Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

David Hollanders

The Political Economy of Intergenerational Risk Sharing

Discussion Paper 01/2009-011

The political economy of intergenerational risk sharing

David Hollanders[1]

Abstract

This paper analyses the political constraints of intergenerational risk sharing. The first result is that the political process generally does not lead to ex ante optimal insurance. The second result is that in a second best political setting PAYG still contributes to intergenerational risk sharing. The third result is that aging increases the discrepancy between first-best and second-best transfers. The source of the inefficiency is that politicians redistribute to larger and easier swayed cohorts. Ex post redistribution to lower incomes still leads to an outcome that from an ex ante point of view is preferable to a situation without intergenerational transfers.

JEL code: D72, E61, H21, H55.

Key words: risk sharing, aging, political economy

1 Introduction

By their nature macro-economic risks cannot be avoided, but their effect can be mitigated by sharing them between generations. The optimal design of intergenerational risk sharing has been investigated by many authors, including Gordon and Varian [1988], Beetsma and Bovenberg [2007] and Gollier [2008]. While this

[1] Tilburg University, department of general economics (d.a.hollanders@uvt.nl). The author thanks Lex Meijdam, as well as participants of Netspar lunchseminar (September 2008), the Nake Day (October 2008), and two Netspar macro-pillar theme meetings (September 2008, February 2009) for helpful comments.

literature focuses on the optimal design, this paper focuses on the political limits of optimal risk sharing instead.

Intergenerational risk sharing necessitates that future generations participate, but unborn generations are not able to agree on such participation. Once the young enter the insurance market, the outcome for the elderly is known and is thus uninsurable. Markets can therefore not implement optimal intergenerational risk sharing. The government can enforce participation and is thereby in the position to ensure the ex ante efficient transfer scheme.

This paper's central observation is that political decisions are not driven by efficiency. Political decisions depend on political support of voters. Political parties are not concerned with insurance but with redistributing towards large and/or easy swayed voters. The institution that is in a position to implement insurance is thus motivated by redistribution instead.

The central question is whether ex ante efficient risk sharing still arises endogenously in a democratic society that is primarily concerned with redistribution. This question is analyzed with a probabilistic voting model where the government gives priority to larger and easier swayed cohorts. The resulting answer is that politically determined allocations generally differ from the optimal allocations. From an ex ante perspective they are nonetheless better than "autarky" with each generation saving for itself. In a democratic political process, some risk sharing still arises. The reason is that low returns on savings for the elderly induces politicians to redistribute from young to older generations. The lower incomes of the latter makes their political support more sensitive to the tax policy, as political support depends linearly on utility and marginal utility is decreasing.

This paper build on the literature on the political economy of Social Security, see Galasso and Profeta [2002] for an overview. This literature focuses mainly on fixed transfers in a non-stochastic environment. Rangel and Zeckhauser [1999] and Demange [2005] are exceptions who consider a stochastic environment using a median voter model. The median voter model has the limitation that the majority generation holds all political power. This generally leads to a wide range of rationalizable equilibria when the median voter is young and the prediction of maximum taxation when the median voter is old. Following D'Amato and Galasso [2008] this paper proposes that a probabilistic voting model is better suited to analyze the political limits of risk sharing and the effect of aging upon that. In this approach a minority is no longer politically powerless and taxation is an important factor for voters but not necessarily the only one.

The main difference between this paper and D'Amato and Galasso is the differing question. The question here is to what extend a political process skewed towards redistribution still generates ex ante insurance. Therefore the welfare concept is ex ante efficiency and the efficient allocations are compared to the political process. Any deviation between the two may be considered the (unavoidable) costs of institutions. This deviation arises from the political clout of large cohorts.

D'Amato and Galasso instead compare maximization of a Social Welfare Function with the decisions of their government. Their normative benchmark then includes redistribution towards the first generation. Their government is a Stackelberg leader vis-à-vis future governments, not only anticipating future governments will "bail-out" the currently young generations out but actively exploiting it. Their question is therefore to what extent "exploiting" behavior of current governments -and not the political influence of large groups- aggravates redistribution -and not how it affects risk sharing per se.

The different approaches further determine the effect of aging on Social Security. D'Amato and Galasso predict that aging decreases contributions on Social Security, as a lower rate of return of PAYG decreases the scope to exploit future generations. This paper argues that the political clout of the elderly dwarfs all other effects and aging increases the discrepancy between the optimal transfer and the political feasible one.

This paper abstracts from incentive effects and general equilibrium effects, see Sánchez-Marcos and Sánchez-Martin [2006] and Kruger and Kubler [2006] for important contributions. Here not the exact degree of risk sharing is the issue but whether politics allows whatever is efficient to be executed.

The rest of the paper is organized as follows. The second section introduces the overlapping generation model with probabilistic voting and discusses the different agents in the model. The third section determines and discusses the first-best and second-best results in two different settings; first when savings are fixed and second when savings are endogenous. The last section concludes.

2. The model

In the next section two separate cases are considered; the case when savings are fixed and exogenous and the case when savings are endogenous. This section discusses the latter, more general setting.

The model consists of overlapping generations that live two periods. Wages and capital returns are determined on international markets and are exogenous from the perspective of the small, open economy considered here. People work in the first period and are retired in the second period. In the first period people inelastically supply one unit of labor, for which they receive a wage w that is normalized to one. Government can set a lump-sum tax τ distributing the proceedings to the retired generation. The after-tax wage can be either consumed or saved. Retirees consume accrued savings and any tax-financed benefits.

