Preliminary Draft
Please do not cite or quote without permission.

Welfare and Generational Equity in Sustainable Unfunded Pension Systems

Alan J. Auerbach
Economics and Law
University of California, Berkeley

Ronald Lee
Demography and Economics
University of California, Berkeley

October 2007

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 the grant R37-AG025488. 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 provided valuable comment and advice that helped us model the actual Swedish system. The authors gratefully acknowledge the assistance of Erin Metcalf, Anne Moore and Carl Boe.

Introduction

Actual and projected population aging threatens the financial viability of Pay-As-You-Go (PAYGO) pension programs in many countries. These programs promise a level of benefits that cannot be sustained given current tax rates, so deep structural reforms are expected and in some countries have already occurred. Beyond the problem of fiscal instability, most PAYGO programs are Defined Benefit and create incentives for early retirement (Gruber and Wise, 1999) and may distort labor supply decisions over the whole life cycle. Furthermore, because these programs create pension wealth that is not backed up by assets, they may weaken incentives to save and thereby lead to a lower capital stock.

Reforms that adjust the general level of taxes or benefits in PAYGO programs can address the problem of fiscal sustainability, but would have little effect on incentives for work or saving. A new kind of pension program, called Notional Defined Contribution or Non-financial Defined Contribution (NDC), is intended to address both 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).

NDC plans are intended to mimic the structure and incentives of Defined Contribution plans in that individuals contribute to their own 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. The specified rate of return earned by the accounts and paid by the annuity is in principle linked to the growth rate of total wages, which is the sustainable rate of return in a steady state PAYGO system. The terms of the annuity reflect current mortality conditions, so effectively benefits are indexed to life expectancy. These provisions should make NDC systems fiscally sustainable. 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 Defined Benefit systems. Because these NDC systems are PAYGO and do not hold significant assets, they are easier to implement then funded systems since the painful transitional costs, which can total one, two or three years worth of GDP, do not have to be borne. However, the other side of this coin is that NDC programs would still be expected to weaken the incentive to save and to reduce the aggregate capital stock as would other kinds of PAYGO programs.

Many actual NDC programs, including Sweden's, set the rate of return equal to the growth rate of the wage rate rather than the growth rate of total wages, so that it does not reflect the growth rate of the labor force which in many countries is expected to be negative. Demographic change and decline will lead such programs into fiscal problems, so they include additional mechanisms to achieve fiscal sustainability.

The costs of different reform plans will have different consequences for generations during the transition to the new system, resulting from the interaction of the reform plan with the particular initial demographic conditions of each country such as the baby boom and bust in the United States. In this paper, our goal is not to analyze generational risk sharing during this specific transitional period for any particular country. Rather we analyze the generational uncertainty and risk sharing in a more general context of economic and demographic uncertainty, with the goal of finding more general properties of the pension systems that do not result from some particular demographic circumstances or transition policy strategies.

Any pension plan must operate in an environment of demographic and economic uncertainty. Fiscally sustainable plans must make adjustments in light of this changing environment. Different plans will make different adjustments, leading to differences in the way that risks are shared among generations. These differences will influence the level of uncertainty faced by a typical generation, and also affect the extent to which welfare varies across generations. This paper modifies a stochastic forecasting model (Lee and Tuljapurkar, 1998, and Lee *et al*., 2003) to investigate both these consequences of plan structure. An earlier paper (Auerbach and Lee, 2007) used the same setup to investigate the fiscal stability and sustainability of different PAYGO pension plans, and we will build on those results. Here we will restrict our attention to plans that are fiscally sustainable.

We will consider a number of actual and hypothetical PAYGO pension structures, including: 1) versions of the US Social Security system in which taxes or benefits are adjusted annually to maintain fiscal balance with zero debt or assets in every period; 2) the actual Swedish NDC system, together with several modifications of it developed in our earlier paper; 3) the actual reformed German system[1] For each sustainable system, we will consider some descriptive measures of uncertainty in outcomes for generations, and we will also consider expected utility measures based on simplifying assumptions. Finally, we will compare their performance using a new measure of intergenerational equity which weights more heavily differences in outcomes for neighboring generations than for distant ones (Auerbach and Hassett 2002).

[1] We have also simulated the actual US Social Security system, but because it is not sustainable under current law there is little interest in comparisons to sustainable systems.

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 the representative individual in each generation. Since we do not consider individual heterogeneity within a generation, many details of a pension plan are irrelevant. Nor do we model behavioral responses to differences in pension plans. Of course, these aspects of pension structure and behavioral response are very important in their own right, but in this paper we simplify drastically in order to focus on macro uncertainty and inter-generational differences. In our discussion of various pension plans below, we will therefore only discuss those aspects which are relevant for this undertaking.

One way in which we simplify is to impose certain common characteristics across the different pension systems. We scale all pension systems so that they have average contributions of the same scale as the current US system, and assume that all individuals work until age 67, the long-run normal retirement age under current law, and that all individuals are retired thereafter. Modifications that these assumptions impose on the different schemes are discussed further below.

US Social Security

We will consider three plans derived from the current law US Social Security program. For the Old Age and Survivors (OASI) portion of this system, the current law tax rate is set at 10.6 percent of taxable payroll, and we set the tax rate at this level for all of the systems we consider in this paper.[2] The current law average benefit level for a generation is set as a certain percentage

[2] For systems in which contributions fluctuate in order to maintain budget balance, as under the US "Tax Adjust" plan, we scale the system so that the average tax rate over all trajectories is 10.6 percent.

of the average real wage at age 60 with some upward adjustment that depends on the extent to which the generation continues to work past the early age at retirement at age 62. As the population ages and the ratio of retirees to workers rises, the balance of tax revenues and benefit costs will change so the system is not fiscally stable. We create fiscally sustainable systems by modifying this current law structure.

In the “Tax Adjust” program, the age schedule of taxes is adjusted by a multiplicative factor each period so as to produce the amount of revenue needed to cover that period’s benefits, so that the system is in perfect balance each period. In the “Benefit Adjust” program, the age schedule of benefits is similarly adjusted each period so that the costs of benefits exactly equal that period’s tax revenues. In the “50-50 Adjust” program, balance is achieved half by adjusting taxes and half by adjusting benefits. Further details will be provided when the simulations are discussed.

