Journal of Public Economics 95 (2011) 16–27

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Welfare and generational equity in sustainable unfunded pension systems

Alan J. Auerbach [a,*], Ronald Lee [b]

[a] Department of Economics, University of California, 508-1 Evans Hall, Berkeley, CA 94720-3880, United States
[b] Department of Demography, University of California, 2232 Piedmont Ave, Berkeley, CA 94720-2120, United States

ARTICLE INFO

Article history:
Received 11 June 2009
Received in revised form 11 September 2010
Accepted 15 September 2010
Available online 25 September 2010

Keywords:
Pensions
PAYGO
Social security

ABSTRACT

Using stochastic simulations we analyze how public pension structures spread the risks arising from demographic and economic shocks across generations. We consider several actual and hypothetical sustainable PAYGO pension structures, including: (1) versions of the US Social Security system with annual adjustments of taxes or benefits to maintain fiscal balance; (2) Sweden's Notional Defined Contribution system and several variants developed to improve fiscal stability; and (3) the German system, which also includes annual adjustments to maintain fiscal balance. For each system, we present descriptive measures of uncertainty in representative outcomes for a typical generation and across generations. We then estimate expected utility for generations based on simplifying assumptions and incorporate these expected utility calculations in an overall social welfare measure. Using a horizontal equity index, we also compare the different systems' performance in terms of how neighboring generations are treated.

While the actual Swedish system smoothes stochastic fluctuations more than any other and produces the highest degree of horizontal equity, it does so by accumulating a buffer stock of assets that alleviates the need for frequent adjustments. In terms of social welfare, this accumulation of assets leads to a lower average rate of return that more than offsets the benefits of risk reduction, leaving systems with more frequent adjustments that spread risks broadly among generations as those most preferred.

1. Introduction

Population aging threatens the financial viability of Pay-As-You-Go (PAYGO) pension programs in many countries, and demographic fluctuations may lead to generational inequities. The old-age dependency ratio in the United States is projected to rise by 80% between 2010 and 2050, to more than double in Japan, and to rise by 55 and 70% in Sweden and Germany (United Nations, 2007).[1] Baby booms and busts have characterized the recent past of industrial nations, and fertility has recently risen in many countries with formerly hyper-low fertility (Goldstein et al., 2009). PAYGO programs are typically not structured to deal automatically with changing old-age dependency ratios, and as a consequence now promise a level of benefits that cannot be sustained at current tax rates. Thus, deep structural reforms are expected and in some countries have already occurred.

Beyond the problem of fiscal instability, most PAYGO programs are of a Defined Benefit (DB) structure and create incentives for early retirement (Gruber and Wise, 1999) and distort labor supply decisions over the whole life cycle. Furthermore, in creating unfunded pension claims, these programs may weaken incentives to save and thereby reduce national wealth. Reforms that adjust the general level of taxes or benefits in PAYGO programs can address the problem of fiscal sustainability, but reforms that maintain existing program structures might have little effect on incentives for work or saving. Moreover, such reforms provide no permanent solution to fiscal imbalances given subsequent demographic and economic shocks.

Any PAYGO system tends to reduce national saving by creating transfer wealth, but transition to a funded system involves potentially large burdens on transition generations, and therefore may not be a politically viable reform option. Even within the PAYGO framework, however, there are alternatives that might be more attractive than a simple realignment of taxes and benefits. A new kind of pension program, called Notional, or Non-financial, Defined Contribution (NDC), is intended to address both permanent fiscal stability and labor supply incentives. Sweden has developed and implemented an NDC system and some other countries have followed suit including Italy, Poland, Latvia, Mongolia and the Kyrgyz Republic. Germany has recently adopted pension reforms that reflect some of the NDC principles, and France is also considering doing so (Legros, 2003; Holtzmann and Palmer, 2005).

The basic approach of NDC plans is to mimic the structure and incentives of funded Defined Contribution (DC) plans, such as 401(k) plans in the United States. As in actual DC plans, individuals contribute

* Corresponding author. Tel.: +1 510 643 0711; fax: +1 510 643 0413.
E-mail addresses: auerbach@econ.berkeley.edu (A.J. Auerbach), rlee@demog.berkeley.edu (R. Lee).

[1] These figures are for increases. The actual levels also differ widely with a projected old-age dependency ratio for the US in 2050 at 33.3 persons aged 65 and over per 100 persons age 15 to 64. Comparable figures for Japan, Sweden and Germany are 70.8, 41.8 and 50.1.

0047-2727/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpubeco.2010.09.008

to their own, *notional* accounts which yield a specified rate of return and are converted into annuities yielding a specified rate of return at an age chosen by the individual but above some stated minimum age. However, in the NDC plans, the specified rate of return earned annually by the accounts and paid by the annuity is generally linked to the growth rate of wages, which given the PAYGO setting should help make NDC systems more fiscally sustainable without frequent active policy interventions. Furthermore, the individual accounts based on individual contributions and explicit rates of return can reduce the distortion of work incentives if workers view their benefits and taxes as more closely linked than under traditional DB systems, in which the linkage of taxes and benefits may be minimal or not transparent.

NDC or related plans are designed to provide fiscal stability by incorporating automatic adjustments of benefits and, in some cases, taxes, in response to economic and demographic shocks. In an earlier paper (Auerbach and Lee, 2009) we explored the effects of such adjustments on fiscal stability. Fiscal stability is, of course, a desirable feature, but how the stability is achieved will affect the manner in which the risks associated with shocks are spread among generations. This risk-spreading will influence individual and social welfare, and in this paper we consider the performance of different NDC plan variants in this regard.

There are many respects in which risk-spreading might differ among plans, including the degree to which shocks are spread among generations, and the relative treatment of generations of different sizes, of contiguous generations, of workers and retirees, and of current and future generations. It is difficult to summarize these effects using one simple measure of plan performance, and so we consider several measures in order to shed light on the welfare effects of different plans. To do so, we use a modified version of the stochastic forecasting model developed by Lee and Tuljapurkar (1998), and Lee et al. (2003) to generate a large set of sample paths under different social security systems, each sample path corresponding to a different realization of economic and demographic shocks.

We will consider several actual and hypothetical PAYGO pension structures, including: (1) the actual Swedish NDC system, together with several modifications of it developed in our earlier paper; (2) the actual reformed German system, which maintains annual fiscal balance using a combination of tax and benefit adjustments; and (3) hypothetical versions of the US Social Security system in which taxes, benefits, or both are adjusted annually to maintain fiscal balance with zero debt or assets.[2] Some of these structures are currently in use, and others are hypothetical extensions of existing plans, but all are modeled with a high degree of realism.[3]

The remainder of the paper is organized as follows. The next section describes the plans we analyze, focusing on the key characteristics that will affect the spreading of risks. Section 3 provides a brief description of our stochastic simulation model and how it has been adapted for the current project. Section 4 describes the measures we use to evaluate the different systems' welfare effects, and provides the results of our analysis, and Section 5 offers a summary of our findings and some concluding comments.

2. Varieties of PAYGO pension plans

Our simulation model is based on the average age profiles of tax payments and benefit receipts for surviving members of the population, which we might interpret as referring to each generation's representative individual. Since we do not consider within-generation heterogeneity, many pension plan details are irrelevant and therefore are not discussed below. Nor do we model behavioral responses to differences in pension plans, instead considering alternative assumptions regarding the valuation of resources that may be seen as corresponding to different assumptions regarding behavioral responses. Of course, these aspects of pension structure and behavioral response are important in their own right, but in this paper we simplify drastically in order to focus on macro uncertainty and intergenerational differences, and to include rich detail on pension systems and the possible array of economic and demographic shocks.

To facilitate comparisons of different pension systems, we hold certain characteristics constant across them. We start by scaling the contribution level of each pension system to equal 10.6% of taxable payroll when averaged over all trajectories, corresponding to the Old Age and Survivors (OASI) portion of the current US system. We assume that all individuals work until age 67, the long-run normal retirement age under current US law, and that all individuals are retired thereafter. Further details of the different systems, as we model them, now follow.

2.1. US social security

To the tax system just discussed, we add benefits based on current US profiles, estimated as described below in the section describing the simulation model. Because the current US Social Security system is not sustainable, we will consider three alternative versions that would maintain its fiscal balance. None is intended to characterize the actual process of adjustment that will occur, which is of course very difficult to predict, but by considering alternative adjustment mechanisms we hope to trace out the range of possible risk-sharing outcomes implicit in the current law US program.

In the “Tax Adjust” variant, the age schedule of taxes is adjusted by a multiplicative factor each year to produce the revenue needed to cover that year's benefits and thus keep the system in perfect balance each year. In the “Benefit Adjust” variant, the age schedule of benefits is similarly adjusted each year so that the benefits exactly equal that year's tax revenues. In the “50–50 Adjust” variant, balance is achieved half by adjusting taxes and half by adjusting benefits. These three mechanisms will differ in terms of how shocks are spread among cohorts, with the tax adjust variant, at one extreme, focused exclusively on younger, working age cohorts, and the benefit adjust variant focused only on older, retired cohorts.