There are two states of the economy, one state (state L) in which gross capital return is low, and another state (state H) in which it is high. Capital return is Bernouilli distributed with $r^L < 1 < r^H$, where the superscript indicates the state of the economy. State L occurs with probability $0 \leqq \lambda \leqq 1$, and state H occurs with complementary probability $1 - \lambda$. There is constant geometric population growth n: $N_t = (1 + n)N_{t-1}$, where N_t is the number of young agents in period t and $n > -1$. The agents in the model are now discussed in turn.

2.1 Individuals

Agents born at time t have a time-separable utility function with felicity functions exhibiting constant relative risk aversion (CRRA):

$$V(c_t^y, c_t^o) = u(c_t^y) + u(c_t^o) = \frac{(c_t^y)^{1-\sigma}}{1-\sigma} + \frac{(c_t^o)^{1-\sigma}}{1-\sigma} \qquad \sigma \geqq 0$$

Here c_t^y and c_t^o represent consumption in the first and second period respectively of an agent born at time t. There is no time preference in this model. This could easily be included, but that would not add to the conclusion nor change it.

The young maximize expected life-time utility with respect to savings, taking taxes in the current period as given. For young agents, this boils down to maximizing $u(c_t^y) + Eu(c_t^o)$ subject to $w = c_t^y + \tau_t + s_t$ and given expectations on the tax rate in the next period (addressed below). The constraint follows from the life-time budget of the individual. Consumption in old age depends on the state of the economy and expected second period consumption is given by:

$$E(c_t^o) = \lambda[s_t r^L + (1+n)\widehat{\tau}_{t+1}^L] + (1-\lambda)[s_t r^H + (1+n)\widehat{\tau}_{t+1}^H]$$

The probability distribution of capital market returns is common knowledge; the distribution of taxation the next period is not a priori clear. At time t, a young individual has to form expectations, where $\widehat{\tau}_{t+1}^i$ denotes the expectation of someone in period t about tax-level in state i $(i = L, H)$ at time$t+1$. The model is closed with rational expectations: agents' conjecture about the probability distribution of taxes is correct.

Higher *aggregate* savings lead to lower benefits the next period, as the government does not take into account whether low income of the old resulted from low returns on savings or low savings. *Individual* savings have a neglectable effect on aggregate savings, and absent coordination agents will therefore not take into account the effect of current savings on future benefits.

Older generations leave no bequests and have no decision to make. They consume their entire accrued savings and their benefits, indicated by $b_t \equiv (1+n)\tau_t$.

2.2. First-best: optimal intergenerational risk sharing

In line with, among others, Ball and Mankiw [2001], ex ante efficiency is used as the welfare concept. Ex ante efficiency evaluates utility of agents prior to birth, when agents do not know in what state they are born. A policy may affect utility positively in one state of nature but negatively in another one. A policy is ex ante efficient if it maximizes the expected utility of an unborn individuals.

The solution technique is taken from Van Hemert [2005]. Let H^k denote the set of all possible k-histories in k subsequent periods. As there are two possible states of the world in each period, there are 2^k elements in H^k. Consider maximizing utility of a generation, conditioning on the last k periods only. Denoting two particular k-histories by h and h^+, ex ante utility is maximized by:

$$\begin{aligned}\overline{V}^k &= \max_{s(h),\tau(h)} \quad \sum_{h\in H^k}\sum_{h^+\in H^k} P[h_t = h, h_{t+1} = h^+] \times \\ & \quad [u(w - \tau(h) - s(h)) + u((1+n)\tau(h^+) + s(h)R(h^+))]\end{aligned}$$

$R(h), s(h), \tau(h)$ denote the interest rate, savings and taxes that occur in the particular k-history h; taxation and savings are contingent on the history the economy (the previous states), including the current state.

As an example, consider $k = 2$ histories. There are four possibilities for both h and h^+, namely (L, L), (L, H), (H, L), (H, H), where the last entry denotes the last time period. Note that h_t denotes the k-history at time t, whereas h_{t+1} denotes the k-history at time t+1. $P[h_t = h, h_{t+1} = h^+]$ denotes the probability that $h_t = h$ and $h_{t+1} = h^+$. This probability may equal zero, for example when $h_t = (L, L)$ and $h_{t+1} = (H, L)$; the first history indicates that state L occurred at time t and the latter indicates that H occurred at that time.

The maximization procedure takes ever more periods into account until convergence. The number of k-histories after which convergence emerges, indicates over how many generations risk is effectively spread.

2.3. Second-best: political redistribution

The government has the ability to tax the young generation and transfer the collected contributions to the old. Government runs a balanced budget with total contributions equaling total benefits. The incumbent party maximizes the following function:

$$\begin{aligned} W(\tau_t|s_{t-1}, s_t) &= \phi(1+n)E[V(c_t^y, c_t^o)] + u(c_{t-1}^o) \\ \text{s.t. } (1+n)\tau_t &= b_t \end{aligned}$$

The benefits per retiree at time t are again indicated by b_t. This equation results from probabilistic voting, see Persson and Tabellini [2000]. With probabilistic voting two competing vote-seeking political parties state their preferred tax-rate τ before elections and commit to this policy. Voters take the stated tax-rate into account while also considering a second and fixed characteristic of the political parties. This characteristic cannot be changed by the party and it may be interpreted as party ideology or charisma of the political leader. The more the fixed component matters, the fewer voters can be swayed by a change of policy.

