

CONGRESS OF THE UNITED STATES
CONGRESSIONAL BUDGET OFFICE

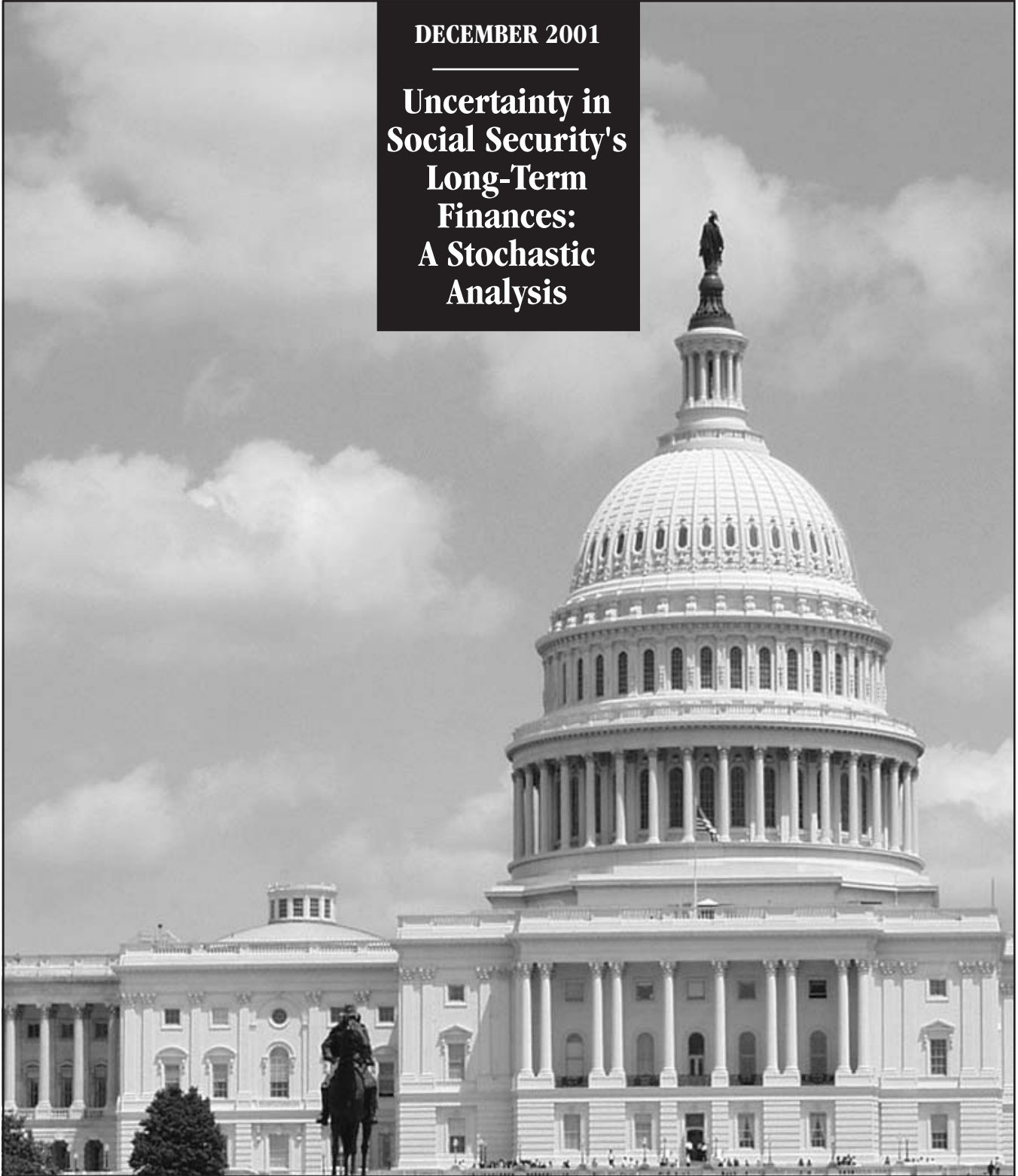
A

CBO

PAPER

DECEMBER 2001

**Uncertainty in
Social Security's
Long-Term
Finances:
A Stochastic
Analysis**



UNCERTAINTY IN SOCIAL SECURITY'S
LONG-TERM FINANCES:
A STOCHASTIC ANALYSIS

December 2001

Congress of the United States
Congressional Budget Office

PREFACE

Consideration of proposals to reform Social Security must take into account the uncertainty in any forecast of Social Security's finances—especially over the 75-year time frame used by the program's trustees. This Congressional Budget Office (CBO) paper provides an overview of that uncertainty. It discusses how the Social Security Administration (SSA) projects the program's finances and illustrates their uncertainty. Then, using time-series analysis of historical data and CBO's new Long-Term Actuarial Model, the paper provides quantitative estimates of the program's uncertainty. Those estimates include ranges of probability for the economic and demographic variables that underlie SSA's projections as well as for Social Security's finances over 75 years.

Noah Meyerson, John Sabelhaus, Michael Simpson, and Joel Smith of CBO's Long-Term Modeling Group (LTMG) wrote the paper. Along with Amy Rehder Harris and Josh O'Harra of LTMG, they also developed the Long-Term Actuarial Model. Paul Burnham, Bob Dennis, Doug Hamilton, Arlene Holen, Steve Lieberman, Deborah Lucas, and Ralph Smith of CBO reviewed the paper and provided helpful comments, as did members of CBO's Long-Term Modeling Advisory Group.

Many analysts in SSA's Office of the Chief Actuary gave a great deal of time to help CBO understand the agency's projection techniques. This project would not have been possible without their assistance.

Joseph Foote and Christian Spoor edited the paper, and Christine Bogusz proof-read it. Kathryn Winstead produced the cover, Lenny Skutnik produced the printed copies, and Annette Kalicki prepared the electronic versions for CBO's Web site (www.cbo.gov).



Dan L. Crippen
Director

December 2001

CONTENTS

	SUMMARY	ix
I	INTRODUCTION	1
	Goals for Studying Trust Fund Uncertainty	2
	How Assumptions About Inputs Determine Social Security's Projected Finances	4
	What Types of Uncertainty Are Not Being Captured?	9
II	SOCIAL SECURITY'S INTERMEDIATE PROJECTIONS	11
	Demographic Inputs and Population Projections	11
	Economic Inputs	14
	Projecting Social Security's Finances Using the Long-Term Actuarial Model	22
III	EVALUATING UNCERTAINTY USING HIGH-COST AND LOW-COST SCENARIOS	25
	SSA's Ranges of Values for Input Assumptions	25
	How Varying the Input Assumptions Affects Social Security's Projected Finances	28
	Concerns About the Scenario-Based Approach	34
IV	MEASURING UNCERTAINTY ABOUT INPUT ASSUMPTIONS	35
	Producing a Forecast with Time-Series and Monte Carlo Techniques	36
	Probability Distributions for Demographic and Economic Inputs	41

V	UNCERTAINTY ABOUT SOCIAL SECURITY'S LONG-TERM FINANCES	55
	Base-Case Estimates of Trust Fund Uncertainty	56
	Sources of Trust Fund Variability	62
	The Effect of Random Changes in the Central Tendency for Fertility	64
	The Effect of Using Bootstrapped Random Shocks	66
APPENDIX	ESTIMATES OF TIME-SERIES EQUATIONS FOR INPUT ASSUMPTIONS	69

TABLES

S-1.	Estimated Uncertainty About Trust Fund Outcomes	xii
1.	The Social Security Administration's Ranges of Ultimate Values for Nine Primary Inputs	26
2.	Sensitivity of Social Security's 75-Year Actuarial Balance to Changes in Ultimate Input Values	30
3.	Sensitivity of Social Security's Annual Actuarial Balances to Changes in Ultimate Input Values	32
4.	Ranges of Uncertainty for Inputs	43
5.	Estimated Uncertainty About Trust Fund Outcomes in CBO's Base Case	58
6.	Sources of Uncertainty About Social Security's Finances	62
7.	Estimated Uncertainty About Trust Fund Outcomes Assuming Random Changes in the Central Tendency for Fertility	65
8.	Estimated Uncertainty About Trust Fund Outcomes Using Bootstrapped Random Shocks	67
A-1.	Estimated Coefficients for Mortality Reductions	70
A-2.	Estimated Coefficients for CBO's Economic Model	72

FIGURES

S-1.	Probability Distribution of the Balance of the Social Security Trust Funds	xiv
1.	How Inputs Affect the Balance of the Social Security Trust Funds in CBO's Long-Term Actuarial Model	5
2.	Intermediate Projections for the Overall Rate of Fertility	12
3.	Intermediate Projections for the Overall Rate of Mortality Improvement	13

4.	Intermediate Projections for the Level of Immigration	15
5.	Intermediate Projections for the Rate of Real Wage Growth	16
6.	Intermediate Projections for the Inflation Rate	17
7.	Intermediate Projections for the Unemployment Rate	18
8.	Intermediate Projections for the Real Interest Rate on Assets in the Social Security Trust Funds	19
9.	Intermediate Projections for the Rate of Disability Incidence	20
10.	Intermediate Projections for the Rate of Disability Termination	21
11.	Uncertainty Bands for the Rate of Real Wage Growth	44
12.	Uncertainty Bands for the Overall Rate of Mortality Improvement	45
13.	Uncertainty Bands for the Unemployment Rate	47
14.	Uncertainty Bands for the Inflation Rate	48
15.	Uncertainty Bands for the Real Interest Rate on Assets in the Social Security Trust Funds	49
16.	Uncertainty Bands for the Level of Immigration	50
17.	Uncertainty Bands for the Rate of Disability Incidence	51
18.	Uncertainty Bands for the Rate of Disability Termination	52
19.	Uncertainty Bands for the Overall Rate of Fertility	53
20.	Uncertainty Bands for the Overall Rate of Fertility Using the First-Differenced Specification	54
21.	Probability Distribution of the Balance of the Social Security Trust Funds	61
22.	Probability Distribution of the Balance of the Social Security Trust Funds Assuming Random Changes in Central Tendency for Fertility	66

SUMMARY

Currently, the government collects more in revenues from Social Security taxes than it pays in Social Security benefits. But projections by the Social Security Administration (SSA) imply that benefits will outpace revenues within a few decades and that the gap between the two will widen each year after that. Those projections, however, depend on assumptions about demographic and economic trends, which serve as inputs to models of Social Security. Because the input assumptions are uncertain, the projections of Social Security's finances are also uncertain.

A question of interest to both policymakers and modelers is, Just how uncertain are those projections? In this analysis, the Congressional Budget Office (CBO) uses its new Long-Term Actuarial Model to construct probability distributions—bands of uncertainty—around SSA's point estimates for the next 75 years. The analysis indicates that although the gap between Social Security's annual spending and revenues is projected to grow steadily through 2075, that growth is overshadowed by an increase in the uncertainty of the projections. As a result, the only outcome in 2075 that can be predicted with confidence is that a gap between spending and revenues will exist; whether the gap will be small or immense is impossible to say with any certainty.

CBO's analysis also suggests that there is a 90 percent chance that the Social Security trust funds will remain solvent through 2029—but only a 10 percent chance that they will still show a positive balance by 2054. SSA's most optimistic scenario, by contrast, shows the trust funds remaining solvent through 2075. However, CBO's analysis suggests that the probability of that happening is just 1 percent.

Although projections of demographic and economic inputs will always be uncertain, Social Security's rules could be changed in ways that would make the program less sensitive to variation in economic and demographic trends. As the Congress examines various proposals to alter Social Security, it needs to take into account not only the expected effects of the proposals but also how they would change the stability of the program. This analysis does not examine specific proposals, but it provides useful background by estimating the uncertainty that exists under current law.

HOW SSA PROJECTS SOCIAL SECURITY'S LONG-TERM FINANCES

In making its projections, SSA considers nine key inputs. Three are demographic: the fertility rate, the rate of mortality improvement (the annual change in the mortality

rate), and the level of immigration. The other inputs are economic and behavioral: the growth of real (inflation-adjusted) wages, inflation, the unemployment rate, the real interest rate on assets in the Social Security trust funds, and the rates at which people join and leave the rolls for federal Disability Insurance.

SSA projects values for those nine inputs by examining historical data and making judgments about the relevance of past data to future trends. SSA uses its “best estimates” for its baseline, or intermediate-cost, scenario. The agency also illustrates the possible range of future outcomes by making high-cost (pessimistic) and low-cost (optimistic) assumptions.

That scenario-based method offers a general sense of the range of possibilities, but it has several limitations. First, the high- and low-cost estimates are not generated with a statistical model or associated with any probabilistic interpretation; they are simply intended, according to SSA, to be “illustrative.” Second, all three scenarios unrealistically assume that each input will have the same value in every year (after phase-in periods of five to 25 years), ignoring annual or cyclical changes. Finally, in the high- and low-cost scenarios, all inputs vary in a single direction—either all better or all worse than in the intermediate scenario. It would be more realistic to assume that some inputs will turn out better than expected and others will turn out worse.

CBO'S ANALYSIS OF UNCERTAINTY

CBO's Long-Term Actuarial Model (LTAM) allows a more comprehensive analysis of uncertainty than SSA's scenario-based approach does. Although LTAM does not always operate at the same level of detail as SSA's model, it was written to follow the same methodology, and tests show that its estimates are consistent with those produced by SSA. Unlike SSA's model, however, LTAM consists of a single, integrated computer program. Therefore, LTAM can be run repeatedly under a range of assumptions about future events to obtain a full distribution of possible outcomes.

Projecting Uncertainty About Inputs

To make quantitative estimates of the uncertainty of Social Security's finances, analysts need to estimate the uncertainty of the major input assumptions. In this analysis, CBO begins with SSA's expected values for those inputs over 75 years and projects uncertainty bands around them (without proposing alternative expected values).

Standard statistical techniques—time-series modeling—allow economists to quantify future uncertainty on the basis of historical data. Once analysts have chosen appropriate equations to describe the inputs, annual changes in the value of an input can be expressed as a combination of several factors: random shocks, a systematic

relationship to the input's own past values, and, in some cases, a systematic relationship to other inputs. (Inflation, unemployment, and the real interest rate on Social Security assets are clearly related in the historical data, so CBO modeled those three inputs together.) Although there is no single way to apply time-series techniques correctly, in most cases the appropriate equation is readily apparent. The estimated equations can then be used to generate probability distributions (the width of the uncertainty bands) for future values.

The estimates of uncertainty presented in this paper should be considered lower bounds. Besides the nine main inputs, assumptions about other factors contribute to the uncertainty of Social Security projections. Those other factors include labor force participation and retirement patterns as well as family structure (and thus the level of Social Security benefits received by the children and spouses of workers). Those assumptions, however, are not considered in this analysis.