2.2. Swedish notional defined contribution system

The actual Swedish NDC system, described more fully in our earlier paper and in Holtzmann and Palmer (2005),[4] specifies a rate of return earned on accounts in each year t equal to the contemporaneous growth rate of the wage rate, g_t. At the date of retirement, the individual's account is converted into an annuity based on the account's balance as of that date. The terms of the annuity reflect mortality conditions at the time of conversion and an assumed wage growth rate, but the annuity is adjusted *ex post* for deviations of wage growth from this assumed rate.[5] Because the sustainable steady state PAYGO rate of return is the growth rate of total wages, $g+n$ (where n is the labor force growth rate), this system may have stability

[2] Because it is convenient for technical reasons to carry out simulations of the Swedish systems beginning with a small initial asset balance, we start all systems with the same initial balance to keep them on an equal footing.

[3] Our model of the actual US system accurately matches projections of the Actuaries of the US Social Security system when similar deterministic demographic and economic trajectories are assumed, for example.

[4] Also see Swedish Social Insurance Agency (2008).

[5] The actual Swedish system uses a pre-specified expected wage growth rate of 0.016, but we use 0.011, the underlying average rate of labor productivity growth in our simulations. We also adjust the annuity level to reflect changes in mortality after retirement, which is not a feature of the actual Swedish system. We found in our earlier paper that this post-retirement updating had only a minor impact on system stability.

problems. Recognizing this, the system's designer also included a balancing mechanism that goes into effect when program assets fall below a certain level, reducing the rate of return earned on accounts and annuities until asset levels are restored. In this way, demographic shocks enter the system indirectly and will be spread in a complex manner across current and future cohorts. Productivity shocks, which affect wage-rate growth, are absorbed entirely by benefits, but not simply those of current beneficiaries, because the shocks will also influence workers' notional pension wealth accumulation.

In the Swedish system, the balancing mechanism is based on a "balance ratio" b:

$$b = \frac{F + C}{NPW + P} \tag{1}$$

In this expression, F is the level (possibly negative) of financial assets, similar to the US Social Security trust fund. C is a new measure called the "Contribution Asset" which is defined to approximate the present value of future tax payments by participants. It equals the three-year average value of tax payments multiplied by the expected length of time between tax payments and receipt of benefits which is roughly thirty years.[6] The denominator of expression (1) is the total pension liability of the system, equal to notional pension wealth of current workers (NPW) plus an approximation of the present value of benefits due to current retirees (P). This balance measure does not rely on explicit projections of demographic and economic variables and is therefore not based on all the available information. On the other hand, as it is based only on current cross sectional data, it is seen as less likely to be distorted by political pressures.

If the ratio b falls below 1.0 then the balancing mechanism becomes operative and a brake is activated, reducing the gross rate of return $1+g_t$ used in computing notional pension wealth accruals and the growth of annuity payments by the factor b_t, to $(1+g_t)b_t$. The mechanism remains active until the product of the balance levels b_t for all years since activation first reaches or exceeds 1.0, that is, it applies in year $s>t$ if $\prod_{v=t}^{s} b_v<1.0$. The purpose of this condition is to ensure that the balancing mechanism has no long term effect on the level of benefits; that is, there is a catch-up period with $b>1$ until the initial slowdown in the growth of notional pension wealth and annuity payments is reversed. The fiscal problem is thus resolved by a temporary drop in benefits below the unrestrained trajectory rather than a move to a lower trajectory. Note that the balancing mechanism is asymmetric, and does not prevent the unlimited growth of the trust fund F.[7]

We simulate the Swedish system just described. However, we also construct and examine alternative NDC systems that have modified, symmetric versions of the balancing mechanism, a rate of return of $n+g$, or both. When the rate of return is $n+g$ rather than g the system tracks the varying demographic context and therefore should have less need of the balancing mechanism.[8] Using a symmetric balancing mechanism avoids what one may view as an unrealistic degree of asset accumulation. Because we impose the mechanism symmetrically, it is not necessary to incorporate the catch-up phase described above; we simply adjust benefits downward when $b<1$ and upward when $b>1$. We also consider another potential variation in the balancing mechanism, relating to the speed of adjustment. Let r_t^a be the adjusted *net* rate of return earned by the notional accounts, that is, the adjusted gross rate of return minus 1. In the Swedish system this is $r_t^a=(1+r_t)b_t-1$, where $r=g$. To soften the impact of this adjustment we introduce a scaling factor A, between 0 and 1, as follows:

$$r_t^a = (1 + r_t)[1 + A(b_t-1)]-1 \tag{2}$$

with $A=1$ under the original Swedish system and $A<1$ associated with a smaller immediate adjustment. We consider variants of the symmetric mechanism in which $A=0.5$ and $A=1$.

We thus consider three modifications of the original Swedish system: accounting for labor force growth in computing the normal rate of return; making the balancing mechanism symmetric; and reducing the strength of the balancing mechanism. All three should influence the manner in which economic and demographic shocks are spread among cohorts, but in a complex manner that is difficult to predict in advance of considering the simulations. These effects will be explored using three alternative variants of the Swedish system, all with a symmetric balancing mechanism, with the variables (r, A) equal to (g,1), (g,0.5), and ($n+g$, 0.5).

2.3. German system

Each beneficiary i receives a payment in year t equal to:

$$B_{t.i} = PV_t{}^*EP_i{}^*AA_i \tag{3}$$

where PV_t is the current pension value in year t, AA_i is an actuarial adjustment based on when the pensioner retired, and EP_i is the individual's "earning points" collected until retirement, with earnings points an increasing function of an individual's earnings relative to that of the average-wage individual for each year the individual worked. (For more details, see Börsch-Supan et al. (2003); Börsch-Supan and Wilke (2004); Ludwig and Reiter (2010)). As we are ignoring intragenerational heterogeneity, we fix EP_i at the (constant) assumed number of years of labor force participation, Y, and set AA_i equal to 1. Thus, the benefits formula reduces to[9]:

$$B_{t.i} = B_t = Y^*PV_t \tag{3'}$$

Note that pensions are set up so that retirees of different ages get the same benefit. This is but one element of the German system that will influence how shocks are spread.

The pension in year t evolves according to:

$$PV_t = PV_{t-1}{}^*\frac{AGW_{t-1}(1-CR_{t-1})}{AGW_{t-2}(1-CR_{t-2})}{}^*\left[1-\alpha^*\left(\frac{OA_{t-1}-OA_{t-2}}{OA_{t-2}}\right)\right] \tag{4}$$

where CR is the pension "contribution rate" (i.e., the social security payroll tax), AGW is the average gross earnings of employees, and OA_t is the old-age dependency ratio (OADR), defined as the ratio of population over 65 to population aged 15–64 in year t. The parameter $\alpha=0.25$, which we also use in our analysis. In our terms, pension benefit growth is linked to g, but adjusted for fluctuations in CR and OA. Substituting Eq. (3′) into Eq. (4), we get:

$$B_t = B_{t-1}{}^*\frac{AGW_{t-1}(1-CR_{t-1})}{AGW_{t-2}(1-CR_{t-2})}{}^*\left[1-\alpha^*\left(\frac{OA_{t-1}-OA_{t-2}}{OA_{t-2}}\right)\right] \tag{5}$$

Expression (5) describes the evolution of benefits. Taxes are adjusted each year as a residual so that taxes and benefits are equal in

[6] This measure would exactly equal this present value in steady state, with discounting at the rate $n+g$, and therefore would equal the system's ability to meet future pension obligations through taxes (Settergren and Mikula, 2005; Lee, 2005).

[7] Note also that while it is mathematically possible for b to fall below zero, this could not meaningfully happen because it would entail more than complete confiscation of pension wealth and benefits.

[8] We in fact found this to be so in our earlier paper. While even a system based on $n+g$ would on certain stochastic trajectories require some further intervention to preserve stability, the strength and frequency of these interventions were reduced by incorporating labor force growth in the annual benefit adjustments.

[9] In the end, the value assumed for Y does not matter, because we scale the size of the system so that the average tax rate equals that of the US system.

the aggregate, and the level of benefits is determined by our requirement, mentioned above, that the system's average tax rate over time equals that for the United States.

The German system differs from the others in that the pension benefit of all generations rises after retirement in proportion to preceding wage growth, *g*. However, benefits for a generation also vary over time in inverse proportion to the tax rate ("contribution rate"), so that beneficiaries share the pain or the gain of fiscal adjustment with taxpayers. Demography also influences the benefit through the rate of change in the old-age dependency ratio. As in the US 50–50 Adjust system, some of the adjustment occurs through the tax system and some through the benefit system, although the relative proportions within the German system are not clear from its specification.