Swedish Notional Defined Contribution System

We describe the current Swedish system relatively briefly here. Our earlier paper provides a more detailed discussion and analysis of this system, and descriptions can also be found in Holtzmann and Palmer, eds. (2005).

The actual Swedish NDC program specifies that the rate of return earned on accounts in each year t equals the contemporaneous growth rate of the wage level, g_t. Because the sustainable rate of return in a steady state PAYGO system is $g+n$, where n is the growth rate of the labor force, this system may have problems with stability. Recognizing this, the program also has a balancing mechanism, or “brake” (to be described shortly), that goes into effect when program assets reach a certain level, reducing the rate of return earned on accounts and annuities until asset levels are restored. In this way, demographic variations influence the system through the back door. Therefore the effects of economic and demographic variations are likely to be distributed across the generations quite differently than in the Social Security case.

At the time of retirement, which we assume to be at age 67, the individual's account is converted into an annuity based on the account's notional balance as of that date. The terms of the annuity reflect mortality conditions at the time of conversion and a rate of return which is set equal to a pre-specified expected rate of wage growth taken in the Swedish system to be .016, but in our simulations taken to be the actual growth rate at the time, which has a mean value of .011, the underlying average rate of labor productivity growth in our simulations. The benefit stream is then calculated to be constant over the individual's (or in our case, the generation's) remaining years, but it is adjusted for *ex post* deviations of wage growth from the assumed rate In our simulations we also adjust the annuity level to reflect changes in mortality after retirement, although we found in our earlier paper that including this post-retirement updating had a minor impact on system stability..

Those who designed the Swedish system expected that with a rate of return equal to g and with the benefit actuarially reflecting mortality conditions at the time of retirement the system would be fiscally quite stable. Nonetheless they realized that it could still go into debt, for example if the population growth rate were negative for a sustained period, so they designed a brake, which would automatically reduce the rate of return under certain conditions, and thereby guarantee long term fiscal stability. To this end they defined a balance ratio b:

$$(1) \qquad b = \frac{F + C}{NPW + P}$$

In this expression, F is the level of a financial asset, similar to the trust fund of US Social Security, but able to take negative values in case of debt. 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 times the expected length of time between tax payments and receipt of benefits which is roughly thirty years. 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, 2005; Lee, 2005).

The denominator of the expression is the total pension liability of the system. The first term, *NPW*, is the present value of obligations to current workers. The second term, *P*, approximates the present value of benefits due to current retirees. This balance measure does not rely on explicit projections of demographic and economic variables which might be distorted by political pressures, and can be calculated from current cross sectional data.

If the balance b falls below 1.0 then the brake is activated. When the brake is active, the gross rate of return $1+g_t$ that is used in computing both accruals of notional pension wealth and the growth of annuity payments is reduced by a factor of $b_t < 1.0$, that is the gross rate of return becomes $(1+g_t)b_t$. When the brake is not in place, the rate at which pension benefits grow is $(1+g_t)/.016$, where .016 is the assumed central value of g. That is, if wages grow more rapidly than .016 then the benefit is raised. Multiplication by b_t makes it more likely that the level of benefits will decline rather than rise.

The brake remains active until the product of the balance levels b_t for all the years since it is activated first 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 for continuing activation of the brake is to ensure that the temporary application of the brake has no long term effect on the level of benefits. That is, there is a catch-up period while the brake is in effect, 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 reducing the level of benefits only temporarily while the brake is active, the saving coming from a temporary drop below the unrestrained benefits trajectory rather than a move to a lower benefits trajectory.

Note that the brake is asymmetric, and does not prevent the unlimited growth of the trust fund F. 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.

We have simulated the Swedish system with the brake as just described, and with a rate of return of g. However, we have also defined alternative NDC systems that are closely modeled on the Swedish system but which have modified versions of the brake, or which have a rate of return of $n+g$, or both. We will now explain these modifications.

By setting the rate of return to $n+g$ rather than g, a system should track the varying demographic context and therefore have less need of the brake. 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 benefits adjustments.

We have noted that the Swedish brake is asymmetric. In considering alternatives to the existing Swedish system, we will focus on systems in which there is a symmetric brake, that is both a brake and an accelerator, such that the rate of return rises when b>1.0 in addition to falling when b<1.0. Because we impose the brake 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 have also developed systems with a somewhat different braking structure. Let r_t^a be the adjusted net rate of return. In the Swedish system this is $r_t^a = (1 + r_t)b_t - 1$; that is, it equals the adjusted gross rate of return minus 1. At low values of *b*, this takes a fierce bite out of the pension, and if *b*=0, which is possible, the net rate of return is -1. To soften the action of the brake we introduce a scaling factor *A*, between 0 and 1, as follows:

(2) $$r_t^a = (1 + r_t)[1 + A(b_t - 1)] - 1$$

When *A*=1, this reduces to the original Swedish brake. When *A* is less than 1, the adjusted rate of return varies less than in proportion to the b_t, and full confiscation does not occur until *b* falls to 1-1/*A* <0. If we set A=0 then there is no brake at all. In some of our simulations we use a brake with *A*=.5, which does not become fully confiscatory unless *b* falls to -1. In our simulations or *r*=*g* systems this almost never happens: for 2 out of 1000 trajectories of 500 years each with the asymmetric brake, and 7 out of 1000 with the asymmetric brake.

German System

A simplified description of the German pension system and our implementation of it follow. For more details, see Börsch-Supan and Wilke (2003), Börsch-Supan, Reil-Held and Wilke (2003), and Ludwig and Reiter (2006). Under this system, each pension beneficiary *i* receives a payment in year *t* equal to:

(3) $$B_{t.i} = PV_t * EP_i * AA_i$$

where PV_t is the current pension value in year *t*, EP_i is the individual's "earning points" collected until retirement, and AA_i is an actuarial adjustment based on when the pensioner retired. Earnings points are an increasing function of an individual's earnings, scaled to equal 1.0 for the average-wage individual for each year the individual worked. Given that we are ignoring intra-heterogeneity, we set EP_i constant at the assumed number of years of labor force participation. The actuarial adjustment AA_i scales benefits up or down according to the age at which the individual retires. We just assume retirement at 67 and set $AA_i = 1$. Thus, the benefits formula reduces to:

(3') $B_{t.i} = B_t = Y*PV_t$

where Y is the number of years worked until retirement. (In the end, the value assumed for Y does not matter, because we scale the size of the system to conform so that the average tax rate equals that of the US system.) Note that pensions are set up so that retirees of different ages get the same benefit, since earnings points are based on wages relative to the average wage when one worked, not on one's absolute wages.

