individual's agreement- is conditioned on the current state and on the expectations on pension benefits, which are affected both by the (exogenous) evolution of the economy and the support of next generations to the rule. Sustainability defines an equilibrium for the political process and accounts for this forward-looking aspect. A contribution rule is *sustainable* if in each state the current generation supports the rate dictated by the rule when it assumes that future generations will support the rule. In some economies, no (positive) contribution rule is sustainable. The main goal of the paper is to study how the characteristics of the economy -process of the shocks, individuals risk aversion- and the set of available financial instruments affect the existence and the design of a sustainable rule.

Section 3 considers 'short term' financial instruments that allow individuals *within* a generation to exchange the risks they face. The exchange possibilities vary with the set of the financial instruments, and are maximal when the markets contingent on the macro-economic state in next period are complete. Without redistribution nor uncertainty, a simple criterion for a sustainable rule to exist is that the population growth rate be greater than the return to capital. This criterion is extended to our stochastic setup as requiring that the maximal eigenvalue of a given matrix is larger than 1. The matrix is made up of the marginal rates of substitution between the future and current states weighted by the ratio of the rate of growth of population to the return of capital and adjusted to account for the benefit rule. The eigenvalue criterion is a useful tool for studying how the risk characteristics of the shocks and risk aversion affect the existence and the design of a sustainable rule. The inter-temporal correlation of the shocks plays an important role and invalidates a "myopic" comparison at a given date between population growth and investment return. Also, under some conditions, the eigenvalue is easy to interpret as it is equal to the expectation of the ratio of the rate of growth of population to the return on capital next period properly adjusted by the risk aversion and the redistribution factor of the decisive voter. In particular increasing risk aversion increases the chances for a sustainable rule to exist.

Sustainability is affected by the available financial instruments. Without any instrument, only the decisive voter's preferences (in particular his risk aversion) matter. With complete short term instruments instead, all preferences matter as they determine the supporting Arrow-Debreu prices, hence the eigenvalue condition (since the matrix is made up of marginal rates of substitution). Whatever situation, under plausible assumptions about the correlation between population growth and investment returns, the sharing of macro-economic risks between generations favors the possibility of a sustainable rule. The intuition is that, by providing benefits that are linked to labor earnings, a payg system is a tool for improving risk sharing across generations, risk sharing that cannot be provided by short term financial instruments. As discussed below, similar results have been obtained in a full commitment setting in which a planner can choose future contribution rates.

Section 4 examines the interaction between governmental debt and the sustainability of a payg system. In a two-period lived overlapping generations economy, a governmental rolled-over debt performs intergenerational transfers as an unfunded system, as pointed out by Diamond (1965). The newly issued debt is bought by young individuals and the collected amount is used to reimburse the mature debt, which is held by the old generation. In fact, in a risk-less economy, the returns on an unfunded system and on debt are both comparable so that sustainability imposes very strong conditions. In a stochastic framework, debt returns, which include future prices, and pension benefits, which are indexed on current wages, are *a priori* not proportional. Due to differences in their risk profile, the two instruments might not be comparable and could coexist. Our results show however that rolled-over governmental debt severely restricts the possibilities of redistribution of a sustainable payg if there are no frictions on exchanges. In particular, if there are no short sales constraints, a redistributive sustainable payg system cannot coexist with a debt that has a positive value (Theorem 3).

This work can be viewed as extending the economic analysis of intergenerational risk sharing by adding political constraints. Distinct informational or institutional settings call for distinct welfare criteria ranging from *interim* to *ex ante* optimality (Demange 2002). The *interim* concept evaluates the welfare of individuals at birth conditional on their information (Muench 1977). As such it properly accounts for informational constraints and their impact on individuals decisions. The possibility for reaching *interim* Pareto optimal allocations through a payg system or through money has been investigated by various authors (e.g. Peled 1984, Manuelli 1990, Demange and Laroque 1999, Chattopadhyay and Gottardi 1999). With a representative agent (or with short term complete markets) this possibility is characterized by an eigenvalue condition that extends to a stochastic setup the usual condition of (dynamical) inefficiency of the 'autarky' equilibrium (i.e. without intergenerational transfers). With a representative agent, a sustainable rule is akin to money and our eigenvalue condition coincides with that found in this literature. This is no longer true with multiple agents. Due to the distortions generated by the political process and the redistribution operated by a payg system, a rule may be sustainable even though the autarky is efficient and performs different transfers than money.

Intergenerational risk sharing has also been investigated in a planner setting with full commitment and no constraints, neither political nor informational. Generations who are not alive at the same time cannot share their risks through markets. A planner thus can design a payg that improves the *ex ante* welfare of all generations over the laissez-faire situation. Theoretical studies include Gordon and Varian (1988), Bohn (1998), Shiller (1999), Ball and Mankiw (2001). De Menil, Murtin and Sheshinsky (2006), and van Hemert (2005) provide estimations of *ex ante* Pareto improving reforms for several countries. In a setting as here in which full commitment is excluded and information becomes available over time, a planned system is not implementable. At the time a new born generation is asked to vote, it may altogether refuse to contribute the amount that was planned previously on the basis of the information that has been revealed on the current state. As in most insurance problems, information may have a perverse effect on insurance possibilities. It is therefore unclear whether intergenerational risk sharing is still relevant in assessing the sustainability of a payg system. According to our results, however, intergenerational risk sharing may still promote sustainability even if information and the associated political constraints are taken into account.

This paper is also related to the vast literature that aims at explaining intergenerational transfers without invoking altruism or commitment. In a seminal paper, Hammond (1975) points out the fundamental role played by expectations in explaining why workers agree to finance the pensions of the current retirees. Considering a 'pension game' with a prisoner dilemma flavor, defecting from performing a prescribed transfer is a dominant strategy for generations in the finite horizon version of the game. As a result, transfers can be sustained only if current generations expect the system to last forever. Apart from 'good' expectations along the equilibrium path, the threat of collapse, the possibility of punishment, and the introduction of institutions that are costly to change promote the support to the system (Browning 1975, Kandori 1992, Esteban and Sakovicks 1993). Once a system is in place, the threat of a total collapse creates a distortion in favor of the system. Middle-aged individuals, treating their past contributions as sunk costs, may find it in their interest to vote in favor of the system even though, overall, they end up losing from the system. In an economy calibrated on the US data with four period lived generations, Cooley and Soares (1999) find that the effect is important enough to lead US citizens to support an unfunded system even though voters account for general equilibrium effects. However, the role played by the threat of collapse and the possibility of a punishment inflicted forever to all generations may seem unrealistic. To limit the possibility of infinite punishment, Forni (2004) considers in a deterministic economy Markovian equilibria in which the state is the level of capital. To compare with this literature, our model can be seen as a pension game in a stochastic environment with the following features. Since individuals are two-period lived, there are no distortions due to past contributions. Individuals are small, and do not take into account general equilibrium or signaling effect following a change in the payg system. Furthermore, the strategy choice is the scale of the system, hence not reduced to the acceptance or the rejection of the system, and stationary equilibria are considered. The threat of unrealistic punishment is thus avoided. These assumptions are of course debatable but examining their implications in a stochastic model that is simple enough to understand the underlying mechanisms is worthwhile.

Finally, the paper is also related to the 'political approach' to social security, initiated by Browning (1975), which aims to explain the factors determining the size of a system and to assess the impact of the redistribution performed by a system by considering various decision-making processes such as planner, median voter, lobbies.

Casamatta *et al.* (2000) for instance analyze the determination of the scale of the system in a majority game, taking the intra-generational redistribution fixed. Their aim is to explain the 'puzzle' according to which Beveridgean systems, with large redistribution, are often of smaller size than the more Bismarckian ones.[3] This paper can be seen as extending to a stochastic setting (and a non necessarily majority process) their modeling approach. A joint determination of the size and the redistributiveness of the system in a majority game faces the difficulty that a majority winner may fail to exist (see however Conde-Ruiz and Profeta (2007) who propose such a joint determination by considering an issue-by-issue equilibrium). Our focus on the size of the social security system is justified by the fact that, in most countries, the intra-generational redistribution is not much discussed in contrast to its size.

The paper is organized as follows. Section 2 sets up the model and defines the concept of sustainable rule. Section 3 studies the existence and properties of a sustainable rule when short term financial securities are available. Section 4 discusses extensions to mixed systems, multi-lived individuals, and the introduction of capital. The interaction between sustainability and governmental debt is studied in Section 5. Proofs are gathered in the final Section.

2 The model

2.1 The economy

The overlapping generations economy is populated with heterogeneous individuals in each generation and is subject to macro-economic uncertainty. To keep the structure simple there is a single good that can be either consumed or invested and the investment and labor productivities are exogenous.

Individuals. Each generation is composed with I-types of individuals indexed by i, $i = 1, \ldots, I$, that grow at an identical rate, which is the population growth rate. Individuals live for two periods, supply a fixed quantity of labor when young, and retire when old. Each i is characterized by a productivity parameter, θ^i, which determines the individual's wage in conjunction with the state of the economy as made precise below. Individuals are not altruistic. Individual i' preferences are represented by utility function U^i defined over positive consumption plans[4] $(c_y, \widetilde{c}_o)$, where c_y and c_o denote consumption when young and old respectively. Function U^i is concave, strictly

[3] The puzzle arises under the assumption that the system is determined by majority. The reason is that, in simple risk-less economies without any borrowing constraint, the median voter is an individual with median income. Given that median income is typically smaller than average income, the median voter benefits from intra-generational redistribution. Hence, the more redistributive the system is, the larger the size of the system should be. In the presence of borrowing constraints, the median voter may not be an individual with median income, which solves the puzzle.

[4] $\widetilde{x}$ denotes a random variable and x its realization.

increasing in each argument and continuously differentiable. It can be of the Von Neumann Morgenstern type or of the recursive type, that is, dropping index, of the form $E[u(c_y, \widetilde{c}_o)]$ or $u(c_y) + \delta u(v^{-1}(Ev(\widetilde{c}_o))$. Furthermore, to avoid corner solutions, Inada conditions are assumed: $\lim_{c\to 0} U_c^{i'}(c_y, c_o) = \infty$ for $c = c_y$ or c_o.

Macro-economic shocks. The rate of growth of the population, the average labor productivity, and the rate of return on capital are subject to shocks at each date. The rate of growth of the population between dates $t-1$ and t is denoted by γ_t. Since labor is supplied inelastically and by the young generation only, the ratio workers to retirees at t is also equal to γ_t. Labor productivity at date t, denoted by w_t, determines the level of labor earnings at that date. The wage income of a worker with characteristic θ at t is given by θw_t. Normalizing the distribution of the parameter θ across workers to 1, w_t stands for the *average* wage income at date t. The good may be transferred from one period to the next through a linear random technology with return $\widetilde{\rho}_{t+1}$: s_t units invested at date t yield $\widetilde{\rho}_{t+1} s_t$ in period $t+1$.

The variables $(\widetilde{\gamma}_t, \widetilde{w}_t, \widetilde{\rho}_t)$ are assumed to follow a first order Markovian process: the distribution of the shocks at date $t+1$ conditional on what happened at date t is invariant over time. We shall call the realization of the shock $e = (\widetilde{\gamma}, \widetilde{w}, \widetilde{\rho})$ at some date the *state* of the economy. For simplicity, the state space E is assumed to be finite (the Markov assumption and the choice of the state space will be discussed at the end of this section). The transition probability from state e to state $e_+ = (\gamma_+, w_+, \rho_+)$ is denoted by $\Pr(e_+|e)$. We assume that the process has a unique invariant distribution (which holds true for example if all transition probabilities are positive).

The realized state at t, (γ_t, w_t, ρ_t) is observed at time t, and the transition probabilities are known. Thus, young individuals correctly expect the distribution of the states at the subsequent period conditional on the observed current state to be given by Pr.

