Volume Title: Frontiers in the Economics of Aging

Volume Author/Editor: David A. Wise, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-90304-4

Volume URL: http://www.nber.org/books/wise98-1

Publication Date: January 1998

Chapter Title: Stochastic Forecasts for Social Security

Chapter Author: Ronald Lee, Shripad Tuljapurkar

Chapter URL: http://www.nber.org/chapters/c7307

Chapter pages in book: (p. 393 - 428)

9 Stochastic Forecasts for Social Security

Ronald Lee and Shripad Tuljapurkar

Population aging is projected to have a major impact on the federal budget in the next century, in part through its effects on health costs through Medicare and Medicaid, and in part through its effect on the retirement system. Despite the inevitability of the aging of the baby boom, and its dramatic effect on the old age dependency ratio, a great deal of uncertainty remains about the extent of future population aging. On the one hand, we do not know how rapidly mortality will decline and how long people will be living, and on the other hand, we do not know what fertility will be, and therefore we do not know how large the labor force will be in the future. Immigration adds another layer of uncertainty, but we do not consider it in this paper. In addition to these demographic sources of uncertainty, there are economic variables with important effects on the future finances of the social security system, notably the rate of growth of productivity or real wages and the level of the real interest rate. Rational planning for the next century must somehow take into account not just our best guesses about the future but also our best assessments of the uncertainty surrounding these guesses.

Scenario-based forecasts are widely used to express the uncertainty of long-term forecasts. In these, the forecaster chooses, for each variable, a medium or best-guess trajectory, as well as high and low trajectories. Then one of these trajectories for each variable is grouped with others in a scenario, or collection

Ronald Lee is professor of demography and economics at the University of California, Berkeley. Shripad Tuljapurkar is president of Mountain View Research.

The authors are grateful to Bryan Lincoln, Timothy Miller, and Carl Boe for their research contributions to this project, and to Alan Auerbach and David Wise for helpful discussions at the planning stage. Participants in the Boulders NBER meeting also made helpful comments. Lee's research for this paper was funded by National Institute on Aging (NIA) grant AG11761. Tuljapurkar's research for this paper was funded by NICHD grant HD32124. The authors also acknowledge support by Berkeley's NIA-funded Center for the Economics and Demography of Aging.

of trajectories. These scenarios may, in turn, be described as "high," "medium," or "low," or by other terms such as "optimistic" and "pessimistic." A high scenario would typically be based on a high trajectory for fertility and net immigration and a low trajectory for mortality; this combination would yield high population growth. Alternatively, a "low cost" scenario for social security would bundle together high fertility and net immigration with high, rather than low, mortality; this scenario would generate the lowest old age dependency ratio. For the social security forecast, this low-cost demographic scenario would then be combined with high trajectories for productivity growth and (perhaps) for interest rates.

The scenario-based approach has several features worth noting. First, no probability is assigned to any of the trajectories or to the range that they cover, or to a scenario. Second, the trajectories are always high, or always low, or always middle; this means that fluctuations and structural shifts are assumed away. Third, the combinations of trajectories in scenarios are fixed to produce extreme outcomes. If there is some probability that fertility will follow the high trajectory, that must be higher than the probability that mortality will simultaneously follow its low trajectory; and that joint probability must be greater than the probability that productivity growth will simultaneously follow its high trajectory. Thus, even if we could attach some rough probability to the individual trajectories, we would have no clear way to attach probabilities to their combination. Fourth, any effort to attach probabilities to trajectories or to scenario outcomes would have to ignore many internal contradictions. For example, uncertainty about fertility is compounded by uncertainty about numbers of reproductive age women, but much of this uncertainty may cancel when it is transformed into uncertainty about numbers of annual births. Some uncertainty about annual births will cancel when it is summed into uncertainty about age group sizes or total population sizes. When demographic uncertainty is added to uncertainty from other sources, further cancellation should take place. Uncertainty is not additive in the way that it is assumed to be in scenario-based forecasts. These problems with the usual approach are very serious and have potentially serious consequences for planning and policy formation.

In this paper, we build on earlier work (Lee and Carter 1992; Lee 1993; Lee and Tuljapurkar 1994; Tuljapurkar and Lee, in press) to develop stochastic forecasts of the social security trust fund (assuming currently legislated changes in taxes and benefits), of summary actuarial balances, and of the "pay as you go" (PAYGO) payroll tax rate. These forecasts are based on stochastic models for fertility, mortality, productivity growth, and interest rates, which are fitted on historical data. For interest rates and productivity growth, the fitted models are constrained to conform in long-run expected value to the middle assumptions of the Social Security Administration (SSA; see Board of Trustees 1995). We have not attempted to build a simulation model for social security that is realistic in detail, which would be a major task. Instead we have sought to build a simple model that captures the core features of social security, and that can reproduce the social security forecasts when run with SSA specifications.

The Lee-Tuljapurkar (1994) stochastic population forecasts were used by the Congressional Budget Office (CBO) in its 1996 and 1997 reports to generate stochastic forecasts of the federal budget over the long term (CBO 1996). For an alternative approach to generating stochastic forecasts of social security finances, see Holmer (1995a, 1995b).

9.1 Forecasts of Fertility and Mortality

The evolution of the population depends on fertility, mortality, and net migration. In this analysis, we take fertility and mortality to be stochastic but have taken net migration to be fixed at the future levels assumed by the SSA. According to the sensitivity analysis done by the SSA (see Board of Trustees 1995, 32–42), in 2070 the effect on projected actuarial balance of the range of migration assumptions is only one-third as great as the effect of the fertility range, and only one-fifth as great as the range of the mortality assumptions.

Because our methods for forecasting fertility and mortality have been described in detail elsewhere (Lee and Carter 1992; Lee 1993), we only sketch them here. Both fertility and mortality are fitted by a model of the form

$$m_{x,t} = a_x + k_t b_x + \varepsilon_{x,t},$$

where $m_{x,t}$ is fertility of women aged x in year t, or the logarithm of the central death rate for age x in year t; a_x is an additive age-specific constant, reflecting the general shape of the age schedule; k_t is a period-specific index of the general level of fertility or mortality; b_x indicates the responsiveness of fertility or mortality at age x to variations in the general level; and $\varepsilon_{x,t}$ is the error in this approximation to the actual age schedule. There are some important issues in the estimation of this model, but they will not be discussed here. The key for forecasting is the period index, k. This is modeled as a stochastic time-series process, and the model is fitted to historical data. In the stochastic demographic forecasts, the stochastic model is used either to characterize the probability distributions for future fertility and mortality in analytic work or to generate sample paths for fertility and mortality through stochastic simulation. In either case, the coefficients a_x and b_x are then used to generate the age schedules for each t. In the case of fertility, it is necessary to constrain the model based on outside information; otherwise, the long-run behavior of the forecast is unsatisfactory, for reasons discussed later (see Lee 1993). The simplest and most satisfactory approach has been to constrain the long-term expected value of the fertility process to equal a specified level, such as 2.1 for the total fertility rate (TFR). Then the model conveys information about the variance and autocovariance of fertility around this long-term mean.

9.2 Forecasts of Productivity Growth and Interest Rates

The finances of the social security system depend not only on demographic variables but also on a whole range of economic and health variables: productivity growth rates, interest rates, inflation, disability uptake rates, and disability departure rates, to name only the ones for which the SSA performs a sensitivity analysis. From this list, we have chosen to focus on productivity growth rates and real interest rates. This choice is based partly on the sensitivity analysis reported in Board of Trustees (1995) and partly on the conceptual centrality of these variables for the retirement system.

Our analysis treats fertility, mortality, productivity growth, and the interest rate as stochastic. In addition to these four factors, the SSA performs sensitivity tests on assumptions about net migration, the rate of inflation, disability incidence, and disability termination. Analysis of the SSA projections and sensitivity tests (Board of Trustees 1995) indicates that the four variables we treat as stochastic account for 63, 70, and 76 percent of the width of the SSA low-cost–high-cost range in 2020, 2045, and 2070, respectively.[1]

In a pure PAYGO system, the interest rate would be irrelevant. However, the U.S. system has become partially funded, and interest on its growing reserve fund is an important inflow to the system. Also, in forecasts that imply negative reserve fund balances, it is appropriate to incorporate the cost of borrowing to cover the deficit.

Retirees receive benefits based on their earnings histories. When productivity growth is rapid, these histories are on average much lower than are the wages of the current workers. For this reason, rapid productivity growth reduces the tax rate necessary to fund the benefits of current retirees. The average age of receiving benefits is estimated to be 71 years old around 1990, whereas the level of benefits of a cohort is set by the level of the real wage when the cohort turns 60 years old. This 11-year age difference leverages the effects of changes in the rate of productivity growth.

Productivity growth and interest rates are economic variables that could be forecast in some highly structured way, based on the marginal products of labor and capital in an aggregate production function, with labor supply driven by the demographic forecasts, and with capital accumulation depending in part on savings behavior that might in turn be driven by our forecasts of demographic change, drawing on either life cycle saving theory or estimated age profiles of saving and dissaving. Such an approach could certainly be implemented. However, demographic change has shown limited power to explain saving rate variations in the past (see Aaron, Bosworth, and Burtless 1989); domestic saving does not explain all variations in capital accumulation, due in part to international capital flows, nor does the capital-labor ratio explain all

1. These are calculated from ranges in the sensitivity tests reported in Board of Trustees (1995, 132–42) and are for the summary actuarial balance. The total range when all vary is very close to the sum of the individual ranges when one varies and all others are held at their middle values.

variation in real interest rates; the capital-labor ratio is only one influence on labor productivity; technological progress is a great unknown; and on the whole, we have little confidence that such a structured approach would actually be very useful. Instead, we will simply model productivity growth and interest rates as single time series, without attempting to relate them to the economic structure believed to generate them. Provided that the trust fund does not become too large in absolute value relative to GNP, our no-feedback assumption may not be too bad. In principle we could model and forecast them as jointly evolving (e.g., VAR or cointegrated processes), but our empirical analysis turned up surprisingly little evidence that they were associated.

Productivity was measured as output per worker-hour, from the Bureau of Labor Statistics. This measure has a longer available time series than the alternatives, although there are sometimes fairly substantial differences between its rate of change and that of the social security covered wage series. In modeling productivity growth, we take into account the effect of changing age composition of the labor force on aggregate wages or productivity. The cross-sectional age profile of wages has a characteristic inverted U shape, peaking in the late 40s. This means that when there are disproportionately many young workers, for example, as in the 1970s and 1980s, there is downward pressure on aggregate wage levels or productivity per worker, arising purely from the compositional effect. We purged our productivity growth measure of age composition effects by dividing it by the summed age-by-age products of the average age profile of earnings times the year-specific proportional age-sex composition of the labor force.[2] Figure 9.1 shows the real productivity growth rate series from 1948 to 1994 *(line)* and the growth rate series with adjustment for demographic composition effects *(diamonds)*. Demographic adjustment typically alters the growth rate by less than ±.5 percent. In addition, there may be a twist in the age distribution of earnings due to the well-known cohort size effect. However, we view this as a second-order effect and do not try to incorporate it.

Statistical time-series methods are not intended for long-term forecasts. Their purpose is instead to fit a highly simplified and parsimonious linear model with very short memory, which will mimic the behavior of the true generating process for a few steps into the future. Standard diagnostic and modeling procedures typically indicate that one should at least first-difference the time series before fitting it as an ARMA process. The first-differencing means that the original process is modeled as some sort of random walk, with a

2. The average wage profile by age and sex was calculated from the 1990 March Current Population Survey (CPS) supplement. The data on age-sex composition of the labor force are also taken from the CPS for 1948–94. If $N(i,j,t)$ is the number in the labor force for age category i, sex j, and year t, and the average wage from the 1990 CPS is $W(i,j)$, then the index $I(t)$ is

$$I(t) = \sum_i \sum_j W(i,j)N(i,j,t) / \sum_i \sum_j N(i,j,t).$$

The productivity series is then divided by $I(t)$ to remove the effects of changing age-sex composition of the labor force.

