Intergenerational Risk Sharing via Social Security When Financial Markets are Incomplete*

Dirk Krueger
And
Felix Kubler
Stanford University

December 2001

The Stanford Institute for Economic Policy Research at Stanford University supports research bearing on economic and public policy issues. The SIEPR Discussion Paper Series reports on research and policy analysis conducted by researchers affiliated with the Institute. Working papers in this series reflect the views of the authors and not necessarily those of the Stanford Institute for Economic Policy Research or Stanford University.
Intergenerational Risk Sharing via Social Security when Financial Markets are Incomplete

Dirk Krueger                Felix Kubler
Dept. of Economics          Dept. of Economics
Stanford University        Stanford University

December 10, 2001

Abstract

This paper develops an overlapping generations model with stochastic production and incomplete markets to assess whether the introduction of an unfunded social security system can lead to a Pareto improvement, even if the initial equilibrium is neither production-inefficient in the spirit of Diamond (1965) nor dynamically inefficient in the spirit of Samuelson (1957).

When returns to capital and wages are imperfectly correlated and subject to aggregate shocks, then the consumption variance of all generations can be reduced if private markets or government policies enable them to pool their labor and capital incomes. A social security system that endows retired households with a claim to labor income may serve as an effective tool to share aggregate risk between generations, in the absence of financial securities that could serve a similar purpose.

We construct numerical examples which demonstrate that the intergenerational risk sharing role of social security can be sufficient to warrant its introduction on welfare grounds. For a realistically calibrated economy, however, we find that this role is insufficiently strong quantitatively to offset the negative effects an unfunded system has on capital accumulation and thus does not constitute a Pareto improving policy.

*We thank K. Judd, L. Kotlikoff, S. Pauli, J. Soares and K. Smetters for helpful suggestions, and gratefully acknowledge financial support under NSF grant SES-0004376. All remaining errors are our own.
1 Introduction

The current U.S. pay-as-you go social security system was introduced in 1935, partly as a response to the impoverishment of an entire generation during the great depression, the biggest negative aggregate shock the U.S. economy has experienced so far. In the current political and academic debate about social security reform one of the major concerns cited by the opponents of a reform towards a funded system is the risk of low returns to savings for an entire generation due to a large and long unfavorable aggregate shock (see Aaron et al. (2001) or Burtless (2001) for a discussion).

Despite the fact that the role an unfunded social security system may play in facilitating the allocation of aggregate risk among generations was a key consideration in its introduction and is a key discussion point in its current reform debate, academic research on the role of social security in models with aggregate uncertainty, which affects different generations along different margins, remains limited.

As Bohn (1998, 1999) has argued, if returns to capital and wages are imperfectly correlated and subject to aggregate shocks, then the consumption variance of all generations can be reduced if private markets or government policies enable them to pool their labor and capital incomes. A social security system that endows retired households with a claim to labor income may serve as such an effective tool to share aggregate risk between generations, in the absence of financial securities that achieve the same risk allocation.

It is the goal of this paper to evaluate the quantitative importance of this intergenerational risk sharing role of social security, in comparison to the more traditional arguments of reducing overaccumulation of capital and intergenerational insurance and redistribution. In order to give this question historical content we ask whether, in a situation like in 1935, after a large adverse macroeconomic shock, it is possible to justify the introduction of a unfunded, redistributive pay-as-you go social security system on the normative grounds that it provides a welfare improvement, in the sense of ex-interim Pareto efficiency. That is, we ask whether it is possible that the economy in 1935 was in an aggregate state such that the introduction of a unfunded, redistributive pay-as-you go social security system, as implemented by the Roosevelt administration increased utility of all generations then alive and expected utility for all future generations to be born, where a generation is not only identified by the time of birth, but also by the node of the event tree at which it is born.

We find that in realistically calibrated economies intergenerational risk
sharing alone is unlikely to provide, based on a ex-interim Pareto criterion, a rationale for the introduction of an unfunded social security system. However, this result depends crucially on the calibration as well as on the assumption that we consider a closed economy. We also provide numerical examples where social security leads to considerable, Pareto-improving, welfare gains, both in a closed economy and in a small open economy.

The introduction of an unfunded redistributive, payroll-tax financed social security system affects the competitive equilibrium in a number of ways. It reduces the capital accumulation in the economy and distorts the labor-leisure decision, but also may provide intragenerational and intergenerational risk-sharing and redistribution. Since the main focus of the paper is evaluate the quantitative intergenerational risk sharing role of social security, we consider an economy in which all other effects of introducing an unfunded redistributive social security system, besides the adverse effects on capital accumulation are absent. This economy is populated by three generations at each point of time and there is no heterogeneity within each generation (and therefore no intragenerational redistribution or insurance motive of social security). The stochastic labor endowment is inelastically supplied. The only asset available for trade between generations is physical capital, which yields stochastic returns. The introduction of physical capital with stochastic returns has a two-fold motivation: first it enables intergenerational trade, so that social security is not the only mechanism of consumption smoothing over time (as in Rangel and Zeckhauser (1997), Bohn (1999) and many others), and second, it allows us to re-examine Diamond’s (1965) discussion of production inefficiency of competitive equilibrium in a stochastic environment.

In order to discuss efficiency properties of equilibrium allocations it is crucial to make precise the different notions of efficiency used in this paper. By Pareto efficiency we mean Pareto efficiency in the ex interim sense as defined above. We say that an allocation is production efficient if there exist is no alternative allocation for which aggregate consumption is weakly higher at every node and strictly higher at some node of the event tree. We attribute this notion of efficiency to Diamond (1965) and Zilcha (1990).

Finally, we say an allocation is dynamically efficient, if there is no set of intergenerational transfers of consumption that lie in the marketed subspan of the consumption set and that make every generation as least as well off and some generation strictly better off (again in an ex interim sense). We attribute this notion of efficiency to Samuelson (1957) and Demange (2001).

In this paper we want to focus on cases where the level of capital ac-
cumulation is production-efficient and where the consumption allocation is
dynamically efficient. If markets are sequentially complete, in that agents
can trade a full set of claims contingent on the realization of aggregate un-

certainty then an equilibrium allocation is Pareto efficient if and only if it
is dynamically efficient. The presence of incomplete markets, however leads
to a possible divergence of these efficiency criteria and the introduction of
an unfunded social security system can be Pareto improving.

We focus on dynamically efficient allocations because empirical evidence
seems to suggest that the US economy is dynamically efficient (see Abel et
al. (1989)) and because we want to separate intergenerational risk sharing
from inefficiency of the Samuelson type.

Via numerical examples we then demonstrate that the introduction of
social security into the competitive economy can lead to a welfare improving
consumption allocation in the Pareto sense, even though the original alloca-
tion was not dynamically inefficient. This Pareto improvement is purely due
to enhanced intergenerational risk sharing of imperfectly correlated shocks
to the returns to labor and capital. We then argue, however, that for real-
istically calibrated examples the risk sharing benefits of an unfunded social
security system tend to be dominated by its negative affect on capital ac-
cumulation. As we show, if one is willing to make a small open economy
assumption and thus shut down the latter effect, then even for realistically
calibrated examples the introduction of an unfunded social security system
may constitute a Pareto improvement.

The paper is organized as follows. In the next section we describe the
model, define equilibrium and efficient allocations, and describe the numer-
ical algorithm to compute a (functional rational expectations) equilibrium.
Section 3 relates our paper to the existing literature and summarizes theo-
retical results that can be established about the efficiency of equilibrium in
our model, for the various notions of efficiency defined before. In Section 4
we provide a first quantitative example that demonstrates the potential for
social security to provide a Pareto improvement by enhancing intergenera-
tional risk sharing and discuss whether the same Pareto improvement could
be achieved with financial innovation (the introduction of additional assets)
rather than the introduction of an unfunded social security system. Sec-
tion 5 discusses the calibration of the model and Section 6 summarizes our
main results, first for a closed and then for a small open economy version of
our model. Final conclusions are contained in Section 7 and all figures are
contained in the appendix.
2 The Economic Model

Time is discrete and extends from \( t = 0, \ldots, \infty \). Aggregate uncertainty is represented by an event tree. The root of the tree is given by some fixed event \( z_0 \). Each node of the tree is a history of exogenous shocks to the economy \( z^t = (z_0, z_1, \ldots, z_t) \). Let by \( \pi_t(z^t) \) denote the probability that the node \( z^t \) occurs.

