Uncertainty in Social Security Trust Fund Projections

Abstract - *This paper presents measures of uncertainty about Social Security Trust Fund projections based on the new Long–Term Actuarial Model (LTAM) being developed at the Congressional Budget Office. Measuring the variance in Social Security outcomes involves three steps: specifying a model, characterizing uncertainty about model inputs, and generating useful measures of the uncertainty about model outputs. There are significant trade–offs to be made at each step, which can affect measured uncertainty in important ways. The LTAM framework is a promising approach for reconciling differences in other studies of uncertainty about long–term Social Security finances.*

INTRODUCTION

Current projections for the Old Age Survivor's and Disability Insurance (OASDI) program suggest a dramatic change in financing fundamentals starting in about ten years. As the baby boom cohort begins to retire around 2010, the current excess of payroll tax revenues over benefits paid is expected to dwindle and eventually turn negative. These projections are, of course, based on a number of demographic and economic assumptions, all of which are uncertain. This paper considers the issue of uncertainty of OASDI projections using the new Long–Term Actuarial Model (LTAM) being developed at the Congressional Budget Office.[1]

Evaluating uncertainty about Social Security projections is not a new idea. The Office of the Chief Actuary (OCACT) at the Social Security Administration (SSA) annually produces both baseline and alternative (low cost and high cost) projections when evaluating OASDI finances. Those alternatives are intended to give some sense of the uncertainty about the long–run forecasts, although the range of outcomes is not associated with any specific probabilistic interpretation. The lack of specific probabilistic interpretations (such as confidence intervals and standard deviations) from OCACT for standard financial measures has led the Social Security Advisory Board (1999) to advocate more effort in this direction,

Noah Meyerson & John Sabelhaus
Congressional Budget Office, Washington, D.C. 20515

National Tax Journal
Vol. LIII, No. 3, Part 1

[1] The LTAM model is being developed by the Long-Term Modeling Group at CBO, which includes Amy Rehder Harris, Michael Simpson, and Joel Smith, and the authors. We are grateful to them and other colleagues at CBO who have supported the LTAM project in general and this paper in particular.

and has also led researchers outside of OCACT (for example, Holmer (1996; 1999; 2000) and Lee and Tuljapurkar (1998a; 1998b)) to develop their own models in order to measure uncertainty about OASDI projections.

The first principle for evaluating alternative approaches to measuring uncertainty is the ability of a given measure to inform policy decisions. Expected values about input assumptions (and model structure) are not sufficient to make good policy; two policies with the same expected values for system finances but different variances for the outcomes are obviously not equivalent. However, variance in OASDI outcomes arises for two reasons—uncertainty about input assumptions and the sensitivity of the OASDI system to changes in those inputs—and those have different implications for policy. Both estimates of uncertainty about inputs and model specification are subject to error, so the overall calculated variance in trust fund projections will reflect both sources of error.

The first (and most widely studied) source of variance in OASDI projections is that which stems directly from uncertainty about the input assumptions. For OASDI, there are several demographic assumptions (mortality improvement, fertility, and immigration) and economic assumptions (inflation, interest rates, wage growth, unemployment, disability incidence, and termination) that go into making a projection, all of which are uncertain. Although historical data exists for these inputs, the extent of uncertainty about the future path for those inputs is still subject to interpretation. Indeed, it can be argued that some approaches to using historical data for measuring how much uncertainty exists about certain input assumptions may be suspect because secular changes have occurred in the underlying process generating that assumption. For example, some might claim (with more certainty than historical data suggest) that fertility rates will probably not double from the levels today to those that existed during the baby boom era.

The second source of variability is from model specification (meaning the OASDI forecasting model, as distinct from the model(s) that generate values for the inputs). When making projections, it seems obvious that starting with the best forecasting model is a good idea. However, a model which reflects everything that is known about a system as complex as OASDI by definition will never get built, and even feasible models (such as OCACT's) are too complex to be solved repeatedly, as in a Monte Carlo simulation. In the real world researchers build simplified models, and the degree of simplification may limit the ability to measure uncertainty because some input assumptions are not incorporated in the stripped–down model. Thus, two different forecasting models may generate two different variance estimates for outcomes even if the variability of input assumptions are the same.

The fact that there are two sources of variance in outcomes complicates estimates of the overall level and relative sources of variance in program projections. Both are important, but the source of uncertainty across the various input assumptions is especially interesting because policy choices may depend on estimated sensitivity. One set of policy rules may have "shock absorber" properties with respect to an input assumption, meaning the long–run financial viability of the system is not much affected by values for that assumption. For example, under current rules, if wage growth rises payroll tax revenues will immediately increase, but benefits paid will also eventually rise and (in the long–run) offset the increased revenues, leaving the overall finances basically the same. There is a small improvement in system finances because of the timing, but that is secondary. Changes in mortality rates do not have this "shock absorber" property under current rules, however, because retirement age is not explicitly linked to expected longevity.

Clearly, there is a direct relationship between the estimate of uncertainty about an input assumption and the degree of outcome sensitivity with respect to that input. If the OASDI system will automatically absorb shocks from a given input, the inability to measure the uncertainty about that input is not important for estimating uncertainty about system finances, though it may be important for measuring outcomes like rates of return to individuals in the system.