Remark. Productivity growth can be handled with by considering γ_t as the growth of effective labor. Then w_t stands for the transitory shocks. That $\widetilde{\gamma}_t$ is possibly perceived as random at date $t-1$ is far from unrealistic. In most developed countries, especially european, the decline in the fertility rate and the sharp increase in life expectancy were not expected to be so severe. Also labor participation is rather unpredictable, owing to changes in behavior such as women working decision, or in legislation affecting the retirement date and the number of working hours for instance.[5]

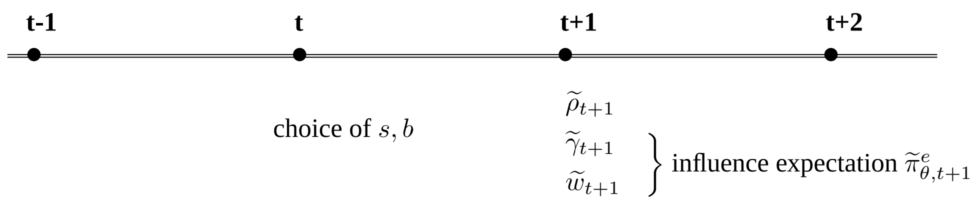
Individuals' budget constraints. Individuals' budget constraints and risks are affected by the available financial instruments and social security system. In addition to the constant returns to scale technology, individuals may have access to financial markets. Several cases will be considered, with short term securities in section 3 or government bonds in section 5. Let p_t denote the securities prices at the current period and $\widetilde{a}_{t+1}$ denote the payoffs yielded by these securities next period (both are vectors if there are multiple securities).

[5]This point also suggests that the ratio workers to retirees is in part endogenous, sensitive to some policies, in particular to social security. This aspect is not addressed here.

Social security is compulsory, levied through a tax bearing on labor income. Let τ_t be the contribution tax rate at date t. Given the labor productivity shock, w_t, each $\theta-$worker contributes $\tau_t \theta w_t$ to the system. In return, he expects to receive some pension benefits $\widetilde{\pi}^e_{\theta,t+1}$ at retirement. How pensions are distributed and how expectations are formed will be described in the next sections.

Given τ_t, p_t, $\widetilde{a}_{t+1}$ and the expected benefits $\widetilde{\pi}^e_{\theta,t+1}$, the budget constraints faced by a $\theta-$individual who invests s in the technology and buys portfolio b are:

$$\begin{cases} c^y + s + b.p_t = (1-\tau_t)\theta w_t, s \geq 0 \\ \widetilde{c}^o = s\widetilde{\rho}_{t+1} + b.\widetilde{a}_{t+1} + \widetilde{\pi}^e_{\theta,t+1} \end{cases} \tag{1}$$



In the first period, labor income is used to consume, to invest in the technology, to buy (or sell) financial securities, and finally to contribute $\tau_t \theta w_t$ to the pension system. In the second period, the individual retires and consumes all of his resources. The decisions when young depend not only on the distribution of next investment return but also on next population growth and wage level if a pension system is in place, as depicted in the diagram.

2.2 Unfunded pension schemes

The analysis focuses on unfunded pension systems, except in Section 4 which considers systems that are partially funded. In an unfunded system, the collected amount at any date is fully transferred to retirees. At date t, given the labor productivity shock, w_t, and the current tax level, τ_t, workers contribute on average $\tau_t w_t$. Therefore, if population has grown by the factor γ_t between $t-1$ and t, the per head *average* pension benefits are equal to

$$\pi_t = \gamma_t \tau_t w_t. \tag{2}$$

The pension benefits received by a particular retiree may differ from the average level π_t, except in a beveridgean system. In a bismarckian (also called purely contributive) system for instance, pension benefits are proportional to contributions, equal to $\theta \pi_t$ for a $\theta-$individual. More generally, a rule determines the pension benefits in relation to previous contributions. This rule is described here through the *redistributive factors*, $\mu(\theta)$, which give the distortion with respect to a bismarckian system. More precisely, the benefits $\pi_{\theta,t}$ received by a $\theta-$retiree are given by:[6]

$$\pi_{\theta,t} = \theta\mu(\theta)\pi_t = \theta\mu(\theta)\gamma_t\tau_t w_t. \tag{3}$$

The function μ is positive, non-increasing (to describe redistribution), and satisfies $\sum_i \theta_i\mu(\theta_i) = 1$ so as to ensure budget balance. For instance a system that combines bismarckian and a beveridgean systems in fixed proportions as considered by Casamatta *et al* (2000) is described by $\mu(\theta) = (\alpha\theta + (1-\alpha))/\theta$ for some α between 0 and 1.

2.3 Sustainable contribution rule

In most countries, the redistributive function of the payg system is not much discussed and is rather stable. This justifies to analyze the political support to the level of the contribution rates taking the benefit rule as given. Thus the redistributive factors μ are fixed throughout the paper. We focus our attention on systems that specify how the contribution rate is adjusted in function of the state of the economy. Such a system is described by a *contribution rule* $\boldsymbol{\tau} = (\tau(e))$. Given $\boldsymbol{\tau}$, the contribution rate at time t is set to τ_t equal to $\tau(e_t)$. We shall focus on rules with strictly positive rates, written as $\boldsymbol{\tau} > 0$ (the results are easily modified to rules with some null rates).

A contribution rule is interpreted as the pension system designed and announced by the social security institution. To analyze political risk, the contribution rule is subject to approval by each new generation. Equilibrium requires *each* generation to support the system. In our stationary framework, this amounts to require that the contribution rule is approved in each state. It remains to make precise what we mean by political approval.

The political support is modeled as a pension game, the features of which are justified as follows. The future contribution rule cannot be freely chosen by the current generation (next period rates would be set at their maximum, which does not make much sense and would never be fulfilled). Future rates cannot be considered as fixed either because then the current generation would always choose a null contribution rate. We consider a process that avoids these drawbacks and retains the main point of discussion about pension systems, namely their general level: it is the scale at which the rule is applied that is subject to approval.[7] Specifically, let $\boldsymbol{\tau}$ be a contribution rule announced by the social security institution. Voters are asked the scale at which they would like the rule to be implemented. Changing the scale leads to changes in the level of their current contributions (which depends on the current state) and in the level of their benefits next period in the same proportion. Thus, given the current state e, individuals expect the choice of a scale level λ to change the current rate $\boldsymbol{\tau}(e)$ into $\lambda\boldsymbol{\tau}(e)$ and the (contingent) rates next period $\boldsymbol{\tau}$ into $\lambda\boldsymbol{\tau}$. A rule $\boldsymbol{\tau}$ is said to be sustainable if no generation wants to change the scale, that is if λ equal to 1 is an equilibrium in each state.

[6]The replacement ratio depends both on the redistributive factor and the contribution level. For wages that gradually adjust along the working period, a θ-worker born at $t-1$ earns θw_t before retirement so that the replacement ratio is equal to $\gamma_t\tau_t\mu(\theta)$.

[7]As discussed in the introduction, considering the strategy choice to be the scale of the system instead of a zero-one decision avoids giving an unrealistic role to the threat of collapse.

An interpretation of this process is that, given a contribution rule, the social security commits to change the scale forever (or at least in the current and subsequent periods) according to the outcome of the voting process and voters believe it. The commitment is credible and the rule will indeed be implemented without modifications only if the rule is sustainable (see another interpretation following the formal definition). Then the main questions we ask are the following ones. Under which conditions is there a sustainable rule $\boldsymbol{\tau}$? What are the main factors that favor sustainability ?

Let us formalize a voter's decision problem (dropping unnecessary index). Given the rule $\boldsymbol{\tau}$ and the state e, a θ-individual contributes $\theta(\tau w)(e)$ and expects to receive pension benefits equal to $\theta\mu(\theta)(\gamma w\tau)(e_+)$ if state e_+ realizes. For an individual born in a given state, changing the scale leads to changes in the level of his current positive contribution rate *and* in the level of the benefits next period in the same proportion. The choice of a scale level λ leads to multiply all these amounts by λ. This modifies the successive budget constraints as given by (1) and leads to an adjustment in consumption and investment decisions. (we use here that the voter does not account for the impact that a change in the scale may have on equilibrium asset prices). Specifically, a θ-voter born in state e anticipates by choosing λ the indirect utility level $V(\lambda, e)$ defined by :

$$\begin{cases} V(\lambda, e) = \max_{c_y, s \geq 0, b} \;\; U(c^y, \widetilde{c}^o) \\ c^y + s + b.p = (1 - \lambda\tau(e))\theta w(e), \\ \widetilde{c}^o = s\widetilde{\rho} + b.\widetilde{a} + \lambda\theta\mu(\theta)(\gamma w\tau)(e_+) \end{cases} \tag{4}$$

It remains to determine the outcome of the voting process. To be not too specific, the outcome is described by the choice of a "decisive" voter, who is a member of the new born generation. The decision process by which the decisive voter is chosen is left unspecified. In case of a majority process for instance, the decisive voter is the median voter corresponding to the state (who is shown to exist). A sustainable rule is defined as a rule that will indeed be implemented because no decisive voter wants to change the scale of the system.

Definition 1 *A contribution rule $\boldsymbol{\tau} = (\tau(e))$, $\boldsymbol{\tau} > 0$, is said to be sustainable if the decisive voter agrees on the scale level in the sense that the indirect utility function $V(\lambda, e)$ defined by (4) is maximized at $\lambda = 1$.*

The design is sustainable if the decisive voter agrees in each possible state. Since the decisive voter can perform the computation whatever the state, future approval is credible. Put differently, sustainability is self enforcing: the current rate is approved when next generation is expected to contribute according to $\boldsymbol{\tau}$, and next generation will indeed agree to contribute that much, in each possible state, given that it expects the following one to contribute according to $\boldsymbol{\tau}$ and so on at infinitum. As discussed in the introduction, the decisive individual does not take into account general equilibrium or signaling effect following a change in the payg system.

The sustainability concept can be defined without considering stationary and announced contribution rules. Sustainability then refers to a rational expectations equilibrium. When asked to vote, expectations on how the vote may change future benefits is crucial. Assume that individuals think in term of the scale of the system leads to link the current choice to next benefits in a way similar to the one we have defined. Basically voters *expect* that the level of the benefits next period would be changed in the same proportion as the *current* positive contribution rate. In this interpretation, the political support depends on expectations and as such may be thought more fragile than a voting process on a contribution rule. This makes the support to a payg system akin to the support to money.

To illustrate the model, consider a majority voting game in which individuals vote on the scale λ on the contribution rates. Old individuals trivially prefer the maximal scale level defined by $\lambda\tau = 1$. As for young individuals, their preferences over scale levels are represented by the indirect utility function V defined in (4) (possibly different V for different individuals). Since each V is concave hence single-peaked, a median voter exists. The contribution rule is sustainable when the median voter agrees on it in each state. A majority of the population would like to increase the scale and another would like to decrease it.

Discussion about the stationary framework By considering a Markovian structure for the economy, we focus on recurrent states. The stationarity assumption has a long tradition in economic theory (see the literature mentioned in the introduction starting from Muench 1977 and Peled 1984 and more recent work such as the calibrated US economy in Geanakoplos *et al* 2004). Many countries have been experiencing over the last half century a once-for-all shock in their demography that also calls for a (politically sustainable) transition. Extending the analysis to the case where there are both recurrent and transient shocks would be much interesting.

Restricting to stationary policies leaves flexibility. Although the state must include the vector of current shocks, which is the minimal data on which decisions are made, the analysis extends easily to a situation where the state is 'enlarged'. If for instance contribution rates are contingent not only on the current shocks but also on some lagged shocks, then individuals' decisions would also depend on this enlarged state. Everything goes through provided that the state space, say E^*, is finite and that there is a unique invariant distribution the support of which is the whole space E^*. Also, there are sustainable rules that may depend on a state that is not related to the 'fundamentals'. In a risk-less economy for instance where a stationary payg system amounts to money with constant price, equilibria with sunspots may exist (see Azariadis and Galasso 2002).