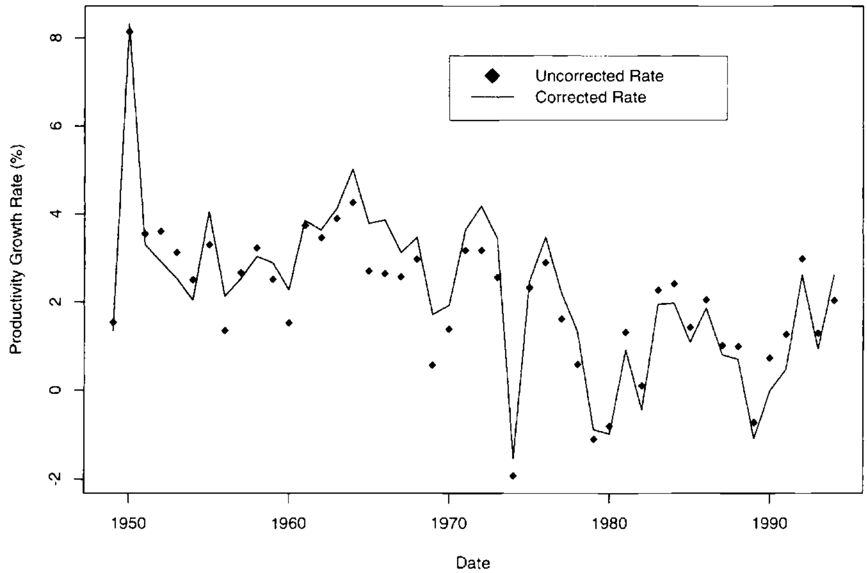

Fig. 9.1 Rate of productivity growth, with and without demographic adjustment, 1948–94

Note: Productivity is measured as output per worker-hour. Demographic adjustment removes the influence of the changing age-sex composition of the labor force, using the average age-sex earnings schedule over the period.

complicated innovation term and with drift. This often works well for short-run forecasts. For long-run forecasts, however, the drift (arising from any constant term in the differenced process) may lead to absurd levels at the same time that the random walk (or integrated) nature of the model leads to huge variances. For these reasons, long-term forecasts from these models can often be rejected as absurd. For example, in one of our fitted models, the forecast of the productivity growth rate for the year 2070 was 20 percent per year, plus or minus 60 percent.

The standard diagnostic and modeling procedures go quickly to first-differencing because it is simple, not because the data indicate that the process is truly a random walk. An alternative autoregressive model may fit just as well yet behave entirely differently in the long run, with no long-term trend and much smaller variance, relative to the integrated process. There is often too little information in the historical data to indicate which model is more correct. It may be necessary to turn to outside information, which is the route we take here.[3]

For some of our series, there is an additional problem: the consensus view

3. The level of productivity, of course, has a very strong trend and must be differenced (after taking the logarithm) to obtain the rate of productivity growth. The discussion in this section addresses the question of whether the rate of productivity growth should then be differenced once more, before modeling.

is that there have been structural changes in recent years such that the means of the fertility, productivity, and interest rate series will be different from the average of their past values. We initially attempted to deal with these purported structural changes formally, by fitting state space models with random trends (Harvey 1989). These efforts did not lead to plausible forecasts. After much experimentation, we found that apparently satisfactory models and forecasts were obtained by prespecifying the long-term means of the series, rather than estimating them from the data. In particular, we chose the long-term means of the real productivity growth and real interest rate series to equal the middle values assumed by the SSA: 1 percent per year and 2.3 percent, respectively. For better or for worse, this guarantees that the expected values of the forecasts will converge to the prespecified values, insuring consistency with the SSA middle scenarios. Since we wish to focus on the uncertainty rather than on the mean values, this feature is acceptable. Furthermore, with constrained means and an autoregressive (rather than differenced) specification, the probability bands for the forecasts of these rates appear very plausible, as we shall see below.

The trust fund for social security is held in special government securities, described as follows: "By law, the securities issued to Social Security for investment of the reserve fund bear an interest rate equal to the average yield, as of the last day of the prior month, on all outstanding Federal securities that are not due to mature for at least four years" (Foster 1994, 21). We constructed a historical interest rate series using the interest rate series for this special issue from 1961 to the present. Before 1961, social security interest rates were governed by different policies. Ideally, we would have constructed an equivalent measure for the period before 1961, since it would be most informative for forecasting purposes, whether or not it was actually the rate earned during the earlier period. In practice, having experimented with weighted averages of the three-month Treasury bill rate and the 10-year rate for best fit after 1961, we settled on using a within-sample regression fitted to predict the pre-1961 special issue rate from the simple three-month rate (bank discount basis). A Chow test accepted the null hypothesis that models fitted separately to the pre- and post-1961 data were the same. To convert these nominal interest rates into real rates, we used the Consumer Price Index for Urban Consumers (CPI-U).

As discussed earlier, we experimented with standard ARIMA models, random trend (structural time series) models, and ARMA constrained mean models (ARMA-CM), and chose the ARMA-CM models for the preferred estimates. To fit these models, one simply subtracts the constrained mean from the variable to be modeled and then fits an ARMA model in the usual way, but with no constant term.

The fitted model for productivity growth rate (PGR) is as follows:

$$PGR_t - .01 = .607(PGR_{t-1} - .01) + \varepsilon_t,$$

$$\sigma_\varepsilon = .021, \qquad R^2 = .50.$$

The fitted model for the real interest rate (I) is as follows:

$$I_t - .023 = .735(I_{t-1} - .023) + \nu_t,$$

$$\sigma_\nu = .018, \qquad R^2 = .17.$$

A standard ARIMA model leads to a long-term productivity growth rate forecast of .0210 for the raw series and .0225 for the series that was adjusted for age-sex structure, with 95 percent probability intervals $\pm.065$ and $\pm.072$, respectively. Our Kalman filter estimates of a random trend model yielded mean growth rates of about .016, with an interval of $\pm.16$. In both cases, the conventional mean forecast remains substantially higher than the social security mean forecast, and higher than many analysts would view as plausible. Additionally, the interval width for the Kalman filter estimate is greater than we find reasonable. The social security forecasts, as well as many analysts, assume that the productivity growth rate has undergone a structural decline to 1 percent per year. We therefore took this as the mean for our ARMA-CM models. These have mean productivity growth rates of .01, by assumption, with empirical interval ranges of .09 and .09, respectively.[4] Note that estimated errors are greater when there is no constant in the fitted model, leading to wider intervals than in the conventional models. Figure 9.2 shows the historical series of corrected growth rates together with the model forecast and 95 percent probability bands.

The SSA uses a productivity growth rate range of .005 to .015 per year, for an interval width of .01, only one-ninth as wide a range as in our preferred model, the ARMA-CM. But it is very important to realize that these intervals are not comparable. The SSA interval is intended to bracket some notion of the long-term tendency in the rate, while ours is intended to bracket year-to-year fluctuations. Over the time period 1948–94, the rate of real productivity growth had a high of more than 8 percent per year and a low of almost negative 2 percent, for a range of about 10 percent. This compares quite closely to the interval width of 9 percent estimated by the models[5] and obviously far exceeds the range assumed in the SSA forecast. For purposes of comparison, we also calculated the 95 percent probability interval for the average value of the productivity growth rate up to 2070. Instead of having a width of .09, it has a width of .0244, only about a quarter as great. It is narrower because a great deal of the year-to-year variation cancels out in the average. Note that the greater the positive autocorrelation, the less cancellation there will be, and the more similar will be the averaged and unaveraged brackets.

The SSA forecasts a real interest rate of .023, with a range of .015 to .030. Our ARIMA forecasts (actually a simple 1,0,0 model was chosen) have a long-

4. I.e., the difference between the upper and lower 95 percent bounds is .09.

5. We would expect to encounter a value at least as great as the 95 percent upper bound about once in a realization as long as 1948–94, and the same for the lower bound: each should be attained by about 2.5 percent, or one-fortieth, of the observations.

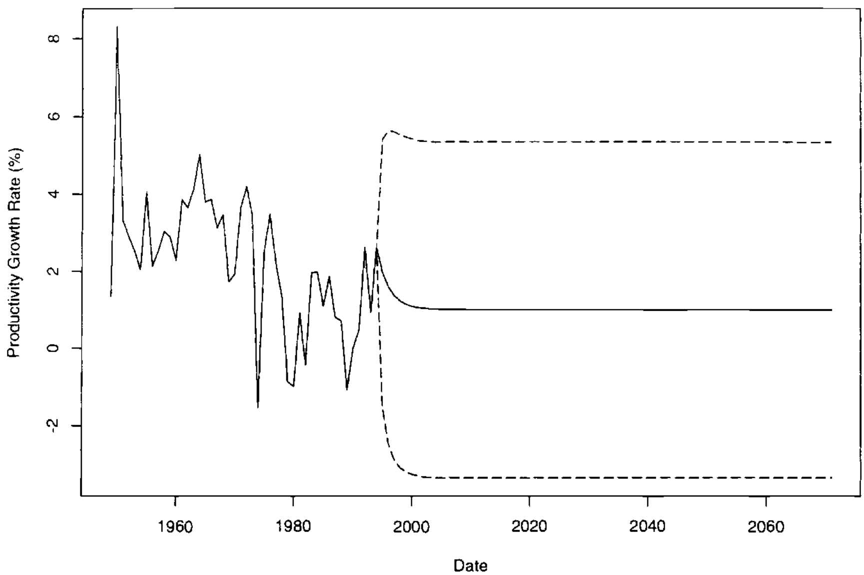

Fig. 9.2 Demographically adjusted productivity growth rates, actual and forecasted, with 95 percent probability interval, 1948–2070

Note: The long-term rate of productivity growth has been constrained to be 1 percent annually. See text for details.

run mean of .022, close to that of the SSA, with an interval width of about .12. The ARMA-CM estimates, with a constrained mean of .023 to match the SSA, also have an interval width of .12. These interval widths are very large, but the maximum-minimum range in the data series from 1948 to 1994 is even wider at 15 percent. As before, it is useful to calculate the interval width for the cumulative average values in 2070, and these are far narrower, at .045, or about a third the width. Figure 9.3 shows the historical series, the forecast, and the 95 percent probability interval for the special issue real interest rate.

The upshot, then, is that our preferred forecasts of productivity growth rates and interest rates have the same means as the SSA forecasts by assumption and are stochastic with 95 percent probability intervals for the cumulative averages that are about two and a half or three times as wide as the SSA scenario ranges. Like the SSA, we find forecasts of the real interest rate to be considerably more uncertain than those of the productivity growth rate.

9.3 The Tax and Benefit Schedules

The age schedules of payroll tax payments and receipt of benefits, when multiplied times the forecasted population age distribution and summed, determine the main flows into and out of the system. It is important to give careful consideration to the shapes and levels of these schedules.