If the overall OASDI system finances are sensitive to changes in an input, then the extent of (measured or assumed) variance in that input will determine the variance of the overall system outcomes (with respect to that input). If one analyst infers there is little variance in mortality improvements, while a second analyst finds there is significant variance, and they use an otherwise identical forecasting model for system finances, the second analyst will obviously find a wider variance in program outcomes. This issue could be important because it is easy to imagine policy changes being recommended because of how much they would reduce (estimated) variance in the trust fund projections.

Distinguishing the two sources of uncertainty is a main organizing principle for this paper. The next section discusses the specification of a projection model, with a particular focus on the OCACT techniques as a reference point. The third section shows the results of comparing LTAM input responses with those from OCACT's model. The model specifiction dictates which input assumptions (and/or model characteristics) can be varied, but the issues of how to vary the assumptions in a Monte Carlo simulation is addressed separately in the fourth section. The fifth section discusses an interesting aspect of stochastically-generated forecasts; average outputs from those projections are not necessarily equal to the means from deterministic solutions of the same model. The last section presents estimates of OASDI system variance using the LTAM model and several alternative specifications for variability in the input assumptions.

MODELING LONG-TERM SOCIAL SECURITY FINANCES

The specification of a forecasting model (given values for the input parameters) generally receives limited consideration in studies of uncertainty about long run Social Security finances—most of the attention is devoted to estimating the variance of the input parameters. The premise is that one should develop a forecasting tool which adequately captures how changes in assumptions will affect the OASDI outcomes of interest, then vary the assumptions stochastically to generate ranges for those outcomes. However, the choice of model structure does involve a crucial tradeoff: the ability to accurately measure the sensitivity of system finances with respect to changes in the input parameters versus the ability to build and solve the model quickly under alternative values for those parameters. In this section that tradeoff is described in some detail; the reference point is the OCACT forecast procedure, and the comparison is with other models used for studying uncertainty, including the LTAM model used in this paper.

Given what goes into generating a baseline OASDI forecast, OCACT currently has little choice but to use the scenario approach for measuring uncertainty, because the number of times they can solve their model under alternative assumptions is limited. To some extent it is not even meaningful to speak of running the OCACT "model" over and over again with alternative parameter assumptions—

the OCACT projection methodology involves a sequence of several models which are run by individual analysts, using a variety of different software applications, each suited to the particular model being implemented. Thus, each alternative simulation involves a number of steps for a series of people, not just a series of steps in one linked set of computer code.

The steps that go into generating an OASDI forecast are easy to list, but difficult to describe concisely and with any reasonable level of detail.[2] OCACT first projects population by detailed age, sex, and marital status groups; then projects employment, payroll tax revenues, number of beneficiaries, and average benefits by combining economic assumptions with the underlying demographics; then finally computes trust fund outcomes by summing payroll tax revenues, interest received, other inflows and subtracting benefits paid, administrative costs, and other outflows. For all but the demographic projections, there are also generally distinct short and long run projection methodologies.

There is nothing inherent in the OCACT demographic modeling that rules out stochastic analysis, but it is easy to see why it is complicated to introduce randomness. Incrementing the cells in the population matrix involves subtracting deaths by age and sex, adding births by sex, adding immigrants by age and sex, and distributing the population within each age–sex group by marital status in each forecast year. Each of the four demographic processes is itself a model which builds on the outputs from previous models. For example, fertility rates are set exogenously, but the level of fertility will depend on fertility rates and the female child–bearing population, which depends on the number of female deaths in the child–bearing age groups. Marital status for a given age–sex group depends on the number of (available) people in the age–sex cell and the patterns of marriage across the age distribution. Thus, to make this process stochastic, all of the models have to be linked together.

Some parts of the OCACT projection process would be even more difficult (if not impossible) to make stochastic without significant restructuring. For example, the macroeconomic and labor force/employment sector of the model is actually managed by a time–series software package. The macro model takes the population by age, sex, and marital status as given, and generates projections for the number of covered workers and aggregate earnings for the forecast period. Aggregate earnings feed into a calculation of taxable payroll (based on the taxable maximum OASDI parameter) and those outcomes are fed directly into a model (again, in a different software program than the macro model) which analyzes OASDI program outcomes. The covered–worker outcomes (by age and sex) also flow into the computer program which determines expected benefit awards.

The covered–worker model outcomes affect OAI and DI worker average benefits in the OCACT projection sequence (as they should) because benefits are estimated using a micro–simulation based on the Continuous Wage History Sample (CWHS) longitudinal data set which is calibrated to be consistent with the estimated group–level work histories. The micro–simulation starts with samples of new OASI and DI beneficiaries in 1996 and alters the micro–level data so that the individual work histories are consistent with the overall patterns by age and sex for the relevant cohort, where those group–level employment patterns are taken as given from the macro/labor market model. Benefits are also affected by

[2] For a highly technical and fairly comprehensive description of OCACT projection techniques, see Frees (1999).

average wage growth in the economy, which is also solved for in the macro/labor market model.

The new benefit awards (by age and sex) solved for in the micro–simulation model then feed into other parts of the model. Auxiliary benefit levels for the 25 or so different benefit recipient programs are set relative to the insured OAI and DI worker average benefits. Then, average benefits for both insured and auxiliary beneficiaries are multiplied by the number of beneficiaries (independently solved for in another series of programs) to solve for total benefits paid. The outputs of the various models are pulled together by a final set of programs, which add other components of income and cost (such as administrative expenses and revenue from income taxes on benefits) and then increment trust fund balances.