Considering an infinite state space introduces technical difficulties. Such difficulties have been partly solved by Demange and Laroque (2000) in a similar single good economy setting. Results do not change fundamentally as long as the state is exogenous. Most often however, one has to deal with an infinite state space because of an endogenous stock variable such as capital. I come back to this point in Section 4.

First order conditions on sustainable rules Let us introduce some notation. Individual i's consumption and portfolio decisions when young are described by functions of the state e at birth, c_y^i, s^i, and b^i, and when old by a function c_o^i of both states e and e_+ that realize during i's lifetime. Let $q^i(e_+|e)$ denote i's marginal rate of substitution between current consumption in state e and consumption next period in state e_+. Dropping index i, the marginal rate of substitution writes for an expected utility function u as

$$q(e_+|e) = \frac{u_{c_o}(c_y(e), c_o(e, e_+))}{E[u_{c_y}(c_y(e), c_o(e, \widetilde{e}_+))|e]} \text{Pr}(e_+|e) \tag{5}$$

and for a recursive utility, $U(c_y, \widetilde{c}_o) = u(c_y) + \delta u(v^{-1}(Ev(\widetilde{c}_o))$, as

$$q(e_+|e) = \delta \frac{u'(\hat{c}_o(e))}{u'(c_y(e))} \frac{v'(c_o(e, e_+))}{v'(\hat{c}_o(e))} \text{Pr}(e_+|e) \tag{6}$$

where $\hat{c}_o(e)$ is the certain equivalent[8] of $\widetilde{c}_o(e, e_+)$ knowing current state e. Finally, the decisive voter characteristic in state e is denoted by $\theta^d(e)$ and the associated redistributive factor by $\mu^d(e)$, that is $\mu^d(e) = \mu(\theta^d(e))$.

By the concavity of the function V in λ, the agreement of the decisive voter is equivalent to the first order condition, $V'(1, e) = 0$ (agreement needed only in a state with a positive rate). Denoting $q^d_{\boldsymbol{\tau}}(e_+|e)$ the voter's marginal rate of substitution at his optimal consumption plan given rule $\boldsymbol{\tau}$, sustainability is characterized as follows.

The contribution rule $\boldsymbol{\tau} > 0$ is sustainable if and only if in each state e

$$\tau(e)w(e) = \mu^d(e) \sum_{e_+} q^d_{\boldsymbol{\tau}}(e_+|e)(\gamma \tau w)(e_+). \tag{7}$$

These conditions (7) can be put in matrix form :

$$\boldsymbol{\tau} = M_{\boldsymbol{\tau}} \boldsymbol{\tau} \tag{8}$$

in which the positive matrix $M_{\boldsymbol{\tau}}$ is defined by

$$M_{\boldsymbol{\tau}}(e, e_+) = \mu^d(e)\gamma(e_+) \frac{w(e_+)}{w(e)} q^d_{\boldsymbol{\tau}}(e_+|e) \tag{9}$$

The system (8), which has as many unknowns as equations, make clear that sustainability imposes strong conditions on how the rates should vary with the state. Also it allows us to analyze the conditions under which a sustainable rule exists. Different sources of risks -on investment and pension benefits- affect income at retirement. Young individuals' saving behavior and support to social security are influenced by the insurance possibilities offered by financial securities. The next section assumes that only one-period lived securities are available.

3 Sustainability with one-period lived financial securities

Exchanges of one-period lived securities take place between contemporaries. Thus, there are no transfers between generations when there is no payg system (i.e. $\boldsymbol{\tau} = 0$). Such a situation is referred to as 'autarky'. We show that

[8]The utility level derived from $(c_y, \widetilde{c}_o)$ can be written as $u(c_y) + \delta v(\hat{c}_o)$, the inter-temporal utility derived from the risk-free consumption plan $(c_y, \hat{c}_o)$ where $\hat{c}_o$ is the certain equivalent of $\widetilde{c}_o$ under the Von Neumann Morgenstern utility v: $v(\hat{c}_o) = Ev(\widetilde{c}_o)$.

the existence of a sustainable rule is much related to the properties of the autarky equilibrium. This equilibrium depends on the available financial instruments. The more tractable situation is the one in which no financial securities are available. Although not the more realistic situation, it is instructive to start with it.

3.1 Sustainable rule without financial securities

Individuals can only invest in the technology, which yields the exogenous rate of return $\widetilde{\rho}$. This gives the following budget constraints for a young $\theta-$worker born in state e under rule $\boldsymbol{\tau}$:

$$\begin{cases} c_y + s = (1-\tau(e))\theta w(e) \\ c_o(e_+) = \rho(e_+)s + \theta\mu(\theta)(\gamma\tau w)(e_+) \text{ for each } e_+ \end{cases} \tag{10}$$

Given $\boldsymbol{\tau}$, each individual chooses to invest some non negative amount s so as to maximize his utility conditional on the current state e under (10). The market for the consumption good is automatically balanced.

Theorem 1 *Consider a stationary economy without financial securities. Let $M_{\mathbf{0}}$ be the matrix of the decisive voters' weighted inter-temporal rates of substitution defined in (9) computed at the autarky equilibrium. There is a sustainable rule if and only if the largest eigenvalue of $M_{\mathbf{0}}$ is above 1.*

In a deterministic economy the eigenvalue condition is easy to understand. At autarky, individuals can only invest in the technology to get some income at retirement. Therefore, the marginal rate of substitution is equalized to the reciprocal of ρ, so that the unique element of matrix $M_{\mathbf{0}}$ is equal to $\mu^d\gamma/\rho$. Hence the eigenvalue condition simply says that, from the point of view of the decisive voter, the (risk-less) return of the payg system dominates the technology return.[9] In a stochastic environment, returns cannot be compared so easily: they are risky, and are moreover endogenous through the variations in contribution rates. To understand and interpret the eigenvalue condition, consider the introduction of small contribution rates "in the direction" of $\boldsymbol{\tau}_0$, i.e. of the form $\epsilon\boldsymbol{\tau}_0$ for $\epsilon > 0$. Consumption levels are changed but, from the envelope theorem, taking ϵ small enough, their impact is negligible. The impact on the decisive voters' utility born in state e is therefore proportional to

$$\mu^d(e)\sum_{e_+} q_0^d(e_+|e)\gamma(e_+)\tau_0(e_+)w(e_+) - \tau_0(e)w(e), \tag{11}$$

which, up to $w(e)$, is equal to $(M_{\mathbf{0}}\boldsymbol{\tau}_0 - \boldsymbol{\tau}_0)(e)$. Therefore small contribution rates in the direction of $\boldsymbol{\tau}_0$ make decisive voters better off than at autarky *whatever the state at birth* if the inequality $M_{\mathbf{0}}\boldsymbol{\tau}_0 > \boldsymbol{\tau}_0$ is satisfied. From a well known result on positive matrices, such direction exists if and only if the matrix $M_{\mathbf{0}}$ has an eigenvalue larger

[9] A positive sustainable rate maximizes $U^d(\theta^d w(1-\tau), \mu^d\theta^d\gamma\tau w)$. If $\mu^d\gamma < \rho$, the decisive voter prefers to invest in the technology, hence he chooses $\tau = 0$.

than 1. This readily gives an interpretation of the eigenvalue condition: there is a rule that is Pareto improving for decisive voters over autarky. Such a rule however has few chances to be sustainable (since $\boldsymbol{\tau}_0$, or any system proportional to it, does not satisfy $\boldsymbol{\tau} = M_{\boldsymbol{\tau}}\boldsymbol{\tau}$ in general). The proof of existence, under the eigenvalue condition, relies on a fixed point argument.

Why is the eigenvalue condition necessary? Whatever $\boldsymbol{\tau}$, the decisive voter can choose his autarky utility level[10] by setting λ equal to 0. This readily implies that whatever the state at birth, the decisive voter is better off at a positive sustainable system $\boldsymbol{\tau}$ than at autarky. By a concavity argument, decisive voters are also made better off by the introduction of small contribution rates "in the direction" of $\boldsymbol{\tau}$. Thus expression () at $\boldsymbol{\tau}$ is positive in each state, i.e. $M_{\mathbf{0}}\boldsymbol{\tau} > \boldsymbol{\tau}$. Since $M_{\mathbf{0}}$ is a positive matrix and $\boldsymbol{\tau}$ is positive, the eigenvalue is larger than 1.

A similar eigenvalue condition appears in the literature that studies *interim* Pareto optimal allocations (for which no other feasible allocation gives a larger expected utility to each individual at birth whatever the state). We show that these results coincide with ours when there are no distortions due to redistribution or to the lack of exchange risks between contemporaries, that is when there is a representative individual per generation. From the just above argument, there is a sustainable rule only if the autarky situation is not *interim* Pareto efficient (since the decisive voter can only be the representative agent). Hence, the existence of a sustainable rule coincides with the inefficiency of the autarky equilibrium, which explains why it is characterized by the same eigenvalue condition as that obtained in the above referred literature. Furthermore, one can also infer that a sustainable rule leads to an *interim* Pareto optimal allocation in that setting. Using similar arguments as in Demange and Laroque (1999) on voluntary subscriptions to a payg, a sustainable rule is akin to an equilibrium on a money asset, which is known to be *interim* Pareto optimal (Peled 1984).

These correspondences do no extend to heterogeneous individuals because of the redistributive factors and the difference in individuals contingent values. As a result, a sustainable rule may not lead to an *interim* Pareto improvement over autarky. Furthermore, the inefficiency of the autarky equilibrium is typically not characterized by a single number, the eigenvalue, that relates to the efficiency of intergenerational transfers from the young individuals towards the old ones (see Demange 2002). Finally, a sustainable rule cannot be the interpreted as an equilibrium with a money asset. Since with adequate financial securities, individuals contingent values are equalized, we shall investigate again this question in next section.

The characterization in Theorem 1 is useful to understand the determinants that favor sustainability. Increasing any element of the matrix $M_{\mathbf{0}}$ increases the maximal eigenvalue. Hence, not surprisingly, whether there is a sustainable rule depends positively on the decisive voters redistributive factors, and on the population growth rate (note that the marginal rates of substitution at autarky are independent of γ). We analyze further how the stochastic

[10] We make use here that changing the level of the contributions only affects the distribution of endowments and has no price effect. Such an argument does not extend when there are financial securities.

process and risk aversion influences sustainability.

3.2 The impact of the stochastic process and risk aversion

Stochastic process The stochastic process, that is the transition probability Pr, influences sustainability. We first illustrate this influence with the extreme case of highly persistent states, for which the transition matrix is close to the identity matrix.

Corollary *For highly persistent states, a sustainable rule exists if the maximum of* $(\mu^d \frac{\gamma}{\rho})(e)$ *over the states is slightly larger than 1.*

With[11] persistent states, next state is known almost surely to be the current one. If it was to be known perfectly, the standard criterion of comparison between the rate of population growth and the technology rate of return, adjusted by the redistribution factor, $\mu^d(e)\gamma(e) > \rho(e)$, would have to be satisfied to achieve sustainability. With some uncertainty, the criterion needs to be satisfied in one state only according to the corollary. Sustainability is achieved by setting a low contribution rate in states where the criterion is not satisfied, if any, much lower than in states where the criterion holds, as illustrated in next example.

Example 1 There are only two states, h and l with values (γ_h, ρ_h) and (γ_l, ρ_l) respectively and identical w. Let ratio γ/ρ be the largest in state h. Take also $\mu^d(e) = 1$ and drop index d (the analysis is directly transposed to a constant μ^d). Let p_h denote $\Pr(h|h)$, that is the probability of a future high ratio γ_h/ρ_h conditional on a current high one, and similarly for p_l. To analyze the impact of the process, let us write $q(e_+|e) = Pr(e_+|e)\text{mr}(e_+|e)$ where $\text{mr}(|)$ denotes the ratio of marginal utilities in the corresponding states.