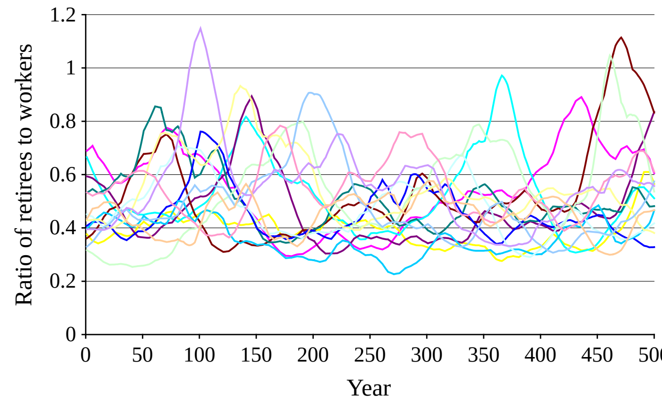

Fig. 1. Ratio of retirees to workers, 15 sample paths.

3. The stochastic simulations

To evaluate and compare the risk-sharing characteristics of these public pension systems, we build upon a stochastic simulation model developed in earlier work to generate stochastic long term forecasts of the US Social Security system (Lee and Tuljapurkar, 1998; Lee et al., 2003; see also the approaches in Alho et al., 2005, and Alho et al., 2008).

In the population model, the log of each age-specific mortality rate is taken to be a linear function of a single mortality index that is in turn modeled as a random walk with drift, based on Lee and Carter (1992). Fertility is modeled as an ARIMA process with a pre-assigned long term mean of 1.95 births per woman (the period Total Fertility Rate) as assumed in the 2004 Social Security Trustees Report (Board of Trustees, 2004). Immigration is taken as deterministically given by the intermediate assumption in that report.[10] The fertility and mortality processes are fit to historical US data. To give a sense of the demographic variations that result from this approach, Fig. 1 plots the ratio of retirees to workers along 15 randomly chosen sample paths. Because the time series models for fertility and mortality were fit to US data, these simulated sample paths reflect randomly occurring low-frequency fluctuations something like the US baby boom and baby bust, with consequent effects on the OADR. Mortality variations have a less profound effect on the age distributions because the variance in mortality is far smaller than in fertility, and because a mortality variation affects every age, while fertility affects only births.

The modeling of the US Social Security system builds on the stochastic population model using cross-section estimates of age profiles of labor earnings, tax payments and benefit receipts. The age profiles of labor earnings and tax payments are multiplicatively shifted from period to period by a time series of labor productivity growth. A time series model is fitted to the historical time series of productivity since 1950, purged of the influence of changes in the age composition of the labor force. The same simulation also can generate paths of age-specific benefits levels, since these depend on average wages at age 60 for each generation, which are in turn determined by stochastic productivity. The modified version of the US system we consider here has payroll taxes set equal to zero above age 67 and benefits set equal to zero below age 67, since we are assuming that retirement under each system occurs at age 67. Benefits after age 67 are based on a simplified version of the actual US formula of providing a replacement rate of average indexed monthly earnings that is then indexed to the price level after retirement.

The trust fund for the US system is set at a small initial balance that is maintained constant as a share of payroll. For all systems, we accumulate balances using a time series model of the real interest rate earned on Social Security special issue bond holdings. The interest rate is modeled as a stationary series, as is the productivity growth rate, with their evolution modeled using a two-variable VAR. The productivity growth rate has a positive mean of 1.1% (based on the Actuary's assumption in Board of Trustees, 2004, for the growth rate of covered real wage), so productivity, earnings, taxes and benefits all trend upwards.[11]

We abstract from the particular demographic situation in the United States today so that we can derive results that are of more general applicability. We modify the model described above so it converges to an approximately stationary stochastic population distribution, accomplishing this by setting the mortality drift term to zero.[12] Because the level of mortality follows a random walk the population variance increases very slowly over time, and therefore this stochastic equilibrium is only approximately stationary, but in practice we have found that the nonstationarity of mortality is negligible.[13]

While this description of the model has glossed over many details, it should convey a general idea of how we simulate the pension systems.[14] For the actual simulations, we consider a common sample of 1000 randomly drawn trajectories, which in practice appears to be a large enough number to calculate stochastic distributions reasonably accurately. For each trajectory, we start with initial conditions based on long-run average values of different state variables and run the model for a "pre-sample" period of 100 years to generate histories needed for certain pension calculations. We then follow the paths for an additional 500 years, allowing us to examine the welfare of nearly 400 cohorts over their entire lives, which are assumed to extend for a maximum of 106 years. We consider large numbers of cohorts for each trajectory so that we can take account of transition effects of particular systems (notably the actual Swedish system, with its tendency toward asset accumulation) and allow for the full effects of

[10] See Lee et al., 2004 for treatment of immigration as stochastic in this framework.

[11] The Trustees Reports distinguish between the growth rate of productivity and of the covered real wage, with the gap reflecting changes in hours worked and the proportion of compensation that is fringe benefits. We abstract from these issues and simply use the assumed growth rate of the covered real wage, 0.011, which we will henceforth refer interchangeably to as the productivity growth rate or growth rate of the wage.

[12] The expected value of fertility is below replacement level, so in the long-run the population converges to a level at which the deficit of births is just balanced by the deterministic inflow of migration.

[13] Note that while productivity growth is stationary, productivity levels are not. We take this nonstationarity into account below when constructing summary measures that aggregate across generations.

[14] Further details of this simulation methodology are provided in our earlier paper.

Table 1
Summary statistics for eight pension plans based on 1000 stochastic paths of 500 years.

	US tax adjust	US benefit adjust	US 50–50 adjust	NDC Sweden	NDC (g, $A=1$) symmetric	NDC (g, $A=0.5$) symmetric	NDC ($n+g$, $A=0.5$) symmetric	German
NPV ($\gamma=0$)								
Mean	−0.0593	−0.0605	−0.0599	−0.0627	−0.0576	−0.0572	−0.0577	−0.0602
Variance across								
Trajectories	0.00165	0.00106	0.00122	0.00098	0.00100	0.00103	0.00103	0.00143
Generations	0.00101	0.00046	0.00061	0.00034	0.00046	0.00049	0.00049	0.00074
EU								
$\gamma=3$	−0.0809	−0.0782	−0.0792	−0.0801	−0.0753	−0.0751	−0.0758	−0.0792
$\gamma=5$	−0.1320	−0.1270	−0.1291	−0.1300	−0.1233	−0.1232	−0.1244	−0.1295
$\gamma=3$, ret. w	−0.0875	−0.0856	−0.0862	−0.0866	−0.0825	−0.0824	−0.0831	−0.0857
$\gamma=3, 5$	−0.1039	−0.0993	−0.1010	−0.1057	−0.0913	−0.0908	−0.0935	−0.1045
HE								
NPV	0.999621	0.999893	0.999848	0.999962	0.999895	0.999885	0.999882	0.999819

persistent shocks to be considered. Long trajectories also improve the precision of our estimates of distributions, conditional on the number of trajectories studied.

4. Results

In presenting results based on the simulations described above, we utilize several complementary and related measures to characterize the performance of the different systems. The first, the Net Present Value of expected lifetime taxes minus benefits relative to the expected present value of a cohort's lifetime labor earnings, is a common measure of pension program outcomes and therefore a useful place to start. However, to consider adequately the impact of risk, we need a measure of expected utility, which we consider next. Expected utility provides a good measure of lifetime welfare, but to take into account the effects on all generations, including those alive during transition periods, we need a social welfare function, which we also consider. Finally, to reflect the possibility that simple aggregation across cohorts may provide an inadequate picture of welfare effects when disparities among cohorts of approximately the same year of birth are seen as particularly undesirable, we also present a measure of horizontal equity constructed to provide information on the extent of such disparities. Tables 1 and 2 present these measures for the different public pension systems, which include three sustainable variants of the US system, the actual Swedish system along with three variants of it with different specifications of the rate of return and the balancing mechanism, and the actual German system.

Table 1 aggregates measures across generations weighting by generation size. We believe that population-weighted measures are generally more informative about welfare effects. If we did not weight by generation size when aggregating, for example, then a program that paid lower benefits to larger generations would appear to perform better than one that did not, but the welfare implications would be unclear because we would be ignoring the larger number of individuals facing the adverse adjustment. However, because differences between weighted and unweighted results may be informative about the relative treatment of large and small cohorts under the different systems, Table A1 in the Appendix presents the comparable calculations unweighted by cohort size.