The pension in year t evolves according to:

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

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 α is now set to 0.25. In our terms, pension benefit growth is linked to g, but adjusted for fluctuations in CR and OA.

Substituting (3') into (4), we get:

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

Expression (5) describes the calculation of benefits. Taxes are adjusted each year as a residual so that taxes and benefits are equal in the aggregate. As to the system's scale, we adjust the initial level of benefits so that the average level of taxes across time and trajectories is the same as under the benchmark US system.

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. Another factor reduces benefits if the old-age dependency ratio is rising.

Discussion of These Systems

We have simplified our representation of these systems to focus on their intergenerational risk sharing characteristics, but even so they remain complex, particularly the German and the NDC plans. Nonetheless, we can highlight one key difference among them. In the US Tax Adjust system, the benefits are fixed and the burden of fiscal adjustment falls entirely on the taxpayers. In contrast, in the US Benefit Adjust and all the NDC systems, the tax rate is fixed and the burden of fiscal adjustment falls entirely on the beneficiaries. Finally, in the German system and in the US 50-50 system, the burden of adjustment is shared by tax payers and beneficiaries.

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 by Lee and collaborators (see also the approaches in Alho *et al*., 2005, and Alho *et al*., in press).

In this earlier work, Lee and Tuljapurkar developed a model to generate stochastic long-term forecasts of the US Social Security system (Lee and Tuljapurkar, 1998; Lee *et al*., 2003). In this model, the log of each age specific mortality rate is 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, henceforth TR04, as discussed in Lee (1993) and Lee and Tuljapurkar (1994). Immigration is taken as deterministically given by the intermediate assumption in the Social Security Trustees Report of 2004 (see Lee *et al*., 2004 for treatment of immigration as stochastic in this framework). The fertility and mortality processes are fit to historical US data. Together these assumptions and models are the basis for stochastic simulations of the population of the United States (Lee and Tuljapurkar, 1994). Monte Carlo methods were used to generate a large number of stochastic sample paths, and from these the probability distribution of demographic outcomes such as the OADR (population age 65 and over divided by population age 20 to 64)[3] were calculated.

To give a sense of the demographic variations that result from this approach, Figure 1 plots the simulated OADR along 15 randomly chosen sample paths from among the 1000 we use for our analysis. Because the time series models for fertility and mortality were fit to US data,

[3] Note that this is a slightly different definition of the OADR than that used by the German system, discussed earlier, which starts from age 15 in computing the denominator.

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.

The Social Security model builds on the stochastic population using cross-sectionally estimated 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. A Monte Carlo simulation generates sample paths of productivity, and earnings and tax payments by age. The same simulation can also generate sample 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, consistent with our modeling assumption that retirement under each system occurs at age 67. While the full model includes a procedure for tracking the Social Security trust fund, in the versions of the US system analyzed here, the trust fund is uniformly equal to zero. For systems that can accumulate financial assets and liabilities – variants of the Swedish system – 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 (they are modeled jointly using VAR). The productivity growth rate has a positive mean of 1.1 percent (based on the

Actuary's assumption in TR04), so productivity, earnings, taxes and benefits are all trending upwards.

For our purposes we wish to abstract from the particular demographic situation in the United States today so that we may derive results that are of more general applicability. We modify the model described above so it converges to an approximately stationary stochastic population distribution. This is accomplished by setting the mortality drift term to zero. 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. 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. More important is the nonstationarity of productivity levels. Productivity growth is stationary, but the productivity levels to which it cumulates are not.

While this description of the model glossed over many details, it should convey a general idea of how we simulate the pension systems. For the actual simulations, we consider a sample of 1000 randomly drawn trajectories, using the same sample of 1000 for each of the pension systems considered. 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 and to reduce the relevance of the initial conditions. 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.[4]

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

Previous Results on Fiscal Stability

Our earlier paper used the same stochastic simulations to consider the fiscal stability of the basic Swedish system and several variants, including some analyzed in this paper as well. We found that the current Swedish NDC system with its asymmetric brake effectively avoids accumulating large debt but sometimes does accumulate substantial assets which could otherwise have been returned as higher benefits for plan participants. We developed a symmetric version of the brake with a parameter regulating its severity of action, already described above. This symmetric brake insures two-sided fiscal stability for the NDC system, based like the existing Swedish system on the growth rate of the wage, g, even if it operates less severely than the current Swedish asymmetric brake. As discussed above, we also developed an NDC system in which the stipulated rate of return responds to demographic shocks as well as to variations in the productivity growth rate. This system is inherently more stable than the system based on the wage rate alone when the two systems are compared in operation without a brake. We also found that economic factors, as well as demographic ones, contribute significantly to the volatility of NDC pension systems.

Measures of Outcomes for Generations

Our earlier paper was largely limited to considering issues of fiscal stability in relation to system design. Here we will restrict our attention to systems that are fiscally stable and examine aspects of risk sharing across generations. This requires different kinds of measures, which we now discuss.

We will begin by considering two common measures of pension program outcomes. One is the Net Present Value (NPV, or simply V in our equations) of expected lifetime contributions or taxes minus benefits. We estimate the NPV for each generation along each simulated sample path for the system, which includes a simulated path for real interest rates. For the generation born in year t and age s in year $u=t+s$, on some particular sample path i, with a real interest rate r_u when this generation is age s; let $F_{u,i}$ be the survival weighted per capita benefit minus tax payment for this generation in year $u=t+s$. Then, as cohorts exist for 106 years, the NPV for this generation on this sample path is given by:

(6) $$V_{t,i} = \sum_{s=t}^{t+106} \left(\prod_{u=t}^{s} (1 + r_{u,i})^{-1} \right) F_{s,i} \; .$$

Averaging $V_{t,i}$ over i, or over both i and t, gives the expected NPV for generation t and the global expected NPV.[5] Because the mean productivity growth rate is positive, taxes and benefits are trending upward, and so is the NPV. For most purposes it is preferable to remove this trend in NPV by expressing it relative to the present value of lifetime earnings which results in a trend free measure.