Projecting Uncertainty About Long-Term Finances

Once CBO computed probability distributions for the inputs, it used Monte Carlo simulation (also called stochastic simulation) to produce probability distributions for various measures of Social Security's finances. In the Monte Carlo simulation, annual values for each input were chosen randomly from the estimated distributions and fed into LTAM. For this analysis, CBO created 1,000 different sets of input projections. The 1,000 simulated paths were examined to draw inferences about the probability distributions of future outcomes.

Analysts commonly use two sets of statistics to assess the long-term financial prospects for Social Security. The first set focuses on the relationship between the program's costs and income in any given year. The second set summarizes the expected adequacy of trust fund balances over a specific period. (For example, the 75-year summary actuarial balance that is often reported basically indicates how far the system is from having a positive balance in the trust funds at the end of 75 years.)

Projections for both types of statistics are sobering. In SSA's intermediate projections, Social Security's costs will begin to exceed its dedicated income (excluding interest on trust fund assets) around 2015, and that gap (measured as a percentage of taxable payroll) will grow to 6.2 percent by 2075. For the summary statistics, SSA predicts that the trust funds will run out of money around 2037 and that the 75-year actuarial deficit is -1.89 percent of taxable payroll.¹ That latter number implies that

1. This analysis is based on Social Security Administration, *The 2000 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds* (March 30, 2000). The 2001 version of that report projects a 75-year actuarial deficit of -1.86 percent and an insolvency date of 2038 but makes no changes to the long-term input assumptions.

SUMMARY TABLE 1. ESTIMATED UNCERTAINTY ABOUT TRUST FUND OUTCOMES

	SSA's Expected Value (Using deterministic model) ^b	CBO's Estimates ^a		
		Expected Value (Using stochastic model)	Standard Deviation	90 Percent Range of Uncertainty
75-Year Summary Measures				
Actuarial Balance	-1.89	-2.18	1.16	-4.17 to -0.34
Cost Rate	15.40	15.67	1.29	13.62 to 17.86
Income Rate	13.51	13.49	0.16	13.24 to 13.79
Annual Measures for 2030				
Actuarial Balance	-4.26	-4.78	2.54	-9.49 to -1.00
Cost Rate	17.35	17.85	2.65	13.90 to 22.86
Income Rate	13.08	13.07	0.11	12.91 to 13.27
Annual Measures for 2075				
Actuarial Balance	-6.18	-8.20	5.26	-18.38 to -0.97
Cost Rate	19.53	21.62	5.55	14.02 to 32.43
Income Rate	13.34	13.42	0.28	13.04 to 13.98

SOURCES: Social Security Administration; Congressional Budget Office.

NOTES: The income rate is the percentage of taxable wages paid into the Social Security system. The cost rate is the percentage of taxable wages paid to beneficiaries. The actuarial balance is the difference between the income rate and the cost rate.

The 75-year summary measures incorporate interest on assets in the Social Security trust funds, which is excluded from the annual measures.

a. Based on 1,000 Monte Carlo simulations using CBO's Long-Term Actuarial Model.

b. Based on SSA's intermediate assumptions.

achieving solvency through 2075 will require an immediate tax increase equal to 1.89 percent of taxable payroll (say, from the current payroll tax rate of 12.4 percent to 14.29 percent), an equivalent cut in Social Security benefits, or some combination of the two. (Under SSA's high-cost assumptions, the 75-year summary actuarial deficit is -5.00 percent; under its low-cost assumptions, SSA projects an actuarial surplus of 0.38 percent.)

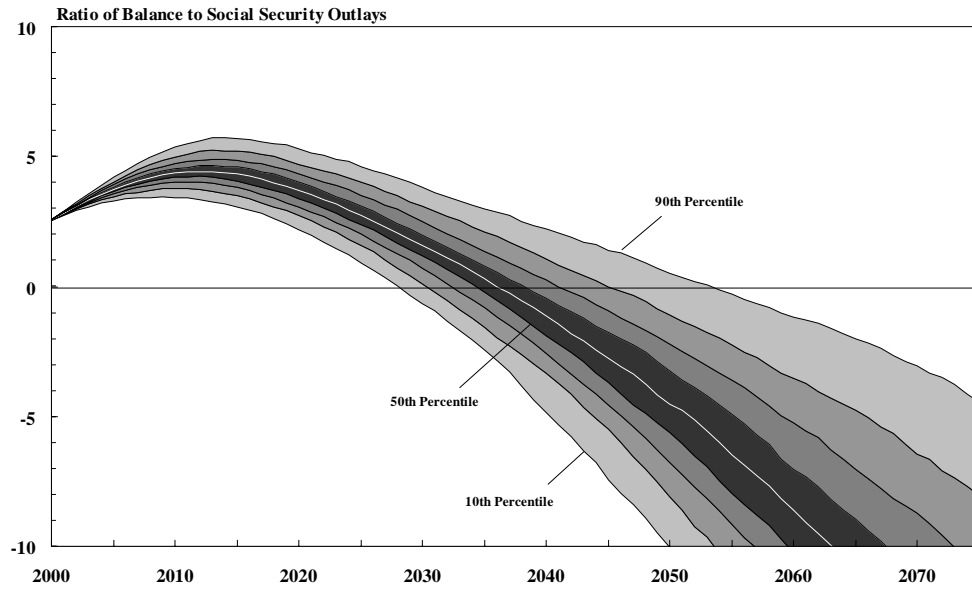
CBO's stochastic model provides a fuller picture of the uncertainty of those statistics. Although relatively little uncertainty is apparent in the near term, the variation becomes huge by the end of the projection period. In the intermediate case, yearly costs are expected to exceed income by 4.26 percent of taxable payroll in 2030 and by 6.18 percent in 2075. But that growth in the gap between annual costs and income is dwarfed by the growth in uncertainty. In 2030, the 90 percent range of uncertainty is -9.49 to -1.00 percent (meaning that there is a 90 percent chance that the actual value will fall within that range), whereas in 2075, the range is more than twice as wide: -18.38 to -0.97 percent (see Summary Table 1). After 75 years, the crystal ball becomes quite cloudy. The most that can be said with any confidence about 2075 is that a gap between costs and income is virtually inevitable, although it is anyone's guess whether that gap will be small or very large.

Less uncertainty surrounds the 75-year summary actuarial balance; in 90 percent of the simulated cases, it falls between -0.34 and -4.17 percent. Future program costs are the chief source of uncertainty about the actuarial balance. The 75-year summary cost rate (Social Security benefits as a percentage of taxable payroll) has a 90 percent probability of ranging between 13.62 and 17.86 percent. In contrast, the 75-year income rate (Social Security revenues as a percentage of taxable payroll) has a relatively narrow range of uncertainty: 13.24 to 13.79 percent. Because the program's payroll tax rates are fixed, the only variation in the income rate arises from the taxation of Social Security benefits, which provides a small portion of revenues.

A final measure is the ratio of trust fund assets to Social Security's annual spending. In CBO's base case, the probability distribution of that ratio widens substantially over time, as shown in Summary Figure 1. That figure also illustrates the range of uncertainty about the date of trust fund insolvency. CBO estimates that there is a 90 percent chance that the trust funds will be solvent at least through 2029 (in Summary Figure 1, the year in which the 10th percentile path crosses the zero line). At the other extreme, the trust funds have only a 10 percent chance of staying solvent through 2054 (when the 90th percentile path crosses zero).

In SSA's low-cost scenario, the trust funds remain solvent throughout the 75-year projection period, suggesting that long-term solvency is a plausible outcome. CBO's stochastic analysis, in contrast, shows that the trust funds face a 99 percent probability of exhaustion before the end of the 75-year period.

SUMMARY FIGURE 1. PROBABILITY DISTRIBUTION OF THE BALANCE OF THE SOCIAL SECURITY TRUST FUNDS



SOURCE: Congressional Budget Office.

NOTE: In each year, there is a 10 percent chance that the balance will fall within each shaded band, a 10 percent chance that it will fall below the lowest band, and a 10 percent chance that it will fall above the highest band.

CHAPTER I

INTRODUCTION

In considering the actuarial aspects of operations in the longer future, the unreliability of any specific estimates is of such degree that only the general course of financial development of the program may be indicated.

—First report of the Board of Trustees of the Federal Old-Age and Survivors Insurance Trust Fund, January 3, 1941

The Social Security Administration (SSA) projects that Social Security’s Old-Age and Survivors Insurance and Disability Insurance Trust Funds will be exhausted within a few decades. After that, it predicts, a permanent imbalance will exist between the benefits paid by Social Security and the revenues collected for the program.¹ That projection is based on assumptions about long-term values for the demographic and economic variables (mortality improvement, fertility, immigration, inflation, unemployment, real wage growth, interest rates, and disability patterns) that affect Social Security revenues and benefits. Because a projection of the Social Security system’s finances is based on assumptions, it is by nature uncertain.

The Congressional Budget Office (CBO) is one of several organizations that study the financial uncertainty facing Social Security. SSA publishes “high-cost” and “low-cost” scenarios as part of its annual 75-year projections of the program’s finances; other analysts have created Social Security models and used them to make stochastic projections.² CBO has developed a model—called the Long-Term Actuarial Model—that employs a unique combination of modeling detail and an easily managed programming format. That combination allows users to vary numerous inputs to the model and also assign probabilistic interpretations to the range of possible outcomes for the balance of the Social Security trust funds. The model uses inferred uncertainty

-
1. Social Security Administration, *The 2000 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds* (March 30, 2000).
 2. See, for example, Ronald Lee and Shripad Tuljapurkar, “Stochastic Forecasts for Social Security,” in David Wise, ed., *Frontiers in the Economics of Aging* (Chicago: University of Chicago Press, 1998), pp. 393-420; and Craig Copeland, Jack VanDerhei, and Dallas L. Salisbury, *Social Security Reform: Evaluating Current Proposals*, Issue Brief No. 210 (Washington, D.C.: Employee Benefit Research Institute, June 1999). In a stochastic model, every input variable is assigned a range of possible values; the model is run many times (each time drawing values for the inputs from their respective ranges) and yields a set of results with varying probabilities. In a deterministic model, such as the one used by SSA, each input is assigned a single value for each year, and the model produces a single result.

about each input assumption (based on historical data) to compute the implied uncertainty about Social Security's future finances.

GOALS FOR STUDYING TRUST FUND UNCERTAINTY

Each year, SSA reports its long-term estimates of the actuarial status of the Social Security program to the Congress. Those estimates allow policymakers to anticipate imbalances or other structural difficulties with the program and to respond far in advance. However, policy choices should take into account not only expected effects but also how those choices might alter the stability of the Social Security system.

Why Consider Uncertainty?

An important measure of system finances is whether Social Security is in "close actuarial balance"—meaning that the program's revenues will be large enough to pay expected benefits over a 75-year period. According to SSA's baseline (or intermediate) assumptions, that condition is not currently being met.³ Equally important, however, is an understanding of the uncertainty of such projections.

In general, a system with more certainty is preferable because it allows plans to be set in advance. Simply put, uncertainty carries a cost. To make fully informed decisions about Social Security, policymakers should know both the expected path of various indicators—such as the date when the trust funds will be exhausted—and the best available probability distribution of those indicators.

Consider a scenario in which the trust funds are projected to have a positive balance for 50 years but also a 25 percent chance that the balance will turn negative within two decades. Such a system would require more-urgent action than one in which a positive balance was expected to last for just 45 years but there was only a 5 percent chance that the balance would turn negative within 20 years.

Uncertainty and System Structure

Uncertainty about the finances of the Social Security system comes from two factors: uncertainty about inputs and the system's sensitivity to changes in those inputs. The first factor is relatively intractable; despite advances in economic forecasting, accurately predicting national macroeconomic and demographic trends over the long run is impossible. However, the second factor—system sensitivity—can be changed.

3. Social Security Administration, *The 2000 Annual Report of the Board of Trustees*, Section II.F.

Indexing the parameters of the Social Security system to uncertain inputs (so that the system changes along with those inputs) can increase stability. For example, the formula used to compute initial benefits is indexed to the nominal growth of wages, and recipients' current benefits are indexed to inflation. That indexing greatly reduces the system's sensitivity to changes in wage growth and inflation.⁴

The Social Security system is not indexed to demographic characteristics, however. As a result, an unexpectedly low rate of mortality or fertility will increase the system's cost rate (the ratio of benefits to taxable wages). Many analysts have proposed indexing Social Security's retirement age to mortality rates as a way to reduce benefit levels and the system's sensitivity to changes in mortality.

Redistribution of Risk

As analysts continually remind policymakers, the balance of the trust funds should not be the only indicator of the condition of Social Security, nor should effects on the trust funds be the only measure for evaluating a policy proposal. Any analysis of proposals to change Social Security must consider the expected benefits, costs, and risks borne by workers, beneficiaries, and other parts of the federal government. For example, although investing trust fund assets in corporate stocks might increase expected returns, it would also increase the riskiness of the system.⁵ In that case, the cost of the increased risk, as valued by financial markets, would exactly equal the expected increase in returns. Thus, the risk-adjusted value of the returns would not change. Individual investment accounts might increase that risk even more, depending on how they were structured, and could also redistribute risk among age and income groups. Finally, as with any program administered by the government, the additional element of political risk (that lawmakers will alter the program) is always present. This paper does not directly address the risk assumed by various parties, but it should be read with that broader context in mind.

Policymakers could eliminate uncertainty about the actuarial balance of the trust funds in a number of simple ways. For instance, after the trust funds were exhausted, the system could move to a pure pay-as-you-go approach—in which revenues equaled outlays each year—by annually cutting benefits, raising taxes, or transferring money from the Treasury. But although those policies would eliminate uncertainty about the actuarial balance, they would also crudely shift risk onto beneficiaries or workers.

4. The Social Security system remains somewhat sensitive to changes in wage growth and inflation because of timing lags between when those changes occur and when they are reflected in recipients' benefits.

5. See Amy Rehder Harris, Noah Meyerson, and Joel Smith, "Social Insecurity? The Effects of Equity Investments on Social Security Finances," *National Tax Journal* (September 2001).

Such restructuring of the Social Security system cannot eliminate or even reduce total financial risk. It can only shift that risk, as private insurance systems do. How the risk is redistributed is important, however. The shift should be made in such a way that negative and positive uncertain outcomes are balanced (in other words, risks should be negatively correlated). That result would be one major advantage of indexing the retirement age to mortality—the reduction in benefits would be correlated with increases in life expectancy and corresponding increases in workers' earning capacity.

HOW ASSUMPTIONS ABOUT INPUTS DETERMINE SOCIAL SECURITY'S PROJECTED FINANCES

Projecting Social Security's finances decades into the future requires two things: a model that shows how key economic and demographic variables interact with policy rules to determine financial flows, and assumptions about the annual values for those variables. The structure of the model determines how changes in assumptions affect the system's finances and thus how uncertainty about those assumptions results in uncertainty about the financial status of Social Security.