At autarky, individuals save in the technology, since otherwise they would not consume when old. Therefore the first order conditions $p_h\text{mr}(h|h)\rho_h + (1-p_h)\text{mr}(l|h)\rho_l = 1$ and $(1-p_l)\text{mr}(h|l)\rho_h + p_l\text{mr}(l|l)\rho_l = 1$ are satisfied. Matrix $M_\mathbf{0}$ writes

$$M_\mathbf{0} = \begin{pmatrix} \gamma_h p_h \text{mr}(h|h) & \gamma_l(1-p_h)\text{mr}(l|h) \\ \gamma_h(1-p_l)\text{mr}(h|l) & \gamma_l p_l \text{mr}(l|l) \end{pmatrix}$$

whose maximal eigenvalue is equal to $S/2 + \sqrt{(S/2)^2 - det}$ where S is the sum of the diagonal elements and det the determinant. This value, hence the existence of a sustainable rule, is much related to the stochastic process. To see this, consider the two extreme cases of either persistent or switching states. Whatever case, next state is known

[11] The proof is as follows. At autarky, individuals save in the technology, since otherwise they would not consume when old. Therefore the first order conditions on investment $\sum q(e_+|e)\rho(e_+) = 1$. For highly persistent states all terms off the diagonal of matrix $M_\mathbf{0}$ are close to zero, which implies that $q(e|e)$ is close to $1/\rho(e)$ and the e diagonal element close to $\mu^d(e)\frac{\gamma}{\rho}(e)$.

almost surely. If it is state h (resp. l), the ratio $\mathrm{mr}(h|.)$ is equalized to $1/\rho_h$ (resp. $1/\rho_l$) (which explains why risk aversion does not matter in the following discussion). This gives a value close to γ_h/ρ_h for (almost) persistent states, and $\sqrt{\gamma_h\gamma_l/\rho_h\rho_l}$ for switching states. Let us see why.

(1) States are persistent if p_h and p_d are close to 1. Sustainability condition is close to $\gamma_h > \rho_h$, which is compatible with $\gamma_l < \rho_l$. The sustainable rule is as follows. The contribution rate in the low state is positive, but much smaller than in the high state. Accordingly, if the current state is low, whereas the return on the payg system can be smaller than the technology return (if $\gamma_l < \rho_l$) with high probability p_l, this is compensated by the large return obtained if next state is high. If the current state is high, the return on the payg system is larger than ρ with a large probability.

(2) In the case of a perfect switch between the states (p_h and p_l are null), there is a sustainable rule if

$$\gamma_h\gamma_l > \rho_h\rho_l. \tag{12}$$

Note that condition (12) is compatible with $\gamma_l < \rho_l$. In that case, in state h next population growth is smaller for sure than the technology return. To understand why a sustainable system nevertheless exists under (12), consider small contribution rates in the "direction" of $(\tau_0(h), \tau_0(l)) = (\sqrt{\gamma_l\rho_h}, \sqrt{\gamma_h\rho_l})$. We show that they improve the decisive voter's welfare over autarky in each state. In state h, the return on the payg system is risk-less equal to $\gamma_l\tau_0(l)/\tau_0(h)$, that is equal to $\sqrt{\gamma_h\gamma_l\rho_l/\rho_h}$. It is larger than the technology return, ρ_l under (12): the current contribution is sufficiently low compared to the next one so that the payg system becomes attractive. Similarly in state l, the decisive voter faces the return $\sqrt{\gamma_h\gamma_l\rho_h/\rho_l}$, larger than ρ_h.

This case shows that a "myopic" comparison at a given date between population growth and investment return, even risk-less, is not sufficient to assess sustainability. It should be clear that allowing contribution rates to be adjusted with the economic state is essential. ■

The impact of risk aversion To analyze the role of risk aversion, it is convenient to assume voters' utility functions to be recursive, given by $U(c_y, \widetilde{c}_o) = u(c_y) + \delta u(v^{-1}(Ev(\widetilde{c}_o))$. Inter-temporal substitution, which depends on u, is independent of risk aversion, which depends on v. This independence allows us to perform comparative statics on the sustainability condition with respect to risk aversion only. As stated by next proposition, under some additional assumptions, the maximal eigenvalue can be computed and interpreted.

Proposition 1 *Assume the decisive voter's utility to be recursive with v homothetic, states to be independent across periods, and $\mu^d(e)$ to be constant across states equal to μ^d. Then the maximal eigenvalue of $M_{\mathbf{0}}$ is given by*

$$\mu^d E[\frac{\widetilde{\gamma}}{\widetilde{\rho}}\frac{v'(\widetilde{\rho})\widetilde{\rho}}{E[v'(\widetilde{\rho})\widetilde{\rho}]}] \tag{13}$$

Assume in addition that $E[\widetilde{\gamma}|\rho]/\rho$ is non-increasing, The eigenvalue increases with risk aversion: the more risk averse the decisive individual, the more likely a sustainable rule to exist.

From (13), with risk neutral individuals, the eigenvalue is equal to $\mu^d E[\widetilde{\gamma}]/E[\widetilde{\rho}]$: the condition for sustainability in a risk-less economy is simply transposed by considering expected values for population growth and investment return. With risk averse individuals, the maximal eigenvalue is μ^d times the expected value of the ratio $\widetilde{\gamma}/\widetilde{\rho}$ under a "risk-neutral" probability that accounts for risk aversion (since the function $v'(\rho)\rho/E[v'(\widetilde{\rho})\widetilde{\rho}]$ defines a density). Increasing risk aversion distorts this density by putting more weight on low values of ρ. Note that the additional condition that $E[\widetilde{\gamma}|\rho]/\rho$ be non-increasing in ρ is plausible. It holds under independence between $\widetilde{\rho}$ and $\widetilde{\gamma}$ for instance. That, under this condition, risk aversion favors the existence of a sustainable rule is quite easy to interpret. The introduction of a payg system provides pension benefits that allow for a partial hedge against investment risk, which encourages risk averse individuals to support the system. The effect of risk aversion will be illustrated further by example 2 in next section.

3.3 Financial securities

In the previous section, we have shown that the risk sharing opportunities provided by a payg system may favor its sustainability. The result however is obtained when there are no financial markets at all. Owing to differences in revenues and tastes -in particular in preferences for present consumption and attitudes towards risk- young individuals within a generation may benefit from exchanging among themselves. To which extent is the previous analysis driven by the lack of risk sharing tools within a generation and the fact that a payg system may be a substitute to them? To investigate this question, this section introduces financial securities that allow young individuals to exchange these risks.

Securities are one-period lived. They are traded at one date in exchange of a (final) payoff at the subsequent date. Thus only the young individuals within a generation exchange the securities. There are K securities, each one described by its non negative payoff in each state, $a_k(e)$ for security k if state e materializes (the K payoffs vectors can be assumed independent). An equilibrium for the financial securities is established between the young generation 'state by state' as follows. Given contribution rates $\boldsymbol{\tau}$ and the current state e, let p_k be the current price of security k in terms of current good. A young $\theta-$individual faces the budget constraints :

$$\begin{cases} c_y + s + \sum_k p_k b_k = (1-\tau(e))\theta w(e) \\ c_o(e_+) = \rho(e_+)s + \sum_k b_k a_k(e_+) + \theta\mu(\theta)(\gamma\tau w)(e_+) \text{ for each } e_+ \end{cases} \tag{14}$$

Each individual chooses to invest and consume so as to maximize his utility conditional on the current state. Thus a securities price equilibrium in state e is a K-vector, $p_{\boldsymbol{\tau}}(e)$, for which the aggregate demand for the securities is null, namely $\sum_i b_k^i = 0$ for each k. Thanks to Walras law, the market for the consumption good is balanced. Using standard arguments, there is an equilibrium in each state for any $\boldsymbol{\tau}$ with $0 \leq \tau(e) < 1$ (see the proof of

Theorem 2).

The possible multiplicity of equilibria introduces difficulties into the analysis. In case of multiplicity, it is delicate to operate a selection, hence to assess the impact of a pension system. Also, security prices vary as contribution rate varies, and in case of multiplicity, no continuous selection may exist. To avoid these shortcomings, I assume that for each state e, each contribution vector $\boldsymbol{\tau}$, equilibrium price vector $p_{\boldsymbol{\tau}}(e)$ is unique. Let $M^a_{\boldsymbol{\tau}}$ denote the matrix defined by (9) in which the marginal rates of substitution are computed at the unique equilibrium given the K securities payoffs vectors $a = (a_k)$. The agreement conditions (8) required for the sustainability of $\boldsymbol{\tau}$ write as: $\boldsymbol{\tau} = M^a_{\boldsymbol{\tau}}\boldsymbol{\tau}$.

Next theorem uses the matrices at an autarky equilibrium $M^a_{\mathbf{0}}$. Typically, financial securities are exchanged at an equilibrium so that hence these matrices depend on the payoffs a. However, under specific assumptions such as identical homothetic preferences, financial securities are useless at autarky and all autarky equilibria coincide.[12]

Theorem 2 *Consider a stationary economy in which young individuals have access to a set of one-period lived securities with payoffs a. Assume the securities equilibrium to be unique in each state e for each $\boldsymbol{\tau}$, $\boldsymbol{\tau} < 1$. Let $M^a_{\mathbf{0}}$ be the matrix of weighted marginal rates of substitution at the autarky equilibrium.*

If the largest eigenvalue of $M^a_{\mathbf{0}}$ is above 1, then there is a sustainable rule.

If the autarky equilibrium when the securities are traded coincides with that without securities, $M^a_{\mathbf{0}} = M^0_{\mathbf{0}}$, then the converse is true, i.e. a sustainable rule exists only if the largest eigenvalue of $M^a_{\mathbf{0}}$ is above 1.

That the eigenvalue condition is sufficient for a sustainable rule to exist follows from arguments similar to the one used without financial markets. More precisely, an eigenvalue larger than 1 for the matrix M_0 ensures that there is a positive rule that makes decisive voters better off over autarky taking the prevailing prices as given. The existence of a positive sustainable rule follows, relying on a fixed point argument and using that securities prices are continuous in the tax rule (thanks to their uniqueness). The eigenvalue condition however is no longer necessary due to the impact of a payg system on equilibrium securities prices. Even though a decisive voter is surely better off at the prevailing securities prices by contributing at a sustainable rule than by choosing a null rate, it cannot be asserted that he would be better off at the prices prevailing at autarky. This explains why sustainability does not require the eigenvalue condition to be satisfied at autarky if price effects are strong enough.[13]

Short-term complete markets Let us illustrate the results when markets are short-term complete, that is when all opportunities of exchanges between contemporaries are available. There are enough financial securities so that

[12] At autarky, given a set of securities, individuals face the same constraints up to a multiplicative scalar due to their different productivities. Hence, under identical homothetic preferences, individuals' consumption and saving decisions are proportional between each other: financial securities are not traded.

[13] Adapting the argument of Theorem 1, the eigenvalue condition is necessary when the decisive voter is better off at a sustainable rule than at autarky at the autarky equilibrium prices.

any consumption plan contingent on next state can be reached through an appropriate portfolio. Given $\boldsymbol{\tau}$, the securities equilibrium in a state determines a set of contingent prices to which the marginal rates of substitution for all individuals are equalized : $q^i(e_+|e) = q_{\boldsymbol{\tau}}(e_+|e)$ for each i where $q_{\boldsymbol{\tau}}(e_+|e)$ is the equilibrium contingent price in state e for one unit of the good in state e_+. A direct argument for obtaining the agreement conditions is simple and instructive. Thanks to complete markets, individuals utility levels are increasing with their lifetime income defined as the value of all incomes evaluated at the contingent prices.[14] The lifetime income of a $\theta-$individual is

$$\theta w(e) + \theta[-\tau(e)w(e) + \mu(\theta)\sum_{e_+} q_{\boldsymbol{\tau}}(e_+|e)(\gamma\tau w)(e_+)]. \quad (15)$$

equal to the sum of the wage and the net value derived from the pension system (the value of the future benefits less the contribution). By choosing the scale λ, the decisive voter multiplies by λ the net value he derives from the pension system. The agreement condition in state e follows[15]

$$\tau(e)w(e) = \mu^d(e)\sum_{e_+} q_{\boldsymbol{\tau}}(e_+|e)(\gamma\tau w)(e_+). \quad (16)$$

With a bismarckian model with redistributive factors equal to 1, this condition provides the links between a sustainable rule and an equilibrium with a money asset. Let quantity of money grow at the same rate as the population. Condition (16) is the equilibrium condition for money where the price of the money asset in state e is $\tau(e)w(e)$. Clearly, this interpretation does not extend to redistributive systems. ∎

A comparison of theorems 1 and 2 helps us to assess the impact of financial markets on sustainability. Sustainability is possible without financial markets and fails with complete ones only if the largest eigenvalue of $M^a_{\mathbf{0}}$ is above 1 without markets and less than 1 with complete ones. But, there is no a priori relationships between the two eigenvalues. Nest example illustrates this point.