4.1. Net present values

We begin by considering the Net Present Value (*NPV*) of expected lifetime taxes minus benefits, expressed as a share of the expected present value of a cohort's lifetime labor earnings, both discounted to the cohort's year of birth,

$$NPV_t = \frac{\sum_i \pi_{t,i} \sum_{s=t}^{t+T} \frac{1}{E_t\left[\prod_{u=t}^{s}(1+r_u)\right]} F_{s,i}}{\sum_i \pi_{t,i} \sum_{s=t}^{t+R} \frac{1}{E_t\left[\prod_{u=t}^{s}(1+r_u)\right]} Y_{s,i}}, \tag{6}$$

where $F_{s,i}$ is the survival-weighted per capita benefit minus tax payment for this generation in year s and along trajectory i, $Y_{s,i}$ is survival-weighted income per capita for the generation in year s and along trajectory i, r_u is the interest rate in year u, T is the maximum lifespan (106 in our simulations) and R is the retirement age (67 in our simulations). When trajectories are weighted by population, the weights $\pi_{t,i}$ equal the cohort's initial population along trajectory i as a fraction of the sum of the cohort's initial populations along all trajectories. Otherwise, we set $\pi_{t,i} \equiv 1/N$, where N is the number of trajectories. As discussed in the Appendix below, the *NPV* as described in Eq. (6) measures the impact of the social security system on the cohort's expected utility (expressed as a share of expected lifetime earnings) for the case of risk-neutral preferences.[15]

The first row of Table 1 presents the mean *NPV*s for each system. These mean values are in the range −0.057 to −0.063 as a fraction of lifetime earnings. These negative values are what one would expect for a social security system in a dynamically efficient economy.[16] The mean value for the actual Swedish system is notably lower than that of any other system, as it is penalized by the asymmetric balancing mechanism. The remaining systems fall roughly into two groups, with the three variants of the Swedish system having means between −0.0572 and −0.0577, and the three US variants and the German system having means between −0.0593 and −0.0605. As will be discussed below, these differences in means among plans that do not accumulate assets relate primarily to the treatment of transition

[15] In earlier versions of this paper we also presented results for another common measure of program performance, the system's internal rate of return. The results were generally consistent with those presented here for the *NPV* measure. We omit them here in the interest of space and because they relate less directly to measures of individual and social welfare.

[16] A mature, stable system of intergenerational transfers, such as a PAYGO public pension system, yields a rate of return equal to the growth rate of the labor force plus the growth rate of productivity. In general, the rate of return to capital and the market interest rate will be greater than the rate of return to a PAYGO pension system, and therefore the *NPV* of the benefit-tax stream will be negative.

Table 2
Social welfare calculations for eight pension plans.

	US tax adjust	US benefit adjust	US 50–50 adjust	NDC Sweden	NDC (g, $A=1$) symmetric	NDC (g, $A=0.5$) symmetric	NDC ($n+g$, $A=0.5$) symmetric	German
Unadjusted								
$\gamma=0$	0.00140	0.00140	0.00140	−0.00878	0.00186	0.00189	0.00181	0.00140
$\gamma=3$	−0.00406	−0.00308	−0.00273	−0.01175	−0.00226	−0.00226	−0.00252	−0.00287
$\gamma=5$	−0.02109	−0.01923	−0.01890	−0.02645	−0.01788	−0.01792	−0.01840	−0.01921
$\gamma=3$, ret. w	−0.01014	−0.00963	−0.00911	−0.01787	−0.00844	−0.00841	−0.00878	−0.00882
$\gamma=3, 5$	−0.00120	−0.00156	−0.00007	−0.00990	−0.00022	−0.00026	−0.00074	−0.00110
Adjusted for initial differences under risk-neutrality								
$\gamma=0$	0.00186	0.00186	0.00186	−0.00878	0.00186	0.00186	0.00186	0.00186
$\gamma=3$	−0.00360	−0.00262	−0.00227	−0.01175	−0.00226	−0.00229	−0.00248	−0.00241
$\gamma=5$	−0.02063	−0.01877	−0.01844	−0.02645	−0.01788	−0.01796	−0.01835	−0.01875
$\gamma=3$, ret. w	−0.00968	−0.00917	−0.00865	−0.01787	−0.00844	−0.00844	−0.00873	−0.00836
$\gamma=3, 5$	−0.00074	−0.00110	0.00039	−0.00990	−0.00022	−0.00029	−0.00070	−0.00064

generations, who are not yet included as we focus on cohorts with complete lifetimes.

The next two lines of Table 1 present two variance measures for the *NPV*, first across trajectories (averaged across cohorts), and then across cohorts (averaged across trajectories). The first variance measure is more relevant when considering the impact of social security on risk-averse households, while the second measure conveys some information about how different generations fare along any particular trajectory, which might also be useful for social welfare evaluations where horizontal equity is a concern. The two measures of variance provide similar rankings of the different systems. The actual Swedish system has the lowest variance across trajectories and across generations, the second attribute being particularly evident. Intuitively, the Swedish system achieves this stability at the cost of having the lowest mean, because its asymmetric balancing mechanism lets assets accumulate on some sample paths and thereby reduces the need for taxes or benefits to respond to shocks.

Of the remaining systems, the ranking in terms of variance shows four plans grouped next after the Swedish system: the US Benefit Adjust and the three variants of the Swedish system. These are followed by the US 50–50 system and the German system, with the US Tax Adjust system showing the highest variance across trajectories and generations. Note that the systems with higher variances are those that rely more on tax adjustments, with the system relying solely on tax adjustment being the highest. Presumably, this reflects the fact that adjustments occurring earlier in life are less heavily discounted than those occurring later, an issue to which we will return below when considering overall social welfare measures. In this case also, the consideration of transition generations makes a difference.

We will also consider how the plans vary in the relative burdens they impose on different-size cohorts, by comparing the results in Table 1, which are weighted by generation size, to those in Table A1 in the Appendix, which are not. We find that population weighting raises average values of the *NPV* for plans that depend wholly or in part on tax adjustments. This makes sense: it is advantageous to be in a large cohort because the tax burden is shared among more taxpayers and the tax rate will be lower, other things equal. With the US Benefit Adjust system, population weighting actually lowers the *NPV*, because adjustments must be larger when the old-age dependency ratio is higher. An interesting contrast, however, is provided by the *NDC* plans, particularly the three variants with a symmetric balancing mechanism. For two of these plans (those using only the growth of the wage rate in computing the rate of return), weighting actually raises the mean *NPV*. This suggests that large cohorts fare somewhat better, even though the adjustments occur only on the benefit side. This result reminds us that there are two issues associated with fiscal adjustment, not just whether benefits or taxes are adjusted, but also *whose* benefits or taxes. Under the *NDC* plans, benefit adjustments effected at a particular date – through changes in the rate of notional account accumulations – may fall on cohorts quite far from retirement, i.e., the same cohorts that would be hit by tax adjustments.

4.2. Expected utility

We have considered both mean values and variances of the *NPV* measures as a way of describing the trade-offs of the different systems. However, these summary measures cannot adequately characterize impacts on individual welfare, because one cannot weigh the trade-off between mean and variance without an explicit utility function, and even then one needs more than these two moments to assess the impact on utility. For example, a high variance in the *NPV* could be helpful if the upper tail of its distribution coincides with states of nature in which wage growth is below its mean and thereby insures, rather than exacerbates, lifetime income risk. We have, therefore, developed a methodology for approximating the incremental impact on expected utility of any particular pension system.

Our simulations cover only the pension system and do not include saving, asset income, or non-pension taxes. A full expected utility calculation would need to take these other elements into account in a very complex dynamic programming problem. As an alternative, we develop a methodology that, although being partial-equilibrium in nature, is designed to capture the potential effects of individual responses. In particular, we derive a local approximation of the impact of different systems on expected utility using simplifying assumptions that relate the marginal utility of consumption along a trajectory to the level of risk aversion and the level of wages along that trajectory. The Appendix describes this methodology in detail. As already mentioned, this method yields the *NPV* measure under risk-neutrality.[17] Also, although we do not take into account the general equilibrium effects of pension-system differences on factor prices, the stochastic processes for wages and rates of return could, in principle, be linked to demographic variables and thereby incorporate the impact of demographic changes on factor prices.[18]

[17] Because we do not consider intragenerational heterogeneity, we cannot analyze the extent to which different systems spread risks within cohorts, for example by redistributing from high-income to low-income individuals. While this aspect of pension systems is also interesting, it is one that has received more attention in the literature.

[18] We say “in principle” because our estimated stochastic processes for wages and interest rates failed to show any linkages to fertility and mortality.

The second panel of Table 1 provides our estimates of Expected Utility (*EU*) for the various plans, under different assumptions about the coefficient of relative risk aversion and the determination of marginal utility in retirement. As with the *NPV* estimates just considered, these are computed by averaging over all complete generations, weighted by population size.

The first row of this panel presents results for the case in which the coefficient of risk aversion, γ, equals 3 throughout life. Looking at this row, we can see a number of significant changes in the relative standing of the different plans, relative to the mean *NPV*s. First, the actual Swedish system gains relative to the other plans, reflecting the smoothing of outcomes that its asset accumulation permits. Second, performance more generally reflects the *NPV* variances already discussed, with the US Tax Adjust plan showing the biggest drop relative to its *NPV* and the other plans that include tax adjustments (US 50–50 Adjust and German) doing only somewhat better. Indeed, the US Tax Adjust plan has a lower value of *EU* than the Swedish plan, suggesting that, at least for the generations taken into account, the smoothing under the Swedish system outweighs the penalty imposed by asset accumulation. Third, the differences among the three NDC plan variants are little changed, reflecting the fact that their *NPV* variances were similar.