A value of ϕ larger than 1 indicates that young voters are relatively less inclined to vote ideologically. They are more responsive to policy changes than older voters and more important to politicians because of it. As the parties are symmetric, they face the same maximization problem. Therefore their chosen policies converge, the outcome of which is given by the equation, see again Persson and Tabellini.

Following Meijdam and Verbon [1996], government and agents take each other's action as given. In each period the young maximize expected life-time utility taking taxes as given, and the government maximizes the probability of being elected taking savings as given.

The outcome of the two separate maximization-problems of individuals and government is equivalent to maximizing $W(.)$ simultaneously with respect to *both* savings and taxes together, that is:

$$\max_{s,\tau} \quad W_t(\tau_t|s_t, s_{t-1})$$
$$\text{subject to } (1+n)\tau_t = b_t \text{ and subject to } w = c_t^y + \tau_t + s_t$$

Government can do no better as their objective function is maximized, but given the maximizing value of taxation, maximizing $W(.)$ coincides with maximizing life-time utility of the younger cohort. The reason is that utility of the young cohort enters additively. Complicating point in solving this equation is that the optimal values of savings and taxation depend on savings a period earlier and on the current (rational) expectations of taxes a period later, which in turn coincides with actual taxes. The appendix shows existence of a unique equilibrium.

An alternative for the probability voting model is the median voter model. This set-up has been studied by Browning [1975], Breyer and Stolte [2001], Boldrin and Rustichini [2000], Conesa and Krueger [1999], Cooley and Soares [1996, 1999], Rangel and Zeckhauser [1999], and Sjoblom [1985]. Which approach is more suitable depends on the particular question posted. In the case of population growth and a young decisive voter, the median voter model shows the importance of coordination for the occurrence of positive transfers. When workers do not trust future generations to replicate their own positive transfer to retirees, an (inefficient) tax of zero arises. The median voter model thereby makes immediate that young only contribute (voluntarily) to PAYG when they anticipate that future generations do likewise.

However, a disadvantage is that it predicts a sharp, discontinuous shift in political power when population growth changes sign; this is a consequence of the lack of political power of even a large minority. When retirees form a majority a taxation of 100% is predicted. When young voters form a majority, there is a wide range of rationalizable results, including zero transfers. Probabilistic voting instead allows a more reasonable and precise analysis of the effect of population growth on efficiency properties of pension politics.

3. First-best versus second-best

This section determines, discusses and compares first- and second-best allocations. First-best allocations refer to the allocations that maximize utility of a steady-state generation. These are the ex ante efficient outcomes which are compared by the outcomes of a political process driven by redistribution. The latter are referred to as the second-best allocations. Before turning to the general case, the case were savings are fixed are considered. This offers insight in the model and allows for a simple analytical solution.

3.1. Fixed savings

In order to exclusively focus on the mechanism of risk sharing, first the case of fixed savings is considered. The problem is further simplified by allowing only non-negative transfers and transfers in state L.

The ex ante efficient transfer when capital return is low, follows from maximizing the following function where $\overline{s}$ denotes fixed and exogenous savings[2]:

$$\max_{\tau} \frac{1}{1-\sigma}\{(w-\overline{s}-\tau)^{1-\sigma} + (\overline{s}r^L + (1+n)\tau)^{1-\sigma}\}$$

The first order condition equals:

$$(w-\overline{s}-\tau)^{-\sigma} = (1+n)(\overline{s}r^L + (1+n)\tau)^{-\sigma}$$

Solving this equation gives the solution:

[2]The expression is derived as follows. Ex ante utility is given by:
$\max_{\tau} \lambda^2[u(w-\overline{s}-\tau) + u(\overline{s}r^L + (1+n)\tau)]+$
$\lambda(1-\lambda)[u(w-\overline{s}-\tau) + u(\overline{s}r^H)]+$
$\lambda(1-\lambda)[u(w-\overline{s}) + u(\overline{s}r^L + (1+n)\tau)]+$
$(1-\lambda)^2[u(w-\overline{s}) + u(\overline{s}r^H)]$
This simplifies to:
$\max_{\tau} u(w-\overline{s}-\tau) + u(\overline{s}r^L + (1+n)\tau)$, as all expressions without τ can be omitted.

$$\tau^{fb} = \frac{w-\overline{s}[1+r^L(1+n)^{-\frac{1}{\sigma}}]}{1+(1+n)^{\frac{\sigma-1}{\sigma}}}$$

The first-best tax increases when wages increase or when savings or capital return in the low state decrease. Higher income for young and a lower income for retirees increase the ability and the need respectively to insure retirees against a shortfall in accrued savings. The effect of population growth is ambiguous; if there are relatively many young workers, this increases the rate of return of PAYG. The other way round, the same income can be insured to the old by a lower contribution of the young, which can consume the remaining. This income effect and insurance effect work in opposite directions and the net effect on the first-best taxation depends on the particular values of the parameters involved.