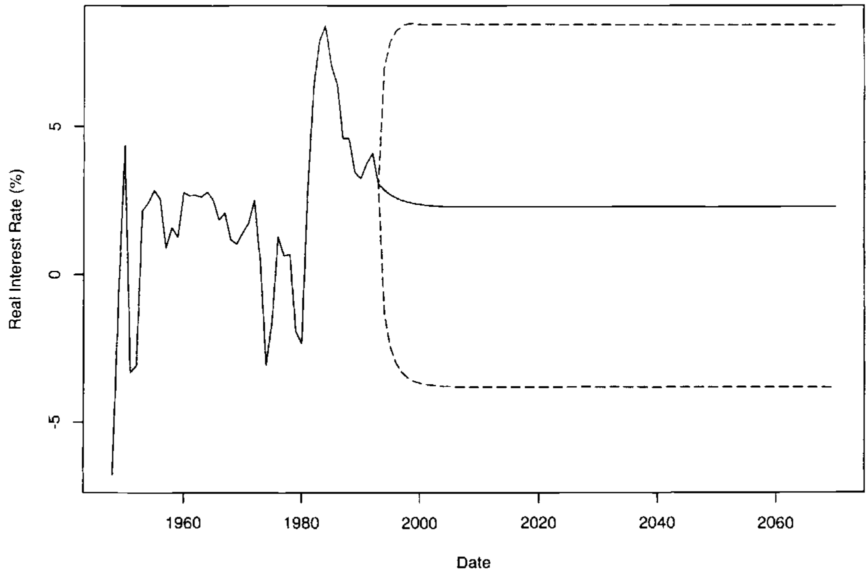

Fig. 9.3 Real interest rate, actual and forecasted, with 95 percent probability interval, 1948–2070

Note: Interest rate is for the special Treasury issue for social security from 1961 to 1994, with estimated values before 1961. It is deflated by CPI-U. The long-term real interest rate is constrained to be 2.3 percent. See text for details.

Our starting point will be the average age schedules of payroll taxes and benefits for 1994. For benefits, these are calculated from data in the *Annual Statistical Supplement* to the *Social Security Bulletin,* and for taxes, from the March CPS.[6] Figure 9.4 plots the age profiles of tax payments per capita for males and females. Figure 9.5 plots the corresponding data for receipt of benefits per capita. In figure 9.5, note the survivor's benefits received at ages under 20, and the survivor's and disability benefits received in growing amounts in the decades before the early retirement age of 62 is reached.

We will take into account the way in which productivity growth will affect the level and shape of the tax and benefit schedules. We will also take into

6. Data on the OASDI payments made by the SSA during 1994 are taken from the *Annual Statistical Supplement* to the *Social Security Bulletin* (SSA 1994). Estimates of OASI and DI payments by single year of age and by sex are generated from these data by assuming a constant per capita benefit level within the broad age categories provided in the published tables and allocating based on Census Bureau estimates of the 1994 U.S. resident population. These sex-age profiles of OASI and DI benefits are then adjusted by a constant factor so that the population-weighted sum equals the total benefit paid by OASI and DI as reported in the *Annual Statistical Supplement.* Data on the amount of OASDI taxes paid are derived from estimates of OASDI-taxable income taken from the March 1994 CPS. The sex-age profile of OASDI taxes paid is then adjusted by a constant factor so that the population-weighted sum equals the total taxes received by the OASI and DI trust funds as reported in the SSA's *Annual Statistical Supplement.*

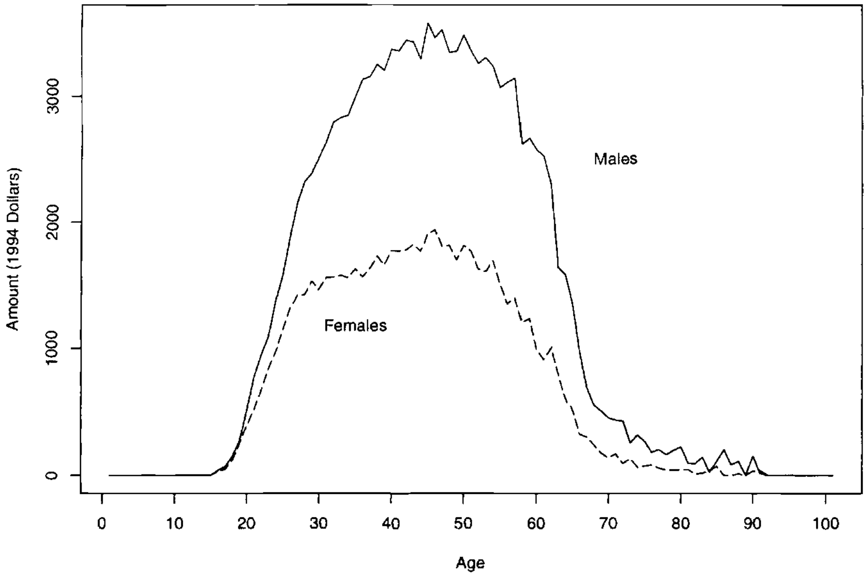

Fig. 9.4 Per capita social security tax payments by age and sex, 1994

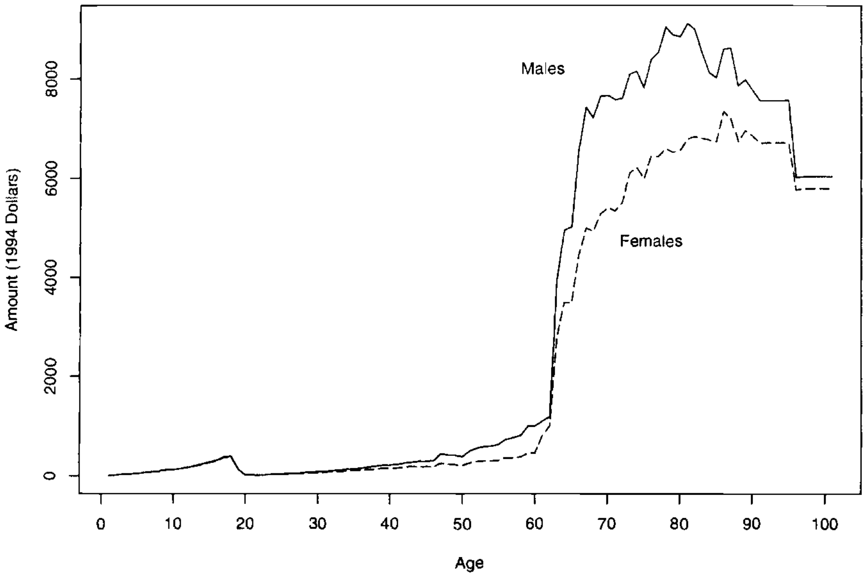

Fig. 9.5 Per capita social security benefits received by age and sex, 1994

account the effect of planned changes in the normal retirement age on these schedules. As mortality declines, there will be fewer survivors to claim benefits, and compositional effects of widowhood and selectivity (to be discussed later) will be reduced. We have not attempted to incorporate these effects of mortality decline. There will doubtless be other changes in the level and shape of these schedules that will occur over the next 75 years arising from changes in labor supply behavior due to causes other than the changes in social security regulations, and perhaps also arising from changes in the age structure of wages due to the changing age distribution of the labor force (young workers becoming less plentiful relative to older workers). There may also be further, currently unscheduled, changes in the social security regulations that will affect taxes and benefits. We have attempted neither to anticipate any of these changes nor to include measures of the uncertainty to which they give rise.

9.3.1 The Effect of Postponing the Normal Retirement Age

First consider the effects of the planned changes in the normal retirement age. The early age at retirement will remain at 62, but the age of retirement with full benefits will be raised gradually from age 65 to age 67. Since benefits will be reduced in an actuarially fair way for those retiring before age 67, benefits at age 62 will be lower than they are now, other things equal. In addition, it is planned to make actuarially fair adjustments to benefits for those who continue working past the normal retirement age, which should lead to some additional postponement of retirement.

A number of econometric studies attempt to assess the effects on retirement behavior of these planned changes (e.g., Lumsdaine and Wise 1994; and studies reviewed in Hurd 1990). Most analysts conclude that the effects of the planned changes will probably be relatively minor, but there is no consensus on details.

We have taken a mechanical approach to adjusting the age schedules. Our procedure is illustrated by the adjustment to benefits for the full two-year shift in the normal retirement age. We assume that at age 61, there will be no change at all in the tax or benefit schedules. At ages 70 and over, as discussed by the SSA, all beneficiaries will have started taking benefits. The levels of taxes and benefits that now obtain at ages 62–70 are assumed to shift in a smooth fashion, toward a profile in which the benefits at a given age before the shift are obtained two years later after the shift. At intermediate points, the amount of the age shift is interpolated proportionately.[7]

7. More generally, let the size of the shift in the normal retirement age be Y years (Y is 2 in the above example). Then the final benefit or tax schedule at age x will interpolate between its starting value at that age and its starting value at age $x + Y$, with an interpolation factor that depends on $(x - 62)/(70 - 62)$ for $62 \leq x \leq 70$. We assume that the new benefit profile phases out the old one evenly over the time period specified under law. The tax profile is adjusted in a similar way to incorporate additional wage-earning years before the new retirement age.

9.3.2 The Effect of Productivity Growth

We assume that the tax cutoff will be raised in proportion to the general level of wages, so we ignore it except to the extent that it is already reflected in the level and shape of the tax payment age profile. Thus we assume that the age schedule of taxes shifts upward proportionately at the same rate as productivity grows. These rising age schedules of payroll taxes are applied to projected population age distributions, thereby capturing the influence of the changing age distribution of the labor force on tax revenues in addition to the influence of productivity growth itself. The procedure restores the age composition effects that were initially purged from our measure of productivity.[8]

The effect of productivity growth on the benefit schedule is more complicated to model. The base-year age profile of benefits reflects many factors:

- Timing of retirement, and the way that the benefit level changes as a consequence
- Effect of widowhood and other factors on the proportion of beneficiaries that are single, since the per capita benefit is lower for married couples with only one primary beneficiary
- Effects of productivity change in the past, since more rapid productivity growth reduces the benefit level of older beneficiaries relative to younger ones
- Selectivity of mortality at higher ages, as those with lower lifetime incomes tend to have shorter lives and lower benefits
- Overgenerous adjustment for inflation led to rapidly rising benefits for some cohorts, and disappointment for the notch generations

We wish to retain the influences of all these factors except the last. In principle we should purge the profile of the effects of the notch generations, since these are relevant only for certain cohorts that will die out before many years have passed. In practice we have not yet attempted to do so.

Our approach is to modify the base-year benefit schedule in such a way that it is constantly updated to reflect the changing effects of past and future productivity change. For a person retiring at any age who turned 60 in year t, benefits depend on the average of his or her 35 highest annual earnings totals, with earnings before age 60 in year $t - s$ multiplied by the factor $w(t)/w(t - s)$, where $w(t)$ refers to economy-wide average wage levels. Because of the averaging and adjusting of wages, it happens that the primary benefit amount will depend almost entirely on the average wage level when the cohort

8. By first purging the measure of age effects, then modeling and forecasting it, and then reintroducing the effects of age composition change on tax revenues, we achieve two goals. First, we derive a measure of productivity growth that should be more amenable to time-series modeling, since once source of long swings has been removed. Second, we obtain a measure that can appropriately be multiplied times the tax age schedule, which would not otherwise be true.

turned age 60, and not on average wage levels in earlier years. To be more exact, for someone who retires at age 62, and whose highest earnings were in the last 35 of those years, the primary benefit will reflect the sum of (a) wages earned in the 33 years through age 59, each year's wages being inflated to the wage levels when the person is 60, and (b) the sum of wages at ages 61 and 62.

We have developed an algorithm to adjust the shape of the benefit schedule to reflect changes in the growth rate of productivity or wages while retaining the other important aspects of the shape of the benefit profile, as catalogued earlier. Let $B(x, t)$ be the benefit level for a person aged $x \geq 62$ in year $t \geq 1995$. We will define the benefit level for $B(x, t)$ recursively in terms of the benefit level at the same age in the previous year, $B(x, t - 1)$. These two benefit levels will differ because those aged x in t retired one year later than those aged x in $t - 1$, and therefore their benefits will be higher by a factor of $w(t - (x - 60))/w(t - 1 - (x - 60))$, or by the rate of productivity growth (or wage growth) in the year just before they turned age 60. Therefore

$$B(x,t + 1) = \{w(t - (x - 62) - 1)/w(t - 1 - (x - 62) - 1)\}\ B(x,t).$$

This simple equation ignores a term that depends on productivity change between ages 60 and 61, but this has very little effect on the results.