The next three rows of the table present some sensitivity analysis. Setting the coefficient of risk aversion, γ, at the higher value of 5 simply enhances the effects just considered. In the next row, we assume that the marginal utility of consumption for retirees is based on the wage rate at the cohort's age of retirement rather than at its entry into the labor force.[19] As discussed in the Appendix, this alternative assumption about marginal utility is more consistent with the case in which retirees rely primarily on transfers received from working generations, rather than their own resources, during retirement. While this variation in assumptions lowers expected utility, it has little impact on the relative attractiveness of the different systems. The final row of the table illustrates what happens when we assume that risk aversion changes over the course of life, in particular assuming that the coefficient of risk aversion is higher ($\gamma = 5$) in retirement than during working life ($\gamma = 3$).[20] Here, there are interesting patterns in the relative performance of the different systems. The smallest reductions in expected utility relative to the case for which $\gamma = 3$ throughout life are for the three NDC systems, while the biggest impact is for the German system. Throughout the sensitivity analysis, though, the NDC systems provide the highest level of expected utility, and the US Tax Adjust system or the German system is the second-lowest.

4.3. Social welfare calculations

As discussed earlier, we scale the different public pension systems to make it easier to compare their long-run impacts. In particular, we keep the average size (in terms of tax payments) the same, to avoid confusing the systems' relative performance with the fact that PAYGO pensions yield below-market rates of return. However, as came up when contrasting the performance of the plans that adjust benefits and those that adjust taxes, systems that impose more of the adjustment risk on benefits also end up imposing more risk on initial transition generations. Therefore, leaving such initial generations out of the calculations may bias our conclusions in favor of systems that adjust benefits, even if all systems are scaled to have the same average size.

In order to deal with this potential bias, we expand our *NPV* and *EU* calculations to include all transition generations alive during our simulation period, including those already alive at the beginning of the transition and those still alive at the end. Following the methodology laid out in the Appendix, we calculate partial values of *NPV* and *EU* for these cohorts and then aggregate these with the corresponding values for complete generations, discounting everything back to the initial year of our simulations to get a single social welfare measure for each system. By construction, our social welfare measure will equal zero under zero risk aversion for any system that maintains a balance of zero assets throughout the simulation, because it then reduces to the discounted sum of benefits less taxes in all years.

The results of these calculations are provided in Table 2. The first row of the table gives the results for the case of risk-neutrality. Except for the actual Swedish system, the results are all slightly positive, because the systems all start with the same small positive asset balance. For the US and German plans, the values are identical, because these plans maintain a constant asset-payroll ratio throughout each simulation trajectory. For the NDC variants of the Swedish system, the values are positive but slightly higher, suggesting that assets fall slightly from their initial value, on average. For the actual Swedish system, the value is negative because this system accumulates assets on average. To neutralize these small differences among the systems related to asset drift, we adjust the numbers in the table by these differences using the NDC (g, $A = 1$) plan as the benchmark for adjustment. The adjusted version of the table is presented in the lower panel of the table. Note that we do not adjust the actual Swedish system, because this system accumulates assets by design.

Focusing on the adjusted values in Table 2, we see once again that the actual Swedish system fares relatively better as risk aversion is taken into account. Unlike in Table 1, however, this system continues to perform much worse even when risk aversion is high. This relatively poor performance when transition generations are included makes sense, because as the Swedish system accumulates assets as a buffer during the initial phase, the balancing mechanism is frequently in place. The initial cohorts therefore are more likely to suffer under this initial adjustment process and to get little benefit from the subsequent reduction in volatility that this buffer provides.[21]

Looking now at the US systems, we see that the apparent advantage of the Benefit Adjust plan has disappeared. This plan now always performs worse than the US 50–50 Adjust plan, which now is the preferred US plan. Taking into account the impact on initial generations causes the Benefit Adjust plan to lose its apparent advantage. Indeed, when risk aversion is higher in old-age ($\gamma = 5$ vs. $\gamma = 3$), the Benefit Adjust plan fares worse than the Tax Adjust plan, because beneficiaries are more risk averse than workers. That the 50–50 Adjust plan performs better than either of the other two also makes sense, because it spreads the impact of each adjustment over more generations than either of the other plans. The German plan, which also distributes its annual adjustment over both beneficiaries and workers, performs only slightly worse than the US 50–50 Adjust plan and always better than the other two US plans.

Indeed, the US 50–50 Adjust plan now performs similarly to the Swedish NDC plans. Although the NDC plans adjust only benefits, they spread each year's adjustment over a larger number of cohorts, including not just current beneficiaries but also future ones, i.e., current workers. Among the Swedish plans, there is a distinct partial ranking. While the strength of the balancing mechanism ($A = 0.5$ vs. $A = 1$) still has virtually no impact, the inclusion of population fluctuations in the rate-of-return calculation ($n + g$ vs. g) has a distinctly *negative* impact on social welfare. Recall our previous finding that an NDC system based on the growth of wages, rather than

[19] In the notation used in expression (A10′), the wage w_{t+R} rather than w_t is used.

[20] There is some evidence for such a pattern of increasing risk aversion, although the evidence is not overwhelming. See Poterba (2001).

[21] At the time of its actual pension reform, Sweden had accumulated large buffer funds that reduced the likelihood of such adverse effects on transition generations by the pension system. However, such accumulations presumably imposed the same type of burden on roughly the same generations outside the pension system.

of the wage rate, is inherently more stable and relies less on the balancing mechanism for stability. Our findings here suggest that the added volatility in benefit accruals under the NDC ($n+g$) system when the balancing mechanism is not in place outweighs the volatility imposed by more frequent application of the balancing mechanism under the NDC (g) system.

In summary, the various plans differ less once we take transition generations into account. But, even leaving aside the actual Swedish system, the differences among plans are not insignificant. For example, the best-performing plan according to the lower panel of Table 2, the NDC (g, $A=1$) plan, provides a level of social welfare that is higher by 0.003–0.3% of lifetime earnings than the US Tax Adjust plan, when risk aversion is high ($\gamma=5$). This is, admittedly, a stark comparison, as the US Tax Adjust plan involves no smoothing of annual shocks across time. But it does suggest that, even among fiscally stable PAYGO plans that spread risk among large numbers of generations, the precise pattern of risk-sharing matters.

4.4. Horizontal equity

To this point we have considered measures of the uncertainty of pension program outcomes and the trade-off of this uncertainty against the mean return. The expected utility measure reflects this trade-off, as does our social welfare measure that also takes into account the expected utility of transition generations. Many would argue that the vector of expected utilities for different cohorts provides all the information needed to evaluate social welfare. But we may care in addition about how generations fare relative to other generations along a particular trajectory.[22] One simple measure of performance in this regard is the variance of the *NPV* among generations along a particular trajectory, already considered. But we sense that such concerns about the relative well-being of different generations relate primarily to the treatment of individuals of similar ages, since these other generations form a likely reference group. For example, the “notch” generations in the United States experienced particularly sudden and large variations in their pension benefits in a way that struck many as unfair.

To reflect these concerns, we provide an additional set of performance measures for the various public pension schemes, based on the horizontal equity measure developed by Auerbach and Hassett (2002). This measure is derived from a social welfare function for which the degree of inequality aversion may differ according to whether individuals are “near” each other, by some measure, or not. One may decompose this social welfare function into different components, one of which reflects the social welfare cost of local tax burden disparities at each income level. We derive the scalar index of horizontal equity by asking what uniform fraction of existing income would deliver the same level of social welfare if all such local disparities were eliminated. This index has a maximum possible value of 1.0, and values closer to 1.0 indicate greater horizontal equity. For example, a value of 0.999 indicates that we would be willing to give up 0.1% of total income to eliminate horizontal inequality.

Measuring horizontal equity requires the specification of the degree of inequality aversion for comparisons among members of a particular reference group and a definition of the reference group itself, in this case with respect to generational proximity. As we claim no particular insight as to which parameters are best here, we simply adapt as closely as possible those used for the base case in Auerbach and Hassett (2002), a CES degree of inequality aversion equal to 2 and a neighborhood based on a normal distribution with standard deviation equal to 0.1 times the total number of generations (i.e., $0.1*396=39.6$). We estimate horizontal equity in relation to our *NPV* calculation,[23] measuring income as the present value of earnings along the trajectory and taxes as the net present value of taxes minus benefits along the trajectory, and then averaging the resulting measures of horizontal equity across trajectories weighted by the trajectory's average cohort size.

The horizontal equity (*HE*) results are displayed in the last line of Table 1; the gap between the reported measure and 1.0 is the fraction of lifetime income that society would be willing to pay in order to remove horizontal inequity. The higher the value in the table, the greater is Horizontal Equity. The results are clear: the actual Swedish system dominates the other systems, its treatment of contemporaneous generations smoother than any other. The NDC systems and US Benefit Adjust systems do the next best, and the US Tax Adjust fares worst.