The second-best allocation follows from the following procedure:

$$\max_{\tau} \frac{1}{1-\sigma}\{\phi(1+n)(w-\overline{s}-\tau)^{1-\sigma} + (\overline{s}r^L + (1+n)\tau)^{1-\sigma}\}$$

The first order condition is:

$$\phi(1+n)(w-\overline{s}-\tau)^{-\sigma} = (1+n)(\overline{s}r^L + (1+n)\tau)^{-\sigma}$$

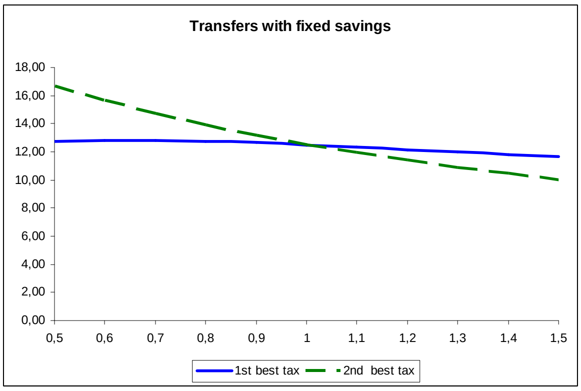
Solving this equation gives the following solution:

$$\tau^{sb} = \frac{w\phi^{-\frac{1}{\sigma}} - \overline{s}[r^L + \phi^{-\frac{1}{\sigma}}]}{1+n+\phi^{-\frac{1}{\sigma}}}$$

This expression gives the transfer decided on by the government. It increases when wages increase or when savings decrease, as poorer voters are easier swayed by a transfer. The reason is that support depends linearly on utility and utility increases faster for lower values of consumption. The net effect of population growth is unambiguously negative. An increase in n leads to more political clout for the young. This effect dominates the lower rate of return of a PAYG-arrangement.

The expressions for the first-best and second-best transfers generally differ but coincide in a special yet meaningful case. When there is an equal number of older and younger voters and they are equally influence able by the tax policy ($\phi = 1 + n = 1$), first-best and second-best transfers coincide and equal $\frac{w - \bar{s}[r^L + 1]}{2}$. This is discussed in more detail below. Generally the transfers do not coincide as the expressions indicate and the following picture illustrates.

Table 1: parameter values in base-line scenario		
wage	w	1
Capital return in state L	r^L	0.5
Capital return in state H	r^H	2
probability of state L	λ	0.5
Coefficient of relative risk aversion	σ	4
relative ideological bias of the young	ϕ	1



Graph 1

Graph 1 gives transfers as a function of population growth for the values given in the table. These values are not calibrated and are for illustrative purposes only. The expected return of the financial asset equals 1.25, so for values $n > 0.25$ the rate of return of a PAYG-arrangement is higher than that of the capital market.

In expected terms the economy is dynamically inefficient from that point on. As savings are fixed this does not affect savings.

Population growth, n, is a main determinant of differences in first-best and second-best taxation, but is not the only one. The different degree to which young and older voters are loyal to a party, captured by ϕ, influences the discrepancy between first-best and second-best as well. A value ϕ exceeding one indicates that the old are relatively more ideological, which may result from habit-formation or party loyalty. A value ϕ lower than one results when the old are more footloose and thereby exert de facto extra political power. This would occur if older voters care relatively more about pensions than the working generation, for whom pensions is an issue that will become relevant years from now. A special case results if $\phi = \frac{1}{1+n}$. Then the cohort-size effect and the ideological factor cancel out and the transfer decided by government is ex ante efficient.

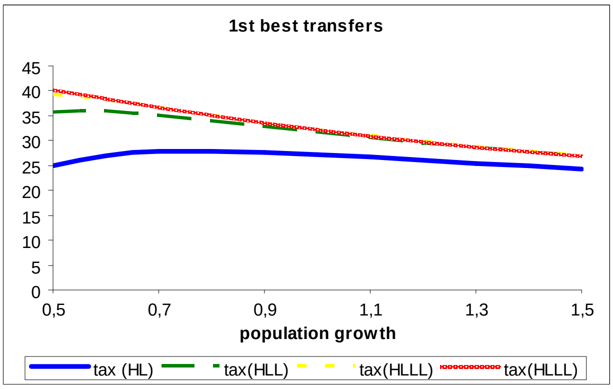
Thus far the tax scheme involved only non-negative transfers and only transfers in the L state with low capital returns. This allowed focusing on the most important insurance-aspect of risk sharing; workers providing support to the elderly when capital markets tumble. This restrictive assumption means that only labor is taxed, not capital or income. While labor is indeed most heavily taxed in most Western countries, this assumption is restrictive nonetheless. If taxation can be both positive and negative in both states of the world, transfers in state L would be the same. First-best and second-best transfers in state H would change and they can be derived analogously with r^H replacing r^L in the expressions given above. This case is not analyzed further; incorporation of taxation in state H influences the size of the welfare effects, but does not change qualitative conclusions. It is however important to note that what is called here first-best is second-best from a more general perspective that considers less restrictive forms of taxation.

3.2. Endogenous savings

The assumption of fixed savings provides insight in the model and allows for analytical solutions. These fixed savings can be understood as mandatory and fixed contributions to a Defined Contribution scheme that cannot be avoided and can also not be increased easily. Nonetheless, fixed savings is a stark assumption. One particular effect is that transfers solely depend on the current state (L or H) instead of the entire history. With endogenous savings this is no longer case, opening up the possibility that discrepancies between taxes spill over to the next period. In this circumstance the assumption that government takes savings as given, becomes relevant. This is further discussed below.

3.2.1. First-best: optimal intergenerational risk sharing

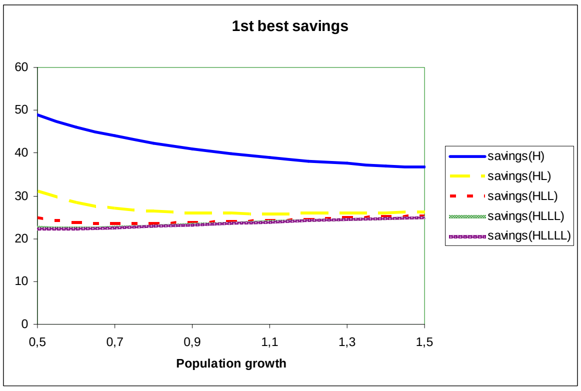
The outcome when savings are endogenous is considered in graphs two and three. This case is solved numerically for the same parameter values as before. Graph 2 shows the ex ante efficient transfers for different histories as a function of population growth. Graph 3 does the same for savings. Both savings and taxation are given in percentages.