The Overall Structure of CBO's Model

CBO's Long-Term Actuarial Model (LTAM) relies on assumptions about nine primary inputs. Eight of them—rates of fertility, mortality, immigration, unemployment, incidence and termination of claims for Disability Insurance (DI), real wage growth, and inflation—affect intermediate demographic and economic variables that determine the accumulation of money in the Social Security trust funds (see Figure 1). The ninth primary input—the interest rate on Social Security assets—affects the trust funds directly. The final output of the model is the annual change in the trust fund balance, based on a simple accounting formula:

$$\begin{aligned} \textit{Trust fund balance this year} = & \textit{trust fund balance last year} + \\ & \textit{interest earned on trust fund assets} + \\ & \textit{Social Security payroll taxes and other revenues} - \\ & \textit{benefits paid and other outlays} \end{aligned}$$

The higher the interest rate, the faster the trust funds grow (assuming a positive balance).

Social Security revenues and outlays equal numbers of people (workers or beneficiaries) multiplied by dollar amounts (average Social Security taxes paid or average benefits). Numbers of people, in turn, are based on population totals, and dollar amounts are based on earnings. Those relationships are explained below.

Projecting Population by Age, Sex, and Marital Status

The equation for projecting total U.S. population is:

$$\begin{aligned} \text{Current-year population} = & \text{population last year} + \\ & \text{current-year births} - \\ & \text{current-year deaths} + \\ & \text{current-year immigration} \end{aligned}$$

The details of the calculations are complex. LTAM begins with a huge matrix that includes counts of people by age, sex, and marital status. Once the modeler has selected future annual values for the mortality and fertility rates and the level of immigration, the model applies the mortality rate to the current population to compute the number of deaths by age and sex; it also applies the fertility rate to the female population to determine the number of births by age of the mother. Those figures, along with the assumed net number of immigrants, are added to last year's population to obtain the new population figure (with a new age and sex distribution). After that, the model distributes the new population among four marital-status groups—single, married, divorced, and widowed—according to age and sex. That process is repeated for each year's projection.

Number of Workers. The model applies three factors to the total working-age population (in this analysis, people ages 15 to 74) to obtain a projection of the total number of workers covered by Social Security:

$$1) \text{ Total working-age population} \times \text{labor force participation rate} = \text{labor force}$$

$$2) \text{ Labor force} \times \text{employment rate} = \text{workers}$$

$$3) \text{ Workers} \times \text{covered-worker rate} = \text{covered workers}$$

Labor Force Participation Rate. Over the past 50 years, the labor force participation rate—the fraction of the working-age population that is employed or looking for employment—has changed significantly.⁶ The proportion of women in the labor force has increased steadily and substantially, though the participation rate for women age 65 or older has stayed the same. The rate for older men is substantially lower now than in 1950, but it has remained level since about 1985.

Although those facts inform projections, many questions remain. For instance, how much further will the labor force participation rate for women increase? Will the

6. Howard Fullerton Jr., "Labor Force Participation: 75 Years of Change, 1950-98 and 1998-2025," *Monthly Labor Review* (December 1999).

long-term trend toward earlier retirement for men continue, or will the rate observed over the past 15 years endure? For the most part, SSA's projections assume that participation rates for various demographic groups will remain at current levels. The 1999 Technical Panel of the Social Security Advisory Board suggested that SSA assume additional increases in women's participation in the workforce and investigate further the effect that changes in pension systems (including Social Security) will have on retirement patterns.⁷ Participation in the labor force is an important source of uncertainty. Still, it is not included as one of the nine input variables in CBO's model, and LTAM does not, at this stage of its development, permit variations of that factor. (Including labor force participation in the quantitative analysis in this paper would be ideal, but it is not possible at this time.)

Currently, LTAM follows SSA's methodology of estimating separate labor force participation rates for 103 age/sex/marital-status groups. Those rates depend on economic factors such as past rates of participation in Social Security, benefit levels, and unemployment rates as well as on social factors such as disability rates, military enrollment, and number of children.

Employment Rate. The employment rate is equal to 1 minus the unemployment rate. The unemployment rate is one of the nine major input variables in CBO's model and is described in more detail in Chapter II.

Covered-Worker Rate. The covered-worker rate is the percentage of employees who work in employment covered by Social Security. Although that rate will increase slightly as older government workers—who are less likely to be covered—retire, it is projected to remain relatively stable over the next 75 years.

Number of Beneficiaries. The Social Security program has about two dozen categories of beneficiaries, including retired workers, disabled workers, widows and widowers, and children. Some of the larger categories are broken down by age and sex, but the major division is between retired workers, their dependents, and the survivors of deceased workers (who receive Old-Age and Survivors Insurance, or OASI, benefits) and disabled workers and their dependents (who receive DI benefits).

The number of OASI beneficiaries is a relatively stable percentage of each age and sex group. As a result, projecting growth in the number of those beneficiaries is fairly easy for a given age- and sex-specific population. Although changes in labor force participation, earnings patterns, and retirement rates affect the number of OASI beneficiaries, the percentage of the elderly in the population has a much greater impact on Social Security's finances.

7. 1999 Technical Panel on Assumptions and Methods, *Report to the Social Security Advisory Board* (November 1999).

The number of DI beneficiaries is much less certain. SSA projects disability rates by age and sex. Because those rates increase substantially with age, the gross (overall) disability rate will rise in the future even without changes in age- and sex-specific disability rates. However, the overall rate depends mainly on the projected rates of disability incidence and termination (described in Chapter II).

Projecting Per Capita Revenue and Benefit Levels

Levels of Social Security revenues and benefits per person depend primarily on the growth of wages. That growth in turn can be separated into two sources: real wage growth (which is effectively determined by productivity growth) and inflation.

Average Revenue Levels. By definition, the amount of Social Security payroll taxes paid per capita can be determined by multiplying the average effective taxable payroll by the statutory tax rate. Under current law, the tax rate is constant; thus, the only uncertainty about average revenue levels comes from the taxable payroll.

CBO's model begins with average taxable earnings in the past year, then increases that number by projected nominal wage growth—the sum of inflation and real wage growth (the first two exogenously projected economic variables).

Growth in the average payroll tax paid by a worker tracks average wage growth very closely, but not exactly. Workers pay Social Security taxes only on amounts below the statutory maximum set for taxable earnings—\$80,400 in 2001. That maximum is indexed to overall wage growth. If the wage distribution changes, the share of earnings below that level also changes. Currently, about 84 percent of covered wages are taxable; that figure is expected to decline, however, by about 0.1 percent per year in the near term. LTAM does not permit users to vary that assumption. Still, any uncertainty about the effect of distributional changes on the growth of taxable payroll is probably small compared with uncertainty about overall wage growth.

Average Benefits. An average benefit level must be projected for each of Social Security's many beneficiary categories. To analyze the effect of changing the complex benefit formula, CBO, like SSA, employs a microsimulation that projects average benefits for newly retired and newly disabled workers. (Microsimulation involves producing exact calculations of benefits for a simulated sample of the population and aggregating the results from the sample. That type of simulation is necessary because the effects of policy changes on different workers vary widely depending on factors such as those workers' wage levels and years of employment.) LTAM takes a sample of thousands of newly entitled beneficiaries from SSA's Continuous Work History Sample, which includes each worker's entire wage history. Using those wage histories, the model calculates a benefit for each worker and then computes averages for newly retired and newly disabled workers.

Because beneficiaries who retire in 2010 or 2050 will have different earnings patterns from today's retirees, the model adjusts the data to account for changing labor force participation patterns. Other changes that might occur between now and then are not simulated.⁸ The input variables in LTAM that affect benefit levels are wage growth, unemployment, inflation, and, to a lesser extent, disability rates.

WHAT TYPES OF UNCERTAINTY ARE NOT BEING CAPTURED?

This paper quantifies the uncertainty in actuarial projections that results from uncertainty in the nine major input assumptions. However, other sources of uncertainty not included in this analysis will also affect Social Security's finances. Thus, the measures of uncertainty described in Chapter V should be considered lower bounds.

Every model is, by necessity, a simplification of events that are likely to occur in the real world. That simplification is particularly true of macroeconomic models, which employ a few variables to represent an entire national economy. Models for Social Security are simpler because explicit rules can be employed to model a specific person's experience. Still, LTAM is only an actuarial model, which generally substitutes averages for actual person-by-person calculations. (CBO is working on a comprehensive microsimulation model that will generate individual-level calculations for benefits and other information.) For the model to function, assumptions about the actuarial processes must be made.

For some factors, CBO assumes that current ratios will remain constant. For example, like SSA, it assumes that the current relationship between average auxiliary benefits (those based on another person's work history) and average worker benefits (those based on one's own work history) will stay the same. (Each type of auxiliary benefit is tracked separately.) That assumption is only an approximation. To the extent that the relationship changes over the next 75 years—and it will—LTAM's projections will be in error. However, CBO expects that error to be small, for two reasons. First, auxiliary benefits represent a minority (about one-quarter) of total Social Security benefits. Second, the ratios used to project auxiliary benefits are bounded between zero and one and are unlikely to be near those bounds. Thus, the magnitude of possible errors is limited.

Other required factors that are held constant in CBO's model include the labor force participation rate, the fraction of workers eligible for benefits, and the rate of retirement. In addition, because Social Security's benefit formulas are progressive (an extra dollar of earnings increases a low-income person's benefit more than it would

8. LTAM uses data on the cohort that claimed benefits in 1997. Those data are adjusted for changes in 1998 through 2075.

increase a high-income person's benefit), shifts in the distribution of income would change average benefit levels and, to a lesser extent, the effective payroll tax base. LTAM assumes, however, that the distribution of wages and benefits will stay relatively constant. That assumption should be suitable for CBO's purposes because any variation is likely to be small, and the Social Security system should not be particularly sensitive to the variation that will arise.

CHAPTER II

SOCIAL SECURITY'S INTERMEDIATE PROJECTIONS

In developing its Long-Term Actuarial Model, the Congressional Budget Office has followed the practices of the Social Security Administration's Office of the Chief Actuary, with certain adaptations. This analysis of the uncertainty of LTAM's projections looks at the same key 75-year assumptions for demographic and economic variables that SSA focuses on in its annual analysis of the solvency of the Social Security trust funds. The demographic variables are

- Fertility rate,
- Rate of mortality improvement, and
- Immigration level.

The economic variables are

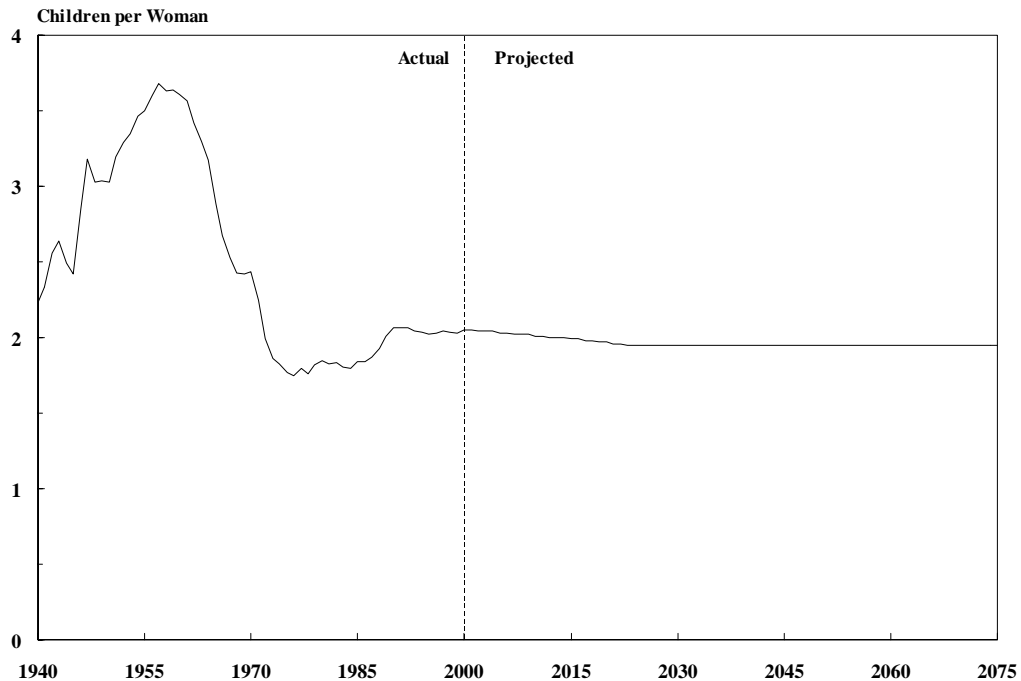
- Real wage growth,
- Inflation,
- Unemployment rate,
- Interest rate,
- Disability incidence rate, and
- Disability termination rate.

Each year, the Social Security Board of Trustees makes 75-year projections for those nine variables. Understanding the challenges it faces in choosing expected values for those inputs over the projection period is a first step in understanding the uncertainty inherent in each of the input assumptions.

DEMOGRAPHIC INPUTS AND POPULATION PROJECTIONS

Because Social Security is basically structured as a pay-as-you-go system, its financial status is primarily determined by the ratio of beneficiaries to workers, which in turn depends on the ratio of elderly to working-age people (known as the aged dependency ratio).

FIGURE 2. INTERMEDIATE PROJECTIONS FOR THE OVERALL RATE OF FERTILITY



SOURCE: Social Security Administration, Office of the Chief Actuary.

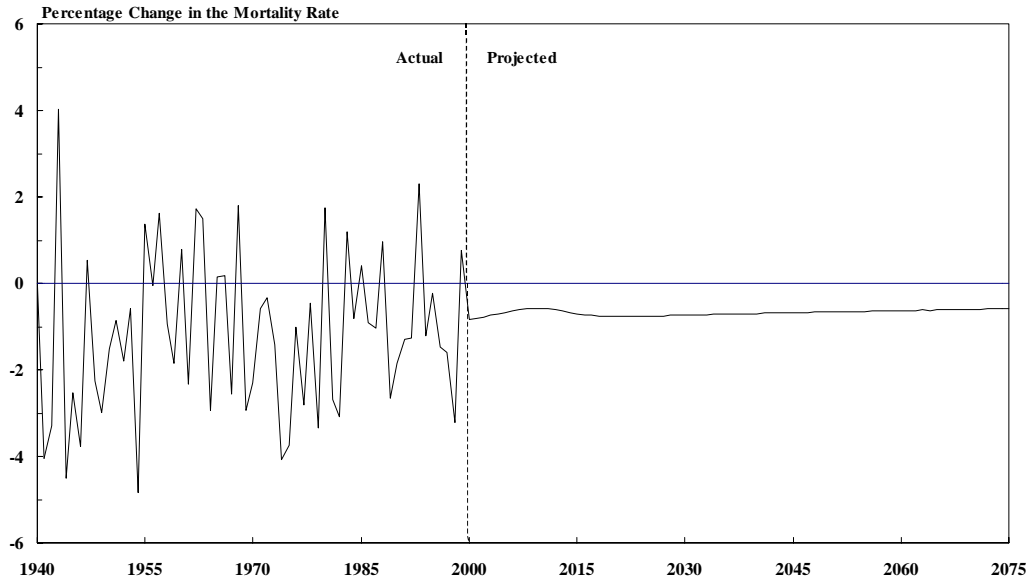
Although parameters of the Social Security program vary automatically with changes in wage growth and inflation, no such adjustments occur when the size or structure of the population changes. Thus, estimates of fertility and mortality (and, to a lesser extent, immigration) are very important in projecting the program's future finances.