How important is it to take horizontal equity into account? According to these calculations, the biggest difference among systems is a fraction 0.00034–0.034% of lifetime income, between the actual Swedish system and the US Tax Adjust system. By comparison, in our adjusted social welfare measures in Table 3, the Swedish system is always at least a fraction 0.006–0.6% of lifetime income worse than the US Tax Adjust system, and this is for the case of very high individual risk aversion ($\gamma=5$). This comparison suggests, at least for the chosen specification of horizontal equity, that the very similar treatment the Swedish system provides to generations with nearby birth years should not be a major factor in evaluating the plan's overall performance.[24]

5. Discussion and conclusions

The NDC systems aim to pay a rate of return to contributors that is warranted by the macroeconomic/demographic environment. However, Sweden, in setting up its system, chose to make that rate of return equal the rate of wage growth, g, rather than $n+g$ which is the rate payable in steady state. Because they also included a balancing mechanism in their system design, if labor force growth should drop below 0 then the balancing mechanism would eventually automatically reduce the rate of return below g. Our analysis shows that this program design insulates participating generations from variations in the economic/demographic environment. The asymmetric balancing mechanism, which reduces the rate of return in some circumstances but never raises it, apparently plays a key role. This arrangement permits the system to accumulate undistributed assets and therefore makes it yield a lower mean *NPV* compared to NDC systems with a symmetric balancing mechanism. But, by accumulating more assets, it avoids having to activate the balancing mechanism and thereby leaves the rate of return more stable along a trajectory. This makes the Swedish system look relatively better when risk aversion is explicitly included in the calculation of expected utility, but the net benefit appears smaller once the welfare of initial transition generations is taken into account, for these are the generations that bear the brunt of the Swedish system's buffer stock accumulation.

Our results suggest that assessment of the effects of programs on different generations needs to take account not only of risk aversion, but also of the treatment of transition generations. This treatment of transition generations is well known to matter for PAYGO systems in

[22] Consider, for example, two generations and two outcomes, “good” and “bad,” which occur with equal probability. In case A, the two generations experience the same outcome (i.e., the combined outcomes for the two generations are good–good and bad–bad, each half the time); in case B, the generations experience opposite outcomes (good–bad and bad–good, each half the time). The expected utility for each generation is the same in case A and case B, but we might conceivably have a social preference for case A over case B.

[23] The measure as originally developed is not easily applied for our expected utility measure with risk aversion.

[24] Reasonable variations in the parameters of the horizontal equity calculation do not change this conclusion. For example, raising the inequality aversion parameter from 2 to 5 lowers all measures of horizontal equity, as one would expect, but increases the difference between the Swedish system and the US Tax Adjust system only to 0.00043 of lifetime income. Reducing the relevant neighborhood using a standard deviation of 15 rather than 39.6 raises all measures of horizontal equity, as cohorts that are closer together are treated more equally by all the systems, and reduces the gap between the same two systems to 0.00013.

terms of the net transfers to such generations, but our methodology neutralizes differences across systems in such transfers. Our findings here relate to *second* moments — to the distribution of risk, not to average resources. Thus, the apparent disadvantages of plans that impose risk on younger generations – those relying more on tax adjustments – tend to vanish once the risk-bearing of transition generations is taken into account. For example, the apparent advantage of the US Benefit Adjust plan over other stable variants of the US system disappears when transition generations are taken into account, and one finds that the US 50–50 Adjust system performs up to the standards of the NDC systems, all systems that distribute annual shocks among both workers and retirees. The German system resembles the US 50–50 Adjust system in many respects but its performance suggests that it places a higher relative burden of risk bearing on workers and spreads risk somewhat less efficiently.

Among the NDC plans, there is relatively little difference apparent until transition generations are taken into account, at which point the NDC(g) systems look somewhat better than the inherently more stable NDC($n+g$) systems. This suggests that shifting more of fiscal adjustment to the balancing mechanisms may also improve risk-spreading.

Our results suggest, then, that spreading risk widely among generations improves welfare, and that the policy of reducing risk through asset accumulation, as the Swedish system does, offers a less attractive approach unless one places extremely high weight on horizontal equity, i.e., on maintaining a very smooth pattern of net benefits from one cohort to the next.

In future work, we hope to look more closely at the differences in performance we have uncovered here by looking at how the different systems distribute different types of shocks among cohorts. This will serve not only to confirm (or correct) some of the intuition provided here, but also to help us understand the extent to which different approaches might be used according to the source, strength and stochastic properties of shocks.

Acknowledgments

This research was supported by the U.S. Social Security Administration through grant #10-P-98363-1-02 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium and by NIA through grant R37AG25247. The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the Federal Government, or the NBER. Ed Palmer and Ole Settergren gave us valuable comments and advice that helped us to model the actual Swedish system. Alex Ludwig provided comments and very helpful information for modeling the German system. We also thank participants in the NBER Summer Institute, the NBER Public Economics Program meeting, including our discussant Peter Diamond, a seminar at Stockholm University, and two anonymous referees for comments on earlier drafts, and gratefully acknowledge the assistance of Anne Moore, Erin Metcalf Johnson, and Carl Boe.

Appendix A. Valuing flows from social security

In theory, the preferred way to evaluate social security is to specify the household's other sources of income and then solve for its optimal consumption and portfolio choice behavior as a function of state variables at each date. As part of this dynamic programming solution, we would also obtain the household's value function as of date t. Solving for the value function in the presence and absence of social security would give us the value the household would place on the social security system. This approach is not feasible, however, because of the very large number of state variables involved and the complicated methods of calculating social security benefits, particularly under the Swedish system and its balancing mechanism. Thus some simplification is necessary. The approach we take here is to treat differences among systems as small deviations and use second-order Taylor approximations in combination with assumptions about the marginal utility of consumption at the point of approximation.[25]

Before specifying our proposed methodology, it is useful to distinguish two ways in which risk aversion will affect the valuation of taxes and benefits:

1. The household will value flows differently in different states of nature (assuming that there is no perfect insurance against variations in productivity, etc.); and
2. The household will be averse to fluctuations in benefits and/or taxes, even absent other sources of income fluctuations.

Proposed methodology

In the presence of uncertainty, the Euler equation for household optimization would imply that the expected marginal utility for $s>t$ would relate to marginal utility in the year of birth, t, by:

$$U'_t = E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]U'_s\right\}. \tag{A1}$$

We don't know these state-contingent marginal utilities without solving the full optimization problem, so we make a simplifying assumption, that marginal utility at date s in state i is proportional to some function of the vector of state variables at date s, (including the level of productivity, interest rates, population, etc.) $x_{s,i}$:

$$U'_{s,i} \sim h_s\left(x_{s,i}\right) = a_s h_s\left(x_{s,i}\right), \tag{A2}$$

where a_s is a constant at date s. The idea here is that incremental additions to or subtractions from resources will vary across states at a given date, these marginal valuations being higher in bad states (e.g., states with low levels of productivity) than in good ones. Note that the function is subscripted by date, indicating that it may differ over age. For example, an individual's marginal utility may be more sensitive to the economy's level of productivity during working years, when most resources come from wages, rather than during retirement years. We will discuss the specification of $h_s(\cdot)$ further below.

Substituting Eq. (A2) into Eq. (A1), we get:

$$U'_t = E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]a_s h_s(x_s)\right\}. \tag{A3}$$

Without loss of generality, we can normalize the utility function so that $a_t=1$ and hence $U'_t=h_t(x_t)$. This normalization then gives us the solutions for a_s at each date s:

$$a_s = \frac{1}{E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]\frac{h_s(x_s)}{h_t(x_t)}\right\}}. \tag{A4}$$

[25] An alternative approach, taken by Ludwig and Reiter (2010) to analyze the German system, is to use a linear-quadratic approximation of the government's problem to derive linear decision rules in response to shocks around a deterministic steady state, and to evaluate the second-order welfare effects of the resulting fluctuations around the deterministic steady state. Our approach does not assume that stochastic fluctuations are small (only that the differences among the systems are small), but instead requires alternative simplifying assumptions regarding marginal valuations along different stochastic trajectories.

From Eqs. (A2) and (A4), we have the solutions for marginal utility,

$$U'_{s,i} = \frac{h_{s,i}\left(x_{s,i}\right)}{E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]\frac{h_s(x_s)}{h_t(x_t)}\right\}}. \tag{A5}$$

Calculating the value of social security taxes and benefits

The value of taxes and benefits for the generation born in year t equals the change in utility associated with the flows of taxes and benefits, or

$$V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} P_{s,i}\left[U\left(c^*_{s,i} + \frac{F_{s,i}}{P_{s,i}}\right) - U\left(c^*_{s,i}\right)\right], \tag{A6}$$

where T is the maximum number of years of life, F equals a cohort's social security flows (either minus taxes or plus benefits) per original member, $c^*_{s,i}$ is some benchmark level of per capita consumption and $P_{s,i}$ is the surviving fraction of the population in state i at date s. The weights $\pi_{t,i}$ equal the cohort's initial population along trajectory i as a fraction of the sum of the cohort's initial populations along all trajectories. (We also provide calculations for the case in which we ignore differences in initial populations within and across cohorts and set $\pi_{t,i} \equiv 1/N$, where N is the number of trajectories.)