This highly simplified description of the OCACT projection sequence does not come close to revealing the true complexity of the models. For example, estimates of mortality rates by age and sex are based on quite complicated procedures that begin with estimates of disease–specific death rates and incidence. Also, short and long–term forecasting techniques for each of the sub–models (except demographics) differ, because recent trends are more likely to influence one's view of the short run, while stability is a guiding principle for the long run. In any case, it is easy to see why synthesizing all of these models into one program unit where one can repetitively draw values for the input assumptions and run Monte Carlo simulations would be very difficult.

That difficulty has led other researchers to develop stripped–down models for projecting OASDI finances in order to measure uncertainty (in particular, Holmer (1996; 1999; 2000) and Lee and Tuljapurkar (1998a; 1998b)). The recent models developed by Holmer (1999; 2000) also incorporate macroeconomic feedback effects (in addition to stochastic capabilities for the input assumptions) which is also made possible by simplifying the projection model. The LTAM model used here for measuring uncertainty about OASDI projections is another example of a stripped–down model capable of making stochastic projections, though it is much closer to the OCACT approach than other models.

The various stripped–down models use different approaches to simplifying the projection process. For example, Lee and Tuljapurkar (1998a; 1998b) use their own model of demographics combined with empirical age–sex profiles of average payroll taxes collected and benefits received, which are aged forward through time using an aggregate productivity growth variable (which is stochastic). This limits the capability of the model to do policy simulations, because, for example, benefits are not linked to an underlying micro data set as in the OCACT approach. But the authors are able to conduct some simple policy experiments, like raising the normal retirement age, by making ad hoc adjustments to average payroll taxes and benefits for the age groups that will be affected. This model also limits the number of assumptions which can be varied, because, for example, disability incidence and termination are not modeled explicitly.

The Holmer (1999; 2000) model does allow for stochastic analysis of all the main OCACT input assumptions: Fertility, immigration, mortality improvement, female and male labor force participation, unemployment rate, inflation rate, productivity growth rate, wage share growth rate, hours worked growth rate, nominal interest rate, and disability incidence and recovery rates are all stochastic. The Holmer (1999; 2000) model seems to follow OCACT projection techniques more closely than Lee and Tuljapurkar (1998a; 1998b), but there are still simplifications. For example, OAI worker benefits are based on group–level labor force participation and average wages for cohorts in the years right before they retire, rather than a micro–simulation using longitudinal work histories.

The initial objective for the LTAM project was to build a model which mimics the crucial pieces of the OCACT procedure, so that if an input or policy variable is changed, the model will automatically respond appropriately. LTAM embodies the crucial machinery for projecting population, labor force, payroll tax revenues, and benefit awards (using a micro–simulation) in a way that is consistent with OCACT. Discrepancies between LTAM and OCACT projections are reflected in a series of calibration factors that are applied in all simulations, so baseline projections will match, and alternative simulations (those with different input assumptions or policy parameters) will yield answers that should be consistent with OCACT's.

As in other stripped–down models, LTAM does not attempt to duplicate much of the effort that goes into an OCACT projection. Rather, the model embodies only the long–run projection techniques, and much of the model is based on exogenous ratios derived from OCACT data files. For example, LTAM solves for the ratio of auxiliary (spouse, widow, child) benefits to underlying OAI or DI worker benefits using OCACT benefit projections, then applies those ratios in simulations. Auxiliary benefits will then change if and only if underlying worker benefits (determined by the micro–simulation) change. Other parts of the model work in a similar way; these exogenous ratios will be replaced with explicit forecasting models as LTAM evolves.

Though the initial goal of building LTAM is generating deviations from OCACT baselines under alternative input assumptions and policy parameters, the model architecture was set up to allow subsequent expansion in at least two dimensions. First, the model is written as a self–contained unit in one piece of software, so it is straightforward to introduce Monte Carlo simulation as was done for this paper. Second, the model is solved year by year (rather than sector by sector, as in the OCACT procedure), which will make it feasible to extend the model to include macro feedbacks and even simultaneity across sectors, which could eventually lead to a full general equilibrium approach.

As of this writing LTAM has basic capabilities for stochastic simulation on four input assumptions: mortality improvement, fertility, immigration, and the real interest rate. Although the current model does allow users to change other assumptions that are on the input assumption list (unemployment, inflation, real wage growth, labor force participation, disability incidence, and termination), those inputs all require the benefits micro–simulation to be re–run, which takes about eight minutes, and is thus prohibitive for useful stochastic runs of 1,000 or so simulations. Speeding up the micro–simulation is a priority for near–term LTAM development work.

SENSITIVITY ANALYSIS USING LTAM

Although the stripped–down models use simpler approaches to estimating how deviations in assumptions cause differences in outcomes, there is a well–established set of criteria used to evaluate whether or not these simplified models are adequate for measuring the stochastic properties of the OASDI forecast. The criteria is whether they can (1) duplicate OCACT baseline projections and (2) duplicate OCACT's sensitivity analysis experiments in which one input assumption is changed and the impact on some summary trust fund outcome is evaluated.[3] This

[3] These sensitivity analysis results are published in the annual *Trustees Reports*.

suggests a general consensus that OCACT techniques (setting aside possible disagreement about input assumptions) generates a good reference baseline and that their sensitivity analysis also adequately represents how the OASDI system will respond to given changes in the inputs.