The shocks are assumed to follow a Markov chain with finite support \( Z \) and with transition matrix \( \Pi \). There are three commodities at each node, labor, a single consumption good and a capital good which can only be used as an input to production.

2.1 Households

The economy is populated by overlapping generations of agents that live for 3 periods. The population growth rate is given by \( n \). In each period \( t \), \( L_t = (1+n)L_{t-1} \) identical new households are born. By \( L_0 = 1 \) let denote the number of newborns in period 0. A household is fully characterized by the node in which she is born \( (z^s) \). To simplify notation, we collect all households which are alive at some node \( z^s \) in a set \( I_{z^s} \) and denote a typical household by \( i \in I_{z^s} \). When there is no ambiguity about the identity of households we will index households simply by their time of birth.

An agent born at node \( z^s \) has non-negative, possibly stochastic labor endowment over her life-cycle, \( (l^0(z_s), l^1(z_{s+1}), l^2(z_{s+2})) \), that we assume to depend only on the current aggregate shock. The price of the consumption good at each date event is normalized to one and at each date event \( z^s \) the household supplies her labor endowment inelastically for a market wage \( w(z^s) \).

Let by \( c^s_t(z^s) \) denote the consumption of an agent born at time \( s \) in period \( t \geq s \). Individuals value consumption according to

\[
E_s \sum_{t=s}^{s+2} \beta^{t-s} u(c^s_t)
\]

where \( u : \mathbb{R}_{++} \rightarrow \mathbb{R} \) is assumed to be strictly increasing, strictly concave, \( C^2 \) and to satisfy the Inada condition \( \lim_{c \to 0} u'(c) = \infty \).

Households have access to a storage technology: they can use one unit of the consumption good to obtain one unit of the capital good next period. We denote the investment of household \( s \) into this technology by \( a^s_t(z^s) \). All agents are born with zero assets, \( a^s_{s-1}(z^{s-1}) = 0 \). We do not restrict
$a^t(z^t) \geq 0$, because we want to permit households to borrow against future labor income. One possible interpretation of this assumption is that there is a bank which acts as an intermediary and which stores the capital good for all households, and each individual household can then borrow from this bank. At time $t$ the household sells its capital goods accumulated from last period, $a^t_{t-1}(z^{t-1})$, to the firm for a market price $r(z^t) > 0$. We assume that the capital good cannot be converted back to the consumption good, so that it is optimal for the household to always sell all the capital goods accumulated from last period. The budget constraint of household $s$ in period $t \geq s$ therefore reads as

$$c^t_s(z^t) + a^t_s(z^t) = r(z^t)a^t_{s-1}(z^t) + (1 - \tau)I^{s-1}(z_t)w(z^t) + I(s)b(z^t)$$

where $\tau$ is the payroll tax to finance social security payments, $b(z^t)$ are the social security benefits received by a retired agent\footnote{Note that benefits $b(z^t)$ only depend on the aggregate event history, but not on individual income, whereas in the actual U.S. systems benefits do depend on individual labor earnings, although in a fairly progressive fashion. There is also a maximum income level beyond which no further social security contributions are levied. Even though our modeling choice may attribute too much intergenerational risk sharing to the social security system by failing to account for the linkage between benefits and contributions, given the progressive nature of the system it may provide a reasonable first approximation, without average lifetime income becoming an additional state variable for each generation.} and $I(s)$ is the indicator function, with $I(s = 3) = 1$ and $I(s) = 0$ for $s = 1, 2$.

To start off the economy we assume that at the root node, i.e. in period zero, there are $L_0/(1+n)^i$ households of ages $i = 0, 1, 2$ who enter the period with given capital holdings $a_{i-1}^{-0}, a_{i-1}^{-1}, a_{i-1}^{-2}$, where by assumption $a_{i-1}^{0} = 0$.

### 2.2 Firms

There is a single representative firm which in each period $t$ uses labor and capital to produce the consumption good according to a constant returns to scale production function $f_t(K, L; z_t)$. Since firms make their decisions on how much capital to buy and how much labor to hire after the realization of the shock $z_t$ they face no uncertainty and simply maximize current period profits. This allows us to abstract from problems which usually occur when one tries to incorporate production into a model with uncertainty and incomplete financial markets (see Magill and Quinzii (1996) for an overview).

In order to do so, however, one needs to make the somewhat non-standard assumption that households have access to a safe technology which
turns period $t-1$ consumption goods into period $t$ capital goods. We also
assume that households sell, rather than rent the capital to firms for con-
sumption goods.\footnote{If firms rent the capital, we have to give households access to a technology that
converts capital goods back to consumption goods as well.} These assumptions are necessary to prevent households
from consuming the capital at the beginning of the period instead of renting
it to the firm in states where the net return to capital is negative.

In the examples below we will always use the following parametric form
for the production function.

$$f_t(K, L) = \xi(z_t)F(K, (1 + g)^tL) + K(1 - \delta(z_t))$$ (3)

where $\eta(.)$ is the stochastic shock to productivity, where $\delta(.)$ may be in-
terpreted as the (possibly) stochastic depreciation rate if $\delta(.) \in [0, 1]$, and
where $F(., .)$ is a constant returns to scale production function. In order
to assure that the economy, on average, grows at rate $g$ we need to assume
that exogenous technological progress is labor augmenting, unless $F$ is of
Cobb-Douglas form. Under this assumption, since shocks to production are
multiplicative as in Diamond (1967), our model is equivalent to a model
where firm rents the capital at the beginning of the previous period and
where households have no storage technologies at their disposal.

\subsection{Government}

The only role the government has is to levy payroll taxes to pay for social
security benefits. We model social security as a defined contribution pay-as-
you-go system that adheres to period by period budget balance, with size
characterized by the payroll tax rate $\tau$. This requires that taxes and benefits satisfy

$$\tau w(z^t)L(z^t) = b(z^t)L_{t-2}$$ (4)

where $L(z^t)$ is total labor input at node $z^t$ and $L_{t-2} = (1 + n)^{t-2}$ is the total
number of retired people in the economy.

\subsection{Markets}

In this simple economy the only markets are spot markets for consumption,
labor and capital, all of which are assumed to be perfectly competitive. Oc-
casionally we will compare the equilibrium welfares to the ones obtained in
a benchmark economy with sequentially complete markets. In this framework, markets are sequentially complete when in each period $t \in \mathbb{Z}$ Arrow securities are traded.