Because we will be evaluating this difference in utilities using a Taylor approximation, it is better to consider small variations, so we will look not at V_t as specified in expression (A6), but at the difference between V_t and the value of some benchmark system of constant taxes and benefits,

$$\Delta V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} P_{s,i}\left[U\left(c^*_{s,i} + \frac{F_{s,i}}{P_{s,i}}\right) - U\left(\frac{c^*_{s,i} + \overline{F}_s}{P_{s,i}}\right)\right]. \tag{A7}$$

Letting $f_{s,i}$ equal per capita flows, and taking second-order Taylor approximations around the benchmark social security system of the utility variations in Eq. (A7), we get:

$$\Delta V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} P_{s,i}\left[U'_{s,i}\cdot\left(f_{s,i} - \overline{f}_s\right) + \frac{1}{2}U''_{s,i}\cdot\left(f_{s,i} - \overline{f}_s\right)^2\right]. \tag{A8}$$

Assuming that households have CES utility with risk-aversion parameter γ, we know that $U'' = -\gamma\frac{U'}{c}$, where c is consumption including the base level of social security flows, $c^* + \overline{f}$, around which the Taylor approximation is being taken. Thus, Eq. (A8) can be rewritten:

$$\Delta V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} P_{s,i} U'_{s,i}\cdot\left[\left(f_{s,i} - \overline{f}_s\right) - \frac{1}{2}\gamma\frac{\left(f_{s,i} - \overline{f}_s\right)^2}{c_{s,i}}\right]. \tag{A9}$$

Without loss of generality, we can drop the term $\overline{f}_s$ when it first appears in Eq. (A9), because this is a constant term that does not vary across social security systems. Dropping this term, using the facts that $f = F/P$ and $c = C/P$, and substituting Eq. (A5) into Eq. (A9) yields:

$$\Delta V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} \left\{\frac{h_{s,i}\left(x_{s,i}\right)}{E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]\frac{h_s(x_s)}{h_t(x_t)}\right\}}\right\}\cdot\left[F_{s,i} - \frac{1}{2}\gamma\frac{\left(F_{s,i} - \overline{F}_s\right)^2}{C_{s,i}}\right], \tag{A10}$$

where C is a generation's consumption per initial individual at the benchmark level.

Expression (A10) will be the basis for our valuation of flows. As discussed in the introduction, risk aversion affects valuation in two ways, through the variation in the value taken by the function h (the term in curly brackets in Eq. (A10)) and through the impact of fluctuations on the flows themselves (the next term).

As a benchmark, note that for risk-neutrality, $h(x) \equiv 1$ and $\gamma = 0$, so Eq. (A10) reduces to:

$$\Delta V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} \frac{1}{E_t\left[\prod_{u=t}^{s}(1+r_u)\right]} F_{s,i}. \tag{A11}$$

That is, under risk-neutrality, we should divide the average flow at each date by the average discount factor. [26] After summing over the trajectories for each generation in Eq. (A11), we will divide by that generation's present expected discounted value of earnings,

$$PDVE_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+R} \frac{1}{E_t\left[\prod_{u=t}^{s}(1+r_u)\right]} Y_{s,i}, \tag{A12}$$

where R is the retirement age and $Y_{s,i}$ is the generation's earnings in year s along path i. This normalization serves two purposes. First, it removes the productivity growth trend to avoid giving more weight to later generations when we calculate an average over generations. Second, it scales ΔV so that it is expressed as a fraction of expected lifetime earnings. We average these generation-specific ratios by average initial generation size to form an average estimate of expected utility.

Parameterization

To implement expression (A10), we need to make three sets of parameter assumptions.

Risk-aversion parameter, γ

We consider four cases, $\gamma = 0$ (neutrality) 3, and 5, and $\gamma = 3$ when working and 5 when retired.

State-contingent valuation

There are a variety of possibilities here. One is to assume that h is related to the contemporaneous wage, $h_{s,i} \sim w_{s,i}^{-\gamma}$, which would make sense if consumption were proportional to labor income. Another is to assume that h is related to the initial wage along the cohort's trajectory, $h_{s,i} \sim w_{t,i}^{-\gamma}$, which would take into account the fact that consumption is financed to some extent by past saving and social security benefits. The approach that we adopt as our base case is to assume that $h_{s,i} \sim w_{s,i}^{-\gamma}$, during working years and $h_{s,i} \sim w_{t,i}^{-\gamma}$ during retirement years. As an alternative assumption, we assume that $h_{s,i}$ when retired is based on the wage in the last year of work as might be more appropriate in an economy in which the well-being of the elderly depends on contemporaneous wages.

There remains the question of how to scale the marginal utilities of different generations. One approach would be to assume a constant utility function over time, in which case successive generations would,

[26] Note: we are implicitly assuming access to annuity markets; had we not, then there would be an extra Ps multiplied by the discount factors, so that Eq. (A11) would have become:

$$\Delta V_t = \sum_i \pi_{t,i} \sum_{s=t}^{t+T} \frac{1}{E_t\left[\prod_{u=t}^{s}(1+r_u)P_s\right]} F_{s,i}.$$

on average, have lower and lower levels of marginal utility as a consequence of trend productivity growth. This approach would tend to make social security systems that transfer resources from future generations to current ones look better than those that do not involve such transfers. Although such transfers would be relatively subtle for the systems we are considering here, we nevertheless wish to avoid confusing intergenerational transfers with risk-sharing. Thus, we scale each generation's marginal utilities by that generation's average initial wage, that is, $h_{s,i}=\left(\frac{w_{s,i}}{\overline{w}_t}\right)^{-\gamma}$ during working years and $h_{s,i}=\left(\frac{w_{t,i}}{\overline{w}_t}\right)^{-\gamma}$ during retirement years.

Benchmark level of consumption, C

Here, we use just two such numbers (relative to trend) to keep fixed across scenarios, one for workers (C^L) and one for retirees (C^R), rather than age-specific values. The calculation of C^L and C^R will be discussed shortly.

In summary, with our parameterization, Eq. (A10) becomes:

$$\Delta V_t = \sum_i \pi_{t,i}\left\{-\sum_{s=t}^{t+R}\frac{\left(\frac{w_{s,i}}{\overline{w}_t}\right)^{-\gamma}\left[T_{s,i}+\frac{\gamma\left(T_{s,i}-\overline{T}_s\right)^2}{2C_s^L}\right]}{E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]\left(\frac{w_s}{w_t}\right)^{-\gamma}\right\}}\right. \qquad \text{(A10')}$$

$$\left.+\sum_{s=t+R+1}^{t+T}\frac{\left(\frac{w_{t,i}}{\overline{w}_t}\right)^{-\gamma}\left[B_{s,i}-\frac{\gamma\left(B_{s,i}-\overline{B}_s\right)^2}{2C_s^R}\right]}{E_t\left\{\left[\prod_{u=t}^{s}(1+r_u)\right]\right\}}\right\},$$

with everything except taxes, $T_{s,i}$, and benefits, $B_{s,i}$, the same across different social security scenarios.[27] For the benchmark values of taxes and benefits, $\overline{T}_s$ and $\overline{B}_s$, we use the average values for the US benefits-adjust system. Note that the average values are indexed by time, because they will follow the trend in productivity. To calculate C_s^L and C_s^R, we compute the ratios of taxes to consumption and adjusted benefits to consumption for the US population in 2003, based on populations aged 20–64 and over 65 (excluding those in nursing homes), respectively, using OASI payroll taxes and benefits and adjusting benefits (and consumption of beneficiaries) downward until they equal taxes in the aggregate.[28] The resulting ratios are 0.103 for taxes relative to non-retiree consumption, and 0.235 for adjusted benefits relative to adjusted retiree consumption. We multiply the inverses of these ratios by $\overline{T}_s$ and $\overline{B}_s$ in a given year to get the values of worker and retiree consumption around which the expected utility approximation is computed.

Finally, as in the case of risk-neutrality, we divide expression (A10′) by the cohort's present expected discounted value of earnings, as given in Eq. (A12), in order to weight the results equally across generations and express them as a share of lifetime earnings.

Estimating social welfare

Our methodology compares social security systems that are normalized to be of the same size (in terms of taxes relative to earnings) because larger systems, which make larger transfers to initial transition generations, provide lower present value returns to the generations that follow. However, even this normalization may fail to neutralize all differences in the relative treatment of transition generations. In particular, systems that rely on immediate adjustments to benefits (in particular the US Benefit Adjust system) will impose greater uncertainty on initial transition generations, and in so doing may impose less uncertainty on subsequent generations. Therefore, such systems may appear more attractive than they actually are if we ignore the effects on the welfare of transition generations. Note that this is not simply an issue of what is being adjusted (taxes or benefits), but rather of which generations are affected. For example, all variants of the Swedish system also rely solely on benefit adjustments, but these adjustments are spread over more generations. That is, while the US Benefit Adjust system achieves current cash flow balance by reducing the benefits of current retirees, the Swedish systems also reduce the future benefits of current workers by reducing their current accumulations of notional pension wealth.

To deal with this issue, we construct an expanded welfare measure that takes account of transition generations, both at the beginning and the end of our simulation period. Our basic approach is to use a discounted sum of the values of ΔV as calculated above and to include as well the partial values for initial and terminal cohorts.