We can also estimate the Implicit or Intrinsic Rate of Return (IRR, or simply ρ in our equations) for each generation on each sample path as that constant value for the real discount rate $r_u = \rho$ for all u, for which the NPV $V_{t,i} = 0$. We obtain a different $\rho_{t,i}$ for each generation and sample path, and can again calculate the global average value. We know that in a deterministic steady state ρ equals $n+g$, but the IRR is a highly nonlinear measure so the average over the stochastic sample paths can be rather different than this. Figure 2 **[to be added]** illustrates IRR for generations along the same 15 sample paths for which Figure 1 showed the OADR.

[5] Back of the envelope calculations indicate a deterministic steady state NPV value of -.04 times the PV of lifetime earnings, based on a thirty year difference in average ages of paying payroll tax and receiving benefits, with mean productivity growth of .011 and mean real interest rate of .03. This in fact matches closely the mean of the simulation results.

In combining results for different generations it may be desirable to weight these by the size of the generation at birth. The structure of some pension plans may result in higher or lower benefits for larger generations, for example. A program that paid lower benefits to larger generations would have a larger global mean IRR or NPV without weighting by size, but the welfare implications would be unclear. For example, under the US Tax Adjust system 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. Under the Benefit Adjust system it is better to be in a small cohort, other things equal, since for a given amount of tax revenues benefits will be shared among fewer and will be higher. For most systems it is not obvious whether large or small generations will be favored, and some might be favored by population weighting for some measures and not for others.

The dispersion of V and ρ about their global means describes the uncertainty faced by a generation, and these can be summarized by their variances which can be compared across pension systems, as we will do later. But we might also want to evaluate these different distributions of outcomes in terms of utility. 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 derive a local approximation of the impact of different systems on expected utility using a series of simplifying assumptions relating 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.

To this point we have considered measures of the uncertainty of pension program outcomes and the tradeoff of this uncertainty against the mean return. The expected utility measure reflects this trade-off, and some might argue that the vector of expected utilities for different cohorts provides all the information needed to evaluate social welfare. But we may additionally care about how generations fare relative to other generations near to them in age along any particular trajectory, 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 percent of total income to eliminate horizontal inequality.

It is important to stress that this measure of horizontal equity is derived as one component of social welfare from a social welfare function that may weigh inequality among similarly situated individuals differently from inequality among other individuals. Thus, it would be double-counting to add this welfare measure to an overall measure of welfare derived from a social welfare function, such as the sum of expected utilities. We provide this measure separately simply to give a sense of how the different systems perform in one particular dimension.

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 adopt 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 .10 times income. We estimate horizontal equity based on our NPV measure and our Expected Utility measure without risk aversion, in each case measuring income as the present value of earnings along the trajectory and taxes and the net present value of taxes minus benefits along the trajectory.

Results

Description of the Uncertainty Faced by Generations

We begin by considering Figure 3, which gives the frequency distribution of NPV outcomes for generations, and Figure 4, which does the same for the IRR. These distributions are weighted by size of generation, but the unweighted distributions look very similar. For each pension structure in each figure, the distribution of outcomes for 396,000 outcomes is plotted (1000 sample paths each with 396 years).

There are a number of interesting points. First, Figure 4 shows the distribution of discount rates ρ needed to make the corresponding distribution of NPVs in Figure 3 collapse to 0. Pulling down the right tail of the NPV distribution requires a larger increase in the discount rate than the decrease needed to pull up the left tail. This explains why the distribution in Figure 4 appears roughly symmetric while the distributions in Figure 3 tend to be right skewed.

Second, Figure 4 shows that the actual Swedish system has a strikingly tighter distribution of the IRR than any other program. Clearly it delivers a lower level of uncertainty, by this measure. However, this comes at the cost of a lower mean IRR because its asymmetric brake allows surplus funds to accumulate. Comparison to NDC(*g*) systems (those whose rate of return is based on the growth in the wage rate, *g*) with a symmetric brake (*A*=.5 or 1.0) shows that setting the rate of return based only on wage growth while ignoring demographic change is not sufficient to achieve this low level of uncertainty. The surpluses that accumulate with the asymmetric brake provide a buffer in bad times such that the brake seldom comes into play. Aside from the Swedish distribution, the others look quite similar, although the German and 50-50 US system have lower modes than the others.

Third, the systems that adjust taxes or share the adjustments between taxes and benefits, as happens in the US Tax Adjust or 50-50 systems and the German system, all have thicker left tails for the NPV distribution. Those that rely entirely on adjustment of benefits like all the NDC systems and the US Benefit Adjust have severely limited left tails. This must reflect the greater effect of discounting on benefits, which occur at a later age than tax payments. The Tax Adjust system, by contrast, has the thickest left tail because variations in taxes have the greatest effect on the NPV since they occur at a younger age. The German system and the 50-50 US system have very similar distributions of NPV.

Means and Variances

For reasons discussed earlier, we will present results both for generations weighted by their size at birth, in Table 1, and unweighted by size, in Table 2. We believe that the population-weighted measures are most relevant, since results are otherwise strongly influenced by whether the pension structure favors larger or smaller generations. The means and variances for the distributions plotted in Figures 1 and 2 are given in Table 1, and those of the unweighted distributions are given in Table 2. We will start by considering the results for the NPV.

The mean NPVs are in the range -.041 to -.035 as a fraction of lifetime earnings, as expected. Because the NPV is a nonlinear outcome measure which interacts with stochastic fluctuations and pension designs, its mean value can vary from pension system to pension system and these variations are not in themselves informative about welfare. The actual Swedish plan clearly has the lowest variance for the NPV both with and without population weighting, followed by the NDC(*g*) systems with symmetric brakes. All the NDC systems dominate the German and US plans with respect to variance. The Swedish system achieves this stability of NPV at the cost of having the lowest mean NPV for the population-weighted case, because its asymmetric brake lets assets accumulate on some sample paths.

The mean IRRs are in the range .010 to .013 with the variation due again to the specific nonlinear interactions of the plans, the disturbances, and this particular measure. In a deterministic steady state all would yield the same IRR, .011. However, the IRR for the Swedish system is penalized by the asymmetric brake and is slightly lower than the others. Once again the Swedish system clearly has the lowest variance with and without population weighting, but the other NDC systems have high variances and the NDC(*g*) with A=.5 has an extremely high variance.

Based on these descriptive statistics the Swedish system stands out for its low uncertainty but also for slightly lower mean performance. The expected utility measures to which we turn shortly take both into account in constructing a single outcome measure.