2.5 Equilibrium

Definition 1 A competitive equilibrium, given initial conditions $z_0, (a_{s-1}^t)_{s=1}^0$, is a collection of choices for households $(c_t^s(z_t), a_t^s(z_t))_{t=1}^{s+2}$, for all $t, z_t$ for the representative firm $\{K(z_t), L(z_t)\}$, a policy $\{\tau, b(z_t)\}$ as well as prices $\{r(z_t), w(z_t)\}$ for all $z_t, t = 0, \ldots, \infty$, such that

1. Given $\{r(z_t), w(z_t)\}_{t=0}^{\infty}$, and $\{\tau, b(z_t)\}$ for all $z_t$ the choices $(c_t^s(z_t), a_t^s(z_t))_{t=1}^{s+2}$ maximize (1), subject to (2)

2. Given $r(z_t), w(z_t)$ the firm maximizes profits, i.e.

$$ (K(z_t), L(z_t)) \in \arg \max_{K_t, L_t \geq 0} f_t(K_t, L_t, z_t) - r(z_t)K_t - w(z_t)L_t \quad (5) $$

3. The government policy satisfies (4)

4. All markets clear: For all $t, z_t$

$$ L(z_t) = (1 + n)^t \sum_{s=0}^{2} \frac{I^s(z_t)}{(1 + n)^s}, \quad K(z_t) = (1 + n)^t \sum_{s=1}^{2} \frac{a_{t-s}^s(z_{t-1})}{(1 + n)^s} $$

$$ (1 + n)^t \sum_{s=0}^{2} \frac{c_t^{-s}(z_t) + a_t^{-s}(z_t)}{(1 + n)^s} = f_t(K(z_t), L(z_t), z_t) $$

Note that by Walras law market clearing in the labor and capital market imply market clearing in the consumption goods market. Note furthermore that with the assumptions on the parametric form of the production function concavity and differentiability of $F$ imply that equilibrium prices satisfy

$$ w(z_t) = (1 + g)^t \xi(z_t)F_L(K(z_t), (1 + g)^t L(z_t)) \quad (6) $$

$$ r(z_t) = \xi(z_t)F_K(K(z_t), (1 + g)^t L(z_t)) + (1 - \delta(z_t)) \quad (7) $$
2.6 Stationary Equilibrium

In order to solve for the equilibrium numerically using recursive techniques we first have to de-trend the economy by deterministic population growth and technological progress. In order to do so we now make the assumption that the period utility function is of CRRA form

\[
    u(c) = \begin{cases} 
    \frac{c^{1-\sigma}}{1-\sigma} & \text{if } \sigma \neq 1 \\
    \log(c) & \text{if } \sigma = 1
    \end{cases}
\]

Defining, for an arbitrary agent

\[
    \gamma_t^s(z^t) = \frac{c_t^s(z^t)}{(1+g)^t}
\]

\[
    \alpha_t^s(z^t) = \frac{a_t^s(z^t)}{(1+g)^t},
\]

the lifetime utility function of an agent born into node \( z^s \) can be rewritten as

\[
    U(c, z^s) = \sum_{t=s}^{s+2} \sum_{z^t | z^s} \beta^{t-s} \pi_t(z^t | z^s) \left[ \frac{c_t^s(z^t)}{(1+g)^t} \right]^{1-\sigma} \frac{1}{1-\sigma}
\]

\[
    = (1+g)^{1-\sigma} \sum_{t=s}^{s+2} \sum_{z^t | z^s} \left( (1+g)^{1-\sigma} \beta \right)^{t-s} \pi_t(z^t | z^s) \left[ \gamma_t^s(z^t) \right]^{1-\sigma} \frac{1}{1-\sigma}
\]

\[
    = A_0 \sum_{t=s}^{s+2} \sum_{z^t | z^s} \beta^{t-s} \pi_t(z^t | z^s) \left[ \gamma_t^s(z^t) \right]^{1-\sigma} \frac{1}{1-\sigma} = \tilde{U}(\gamma, z^s)
\]

where \( \tilde{\beta} = (1+g)^{1-\sigma} \beta \). The budget constraint can be rewritten as, dividing by \( (1+g)^t \)

\[
    \gamma_t^s(z^t) + \alpha_t^s(z^t) = r(z^t) \frac{\alpha_t^{s-1}(z^t)}{1+g} + (1-\tau) \beta^{t-s}(z_t) \omega(z^t) + I(s) v(z^t)
\]

where \( \omega(z^t) = \frac{w(z^t)}{(1+g)^t} \) is the growth-adjusted wage rate and \( v(z^t) = \frac{b(z^t)}{(1+g)^t} \) is the growth-adjusted social security benefit.

In the firms problem, defining

\[
    l(z^t) = \frac{L(z^t)}{(1+n)^t}
\]

\[
    \kappa(z^t) = \frac{K(z^t)}{(1+g)^t L(z^t)} = \frac{K(z^t)}{[(1+g)(1+n)]^t l(z^t)}
\]
we rewrite (6) and (7) as
\[
\omega(z^t) = \xi(z_t)F_L(\kappa(z^t), 1) \quad (8)
\]
\[
r(z^t) = \xi(z_t)F_K(\kappa(z^t), 1) + (1 - \delta(z_t)) \quad (9)
\]
The equilibrium conditions become
\[
l(z^t) = \sum_{s=0}^{2} \frac{l^s(z_t)}{(1 + n)^s} \quad (10)
\]
\[
(1 + n)(1 + g)\kappa(z^t)l(z^t) = \sum_{s=1}^{2} \frac{\alpha_{t-s}^t(z^{t-1})}{(1 + n)^{s-1}} \quad (11)
\]
\[
\sum_{s=0}^{2} \frac{\gamma_{t-s}^t(z^t) + \alpha_{t-s}^t(z^t)}{(1 + n)^s} = l(z^t) \left[ F(\kappa(z^t), 1) + (1 - \delta(z^t)) \right] \quad (12)
\]
Finally, the government budget constraint becomes, dividing both sides by \([(1 + n)(1 + g)]^t\)
\[
\tau \omega(z^t)l(z^t) = \frac{v(z^t)}{(1 + n)^2}
\]
The Euler equations from the individuals’ optimization problem read as
\[
\left( \gamma_{t}^s(z^t) \right)^{-\sigma} = \beta \sum_{z_{t+1}} \pi(z_{t+1}|z_t) \frac{r(z^t, z_{t+1})}{1 + g} \left( \gamma_{t+1}^s(z^t, z_{t+1}) \right)^{-\sigma} \quad (13)
\]
where \(r(z^t, z_{t+1})\) is given from (9), with \(\kappa\) and \(l\) defined from (10) and (11). Our numerical algorithm will operate on the recursive version of these Euler equations.

2.7 Functional Rational Expectation Equilibrium

In order to compute equilibrium allocations numerically we have to reformulate the Euler equations recursively. In order to make precise the object we are actually computing we now define a Functional Rational Expectations Equilibrium (FREE), following the approach of Spear (1988).

A FREE is abstractly defined as a collection of intervals \([\underline{\alpha}_i, \overline{\alpha}_i]\), for \(i = 1, 2\) and smooth functions \(\theta_i^t : [\underline{\alpha}_1, \overline{\alpha}_1] \times [\underline{\alpha}_2, \overline{\alpha}_2] \rightarrow [\underline{\alpha}_i, \overline{\alpha}_i]\) for \(i = 1, 2\) and all possible shocks \(z \in \mathcal{Z}\), which map the current distribution of asset holdings into asset holdings tomorrow for an agent of age \(i\), given that today’s aggregate shock is \(z\). For these intervals and functions to be a
FREE it has to be the case that for all initial $z \in \mathcal{Z}$ and for all $(\alpha_{t-1}, \alpha_{t-2}) \in [\underline{\alpha}_1, \overline{\alpha}_1] \times [\underline{\alpha}_2, \overline{\alpha}_2]$ there exists a competitive equilibrium which satisfies

$$a_t^s(z^t) = \theta_t^{t-s}(a_{t-1}^{t-1}(z^{t-1}), a_{t-2}^{t-2}(z^{t-1}))$$

for all $z^t$ and all $s = t - 1, t$. Note that, since agents don’t leave bequests, agents of age 3 choose $\alpha_t^{t-3}(z^t) \equiv 0$ and thus we do not have to solve for their optimal capital holdings explicitly.