Example 2 Consider an economy as in section 3.2, with inter-temporal independent shocks and recursive preferences. From Proposition 1, without financial markets, the eigenvalue at autarky equals $\mu^d E[\frac{\widetilde{\gamma}}{\widetilde{\rho}}\frac{v'(\widetilde{\rho})\widetilde{\rho}}{E[v'(\widetilde{\rho})\widetilde{\rho}]}]$ in which v represents the decisive voter's attitudes towards risk. With complete markets, the autarky equilibrium depends on the preferences of all individuals. We consider two cases.

(1) Individuals share the same preferences. No trade occurs at autarky. As a result, the sustainability condition is identical with complete financial markets or without any. To fix the idea take u with constant elasticity β and v with constant relative risk aversion coefficient α: $u(c) = \frac{1}{1-\beta}c^{1-\beta}$ and $v(c) = \frac{1}{1-\alpha}c^{1-\alpha}$. Assume also that

[14]More precisely, the budget equations (14) are equivalent to the single lifetime budget constraint that equalizes the value of consumption levels and net investment, $c_y + s + \sum_{e_+} q_{\boldsymbol{\tau}}(e_+|e)(c_o(e_+) - \rho(e_+)s)$, to the lifetime income as defined by (15).

[15] Observe that the condition implies that if the decisive voter is chosen by majority, the median voter is a voter with median income.

$(ln\widetilde{\rho}, ln\widetilde{\gamma})$ is a gaussian vector. Easy computation [16] yields

$$E[\frac{\widetilde{\gamma}}{\widetilde{\rho}}\frac{v'(\widetilde{\rho})\widetilde{\rho}}{E[v'(\widetilde{\rho})\widetilde{\rho}]}] = \frac{E[\widetilde{\gamma}]}{E[\widetilde{\rho}]}\exp\alpha[\text{ var}(ln\widetilde{\rho}) - \text{ cov}(ln\widetilde{\rho}, ln\widetilde{\gamma})]. \quad (17)$$

The exponential term reflects the impact of risk aversion. The condition in 2 of Proposition 1 -$\frac{E[\widetilde{\gamma}|\rho]}{\rho}$ is non-increasing- is equivalent to the positivity of argument in the exponential. *Risk aversion may substantially affect a criterion based on the simple comparison between expected growth and expected return*. If $\widetilde{\rho}$ and $\widetilde{\gamma}$ are independent for instance, the correcting term, $\exp\alpha[\text{ var}(ln\widetilde{\rho})]$, may be far from negligible. As expected, the larger the risk aversion coefficient and the more risky the investment return, the more chance a sustainable rule to exist.

(2) There is a risk neutral individual in the economy. Complete markets have a large effect because they allow risk averse individuals to be fully insured at a fair price. In particular, the decisive voter gets a constant consumption level whatever the realized shock. Easy computation gives that the eigenvalue is equal to $\mu^d E[\widetilde{\gamma}]/E[\widetilde{\rho}]$. In the more likely case in which population growth and technology return are not too much positively correlated, this eigenvalue is smaller than the corresponding one without markets. In that case, as explained in the previous section, in the absence of financial markets, the risk profile of the contribution rule provides an insurance against the risky investment return, and this is why a sustainable rule may exist even though $\mu^d E[\widetilde{\gamma}] < E[\widetilde{\rho}]$. If insurance against investment risk is provided for free by contemporary risk neutral individuals, then only expected returns matter. Assuming a risk neutral market is however far from a natural assumption, given the risk premium observed on stock markets.

3.4 Welfare properties

As already said, a sustainable contribution rule may not lead to an *interim* Pareto improvement over autarky due to several distortions, the redistribution performed by the system, the decision process, which aggregates individuals preferences in a specific way, and the price effects of the system that are not taken into account by the decisive voter. The impact of a sustainable system on welfare may however be partially assessed under some conditions.

Proposition 2 *Assume identical homothetic preferences and short-term securities. At a sustainable rule, in each state, each individual with wage equal or lower than the decisive voter's wage is better off than at autarky. In particular, with complete short term markets and under majority rule, the decisive voter is a voter with median income, and at least half of the voters are made better off at a sustainable rule.*

[16] Letting $X = \gamma\rho^{-\alpha}$ and $Y = \rho^{1-\alpha}$, one has to compute $E\widetilde{X}/E\widetilde{Y}$. Since each variable is log normal, this ratio is given by $\exp[ElnX̃ - Eln\widetilde{Y} + 0.5\text{ var}(ln\widetilde{X}) - 0.5\text{ var}(ln\widetilde{Y})]$, and (17) follows. To determine when $\frac{E[\widetilde{\gamma}|\rho]}{\rho}$ is non-increasing note that the law of $(ln\widetilde{\gamma})$ conditional on $ln(\rho)$ is normal. Thus $E[\widetilde{\gamma}|\rho] = \exp[Eln[\widetilde{\gamma}|\rho] + 0.5\text{ var}(ln\widetilde{\gamma}|\rho]$ We have $E[ln\widetilde{\gamma}|\rho] = E[ln\widetilde{\gamma}|ln\rho] = bln\rho + cste$ where $b = \text{ cov}(ln\widetilde{\rho}, ln\widetilde{\gamma})/\text{ var}(ln\widetilde{\rho})$ and the variance term $\text{var}(ln\widetilde{\gamma}|\rho]$ is independent of ρ. This gives $E[\widetilde{\gamma}|\rho]/\rho = K\rho^{b-1}$, which is non increasing when $[\text{ var}(ln\widetilde{\rho}) - \text{ cov}(ln\widetilde{\rho}, ln\widetilde{\gamma})] > 0$, i.e. the argument in the exponential is positive.

Proposition 2 points out the impact of the redistributive factors. In particular, under a Bismarckian system, all individuals are better off at a sustainable system (since redistributive factors are equal to 1) but a sustainable rule is less likely to exist than with a redistributive system. More generally, the more redistributive the system is to decisive voters, the more likely it is that a sustainable rule exists, but it may be inefficient and wealthy people may be made worse off.

The proof of Proposition 2 goes as follows. Let preferences be homothetic and identical. At autarky, all equilibria with or without securities coincide. As a result, a decisive voter is surely better off at a sustainable rule than at the autarky equilibrium without securities (since he can choose a null scale level and not to trade the securities). The introduction of a payg system has a differential impact on individuals because of the redistribution. This impact is however the same for a θ-individual who has the same redistributive factor than that of the decisive voter. Hence this $\theta-$individual is made better off over autarky. If he benefits more from redistribution than the decisive voter, that is if $\mu(\theta) > \mu(\theta^d)$, he can only be made better off. The last property relies on the fact that under short term complete markets, a median voter is a voter with median income (see footnote 15).

With intra-generational heterogeneity and no complete financial markets, there is little hope to assess the welfare impact of a sustainable payg system. Under restrictive conditions however, such as identical homothetic preferences and no financial markets, a sustainable rule makes every worker whose wage is lower than the decisive individual better off.

4 Extensions

This section presents some extensions of the basic model. To simplify the presentation, we assume that there are no financial securities.

Mixed systems So far the analysis has been restricted to fully unfunded systems. A natural question is whether a more flexible pension system, partly funded, would significantly change the political support to the system. The funded part is not a perfect substitute to direct investment into the technology. Individuals might be unable to undo the funded part because short selling the technology is impossible (the absence of financial markets). Thus they may end up with a (compulsory) investment through the funded part of the system that is too large with respect to their needs. Furthermore, even funded, a system offers returns that differ from those of direct investment owing to redistribution.

A system is said to be $\alpha-$mixed if a fixed proportion α of the contributions is invested into the technology at each period and the return to these investments are redistributed at the subsequent period. Form budget balance, the average pension benefits per head are equal to

$$\pi_t = (1-\alpha)\gamma_t \tau_t w_t + \alpha \tau_{t-1} w_{t-1} \rho_t. \tag{18}$$

The first term on the right hand side corresponds to the part of the contributions directly paid to the retirees, namely the unfunded part, and the second term to the returns on the amount invested the previous period.

Proposition 3 *Consider a stationary economy without financial securities. If a fully unfunded system exists, then a sustainable $\alpha-$mixed system exists as well.*

The intuition for this result is the following one. Recall that the decisive voter is assumed not to be harmed by redistribution, that is the redistributive factor is not less than 1 at the voter's characteristic. Thus, from the voter's point of view, the return on the funded part of a mixed system is at least as large as the one obtained from direct investment, which enhances his support to the system.

Multi-period lived individuals The analysis can easily accommodate a multi-period life model in the absence of financial markets. To simplify, let individuals live three periods, work when young and middle-aged. The decisive voter is a "young" individual.[17] Consider a sequence of states (e, e_+, e_{++}). The ratio workers to retirees in state e_{++} is equal to $\gamma_2(e_+, e_{++}) = \gamma(e_+)(1+\gamma(e_{++}))$. Let $q^d_{1\boldsymbol{\tau}}(e, e_+)$ (resp. $q^d_{2\boldsymbol{\tau}}(e, e_+, e_{++})$) denote the marginal rate of substitution between consumption when young in state e and consumption when middle-aged if state e_+ realizes (resp. at retirement if states e_+, e_{++} realize). Since the indirect utility of a decisive voter is concave in the scale level, the sustainability condition of the payg system $\boldsymbol{\tau} > 0$ requires that in each state e

$$\tau(e)w(e) + \sum_{e_+} q^d_{1\boldsymbol{\tau}}(e, e_+)\tau(e_+)w(e_+) = \mu^d(e) \sum_{e_{++}} q^d_{2\boldsymbol{\tau}}(e, e_+, e_{++})\gamma_2(e_+, e_{++})\tau(e_{++})w(e_{++})].$$

Dividing by $w(e)$ in each state, it can be put in matrix form as: $[I + M_{1\boldsymbol{\tau}}]\boldsymbol{\tau} = M_{2\boldsymbol{\tau}}\boldsymbol{\tau}$, where I is the identity matrix, and $M_{1\boldsymbol{\tau}}$ are $M_{2\boldsymbol{\tau}}$ are positive matrices corresponding respectively to the second period of contribution and the retirement period.

An adaptation of the previous proofs yields that a sustainable payg system exists if (and only if) there is a Pareto improving direction for the decisive voters over autarky, that is a positive $\boldsymbol{\tau}_0$ that satisfies $(I + M_{1\boldsymbol{\tau}_0})\boldsymbol{\tau}_0 < M_{2\boldsymbol{\tau}_0}\boldsymbol{\tau}_0$. This property however cannot be stated in terms of an eigenvalue because the inverse of matrix $I + M_{1\boldsymbol{\tau}_0}$ (which is likely to exist) is not necessarily positive.

Capital Krueger and Kubler (2006) investigates whether (non redistributive) an unfunded system may yield a Pareto improvement in a calibrated model with multi-period lived agents. At the autarky situation the economy

[17] One could also assume a middle-aged voter. If such a voter takes his past contributions as fixed or "sunk", the return to social security from his point of view is increased. As a result, the support for a high tax level is increased, as shown in Browning (1975).

is dynamically efficient (in a sense to be made precise owing to incomplete markets; in our simple economy, dynamical efficiency holds if the return on capital is not always smaller than the population growth rate). They find that a large improvement is possible when the capital return is exogenous but that the gains disappear once the crowding effect of capital is taken into account.