Fertility

The fertility rate in a specific year is the number of children that would be born to a woman over her lifetime at the average birth rates for women of different ages observed in that year. Currently, the fertility rate is about 2.06 (see Figure 2). SSA's intermediate assumption is that the rate will fall steadily, but only slightly, over the next 25 years to 1.95, after which it will stay at that level. Since 1972, fertility rates have remained roughly constant, on average, although they have been much lower than the very high rates of the baby-boom era (1946 to 1964).

The projection of a 1.95 fertility rate assumes that changing social standards and the availability of oral contraceptives make it very unlikely that fertility rates will return to the levels of more than 3.0 seen during the baby boom. Actuaries at SSA

FIGURE 3. INTERMEDIATE PROJECTIONS FOR THE OVERALL RATE OF MORTALITY IMPROVEMENT



SOURCE: Social Security Administration, Office of the Chief Actuary.

NOTE: For this figure, mortality improvements have been aggregated across age and sex groups, resulting in a measure of the change in what demographers refer to as the "central" death rate.

are making a reasonable inference when they suggest that there are no grounds to expect big changes in fertility behavior in the next 75 years.

Nevertheless, the fact that little basis exists for projecting fertility, coupled with the fact that fertility is a chief source of financial instability in Social Security, is a reason to directly consider the uncertainty that surrounds fertility. As shown in later chapters, however, the degree of uncertainty about future fertility depends to some extent on how the historical tea leaves are read.

Mortality Improvement

Mortality projections start with a matrix of current mortality rates, by age and sex. (The mortality rate is the portion of people of a given age and sex who die in a given year.) Historically, mortality rates fall when living conditions and medical care improve; in the absence of a major new disease or significant social disruption, they are expected to continue declining. The change in mortality—the rate of mortality improvement—varies significantly from year to year (see Figure 3). However, it is negative in most years, indicating a general decline in mortality rates.

The difficulty in making an intermediate projection for mortality improvement lies in determining overall trends in historical variability and anticipating whether the *rate* of mortality improvement will change in the future. SSA projects age- and sex-specific improvement rates (for 21 age groups for each sex). SSA's model assumes that those rates will rise steadily for the next 25 years and remain constant thereafter. In contrast, it expects the overall rate of improvement, which is adjusted by age and sex, to dip slightly and then gradually return to roughly the current level by 2075 (see Figure 3).

SSA's projected overall rate of mortality improvement is approximately equal to the average of recent annual rates. Examining data since 1940 shows huge short-term variation as well as longer-term changes in mortality rates, with relatively large declines from 1940 to the mid-1950s that were repeated in the 1970s, and a stable rate of decline during the 1960s. From 1960 to 1968, the overall mortality rate decreased by an average of 0.2 percent a year. Over the next 14 years, however, the average drop was 1.9 percent. Since 1983, the mortality rate has fallen by an average of 0.8 percent a year.

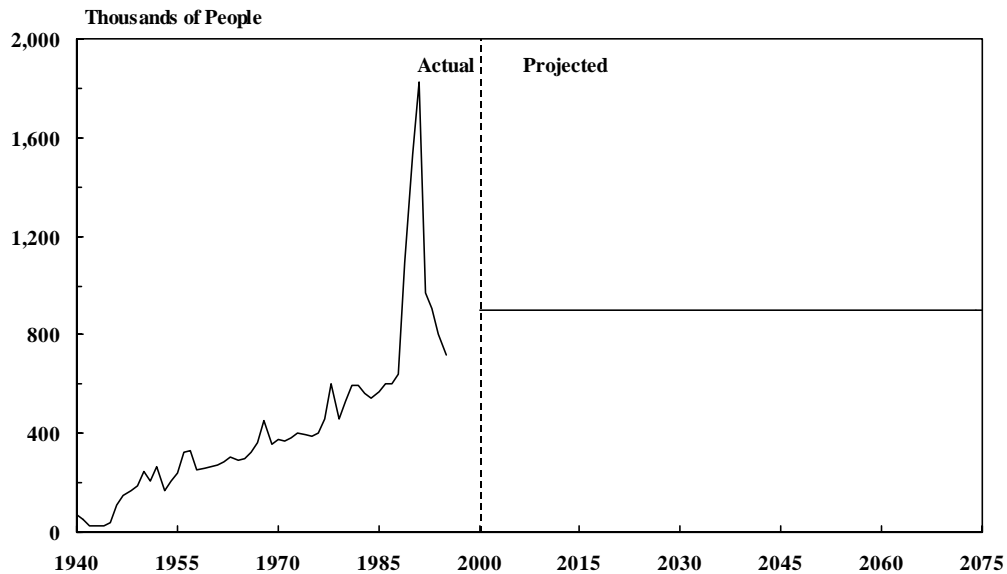
Immigration

SSA's intermediate assumption for the level of immigration is 900,000 people in each of the projection years. At first glance, the clear trend of growth in immigration over the past 50 years makes that flat level appear unlikely (see Figure 4). Moreover, it may seem natural to think that immigration will increase with the total population. However, unlike fertility and mortality rates, the number of immigrants to the United States is governed partly by law, and Social Security's baseline is calculated on the basis of current law.

ECONOMIC INPUTS

The six economic inputs that go into financial projections for Social Security can be classified in three categories: variables that affect how much workers earn and receive in benefits (wage growth, inflation, and unemployment); variables that affect the growth of the trust funds over time (the interest rate on Social Security assets); and variables that affect the ratio of beneficiaries to workers (disability incidence and termination rates).

FIGURE 4. INTERMEDIATE PROJECTIONS FOR THE LEVEL OF IMMIGRATION



SOURCE: Social Security Administration, Office of the Chief Actuary.

NOTE: Recent historical data for the level of immigration are unavailable.

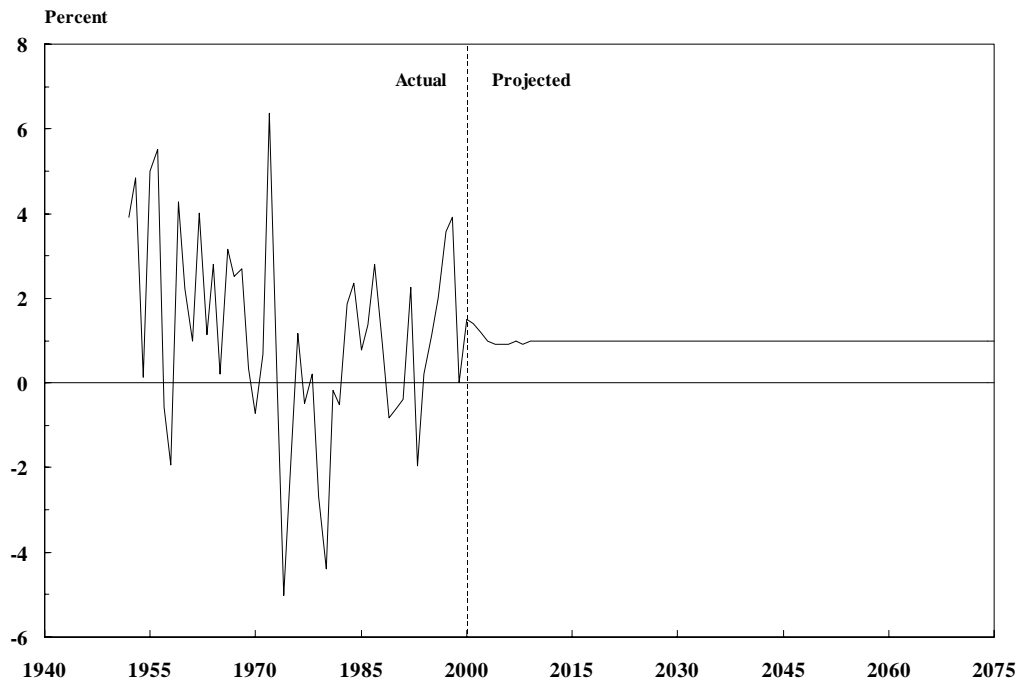
Real Wage Growth

The nominal rate of wage growth used to compute taxable wages and determine benefits in Social Security projections is the sum of the inflation rate and real wage growth. Although real wage growth depends mainly on productivity growth, factors such as the average number of hours worked and the level of nonwage compensation also matter. SSA analyzes and projects each component of wage growth separately.

Real wage growth varies significantly both from year to year and over longer time periods (see Figure 5). Most of that variability is attributable to changes in productivity growth. (Other factors, such as average hours worked and the ratio of taxable wages to total compensation, are more stable.) Because productivity growth is one of the great unexplained variables in economics, real wages are hard to project with precision.

The most difficult part of forecasting real wage growth is guessing whether future growth will be consistent with that of the previous few years, the past few decades, or the entire postwar period—all of which are very different. The past few years have been characterized by a boom in real wage growth and productivity, with sustained rates of increase not seen since the early 1960s. But is that recent boom an indication

FIGURE 5. INTERMEDIATE PROJECTIONS FOR THE RATE OF REAL WAGE GROWTH



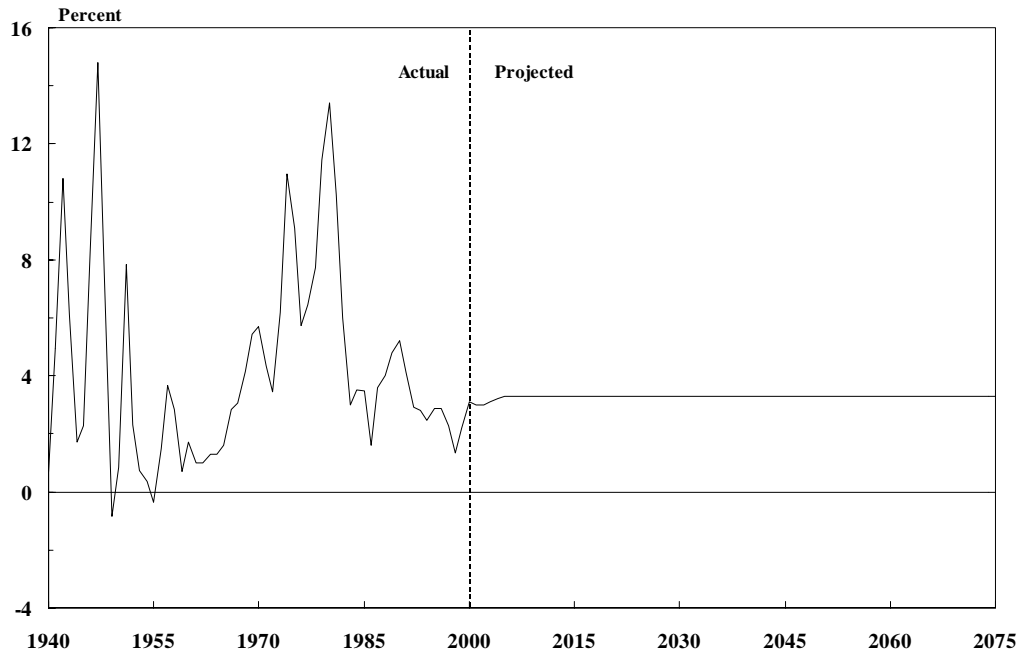
SOURCES: Department of Commerce, Bureau of Economic Analysis; Social Security Administration, Office of the Chief Actuary.

that the United States is returning to an era of strong growth or simply an aberration from the lower level of average growth that appears to have been in place from the early 1970s through the mid-1990s?

SSA has taken a fairly conservative view of recent productivity gains. Thus, its intermediate projection of average real wage growth over the next 75 years—about 1.0 percent per year—is consistent with the idea that the United States will return to the lower rates of growth seen since the 1970s.

The assumption about real wage growth is crucial because, in the short term, high real wage growth causes Social Security revenues to rise but has little impact on benefits. Thus, it improves the program's financial balance. Long-term finances are not quite as sensitive to changes in the rate of real wage growth because initial Social Security benefits are indexed to such growth. Nevertheless, higher wage growth promotes long-term solvency because it causes an immediate rise in payroll tax revenues but a delayed increase in benefits, since only new benefits are indexed to wages. (After a Social Security recipient begins collecting benefits, those payments rise each year with prices.)

FIGURE 6. INTERMEDIATE PROJECTIONS FOR THE INFLATION RATE



SOURCES: Department of Labor, Bureau of Labor Statistics; Social Security Administration, Office of the Chief Actuary.

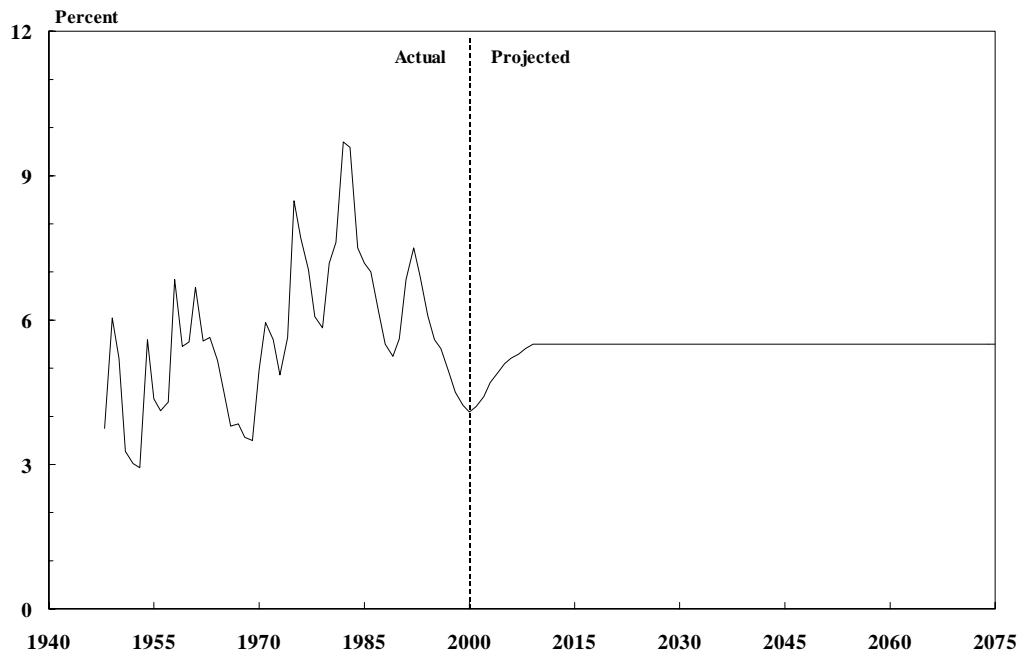
NOTE: In this figure, inflation is measured as the change in the consumer price index for wage earners (CPI-W).