Since agents face a finite dimensional, convex optimization problem the Euler equations are necessary and sufficient for agents’ optimality. A functional rational expectations equilibrium can therefore be characterized as the solution to a functional equation derived from the recursive version of the Euler equations (13). Given policy functions $(\theta_z^i)_{i=1,2}$ for assets saved for next period, with $\theta_z^1(.) = 0$ and some vector of asset holdings at the beginning of the current period $\Theta = (\Theta_1, \Theta_2) \in [\underline{\Theta}_1, \overline{\Theta}_1] \times [\underline{\Theta}_2, \overline{\Theta}_2]$, let $\Theta_0 = 0$ and define (normalized) consumption of agent $i$ as

$$\bar{c}_i(\Theta, z) = \frac{\Theta_{i-1}r(\Theta, z)}{1 + g} + (1 - \tau)l^i(z)\omega(\Theta, z) - \theta_z^i(\Theta) + I(s)v(\Theta, z) \quad \text{for } i = 1, 2, 3$$

Also, denote by

$$\Theta_+ = (\theta_z^1(\Theta), \theta_z^2(\Theta)) \quad \text{with } \Theta_0 = 0$$

the vector of asset holdings at the beginning of the next period, consistent with the policy functions $(\theta_z^i)_{i=1,2}$ as well as current asset holdings $\Theta$. Using the asset market clearing condition in the firms optimization conditions we obtain prices, as functions of the current state, as

$$r(\Theta, z) = \xi(z)F_K(\kappa(\Theta, z), 1) + (1 - \delta(z))$$

$$\omega(\Theta, z) = \xi(z)F_L(\kappa(\Theta, z), 1)$$

where

$$\kappa(\Theta, z) = \frac{1}{(1 + n)(1 + g)l(z)} \sum_{i=1}^2 \frac{\Theta_i}{(1 + n)^i}$$

with

$$l(z) = \sum_{s=0}^2 \frac{l^i(z)}{(1 + n)^s}$$

Given the interval bounds $[\underline{\alpha}_i, \overline{\alpha}_i]$ a FREE is then a list of functions $(\theta_z^i)_{i=1,2}$ that solve the set of functional equations

$$(\bar{c}_i(\Theta, z))^{-\sigma} = \tilde{\beta} \sum_{z'} \pi(z'|z) \frac{r(\Theta, z)}{1 + g} (\bar{c}_{i+1}(\Theta_+, z'))^{-\sigma} \quad \text{for all } i = 1, 2 \quad (14)$$

and for all $\Theta \in [\underline{\Theta}_1, \overline{\Theta}_1] \times [\underline{\Theta}_2, \overline{\Theta}_2]$. 

11
2.8 Welfare and Efficiency of Competitive Equilibrium

In order to make precise the welfare consequences of different social security reforms we now define several efficiency criteria along which we compare allocations arising from these reforms. Since lifetime utilities in the original and the growth-adjusted equilibrium differ by a monotonically increasing transformation, comparing either allocations will yield the same welfare conclusions, so we focus on the original allocations \( c \), keeping in mind that our computational algorithm delivers growth-deflated consumption processes \( \gamma \).

Define as expected utility of an agent born into event history \( z^t \)

\[
U(z^t, c) = u(c_t(z^t)) + \sum_{s=1}^{2} \sum_{z^{t+s}|z^t} \beta^s \pi(z^{t+s}|z^t) u(c_{t+s}(z^{t+s}))
\]

As we mentioned in the introduction we want to focus on the notion of ex interim (conditional) efficiency in this paper. First we define a feasible allocation

**Definition 2** An allocation \( (c, K) \) is feasible if \( c \geq 0, K \geq 0 \) and

\[
C(z^t) + K(z^{t+1}) = f_t(K(z^t), L(z^t), z^t)
\]

for all \( t \) and \( z^t \), where

\[
L(z^t) = (1 + n)^t \sum_{s=0}^{2} \frac{l^s(z^t)}{(1 + n)^s}
\]

is aggregate, exogenously given labor supply and

\[
C(z^t) = (1 + n)^t \sum_{s=0}^{2} \frac{c^{t-s}(z^t)}{(1 + n)^s}
\]

is aggregate consumption.

Now we can define our notion of efficiency.

**Definition 3** A feasible allocation \( (c, K) \) is ex interim efficient if there is no other feasible allocation \( (\hat{c}, \hat{K}) \) such that

\[
U(z^t; \hat{c}) \geq U(z^t; c) \text{ for all } t, z^t
\]

\[
U(z^t; \hat{c}) > U(z^t; c) \text{ for some } t, z^t
\]
A stronger notion of efficiency would be to consider ex ante efficiency as for example in Bohn (2000). However, as mentioned in the introduction we want to focus on the ex interim notion, as any policy reform that provides a Pareto improvement in an ex-interim sense, evidently provides a Pareto improvement in the ex-ante sense. That is, by considering ex-interim efficiency we demand a lot from a particular policy innovation; if we can in fact find such a policy reform, we would conclude that its welfare benefits are not subject to the discussion of having specified the right welfare criterion.\(^3\)

### 2.9 The Thought Experiment

It is well known that in our setup equilibrium allocations are generally suboptimal and that a central planner could find Pareto-improving transfers. However, when discussing possible risk-sharing benefits of a pay-as-you-go social security system we are focusing on a particular intervention. We certainly do not argue that a social security system guarantees full efficiency and we do not attempt to explain why this particular system is in place. We do want to examine if it can be Pareto-improving to introduce social security.

In this paper we will be interested in the following comparative statics exercise: Suppose that in an equilibrium of the economy for a payroll tax rate \(\tau = 0\) at some date-event \(z^t\), there is an unanticipated increase of \(\tau\). What are the welfare effects for all individuals born at \(z^t\) and at all successor nodes?

In order to determine whether an introduction of a social security system (i.e. an increase of \(\tau\) to a positive value) is ex interim Pareto-improving, one needs to compare welfare at infinitely many nodes. In the examples below, we report welfare gains for the next 3 periods (which corresponds to 60 years) and verify that the qualitative conclusions remain the same over 5-6 periods. It turns out that in our examples we do not have to consider more periods.

In any F.R.E.E., the welfare of a newly born depends solely on the aggregate state (i.e. exogenous shock, aggregate capital and the share of capital held by the middle aged) when born. Clearly the endogenous state at \(z^t\) will only depend on \(z_{t-1}\) as well as on the aggregate state at node \(z^{t-2}\).

\(^3\)Reversely, of course, one can argue that the failure of a particular policy reform to generate an ex-interim Pareto improvement may just reflect an overly ambitious welfare criterion rather than a shortcoming of the reform. The discussion to follow should be read with this caveat in mind.
In all the examples we consider an introduction of social security leads to a reduction of capital accumulation. After a transition, which takes around 3-4 periods, the effects of $K(z^{t-2})$ on the current state at $z^t$ are extremely small and can be neglected. In this case, only the current shock and the two previous shocks matter, and, for all practical purposes, social security is Pareto-improving, if and only if it leads to an improvement in welfare for the first 6 periods after its introduction.

3 Incomplete Markets and Intergenerational Risk Sharing

3.1 Is a Pareto Improvement through Social Security Possible?

Ever since Samuelson (1958) and Diamond (1965) it is well known that overlapping generation models can exhibit Pareto suboptimal equilibria. In exchange economies, transfers from young to old agents can be Pareto-improving. In economies with production a reduction in capital accumulation can lead to Pareto improvement through higher aggregate consumption at all future dates. However, the introduction of uncertainty into the basic model adds a variety of complications (even if one focuses on the notion of interim optimality as described above).

Zilcha’s (1990) concept of production efficiency is a natural generalization of Diamond’s characterization of dynamic efficiency in production: an equilibrium is production efficient if there do not exist alternative production plans which lead to higher aggregate consumption at all nodes in a subtree. However, as Barbie et al. (2000) point out, efficiency in production does not necessarily imply Pareto efficiency and there can be room for Pareto-improving interventions even if the allocation is production efficient. It is therefore useful to distinguish between production inefficiency and dynamic inefficiency in the sense of Samuelson (1958): even if an equilibrium is production efficient, there is a theoretical possibility that social security is Pareto improving because the allocation is not dynamically efficient.