Note that when the rate of productivity growth continues at a constant rate that was also the rate of change in the past, then this algorithm simply raises the whole age profile by the same multiplicative factor each year. However, even if the projected future rate is constant, yet differs from the past constant rate, this procedure will change the shape of the benefit profile as time passes, as it should. Generally, more rapid productivity growth will tilt the age profile downward to the right, so that older people get relatively lower benefits than younger retirees. Because in fact productivity growth has recently been slower than in the more distant past, and because it is expected to be slower in the future as well, the age profile of benefits will be made to tilt upward to the right, so that older beneficiaries receive relatively higher benefits.

Figure 9.6 shows the projected age profile of taxes for the years 2005 and 2045, at which time the normal retirement age will have been raised by one full year and by two full years, respectively. It, and figure 9.7, are plotted on the assumption that productivity growth after 1994 is fixed deterministically at 1 percent per year. Figure 9.6 therefore reflects the effects of both productivity growth and the increase in the normal retirement age. Figure 9.7, which shows benefits received by age, is also affected by these factors. Because past productivity growth was generally more rapid than 1 percent, the assumed slowdown of productivity growth tilts the profiles upward toward the right. But the striking increase of benefits with age that is shown in figure 9.7 also reflects compositional change. The average real benefit received per surviving member of a retirement cohort actually rises with age and time following retirement, because survival is selective of those who had higher earnings while working,

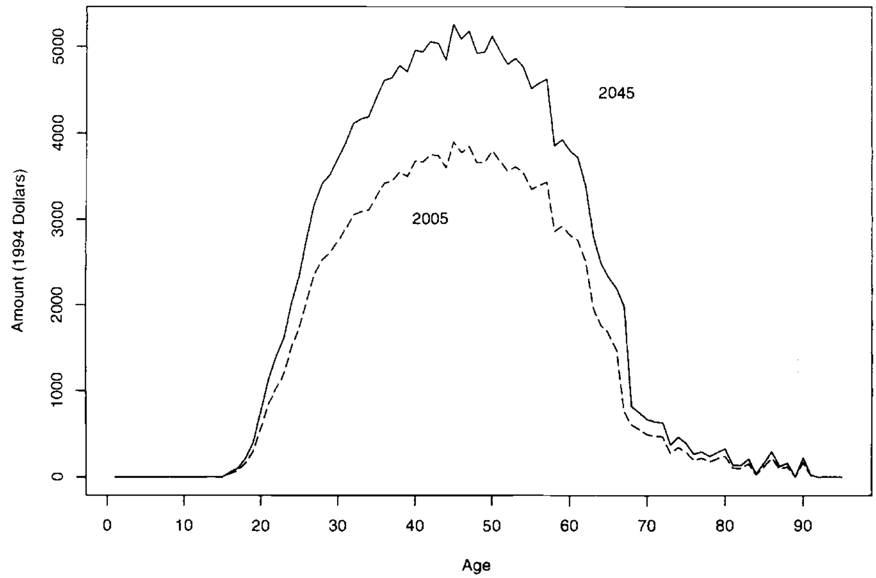

Fig. 9.6 Projected social security taxes for males (in 2005 and 2045), reflecting 1 percent productivity growth rate and legislated increases in retirement age

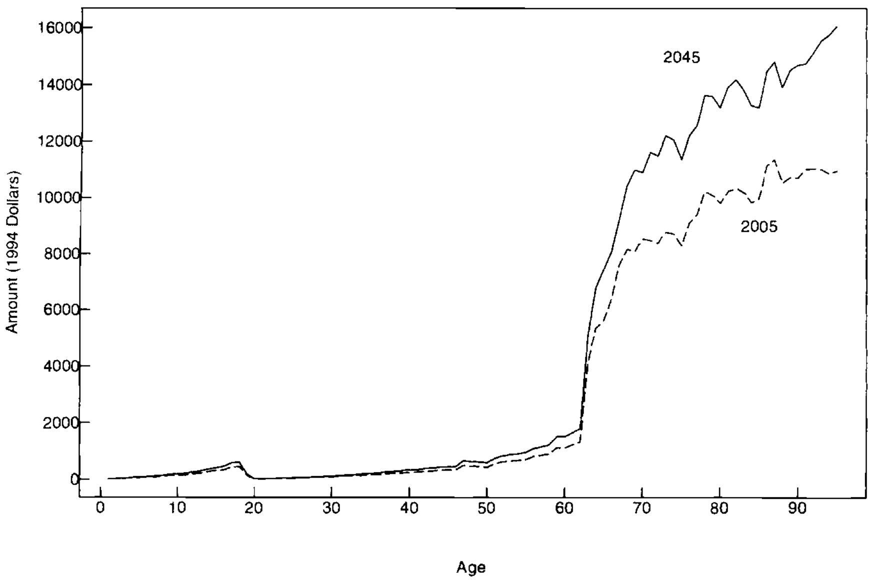

Fig. 9.7 Projected social security benefits received for males (in 2005 and 2045), reflecting 1 percent productivity growth rate and legislated increases in retirement age

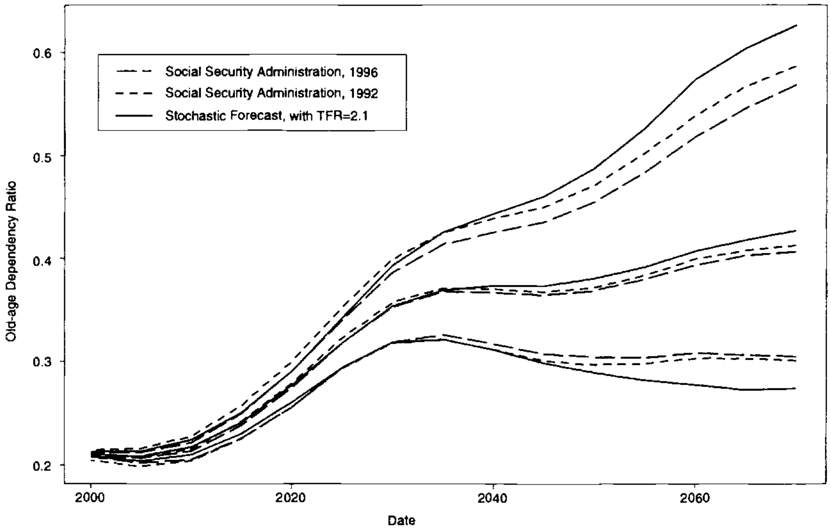

Fig. 9.8 Projected old age dependency ratios to 2070 with brackets
Note: SSA brackets are high- and low-cost variants. Stochastic forecast brackets are 95 percent probability intervals.

and because benefits are higher per capita for those living alone than for those living as married couples, so that increasing widowhood raises the average benefit.

9.4 Forecasts of the Reserve Fund under Legislated Schedule Changes

Having discussed the way in which fertility, mortality, productivity growth, and interest rates are modeled as stochastic processes, we are now ready to discuss the way these are used to generate forecasts of the social security reserve fund. We do this through stochastic simulation. Although it is conceivable that we could derive an analytic solution for the forecasts and their moments, as we have done elsewhere for the population forecasts (Lee and Tuljapurkar 1994), that would be extremely difficult. Stochastic simulation is a straightforward and convenient method for arriving at results for particular models.

The first step is to generate a large set of stochastic simulations of the population from the present to 2070, the end date of our forecasts. We have simulated 750 populations in this way. Our methods follow those in Lee and Tuljapurkar (1994), except that we have increased the migration levels to the intermediate levels in SSA forecasts (Board of Trustees 1996) and have used the SSA initial population for midyear 1994 as the starting population.

Figure 9.8 plots the 95 percent projection intervals for the old age dependency ratio, containing 95 percent of our simulated trajectories. Also plotted

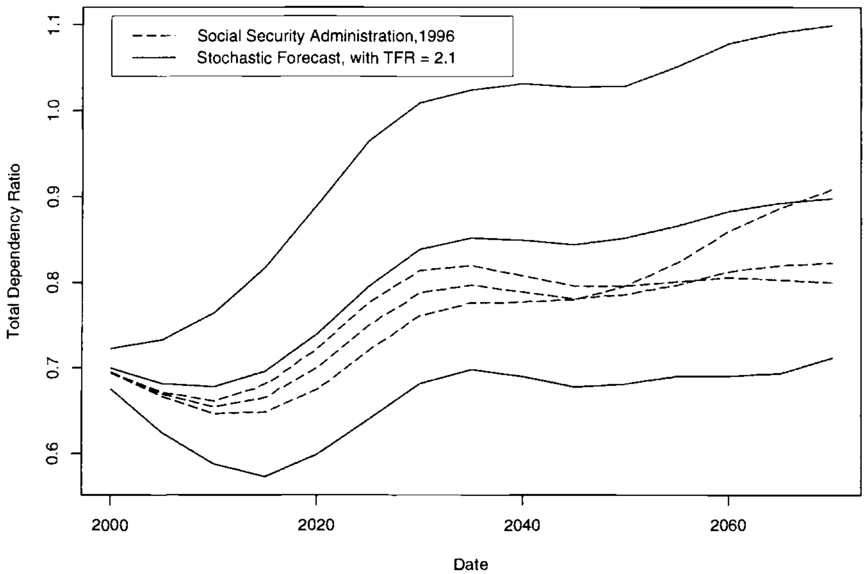

Fig. 9.9 Projected total dependency ratios to 2070 with brackets
Note: SSA brackets are high- and low-cost variants. Stochastic forecast brackets are 95 percent probability intervals.

are the SSA 1992 and SSA 1996 projections and intervals (Board of Trustees 1992, 1996). Their projections and ours are evidently quite similar, with our higher fertility assumption compensating for our lower mortality forecast. It is particularly interesting to note that the fixed scenarios used by the SSA generate a range of dependency ratios that is almost as wide as the 95 percent intervals in the stochastic projection.

Any comfort produced by the similarities in figure 9.8 is dispelled by figure 9.9, which displays the total dependency ratio. Whereas the stochastic projections again yield wide projection intervals in distant years, the SSA scenarios are very close to each other throughout the 75-year forecast period. This remarkable closeness results from intrinsic contradictions in the scenario-based approach. In order to indicate what is construed to be an appropriate amount of uncertainty in the old age dependency ratio, the variable of prime concern to the SSA, low fertility and low mortality are combined in the high-cost scenario, and high fertility and high mortality in the low-cost scenario. But low fertility means few children, while low mortality means many elderly, so that when they are combined in the total dependency ratio, they offset one another. For this reason, the SSA high and low total dependency ratios are very close to one another.

For each simulation, we also stochastically simulate a trajectory for real productivity growth and another for interest rates. The productivity growth rate

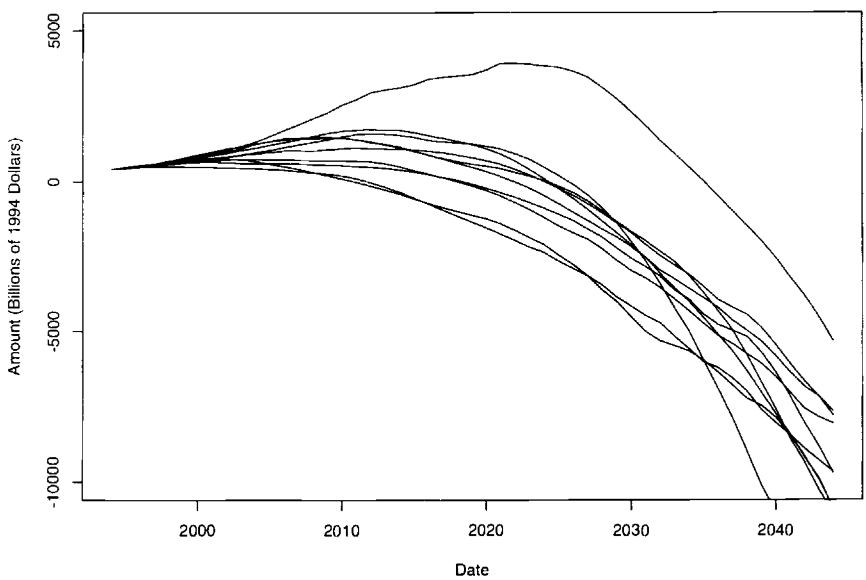

Fig. 9.10 Ten stochastically simulated sample paths of the OASDI trust fund balance

simulation is then used in the ways discussed to alter the tax and benefit profiles, which are next applied to the population age distribution simulation for that year. That generates gross flows of payroll taxes and benefit payments. Additional inflows to the reserve fund come from the simulated interest rate times the reserve fund at midyear, plus revenues from income taxes on benefits.[9] Additional outflows are for administrative costs of the system, and of the separate railroad fund. The difference between total inflows and total outflows, the net flow, is added to the value of the reserve fund. In this way, the state of the system is updated each year, and the level of the reserve fund is calculated for one particular simulated sample path. This process is repeated for each of the 750 stochastic simulations of population, interest rate, and productivity growth. Each simulation assumes that no policy action is taken to adjust tax or benefit rates in response to trends in the reserve fund.