The Effect of Population Weighting on NPV and IRR Outcomes

Comparing Tables 1 and 2 we can calculate that with the US Tax Adjust system, population weighting raises both the NPV and the IRR by about 15 percent relative to their values with no weighting, consistent with the earlier discussion about population weighting.

With the US Benefit Adjust system, population weighting reduces both the NPV and the IRR by about 2%. The US 50-50 system lies in between, with smaller increases in the average NPV and IRR than the US Tax Adjust plan. The German system, which resembles the 50-50 plan in that it involves zero system debt and annual adjustments in both taxes and benefits, follows a similar pattern, although its somewhat larger differences between Tables 1 and 2 (with NPV and IRR rising by 13 and 10 percent, respectively, with population weighting) suggest that the German system lies somewhere in between the US 50-50 system and the US Tax Adjust system in terms of the relative burden of adjustment placed on taxes and benefits. For the Swedish plans, the corresponding changes are generally small and not systematic, as these plans allocate different shares of shock absorption to workers and retirees, depending on each system's detailed specification.

The previous discussion referred to the mean values, but one can also argue that any fluctuation in cost of benefits will require a smaller adjustment in tax rates under Tax Adjust if a taxpaying generation is large, so we might expect the variance of NPV or IRR to be smaller under population weighting for the Tax Adjust system, and indeed it is. Similarly, we might expect that the variance would be less with population weighting under Benefit Adjust because small generations experience greater variations in benefits, and indeed it is so for the NPV measure, but not for the IRR measure. For the German system, as for the US 50-50 system, population weighting reduces the variances of both measures.

Expected Utility

In looking at our NPV and IRR measures, we have considered both mean values and variances 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, as described earlier and developed in more detail in the Appendix. While this methodology is motivated by our interest in how the systems interact with uncertainty and risk aversion, we start by considering measures of expected utility under the assumption of risk neutrality. These estimates are useful in themselves, because they are more tightly related to measures of individual welfare than are the NPV and IRR measures already discussed. Thus, they form a useful basis for evaluating the systems' performance in the absence of risk aversion and provide a benchmark against which to compare the results once risk aversion is added.

Under risk neutrality the expected utility measure gives a clear answer when results are weighted by generation size: the Swedish system performs worst of the eight systems, although only slightly worse than the three versions of the US system. The highest expected utility is achieved by the German system which does slightly better than the NDC systems with symmetric brakes. Continuing with expected utility under population weighting, but now with risk aversion (γ=3), the three NDC systems with symmetric brakes are all superior to the German system, which comes in fourth, and the German system is considerably better than the Swedish system and all versions of the US system.

Our variants of the Swedish system include two NDC(g) versions with different severities of the symmetric brake: A=.5 and 1.0. One might expect that the weaker (A=.5) brake would involve more risk sharing, since response to imbalances is slower. This is the case in Table 1 for both values of γ, although the differences are quite small. In fact, comparison of other outcome measures in Tables 1 and 2 shows that performances of these two systems are generally very similar except that when A=.5 the variance of the IRR becomes large.

While there is no clear winner for Expected Utility with population weighting, since results vary by degree of risk aversion, we can say that all the NDC systems and the German system are better than the actual Swedish system or the various US systems. We also note that with risk aversion, the NDC(g) systems with symmetric brakes move to the fore.

These conclusions change, however, when we don't weight by the size of generation. In this case the NDC(n+g) with symmetric brake (A=.5) does better for all levels of risk aversion. We might expect this program to do well, because its rate of return reflects both productivity growth and labor force growth. It and the other NDC programs with symmetric brakes dominate all other pension forms including the German. Although we think the results in Table 1 are more relevant, the differences between the two tables for the German system suggest that that system's net benefits are more heavily directed at large cohorts than are those of the Swedish systems. This, too, makes sense, once one also compares the outcomes for the US systems in the two tables. The biggest difference in Expected Utility is for the US Tax Adjust System, followed by the German system and the US 50-50 system. As in the case of other criteria, the German system lies between these two US systems in its properties. The comparison of Tables 1 and 2 tells us that systems that rely more heavily on tax adjustments tend to help larger generations more and smaller generations less. One puzzle is why the German system performs so much better under the Expected Utility measures than, say, the US 50-50 system, given these two systems' superficial similarities.

Finally, Table 2 also provides some sensitivity analysis for Expected Utility with risk aversion, showing results for a higher coefficient of risk aversion (γ=5) and under the assumption that the marginal utility of consumption for retirees (see expression (A10') in the Appendix) is based on the wage rate at the cohort's age of retirement rather than at its entry into the labor force.[6] Neither of these variations in assumptions has an important impact on the relative attractiveness of the different systems.

Horizontal Equity

For reasons discussed earlier, we also believe it may be useful to evaluate the programs from the point of view of a social welfare function that weights variations in outcome more heavily when they affect nearby generations differently, since nearby generations seem likely to be the reference group for assessing the fairness of outcomes. This is the purpose of the Horizontal Equity measure. We have not developed a population weighted measure for Horizontal Equity because there are complexities that we have not yet resolved. Table 2 shows the Horizontal Equity results for the NPV and risk-neutral Expected Utility[7]. In the table, the gap between this 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: by both measures, the Swedish system dominates, but by more under the Expected Utility criterion. The NDC(g) systems with symmetric brakes do next best, although the US Benefit Adjust system does about as well under the EU criterion. The

[6] In the notation used in expression (A10'), the wage w_s rather than w_t is used.

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

performance of the Swedish systems based on the wage rate may provide some insight into why the actual Swedish system does not take demographic fluctuations directly into account.

Discussion and Conclusions

The NDC systems aim to pay a rate of return to contributors that is warranted by the macro economic/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 brake mechanism in their system design, if labor force growth should drop below 0 then the brake would eventually automatically reduce the rate of return below g. Our analysis shows that this program design insulates participating generations from some variations in the economic/demographic environment. The system yields a lower variance in NPV and IRR, for example, and it scores best on the Horizontal Equity measures, indicating that it achieves a local smoothness in generational outcomes. The asymmetric brake, which reduces the rate of return in some circumstances but never raises it, apparently plays a key role. To be sure, this arrangement permits the system to accumulate undistributed assets and therefore makes it yield a lower mean NPV and IRR compared to NDC(g) systems with a symmetric brake. But we see there is an unexpected advantage to this arrangement: by accumulating more assets, it avoids having to apply the brake and thereby leaves the rate of return more stable. It therefore has lower variance and better Horizontal Equity than the NDC(g) systems with symmetric brake. In comparisons of the variances, it clearly dominates all the NDC(g) programs with symmetric brakes and these have variances close to the NDC($n+g$) symmetric brake program. It appears, therefore, that the key feature reducing the variance of generational outcomes in the Swedish program is not the g versus $n+g$ feature, but rather the asymmetric brake.