When markets are sequentially complete, however, a sufficiently high return on available assets can be shown to rule out dynamic inefficiency. In this case equilibrium allocations are Pareto optimal and social security can never be Pareto improving. Judging from available data on past returns it is commonly believed that the US economy is dynamically efficient (see e.g.
Abel et al. (1989)). Moreover (and perhaps more convincingly) an asset that promises to pay a non-negligible fraction of aggregate consumption at each state of the world in the future can only have a finite price today if the allocation is dynamically efficient. Since land can be interpreted as such an asset (see Demange (2001)) it seems empirically implausible to consider dynamically inefficient economies.

However, the conclusion that dynamic efficiency implies Pareto efficiency obviously depends crucially on the assumption that markets are sequentially complete. If markets are incomplete allocations will generally not be Pareto efficient, even if land has a finite price and returns of assets are high.

In this paper we focus on economies where markets are not sequentially complete. While it is well known that equilibria will generally not be Pareto efficient, it is unknown whether one-sided intergenerational transfers (like social security) can possibly lead to Pareto-improvements in these models.

We want to present examples where, judging from the rates of return on capital, the allocation is dynamically efficient (i.e. the rates of return would pass Abel et al.’s (1989) test and the price of land could be finite) but where the introduction of a social security system is Pareto improving. For this, we need to review the conditions which imply dynamic efficiency and demonstrate why there can be room for a Pareto improvement through enhanced intergenerational risk sharing even if the allocation is production efficient and dynamically efficient.

### 3.2 Theoretical Results

Demange (2001) generalizes the notion of dynamic efficiency to economies with uncertainty and possibly incomplete security markets.\(^4\) A given sequence of aggregate capital stocks \( (K(z_t)) \) defines the set of achievable allocations. We call an allocation dynamically efficient, if there exists no other allocation in the marketed subspace which constitutes a Pareto improvement. While in finite horizon models or in models with infinitely lived agents, competitive equilibria will always be dynamically efficient, Samuelson’s argument for possible inefficiency can be applied when consider a model with overlapping generations.

\(^4\)Following the GEI-literature (see e.g. Geanakoplos and Polemarchakis (1986)) Demange refers to this concept as C-optimal (‘constrained efficiency’) - however, since a social security system can lead to intergenerational transfers which lie outside of the marketed subspace of the consumption set it can improve on allocations which Demange calls C-optimal and we prefer to call these allocation dynamically efficient.
In the following we characterize dynamically efficient allocations for our economy (the analysis follows directly from Demange (2001) - the only difference is that we consider homothetic economies with growth in population and productivity).

For each value of the shock \( z \), define a production function in intensive units by

\[
\phi(\kappa; z) = \xi(z) F(\kappa, 1) - (1 + \delta(z)) \kappa
\]

Define a supporting price system \( (q(z_t)) \) by \( q(z_0) = 1 \) and

\[
E(q(z^t)) \frac{\partial \phi(\kappa(z^t), z)}{\partial \kappa}|_{z^{t-1}} = q(z^{t-1})(1 + n)(1 + g)
\]

Evidently, since markets are not sequentially complete there are several supporting price systems. We collect them in a set \( Q \). The following proposition (Theorem 1 in Demange (2001)) characterizes dynamic efficiency:

**Proposition 4** An equilibrium allocation is dynamically efficient if

\[
\lim_{t \to \infty} \inf_{q \in Q} E_0 \left( \sum_{s=t}^{t+2} q(s) \right) = 0
\]

The proposition is remarkable since it is sufficient for optimality that the infimum over all supporting prices tends to zero. This implies that we can verify dynamic efficiency if we find some supporting price system that satisfies condition (15).

For general economies this condition can obviously not be easily verified since it involves prices ‘at infinity’. However, in our framework we focus on functional rational expectations equilibria as defined in Section 2.7 above and a sufficient condition for dynamic efficiency is the following. For a given time horizon \( T \) define the \( T \) - expected discounted present value by

\[
R(T) = E \left[ \Pi_{s=1}^{T} \frac{(1 + n)(1 + g)}{1 + r(z^s)} \right].
\]

If for large enough \( T \) this present value is less than 1, the associated equilibrium is dynamically efficient. The following lemma makes this sufficient condition precise.

**Lemma 5** A FREE is dynamically efficient if there exists a \( T > 0 \) such that for all initial conditions

\[
(z, \Theta) \in Z \times [\alpha_1, \bar{\alpha}_1] \times [\alpha_2, \bar{\alpha}_2]
\]
the resulting equilibrium returns satisfy

\[ R(T) < 1. \]

**Proof.** By definition of a FREE all \( z_T, \Theta(z^T) \) will lie in \( Z \times [\alpha_1, \bar{\alpha}_1] \times [\alpha_2, \bar{\alpha}_2] \) themselves and can be viewed as initial conditions as well. Therefore it follows that \( R(iT) \to 0 \) as \( i \to \infty \) Defining

\[ \tilde{q}(z^{t+1}) = \frac{(1 + n)(1 + g)\tilde{q}(z^t)}{1 + r(z^{t+1})} \]

then implies the sufficient condition (15). QED

In applications below it often suffices to consider \( T = 2 \). Note that with Jensen’s inequality for \( T = 1 \), the lemma implies that the allocation is dynamically efficient if the conditional expected returns to capital lies above \( (1+r)(1+n) \) for all possible states in the invariant set. While this is a strong condition that does not hold in all examples, it actually already holds true for the calibrated examples in Section 6.

When markets are sequentially complete, dynamic efficiency implies full Pareto efficiency. It follows directly from Zilcha (1990) that independently of the market structure allocations are also production efficient if (15) holds true. In the examples below we show that without complete markets allocations might be dynamically efficient but not Pareto efficient and that the introduction of social security might be Pareto-improving because it enhances intergenerational risk sharing.

As mentioned above, we focus on dynamically efficient allocations for two reasons. Observed returns on risky capital seem to indicate that the US economy is dynamically efficient (e.g. Abel et al. (1996)). Secondly, as Demange (2001) shows, the presence of ‘land’, i.e. an infinitely lived assets which pays out a non-negligible fraction of aggregate consumption at all states and times ensures that condition (15) must hold.

Even though we do not model land explicitly in our analysis, it is important to point out that our result does not crucially depend on the absence of ‘land’. Although a finite price of land implies dynamic efficiency, this does not imply full Pareto efficiency in our setup. While it is well known that in models with two-period lived agents (and no within generation heterogeneity) land implies full efficiency even when markets are incomplete (see e.g. Demange (2001)), it is subject to future research to examine if there are realistic cases where there is room for intergenerational risk sharing in economies with land.
Finally, a possible caveat in the interpretation of dynamic inefficiency has to be pointed out: when markets are sequentially complete, whether or not the allocation is dynamically efficient cannot be judged by the returns to risky capital alone. In order to verify whether conditions (15) holds one has to recover the unique supporting prices \( q(z^t) \) from the knowledge of returns to all securities. In particular, it is very well possible that with sequentially complete markets the allocation is dynamically inefficient although the return process to risky capital satisfies the condition in Lemma 1. Furthermore, as we show below, it is possible that the allocation was dynamically efficient with incomplete markets while the introduction of new securities leads to higher savings rates which ‘push’ the economy into a dynamically inefficient region.

4 Pareto Improvement from Social Security: An Example

In this section we present an example where even though the economy is dynamically efficient the introduction of an unfunded social security system leads to significant welfare improvements. The example considered is not meant to be a description of the US economy but simply an illustration of our main point.

For simplicity we assume that there is no population or technology growth.

The production function is of the Cobb-Douglas form with a labor share of \( \alpha = 0.3 \). There are two possible exogenous shocks which are i.i.d. with probability 1/2. They only affect depreciation \( \delta \). We assume

\[
\delta(z_1) = 1.0, \quad \delta(z_2) = 0.2 \quad \text{and} \quad \xi(z_1) = \xi(z_2) = 15
\]

Individuals are endowed with 6 units of the consumption good when young and 4 units when middle-aged. They have no endowments when old. Agents’ risk aversion is \( \sigma = 4 \) and they do not discount the future, \( \beta = 1 \).