Table 1 compares the effect of changing several input assumptions in both the OCACT and LTAM models. The table shows the 75 year actuarial balance, which is (roughly) the present value of OASDI costs divided by taxable payroll less the present value of OASDI income divided by taxable payroll.[4] The baseline OCACT actuarial balance of –2.07 percent implies that payroll tax rates would have to be raised (or cost rates would have to be cut) by 2.07 percent of payroll to put the system in balance over the 75 year period. The estimate of –2.03 percent from LTAM differs only because of some approximations built into the model.[5]

For each of the inputs in Table 1, actuarial balance estimates are produced for the range of that input used in the low,

TABLE 1
SENSITIVITY OF 75–YEAR ACTUARIAL BALANCE TO CHANGES IN INPUT ASSUMPTIONS: SOCIAL SECURITY ADMINISTRATION (OCACT) AND CONGRESSIONAL BUDGET OFFICE (LTAM) ESTIMATES

		Low Cost	Intermediate Cost	High Cost
Mortality Improvement				
	OCACT	–1.47	–2.07	–2.75
	LTAM	–1.72	–2.03	–2.51
Fertility				
	OCACT	–1.74	–2.07	–2.42
	LTAM	–1.67	–2.03	–2.40
Immigration				
	OCACT	–1.90	–2.07	–2.18
	LTAM	–1.90	–2.03	–2.12
Interest Rate				
	OCACT	–1.59	–2.07	–2.64
	LTAM	–1.53	–2.03	–2.62
Wage Growth				
	OCACT	–1.55	–2.07	–2.57
	LTAM	–1.62	–2.03	–2.46
Inflation (CPI–W)				
	OCACT	–1.84	–2.07	–2.29
	LTAM	–1.85	–2.03	–2.36
Disability Incidence				
	OCACT	–1.77	–2.07	–2.35
	LTAM	–1.36	–2.03	–2.74
Disability Termination				
	OCACT	–2.01	–2.07	–2.12
	LTAM	–1.87	–2.03	–2.11

Source: OCACT values from 1999 *Trustees Report*. LTAM values are author's calculations using Congressional Budget Office Long–Term Actuarial Model (LTAM).

Note: All values are a percent of taxable payroll.

[4] The reason these are "roughly" true is that the actual measure also considers the balance currently in the OASDI trust fund and the actuarial requirement that the fund have one year's worth of outlays left at the end of the valuation period.

[5] Both of these values relate to the 1999 *Trustees Report* estimates. LTAM does not yet have 2000 *Trustees Report* input data, so the OCACT 1999 values were used to keep things comparable. The 2000 actuarial balance is slightly less negative than the 1999 value.

medium, and high cost scenarios produced by OCACT. Most of those ranges are easily reproduced in LTAM—for example, the fertility rate ranges from 1.6 (high cost) to 2.2 (low cost) per woman, and since LTAM uses the same basic fertility model as OCACT, that parameter can be varied directly. The same is true of immigration, the interest rate, wage growth, and inflation. Mortality improvement is somewhat more difficult to reproduce, because the range is for "overall" mortality improvement, yet the rate of improvement in the OCACT alternatives and LTAM is also differentiated by age and sex. To generate the table, a range of mortality improvement for LTAM was chosen such that the average of male and female life expectancy matched the range published by OCACT.

In general the sensitivity of LTAM compares well to the sensitivity of the OCACT model with respect to key input assumptions. The first four inputs will be varied in the Monte Carlo simulation presented below. Within those four, the results for fertility, immigration, and the interest rate are basically identical across the two models, while the results for mortality are close but show some indication that LTAM is not allocating the improvements in mortality across age and sex groups appropriately, which indicates this part of the model requires some attention. In general, the actuarial balance is not as sensitive to changes in life expectancy in LTAM as it is in the OCACT model.

The other four input assumptions in the bottom half of Table 1 are not currently in the stochastic simulations for LTAM, because changing any of these will lead to a change in average benefits, which causes the internal micro–simulation to be invoked. It is straightforward to run a couple of simulations on each input for the purpose of doing sensitivity analysis, but repeated draws at eight minutes per solution are currently infeasible. (Without the micro–simulation the model solves in a little less than one second). Improving the speed of the micro–simulation in order to make these inputs stochastic is an important development area for the near–term.

Table 1 shows that LTAM does a pretty good job capturing the sensitivity of the actuarial balance with respect to changes in the non–stochastic assumptions, though the match is not as close as it was with the demographic variables, especially for disability. This makes sense, because the demographic simulations effectively (through the calibration procedure) take all of the economic variables (for example, payroll taxes per worker, number of workers relative to population, average benefits, and number of beneficiaries relative to population) as given. If simulated changes in population by age and sex match the changes in the OCACT model, the overall system outcomes will match. It is much more difficult to match on inputs like wage growth, because this feeds right into some of the more complicated parts of the model. Wage growth affects the benefits micro–simulation directly, interacts with the taxable maximum in a non–linear way to determine total taxable payroll, and even affects labor force participation of older workers as benefits change.