This stability of the actual Swedish system is not worth the cost to participants, based on our Expected Utility calculations. If we now focus on the Expected Utility measure for the population weighted case, the Swedish model ranks only fifth out of the eight considered, with and without risk aversion. The best program with no risk aversion is the German. The best with risk aversion are the two NDC(*g*) programs with symmetric brakes.

In sum, the actual Swedish system provides the most stability and the greatest degree of Horizontal Equity. NPV and IRR measures favor one of the three symmetric Swedish systems, as does Expected Utility when there is risk aversion. The German system fares best (under population weighting) under the Expected Utility criterion without risk aversion.

It is notable that the three US Adjust systems are never close to best under any criterion. The distinctive feature of these three systems is that each period they adjust taxes or benefits to achieve perfect balance. Because this balancing is always done cross-sectionally, by proportionally adjusting all tax or benefit schedules, there is no attention to longitudinal outcomes for generations. It may be for this reason that these systems do relatively poorly when assessed on generational outcome measures. The NDC type systems, by contrast, are based on some concept of longitudinal actuarial fairness, and as a consequence the year to year balances are permitted to vary to some degree (Auerbach and Lee, 2007). Even the German system, which has annual adjustments to both taxes and benefits, fares better than the US 50-50 system that also does so, suggesting that the details of how the adjustments occur matter.

Ultimately, the decision as to what criterion matters most must be political. We believe, however, that simulations of the kinds we have undertaken can help to inform the design of unfunded public pension systems and clarify the issues at stake in their choice.

Appendix. Valuing Flows from Social Security

In theory, the correct way to evaluate social security is to specify the household's other sources of income and then solve for its optimal consumption 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 relationship of social security benefits and taxes to past economic and demographic factors. One alternative approach would be to assume some rule-of-thumb saving behavior based on past variables such as wealth and current income, to solve for the consumption path along each trajectory, and then to calculate expected utility by applying a utility function to each consumption trajectory and then aggregating across trajectories. But even this approach would involve substantial computations, requiring, for each social security system variant and each specification of the household's decision rule, the solution for lifetime consumption and saving for each of several hundred cohorts on 1000 separate trajectories. The methodology we propose instead introduces some relevant factors with the aim of getting some idea of the impact of risk aversion on our estimates.

Before specifying the 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:

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

We don't know these state-contingent marginal utilities without solving the full optimization, 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.) $x_{s,i}$:

$$\text{(A2)} \quad U'_{s,i} \sim h_s(x_{s,i}) = a_s h_s(x_{s,i}),$$

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 (A2) into (A1), we get:

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

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:

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

These solutions for a_s can be calculated under the assumption that each path has equal probability $1/N$, where N is the number of paths. From (A2) and (A4), we have the solutions for marginal utility,

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

Calculating the Value of Social Security Taxes and Benefits

The value of taxes and benefits equals the change in utility associated with these flows, normalized so as to be expressed in terms of date-t dollars, or

(A6) $$V_t = \sum_i \frac{1}{N}\sum_{s=t}^{t+T} P_{s,i}\left[U\left(c^*_{s,i} + \frac{F_{s,i}}{P_{s,i}}\right) - U\left(c^*_{s,i}\right)\right],$$

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. 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,

$$\text{(A7)} \quad \Delta V_t = \sum_i \frac{1}{N} \sum_{s=t}^{t+T} P_{s,i} \left[U\left(c^*_{s,i} + \frac{F_{s,i}}{P_{s,i}} \right) - U\left(c^*_{s,i} + \frac{\overline{F}_s}{P_{s,i}} \right) \right].$$

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 (A7), we get:

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

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

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

Without loss of generality, we can drop the term $\bar{f}_s$ when it first appears in (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 (A5) into (A9) yields:

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

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 A10) and through the impact of fluctuations on the flows themselves (the next term).

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

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

That is, under risk-neutrality, we should divide the average flow at each date by the average discount factor. [8] After summing over the N trajectories for each generation in (A11), we will divide by that generation's average initial wage, in order to remove the productivity growth trend to avoid giving more weight to later generations when we calculate an average over generations.

Parameterization

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

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

$$\text{(A11}'\text{)}\quad \Delta V_t = \sum_i \frac{1}{N}\sum_{s=t}^{t+T}\frac{1}{E_t\left[\prod_{u=t}^{s}(1+r_u)P_s\right]}F_{s,i}\ .$$

Risk-aversion parameter, γ

We consider two cases, $\gamma = 0$ (neutrality) and $\gamma = 3$.

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. 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, 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.

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

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

with everything except taxes, $T_{s,i}$, and benefits, $B_{s,i}$, the same across different social security scenarios. 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[9]. The resulting ratios are .103 for taxes relative to non-retiree consumption, and .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

[9] 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.

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 average initial wage, $\overline{w}_t$, in order to weight the results equally across generations.

References

Alho, Juha M., Svend E. Hougaard Jensen and Jukka Lassila eds. (in press) *Uncertain Demographics and Fiscal Sustainability* (Cambridge University Press).

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

Auerbach, Alan J., and Kevin Hassett (2002) "A New Measure of Horizontal Equity," *American Economic Review* (September) pp. 1116-1125.

Auerbach, Alan J. and Ronald D. Lee (2007) "Notional Defined Contribution Pension Systems in a Stochastic Context: Design and Stability" (January 2, 2007). *Berkeley Program in Law & Economics, Working Paper Series.* Paper 215. http://repositories.cdlib.org/blewp/art215

Börsch-Supan, Axel, and Christina B. Wilke, 2003, "The German Public Pension System: How it Was, How it Will Be"

Börsch-Supan, Axel, Annette Reil-Held and Christina Benita Wilke, 2003, "How to make a Defined Benefit System Sustainable: The 'Sustainability Factor' in the German Benefit Indexation Formula"

Gruber, Jonathan, and David A. Wise (1999) *Social Security Around the World*. University of Chicago Press: Chicago.