Introducing an endogenous investment return through a production function is not difficult in a two-period lived agents when investment lasts for one period. The productivity shock ρ_t bears on the constant returns to scale technology which combines capital accumulated in the previous period with current labor as described by a neoclassical production function. The marginal productivity on capital is given by $r_t = f'_k(k_t, \rho_t)$ in which k_t is the stock of capital per head of the new generation ($k_t = s_{t-1}/\gamma_t$ where s_{t-1} is the investment of physical capital per head of the generation $t-1$). Under appropriate assumptions, a sustainable payg exists if the decisive voter can be made better off at autarky. (From an empirical perspective, since in most countries a payg system is in place, the autarky values cannot be directly estimated and have to be predicted. Such a prediction is common in the models that aim at assessing the impact of extinguishing a payg system.) A sustainable contribution rule is affected by capital through the impact on the return to capital. Crowding effects are likely to decrease the rates at sustainable rule. The reason is that the return on capital is likely to increase with $\boldsymbol{\tau}$, which decreases the marginal rates of substitution hence the entries in matrix $M_{\boldsymbol{\tau}}$, hence the solution to $\boldsymbol{\tau} = M_{\boldsymbol{\tau}}\boldsymbol{\tau}$.

Introducing capital into our analysis is worth another study. With capital that lasts for more than one period, the environment faced by an individual at birth depends not only on the shocks at the current date but also on the previous ones through inherited capital. One way to keep a Markovian framework is to include the accumulated capital stock, along with the exogenous shocks in the state variable. The state takes an infinite number of values, and we could rely on the technics used in Demange and Laroque (2000) in order to get some necessary conditions. However even in a model with a representative agent, optimality fails because an individual does not properly take into account the influence of a payg system on capital accumulation and on the distribution of states in the future (put differently, the distribution the state, which is endogenous, may not be optimal). Furthermore some existence issues may arise, as shown by Spear (1985) for instance.

5 Sustainability with rolled over debt

This section analyzes sustainability when governmental debt is issued and rolled over. Governmental debt performs intergenerational transfers as a payg system but these transfers depend on the price of debt, which is endogenous. To simplify the presentation, we suppose that no financial instrument is available in addition to debt. The results however carry over to a more general set up.

The government is assumed to issue at each date bonds that mature at the subsequent date. At t, a unit of bond promises a (possibly random) revenue next period denoted by $\widetilde{a}_{t+1}$. The total amount of debt is normalized at each date so that the number of shares is equal to the size of the young generation. Equivalently the number of units of bonds per young is equal to one. Furthermore debt is rolled over, without using any tax instrument. Therefore, at any point in time, the payments to bondholders are covered by the newly issued debt. This gives the balance equation at time t : $n_{t-1}a_{t-1} = n_t Q_t$, in which n_t is the current size of the population, and Q_t the price of one unit of bond. Dividing by the population size of generation $t-1$ yields $a_t = \gamma_t Q_t$. This equality says that *without taxation the payoff promised by debt is constrained to be equal to the future price of debt multiplied by population growth.* In a stochastic economy, there are few chances for this payoff to be risk-less since population growth may be uncertain and the price of debt, which is determined by equilibrium forces, is likely to vary with the state.

Both the contribution rate and the price of debt are time invariant functions of the current state e, described respectively by $\boldsymbol{\tau} = (\tau(e))$ and $\boldsymbol{Q} = (Q(e))$. We look for an equilibrium in which expectations are correct: young agents form correct expectations on the distribution of the future state $\widetilde{e}_+$, that is on wage, population growth and investment return, conditional on the current state, and furthermore they infer the distribution of the endogenous variables, debt price and contribution rate. Given the price function $\boldsymbol{Q}$ for debt, b units of debt yield $b\gamma(e_+)q(e_+)$ in state e_+ to its owner. Thus, under correct expectations, the present and future budget constraints of a $\theta-$worker born in state e are given by :

$$\begin{cases} c_y + s + bQ(e) = (1-\tau(e))\theta w, \;\; s \geq 0, \, b \geq \underline{b} \\ \widetilde{c}_o = s\rho(e_+) + b\gamma(e_+)Q(e_+) + \theta\mu(\theta)(\gamma\tau w)(e_+) \text{ each } e_+ \text{ in } E \end{cases} \tag{19}$$

in which $\underline{b}$ represents possible short sale constraints. Without such constraints, $\underline{b}$ is set to $-\infty$. This yields the following definition.

Definition 2 *An equilibrium with rolled over debt and sustainable payg is defined by debt prices $\boldsymbol{Q} = (Q(e))$, contribution rates $\boldsymbol{\tau} = (\tau(e))$ both nonnegative, and consumption plans $c_y^i(e), (c_o^i(e, e_+))_{e_+\in E}$ for each i, each e in E, satisfying the following conditions in each state e :*

1. *for each i, $c_y^i(e), (c_o^i(e, e_+))_{e_+\in E}$ maximizes $E[u^i(c_y, \widetilde{c}_o)|e]$ over the constraints (19) for $\theta = \theta^i$*
2. *the bond market clears: $\sum_i b^i(e) = 1$*
3. *the decisive voter agrees on $\tau(e)$, his expectations on average next period pensions being given by $(\gamma\tau w)(\widetilde{e}_+)$ conditional on e.*

Condition *1* states standard rational behavior under correct expectations. Condition *2* is the market clearing condition for debt since the total number of shares is equal to the population size. By Walras law, the market for the good clears as well. Conditions *1* and *2* together say that given contribution rates $\boldsymbol{\tau}$ a rational expectations equilibrium with debt obtains. Condition *3* states the decisive voters' agreement, still under correct expectations.

A sustainable contribution rule in an economy without financial securities, as considered in section 3, gives rise to an equilibrium in which debt has no value (i.e. $Q = 0$). We are interested here in situations in which both contribution rates and debt prices are positive (if any). If debt has a value, the transfer from young to old agents is endogenous through the (non zero) price of debt. More precisely, on average per young, the amount $(\tau w + Q)(e)$ of consumption good is transferred in state e.

Theorem 3 *Assume that debt is rolled over and that there are no short sales constraint on debt ($\underline{b} = -\infty$). At an equilibrium with a positive sustainable rule and positive debt prices :*

1. *the decisive voter is not subsidized, i.e. $\mu^d(e) = 1$ in each state,*
2. *the returns of the payg and debt are identical for any non subsidized individual: for some positive k, $Q(e) = k\boldsymbol{\tau}(e)w(e)$ in each state e.*

Thus, without short sales constraints, the returns of the infinite-lived securities, debt and payg, must be equalized on average at an equilibrium with sustainable payg. In contrast with the case of complete markets considered in next section, the availability of governmental debt without constraints affects drastically the results since redistribution is severely constrained. The result is trivial in a deterministic set up: from the point of view of the decisive agent, payg and debt offer both risk-free returns, respectively γ and $\mu\gamma$ which should coincide to avoid arbitrage opportunities. The result is much more surprising in a stochastic set up in which there is a priori room for two securities with different returns. That debt and payg returns are equalized on average and from the point of view of the decisive voter does not however preclude redistribution, even if quite limited. Workers with an income lower than that of the decisive voter may be subsidized, and those with a larger income be taxed.

Concluding remarks This paper has analyzed the sustainability of an unfunded social security system in a risky environment. Our results merge the economic and political analysis of such systems. Important features that have been extensively studied elsewhere have been assumed away. Some would not dramatically change the qualitative results. For example, we have not considered individual risks, especially the risk of living old. Accounting for individuals' longevity risk tends to favor a payg system. A payg system provides an annuity to retirees, thereby insuring them against this risk, and by making insurance compulsory, avoids the usual problems encountered in markets with asymmetric information. As documented by various studies, the premium associated to longevity risk is roughly 5% (see Brown, Mitchell, and Poterba 2001). To take into account of this premium in our model, we could add an extra return on a payg, which would clearly favor sustainability. On the other hand, because of inelastic labor, the analysis neglects the standard distortionary effects of taxation on labor supply.

Whatever restrictions, we think that the main results are quite robust. Macro-economic risks modify substantially the analysis of unfunded systems both from the political sustainability and welfare point of views. Governmental debt limits seriously the possibility of redistribution in the absence of short sales constraints. Finally, the design of the system, in particular allowing contributions rates to be contingent on the state of the economy, plays an essential role in promoting sustainability.

6 Proofs

In what follows, $x > y$ (resp. $x \geq y$) for two vectors x and y denotes a component wise strict (resp. large) inequality.

Proof of Theorem 1

Sufficient condition[18] Define the function $W(\boldsymbol{\tau})$ for $0 \leq \boldsymbol{\tau} < 1$.

$$W(\boldsymbol{\tau})(e) = (M_{\boldsymbol{\tau}}\boldsymbol{\tau})(e) = \mu^d(e) \sum_{e_+} q^d_{\boldsymbol{\tau}}(e_+|e)(\gamma \tau w)(e_+)/w(e), \tag{20}$$

which gives the marginal benefit in terms of the present good derived by the decisive voter in state e from rule $\boldsymbol{\tau}$. Function W is continuous.

The sustainability condition writes as $W(\boldsymbol{\tau}) = M_{\boldsymbol{\tau}}\boldsymbol{\tau} = \boldsymbol{\tau}$. Thus a positive fixed point of W is a sustainable rate. In order to use a fixed point theorem we first need to restrict W to strictly positive rates since the null vector is a fixed point of W. We make use here of the eigenvalue assumption to show that there exists a contribution vector $\boldsymbol{\tau}_0$ such that $W(\boldsymbol{\tau}_0) > \boldsymbol{\tau}_0$. To see this note that there is $\boldsymbol{\tau}_1 > 0$ such that $M_{\mathbf{0}}\boldsymbol{\tau}_1 > \boldsymbol{\tau}_1$. By continuity, for ϵ positive small enough $M_{\epsilon\boldsymbol{\tau}_1}\boldsymbol{\tau}_1 > \boldsymbol{\tau}_1$. Multiplying by ϵ gives $W(\epsilon\boldsymbol{\tau}_1) > \epsilon\boldsymbol{\tau}_1$ so that $\boldsymbol{\tau}_0 = \epsilon\boldsymbol{\tau}_1$ satisfies the desired inequality.

Now let $T = \{\boldsymbol{\tau} = (\tau(e)), \tau_0(e) \leq \tau(e) \leq 1\}$. The function W is defined by (20) for contribution rates smaller than 1. We show that it can be extended by continuity to any $\boldsymbol{\tau}$ in T by setting $W(\boldsymbol{\tau})(e) = 0$ if $\boldsymbol{\tau}(e) = 1$. To see this, let a sequence $(\boldsymbol{\tau}_n)$, $\boldsymbol{\tau}_n < 1$ converging to some $\boldsymbol{\tau}$ with $\boldsymbol{\tau}(e) = 1$ in state e. In that state, the decisive voter's endowment of the good when young tends to zero hence the consumption level as well (since investment is nonnegative and there are no borrowing possibilities through financial securities). When old instead, consumption levels equal endowments which are positive in any subsequent state e_+ since $0 < \tau_0(e_+) \leq \tau(e_+)$. This implies that all marginal rates of substitution converge to zero, hence also $W(\boldsymbol{\tau}_n)(e)$. It suffices then to apply the following lemma. ∎

Lemma 1 *Let* $T = \{\boldsymbol{\tau} = (\tau(e)), \tau_0(e) \leq \tau(e) \leq 1\}$, *for a vector* $\boldsymbol{\tau}_0$, $0 < \boldsymbol{\tau}_0 < 1$. *Let* W *be a continuous*

[18] We give a proof that is somewhat more complicated than the one sketched in the text because it extends to the case with financial markets.

function on T that satisfies the boundary conditions: $W(\boldsymbol{\tau}_0) > \boldsymbol{\tau}_0$ and $W(\boldsymbol{\tau})(e) = 0$ for any $\boldsymbol{\tau}$ in T for which $\boldsymbol{\tau}(e) = 1$. Then W has a fixed point.