CHARACTERIZING UNCERTAINTY ABOUT MODEL INPUTS

Given a model which is able to replicate sensitivity with respect to changes in input assumptions, the second step in studying variance of OASDI projections is characterizing uncertainty about those input assumptions. The first decision when characterizing uncertainty is to choose between "scenario" and Monte Carlo approaches to measuring variability in outcomes. There is little doubt that more information can be gleaned using Monte Carlo simulation, but the choice may be constrained by the nature of the projection model. The second decision is whether to specify a single "ultimate" long-run outcome or annual values for each input assumption. Finally, one has to choose ad hoc or empirical methods to measure variability in the inputs; if the empirical approach is taken, the method used to recover estimates of variance (and correlation with other inputs) from historical data is crucial, because different approaches may lead to very different answers about the variance of input assumptions.

In order to give a sense of how far the various OASDI outcomes (summary income and cost rates, actuarial balance) might deviate from the baseline (mid-range) values, OCACT bundles what it considers to be a reasonable range for assumptions into "low cost" and "high cost" scenarios, and thus produces a total of three projections. This scenario approach was critiqued by the 1999 Technical Advisory Panel on four grounds: it assumes that trajectories for the inputs are always high or low (ruling out boom/bust patterns), it assumes correlations between the outcomes are rigid (all assumptions high cost or all assumptions low cost), it is possible that the likelihood a given assumption will fall within the high/low range will be much less than the likelihood the summary measure will fall within the high/low range, and OCACT does not explicitly assign probability distributions for inputs.

The first two critiques are explicitly about how to estimate variability in inputs, and the last two are about why those variance estimates are necessary to assign probabilistic interpretations to outcomes. It is true that one needs input probabilities to assign output probabilities, but it is also true that Monte Carlo simulation (or systematic evaluation of each possible probability-weighted "state" for the inputs) is required (in the multivariate input case) in order to derive variances for the outcomes. That is, even if we know the joint probability distribution of the inputs, evaluating the system with just two points from that joint distribution is not sufficient to draw conclusions about the probability distribution for the outcome being analyzed. It is sort of a moot point that OCACT does not explicitly state what the assumed probability distribution is for each input, because the requisite steps for making inferences about the distribution of outcomes involves repeated solutions to the model that are infeasible.

Given the capability to run Monte Carlo simulations, one has to choose ultimate or annual draws for the inputs and be explicit about correlations between those variables in order to run simulations. The current version of LTAM is designed to be solved in a non-stochastic mode using the same input "levers" as the OCACT model, and thus the assumptions are all modeled as ultimate values (with linear interpolation between the first year and the ultimate value in a fixed year; for example fertility hits its ultimate value in 2023). Thus, there is only one stochastic draw per input assumption per simulation.

The probability distributions for the inputs are also specified in a very ad hoc way. The model assumes that the range for low and high cost values of each input represents some points on a normal distribution for that input. For example, the average of the gap between the low-medium and high-medium range for a variable is taken as one or two standard deviations in the simulation. The correlations for the variables are set to either one (meaning all four move towards low cost or high cost together) or zero (meaning there is no correlation).

The ad hoc specification for the probability distributions is a useful place to start, but it is obvious that empirical investigation of the input assumptions is a crucial area for LTAM development work. There is a substantial literature on measuring variability of input assumptions for

OASDI, including Foster (1994), Sze (1996), Holmer (1996), Frees, et al. (1997), and the previously mentioned studies by Lee and Tuljapurkar (1998a; 1999b) and Holmer (1999; 2000). In general, these analyses are based on time–series modeling of the input processes, and the findings raise a number of issues about the future work planned for LTAM. In particular, one can draw different inferences about the stochastic properties of input assumptions using the same historical data, and that will affect conclusions about overall variability in OASDI outcomes as well as the sensitivity of the baseline with respect to particular inputs.

A good example of this issue arises when modeling fertility. Attempts to fit unconstrained time–series models to fertility rates (see, for example, Lee and Tuljapurkar (1994)) imply unrealistic projections—negative fertility rates in the future in particular—because the downward trend following the baby boom era is so pronounced. Therefore the model has to be constrained in some way. The method chosen by Lee and Tuljapurkar (1994; 1998a; 1998b) is to fix the long–run mean for the fertility rate at the OCACT mid–point (currently 1.9 children per woman) and interpret deviations from that trend as (autocorrelated) variability.

This approach to inferring the variability in fertility rates can certainly be questioned. If one believes that structural changes in marital patterns, labor force activity of women, living arrangements, acceptance of birth control, or other determinants has dramatically reduced the chance that fertility rates will climb back to baby boom levels, then the pure time–series estimate of the variance will be biased upwards. Of course, the only way to test this is to build a model of fertility that includes those structural determinants and then compute the residual variance. This opens a can of worms, however, because a comprehensive variance analysis would then require consideration of uncertainty about the structural determinants and even the uncertainty about the coefficients on those determinants.

The time–series approach can also lead to conclusions which work in the other direction. For example, Lee and Tuljapurkar (1994; 1998a; 1998b) find that a trend–based model of mortality improvement fits the historical data quite well. That model leads them to two conclusions: first, the authors conclude that the overall OCACT rate of mortality improvement in the mid–range assumptions is too low, because it implies a slowdown from trend, and second, the authors (implicitly) conclude the variance of OASDI outcomes with respect to uncertainty about mortality reduction is relatively low, because they find that the overall variance of OASDI outcomes when just mortality is stochastic is relatively low. OCACT, on the other hand, builds their projections of mortality improvement by considering disease–specific death rates, so more than just trend analysis is involved.