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

Lee, Ronald (1993) "Modeling and Forecasting the Time Series of US Fertility: Age Patterns, Range, and Ultimate Level," *International Journal of Forecasting* v.9 pp.187-202.

Lee, Ronald (2007) National Transfer Account Project, data for United States, 2003. http://www.ntaccounts.org.

Lee, Ronald and Lawrence Carter (1992) "Modeling and Forecasting U.S. Mortality," *Journal of the American Statistical Association* v.87 n.419 (September 1992), pp.659-671, and "Rejoinder," same issue, pp.674-675.

Lee, Ronald and Shripad Tuljapurkar (1994), "Stochastic Population Projections for the United States: Beyond High, Medium and Low," *Journal of the American Statistical Association*, v.89 n.428 (December) 1175-1189.

Lee, Ronald and Shripad Tuljapurkar (1998) "Stochastic Forecasts for Social Security," in David Wise, ed., *Frontiers in the Economics of Aging*. Chicago: University of Chicago Press, pp. 393-420.

Lee, Ronald D., Michael W. Anderson, and Shripad Tuljapurkar (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'" in Holtzmann R. and Palmer E. (Eds.) (2005) *Non-Financial Defined Contribution (NDC) Pension Schemes: Concept, Issues, Implementation, Prospects*. World Bank, Washington D.C.

Lee, Ronald, Timothy Miller and Michael Anderson (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 and Michael Reiter, 2006, "Sharing Demographic Risk – Who is Afraid of the Baby Bust?" paper presented at the NBER Woodstock conference, October.

Table 1
Summary Statistics for Eight Pension Plans Based on 1000 Stochastic Paths of 400 Years

(Weighted by Generation Size)

	US - Tax Adjust (Scaled)	US - Benefits Adjust	US - 50-50 Adjust (scaled)	NDC Sweden	NDC(*g*) Symmetric Brake A=.5	NDC(*g*) Symmetric Brake A=1	NDC(*n*+*g*) Symmetric Brake A=.5	German Model
NPV								
Mean	-0.03972	-0.03969	-0.03971	-0.04075	-0.03544	-0.03580	-0.03582	-0.03478
Variance	0.00246	0.00210	0.00216	0.00162	0.00180	0.00179	0.00182	0.00192
IRR								
Mean	0.01258	0.01097	0.01203	0.01044	0.01252	0.01268	0.01309	0.01208
Median	0.01283	0.01124	0.01215	0.01079	0.01352	0.01331	0.01316	0.01223
Variance	0.0000492	0.0000556	0.0000397	0.0000233	0.0003815	0.0000911	0.0000522	0.0000445
EU								
$\gamma = 0$	-23.38	-23.77	-23.52	-23.95	-21.82	-21.99	-22.01	-21.50
$\gamma = 3$	-56.67	-53.82	-55.36	-53.61	-50.38	-50.52	-50.83	-51.54

Table 2
Summary Statistics for Eight Pension Plans Based on 1000 Stochastic Paths of 400 Years
(Not Weighted by Generation Size)

	US - Tax Adjust (Scaled)	**US - Benefits Adjust**	**US - 50-50 Adjust (scaled)**	**NDC Sweden**	**NDC(*g*) Symmetric Brake A=.5**	**NDC(*g*) Symmetric Brake A=1**	**NDC(*n*+*g*) Symmetric Brake A=.5**	**German Model**
NPV								
Mean	-0.04701	-0.03890	-0.04416	-0.03965	-0.03597	-0.03593	-0.03476	-0.04015
Variance	0.00263	0.00226	0.00231	0.00170	0.00193	0.00193	0.00196	0.00203
IRR								
Mean	0.01078	0.01112	0.01091	0.00996	0.01062	0.01142	0.01332	0.01090
Median	0.01101	0.01140	0.01108	0.01092	0.01296	0.01295	0.01343	0.01105
Variance	0.0000517	0.0000552	0.0000417	0.0000245	0.0005911	0.0001683	0.0000530	0.0000454
EU								
$\gamma = 0$	-28.14	-25.16	-27.09	-25.15	-23.63	-23.64	-23.13	-25.67
$\gamma = 3$	-68.16	-59.55	-64.77	-59.25	-56.95	-56.77	-56.16	-61.32
$\gamma = 5$	-708.87	-654.53	-687.78	-657.22	-633.46	-632.34	-630.45	-673.71
$\gamma = 3$, **ret.** w	-529.07	-474.71	-508.78	-477.11	-456.84	-456.84	-454.06	-488.04
HE								
NPV	0.998639	0.998897	0.998825	0.999016	0.998908	0.998911	0.998882	0.998857
EU	0.999487	0.999902	0.999759	0.999952	0.999893	0.999902	0.999834	0.999809