Inflation

Prices, as measured by the consumer price index for wage earners (CPI-W), increased at an average rate of 4.4 percent a year during the four decades between 1959 and 1998, but by only 3.2 percent a year during the decade from 1989 to 1998. In projecting that inflation will rise to only 3.3 percent by 2005 and remain at that level thereafter, SSA may be giving too much weight to recent experience (see Figure 6). However, the inflation booms of the 1940s and 1970s were clearly linked to events—such as World War II and the OPEC oil embargo—that are unlikely to recur, which makes SSA's assumption seem more reasonable.

Assumptions about inflation have an important impact on Social Security's finances, but not in the way that most people would think. Social Security benefits are indexed by law to inflation in the previous year, but wages are indexed implicitly to current inflation (because nominal wage growth includes inflation). Thus, a rise in inflation increases revenues immediately but raises benefits only in the future. Although high inflation is generally considered harmful—especially for the elderly, who

FIGURE 7. INTERMEDIATE PROJECTIONS FOR THE UNEMPLOYMENT RATE



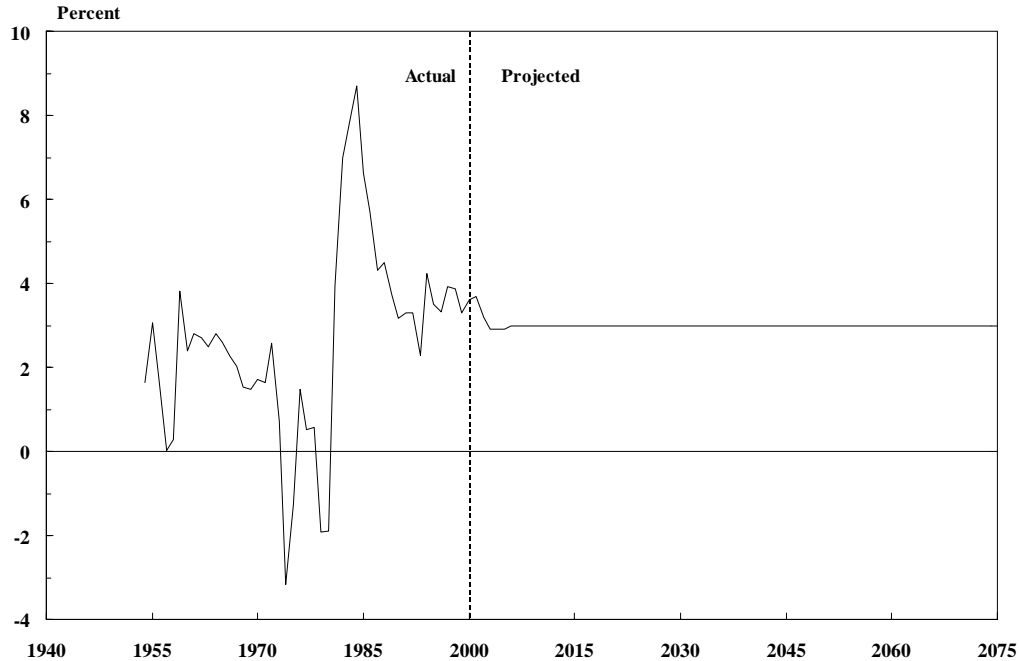
SOURCES: Department of Labor, Bureau of Labor Statistics; Social Security Administration, Office of the Chief Actuary.

often receive pensions that are not indexed to prices—inflation actually improves the actuarial status of Social Security.

Unemployment

The unemployment rate is the percentage of the labor force that is not employed but is actively looking for work. As with any volatile macroeconomic variable, projecting unemployment in the short run is very difficult, although a long-run tendency around fairly stable averages seems to exist (see Figure 7). The positive experience of the past few years has been atypical, and it would be optimistic to assume that unemployment will remain at such a historically low level. Still, recent history prompted SSA to reduce its long-term projection of unemployment from 6.0 percent in its 1998 report to 5.5 percent in its 1999 report. Currently, SSA assumes that the unemployment rate will rise smoothly from 4.1 percent in 2000 to 5.5 percent in 2009 and remain at that level thereafter.

FIGURE 8. INTERMEDIATE PROJECTIONS FOR THE REAL INTEREST RATE ON ASSETS IN THE SOCIAL SECURITY TRUST FUNDS



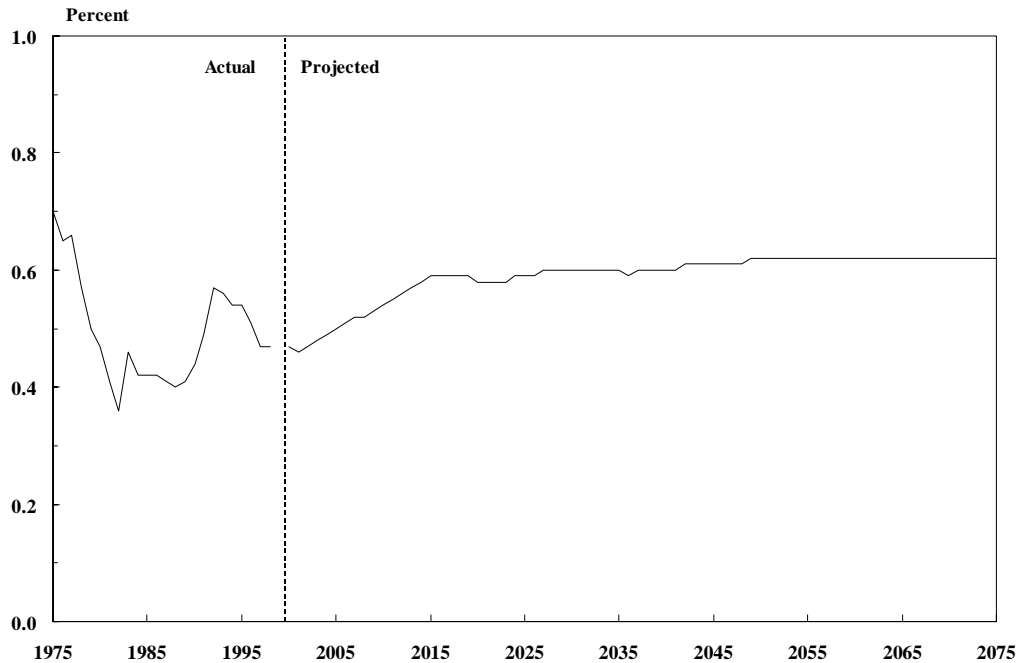
SOURCE: Social Security Administration, Office of the Chief Actuary.

Interest Rate

SSA lists as one of its key inputs the nominal interest rate on assets held in the Social Security trust funds. Under standard economic theory, nominal interest rates are the sum of inflation and the real interest rate—the return investors would receive if inflation were zero. Because inflation is a separate input, it is appropriate to describe returns on trust fund assets in terms of the real interest rate and to sum the two assumptions to compute nominal returns on those assets.

SSA projects a real interest rate for the trust funds of 3.0 percent per year over the next 75 years (see Figure 8). Although that rate is significantly higher than the historical average (about 2.4 percent since 1926), it is much lower than the average over the past few decades (4.2 percent since 1980 and 3.4 percent since 1990). SSA's assumption is consistent with two offsetting observations: first, real interest rates were negative during the 1970s because the high inflation of that period surprised investors, and second, nominal interest rates take time to reflect new information about inflation. The spike in real interest rates during the early 1980s was associated with the extremely tight monetary policy of that time (which was put in

FIGURE 9. INTERMEDIATE PROJECTIONS FOR THE RATE OF DISABILITY INCIDENCE



SOURCES: Congressional Budget Office; Social Security Administration, Office of the Chief Actuary.

NOTE: Recent historical data for the rate of disability incidence are unavailable.

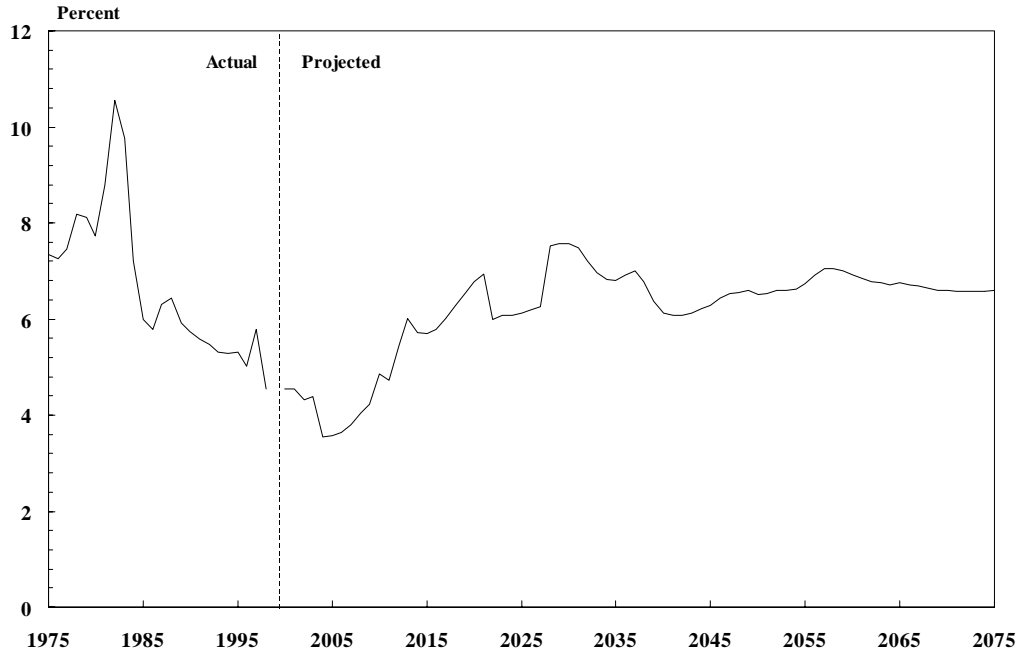
place to rein in inflation, which had galloped out of control). Thus, unless inflationary shocks and tight monetary policy recur, there is no reason to suspect that the real interest rate will diverge from that experienced during more normal times—about 3.0 percent.

Disability Incidence and Termination

SSA projects separate rates for the onset and ending of disability. The onset (or incidence) rate is the percentage of nondisabled workers covered by Disability Insurance who become disabled in a given year. The termination rate is the percentage of disabled workers who leave the DI rolls in a given year—because they either recover, die, or begin receiving retired-worker benefits instead (which occurs automatically when they reach Social Security's normal retirement age).

To project the prevalence of disability, SSA starts by projecting future trends in morbidity (the incidence of disease). Like mortality rates, morbidity rates are ex-

FIGURE 10. INTERMEDIATE PROJECTIONS FOR THE RATE OF DISABILITY TERMINATION



SOURCES: Congressional Budget Office; Social Security Administration, Office of the Chief Actuary.

NOTE: Recent historical data for the rate of disability termination are unavailable.

pected to continue to fall.¹ Still, projecting disability status is more complicated than that because it involves not only morbidity but, perhaps more important, government policy (which varies even in the absence of legislation), changes in the labor market's demand for certain skills, and social definitions of disability.

Thus, although rates of morbidity are improving, rates of disability incidence have not changed steadily over time (see Figure 9). The overall rate of incidence has declined in the past several years; however, SSA expects it to rise over the next few decades and then level out.

Rates of disability termination are also affected by underlying changes in demographics, because death rates play a part in determining how many people leave the disability rolls each year. The termination rate has fallen in the past few years, but that change was influenced by policy and social norms as well as by underlying demographics (see Figure 10). SSA expects that the termination rate will rise in the short term and thereafter be influenced mainly by demographics.

1. See, for example, E. Crimmins, Y. Saito, and D. Ingegneri, "Trends in Disability-Free Life Expectancy in the United States, 1970-1990," *Population and Development Review*, vol. 23 (1997), pp. 555-572.

PROJECTING SOCIAL SECURITY'S FINANCES USING THE LONG-TERM ACTUARIAL MODEL

A number of questions arise about the choice of input assumptions used to project the Social Security system's finances. Indeed, much of the focus of every year's trustees' report is on how the values for the inputs are determined. As noted above, historical data can be interpreted in many ways, and inferences about future values that are based on those data can also vary. In this analysis, CBO does not question the underlying choices of intermediate assumptions made by the trustees. Rather, it employs a model of trust fund accumulation to measure how uncertainty about those inputs leads to uncertainty about Social Security's finances.

CBO's new analytical tool, LTAM, is capable of generating estimates that basically match SSA's projections for a particular set of inputs. In other words, given values for fertility, mortality improvement, immigration, real wage growth, inflation, unemployment, the real interest rate, disability incidence, and disability termination, LTAM can generate projections of the trust fund balance that generally match those produced by SSA.

Moreover, the LTAM approach allows uncertainty to be studied in a way that is not possible with SSA's models. At SSA, the various pieces that go into a trust fund projection—population by demographic group, average wages, benefits, and other variables—are studied separately using analytical tools developed with different software and sometimes different computer systems. Results from each step are combined only at the end when trust fund calculations are made. In contrast, LTAM pulls all sectors together in one compiled model, which can be solved repeatedly (and very quickly) using any combination of input assumptions. That feature makes possible the uncertainty analysis described in the rest of this paper.

The cost of pulling projection techniques together into one model is a certain amount of simplification relative to SSA's projection techniques—LTAM lacks some of the detail of SSA's models. Still, it can simulate how the Social Security system will respond to changes both in assumptions and in most policy parameters. For key calculations—computing the number of deaths using a given mortality rate, or solving for average worker benefits using the microsimulation approach—LTAM attempts to follow the same modeling strategy used by SSA. The criterion for successful model development at CBO was ensuring that the answers for the baseline values were about the same as those produced by SSA's models. Given inputs and outputs from SSA, the model developed at CBO does a good job of mimicking those answers.²

2. To the extent that minor differences remain, CBO uses "calibration" factors in various parts of the model to produce exact matches of SSA's baseline projections.

Some elements of the LTAM projections, however, do not reproduce SSA's methods. Rather, they use SSA outputs as inputs to the model. For example, the ratio of auxiliary benefits to worker benefits is held fixed at ratios computed directly from Social Security output files. That practice is necessarily limited to parts of the model that would not vary (or would have a negligible change) with new policies or new inputs, because LTAM is primarily designed to capture changes in the Social Security system's finances that occur with changes in input assumptions or policy parameters.

Conclusions about whether an approximating model such as LTAM can show how system finances will change when inputs change—a prerequisite for the type of uncertainty analysis conducted here—should be based on evidence, not assertions. Such a basis is possible because results from LTAM can be compared with the results that SSA publishes in its “high-” and “low-cost” scenario analysis, which is discussed in the next chapter.