Although it is clear that the next step for improving the LTAM stochastic simulation capability is to work with empirical models of input uncertainty, these observations about time–series analysis suggest some caution should be exercised. At a minimum, some sensitivity analysis of the time–series specification is in order. It is also likely (see Holmer (2000) in particular) that for some of the economic input assumptions it is important to model the correlations between the inputs, using either a Vector Auto Regression (VAR) or explicit macro–model framework to draw inferences about variability in the inputs.

STOCHASTIC BIAS

Before presenting the estimates of uncertainty about OASDI finances using LTAM, it is important to address one phenomenon that can occur when a model like this is used in a stochastic setting. It seems natural that the expected outcome of a model should be the same in a stochastic and non–stochastic setting, but that is not the case for some previous studies and (to some extent) for the LTAM results presented here. This section explains why "stochastic bias" will affect the outcomes in these types of simulations.

Consider the following: Y is the OASDI outcome of interest, say the 75 year or annual actuarial balance. X_s is the vector of inputs in state s, π_s is probability that state s will occur, and $f(\,)$ is the model that relates outputs to inputs. Therefore, $Y_s = f(X_s)$ is the outcome for OASDI in state s. When OCACT makes its medium–cost projection, it plugs in the expected value for inputs, $E(X) = \Sigma_s \pi_s X_s$, and solves for the expected outcome using $E(Y) = f(E(X)) = f(\Sigma_s \pi_s X_s)$. In a stochastic setting, however, the expected value of the OASDI outcome is computed using $E(Y) = \Sigma_s \pi_s f(X_s)$, where each π_s is set equal to the inverse of the number of stochastic draws.

The non–stochastic and stochastic means for the OASDI outcome will match if the OADSI projection model (that is, $f(X)$ itself) is probability–weighted symmetric in X, so that equal probability values of X on either side of its expected values cause the same (but opposite signed) change in the outcome of interest. This symmetry occurs, for example, if $f(X)$ is linear with respect to inputs or the probability distributions for the inputs and sensitivity of $f(X)$ around the expected value for X are both symmetric.

The possibility of stochastic bias is raised because Holmer (1999), for example, finds a significant change in the estimated actuarial balance when his model is solved stochastically and deterministically. His estimated 75 year actuarial balance worsens from –2.41 percent to –2.95 percent when the inputs are made stochastic, and the only possible explanations are the phenomenon of asymmetry described above (which Holmer implicitly argues for in his description) and the possibility that some drift in the expected values for the inputs is (inadvertently) introduced through his specification for the input assumption equations (perhaps the correlations between variables). In that case, it is the $E(X)$ that has changed, which is a separate problem.

Table 2 shows that there is a possibility of stochastic bias in LTAM. The effect on the 75 year actuarial balance from varying each of the four input assumptions is compared over the range moving from low to intermediate cost, and from inter-

TABLE 2
CONDITIONAL SENSITIVITY OF 75–YEAR ACTUARIAL BALANCE IN LTAM TO CHANGES IN INPUT ASSUMPTIONS

		Change in Actuarial Balance When All Other Inputs Set To...		
Assumption Varied	Range Varied	Low Cost	Intermediate Cost	High Cost
Mortality	Low->Intermediate	0.26	0.31	0.39
	Intermediate->High	0.39	0.48	0.60
Fertility	Low->Intermediate	0.28	0.36	0.50
	Intermediate->High	0.28	0.37	0.51
Immigration	Low->Intermediate	0.12	0.13	0.17
	Intermediate->High	0.07	0.09	0.10
Interest Rate	Low->Intermediate	0.37	0.50	0.67
	Intermediate->High	0.41	0.59	0.82

Source: Congressional Budget Office Long–Term Actuarial Model (LTAM).

Note: All values are a percent of taxable payroll.

mediate to high. If those are equal, the system response is basically linear. Each of these comparisons is done with all other assumptions set to intermediate cost (which corresponds to differences between the columns in Table 1), low cost, or high cost. The differences in pairs within each column imply some non–linearity; the effect of mortality improvement when all other inputs are set to intermediate is .31 percent of payroll over the low to intermediate cost range, and .48 percent of payroll over the intermediate to high range.

The estimated non–linearities in Table 2 are noticeable, but actually turn out to have only a small impact on the mean estimates of system outcomes presented in the next section. That conclusion is entirely dependent on the assumption about input variance. If the input variance is increased significantly, the non-linearities are more pronounced, and the level of stochastic bias rises.

VARIANCE ESTIMATES FOR OASDI USING LTAM

This section presents estimates of variability in OASDI projections using LTAM. The estimates are generated using very simple assumptions about the variability of inputs. Ultimate values for each of the four inputs are first assumed to be independent and normally distributed, with means equal to the OCACT intermediate values and base case standard deviations equal to the average of the gap between intermediate/low and intermediate/high ranges for the assumption. Two of the variants run involve shrinking the standard deviations so the OCACT scenario range represents two standard deviations, and making the inputs perfectly correlated rather than independent.

Table 3 shows several output statistics for three sets of model solutions: deterministic, and the independent draws (for 1,000 simulations each) where the standard deviations are set to the base case and one–half the base case. The first observation is that there is a small (but statistically significant amount) of stochastic bias in the base case—the mean estimate for the 75 year actuarial balance falls from –2.03 in the deterministic case to –2.05. The other statistics are also affected, especially the measures from the end of the projection period in 2075.