For full generations whose birth occurs after our initial simulation year 0 and whose final death is before year L (the last year of our computation), we simply take the values of ΔV but discount them back to year 0; expression (A10′) becomes:

$$\Delta V_t^F = \sum_i \pi_{t,i}\sum_{s=0}^{t+T}\left\{\frac{h_{s,i}\left(x_{s,i}\right)}{E_0\left\{\left[\prod_{u=0}^{s}(1+r_u)\right]\frac{h_s(x_s)}{h_t(x_t)}\right\}}\right\}\cdot\left[F_{s,i}-\frac{1}{2}\gamma\frac{\left(F_{s,i}-\overline{F}_s\right)^2}{C_{s,i}}\right] \qquad \text{(A13)}$$

For generations born in year $t>L-T$, which will still be alive at the end of our period of measurement, we compute partial sums, starting in the year of birth and going through year L, and discount from the year of birth back to year 0:

$$\Delta V_t^P = \sum_i \pi_{t,i}\sum_{s=t}^{L}\left\{\frac{h_{s,i}\left(x_{s,i}\right)}{E_0\left\{\left[\prod_{u=0}^{s}(1+r_u)\right]\frac{h_s(x_s)}{h_t(x_t)}\right\}}\right\}\cdot\left[F_{s,i}-\frac{1}{2}\gamma\frac{\left(F_{s,i}-\overline{F}_s\right)^2}{C_{s,i}}\right] \qquad \text{(A14)}$$

For existing generations, born in year t<0, who are already alive as of year 0, we define a partial value of ΔV, say ΔV^E, as

$$\Delta V_t^E = \sum_i \pi_{0,i}\sum_{s=0}^{t+T}\left\{\frac{h_{s,i}\left(x_{s,i}\right)}{E_0\left\{\left[\prod_{u=0}^{s}(1+r_u)\right]\frac{h_s(x_s)}{h_0(x_0)}\right\}}\right\}\cdot\left[F_{s,i}-\frac{1}{2}\gamma\frac{\left(F_{s,i}-\overline{F}_s\right)^2}{C_{s,i}}\right] \qquad \text{(A15)}$$

That is, we treat these cohorts as if they are born in year 0 with the surviving population as of year 0 and with $T+t<T$ years of life remaining.[29]

We sum over the three groups to get the sum of the ΔVs, weighting each generation by its average initial size, and then divide by the sum of the present values of earnings for the different cohorts, computed for the three groups in parallel fashion.

[27] Note that in case where γ is allowed to differ between work years and retirement, the discount factor in the second term on the right-hand side of Eq. (A10′) will include another term that reflects both values of γ.

[28] The data for this calculation are based on the 2003 Consumer Expenditure Survey and other sources of data as detailed in the US National Transfer Accounts (NTA) at http://www.schemearts.com/proj/nta.

[29] Note that the terms $h_{x,i}(x_{x,i})$ are normalized relative to year 0 wages in this case, to be consistent with the calculation starting in year 0.

Table A1
Summary statistics for eight pension plans based on 1000 stochastic paths of 500 years.

	US tax adjust	US benefit adjust	US 50–50 adjust	NDC Sweden	NDC (g, $A=1$) symmetric	NDC (g, $A=0.5$) symmetric	NDC ($n+g$, $A=0.5$) symmetric	German
(Not weighted by generation size)								
NPV ($\gamma=0$)								
Mean	−0.0669	−0.0601	−0.0637	−0.0620	−0.0583	−0.0582	−0.0570	−0.0673
Variance across trajectories	0.00206	0.00102	0.00137	0.00095	0.00098	0.00102	0.00101	0.00178
Generations	0.00130	0.00048	0.00074	0.00039	0.00048	0.00051	0.00051	0.00095
EU								
$\gamma=3$	−0.0892	−0.0789	−0.0838	−0.0807	−0.0773	−0.0776	−0.0763	−0.0864
$\gamma=5$	−0.1488	−0.1379	−0.1430	−0.1404	−0.1359	−0.1366	−0.1356	−0.1444
$\gamma=3$, ret. w	−0.0957	−0.0859	−0.0906	−0.0871	−0.0841	−0.0844	−0.0832	−0.0928
$\gamma=3, 5$	−0.1288	−0.1158	−0.1219	−0.1220	−0.1108	−0.1122	−0.1102	−0.1256
HE								
NPV	0.999563	0.999897	0.999834	0.999952	0.999902	0.999893	0.999880	0.999788

References

Alho, Juha M., Svend, E., Jensen, Hougaard, Lassila, Jukka (Eds.), 2008. Uncertain Demographics and Fiscal Sustainability. Cambridge University Press.

Alho, J.M., Lassila, J., Valkonen, T., 2005. Demographic Uncertainty and Evaluation of Sustainability of Pension Systems. In: Holtzmann, R., Palmer, E. (Eds.), Non-Financial Defined Contribution (NDC) Pension Schemes: Concept, Issues, Implementation, Prospects. World Bank, Washington D.C, pp. 95–115.

Auerbach, Alan J., Hassett, Kevin, 2002. A new measure of horizontal equity. The American Economic Review 1116–1125 September.

Auerbach, Alan J., Lee, Ronald D., 2009. Notional Defined Contribution Pension Systems in a Stochastic Context: Design and Stability. In: Brown, J., Liebman, J., Wise, D. (Eds.), Social Security Policy in a Changing Environment. University of Chicago Press, Chicago, pp. 43–68.

Board of Trustees, 2004. Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds. The 2004 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds. U.S. Government Printing Office, Washington, D.C.

Börsch-Supan, Axel, Wilke, Christina B., 2004. The German Public Pension System: How it Was, How it Will Be. NBER Working Paper No. 10525.

Börsch-Supan, Axel, Reil-Held, Annette, Wilke, Christina Benita, 2003. How to Make a Defined Benefit System Sustainable: The 'Sustainability Factor' in the German Benefit Indexation Formula Arbeitspapier 37-2003. MEA - Mannheimer Forschungsinstitut Ökonomie und Demographischer Wandel, Universität Mannheim.

Goldstein, Joshua R., Sobotka, Tomáš, Jasilioniene, Aiva, 2009. The end of "lowest-low" fertility? Population and Development Review 35 (4).

Gruber, Jonathan, Wise, David A., 1999. Social Security Around the World. University of Chicago Press, Chicago.

Holtzmann, R., Palmer, E. (Eds.), 2005. Non-Financial Defined Contribution (NDC) Pension Schemes: Concept, Issues, Implementation, Prospects. World Bank, Washington D.C.

Lee, Ronald, Carter, Lawrence, 1992. Modeling and forecasting U.S. mortality September 1992 Journal of the American Statistical Association 87 (419), 659–671 and "Rejoinder," same issue, pp. 674–675.

Lee, Ronald, Tuljapurkar, Shripad, 1998. Stochastic Forecasts for Social Security. In: Wise, David (Ed.), Frontiers in the Economics of Aging. University of Chicago Press, Chicago, pp. 393–420.

Lee, RonaldD., Anderson, Michael W., Tuljapurkar, Shripad, 2003. Stochastic Forecasts of the Social Security Trust Fund. Report for the Social Security Administration (January) posted on the web site of the Office of the Actuary.

Lee, Ronald, 2005. Comments on The Rate of Return of Pay As You Go Pension Systems. 2005 In: Holtzmann, R., Palmer, E. (Eds.), Non-Financial Defined Contribution (NDC) Pension Schemes: Concept, Issues, Implementation, Prospects. World Bank, Washington D.C.

Lee, Ronald, Miller, Timothy, Anderson, Michael, 2004. Stochastic Infinite Horizon Forecasts for Social Security and Related Studies. National Bureau of Economic Research. Working Paper No. 10917 (http://papers.nber.org/papers/w10917.pdf).

Legros, Florence, 2003. Notional Defined Contribution: A Comparison of the French and German Point Systems. paper presented at the World Bank and RFV Conference on NDC Pension Schemes, Sandhamn, Sweden Sept. 28–30, 2003.

Ludwig, Alexander, Reiter, Michael, 2010. Sharing demographic risk — who is afraid of the baby bust? American Economic Journal: Economic Policy 2 (4), 83–118.

Poterba, James M., 2001. Demographic structure and asset returns. The Review of Economics and Statistics 83 (4), 565–584 November.

Settergren, Ole, Mikula, Boguslaw D., 2005. The rate of return of pay-as-you-go pension systems: a more exact consumption-loan model of interest. Journal of Pension Economics and Finance 4 (2), 115–138. doi:10.1017/S1474747205002064.

Swedish Social Insurance Agency, 2008. Orange Report: The Annual Report of the Swedish Pension System. available online at http://www.forsakringskassan.se/irj/go/km/docs/fk_publishing/Dokument/Publikationer/%C3%85rsredovisningar/Orange%20Rapport%202008%20engelsk.pdf.

United Nations, Department of Economic Social Affairs Population Division, 2007. World Population Aging 2007 ST/ESA/SER.A/260.