4.1 Welfare Consequences from the Introduction of Unfunded Social Security

In this economy, without social security, the conditional expected returns to capital always lie above the population growth rate \( n \) plus the average growth rate of the economy \( g \).
We consider an introduction of social security with a social security tax of 8 percent, \( \tau = 0.08 \). Welfare gains (in terms of wealth equivalent) for generations born after introduction in shock 1 are summarized in Figure 1, where at each node of the event tree we indicate the sequence of shocks and the welfare gains in percent.\(^5\)

The figure shows that even 5 periods after the introduction of social security, the welfare gains are around 3 percent, at which level they stabilize for future generations.

4.2 Financial Innovations versus Unfunded Social Security

One could argue that instead of introducing an unfunded social security system the introduction of new securities (e.g. riskless borrowing) is a better policy recommendation since it leads to sequentially complete markets and possibly full Pareto efficiency.

However, things are not this simple. In our example, the introduction of an additional security (which sequentially completes the markets, since the aggregate shock can only take two values) actually leads to huge welfare losses for all future generations, as shown in Figure 2 (the current old and middle aged gain as usual). We assume that instead of introducing social security, at the same date-event an additional asset, namely a risk-free bond is introduced.

The reason for this counterintuitive result is simple: With incomplete markets, capital is an asset with very bad risk-characteristics. With complete markets, a large part of this risk can be diversified away and the middle aged invest huge amounts in this asset to save for retirement.

It should be noted, however, that this part of the example crucially depends on the absence of land in the model. The allocation after the completion of financial markets is no longer dynamically efficient. We now move to more realistic examples, where the effects turn out to be ambiguous.

5 Calibration

In order to parameterize the model we have to choose the following parameters. We have to characterize the aggregate stochastic process govern-\[^5\]We measure welfare gains by asking what percentage increase in consumption in each date, each node of the event tree would be required to make an agent indifferent between the pre-policy change allocation with the additional consumption and the post-policy reform allocation.
ing total factor productivity and stochastic depreciation, population growth and the life-cycle labor income profile, average economic growth, the capital share in the production function and parameters governing preferences.

5.1 Aggregate Growth and Technology

As population growth rate we choose \( n = 1.1\% \) per annum, which equals the average population growth rate for the US postwar period. Similarly we choose the average growth rate of GDP per hour worked equal to \( g = 1.6\% \), the long-run average for the US. The labor share in the Cobb-Douglas production function is taken to be \( \alpha = 0.3 \).

We assume that aggregate uncertainty is driven by a four-state Markov chain with state space \( Z = \{z_1, z_2, z_3, z_4\} \) and transition matrix \( \pi = (\pi_{ij}) \). Since we want to model both shocks to total factor productivity and to depreciation, a particular state \( z_i \) maps into a combination of low or high TFP and low or high depreciation.

\[
T(z) = \begin{cases} 
1.0 + \nu & \text{for } z \in \{z_1, z_2\} \\
1.0 - \nu & \text{for } z \in \{z_3, z_4\}
\end{cases}
\]

\[
\delta(z) = \begin{cases} 
\bar{\delta} - \psi & \text{for } z \in \{z_1, z_3\} \\
\bar{\delta} + \psi & \text{for } z \in \{z_2, z_4\}
\end{cases}
\]

Here \( \bar{\delta} \) is the average depreciation rate. We set \( \bar{\delta} = 0.7 \), reflecting an average depreciation rate of 6% per year.

We assume that the stochastic process is iid over time. Although it is well-established that aggregate technology shocks are positively correlated at a quarterly or yearly frequency, we are not aware of conclusive evidence indicating such positive serial correlation over 20 year periods. Therefore, as a benchmark, we assume that technology shocks are uncorrelated across time (see also Smetters 2001). Given this assumption the transition matrix is determined by the probabilities \( \pi_i, i = 1, \ldots, 4 \) for a particular aggregate state.

The aggregate state \( z_1 \) is characterized by a good TFP-shock and a good depreciation shock (low depreciation), whereas \( z_4 \) features a bad TFP shock and a bad depreciation shock. In aggregate states \( z_2 \) and \( z_3 \) the TFP-shock and depreciation shock move in opposite direction. We make a symmetry assumption in that \( \pi_1 = \pi_4 \) and \( \pi_2 = \pi_3 \). Given the restriction \( \sum_i \pi_i = 1 \) the matrix \( \pi \) is then uniquely determined by a single number \( \pi_1 \).
The remaining technology parameters to be calibrated to selected observations from US data are thus \((\nu, \psi, \pi_1)\). Their selection will be discussed below.

5.2 Preference Parameters and Labor Endowments

Labor endowments are deterministic and set as \(l^0 = 0.485, l^1 = 0.515\) and \(l^2 = 0\), so that young agents have labor income of 94\%, relative to middle-aged agents, consistent with Hansen (1993), who reports average labor efficiency units of individuals of age 35, relative to those of age 55, of 94\%. Old agents are assumed to retire.

As benchmark value for the coefficient of relative risk aversion we choose \(\varpi = 1\); this value is at the low end of the values commonly used in the macroeconomic and public finance literature (see Imrohoroglu et al. (1995) for an overview) and our welfare numbers thus provide a lower bound for the potentially beneficial intergenerational risk sharing effect of social security. At times we will also report results for \(\varpi = 4\), a value for risk aversion at the higher end of the range commonly employed in the literature.

The time discount factor \(\beta\) will be chosen jointly with \((\nu, \psi, \pi_1)\) to selected US data.

5.3 Social Security

As policies we consider various sizes of the social security system. We assume that all working agents pay a fixed payroll tax \(\tau \geq 0\); benefits are then adjusted to preserve period-by-period budget balance, given the payroll tax rate. We consider various sizes of the social security system, with a benchmark of \(\tau = 0\) (no social security).

5.4 Remaining Parameters and Data

The remaining parameters \((\nu, \psi, \pi_1, \beta)\) are chosen jointly so that the benchmark model competitive equilibrium delivers the following statistics, discussed in detail in Constantinides et al. (2001) and Smetters (2000), which also consider OLG models with aggregate uncertainty, where agents live for only a small number of periods and thus a model period has to be interpreted as 20 years. Note that the first 2 facts are motivated by the discussion in Constantinides et al. (2001), the second two facts by Smetters (2001).

1. An average return on risky capital of about 6.5\% per annum
2. A coefficient of variation of aggregate output of about 0.2

3. A correlation coefficient between wages and returns to capital of about 0.7

4. A coefficient of variation for the return of capital of about 0.87

Loosely speaking, the parameter $\beta$ determines the average return on capital, the shock to TFP, $\nu$, determines the variability of aggregate output, conditional on $\nu$ the shock to depreciation $\psi$ determines the variability of interest rates and the probability $\pi_1$ determines how correlated returns to capital and labor are.\footnote{\textsuperscript{6}Of course it is understood that all four parameters ($\nu, \psi, \pi_1, \beta$) jointly determine all four model statistics.}

The parameters required for model-generated statistics to coincide with the four empirical observations stated above are $(\nu, \psi, \pi_1, \beta) = (0.15, 2.13, 0.499, 0.46)$. We make the following observations. First, the time discount factor $\beta = 0.46$ implies an annual time discount rate of 4%. Second, the fact that $\pi_1 = \pi_4 = 0.499$ is required comes from the fact that returns to labor and capital are highly positively correlated for 20 year time periods in the data. In order for the model to reproduce this observations it has to be very unlikely that TFP-shocks and depreciation shocks of opposite direction occur simultaneously. The relative magnitude of TFP-shocks and depreciation shocks is explained by the fact that returns to risky capital are much more volatile in the data than is aggregate output. Since TFP-shocks affect both returns as well as output directly, the size of these shocks have to be somewhat moderate for output not to be too volatile. Given this, depreciation shock have to be of large magnitude to generate returns to capital that are sufficiently volatile in the model.