TABLE 3
OASDI OUTCOMES IN LTAM BASED ON 1000 INDEPENDENT MONTE CARLO SIMULATIONS FOR MORTALITY IMPROVEMENT, FERTILITY, IMMIGRATION, AND THE REAL INTEREST RATE

		Stochastic Solution, OCACT Input Range Interpreted As...			
		One Standard Deviation		Two Standard Deviations	
Outcome	Deterministic Solution	Mean	Standard Deviation	Mean	Standard Deviation
75–Year Actuarial Balance	–2.03	–2.05	0.78	–2.03	0.39
75–Year Cost Rate	15.56	15.58	0.78	15.56	0.38
75–Year Income Rate	13.52	13.53	0.07	13.53	0.04
Actuarial Balance in 2030	–4.64	–4.65	0.40	–4.64	0.20
Cost Rate in 2030	17.71	17.72	0.42	17.71	0.21
Actuarial Balance in 2075	–6.53	–6.95	3.47	–6.66	1.67
Cost Rate in 2075	19.88	20.32	3.66	20.00	1.76
Aged–Dependency Ratio in 2030	19.80	19.80	0.80	19.80	0.40
Aged–Dependency Ratio in 2075	22.70	23.10	4.40	22.80	2.20

Source: Congressional Budget Office Long–Term Actuarial Model (LTAM).

Note: All values are a percent of taxable payroll, except aged–dependency ratio, which is the fraction of the population aged 65+.

The base case standard deviation for the 75 year actuarial balance is 0.78, suggesting a 95 percent confidence interval of roughly –0.5 percent to –3.5 percent. Of course this estimate (and all the estimates in Table 2) are entirely dependent on the ad hoc assumption about the size of the standard deviation and the independence of the input assumptions. The last two columns in Table 2 show the impact of cutting the standard deviations in half (so the assumption range in the *Trustees Report* is interpreted as two standard deviations). The overall standard deviations fall by almost exactly half, suggesting the system response is very linear over this range. Note also that tightening the standard deviations effectively eliminates the stochastic bias for the summary measures.

Although the absolute level of uncertainty is entirely dependent on the assumptions about input variability, there are some interesting observations about uncertainty for the various time periods considered. The actuarial balance and cost rates for 2030 are much more certain than they are for 2075. Remember, this version of the model does not have uncertainty about economic assumptions (except the real interest rate) so fertility and mortality are the only significant influences, and both of those take time to cause big changes. But, the same variance assumptions that overwhelmingly suggest the system will be in significant deficit in 2030 (actuarial balance of –4.65 percent of payroll, standard deviation of 0.4) also suggest it is unclear what the system will look like in 2075 (actuarial balance of –6.95 percent of payroll, standard deviation of 3.47).

Table 4 shows some other aspects of how the model responds to variability in the inputs, focusing on just two output statistics but with different assumptions about which variables are stochastic or how they are related (all standard deviations are set to base case values). The two statistics considered are the 75 year actuarial balance, and the actuarial balance in 2075. Again the table shows (on the first line) the deterministic solution and (on the second line) the four independent stochastic variables solution. The next four lines show the effect of varying each input individually—all other inputs are fixed at the intermediate values. The rank order of the variability estimates is generally consistent with OCACT univariate sensitivity analysis, though the problems with mortality improvement in LTAM are evident because of the (relatively) weaker response than the interest rate (the standard deviation for mortality should be a little higher, not lower).

TABLE 4
OASDI OUTCOMES IN LTAM BASED ON 1000 MONTE CARLO SIMULATIONS UNDER ALTERNATIVE SPECFICATIONS FOR INPUT ASSUMPTIONS

	Output Statistic			
	75–Year Actuarial Balance		Actuarial Balance in 2075	
Input Assumptions	Mean	Standard Deviation	Mean	Standard Deviation
Deterministic	–2.03	n.a.	–6.53	n.a.
Independent Stochastic Draws	–2.05	0.78	–6.95	3.47
Varying Only:				
Mortality Improvement	–2.01	0.41	–6.45	1.66
Fertility	–2.05	0.37	–6.99	2.99
Immigration	–2.04	0.11	–6.56	0.26
Real Interest Rate	–2.06	0.52	–6.53	n.a.
Perfectly Correlated Stochastic Draws	–2.30	1.53	–7.56	5.37

Source: Congressional Budget Office Long–Term Actuarial Model (LTAM).

Note: All values are a percent of taxable payroll.

TABLE 5
COMPARISON OF UNIVARIATE AND SCENARIO OUTCOMES IN OCACT MODEL

	Low Cost	Intermediate Cost	High Cost
75–Year Cost Rate			
Full Scenario	13.37	13.49	13.62
Sum of Univariate Sensitivity Analyses	13.37	13.49	13.60
75–Year Income Rate			
Full Scenario	13.14	15.56	18.60
Sum of Univariate Sensitivity Analyses	12.74	15.56	18.43
75–Year Actuarial Balance			
Full Scenario	0.23	–2.07	–4.97
Sum of Univariate Sensitivity Analyses	0.63	–2.07	–4.83

Source: 1999 *Trustees Report*.

Note: All values are a percent of taxable payroll.