Graph 2

The results show that optimal transfers are contingent on the state of the economy, as they are by construction. Taxation does not necessarily increase in population growth, possibly counter intuitively. On the one hand, a high implicit rate of return induces higher contributions, on the other hand, the "insurance" for the elderly can be accomplished by lower contributions, exactly because of the higher rate of return. This in turn leaves more scope in the first period for consumption or for savings.

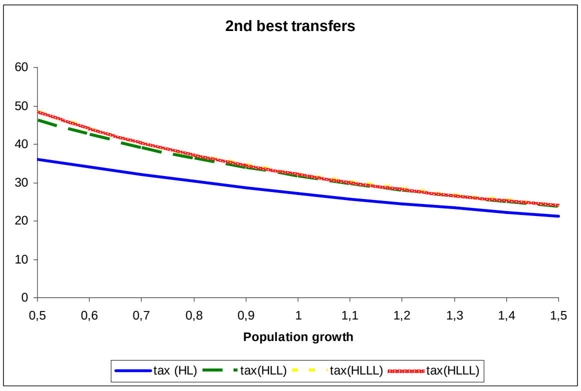
The tax rate increases when more stock market crashes (state L) preceded. Taxation spills over to future periods. This is exactly the principle of risk sharing which spreads losses over all future generations, see Ball and Mankiw. The spill-over takes place via savings. In the presence of a bear market, young contribute to the old, crowding out their own savings. As the young have lower accrued savings the next period, the government in the next period transfers more from the then young to them if market returns are again low.



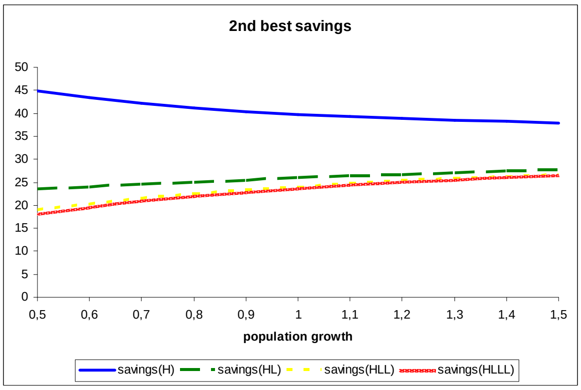
Graph 3

3.2.2. Second-best: political redistribution

The second-best allocation results from the incumbent party maximizing the number of votes by adapting its policy. The graphs show the outcomes for transfers and savings.



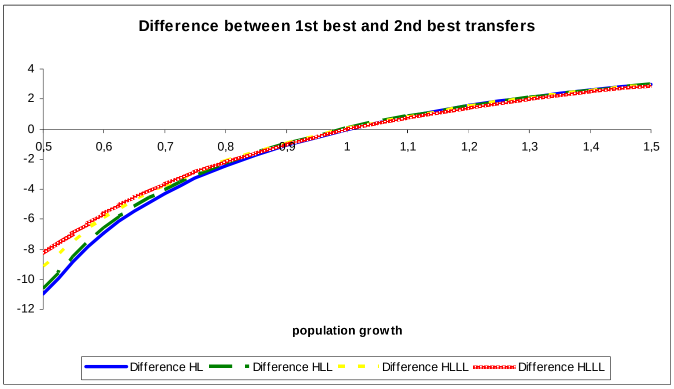
Graph 4



Graph 5

Taxation decreases in population growth. This is the net result of two opposing effects population growth has on contributions. Lower population growth decreases the rate of return of PAYG, thereby making it less attractive to redistribute from the working generation to the older generation. However, lower population growth also increases the political clout of the elderly, which has an upward effect on taxation. With severe aging ($n = 0.5$), the maximum contributions equal 49% whereas efficient taxation equals 40% in that case. This increase in taxes does not imply increased benefits, which equal contributions multiplied by population growth. This reproduces the central result of Breyer and Stolte [2001], who argue that the 'burden' of aging is shared by old and young with lower benefits and higher contributions.

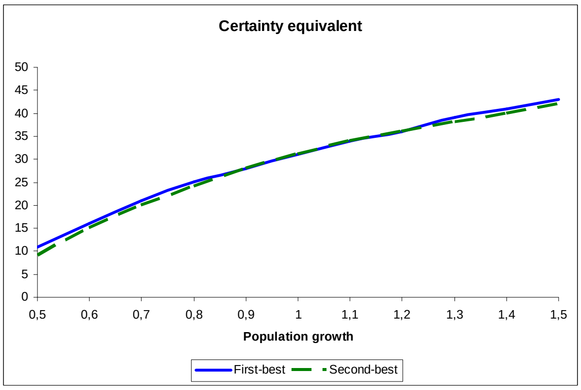
Again taxation increases in the number of previous periods that L occurred. In state L the asset market crashes; this hurts the retirees whereas the working generation is not hit directly. As poorer voters are easier swayed, this results in redistribution from young to old in state L. Workers save less as a result. When state L occurs a second time in a row, workers have saved less and face low returns. This makes them even worse off than the generation that was first hit by low returns. This in turn induces an even higher transfer from the young generation that is working then. Together this leads to higher taxation every time an extra period in state L occurs. This resembles the contingent optimal taxation that likewise increases in the number of preceding market lows. While the direction of first-best and second-best transfers is the same, the precise height differs. This is illustrated in graph 6 which gives the difference between transfers for different contingencies as a function of population growth. As can be seen this is a decreasing function. In the presence of aging politics overshoots, while transfers are 'too low' in the presence of population growth.