The required parameter values do not seem to be implausible, with the exception of the high variance of the depreciation shock. In particular, the size of $\psi = 2.13$ implies that the depreciation rate can be bigger than 100% and smaller than 0%, which makes an economic interpretation of these shocks as depreciation shocks problematic. Alternatively these shocks may be interpreted as shocks to the aggregate production function that do not affect real wages. These shocks are required to be large in order to generate returns to capital with sufficient volatility. Table 1 summarizes the parameterization of the model.
Table 1: Parameterization

<table>
<thead>
<tr>
<th>Par.</th>
<th>( n ) (pa)</th>
<th>( g ) (pa)</th>
<th>( \alpha )</th>
<th>( \pi_1 )</th>
<th>( \nu )</th>
<th>( \psi )</th>
<th>( \delta )</th>
<th>(( \rho_1, \rho_2, \rho_3 ))</th>
<th>( \beta ) (pa)</th>
<th>( \sigma )</th>
<th>( \tau )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val.</td>
<td>1.1%</td>
<td>1.6%</td>
<td>.3</td>
<td>.499</td>
<td>.15</td>
<td>2.13</td>
<td>.7</td>
<td>(.485, .515, 0)</td>
<td>.96</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

6 Results

6.1 The Closed Economy

We now document what the welfare consequences are from unexpectedly introducing a pay-as-you-go social security system into the economy, with size \( \tau = 0.02 \), the initial size of the US social security system at its introduction in 1935.

In the economy with social security the average aggregate capital stock is 9.2% lower than in the benchmark without social security, and average output and wages are 3.1% lower. Figure 3 summarizes the welfare consequences from the reform, if the reform is introduced at a point where the aggregate capital stock is at the high end of the support of the ergodic prereform capital distribution. As in Figure 1 at each node of the event tree we document the welfare gains, measured as percentage consumption equivalent variation, from the reform for a generation that is born at that node; for example \(-1.4\) in node \((2, 4, 3)\) indicates that at agent born in that node would require 1.4% more consumption in each contingency of her life with social security being introduced at node \( z = z_2 \) to be as well off, in terms of expected lifetime utility, as under the situation where social security was not introduced at all. For the initial nodes we also report consumption equivalent variation for generations already alive at the time of the reform.

Not surprisingly all generations already alive benefit from the introduction of the system since they receive benefits without having (fully) contributed. The initial old benefit strongest in states where returns to the capital they have accumulated for retirement are low.

The current middle-aged generation also uniformly benefits from the introduction of social security, because it receives full pensions and has to contribute only in one period. The ranking of states in which this generation benefits most is subtle, however. On one hand states with currently low wages make social security a good deal because small contributions (in absolute terms) stand against average pensions (because of the iid nature of the

\[7\]Conditional on returns, they also benefit from higher current wages because of higher pensions.
shocks). On the other hand there is a strong general equilibrium effect: low wages today mean low savings and thus a low capital stock tomorrow, with high average returns to capital tomorrow which make the implicit return to social security look relatively bad. The same is true for high depreciation shocks today that reduce the capital stock for tomorrow. This effect is sufficiently strong for middle-aged agents to prefer the introduction of social security more in high-wage states $z_1, z_2$ than in low-wage states $z_3, z_4$ and in low-depreciation states $z_1, z_3$ than in high-depreciation states $z_2, z_4$.\footnote{In other words, if recessions are \textit{iid} events, then, ceteris paribus, it is recessions where investments into private capital have high expected returns and social security does not look as attractive.}

Generations born at the time of the reform benefit slightly (between 0.004% and 0.3%) from the introduction of social security even though they face full contributions for their benefits, whereas generations born after the reform suffer losses between 0.8% and 2.2% in terms of consumption equivalent variation.\footnote{For generations born later than 3 periods after the reform welfare losses stabilize between 2% and 2.2%, depending on the current aggregate shock, since, as explained in Section 2, the (stochastic) transition of the economy without social security to one with social security is completed.} The reason for the difference between generations born today and later is explained by endogenous capital accumulation: the introduction of social security reduces private saving and thus capital accumulation, private consumption and wages along the transition to a new stochastic steady state. Since this transition takes some time, this detrimental effect does not hit the current generation in its full extent, but is experienced by future generations to a larger and larger extent, until the transition period is completed and welfare effects from the reform converge.

Finally we observe that the welfare consequences of an introduction of social security are ranked by aggregate shocks as being most beneficial (least harmful) in $z_2$, then in $z_1$, followed by $z_4$ and then $z_3$. The reason for this ranking for newborns is similar as for the current middle-aged, with an added subtlety, since newborns still have 2 more periods to live. A negative shock to current \textit{total factor productivity} affects both the current young and middle-aged agents negatively, depresses savings for both generations, and thus tomorrow’s capital stock is low and the return on capital is high tomorrow. Tomorrow the then middle-aged agents have saved little, therefore save relatively little for the last period of their life and the capital stock is low and returns are high in their last period of life. Thus a bad TFP shock today is associated with high returns to private capital in all future periods of a
currently young generations’ life, which makes social security look relatively bad in these aggregate states (ζ3 and ζ4). On the other hand, a large depreciation shock reduces the capital stock tomorrow and generates high returns tomorrow; but since the current young now save a lot (high returns, current wages unaffected by depreciation shock) the capital stock rebounds quickly and is at average level when the current young are old, so that returns from capital are only mediocre when they retire. So a bad depreciation shock today yields high returns and low wages tomorrow and average returns in the period after that, whereas a good depreciation shock yields low returns and high wages tomorrow and average returns after this. The wage effect dominates and agents benefit from social security more in states with bad depreciation shocks (ζ2, ζ4) than with good depreciation shocks.10

In order to assess the dependence of the welfare consequences from the introduction of social security on the aggregate state of the world, in Figure 4 we report welfare numbers for the same experiment as before, but with the initial capital stock at the time of the reform being at the low, rather than the high end of the pre-reform ergodic capital distribution.

One observes, comparing Figures 3 and 4, that qualitatively the results are somewhat similar, but that the introduction of social security is more favorable to current young and middle-aged generations with initially high, rather than initially low capital stock, since private returns to capital are fairly low (but not dynamically inefficiently low) with a high capital stock and thus social security does not look so unattractive with respect to its implicit return on retirement saving. The figures also show, however, that the initial capital stock hardly plays any role for generations born even one period after the introduction of social security. In summary, judged from its welfare implications a social security reform seems most promising in aggregate conditions characterized by a high aggregate capital stock, bad shocks to the returns to capital and relatively high real wages.

6.2 A Small Open Economy

In our model the beneficial intergenerational risk sharing role associated with the introduction of unfunded social security has to be traded off with the adverse effect on capital accumulation. In a small open economy the latter effect is absent and the effect of social security on the allocation of

10Note that, with respect to depreciation shocks, young agents have a reverse ranking of aggregate shocks than middle-aged agents, since the latter don’t experience wage effects in the next period because they are old then and don’t work.
aggregate risk among generations can be analyzed most clearly.