Figure 9.10 plots 10 randomly chosen simulated sample paths out of the total of 750. These lines often cross one another, indicating that the same line is not always the best or always the worst, as would be the case under scenario-

9. This procedure slightly exaggerates the influence of interest rate movements on the finances of the system in the short run, because once bonds are purchased by the trust fund, the nominal interest rate is fixed for the life of the bond. Subsequent fluctuations in the nominal rate will not affect the money earned on that bond. However, most of the volatility in the real interest rate series that we use is due to variation in the rate of inflation. The rate of inflation affects the real rate of return on bonds no matter when they were purchased, consistent with the specification we use.

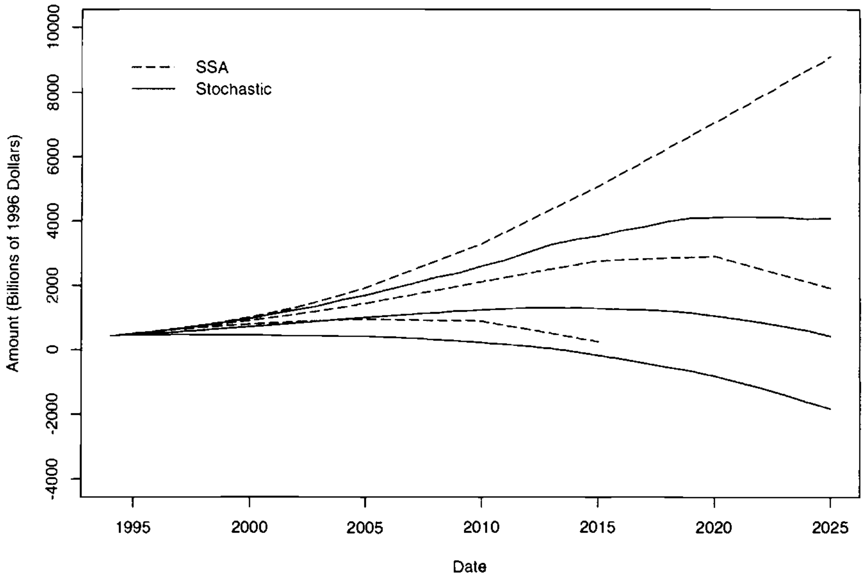

Fig. 9.11 Projected balance of OASDI trust fund (1995 to 2025) with brackets from SSA and stochastic forecasts

Note: SSA brackets are high- and low-cost variants. Stochastic forecast brackets are 95 percent probability intervals.

based forecasting. Based on the full set of 750 (or a larger set of simulations if we wished) we can calculate the mean of sample paths in each year, and we can also find for each year the level of the reserve fund that is greater than 97.5 percent of the paths and the level that is greater than only 2.5 percent of the paths. These two levels then define the 95 percent probability band for each year.

Figure 9.11 shows the forecast only for the first 30 years, up to 2025, for purposes of comparison to the SSA forecast. There are noticeable differences. Although we cannot see it, the simulated mean crosses the line of zero reserves in 2026, four years earlier than in SSA forecasts. The main message, however, is uncertainty about this crossing point: the 95 percent interval includes exhaustion in 2014, as well as exhaustion in 2037, beyond the range of this plot. The level of the reserve fund peaks at $1.3 trillion on average, but the upper bound peaks at $4.0 trillion, and the lower bound at only $0.57 trillion, not much more than it is now.

Figure 9.12 shows the mean and 95 percent probability band for the forecast of the reserve fund, now all the way out to 2070. Because we have not modeled any economic or policy feedbacks and the forecasted fund debt becomes enormous, these forecasts should not be taken at face value. The debt could not actually grow so large without serious consequences for the economy, which

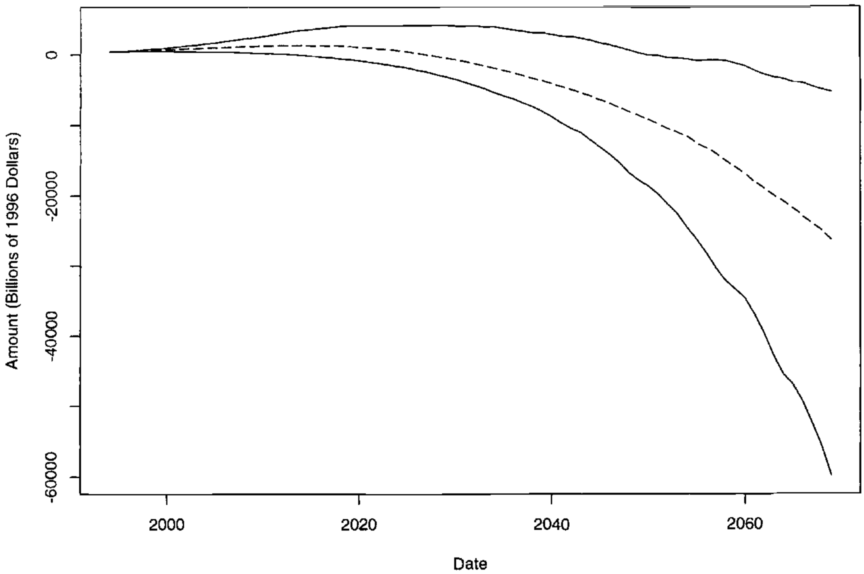

Fig. 9.12 Projected balance of OASDI trust fund (1995 to 2070) with brackets from stochastic forecast

Note: Stochastic forecast brackets (*solid lines*) are 95 percent probability intervals.

would in turn cause changes in interest rates and productivity growth, or lead to policy changes in taxes or benefits.

In figure 9.12, the mean fund balance dips to a debt of $26 trillion by 2070 (all dollars are expressed in 1996 U.S. dollars). The 95 percent range spans debts in 2070 of $6 to $60 trillion. The SSA only publishes fund projections for trajectories above zero, but their low-cost forecast for the fund in 2070 is a positive $9 trillion dollars—a more optimistic projection than ours, despite our assumed expected TFR of 2.1. The most important message from this plot is not the trend in the mean, which should not be far different from other forecasts of the same quantity. The important message is rather the great deal of uncertainty surrounding the mean forecast, and the explicit quantification of the probabilities.

We do not, in fact, mean to suggest that the trust fund will actually go to zero, let alone to debts of trillions of dollars. As the fund begins to fall toward zero, action will be taken to raise the tax schedule, reduce benefits, delay retirement, further tax benefits, invest the reserves in equities yielding higher returns than government bonds, privatize the system, or in some other way prevent the system from going into debt. These forecasts, conditional on the future tax and benefit rules conforming to current plans, are rather intended to shed light on the consequences of those plans. Section 9.5 takes up a different kind of calculation that we believe to be more interesting.

Table 9.1 Summary Actuarial Balance According to SSA and the Current Model (percentage increase in payroll tax rate necessary to restore actuarial balance)

	Summary Actuarial Balance, by Period		
Source of Projection	1995–2020	1995–2045	1995–2070
SSA 1996	.36	−1.39	−2.19
LT: SSA mortality and fertility, deterministic	.82	−1.28	−2.31
LT: SSA mortality, TFR = 2.1, deterministic	.82	−1.25	−2.14
LT: LC mortality, TFR = 2.1, stochastic	.50	−2.00	−3.33

Notes: In deterministic simulations, the discount rate is 2.3 percent; in the stochastic simulation, the discount rate for each sample path is the corresponding simulated rate. "SSA mortality" corresponds to life expectancy of 80.7 years in 2070; "LC mortality," for Lee and Carter (1992) mortality, corresponds to life expectancy of 86.1 years in 2070. "SSA fertility" corresponds to TFR = 1.9. In our simulations, SSA mortality is a mortality trajectory calculated using the Lee-Carter model, but with an imposed rate of decline resulting in the SSA life expectancy level in 2070. Details of the age pattern of decline and the timing of decline differ from SSA assumptions.

9.5 Long-Term Actuarial Balance and Tax Increases

The SSA employs a summarized measure of the long-term status of the trust funds, based on what are called summarized cost rates and summarized income rates. The basic measure, the long-term actuarial balance, is roughly speaking equal to the difference between the present value of the stream of tax revenues and benefit payments over the projection period, divided by the present value of the stream of total payroll. It can be interpreted as the amount by which the payroll tax rate would have to be raised, immediately and permanently, to equalize these two present values, taking into account the existing reserve fund at the start and the target reserve fund at the end. We will use the SSA calculations of actuarial balance as a kind of benchmark against which to test the long-term performance of our model. We will also use this measure as a convenient metric for representing the relative importance of different sources of uncertainty in our stochastic forecasts.

Table 9.1 reports on the calculated summary actuarial balance based on four different models and for each of the four shows the calculated balance over three different periods. The first row reports the balance according to the SSA (Board of Trustees 1996). This report concludes that the system would be in balance through 2070 if the payroll tax rate were immediately and permanently raised from 12.4 percent (for OASDI) to 14.59 percent (= 12.4 + 2.19). We attempt to replicate this result by running a deterministic simulation in which TFR is set at the SSA level of 1.9 children per woman and mortality is assumed to decline at a constant rate (distributed by age according to the Lee and Carter, 1992, model) such as to achieve the SSA projected life expectancy in 2070. Interest rates and productivity growth rates are assumed fixed at the SSA levels of 2.3 percent and 1 percent. The result is shown in the second row, resulting in a figure of −2.31 versus −2.19, which we view as excellent agreement, given our very different approach to projecting the tax and benefits schedules. Additional experiments, not reported here, show that deterministic sensitivity analyses using our model yield results very similar to those reported by SSA.

The third row of table 9.1 shows another deterministic simulation, which differs from the previous one only in that TFR is now set at 2.1, as in Lee and Tuljapurkar (1994). The agreement with SSA is very slightly worse in the medium term, but somewhat closer in the long term. Finally, the last row shows the mean result from the set of stochastic simulations, this time with TFR = 2.1 and mortality declining at the faster pace of Lee and Carter (1992), resulting in roughly twice the gain in life expectancy projected by the SSA. Because of this more rapid mortality decline, the actuarial balance for 1995–2070 is now −3.33, or about 1.2 percent worse than with the SSA mortality.

The SSA (Board of Trustees 1996, 23) gives a range for this summary actuarial balance from +.46 percent to −5.67 percent, with an interval width of 6.13 percent (= .46 + 5.67). The standard deviation of our stochastic summary actuarial balance is 1.58 percent, leading to a 95 percent probability interval ranging from −.17 percent to −6.49 percent, for an interval width of 6.3 percent, very similar to the SSA's. This indicates that the SSA's low-cost–high-cost range for 2070 has an approximate 95 percent probability coverage, assuming it is correctly centered. However, our interval is centered at −3.33, versus −2.19 for theirs, a fairly substantial difference. The difference arises from the more rapid mortality decline in the stochastic simulation.