Proof of Lemma 1 Define correspondence F from T to itself as follows:

- (a) if inequality $W(\boldsymbol{\tau})(e) < \boldsymbol{\tau}(e)$ holds in at least one state then $F(\boldsymbol{\tau}) = \boldsymbol{\tau}_0$,

- (b) otherwise, that is if $W(\boldsymbol{\tau}) \geq \boldsymbol{\tau}$, define

$$F(\boldsymbol{\tau}) = \{t \in T \text{ such that } t(e) = \min(W(\boldsymbol{\tau})(e), 1) \text{ for any } e \text{ with } W(\boldsymbol{\tau})(e) > \boldsymbol{\tau}(e)\}. \tag{21}$$

Thus, in case (b), $t(e)$ can take any value in $[\boldsymbol{\tau}_0(e), 1]$ in a state e for which $W(\boldsymbol{\tau})(e) = \boldsymbol{\tau}(e)$.

An application of Kakutani's theorem will prove that F has a fixed point. Next we show that a fixed point of F is a fixed point of W.

The set T is compact and convex and the correspondence F is convex-valued from T to itself. It is also upper hemicontinuous, thanks to the continuity of the function $W(\boldsymbol{\tau})$. Let $(\boldsymbol{\tau}_n)$ be a sequence converging to $\boldsymbol{\tau}$ and (t_n) a sequence with t_n in $F(\boldsymbol{\tau}_n)$. We show that any limit point t of the sequence (t_n) is in $F(\boldsymbol{\tau})$.

Assume case (a) for $\boldsymbol{\tau}$. $F(\boldsymbol{\tau}) = \boldsymbol{\tau}_0$ and the strict inequality $W(\boldsymbol{\tau})(e) < \boldsymbol{\tau}(e)$ is met in a state e. By continuity of W, the same inequality is met for $\boldsymbol{\tau}_n$ with n sufficiently large hence $t_n = t = \boldsymbol{\tau}_0$. Assume case (b). In a state e in which $W(\boldsymbol{\tau})(e) = \boldsymbol{\tau}(e)$ there is nothing to prove since $t(e)$ is unrestricted. In a state in which $W(\boldsymbol{\tau})(e) > \boldsymbol{\tau}(e)$, the same inequality holds for $\boldsymbol{\tau}_n$ with n sufficiently large. Thus $t_n(e) = \min(W(\boldsymbol{\tau}_n)(e), 1)$ converges to $t(e) = \min(W(\boldsymbol{\tau})(e), 1)$. This proves that any limit point of (t_n) belongs to $F(\boldsymbol{\tau})$. Upper-hemicontinuity follows. By Kakutani's theorem, F has a fixed point, $\boldsymbol{\tau}^* \in F(\boldsymbol{\tau}^*)$.

We show that $\boldsymbol{\tau}^*$ is a fixed point of W. Case (a) does not hold for $\boldsymbol{\tau}^*$. By contradiction, under (a) $F(\boldsymbol{\tau}^*)$ is the singleton $\boldsymbol{\tau}_0$, hence $\boldsymbol{\tau}^*$ is equal to $\boldsymbol{\tau}_0$. The assumption $W(\boldsymbol{\tau}_0) > \boldsymbol{\tau}_0$ then contradicts the conditions (a). Thus case (b) holds and we have $W(\boldsymbol{\tau}^*) \geq \boldsymbol{\tau}^*$. Assume by contradiction that there is a state e for which $W(\boldsymbol{\tau}^*)(e) > \boldsymbol{\tau}^*(e)$. Since $F(\boldsymbol{\tau}^*)$ is given by (21) and $\boldsymbol{\tau}^* \in F(\boldsymbol{\tau}^*)$ we have $\boldsymbol{\tau}^*(e) = \min(W(\boldsymbol{\tau}^*)(e), 1)$. Using $W(\boldsymbol{\tau}^*)(e) > \boldsymbol{\tau}^*(e)$, this gives $\boldsymbol{\tau}^*(e) = \min(\boldsymbol{\tau}^*(e), 1)$, and finally $\boldsymbol{\tau}^*(e) = 1$. But by assumption, $\boldsymbol{\tau}^*(e) = 1$ implies $W(\boldsymbol{\tau}^*)(e) = 0$, which contradicts $W(\boldsymbol{\tau}^*)(e) > \boldsymbol{\tau}^*(e) = 1$. This proves that $W(\boldsymbol{\tau}^*) = \boldsymbol{\tau}^*$. ∎

Necessary condition We need to prove that the existence of a sustainable rule implies the eigenvalue condition. This follows from concavity arguments on the indirect utility achieved by an individual facing contribution rule $\boldsymbol{\tau}$. To simplify notation, drop the individual's index and denote by $V(e, \boldsymbol{\tau})$ the expected utility of an individual born in state e under the rule $\boldsymbol{\tau}$. The proof uses an auxiliary result, which was introduced in Demange and Laroque (1999).

Lemma 2 *The functions $V(e, \boldsymbol{\tau})$ are concave in $\boldsymbol{\tau} = (\tau(e))_{e \in E}$ for all e. Furthermore :*

$$\frac{1}{E(u'_y|e)\theta w(e)} \frac{\partial V(e, \boldsymbol{\tau})}{\partial \tau(e')} = [M_{\boldsymbol{\tau}}(e, e') - \mathbb{1}_{e=e'}]$$

Proof By definition, $V(e,\boldsymbol{\tau})$ is equal to the maximum of $Eu(c_y,c_o)|e$ over the budget constraints. Since the constraints are linear in $\boldsymbol{\tau}$, V is concave in $\boldsymbol{\tau}$. By the envelope theorem one immediately gets :

$$\frac{\partial V(e,\boldsymbol{\tau})}{\partial \tau(e)} = \theta[-E(u'_y|e)w(e) + u'_o[e,e]\gamma(e)\mu(e)w(e)\text{Pr}(e|e)]$$

and for $e_+ \neq e$:

$$\frac{\partial V(e,\boldsymbol{\tau})}{\partial \tau(e_+)} = \theta u'_o[e,e_+]\gamma(e_+)\mu(e)w(e_+)\text{Pr}(e_+|e).$$

The result follows from the definition of the matrix $M_{\boldsymbol{\tau}}$. ∎

Now, consider a sustainable $\boldsymbol{\tau}$. The decisive voter is better off whatever state, so $V(e,\boldsymbol{\tau}) - V(e,0) > 0$. By the concavity of V, we have :

$$\sum_{e'} \frac{\partial V(e,\mathbf{0})}{\partial \tau(e')}\boldsymbol{\tau}(e') \geq V(e,\boldsymbol{\tau}) - V(e,\mathbf{0}) > 0.$$

Dividing inequality in state e by $E(u'_y|e)\theta w(e)$ and using Lemma 2 at **0** yields $(M_{\mathbf{0}} - I)\boldsymbol{\tau} > 0$. By Frobenious theorem (see e.g. in Debreu Herstein [1953, p.601]), the maximal eigenvalue of $M_{\mathbf{0}}$ is strictly larger than 1. ∎

Proof of Proposition 1 I first show that the marginal rates of substitution are given by

$$q(e_+|e) = \frac{v'(\rho(e_+))\text{Pr}(e_+)}{E[v'(\rho(e_+))\rho(e_+)]} \tag{22}$$

Recall that the marginal rates of substitution are given by (6) : $q(e_+|e) = \delta\frac{u'(\hat{c}_o(e))}{u'(c_y(e))}\frac{v'(c_o(e,e_+))\text{Pr}(e_+|e)}{v'(\hat{c}_o(e))}$. At autarky, saving are invested in the risky technology only. Hence $c_o(e,e_+) = \rho(e_+)s(e)$, and investment must be positive to get some consumption when old. This yields the first order condition $\sum_{e_+} q(e_+|e)\rho(e_+) = 1$, and plugging the values of q gives

$$1 = \delta\frac{u'(\hat{c}_o(e))}{u'(c_y(e))}\frac{E[v'(c_o(e,\widetilde{e}_+))\rho(\widetilde{e}_+)|e]}{v'(\hat{c}_o(e))}.$$

Using the obtained expression of $\delta\frac{u'(\hat{c}_o(e))}{u'(c_y(e))}$ we find

$$q(e_+|e) = \frac{v'(c_o(e,e_+))\text{Pr}(e_+|e)}{E[v'(c_o(e,\widetilde{e}_+)\rho(\widetilde{e}_+)|e]}.$$

Now using that v is homothetic, states are independent and $c_o(e,e_+) = \rho(e_+)s(e)$ gives (22). Thus the general term of matrix $M_{\mathbf{0}}$ is

$$M_{\mathbf{0}}(e,e_+) = \mu^d\frac{w(e_+)}{w(e)}\frac{v'(\rho(e_+))}{E[v'(\widetilde{\rho})\widetilde{\rho}]}\gamma(e_+)\text{Pr}(e_+).$$

All rows of $M_{\mathbf{0}}$ are proportional among each other. Hence the eigenvalues of $M_{\mathbf{0}}$ are all null but one, which is positive equal to the sum of the diagonal terms. This is the maximal eigenvalue, equal to $\mu^d E[v'(\widetilde{\rho})\widetilde{\gamma}]/E[v'(\widetilde{\rho})\widetilde{\rho}]$, which can also be written as $\mu^d E[\frac{\widetilde{\gamma}}{\widetilde{\rho}}\frac{v'(\widetilde{\rho})\widetilde{\rho}}{E[v'(\widetilde{\rho})\widetilde{\rho}]}]$. This proves (13).

Consider now the impact of increasing risk aversion on the maximal eigenvalue. To simplify notation, take $\mu^d = 1$. The eigenvalue writes as

$$E[\frac{\widetilde{\gamma}}{\widetilde{\rho}} g_v(\widetilde{\rho})] \text{ where } g_v(\rho) = \frac{v'(\rho)\rho}{E[v'(\widetilde{\rho})\widetilde{\rho}]}.$$

From the property of conditional expectation : $E[\frac{\widetilde{\gamma}}{\widetilde{\rho}} g_v(\widetilde{\rho})] = E[\frac{E[\widetilde{\gamma}|\widetilde{\rho}]}{\widetilde{\rho}} g_v(\widetilde{\rho})]$. Observe that function g_v is a density. Denoting $\phi(\rho) = E\frac{[\widetilde{\gamma}|\rho]}{\rho}$, the eigenvalue is therefore the expectation of ϕ under the distribution G_v whose density is g_v. By assumption the function ϕ is non-increasing. Hence increasing risk aversion increases the eigenvalue if it decreases the distribution $G_.$ in the sense of first-order stochastic dominance. Let function w be more concave than v: $w = f(v)$ for a concave f. We show that the difference $[g_w(\rho) - g_v(\rho)]$ is positive for $\rho < \rho^*$ and negative for $\rho > \rho^*$. This is known to imply that the distribution G_v first-order dominates distribution G_w (or simply observe that $E[\phi(\rho) - \phi(\rho^*)][g_w(\rho) - g_v(\rho)] \geq 0$ and use $E[g_w(\rho) - g_v(\rho)] = 0$). We have $g_w(\rho) = \frac{f'(v(\rho))v'(\rho)\rho}{E[f'(v(\rho))v'(\widetilde{\rho})\widetilde{\rho}]}$. By the intermediate values theorem $E[f'(v(\rho))v'(\widetilde{\rho})\widetilde{\rho}] = f'(v(\rho^*))E[v'(\widetilde{\rho})\widetilde{\rho}]$ for some value ρ^* in the support of the distribution of $\widetilde{\rho}$. Thus the difference writes

$$g_w(\rho) - g_v(\rho) = [\frac{f'(v(\rho))}{f'(v(\rho^*))} - 1]g_v(\rho).$$

which is positive for $\rho < \rho^*$ and negative for $\rho > \rho^*$ since $f'(v(\rho))$ is decreasing with ρ. ■

Proof of Theorem 2 For $\boldsymbol{\tau}$ with $0 \leq \boldsymbol{\tau} < 1$, an equilibrium for financial securities exists in each state e. To show this, given state e consider the following economy $\mathcal{E}_{\boldsymbol{\tau}}(e)$. The consumers are the young individuals. There are $1+K$ goods -the good available today and the securities- and a linear technology described by ρ. In this economy, a θ-individual has preferences defined by

$$V_i(c, b, s) = E[u_i(c_y, \rho(e_+)s + \sum_k b_k(e_+)a_k(e_+) + \theta\mu(\theta)(\gamma\tau w)(e_+))|e]$$

and is subject to the budget constraint $c_y + s + \sum_k p_k b_k = (1 - \tau(e))\theta w(e)$. An equilibrium is given by a securities price vector at which markets clear. For $\boldsymbol{\tau}(e) < 1$, each individual's endowment of the first good (when young) is positive and other goods can be obtained through investment. Since the securities payoffs are nonnegative, prices can be assumed to be positive. Also each security payoff is bounded by some multiple of the technology return, giving an upper bound for its price. Thus, standard arguments yield the existence of an equilibrium.