Graph 6

3.2.3. Welfare comparison

The difference between politically determined transfers and ex ante efficient transfers affects welfare. Ultimately this effect on welfare is what matters, not the effect on the tax rate per se. Graph 7 depicts the certainty equivalent of the first and second-best taxes as a function of population growth. The certainty equivalent follows from a comparison between the ex ante utility in the presence of a transfer and its complete absence. In the latter case, which can be thought of as "autarky", the consumption during retirement age is solely determined by accrued savings. Each young generation decides its own savings which in the baseline case equals 59%. The certainty equivalent is defined as the wage that makes someone in such an autarkic state (with $\tau = 0$) equally well off as someone that has the insurance that the first-best or second-best arrangements provide.



Graph 7

The graph shows that utility increases in population growth, which follows directly from the observation that population growth is the rate of return of a PAYG-scheme. As can also be seen, the first-best allocation is superior except for the case that population size is stable. In other demographic scenarios welfare differences are present, equaling about two percentage points of entire life-time wage income. As the values are not calibrated, not too much should be inferred from the exact numbers. They illustrate that intergenerational risk sharing provided by politics comes with the cost of redistribution.

Ex ante efficiency is the welfare criterion. This can be distinguished from the concept of a social planner maximizing a Social Welfare Function. Ex ante utility assesses utility of someone prior to birth and when the PAYG-system is in place. Expected life-time utility of a 'steady state' generation is subsequently maximized. A steady state generation does not know in which state of the world it is born, except that a PAYG-system is already in place. A Social Welfare Function on the other hand typically maximizes utility of current and future generations explicitly, starting with the retired generation that receives a windfall gain when the PAYG-system originates. Ex ante utility does not consider the welfare gain for the first retired generation. Though this is a real gain for the first generation, this welfare criterion combines risk sharing and redistribution to the first generation (for which good reasons may or may not exist). This paper wants to focus solely on the political effects on risk sharing. The distortionary effect of politics results primarily from the increased political influence of large cohorts.

This can be contrasted with D'Amato and Galasso who use a Social Welfare Function as the welfare criterion. Under a reparameterization this SWF coincides with the political outcome of this paper.[3] D'Amato and Galasso then effectively analyze the effect of a further inefficiency that springs from the Stackelberg behavior of their government (see further below).

3.2.4. The source of inefficiency

As discussed, the first-best and second-best transfers only coincide in a special but meaningful case. The equivalence follows analytically when savings are fixed and was illustrated above when savings are endogenous. In the latter case the equivalence can also be shown mathematically[4]. In other cases, the different ef-

[3]With previous savings $\overline{s}$ given and implementing a transfer scheme in state L, the SWF equals: $u(\overline{s}r^L + (1+n)\tau^L) + (1+n)\rho\{u(w - s^L - \tau^L)$
$+\lambda u(s^L r^L + \tau^{LL}(1+n)) + (1-\lambda)u(s^L r^H))\}+$
$(1+n)^2\rho^2\{u(w - s^{LL} - \tau^{LL})$
$+\lambda u(s^{LL} r^L + \tau^{LLL}(1+n)) + (1-\lambda)u(s^{LL} r^H))\}$

Here ρ is the discount rate the social planner uses for future generations. This function is maximized w.r.t. savings $(s^L, s^{LL}, ..)$ and taxes $(\tau^L, \tau^{LL}, ..)$. Note that only a sequence of states L is considered. Taxes in state H equal zero, so they need not be considered explicitly. After at least one state H a new sequence arises that may lead to different taxation but that is not considered here. With $\rho = \phi$ the function exactly equals a sequence of incumbent parties maximizing political support. The difference is that that function is maximized each period w.r.t. current savings and taxes. However if the optimizing values of the social planner would be proposed, no individual or political party could improve upon that solution. Hence it is also the outcome of the political process.

[4]Suppose histories with k periods are considered. Ex ante efficient transfers are then contingent on $k-1$ periods. Denote by h one particular contingency of $k-1$ periods. Using that $n = 0$ first-best allocations follow from:
$\max_{\tau^h} \lambda p(h) u(w - s^h - \tau^h)+$
$(1-\lambda)p(h)u(w - s^h - \tau^h)+$
$\lambda p(h) u(s^h r^L + \tau^h)+$
$(1-\lambda)p(h)u(s^h r^L + \tau^h)$

Here $p(h)$ denotes the probability of h occurring, and s^h and τ^h denote savings and transfers in that contingency. Contingencies with the first $k-1$ periods equal to h and ending in either state H or L contribute to utility in the first period, whereas contingencies with the last $k-1$

fect of population growth is the source of the discrepancy between first-best and second-best. The reason is that they have a tendency to 'buy' votes by using the transfer mechanism to redistribute to larger and more pragmatic cohorts. In so, politicians overweigh large cohorts.