CHAPTER III
EVALUATING UNCERTAINTY
USING HIGH-COST AND LOW-COST SCENARIOS

Projections for the variables that underlie estimates of Social Security’s long-term finances almost always involve some error, either positive or negative. Indeed, before the uncertainty of those finances can be measured, the width of the error bands (the plus-or-minus range) for each input assumption must be established. Moreover, analysts need to investigate the amount by which the system’s finances change if the values for inputs move within those ranges. The Social Security Administration illustrates the uncertainty of the system’s long-term solvency by using “high-cost” and “low-cost” assumptions about the inputs in its models in addition to its intermediate, or baseline, assumptions (see Table 1). The alternative assumptions can be incorporated into a model separately (to gauge how sensitive Social Security’s finances are to movements in individual assumptions) or together (to assess the overall financial effects if every assumption moved to high- or low-cost values).

SSA’s approach to measuring uncertainty evolves naturally from its method for generating the intermediate projection. To produce that projection, SSA uses historical data about each input to make a best guess of the input’s long-term value, generally allowing for a short transition between the current-year figure and the ultimate value. Although SSA uses the same approach to generate the high- and low-cost scenarios, its projections of alternative ultimate values are generally more pessimistic (or optimistic) than historical experience would suggest. Still, those high- and low-cost values are intended to be within the range of plausible outcomes based on the historical evidence.

SSA’s approach has been criticized on a number of grounds—particularly, that it offers only two alternative scenarios and makes no predictions about the likelihood of those scenarios. Thus, it is impossible to make any sort of probabilistic interpretation of SSA’s estimates of the system’s long-term finances. Developing probabilistic measures requires a different approach to analyzing the historical data about inputs—SSA’s high- and low-cost scenarios are simply insufficient for the job. For example, they cannot be used to go beyond the traditional estimate that the Social Security trust funds will be exhausted by a certain date to consider the probability that the funds will be exhausted at other points in time.

SSA’S RANGES OF VALUES FOR INPUT ASSUMPTIONS

SSA’s high-to-low range for input assumptions builds directly on how the agency projects intermediate assumptions (see Chapter II). The same historical data used to

TABLE 1. THE SOCIAL SECURITY ADMINISTRATION'S RANGES OF ULTIMATE VALUES FOR NINE PRIMARY INPUTS

Input	Long-Term Annual Value for Input Under Different Assumptions About Cost		
	Intermediate Cost	Low Cost	High Cost
Fertility Rate (Number of children per woman)	1.95	2.20	1.70
Rate of Mortality Improvement (Percentage reduction in the mortality rate)	0.7	0.2	1.2
Immigration Level (Thousands of people)	900	1,210	655
Rate of Real Wage Growth (Percent)	1.0	1.5	0.5
Inflation Rate ^a (Percentage change in the consumer price index)	3.3	2.3	4.3
Unemployment Rate (Percent)	5.5	4.5	6.5
Real Interest Rate on Social Security Assets (Percent)	3.0	3.7	2.2
Disability Incidence Rate ^b (Percent)	0.5	0.4	0.6
Disability Termination Rate ^c (Percent)	3.8	4.6	3.1

SOURCES: Social Security Administration; Congressional Budget Office.

- a. SSA's low-cost scenario assumes a relatively strong economy, including a low rate of inflation, even though higher inflation improves the actuarial status of Social Security (by increasing revenues immediately but raising benefits only in the future). Conversely, SSA's high-cost scenario assumes a worse economy and a higher rate of inflation.
- b. SSA's actuaries set separate rates of change for disability incidence for men and women relative to a base period of 1980 to 1984. The rates reported here (which are adjusted for age and sex) are from CBO's Long-Term Actuarial Model (LTAM), are relative to 2000, and generate disability incidences that match those from SSA.
- c. SSA's actuaries set separate rates of change for disability termination by recovery and by death relative to a base period of 1977 to 1980; within the category of termination by death, they set separate rates for men and women. The rates reported here (which are adjusted for age and sex) are from LTAM, are relative to 2000, and generate disability terminations that match those from SSA.

generate expected values for inputs can be used to generate more optimistic or pessimistic—but still plausible—values. The high- and low-cost estimates are not associated with any probabilistic interpretation by SSA; rather, they are intended to be, as the trustees' report says, "illustrative" of how Social Security's finances would change if inputs moved around within the plausible range.

The most common means for determining those ranges is to review actual historical ranges for each variable (shown in Figures 2 through 10 in Chapter II), even when the exact relationship between historical and future variability requires some interpretation. The fertility rate is one example. SSA estimates that the fertility rate over the 75-year projection period will be 1.95, with a range of plus or minus 0.25 for the high- and low-cost estimates. That range may seem narrow compared with fertility rates over the past century, whose range was much larger than 1.70 to 2.20. But it is plausible given the fact that the fertility rate has stayed within a fairly narrow band for the past 35 years or so. The assumption that the fertility rate will remain within that narrow range implies that the Social Security trustees think a structural shift in fertility behavior occurred around 1965 and is unlikely to be reversed. Thus, two different readings of historical data give two very different answers about whether the range is too narrow or about right. (That same problem—whether historical data should be used to measure variability without considering structural change—also occurs when estimating probability-based uncertainty bands for the inputs.)

Structural breaks in a data series are harder to discern for some inputs, largely because of the difficulty in distinguishing long-term trends from annual changes. For example, the rate of mortality improvement adjusted for age and sex has clearly fluctuated significantly from year to year since 1940 (see Figure 3 on page 13). Overall, the average change has been toward longer life spans, a trend that SSA expects will continue at a rate of about 0.5 percent per year. That rate of improvement is slower than the overall average for the past century, however, because (between annual fluctuations) the data indicate a slowdown in the rate of mortality improvement in the past few decades. To some extent, SSA's high-cost value for mortality improvement is consistent with a return to the rapid declines of the early 20th century, whereas its low-cost value implies a continued slowdown in the rate of improvement. Most of the other inputs have characteristics similar to mortality improvement: trends exist but are difficult to separate (at least visually) from annual fluctuations.

In the end, it is impossible to evaluate whether the high- and low-cost error bands are reasonable in a probabilistic sense without using statistical models to interpret the historical data. Using such models would address two issues that the Social Security Advisory Board has raised about how input assumptions vary in the high- and low-cost simulations. Both issues boil down to the criticism that SSA varies inputs in ways that are not consistent with historical data.

First, as in the case of intermediate projections, SSA considers only long-term (ultimate) high- and low-cost values for each input. Those values are generally assumed to be in force within a few years, after a short transition period during which a variable changes smoothly from its most recent historical value to the assumed ultimate value. SSA acknowledges that any entire sequence of annual values would be much more volatile; the expectation is only that, *on average*, an input will stay within the high-to-low-cost error band. None of the three scenarios include any of that volatility, however.

That point is important because, for some inputs, the sequence of values may matter as much as the long-run average. Consider fertility: a baby boom followed by a baby bust might generate the same long-run *average* fertility rate as a flat sequence, but it would produce a very different type of population and have very different effects on the Social Security system.

Second, SSA provides only three scenarios. In each one, every input is always at its intermediate-, high-, or low-cost value. (Technically, SSA assumes that the correlations between the inputs are fixed at 1.) The inputs should be varied, however, to reflect actual correlations in historical data. Only through those correlations can analysts answer such questions as, Is high mortality usually associated with low fertility (as scenarios suggest), or is their correlation in the opposite direction, or is there no correlation at all? Historical data can answer those questions only at an annual frequency. It is not clear that historical correlations are useful if a model is solved using ultimate values.

HOW VARYING THE INPUT ASSUMPTIONS AFFECTS SOCIAL SECURITY'S PROJECTED FINANCES

Using SSA's scenarios to measure uncertainty about the finances of the Social Security system raises an essential question: when input assumptions are systematically varied between high- and low-cost values, what happens to estimates of those finances? The effect of changing any given input (sensitivity analysis) depends on how much the input is changed and on how much impact occurs per unit of change. The overall effect of changing every assumption (scenario analysis) is the sum of the individual effects plus any interaction effects, which may be either reinforcing or offsetting. Although using the scenario approach to measure uncertainty has shortcomings, going through the exercise is instructive because it points out some of the building blocks that can be used to develop better estimates of uncertainty (such as Monte Carlo simulation, which is discussed in the next chapter).

SSA's Sensitivity Analysis

Sensitivity analysis indicates the extent to which Social Security's finances are affected per unit of deviation in a given input. For example, when the fertility rate increases, sensitivity analysis shows how much the trust fund balance or some other financial measure changes per unit of increase. However, sensitivity analysis is not meant to measure uncertainty in the system's finances in any real sense—it is only a tool for constructing overall measures of uncertainty.

Performing a sensitivity analysis (or the scenario analysis described below or a Monte Carlo stochastic analysis) requires an objective measure of financial outcomes. The most widely used measure of the solvency of the Social Security trust funds is the 75-year actuarial balance. That balance currently stands at -1.89 percent of taxable payroll, which means that the present value of expected costs (mostly benefits that will be paid) exceeds the present value of expected income (mostly Social Security payroll taxes that will be collected) by 1.89 percent of the present value of expected taxable payroll.¹ The 75-year actuarial balance is a useful measure because interpreting it for policy purposes is straightforward: payroll tax rates would need to be raised (or benefits cut) immediately and permanently by 1.89 percent of taxable payroll to put the Social Security system into financial balance for the next 75 years.

As part of the trustees' annual report, SSA produces sensitivity estimates for eight of the nine key inputs (all but unemployment). Those estimates measure the effects on the 75-year actuarial balance of using high- or low-cost assumptions for the eight inputs. For example, if the fertility rate is at its low-cost value (2.20) rather than its intermediate value, the 75-year actuarial balance improves by 0.27 percentage points—from -1.89 to -1.62 percent of taxable payroll (see Table 2). Similarly, if the fertility rate is at its high-cost value (1.70), the estimated actuarial balance worsens by 0.30 percentage points—from -1.89 to -2.19 percent of taxable payroll.

The estimated sensitivity of Social Security's finances offers an initial look at how much of the uncertainty in the system can be attributed to each of the key input assumptions. Varying the reduction in overall mortality rates between the high- and low-cost values has the greatest effect on the 75-year actuarial balance, with the real interest rate, real wage growth, disability incidence, fertility, inflation, immigration, and disability termination following in that order. However, because SSA does not attach probabilities to the high- and low-cost values for any input, the estimated sensitivities of the inputs may not be comparable. (That problem is resolved with the Monte Carlo analysis discussed in the rest of this paper.)

1. The analysis in this paper is based on Social Security Administration, *The 2000 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds* (March 30, 2000). The 2001 version of that report projects a 75-year actuarial balance of -1.86, but it makes no changes to the long-term input assumptions.

TABLE 2. SENSITIVITY OF SOCIAL SECURITY'S 75-YEAR ACTUARIAL BALANCE TO CHANGES IN ULTIMATE INPUT VALUES

	Estimated Change in 75-Year Actuarial Balance ^a (As a percentage of taxable payroll)			
	SSA's	SSA's	CBO's	CBO's
	Model Using Low-Cost Value	Model Using High-Cost Value	Model Using Low-Cost Value	Model Using High-Cost Value
Input				
Fertility	0.27	-0.30	0.31	-0.31
Mortality improvement	0.69	-0.79	0.63	-0.74
Immigration	0.14	-0.12	0.18	-0.15
Real wage growth	0.51	-0.50	0.57	-0.55
Inflation ^b	-0.23	0.22	-0.28	0.23
Real interest rate	0.49	-0.58	0.49	-0.57
Disability incidence	0.30	-0.30	0.33	-0.33
Disability termination	<u>0.06</u>	<u>-0.06</u>	<u>0.08</u>	<u>-0.08</u>
Total (Sum of individual changes)	2.23	-2.43	2.31	-2.50
Effect of Changing All Inputs Together	2.27	-3.11	1.96	-3.21

SOURCES: Social Security Administration; Congressional Budget Office.

- a. The 75-year actuarial balance in SSA's intermediate projection published in 2000 is -1.89 percent of taxable payroll.
- b. SSA's low-cost scenario assumes a relatively strong economy, including a low rate of inflation, even though higher inflation improves the actuarial status of Social Security (by increasing revenues immediately but raising benefits only in the future). Conversely, SSA's high-cost scenario assumes a worse economy and a higher rate of inflation.

Although conclusions are limited by the lack of exact probabilistic interpretations, sensitivity analysis can convey some information about the relationship between inputs and the system's overall finances. If an input is varied within the plausible range and the actuarial balance remains near -1.89 percent, one can argue that the Social Security system is robust with respect to that input. If, however, the range of outcomes around -1.89 percent is much larger—either because the plausible range for inputs is wide or because the system responds dramatically to small changes in inputs—it is worth investigating why the system is so sensitive.

For example, one reason mortality has such a great effect on the 75-year actuarial balance is that Social Security has no rules to offset changes in mortality. Retirement ages are effectively fixed by law, so changes in mortality translate directly into changes in how long recipients collect benefits. If the retirement age were tied to mortality, finances would be much less sensitive to such changes. The system is less

sensitive to changes in other inputs—for example, the effect of real wage growth is muted because higher growth rates lead to both higher payroll tax revenues and higher benefits. In that case, the impact on the system’s finances is positive because the government begins receiving the higher taxes before it starts paying the higher benefits.

CBO’s Sensitivity Analysis

Sensitivity analysis is also useful in gauging whether the Congressional Budget Office’s Long-Term Actuarial Model properly simulates how changes in input assumptions affect Social Security’s finances. The model can be solved using techniques that mimic those developed by SSA—with input assumptions set to equal the high- and low-cost values—and thus can generate an alternative estimate of the impact of varying each assumption. The results of that exercise indicate that CBO has achieved its goal of having LTAM be able to capture the effect on system finances of varying each input in a way that mimics SSA’s model. Sensitivity results from LTAM are very close to those from SSA (see Table 2).²

Because LTAM has the same sensitivity as SSA’s model with respect to input assumptions, CBO has not limited its analysis of sensitivity to the measures published in the trustees’ annual report. That report’s present-value 75-year actuarial balance is not ideal as an indicator of financial solvency—the currently projected path for the Social Security system calls for rising trust fund balances that will be rapidly depleted as the baby-boom generation collects benefits. Moreover, analysts expect that the system will be left in poor financial health even after the baby boomers have died because collected taxes will continue to fall well short of benefits. Indeed, although the intermediate projection shows the present value of the cost-to-income gap at 1.89 percent of taxable payroll, the annual level of that gap is expected to be 4.26 percent of payroll in 2030 and 6.18 percent by 2075.