9.6 Fund Exhaustion and Tax Increase

According to the SSA projections, a 2.19 percent increase in the OASDI tax rate should restore the system to long-term balance, under the intermediate set of assumptions. We can use our stochastic simulation (as reported in the last row of table 9.1) to assess the likelihood that exhaustion would occur in any case. Panel *A* of figure 9.13 shows the probability distribution of dates of exhaustion of the OASDI trust fund if taxes were immediately raised by 1 percent, to 13.4 percent. In this case, 93.4 percent of the sample paths reach fund exhaustion before the end of the period in 2070, and the median fund balance in 2070 is −$17 trillion (1996 dollars). Panel *B* shows the corresponding probability distribution assuming an immediate 2 percent increase in the tax rate, to 14.4 percent. In this case, three-quarters of the sample paths still end in exhaustion by 2070, and the median fund balance then is −$8 trillion.

We could search for the permanent tax increase necessary to achieve a 95 percent probability of nonexhaustion by 2070. However, the requirement that there be no subsequent tinkering with taxes and benefits to achieve balance in the future, as more is learned about actual demographic and economic developments, is unappealingly rigid. Instead we pursue a different approach, going to the opposite extreme of adjusting taxes on a year-by-year basis to meet costs.

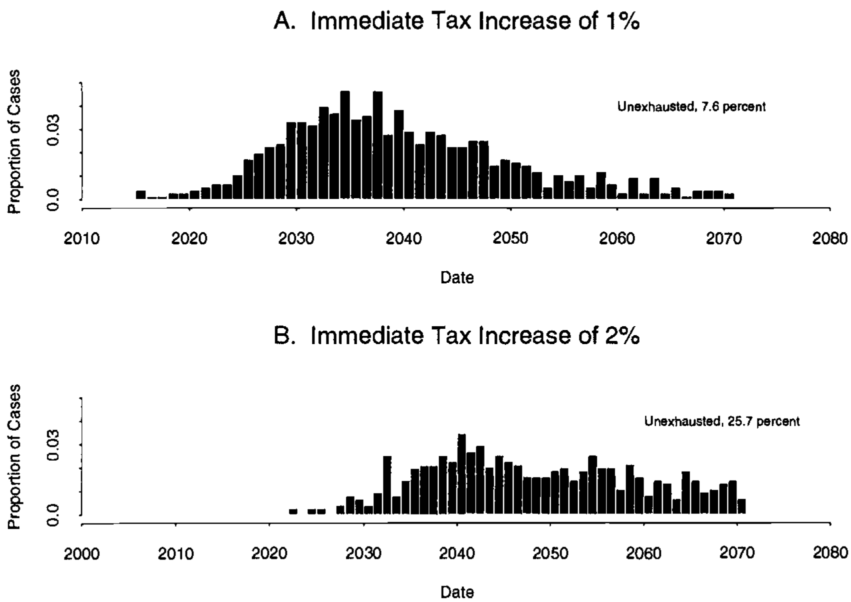

Fig. 9.13 Histograms of dates of exhaustion

9.7 Forecasts of the Payroll Tax Rate under Reserve Fund Constraints

Suppose that instead of fixing a tax rate to hold for the next 75 years, we set the tax rate each year to maintain the reserve fund at a level exactly sufficient to cover one year's benefit payments? We can do this calculation in two ways. One is to leave the currently legislated 12.4 percent tax rate in place until the reserve falls to 100 percent of anticipated year-ahead outflows and thereafter raise or lower the tax rate as necessary to leave the reserve at this level. We call this "eventual pay as you go." The other way is to let the reserve fund immediately fall to 100 percent of outflow (which takes only a couple of years of zero or very low taxes) and thereafter adjust it as necessary. This approach gives us the pure PAYGO rate.

We present results only for the first calculation, eventual PAYGO. We hold constant the 12.4 percent rate until the reserve fund drops to the target level. The results are shown in figure 9.14, which plots lines giving varying probability coverages. For example, in any year, 60 percent of the sample lines lie below the line labeled .6. The median line (not shown) would lie roughly halfway between the .4 and .6 lines. It would remain at 12.4 percent until 2022, rise rapidly as the baby boom retires, and then rise more slowly to 24 percent in 2070. The lines labeled .025 and .975 bound the region with 95 percent probability coverage. The lower .025 bound remains at 12.4 percent until 2043

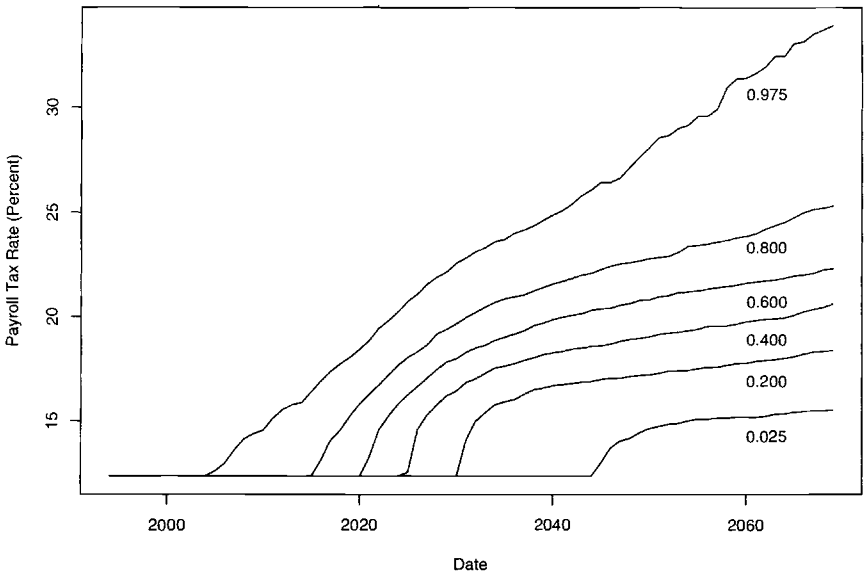

Fig. 9.14 Tax rate for year-ahead balance (distribution by selected quantiles)

and then rises modestly to 16 percent in 2070. The upper 97.5 percent bound rises roughly linearly after 2003, to 34 percent in 2070.

In sum, with the eventual PAYGO system with variable tax rates the median tax rate would double by 2070 and still be rising slowly at that point, reflecting continuing mortality decline. There is a 2.5 percent chance that the tax rate would have to rise to 34 percent of earnings by the end of the forecast in 2070. A 34 percent payroll tax rate, plus additional taxes for Medicare and Medicaid, plus other state, local, and federal taxes, would be a terribly heavy burden.

These results may be compared to annual cost estimates by the SSA (Board of Trustees 1996, 170–71) for 2070, which gives a high-medium-low range of 28.0, 18.8, and 13.1 percent, respectively. The range widths are again similar: 15 percent for SSA versus 18 percent for the stochastic simulation, indicating that the SSA probability coverage would be somewhat less than 95 percent if the range were correctly centered. However, the stochastic simulation gives a median tax in 2070 of 24 percent, versus 18.8 percent for the SSA. The increase from the initial tax of 12.4 percent is nearly twice as great in the stochastic simulation as in the SSA projection. This difference arises from the more rapid decline in mortality incorporated in the stochastic simulation.

9.8 How Much Uncertainty Does Each Component Contribute?

We have seen that there is a great deal of uncertainty about the long-term finances of the system. Where does this uncertainty originate? What would we

Table 9.2 Sources of Uncertainty in Forecasts of Summary Actuarial Balance (standard errors when one variable at a time is stochastic)

	Standard Error of Summary Actuarial Balance, by Period		
Stochastic Variable	1995–2020	1995–2045	1995–2070
Fertility	.02	.27	.86
Mortality	.19	.39	.54
Productivity growth rate	.90	.90	.80
Interest rate	.14	.50	.69
Fertility and mortality	.19	.48	1.02
All	.93	1.18	1.58

Note: The discount rate is 2.3 percent, except when the interest rate is stochastic, in which case the simulated trajectory of interest rates is used for each sample path.

need to know to reduce it? The Board of Trustees (1996) reports a sensitivity analysis for its forecasts, showing the effect on the long-run balance of the system when one factor varies across the specified high-low range, while holding all others at their mean values. The outcome, of course, depends on the sensitivity of the projection to each variable, but it also depends on the amount of change that is examined for each variable—the width of the range. We can avoid arbitrary assumptions about this range by using our stochastic simulation model. Holding all but one of the variables (fertility, mortality, interest rates, and productivity growth rates) fixed at their mean trajectories, we allow the fourth to vary stochastically in the usual way. The resulting widths of the probability bands can then be compared. Results of this exercise are reported in table 9.2, which gives the standard deviation of the summary actuarial balance in each case, for each of three time periods.

From the last column of the table, we see that fertility contributes the greatest uncertainty through 2070, followed by the productivity growth rate, the interest rate, and finally mortality. It is striking that whereas our model puts uncertainty about mortality *last* in importance, the SSA analysis (Board of Trustees 1996, 132–34) puts it *first.* And whereas we put fertility *first* in importance, the SSA puts it *last* (tied with the interest rate).

It is also interesting to note how the relative and absolute contributions to uncertainty change with the period over which actuarial balance is assessed. Fertility makes only a trivial contribution over the period 1995–2020 because there is a long lag between birth and labor force entry, while the productivity growth rate makes by far the strongest contribution over this period. Interest does not matter much because the trust fund is not very large, and mortality, which operates cumulatively on survival rates, has insufficient time to make much difference.

The fifth row of the table shows the standard deviation of the actuarial balance when both fertility and mortality are stochastic but the economic variables are fixed, and the last row gives the result when all four variables are stochastic. Comparing these, we see that over the 25-year horizon, demographic uncertainty generates a standard deviation only one-fifth as wide as the fully stochastic model. However, over a 75-year horizon, demographic uncertainty alone generates a standard deviation almost two-thirds as wide as the fully stochastic model. Evidently, demographic uncertainty becomes far more important over the long run than it is over the shorter run.

9.9 Conclusion

We are still developing the methods used for these stochastic simulations, and there will doubtless be future changes. Nonetheless, this macrosimulation model replicates quite closely the key SSA results, when run in deterministic mode using SSA assumptions. This confirms that the basic simulation model is mechanically sound. Most of the material we present is based on stochastic models of fertility, mortality, interest rates, and productivity growth rates, with expected values that conform closely to the SSA intermediate assumptions. The exception is mortality, for which we believe the SSA projected rate of decline is too slow. Our fitted stochastic mortality model foresees twice as great gains in life expectancy by the year 2070 as do the SSA projections.

We considered long-term stochastic projections of the trust fund balance, the summary actuarial balance, and the PAYGO tax rate. Because we do not incorporate economic or policy feedbacks in our model, outcomes in which large positive or negative trust fund balances occur are bound to be quite unrealistic. With that caveat, our 95 percent probability intervals for 2070 were as follows: for the trust fund, in 1996 dollars, −$6 to −$60 trillion; for the summary actuarial balance, −.2 to −6.5 percent of the present value of payroll; and for the PAYGO balanced budget tax rate in 2070, 16 to 34 percent of payroll.

Because of the more rapid expected rate of mortality decline used in our stochastic simulations, our financial projections are somewhat more pessimistic than those of the SSA. For example, our expected summary actuarial balance is −3.33 percent of the present value of payroll, versus −2.19 for the Board of Trustees (1996). However, the width of the SSA high-cost–low-cost ranges, to which no probability interpretation is attached, are surprisingly close to the width of our 95 percent probability intervals, when compared for horizons of 2020, 2045, and 2070. These widths are much greater than the differences in means between the stochastic simulation projections and the SSA projections, which is reassuring.