The other interesting aspect of the four individual simulations is that the sum of the four individual standard errors (for the 75 year actuarial balance, 1.41) is about twice the overall standard deviation. This is the same basic result as reported by Lee and Tuljapurkar (1998b), and reinforces the idea that only Monte Carlo simulation allows one to construct a cumulative density function for the system outcomes—it cannot be inferred just from the univariate responses.

The last row of Table 4 shows the LTAM solution with perfectly correlated draws for the input assumptions. Two observations jump out from the table. First, the evidence of stochastic bias is quite strong, as the mean 75 year actuarial balance falls from –2.03 (deterministic) to –2.30, and the mean actuarial balance in 2075 falls from –6.53 (deterministic) to –7.56. Second, the estimated standard deviations for the two statistics shown are only slightly larger than the sum of the univariate standard deviations, suggesting that negative and positive draws for all the variables reinforce the overall negative or positive impact on the system finances, but not by much. That possibility is confirmed in Table 5, which shows OCACT's own estimates for 75 year cost rates, income rates, and actuarial balances from their scenario (perfectly correlated errors) and sum of univariate analyses. OCACT also finds that the system is fairly "additive" across the inputs, because the sum of the effects on the 75 year actuarial balance estimate in the low or high cost univariate simulations roughly correspond to the outcomes in the corresponding scenarios. (Also, there is some variation in the scenario, which is not reflected in the univariate analysis, so the actual values are even closer).

CONCLUSIONS

The LTAM project at the Congressional Budget Office is focused on developing a capability to study expected long–term OASDI finances with respect to the effects of both baseline assumptions and policy parameters. In this paper the model was used to analyze the variance in long–term projections by making simple assumptions about the variability of four model inputs: mortality improvement, fertility, immigration, and the real interest rate. The results show that LTAM generates results that are consistent with OCACT model outputs, and that adding uncertainty has predictable and stable effects on the model results.

This paper represents the first efforts at making LTAM stochastic. The next step for the model is to make the remaining input variables (inflation, wage growth, disability incidence and termination, unemployment) stochastic. As noted, that will involve speeding up parts of the solution algorithm, particularly the micro–simulation component, but it is computationally feasible. The larger task is to decide how the variability of all the inputs should be modeled, which involves the use of time–series techniques to recover the variance estimates from historical data. As argued in the paper, however, this step should be taken with a good deal of caution, because historical data may have only limited capacity for predicting future uncertainty.

Acknowledgments

The views expressed here are those of the authors and should not be interpreted as those of the Congressional Budget Office.

REFERENCES

Board of Trustees, Federal Old–Age and Survivors Insurance Disability Insurance Trust Funds.
1999 Annual Report of the Board of Trustees of the Federal Old–Age and Survivors Insurance Disability Insurance Trust Funds. (House Document 106–48). Washington, D.C.: Government Printing Office, March, 1999.

Frees, Edward W. (Jed).
"Summary of Social Security Administration Projections of the OASDI System." Social Security Advisory Panel Working Paper, December, 1999.

Frees, Edward W. (Jed), Yueh–Chuan Kung, Marjorie A. Rosenberg, Virginia R. Young, and Siu–Wai Lai.
"Forecasting Social Security Actuarial Assumptions." *North American Actuarial Journal* 1 No. 4 (October, 1997): 49–84.

Foster, Richard S.
"A Stochastic Evaluation of the Short–Range Economic Assumptions in the 1994 OASDI Trustees Report." Social Security Administration Actuarial Study No. 109. Baltimore: MD: SSA Office of the Actuary, August, 1994.

Holmer, Martin R.
"Demographic Results from SSASIM, a Long–Run Stochastic Simulation Model of Social Security." Appendix A in *Report of the 1994–1995 Advisory Council on Social Security, Volume II.* Washington, D.C.: U.S. Government Printing Office, 1996.

Holmer, Martin R.
"Stochastic Simulation of Economic Growth Effects of Social Security Reform." In *Prospects for Social Security Reform,* edited by Olivia S. Mitchell. Philadelphia: University of Pennsylvania Press for the Pension Research Council, 1999.

Holmer, Martin R.
Introductory Guide to SSASIM. Washington, D.C.: Policy Simulation Group, March, 2000. Available at www.polsim.com.

Lee, Ronald D., and Shripad Tuljapurkar.
"Stochastic Population Projections for the United States: Beyond Low, Medium, High." *Journal of the American Statistical Association* 89 No. 428 (December, 1994): 1175–89.

Lee, Ronald D., and Shripad Tuljapurkar.
"Uncertain Demographic Futures and Social Security Finances." *American Economic Review* 88 No. 2 (May, 1998a): 237–41.

Lee, Ronald D., and Shripad Tuljapurkar.
"Stochastic Forecasts of Social Security." In *Frontiers in the Economics of Aging,* edited by David A. Wise. Chicago: University of Chicago Press, 1998b.

Social Security Advisory Board.
The 1999 Technical Panel on Assumptions and Methods. Washington, D.C.: Social Security Advisory Board, November, 1999.

Sze, Michael.
"Stochastic Simulation of the Financial Status of Social Security Trust Funds in the Next 75 Years." Appendix B in *Report of the 1994–1995 Advisory Council on Social Security, Volume II.* Washington, D.C.: U.S. Government Printing Office, 1996.