The same LTAM simulations used to estimate the present-value actuarial balance under high- and low-cost values for each input assumption also generate annual balances. Striking differences in the underlying sources of uncertainty appear when the focus turns to annual levels in 2075. The ultimate fertility rate—which has little effect through 2030 and moderate present-value effects (as shown in Table 2)—becomes the most important source of uncertainty by 2075 (see Table 3). Specifically, when the fertility rate is varied between its (relatively close) high- and low-cost values with all of the other variables held constant, the estimated gap between Social Security costs

2. Some of the differences may result from subtle distinctions in how the simulations are carried out. For example, CBO chooses to recompute average Social Security benefits for the sample population each time an alternative simulation is run; SSA may only be capturing the first-order effects. Also, CBO has to make some assumptions about fixed interactions between variables. For example, inflation is expected to affect the ultimate real interest rate as well as the ultimate nominal interest rate.

TABLE 3. SENSITIVITY OF SOCIAL SECURITY'S ANNUAL ACTUARIAL BALANCES TO CHANGES IN ULTIMATE INPUT VALUES

	Estimated Change in Annual Actuarial Balance (As a percentage of taxable payroll)			
	In 2030 ^a		In 2075 ^b	
	CBO's Model Using Low-Cost Value	CBO's Model Using High-Cost Value	CBO's Model Using Low-Cost Value	CBO's Model Using High-Cost Value
	Input			
Fertility	0.11	-0.10	2.01	-2.44
Mortality improvement	0.60	-0.65	1.81	-2.16
Immigration	0.40	-0.33	0.35	-0.32
Real wage growth	1.07	-1.16	1.83	-2.22
Inflation ^c	-0.61	0.55	-2.34	1.45
Disability incidence	0.46	-0.47	0.64	-0.68
Disability termination	<u>0.06</u>	<u>-0.05</u>	<u>0.31</u>	<u>-0.35</u>
Total (Sum of individual changes)	2.09	-2.21	4.61	-6.72
Effect of Changing All Inputs Together	1.98	-2.30	4.50	-7.87

SOURCE: Congressional Budget Office.

NOTE: The real interest rate on trust fund assets is excluded because changes in that rate do not affect Social Security's annual costs and income.

- a. The actuarial balance in 2030 in SSA's intermediate projection published in 2000 is -4.26 percent of taxable payroll.
- b. The actuarial balance in 2075 in SSA's intermediate projection published in 2000 is -6.18 percent of taxable payroll.
- c. SSA's low-cost scenario assumes a relatively strong economy, including a low rate of inflation, even though higher inflation improves the actuarial status of Social Security (by increasing revenues immediately but raising benefits only in the future). Conversely, SSA's high-cost scenario assumes a worse economy and a higher rate of inflation.

and income in 2075 rises and falls by more than 2 percent of taxable payroll—or over one-third of the intermediate value of 6.18 percent. Other variables, by contrast, introduce only modest additional uncertainty as the analysis is extended year by year. For example, the impact of real wage growth does not compound dramatically because (as noted above) higher real wage growth increases both revenues and benefits.

Scenario Analysis

Scenario analysis is a first step in determining uncertainty in estimates of the system's finances. If, for instance, the expected funding shortfall for the 75-year horizon is

1.89 percent of taxable payroll, analysts still want to gauge the chances that the actual shortfall will be well above or below that estimate. Assigning probabilities requires a particular type of model (see Chapters IV and V). Instead, SSA uses a scenario approach: after making its intermediate projection, SSA sets all of the inputs to their high or low values and solves the model again. That approach is intended to show how Social Security's finances would be affected if all of the values for inputs moved in the same direction but remained plausible. In that approach, the high and low estimates serve as boundaries.

Conclusions about the overall variability of financial outcomes depend entirely on the range selected for the input assumptions. SSA's choice of a fairly narrow range is consistent with the structure of the high- and low-cost scenarios, in which all of the inputs are set to the most pessimistic or optimistic values to generate a range for financial outcomes. A narrow range of assumptions makes sense because it is difficult to imagine that all of the variables would move in one direction or another—rather, positive and negative changes among inputs would be likely to offset one another.

The high- and low-cost scenarios, which are a standard part of the trustees' report, constitute SSA's only published assessment of the uncertainty of Social Security's future finances. If all of the inputs moved to their low-cost values, the 75-year actuarial balance would rise by 2.27 percent of taxable payroll, and Social Security's funding gap would more than disappear (see Table 2). Conversely, if all inputs equaled their high-cost values, the actuarial balance would worsen by 3.11 percent of taxable payroll, more than doubling the funding gap. Results from LTAM show the same basic pattern.

Comparing outcomes from the scenario approach with the sum of the results from individual sensitivity analyses is useful. Under SSA's calculations, the overall effect of moving to low-cost values is 2.27 percent of taxable payroll, whereas the sum of individual effects is 2.23 percent. Those results suggest that changing all of the inputs simultaneously produces reinforcing effects (though very modest ones). On the high-cost side, the sum of individual effects is much smaller than the results of the scenario analysis, which suggests greater reinforcing effects from changing all inputs together.

Like sensitivity analysis, the scenario approach can also be used to evaluate uncertainty in projections for specific years. SSA expects the actuarial balance in 2030 to be -4.26 percent of taxable payroll. CBO's low-cost estimate shrinks that gap by 1.98 percentage points, whereas the high-cost estimate widens it by 2.30 percentage points (see Table 3). For 2075, the underlying gap is -6.18 percent of payroll, and the low-cost estimate is 4.50 percentage points higher, whereas the high-cost estimate is 7.87 percentage points lower. Note that the same asymmetry of effects that appears in the 75-year present-value estimates—with the high-cost differential being larger than the low-cost differential—occurs in the estimates for 2030 and 2075.

CONCERNS ABOUT THE SCENARIO-BASED APPROACH

SSA's scenario approach provides only limited information about the uncertainty of the agency's Social Security projections. As a result, the 1999 Technical Panel of the Social Security Advisory Board stated, "It is our view that the Social Security Administration must develop different techniques for measuring uncertainty—not merely to refine predictions but to allow policy makers to consider reforms to Social Security that would lessen its sensitivity to adverse economic and demographic trends."³

The panel further noted that "no probabilities are associated with the forecast ranges, so the user has only a vague sense that the forecaster believes it likely that the outcomes will fall within the range." Moreover, even if the high and low input values had some probabilities assigned to them, policymakers would have only part of the story, the panel argued. "For example, the likelihood that fertility will fall within the high-low range in some given year may be much less than the likelihood that the summary actuarial balance will fall within the high-low range."⁴ Clearly, it is not possible to generate probability distributions for Social Security's finances without probability distributions for the inputs.

Probability distributions for inputs, however, do not automatically lead to distributions for outputs because all inputs must be varied to reflect actual historical correlations. The ideal approach would be to assign a probability to every possible combination of paths for input assumptions, solve for the system's finances under each set of paths, and then use the probabilities associated with each set of inputs to create a probability distribution for the outputs. Although it is impossible technically to assign probabilities to every set of outcomes, it is feasible to create an arbitrarily large sample of input combinations, solving each time for system finances, and then evaluate how finances vary within that sample and draw conclusions about the probability distribution of the outcomes. That technique—called stochastic simulation—is the approach recommended by the Social Security Advisory Board. CBO has implemented that approach using its Long-Term Actuarial Model, as described in Chapters IV and V.

3. 1999 Technical Panel on Assumptions and Methods, *Report to the Social Security Advisory Board* (November 1999), p. 32.

4. *Ibid.*, p. 70.

CHAPTER IV

MEASURING UNCERTAINTY ABOUT INPUT ASSUMPTIONS

Given the available historical data for each of the inputs that goes into projections of the Social Security trust funds, actuaries at the Social Security Administration do what sophisticated statisticians do when asked to forecast values for the next 75 years: they assume that future values will follow a pattern consistent with the past. To gauge how sensitive trust fund projections are to the assumed ultimate value for each input, the actuaries use high- and low-cost scenarios, also based on historical data.

That approach does not adequately reflect uncertainty about future trust fund balances, however, for three reasons. First, any average level for an input variable over 75 years is consistent with many possible annual paths (having different fluctuations), with variations in averages over five or 10 years, and with differences in that variable between birth cohorts—all of which can cause variation in the pattern of trust fund accumulations. Second, SSA's actuarial methods do not provide relative measures of the level of uncertainty of each input. Thus, some questions remain unanswered. For instance, is the projection of the unemployment rate or interest rate more uncertain than the projection of fertility? Third, projections by scenario do not incorporate any overall measures of probability for the input scenarios; without such measures, it is impossible to evaluate the likelihood of the results.

Developing better estimates of uncertainty about inputs is essential to developing better estimates of uncertainty about Social Security's finances. To truly measure uncertainty about the future values of inputs—in particular, to estimate probability distributions for annual values—it is appropriate to start (as SSA's actuaries do) with historical data. (The obvious limitation of such data is that future variation may differ from historical variation. To the extent that it does, estimates of uncertainty may be understated or overstated.)

The standard statistical tool for making inferences from historical data is time-series analysis. Such analysis starts by breaking down the historical changes in variables into three main components: annual random shocks that are either positive or negative (but centered around zero), year-to-year correlations in annual values, and random changes in the central tendency of the annual values. Because many variables—such as inflation, unemployment, the real interest rate, real wage growth, and the rate of mortality improvement—seem to have no random change in central tendency over long periods, analysts frequently need to model only the first two sources of change. Whether other variables—including the fertility rate, immigration, and dis-

ability incidence or termination—show changes in central tendency depends on how the historical tea leaves are read.

The decision about whether to incorporate random changes in central tendency is important because it dramatically affects conclusions about the possible range (and thus the probability distribution) of future values. In particular, if no random change in central tendency occurs, outcomes will vary within a probability range that is constant over time. For example, the range of possible outcomes for a variable such as inflation in 2075 would probably be the same as the range of outcomes in 2010. Allowing random changes in central tendency, by contrast, suggests that the range of possible outcomes will widen over time. For example, the range of outcomes for fertility in 2075 could be much wider than the range in 2010 because changes in central tendency generally occur gradually. In the short run, the fertility rate is likely to vary around a fairly predictable central tendency; but in the long run, fundamental social changes could affect average fertility.

PRODUCING A FORECAST WITH TIME-SERIES AND MONTE CARLO TECHNIQUES

The time-series analysis employed in this paper uses historical data to project an input's estimated variability (specifically, its probability distribution) around SSA's intermediate projection for that input. The Congressional Budget Office is interested in estimating the uncertainty of SSA's forecast, not in replacing the forecast itself. For the majority of variables, the projections that CBO generated through time-series analysis are quite similar to those used by the Social Security trustees. Most differences arise because the trustees weight recent experience more heavily. In some cases, the projections also differ because the trustees have used expert judgment about the relevance of past values to future trends.

Time-series analysis uses historical data to project both the future values of variables and the variability around those future values. In the first step, the movements in a variable (or group of variables) are broken down mathematically into the three components described above, and the resulting estimated equation (or set of equations) is used to generate expected future values simply by solving it forward through time.

Given an estimated equation, the second step involves employing computer simulation to generate probability distributions for future outcomes. The procedure used to produce those distributions is called Monte Carlo simulation because it involves making repeated random draws, in a mathematically structured way, from the values for annual shocks (much like what occurs in a casino, where, for example, the proba-

bility of rolling a particular number on a die is one in six).¹ Those annual random draws are plugged into the time-series equation, which then generates a time path of outcomes for the variable in question. By repeating the simulation process many times, analysts can draw inferences about the probability distribution of future outcomes.

Mathematical Models for Projecting a Time Series

Measuring the uncertainty of input variables using time-series analysis is inherently different from developing intermediate values or high- and low-cost scenarios because the statistical techniques for time-series analysis are designed to infer more than a best guess for the range of the long-term average. At one level, the time-series approach uses a fairly simple equation to explain how variables change from year to year. However, because the string of annual changes adds to long-term changes, the technique actually generates both short-term fluctuations and long-term trends simultaneously.

A casual look at the graphs of the inputs being modeled (Figures 2 through 10 in Chapter II) suggests that there is more than one reason why a variable changes at any point in time. One observation is that most of the inputs being modeled—inflation, unemployment, the real interest rate, real wage growth, and mortality improvement—have no trend change in central tendency over the long term. In that case, mathematical models need only incorporate the first two sources of change: annual random shocks and correlations of annual values over time.

Analysts generally assume that the probability distribution of annual random shocks can be approximated with the well-known “normal” pattern. In that standard approach, the values of random shocks have an expected level, or mean—in this case zero—with a symmetric bell-shaped distribution around that expected level. Thus, analysts are much more likely to draw a random shock that is close to the mean than one that is distant.

If the projected outcomes for a variable composed only of a long-run average and annual random shocks were graphed, all of the values would be centered around the average value for the variable because, by definition, the expected value of the random shocks is zero. In addition, the graph would have several features: approximately the same number of high and low values; more values close to the average than far away, because the distribution of the shocks is normal (bell shaped); and finally—a crucial

1. A closely related technique is “bootstrap” simulation, in which random draws are made from the collection of actual shocks that occurred in the past rather than from a theoretical distribution of shocks inferred from historical data. Both techniques are used in this analysis.

distinction—no pattern that connects the values over time. (Outcomes in each year would be independent of the outcomes in the previous year.)²

That description of a variable that has only an average value and annual random shocks does not appear to fit any of the inputs that go into projections of Social Security's finances. Rather, all of those inputs (even the ones with apparently stable long-run central tendencies) seem to move in one direction or another and then stay there for long periods—implying high correlation between outcomes from year to year—before moving back. For example, inflation was generally high in the 1940s, fairly low through the early 1970s, generally high for the next decade or so, and then generally low again (see Figure 6 on page 17). Clearly, variation occurs from year to year, but the outcomes also seem to be correlated over time.³

How much of a particular change is attributable to random shocks and how much to correlation between values over time? Time-series analysis specifies a simple equation for a variable and allows the data to answer that question. In the simplest specifications, the equation relates the current-period value of a variable to three things: a constant term (the central tendency), the value of the variable during the most recent period (in order to capture the correlation over time), and an error term (the random shock).⁴ More-complicated versions of time-series equations involve adding more lagged terms (not just for the most recent period but for two, three, or more previous periods) or employing a “moving average” of error terms, in which random shocks themselves affect outcomes for more than one period.⁵

How can a user tell if the correct equation was chosen to represent the time series being modeled? The answer is to go back to the premise underlying the equation. If

-
2. The time-series description of a series made up only of an average value and annual random shocks is “white noise.”
 3. Inflation is often described as a generalized autoregressive conditional heteroskedasticity (GARCH) process. In that type of process, the variance of errors increases with the level of the variable.
 4. The simplest specification for a variable x_t is:

$$x_t = \alpha + \beta x_{t-1} + \varepsilon_t$$

where t denotes time, α and β are parameters to be estimated, and ε_t is the residual (unexplained error) at time t . As described in the text, α represents the central tendency, and β captures the correlation of values over time. This type of equation can be estimated with standard regression techniques. Note that the derived residual (ε)—which represents implied random shocks—is used to estimate the variance for the random-shock process that feeds into the Monte Carlo simulation described later in this chapter.