According to our simulations, if currently legislated tax and benefit rates are unchanged, there is a 95 percent chance of trust fund exhaustion between 2014 and 2037. There is only a 2.5 percent chance that the trust fund balance in 2070 will exceed −$6 trillion (1996 dollars). If the tax rate were immediately and permanently raised by 1 percent, there would still be a 92 percent probability of trust fund exhaustion before 2070. If the tax were raised by 2 percent, close to the 2.19 percent that the SSA suggests would be needed to achieve long-term balance, we find that there would still be a 75 percent chance that the trust fund would be exhausted before 2070.

We have also computed the implicit rate of return that the generation born in 1980–84, and therefore just about to enter the labor force, will receive through OASDI. We project an expected real rate of return of 1.6 percent, with a 95 percent probability interval of .2 to 2.8 percent.

When we examined the relative contributions of our four stochastic factors to uncertainty in the projections, the results disagreed sharply with the SSA sensitivity analysis. For example, we found fertility to be by far the greatest source of uncertainty about the long-run finances of the system, while the SSA found it to be least important. We found mortality to be least important in the long run, while the SSA found it to be most important. Our results also showed clearly that demography matters most in the long run, where its tidal forces have a cumulative impact.

We are still at early stages of digesting our stochastic forecasts of the finances of the social security system and are still exploring new ways in which these forecasts and experiments might be useful and informative. One promising use, not yet implemented, will be to test the consequences of a range of strategies and policy options for dealing with uncertainty. Is it better to wait and see, adjusting policy continuously as we gain information? Or is it better to accumulate large reserves early on, to provide a buffer against unlikely but possible transitory insults to the system? Or should policy simply be set to deal reasonably with the mean trajectory, ignoring the uncertainty? How about keying the level of benefits to life expectancy at retirement, as in the Swedish system, or tying cohort benefit levels to cohort fertility, as has sometimes been suggested? Stochastic simulations could provide a useful laboratory for testing these alternative policies in relation to goals such as achieving intergenerational equity, avoiding rapid changes in taxes or benefits, and keeping the trust fund above zero.

References

Aaron, Henry, Barry Bosworth, and Gary Burtless. 1989. *Can America afford to grow old? Paying for social security.* Washington, D.C.: Brookings Institution.

Board of Trustees. Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds. 1992. *1992 Annual report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, D.C.: Government Printing Office.

———. 1995. *1995 Annual report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, D.C.: Government Printing Office.

———. 1996. *1996 Annual report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, D.C.: Government Printing Office.

Congressional Budget Office. 1996. *The economic and budget outlook: Fiscal years 1997–2006.* Washington, D.C.: Government Printing Office.

Foster, Richard S. 1994. A stochastic evaluation of the short-range economic assumptions in the 1994 OASDI trustees report. Actuarial Study no. 109 (SSA Publ. no. 11-11502). Washington, D.C.: U.S. Department of Health and Human Services, August.

Harvey, Andrew. 1989. *Forecasting, structural time series models and the Kalman filter.* Cambridge: Cambridge University Press.

Holmer, Martin. 1995a. Briefing on trust-fund asset-allocation policy: Stochastic simulation model. SSA Contract no. 95-22582.

———. 1995b. Overview of SSASIM: A long-run stochastic simulation model of social security. SSA contract no. 95-22582.

Hurd, Michael. 1990. Research on the elderly: Economic status, retirement, and consumption and saving. *Journal of Economic Literature* 28:565–637.

Lee, Ronald D. 1993. Modeling and forecasting the time series of U.S. fertility: Age patterns, range, and ultimate level. *International Journal of Forecasting* 9:187–202.

Lee, Ronald D., and Lawrence Carter. 1992. Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association* 87:659–71.

Lee, Ronald D., and Shripad Tuljapurkar. 1994. Stochastic population forecasts for the U.S.: Beyond high, medium and low. *Journal of the American Statistical Association* 89:1175–89.

Lumsdaine, Robin L., and David A. Wise. 1994. Aging and labor force participation: A Review of trends and explanations. In *Aging in the United States and Japan,* ed. Yukio Noguchi and David Wise, 7–41. Chicago: University of Chicago Press.

Shripad, Tuljapurkar, and Ronald Lee. In press. Demographic uncertainty and the OASDI fund. In *Intergenerational economic relations and demographic change,* ed. Andrew Mason and Georges Tapinos. Oxford: Oxford University Press.

Social Security Administration (SSA). Office of Research and Statistics. 1994. *Social security bulletin annual statistical supplement.* Washington, D.C.: Government Printing Office.

Comment Sylvester J. Schieber

The basic premise in Ronald Lee and Shripad Tuljapurkar's (hereafter L-T) paper is that the planning surrounding major government entitlement commitments in general, and social security in particular, must not only take into account our best estimates about the future costs and cost drivers of these programs but also must consider the uncertainty surrounding those estimates. On the basis of this premise, they develop a set of stochastic forecasts of the future operations of social security.

L-T's approach to forecasting social security operations in the paper has been supported by various groups in the past, including some in advisory roles to the Social Security Administration (SSA). For example, the Technical Panel

Sylvester J. Schieber is vice president of research and information for Watson Wyatt Worldwide.

on Assumptions and Methods that supported the 1994–96 Social Security Advisory Council recommended the "implementation of stochastic analysis procedures for presenting and evaluating the uncertainty in OASDI projections" (Technical Panel 1996, 137). Following on the recommendations of the Technical Panel, the Advisory Council itself recommended that such "modeling should be used as a tool for recognizing explicitly the uncertainty surrounding the . . . demographic and economic assumptions" used in valuing the program's operations (Advisory Council 1997, 22).

While the approach that L-T have recommended and the analysis that they have done may receive relatively widespread support among policy analysts in general, it is still a debated approach within social insurance policy analysis circles, and it is not a path that has been taken by the social security program's actuaries in developing official cost estimates. Robert J. Myers, former chief actuary of the SSA, believes that the deterministic scenario-based modeling that is currently used in projecting social security operations is adequate and that the high- and low-cost estimates give a reasonable range of costs that might be expected under the program.

Myers has laid out a rationale for continued use of the deterministic approach and makes three points to support it. First, he believes that "experienced judgment" is superior to "blind mathematical analysis of past experience" in developing these estimates. Second, he believes that the present practice of lumping together the low-cost (high-cost) assumptions to develop low-cost (high-cost) program projections produces a reasonable "range" of cost estimates in the aggregate. Third, he argues that "any large changes . . . should be phased in over a period of years, so that there is a reasonable certainty that they should be made in their entirety" (Technical Panel 1996, 267–68). In the last point, it is not clear whether Myers is arguing that changes in *assumptions* should be phased in slowly or whether changes in *program costs* should be recognized on a phased basis. To a certain extent, phasing in changes in assumptions over a period of annual valuations will result in the slow recognition of changing cost estimates. Regardless of his precise meaning on the third point, Myers's arguments for staying with the current valuation methodology offer a framework for reviewing the work reported in L-T's paper.

Accuracy of Prior Estimates

One way to judge the quality of the current approach to projecting social security costs or whether it might be desirable to move to a stochastic modeling approach is to review prior projections to see how they have comported with ultimate experience. In this regard, the results are mixed. Estimates of OASI program costs developed by the program's actuaries after the passage of the 1939 amendments to the Social Security Act projected the cost of OASI in 1970 at 6.33 percent of covered payroll based on "original assumptions" and at 8.54 percent of payroll based on "probable maximum cost assumptions." The respective projections for 1980 were 7.21 percent and 10.60 percent of covered payroll (Bronson 1939). The actual cost of OASI benefit payments in 1970 was 6.98 percent of covered payroll (Board of Trustees 1971), and in 1980 it was 9.39 percent (Board of Trustees 1981).

Based on this evidence, it seems the early estimates of long-term program costs were reasonable predictions of the ultimate tax burden to support it. While the SSA's actuaries likely never dreamed of the rapid expansion of benefits during the 1970s, their 1939 "maximum probable cost" estimate for 1980 still bounded the actual cost of the program. The early history of the program and its related cost projections, however, might not be instructive for the future that we now face. It seems there was a general sense among the early advocates of social security that the economy and the public would or could bear cost rates up to 12 percent of covered payroll. The managers of the program were fully cognizant that costs would rise over time as a growing share of the working population attained "insured" status under the program. Over the years, benefits were increased repeatedly but always within the constraints that estimated cost rates allowed. In this environment, the 1939 cost projection for 1970 became a self-fulfilling prophecy. It was accepted from the earliest days of the program as being a reasonable cost, and benefit levels were merely adjusted periodically to move toward it. The target cost rate could be achieved without being exceeded because the program was run on a nominal basis. If costs started to get out of hand, inflationary forces could bring them back into line with acceptable rates. This all changed with the 1972 amendments that automatically indexed benefits and resulted in the double indexing of initial benefit awards for new beneficiaries. These amendments made the program particularly susceptible to price inflation, especially to the extent that it significantly exceeded the rate of growth of wages.

Social Security has always been run largely on a pay-as-you-go basis. Under such a financing regime, revenues must roughly equal expenditures from year to year. In simple mathematical terms,

$$t \cdot N_w \cdot W = N_b \cdot B, \tag{1}$$

where t is the payroll tax rate, N_w the number of workers who are covered by the system, W their average covered wages, N_b the number of beneficiaries, and B their average benefit levels. Stated in another way,

$$t = (N_b/N_w) \cdot (B/W), \tag{2}$$

where the ratio of beneficiaries to workers can be thought of as the dependency ratio and the ratio of benefits to wages might be thought of as the benefits ratio or, in the parlance of retirement plans, the system's average wage replacement level.

Table 9C.1 compares the midrange economic assumptions used in developing the SSA's cost projections for the 1972 Board of Trustees report with actual experience for the five years from 1972 through 1976. As a result of the

Table 9C.1 Comparison of Five-Year Economic Assumptions in the 1972 OASDI Trustees Report with Actual Experience

	CPI Increase (%)		Real Wage Increase (%)		Unemployment Rate (%)	
Year	Assumed[a]	Actual	Assumed[a]	Actual	Assumed[a]	Actual
1972	2.75	3.3	2.25	4.0	4.2	5.6
1973	2.75	6.2	2.25	0.7	4.2	4.9
1974	2.75	11.0	2.25	−3.6	4.2	5.6
1975	2.75	9.1	2.25	−2.5	4.2	8.5
1976	2.75	5.8	2.25	2.5	4.2	7.7
Total	14.53	40.6	11.77	1.0	4.2[b]	6.5[b]

Sources: 1972 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (Washington, D.C.: Government Printing Office, 1972) and *Economic Report of the President* (Washington, D.C.: Government Printing Office, 1982).

[a]Midrange assumptions.

[b]Estimates for five-year unemployment totals and five-year averages.

1972 social security amendments and what was going on in the economy at the time, benefits absolutely exploded relative to wage levels, and the beneficiary-to-worker ratio was adversely affected by the higher than expected unemployment levels. There was a clear breakdown in the reliability of the cost estimation process at this point in social security's history. It is conceivable that stochastic modeling would have helped catch the problems introduced in the 1972 amendments before their passage, but it is also quite possible that it would not have.

The 1977 social security amendments were intended to bring benefits back into line with their more traditional levels and to stabilize the relationship between benefits and wages looking forward, subject to the variations noted by L-T in their analysis. The 1977 amendments proved to be insufficient to deal with the continuing adverse economic conditions of the late 1970s and early 1980s, and the program had to be rebalanced once again in the 1983 amendments. The actuaries' projections after the adoption of the 1983 changes predicted that the trust fund balances would accumulate to more than $20 trillion but would ultimately be depleted in 2063 (Ballantyne 1983). Subsequent projections in the annual Board of Trustees reports have consistently ratcheted down the trust fund accumulation, with the latest projection that it would only accumulate about $2.8 trillion at its peak and would be depleted by 2029 (Board of Trustees 1997). The 1997 intermediate projection actually turned out to be somewhat less optimistic than the pessimistic projection from 1983.