In the political process current political parties take future taxes as given and do not take into account that current taxes (may) influence future taxation. The connection results from the effect of taxation on private savings. If these are crowded out by public transfers, workers will have lower savings during their retirement. This in turn induces the next government to redistribute to these retirees. Higher taxes lead to higher future benefits for current workers. D'Amato and Galasso propose that governments take that into account. Politicians exploit their first-mover advantage by increasing taxes at the expense of private savings, resulting in higher taxes in the future. Taking this effect into account adds another source of inefficiency. Which approach is more reasonable, is an empirical question. Nash behavior is restrictive as political parties do not consider future reactions whatsoever. The Stackelberg-approach is also not unproblematic, as the opposite extreme is assumed: governments know and take fully into account the behavior of future governments. As one period in a two-period OLG model represents 30 years, this is likewise a stark assumption. In reality there will typically be a mixed case. The assumption of Nash behavior allows focusing exlusively on the political effect of cohort size on intergenerational risk sharing.

4. Conclusion

Intergenerational risk sharing cannot be implemented by markets. Only the gov-

periods equal to h and starting in either state H or L contribute to utility in the second period. Together this gives the four parts of the expression. Simplifying that expression gives

$\max_{\tau^h} u(w - s^h - \tau^h) + u(s^h r^L + \tau^h)$

This coincides with the second-best transfers, if first-best and second-best savings are equal. This is in turn assured because for given taxes, the first-best savings decision and the second-best savings decision coincide. As this holds for all contingencies, first-best and second-best transfers coincide if $n = 0$.

ernment or a pension fund with mandatory participation can do so. However, political institutions have an incentive to redistribute. In the realm of intergenerational risk sharing there is thus no insurance without redistribution.

As a result ex ante efficiency is generally out of reach of politics as well. Such are the unavoidable costs of political institutions. This is the first result of the paper: politics generally does not lead to ex ante efficiency. From an ex ante perspective, politics is still better than no transfers at all and this is the second result of the paper. The third result concerns the effect of aging on intergenerational risk sharing. As Bovenberg [2008] states, "the danger facing aging societies is that older voters block the needed reforms. In that case, a conflict arises between the political power of older generations (who depend on public transfers and are risk averse) and the economic power of the younger, working generations (who control the major resource that fuel the modern knowledge-intensive economy –namely, human capital and entrepreneurship). In other words, politics collides with economics." In this paper a majority of retirees cannot outright block a reform, as a median voter model would have it. However, some concern is indeed warranted, as aging increases the discrepancy between first- and second-best outcomes.

Social Security in many countries developed in the 1930s-1950s; for example the USA implemented PAYG-financed Social Security in 1937 when many retirees had suffered the Depression. The resulting decrease in retirements was an important motive for Social Security. Although the paper does not predict a specific moment Social Security is implemented, timing and motivation coincide with an implication of the model that intergenerational transfers are used when retirees have witnessed a severe financial set-back.

The conclusions hold for a wide range of population growth rates and are robust to changes in risk aversion (not shown here) and endogeneity of savings. The current set-up is nonetheless limited in several ways. It does not address incentive and crowding out effects, thereby overestimating the gains of risk sharing. However, by modeling only two generations the potential gain from risk sharing between more generations is underestimated. The model also does not take economic growth into account, further underestimating the return on PAYG. And by ignoring other risk-factors as longevity and productivity the importance of risk sharing is underestimated. Though a welfare gain of 31% in the base-line case is remarkably in line with findings of Van Hemert (33%) and Gollier (25%), it cannot be assessed a priori what the net result of these different effects is. These limitations are recognized and some of them could be further incorporated into the model.

These limitations are however not essential for the main argument. If politics hinders risk-sharing in the relatively simple set-up here, then sharing risks in a more complex environment can certainly not be taken for granted. The aim of this paper is not to determine the optimal tax per se, but to provide an analytical framework to assess whether whatever is optimal arises endogenously in the political process. The bad news is that the democratic process generally does not lead to efficiency. The good news is that politics does a lot better than a situation without any intergenerational transfers.

References

Ball, L. and N. G. Mankiw [2001], 'Intergenerational Risk Sharing in the spirit of Arrow, Debreu, and Rawls, with applications to social security design', NBER working paper 8270.

Beetsma, R. and A. L. Bovenberg [2007], 'Pension systems, intergenerational risk sharing and inflation', CEPR-discussion paper 6089.

Bovenberg, A.L. [2008] 'Grey New World: Europe on the Road to Gerontocracy?', CESifo.

Breyer, F. and K. Stolte [2001], 'Demographic change, endogenous labor supply and the political feasibility of pension reform', *Journal of Population Economics*, (14), 409-424.

Browning, E. K. [1975], 'Why the Social Insurance Budget is Too Large in a Democracy', *Economic Inquiry,* 13(3), 373-388.

Conesa, J. C. and D. Krueger [1999], 'Social Security Reform with Heterogeneous Agents', *Review of Economic Dynamics*, 2, 757-795.

Cooley, T. F. and J. Soares [1999], 'A Positive Theory of Social Security Based on Reputation', *Journal of Political Economy*, 107(1), 135-169.

Cooley, T. F. and J. Soares [1996], 'Will social security survive the baby boom?', Carnegie-Rochester Conference Series on Public Policy 45, 89-121.

D'Amato, M and V. Galasso [2008] 'Political Intergenerational Risk Sharing', IGIER working paper n. 342.

Demange, G. [2005], 'On Sustainable Pay As You Go systems', PSE working paper No. 2005-05.

Galasso, V. and P. Profeta [2002], 'The political economy of social security: a survey', *European Journal of Political Economy*, 18, 1-29.