The assumption on the uniqueness of an equilibrium allows us define the function $W(\boldsymbol{\tau})$ on $0 \leq \boldsymbol{\tau} < 1$ by

$$W(\boldsymbol{\tau})(e) = (M^a_{\boldsymbol{\tau}}\boldsymbol{\tau})(e) = \mu^d(e)\sum_{e_+} q^d_{\boldsymbol{\tau}}(e_+|e)\frac{(\gamma\tau w)(e_+)}{w(e)},$$

which is the marginal benefit in terms of the present good derived by the decisive voter in state e from rule $\boldsymbol{\tau}$. Function W is continuous. The agreement condition is satisfied in state e if $W(\boldsymbol{\tau})(e) = \boldsymbol{\tau}(e)$. Thus a positive fixed point of W is a sustainable rule.

The proof now proceeds as in Theorem 1. First, one restricts W to strictly positive rates: Thanks to the eigenvalue assumption and the continuity of W, there exists a contribution vector $\boldsymbol{\tau}_0$ such that $W(\boldsymbol{\tau}_0) > \boldsymbol{\tau}_0$. Second, the function W, which is defined for contribution rates smaller than 1, is extended by continuity to the whole set $T = \{\boldsymbol{\tau} = (\tau(e)), \tau_0(e) \leq \tau(e) \leq 1\}$ by setting $W(\boldsymbol{\tau})(e) = 0$ if $\boldsymbol{\tau}(e) = 1$. To see this, let a sequence $(\boldsymbol{\tau}_n)$, $\boldsymbol{\tau}_n < 1$ converging to $\boldsymbol{\tau}$ with $\boldsymbol{\tau}(e) = 1$ for some e. In the sequence of economies $\mathcal{E}_{\boldsymbol{\tau}_n}(e)$, the aggregate endowment for the good at the initial date tends to zero and is positive in any subsequent state e_+ (since $\tau_0(e_+) \leq \tau(e_+)$). This implies that individuals' current consumption levels converge to zero. Hence marginal rates of substitution $q^d_{\boldsymbol{\tau}}$ converge to zero as well, and also $W(\boldsymbol{\tau}_n)(e)$. Thus, by Lemma 1, W has a positive fixed point, which gives a sustainable payg.

It remains to show that conversely the eigenvalue condition is necessary for a sustainable rule to exist if the autarky equilibrium with or without securities coincide. Note that the decisive voter is surely better off than at the autarky without securities, since he can choose a null scale level and no trade. A similar argument as developed in the necessity part of Theorem 1 can be used. Let $V(e, \boldsymbol{\tau})$ be the expected utility of the decisive voter born in state e if when the rates are given by $\boldsymbol{\tau}$ and when he has access to financial securities a. Lemma 2 still holds thanks to the envelope theorem. Furthermore, if both autarky equilibria coincide, $V(e, \boldsymbol{\tau}) - V(e, \mathbf{0}) \geq 0$ is also true. The result follows. ■

Proof of Proposition 3 Arguments similar to those used for an unfunded system show that the sustainability of $\alpha-$mixed rule is characterized by the equations

$$\tau(e)w(e) = \mu^d(e) \sum_{e_+} q^d_{\boldsymbol{\tau}}(e, e_+)[(1-\alpha)(\gamma\tau w)(e_+) + \alpha\tau(e)w(e)\rho(e_+)] \text{ in each sate } e$$

and that such a rule exists if for some positive $\boldsymbol{\tau}_0$

$$\tau_0(e)w(e) < \mu^d(e) \sum_{e_+} q^d_{\mathbf{0}}(e, e_+)[(1-\alpha)(\gamma\tau_0 w)(e_+) + \alpha\tau_0(e)w(e)\rho(e_+)] \text{ in each sate } e. \tag{23}$$

Recall that if an unfunded system exists, a positive $\boldsymbol{\tau}_0$ satisfies $\boldsymbol{\tau}_0 < M_{\mathbf{0}}\boldsymbol{\tau}_0$. This vector satisfies the above inequalities because at autarky individuals invest so that $\sum_{e_+} q^d_{\mathbf{0}}(e, e_+)\rho(e_+) = 1$ and furthermore $\mu^d(e) \geq 1$. This proves the claim that a sustainable $\alpha-$mixed system exists whenever an unfunded one does. ■

Proof of Theorem 3 Consider an equilibrium in which the price of debt and contribution rates are both positive. In the absence of short sale constraints, the first order condition on i's debt holding is

$$Q(e) = \sum_{e_+} q^i_{\boldsymbol{\tau}}(e_+|e)\gamma(e_+)Q(e_+) \tag{24}$$

This condition holds for the decisive voter in state e. Using the definition of $M_{\boldsymbol{\tau}}(e, e_+)$, it can be rewritten as

$$\frac{Q(e)}{w(e)}\mu^d(e) = \sum_{e_+} M_{\boldsymbol{\tau}}(e, e_+)\frac{Q(e_+)}{w(e_+)} \tag{25}$$

Define the vector $\boldsymbol{Q}'$ by $q'(e) = q(e)/w(e)$. Recall that $\mu^d(e) \geq 1$ in each state. So $\boldsymbol{Q}'$ satisfies

$$\boldsymbol{Q}' \leq M_{\boldsymbol{\tau}}\boldsymbol{Q}' \text{ and } \boldsymbol{Q}' = M_{\boldsymbol{\tau}}\boldsymbol{Q}' \text{ only if } \mu^d(e) = 1 \text{ in all states } e \tag{26}$$

Moreover, by sustainability, $\boldsymbol{\tau}$ satisfies:

$$\boldsymbol{\tau} = M_{\boldsymbol{\tau}}\boldsymbol{\tau} \text{ and } \boldsymbol{\tau} > 0 \tag{27}$$

We now use well-known results on positive matrices. First (26) implies that $M_{\boldsymbol{\tau}}$ has a positive eigenvector with eigenvalue strictly larger than 1 whenever $\boldsymbol{Q}'$ differs from $M_{\boldsymbol{\tau}}\boldsymbol{Q}'$. Second, (27) implies that the maximal eigenvalue of $M_{\boldsymbol{\tau}}$ is 1 and that all positive eigenvectors are proportional to $\boldsymbol{\tau}$. Hence (26) and (27) imply that $\mu^d(e)$ is identically equal to 1 and that $\boldsymbol{\tau}$ and $\boldsymbol{Q}'$ are proportional. This means that the decisive voter is never subsidized and furthermore that, from his point of view, the return on the payg system and on debt are equalized. ∎

References

Azariadis, C., and V. Galasso (2002) Fiscal constitutions, *Journal of Economic Theory*, Vol. 103, pp. 255-281.

Bohn, H. (1998) Risk sharing in a stochastic overlapping generations economy, mimeo, University of California Santa Barbara.

Brown, J.R., O. Mitchell, and J. M. Poterba (2001) ”The Role of Real Annuities and Indexed Bonds in an Individual Accounts Retirement Program” in *Risk Aspects of Investment-Based Social Security Reform*, eds J. Campbell and M. Feldstein, University of Chicago Press for NBER.

Browning, E. (1975) Why the social insurance budget is too large in a democracy, *Economic inquiry,* Vol. 13 pp. 373-388.

Casamatta, G., H. Cremer, and P. Pestieau (2000) The political economy of social security *Scandinavian Journal of Economics* Vol. 102 pp. 502-522.

Chattopadhyay, S. and P. Gottardi (1999) Stochastic OLG models, market structures, and optimality, *Journal of Economic Theory* Vol.89 pp. 21-67. Conde-Ruiz, J.I., and P. Profeta (2007) The redistributive design of Social Security systems, *The economic Journal* 117, pp. 686-719. Cooley, T., and J. Soares (1999) A Positive Theory of Social Security based on Reputation, *Journal of Political Economy*, Vol 107 pp. 135-160. Demange, G. (2002) On optimality in intergenerational risk sharing, *Economic Theory,* Vol. 20 pp. 1-27.

Demange, G., and G. Laroque (1999) Social Security and demographic Shocks, *Econometrica,* Vol. 67 pp.527-542.

Demange, G. and G. Laroque (2000) Social Security, Optimality and Equilibria in a Stochastic Overlapping Generations Economy, *Journal of Public Economic Theory,*Vol. 2 pp.1-23.

Diamond, P. (1965) National Debt in a Neoclassical Growth Model, *The American Economic Review,* Vol. 55 pp. 1126-1150.

Esteban, J.M., and J. Sakovics (1993) Intertemporal Transfer Institutions, *Journal of Economic Theory,* Vol. 61 pp.189-205.

Forni, L. (2005) Social Security as Markov Equilibrium in OLG Models, *Review of Economic Dynamics,* Vol. 8 pp.178-194.

Geanakoplos, J., M. Magill, and M. Quinzii (2004) Demography and the Long Run predictability of the Stock Market, in *Brookings Papers on Economic Activity*, Vol.1 pp. 241-307

Gordon, R. H. and Varian, H. R. (1988) Intergenerational Risk Sharing, *Journal of Public Economics*, Vol.37 pp. 185-202.

Hammond, P. (1975) Charity: Altruism and cooperative egoism, *Altruism, morality and economic theory*, Russel sage Foundation New York pp.115-131.

Krueger, D., and F. Kubler (2006) Pareto improving Social Security reform when Financial Markets Are Incomplete, *American Economic Review*, Vol.96 pp.737-755.

Kandori, M. (1992) Repeated games played by overlapping generations *Review of Economic Studies* 59 pp. 81-92.

Manuelli, R. (1990) Existence and Optimality of Currency Equilibrium in Stochastic Overlapping Generations Models: The Pure Endowment Case, *Journal of Economic Theory*, Vol. 51 pp. 268-294.

de Ménil, G., F. Murtin, and E. Sheshinski (2006) Planning for the Optimal Mix of Paygo Tax and Funded Savings, *Journal of Pension Economics and Finance*, Vol. 5, No. 1.

Muench, T. (1977) Optimality, the interaction of spot and future markets, and the nonneutrality of money in the Lucas model, *Journal of Economic Theory,* Vol. 15 pp. 325-344.

Peled, D. (1984) Stationary Pareto Optimality of Asset Stochastic Equilibria with Overlapping Generations, *Journal of Economic Theory* Vol. 34, 396-403.

Shiller, R. (1999) Social Security and Institutions for Intergenerational, Intragenerational and International Risk Sharing, *Carnegie–Rochester Series in Public Policy*, Vol. 50 pp.165–204.

Spear, S.E. (1985) Rational Expectations in the Overlapping Generations Model, *Journal of Economic Theory,* 35, No. 2, pp. 251-275.

van Hemert, O. (2005) Optimal intergenerational risk sharing, mimeo Tinbergen Institute, Amsterdam.