Therefore in this section we assume that the total supply of capital is independent of the aggregate domestic savings in the economy. The definition of a competitive equilibrium is as in Definition 1 above, with the crucial difference that we replace the capital market clearing condition by

\[ K(z^t) = \bar{K} \]

for some exogenously given supply of capital \( \bar{K} \). The definition of a FREE is analogous.\(^\text{11}\)

Introducing unfunded social security in this world lowers the average return on savings for retirement with capital from 6.5\% per annum to \((1 + \eta)(1 + g) - 1 = 2.7\%\) per annum for retirement saving via the social security system. However, social security pools labor and capital income risk across generations since returns to labor and capital are imperfectly correlated and thus reduces the variance of consumption for each generation. The welfare consequences of introducing social security then depend on the relative importance of mean consumption and the riskiness of consumption. Note that under the small open economy assumption the mean and variance of returns to capital and labor are independent of the preference parameters \((\sigma, \beta)\). Therefore, for a given \(\beta\), since \(\sigma\) controls risk aversion there exists a one-to-one mapping between \(\sigma\) and the welfare consequences from the introduction of social security, with higher degrees of risk aversion unambiguously associated with higher welfare gains.

In order to make the results in this section comparable to those in the previous section we re-calibrate the model to the same observations. We choose the same \(\beta = 0.46\) as in the previous section (note that \(\beta\) does not affect the return to saving anymore) and pick \(\bar{K}\) so that the annual mean return on capital is 6.5\% per annum. In order to match the same volatility of output, return to capital and correlation between labor and capital returns as before we choose \((\nu, \psi, \pi) = (0.2, 1.55, 0.499)\).

In Table 2 we summarize the welfare consequences of the introduction of social security of size \(\tau = 2\%\) for different degrees of risk aversion. Note that, absent effects of capital accumulation and given the \(iid\) nature of aggregate shocks all generations born after the introduction of the system experience the same welfare consequences as the generation born at the time of the reform, given the same aggregate shock \(z\). Furthermore, since agents are

\(^{11}\)Note that an alternative interpretation that leads to the same results is to assume a production technology that is linear in capital.
born with no capital, the welfare consequences differ only across states for which returns to labor differ, but not across states for which returns to capital differ.\footnote{The welfare consequences for the initial old and middle aged generation evidently depend on these generations’ holdings of capital; we display numbers for holdings in the middle of the pre-reform ergodic distribution. Welfare consequences for other holdings (that have positive probability under the per-reform ergodic distribution) are of similar magnitude and display the same order with respect to the exogenous shock $z$.}

Table 2: Welfare Consequences in Small Open Economy in %

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>Initial Old</th>
<th>Initial Middle Aged</th>
<th>Newborns</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_1$</td>
<td>3.77</td>
<td>0.36</td>
<td>-0.43</td>
</tr>
<tr>
<td>$z_2$</td>
<td>8.30</td>
<td>0.45</td>
<td>-0.37</td>
</tr>
<tr>
<td>$z_3$</td>
<td>3.24</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>$z_4$</td>
<td>10.96</td>
<td>1.53</td>
<td></td>
</tr>
</tbody>
</table>

We find that, not surprisingly, the initial old and middle aged agents have large welfare gains from the introduction of social security because they receive benefits without having contributed to the system at all (the initial old) or having only contributed for one period. For the initial old the benefits are particularly high in states $z_2$ and $z_4$ in which returns to capital are low and wages are high, since their benefits depend on current wages in the economy.

The currently middle-aged agents also benefit from the reform, albeit not as strongly. For them it is states with currently low wages ($z_3$ and $z_4$) that make social security look really attractive since they contribute little in absolute terms (payroll taxes are proportional to wages) and can expect average pensions (since aggregate shocks are iid and the aggregate capital shock does not depend on the history of aggregate shocks).

In order for the introduction of social security to generate a Pareto improvement the welfare consequences for newborn generations are decisive. We see that for moderate risk aversion the reduction in mean returns for part of their retirement portfolio generates welfare losses from social security, whereas for high risk aversion, with $\sigma = 4$, enhanced risk sharing overcompensates the reduction in mean consumption.\footnote{The initial old agents have no risk remaining. The fact that their welfare gains a slightly higher for high risk aversion is entirely due to the fact that for $\sigma = 4$ they have saved less for retirement on average (where averages refer to the pre-reform ergodic distribution) and thus benefit more from the reform. Currently middle-aged agents still face substantial risk, which explains that their welfare gains increase with their degree of risk aversion, for the same intuitive reasons described for the newborns.} So if agents are
fairly risk averse and social security does not crowd out the aggregate capital stock (because of the small open country assumption), then even for a realistically calibrated economy the introduction of a (small) unfunded and redistributive social security system may constitute a Pareto improving policy innovation, since it enhances the sharing of risk across generations.\footnote{We observe that for $\sigma = 1$ the welfare losses are bigger in states $z_1$ and $z_2$ where current wages are high than in states $z_3$ and $z_4$ where current wages are low and social security has a better implicit return. It then may seem surprising that the ordering of states is reversed for $\sigma = 4$. But if agents are very risk averse, since capital is quite risky and therefore a bad asset to hold for agents, young agents short the asset (with $\sigma = 1$ they are savers even when young). The introduction of social security and associated forced saving for retirement makes them short the risky asset even more, in particular in low-wage states $z_3$ and $z_4$. This adverse effect of social security is insufficient to overcome the benefits from enhanced risk sharing, but explains why the welfare gains from the introduction of social security are larger in high-wage states $z_1, z_2$ than in low-wage states $z_3, z_4$.}

7 Conclusion

Can the introduction of an unfunded social security system provide a Pareto improvement by facilitating intergenerational risk sharing? In this paper we argue that, in the presence of incomplete markets, it potentially can do so in a quantitatively important way. However, in a realistically calibrated economy the intergenerational risk sharing role of unfunded social security is dominated in its importance by the adverse effect on capital accumulation arising from the introduction of such a system.

Three immediate avenues for future research remain. Quantitatively, it would be desirable to increase the number of generations populating the economy at a given point of time. This would enable us to interpret the stochastic shocks as real business cycle shocks and calibrate them correspondingly. As we argue in Krueger and Kubler (2001), our computational algorithm can handle up to nine generations, but an alternative numerical algorithm based on an approximation of the cross-sectional wealth distribution (instead of its exact representation, as in the current paper) is needed as the number of generations becomes larger. Such an algorithm has been provided by Krusell and Smith (1998) for infinite horizon models and been applied to overlapping generations economies of the form studied in this paper by Storesletten et al. (2000) and Gourinchas (2001). In our earlier paper we show, however, that the numerical accuracy of this algorithm may suffer substantially if one considers aggregate shocks of the order of magnitude similar to the great depression.
Secondly, the current paper abstracts from several beneficial roles of an unfunded, redistributive social security system. In the presence of incomplete financial markets social security provides a partial substitute for missing insurance markets against idiosyncratic labor income and lifetime uncertainty. On the other hand the distortive effects of payroll taxes on the labor supply decision remain unmodeled as well. The abstract from these features to more clearly isolate the potential magnitude of the beneficial intergenerational risk sharing role of social security. A complete assessment of its relative quantitative importance, compared to the intragenerational risk sharing and distortion effects would require incorporating these effects explicitly, however. Whereas elastic labor supply would add limited complexity to the numerical algorithm by adding simply a control variable, allowing for uninsurable idiosyncratic uncertainty would generate intragenerational heterogeneity, a nontrivial wealth distribution within generations and thus induce the same curse of dimensionality that occurs when expanding the number of generations in the model.

Finally, in this paper we assess whether social security should be introduced under certain aggregate economic conditions. Our normative analysis, however, is silent about the political conflict potentially surrounding the adoption, reform or termination of social security. An extension of the work of Cooley and Soares (1997) and Boldrin and Rustichini (2000) to our environment with aggregate uncertainty would be needed to address the question why, if not mutually beneficial, the US social security system was introduced when it was introduced and who one should expect the major supporters of this reform to be.

References


Figure 1: Consumption Equivalent Welfare Gains from the Introduction of Social Security
Figure 2: Consumption Equivalent Welfare Gains from Completing the Markets
Figure 3: Welfare Consequences of a Social Security Reform for Benchmark Calibration: High Initial Capital Stock
Figure 4: Welfare Consequences of a Social Security Reform for Benchmark Calibration: Small Initial Capital Stock