The point of this discussion is that since social security has matured and the benefit structure been indexed, the cost projections for the program have become somewhat more sensitive to economic and demographic variations than they had been in the past. Recent projections have suggested an unwarranted level of certainty about future costs that have often proved to be badly out of range, and "worst case" scenarios have proved to be optimistic. In short, the projection game has changed, and the old methods no longer seem adequate to the task of providing realistic cost estimates in some cases, or a sense of the vulnerability of those estimates to varying experience under the program.

Stochastic Projections versus Reliance on Experienced Judgment

In Myers's defense of the current SSA projection methods he argues that the "experienced judgment" behind current projections is better than "blind mathematical analysis" of the past. The implication is that stochastic modeling necessarily relies solely on mindless projection of past trends with the historical variation around those trends built into the projections that are being developed. The work by L-T shows that this does not have to be the case. For several of the important variables used in determining social security's long-term costs, L-T constrain their projections to mean distributions that are equivalent or close to the best-guess estimates used by the SSA actuaries in their long-term projections.

However, the results of the L-T stochastic simulations raise a question about the superiority of such projections over the current scenario-based deterministic projections that are formally used in monitoring and projecting social security operations. For example, the close correspondence between the L-T 95 percent probability intervals for the aged dependency ratios and those generated in the high- and low-cost projections developed by the SSA actuaries suggests a similar correspondence in long-term cost projections, given the importance of this dependency ratio in determining costs, as shown in equation (2) above. Indeed, L-T report that the range of their cost estimates from 95 percent probability levels corresponds closely to the range in the low- and high-cost estimates published by the Board of Trustees. L-T's costs are generally higher than those from the SSA largely because of the difference in their assumptions about continued improvements in life expectancy. Earlier work by Ronald Lee and Lawrence Carter (1994) supports the contention that the SSA is being overly optimistic in its assumptions about future life expectancy—that is, it is assuming slower increases in life expectancy than many demographers think is reasonable. The Technical Panel for the 1994–96 Advisory Council was very critical of the assumptions being used by the SSA in its projections (Technical Panel 1996, 251–62).

The criticism that current formal SSA projections are not based on proper assumptions does not necessarily mean that the actuaries' current cost projection methods are inferior to a stochastic approach. But this issue of appropriate assumptions cuts directly against the first of the points that Robert Myers makes in arguing for the current projection methodology over stochastic modeling of the future. If we are to rely on the "experienced judgment" of those doing projections, no matter what their method of projecting that judgment should not blindly ignore the past unless there is a powerful reason for doing

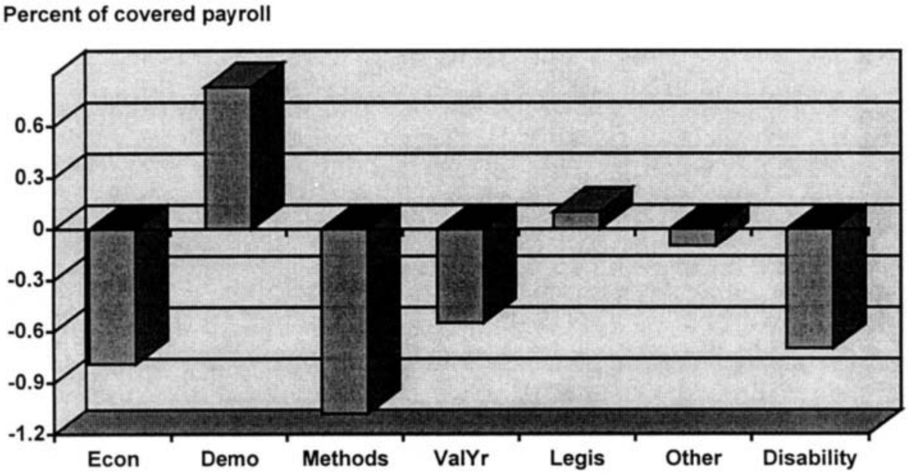

Fig. 9C.1 Changes in 75-year estimated OASDI actuarial balance as a percentage of covered payroll over projection years 1983–95
Source: U.S. Department of the Treasury, Office of the Assistant Secretary for Economic Policy.

so. If the actuaries were using stochastic projection techniques, it is likely that they would use more formal processes in projecting important variables such as life expectancy than they seem to be using now. They would also have to provide a much better rationale than they do now in relying on their "experienced judgment" when such judgment goes against historical trends.

Projecting a Reasonable Range of Outcomes

Robert Myers's second argument for the SSA's staying with its current method of projecting future OASDI costs is that the scenarios project a reasonable range of cost estimates. As noted earlier, the L-T range of projections starting from a base year of 1995 does correspond closely with those of the SSA actuaries. While the range of the two sets of projections may be relatively close, another way to consider the variance in the projections from the respective models is to look at how they might be used in choosing public policies and whether the two approaches would lead to different potential conclusions.

The last time that we undertook a major review of SSA policies that led to significant policy changes was in 1983. After the 1983 social security amendments were adopted, the program actuaries estimated that the program would be in actuarial balance—that is, aggregate revenues plus assets would equal aggregate expenditures—for the next 75 years. Over the years since then, the annual projections of the system have consistently estimated larger and larger actuarial deficits for the program. Figure 9C.1 summarizes, by reason of change, the changes in projected actuarial balances under the formal projections over the period.

A natural question that arises from figure 9C.1 is whether the evolution of SSA projections between 1983 and 1995 would have been any different in a stochastic modeling world than under the deterministic process now used. While it is impossible to be certain, there is some likelihood that the actuaries would have used similar economic and demographic assumptions whether they were doing stochastic or deterministic projections. Beyond the economic and demographic assumptions (Econ and Demo), both models would likely have been plagued by similar experiences. The costs attributed to changing methods (Methods) relate to the discovery of calculation errors or different ways of looking at cost calculations. The change in costs attributed to differences in valuation year (Val Yr) are related to doing 75-year forecasts and the demographic structure of the population; namely, each year we give up one good year in the projection and add a bad one. The cost changes attributed to disability relate to growing incidence rates of such benefits, and it is not clear that explicit modeling of disability on a stochastic basis would have anticipated these increases any more quickly than the actuaries have. L-T are not explicitly stochastically modeling disability at this time, but given the sensitivity of cost estimates to these claims, it would seem like a valuable addition to make to the program.

Given the close correspondence between the L-T and SSA aged dependency ratios noted earlier, the significant divergence in total dependency ratios under the two projection methods is curious. It is easy to understand why L-T would get greater variance in total dependency than the SSA actuaries. The latter, in their low-cost estimates, pair high fertility rates with low improvements in life expectancy, and in their high-cost estimates, pair low fertility rates with high improvements in life expectancy. In some of L-T's simulations, high fertility rates get paired with high improvements in life expectancy resulting in high total dependency rates, a situation that never arises in the actuaries' deterministic projections. The larger variance in the total dependency ratio in the L-T stochastic projections results from the underlying variations in fertility rates in their alternative simulations. But conventional wisdom suggests that the high-fertility scenarios, which would initially drive up the total dependency ratio, should ultimately drive down the aged dependency ratio as the greater number of children being born under the scenario age and enter the workforce. The inconsistency in the results in these two dependency ratios either deserves more analysis or more explanation.

Clearly, stochastic projection models can produce ranges of estimates for programs like social security that are as much within the boundaries of reasonable expectations as the models that are currently being used. They have the possible added benefit that they give a new perspective on how variations in various demographic or economic variables can drive overall program costs. L-T raise this issue in their analysis of the comparative sensitivity of cost estimates to variations in fertility, mortality, productivity growth, and interest rates, although they do not fully explain why their stochastic results are so different from those generated by deterministic projections.

Modeling and Implementing Policy Changes

It is not clear what Robert Myers meant when he wrote in support of the current projection methodology used by the SSA that "large changes . . . should be phased in" gradually. Certainly it makes sense to implement significant policy changes on a phased basis in many cases. It makes no sense, however, to phase in our understanding of the financial obligations that our entitlement programs present and the uncertainties associated with those obligations for our citizens in the future. Policymakers and citizens alike should understand the implications of current policy as soon as policy analysts have relative confidence in their own understanding of these programs.

One of the problems in developing public policy estimates that are bounded by degrees of uncertainty is that policymakers do not like uncertainty. When a problem arises, they want to be able to solve it, or at least to adopt policy changes that they can describe to their constituents as solving it. When social security was in financial crisis in the early 1980s, the deterministic model allowed policymakers to claim that they had solved the problem by adopting the 1983 amendments because the actuaries' projections showed that they had. It is not clear how the same policies would have been greeted under the headline "Congress Adopts Changes That Have a 60 Percent Probability of Solving the Social Security Financing Problem." In retrospect, the public probably would have been better off if something like that had been the message, but putting probability distributions on policy prescriptions may further complicate the process of policy making itself.

We are now at a point in the evolution of our public entitlement programs that there is a fairly widespread consensus that they are seriously underfinanced. The sense of concern about these programs has reached the point that policy analysts are now making proposals to reform them in ways that heretofore were never seriously considered. Many of the proposals that are now being put on the table would likely have significant effects on the economy that go beyond the rebalancing of the entitlement programs themselves.

Possibly the greatest weakness of the deterministic model used to project social security's costs is its relative inability to consider feedback effects that relate to OASDI's overall effects on the economy. Currently, the L-T model suffers from the same problem. In their conclusion the authors note that one potential use of their stochastic model is testing policy options in the context of our ability to limit uncertainty under the programs and proposals to reform them. L-T are absolutely correct in their assessment of the potential usefulness of these kinds of models, but before their model can be used in this way, it must be able to consider feedback effects. They note in the presentation of their results that some of the scenarios they estimate with the current model are implausible because the trust fund deficits become so large that they would undoubtedly have implications for the larger economy. Many of the solutions for the current underfunding of social security anticipate significant increases in trust fund or savings levels that would also have significant macroeconomic effects. These must ultimately be considered in the testing of policy options.

Social security is not the only program around which the issues raised by L-T in this paper might be considered. Under the actuaries' midrange projections, social security benefit claims are expected to rise from around 4.5 percent of GDP today to about 6.5 percent of GDP by 2030. Comparable projections for Medicare would have its claims rising from around 2.5 percent of GDP today to 7.5 percent of GDP by 2030. In reviewing the assumptions used in making this projection, the Medicare actuaries are assuming that the excessive inflation that has plagued this program from its outset will somehow be ameliorated around 2010. Beyond that, they assume the rate of growth in costs beyond pure demographic effects will be at the rate of growth in the economy. In other words, Medicare's actuaries are assuming a significant reduction of inflation in this sector of the economy just as the baby boomers make their largest claims on Medicare and the health delivery system. The uncertainties and risks associated with Medicare far outweigh those in social security and are deserving of the kind of scrutiny that L-T are advocating in their paper.

This work by Ronald Lee and Shripad Tuljapurkar is a good beginning in helping us to understand some of the uncertainty surrounding social security cost projections. The authors should be encouraged to continue the development of their model and to apply it to other public programs.

References

Advisory Council on Social Security. 1997. *Report of the 1994–1996 Advisory Council on Social Security.* Washington, D.C.: Social Security Administration.

Ballantyne, Harry C. 1983. Long-range projections of social security trust fund operations in dollars. *Social Security Administration Actuarial Notes,* no. 117 (October): 2.

Board of Trustees. Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds. 1971. *1971 Annual report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, D.C.: Government Printing Office.

———. 1981. *1981 Annual report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, D.C.: Government Printing Office.

———. 1997. *1997 Annual report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, D.C.: Government Printing Office.

Bronson, Dorance C. 1939. Social Security Board internal memorandum, 31 August. Files of the Social Security Board. National Archives, Washington, D.C.

Lee, Ronald D., and Lawrence Carter. 1994. Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association* 87:659–71.

Technical Panel on Assumptions and Methods. 1996. *Report of the Technical Panel on Assumptions and Methods.* Washington, D.C.: Social Security Administration.