Figure 1. Ratio of Retirees to Workers, 15 Sample Paths

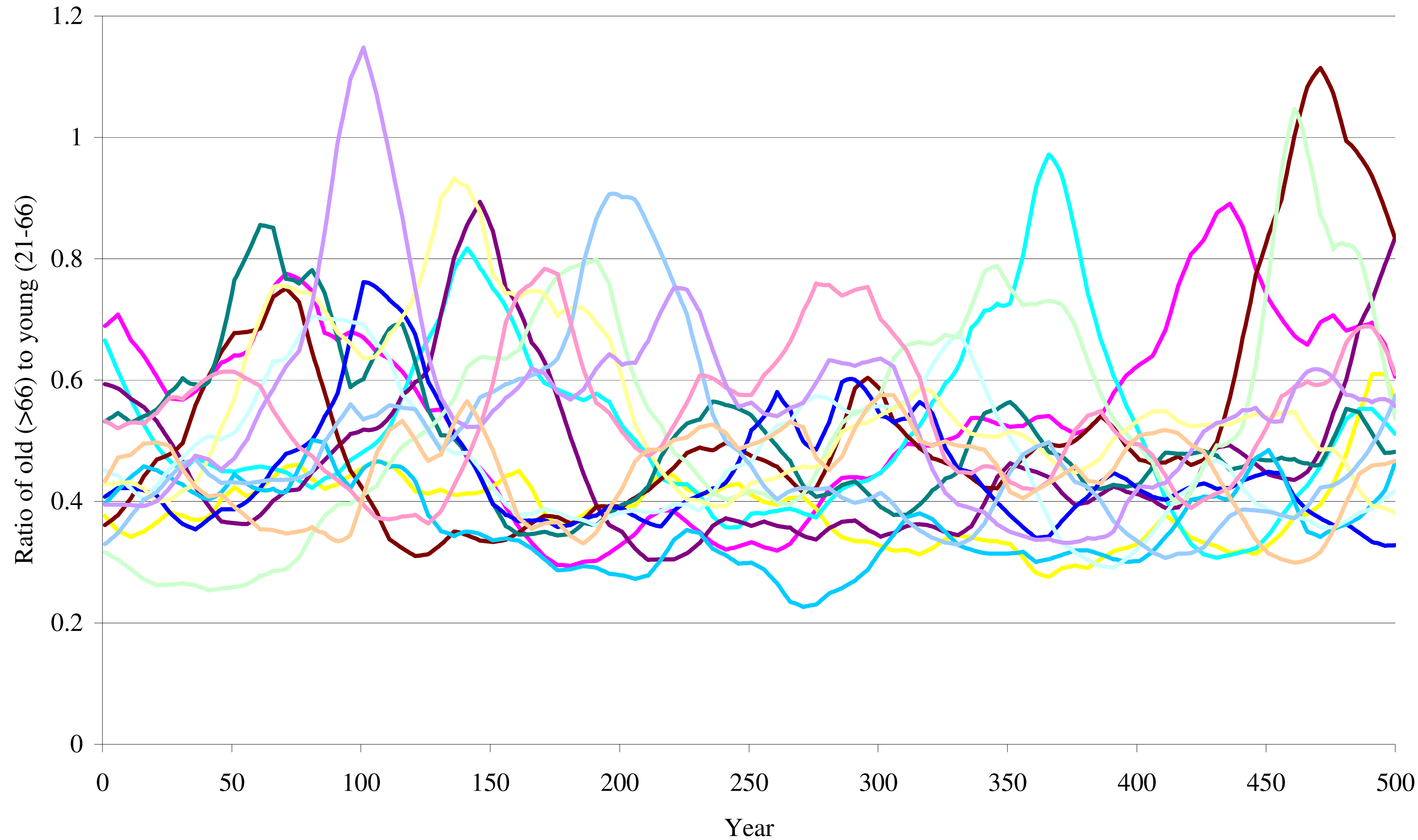

Figure 3. Frequency Distribution of 400,000 NPV Outcomes for Generations, for Eight Pension Systems (Population Weighted)

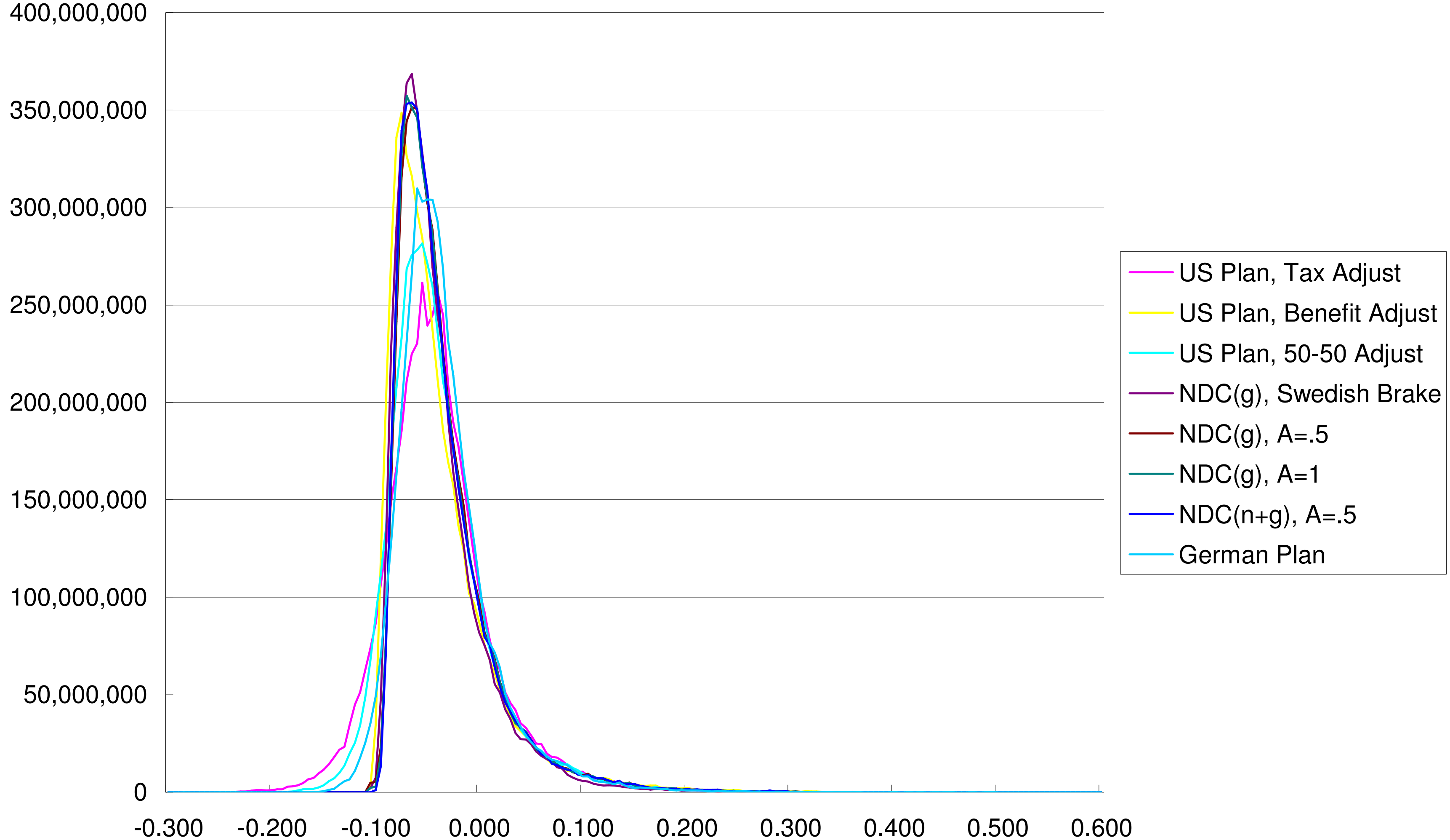

Figure 4. Frequency Distribution of 400,000 IRR Outcomes for Generations, for Eight Pension Systems (Population Weighted)