Gollier, C. [2008], 'Intergenerational risk-sharing and risk-taking of a pension fund', *Journal of Public Economics*, Vol. 92, 1463-1485.

Gordon, R. H. and H. R. Varian [1988], 'Intergenerational Risk Sharing', *Journal of Public Economics*, 37, 185-202.

Hemert, O. van [2005], 'Dynamic Investor Decisions', PhD thesis.

Krueger, D. and F. Kubler [2006], 'Pareto-Improving Social Security Reform when Financial Markets Are Incomplete!?', *American Economic Review*, 96(3), 737-755.

Meijdam, L. and H.A.A. Verbon [1996], 'Aging and political decision making on public pensions', *Journal of Population Economics*, 9, 141-158.

Persson, T. and G. Tabellini [2000] 'Political Economics', *MIT Press*.

Rangel, A. and R. Zeckhauser [1999], 'Can market and voting institutions generate optimal intergenerational risk sharing?', NBER-working paper 6949.

Sánchez-Marcos, V. and A.R. Sánchez-Martin [2006], 'Can social security be welfare improving when there is demographic uncertainty?', *Journal of Economic Dynamics & Control* 30, 1615-1646.

Sjoblom, K. [1985], 'Voting for social security', *Public Choice (*45), 225-240.

Appendix: Existence of a unique solution

The objective is to maximize at each time t the following function. Simplification of the problem by posing $n = 0$ does not alter the problem essentially:

$$\max_{\tau,s} \quad W_t(s_{t-1}, \widehat{\tau}_t^L) = \frac{1}{(1-\sigma)}[r_t s_{t-1} + \tau]^{1-\sigma} + \frac{1}{(1-\sigma)}\{[w - \tau - s]^{1-\sigma} + \lambda[sr^L + \widehat{\tau}_t^L]^{1-\sigma} + (1-\lambda)[sr^H]^{1-\sigma}\}$$

At each time t s_{t-1} and $\widehat{\tau}_t^L$ are given. So there exists a unique solution at each time t by the theorem of Weierstrass and strict concavity of the objective function. However, it needs to be shown that solutions converge in the following sense: there exists an integer l such that taxes and savings conditional on the last l periods equal taxes and savings conditional on the last $l + m > l$ histories for any $m > 0$. So, savings and taxes at time t only depend on (the state of the economy in) the last l periods. This in turn effectively allows 'cutting off' history at one point. For l=3, m=1, this would for example mean $\tau^{HHLL} \approx \tau^{LHLL}$. The superscript here indicates the state in the current period (here L) the and previous periods.

Now taxes are set zero in state H (corresponding to a booming capital market); this makes that size of the savings of retirees do not matter for savings of the young. The old just consume their savings, while the young optimize given their expectation of future taxation, irrespective of the old. So, once state H is reached, only the current state (H) matters and convergence follows a fortiori.

The problem now reduces to showing that there is a period for which $\tau^{i..jL..LL} \approx \tau^{p..qL..LL}$. That is, no matter with what previous savings and expectations state L is reached, after having been in the same state L a finite number of times taxation and savings converge. Note that $W_t(s_{t-1}, \widehat{\tau}_t^L)$ maps previous savings (s_{t-1}) and expected taxes in the future state L ($\widehat{\tau}_t^L$) into current savings and taxes s_t and τ_t. (Expected taxes in the future state H are zero.) Call this mapping informally $f^L(s_{t-1}, \widehat{\tau}_t^L) = (s_t, \tau_t)$. Now, the only candidate for convergence is a fixed point (s^*, τ^*) of this mapping, defined by $f^L(s^*, \tau^*) = (s^*, \tau^*)$. By the fixed point theorem, at least one fixed point exists.

Now, savings converge if and only if taxes converge, that is:

1) $f^L(\widetilde{s}, \tau^*) = (s^{'}, \tau^*) \Longrightarrow \widetilde{s} = s^{'} = s^*$ and 2) $f^L(s^*, \widetilde{\tau}) = (s^*, \tau^{'}) \Longrightarrow \widetilde{\tau} = \tau^{'} = \tau^*$

Ad 1) If $\tau_t = \widehat{\tau}_t^L = \tau^*$, then for any s_{-1} the young face the same maximization problem; previous savings do not enter their maximization problem. Hence $s_t = s^*$. This in turn implies $s_{-1} = s^*$. If not, then the old are strictly worse or strictly better off than if $s_{-1} = s^*$, which would induce higher or lower taxation by the politician, due to strict concaveness of the objective function, contradicting $\tau_t = \tau^*$.

Ad 2) If previous savings are s^*, then taxes expected by the young are τ^*, and consequently $f^L(s^*, \tau^*) = (s^*, \tau^*)$ by definition.

Convergence of either savings or taxes is enough for convergence of the two-dimensional function. This reduces a two-dimensional problem to a one-dimensional problem. The function f^L is hard to derive analytically, but can readily be sketched graphically. As long as properties concerning slope, curvature and the behavior at corner points are given, the equation can be solved by use of phase diagrams. This involves the drawing of a 45-degree line, the intersection of which with the function gives the fixed point. If the solution exists, as indeed it does here, then in the one-dimensional case stability conditions are that the derivative is strictly less than 1 in absolute value. As can be seen from the graphs, shown in the appendix, this is the case.

An important assumption in the above is the zero taxation in state H. This considerably reduces the complexity of the problem. This is done for practical considerations and considering it does not change the intuition of the result. An interpretation of such a restriction is that it is not politically feasible to impose positive taxation during a booming capital market.