5. In the language of time-series econometrics, a process is described in terms of its “AR” and “MA” properties, with “AR” denoting how many lagged terms are included in the equation and “MA” denoting how long the moving average is for the error terms. The simplest equation is an “AR(1),” which has only one lagged term and no moving average. The most complicated process in the list of Social Security inputs is an “ARMA(4,1),” meaning there are four lagged terms and a single-period moving average of errors.

the equation is appropriate, the residuals (error terms) derived from it will have the properties associated with a series of normally distributed random shocks—they will be centered around zero, have the same number of high and low values, have more realizations close to zero than far away, and show no correlation over time. Thus, the time-series approach involves specifying an equation, estimating the parameters with historical data, and testing whether the residuals are consistent with a series of random shocks.⁶

In principle, whether an equation passes that test determines whether unexplained changes in central tendency exist for a variable over time. If an equation for a variable generates residuals that appear to be random shocks, then arguably, no unexplained (random) changes in central tendency exist. All systematic movement in the variable has been captured by the equation, and there is nothing left to explain.⁷

Unfortunately, it is sometimes difficult to tell whether the processes being modeled show random changes in central tendency. The tests used to decide whether derived residuals look like random shocks are not definitive, especially when the time series is short. Thus, the process of deciding whether changes in central tendency have occurred can involve judgment. (As discussed below, CBO chose to use models with and without random changes in central tendency when the evidence seemed unclear.) If the possibility of a nonsystematic changing central tendency is admitted, the simplest approach is to “first difference” the variable in question—that is, to use an equation to describe the *change* in, rather than the *level* of, the variable.

Modeling change rather than level for a variable may seem like a trivial difference, but it has a profound effect on inferences about the bands of uncertainty around the variable. When change is modeled, any random shock *permanently* affects the level of the variable—the shock does not disappear by itself after one period, as in the usual specification. Of course, a random shock in the other direction pushes the level of the variable back in the other direction permanently. Thus, in first-differenced models, the level of a variable at any point in time is the result of cumulative shocks up to that point. Because shocks are all random, any cumulation in one direction pushes the level toward a new central tendency. Thus, uncertainty bands grow over time.

Conclusions about how first-differenced equations differ from level equations result to some extent from how inputs to the model are specified. For example, Figure 5 on page 16 suggests (and statistical tests confirm) that the growth rate of real wages had no change in central tendency in the past five decades. However, a picture of the average level of real wages (which has risen steadily over time) shows a clear

6. The test for random shocks is based on the Durbin-Watson statistic. See the appendix for details.

7. A time-series econometrician would describe this as a “stationary” series. The standard test for stationarity is based on the augmented Dickey-Fuller statistic. See the appendix for details.

change in central tendency.⁸ The growth rate of real wages is effectively a “differenced” version of the level of real wages—if that level was the input being modeled, tests would indicate changes in central tendency. The same relationship exists between the rate of inflation and the price level or between the change in mortality and the central death rate at a particular time.

Using Time-Series Equations to Generate Annual Probability Distributions

Once analysts have produced mathematical equations for an input, they can generate probability distributions for actual annual outcomes. The simplest time-series models imply that annual values depend only on a constant, on the previous period's value (multiplied by a coefficient), and on an annual random shock. Coefficients are generated when the time-series model is estimated using historical data. The extent to which an input varies around the value predicted by the equation indicates the correct size for annual random shocks. Thus, everything is in place to project future values using computer simulation.

The two methods for using estimated equations to generate probability distributions both involve solving repeatedly for annual outcomes using random draws of annual shocks. The first approach is Monte Carlo simulation, which assumes that random shocks are symmetric and follow some distribution—in this case, a normal, or bell-shaped, pattern. Thus, for each year, the mathematical equation for the normal distribution, together with computer-generated random numbers, can be used to pick values for random shocks.

The second approach is “bootstrap” simulation, so-called because it involves “picking itself up by its own bootstraps.” The idea is to use actual residuals generated during the estimation phase to perform the simulation. If 100 data points are used in the estimation, a randomly chosen shock can be selected each year of the simulation, and the probability of drawing any given historical shock is one in 100. The bootstrap approach is a useful alternative to Monte Carlo simulation because it does not require assumptions about the shape of the probability distribution for random shocks.

With Monte Carlo and bootstrap simulations, projecting forward is simple once all of the pieces are in place. Both simulations start with the last actual value, then draw a random value for the annual shock and add that shock to the coefficient multiplied by the last actual value. For the next—second—period, the process is repeated; however, this time the coefficient is multiplied by the outcome for the first simulation period. Thus, annual autocorrelations are built into the projection equation. The pro-

8. In the language of time-series econometrics, the level of real wages is not stationary, but the growth rate of real wages is.

cess is repeated for each year of the simulation period (in this case, the 75-year projection period used by SSA).

Each Monte Carlo or bootstrap simulation yields a possible set of annual outcomes for the variable in question. To move from that set of annual outcomes to a probability distribution for the outcomes in a given year, the process must be repeated many times. The most likely annual outcomes (those near the central tendency) will be realized in many more of the simulations than unlikely annual outcomes (those far away from the central tendency) will be. That principle serves as the basis for inferring probability distributions—if a given outcome occurs five out of 100 times when the random sequence is generated, the probability of that outcome is 5 percent. If some other value occurs 10 times, it is assigned a probability of 10 percent.

PROBABILITY DISTRIBUTIONS FOR DEMOGRAPHIC AND ECONOMIC INPUTS

Time-series analysis is ideal for assigning probability distributions to future values for the nine inputs used in Social Security's financial calculations. The time-series technique allows analysts to infer the likelihood of all possible outcomes solely on the basis of the historical data. Using the technique involves testing various equation specifications for random changes in central tendency, plus one other structural consideration—some variables are modeled in groups, rather than separately, because there are correlations between the outcomes that are likely to continue in the future.

In this analysis, the only input that is modeled using the simplest time-series specification is real wage growth: its model has only one equation and no correlations with other variables (because productivity is not inherently related to any of the other inputs). Mortality improvement would also be simple if separate rates were not being modeled for 21 different age groups for each sex, and if correlations between the separate age-specific error terms were not required. Unemployment, inflation, and the real interest rate are clearly related in the historical data. Thus, those three inputs are modeled in a joint process in which the outcome for any one variable directly affects the other two.

Immigration, disability incidence, and disability termination are all greatly affected by changes in law or policy. As a result, it does not seem appropriate to say that their evolution involves randomly changing central tendencies. With the time-series approach, however, that is one way in which the data can be interpreted.

In the case of fertility, past patterns suggest that outcomes are highly correlated over time, implying that shocks are temporary but last for several decades. However, a reasonable alternative interpretation is that fundamental (and permanent) changes

in the central tendency for fertility have occurred before—at the end of the baby boom, for instance—and thus may occur again.

For this analysis, CBO examined two measures of uncertainty around the expected values for the nine inputs. The first measure is the 5th and 95th percentiles of the values in each year. For example, CBO found that the rate of real wage growth in 2050 was less than -2.38 percent in 5 percent of the simulations and greater than 4.34 percent in another 5 percent of the cases. That range represents the *annual* variation for that year. The second measure of uncertainty is the 5th and 95th percentiles of the *average* value over a specific period. For example, CBO computed average wage growth between 2000 and 2050 for each simulation and then looked at the distribution of those averages.

As expected, average values vary less than annual values do. (For instance, though a reasonable chance exists that the economy will be in a depression in any given year, very little chance exists for a five-year depression and even less chance for a 20-year depression.) In the case of real wage growth, CBO found that average growth between 2000 and 2050 was less than 0.13 percent in 5 percent of the simulations and greater than 1.90 percent in another 5 percent. That range is much narrower than the -2.38 percent to 4.34 percent range for the annual variation in 2050.

The graphs that appear in the rest of this chapter illustrate uncertainty for the nine input assumptions. Those graphs include five lines:

- The solid line in the center represents SSA's intermediate projection;
- The solid lines to either side of it show the 5th and 95th percentiles of *annual* values for the 1,000 paths generated by the Monte Carlo simulations (suggesting that the outcome in any given year will fall between those bands 90 percent of the time); and
- The dotted lines show the 5th and 95th percentiles of the *average* values (from 2000 through the year in question).

For most of the variables, the range between the 5th and 95th percentile values of the 75-year averages is very similar to the range between SSA's high- and low-cost long-term values (see Table 4). Those ranges are not strictly comparable because SSA's long-term values begin as late as 2025, whereas CBO's values cover all 75 years of the projection period. Still, that similarity is striking considering that SSA's high- and low-cost values have no explicit statistical interpretation.

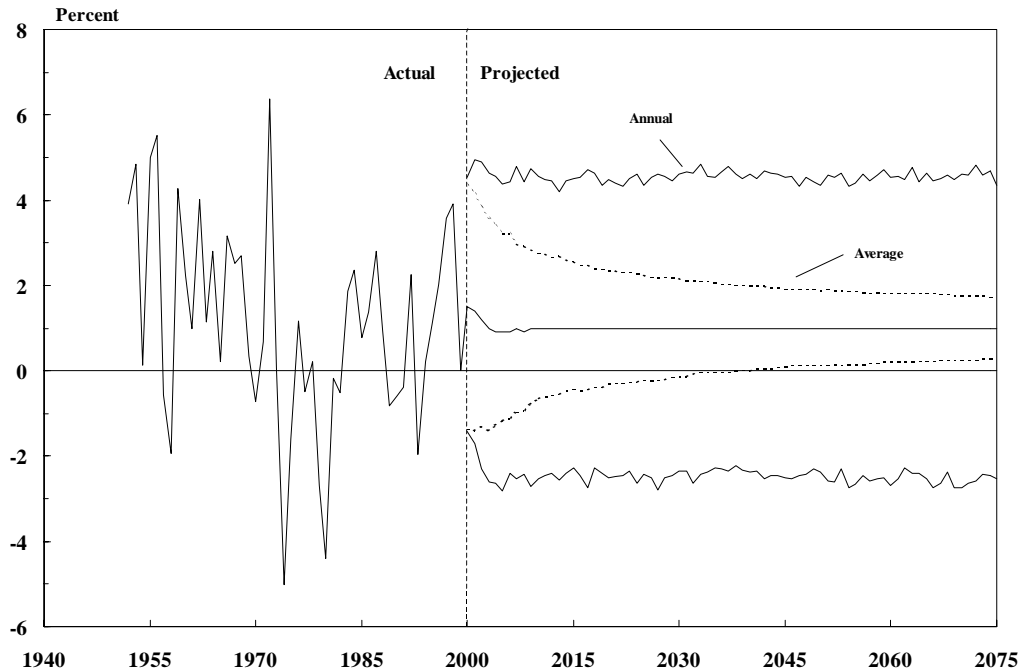
TABLE 4. RANGES OF UNCERTAINTY FOR INPUTS

Input	SSA's Expected Value	Measures of Long-Term Variation	
		SSA ^a	CBO ^b
Fertility Rate (Number of children per woman)	1.95	0.25	0.36
Rate of Mortality Improvement (Percentage reduction in the mortality rate)	0.68	0.45	0.36
Immigration Level (Thousands of people)	900	278	249
Rate of Real Wage Growth (Percent)	1.00	0.50	0.73
Inflation Rate (Percentage change in the consumer price index)	3.30	1.00	1.36
Unemployment Rate (Percent)	5.50	1.00	0.70
Real Interest Rate on Social Security Assets (Percent)	3.00	0.75	0.48
Disability Incidence Rate (Percent)	0.50 ^c	0.07 ^c	0.04
Disability Termination Rate (Percent)	3.83 ^d	0.77 ^d	0.42

SOURCES: Social Security Administration; Congressional Budget Office.

- a. SSA's variation is half of the difference between the ultimate values for each input in the high- and low-cost scenarios.
- b. CBO's variation is half of the difference between the 5th and 95th percentile values for the 2000-2075 average for each input, based on 1,000 Monte Carlo simulations using CBO's Long-Term Actuarial Model.
- c. SSA's actuaries set separate rates of change for disability incidence for men and women relative to a base period of 1980 to 1984. The rates reported here (which are adjusted for age and sex) are from CBO's Long-Term Actuarial Model (LTAM), are relative to 2000, and generate disability incidences that match those from SSA.
- d. SSA's actuaries set separate rates of change for disability termination by recovery and by death relative to a base period of 1977 to 1980; within the category of termination by death, they set separate rates for men and women. The rates reported here (which are adjusted for age and sex) are from LTAM, are relative to 2000, and generate disability terminations that match those from SSA.

FIGURE 11. UNCERTAINTY BANDS FOR THE RATE OF REAL WAGE GROWTH



SOURCES: Department of Commerce, Bureau of Economic Analysis; Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

Real Wage Growth

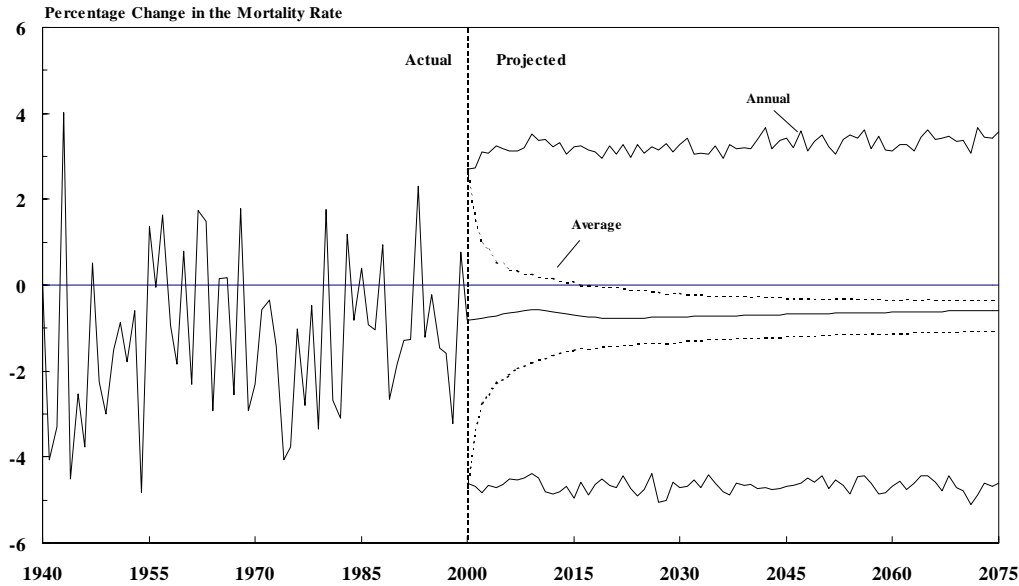
SSA's intermediate assumption for the growth of real wages is 1.0 percent per year over the 75-year projection period. The time-series technique suggests that considerable variation around that value can be expected (see Figure 11).⁹

The equation used to generate paths for real wage growth is the most basic time-series specification. The rate of real wage growth is regressed on a constant and its own lagged value. The resulting error terms pass the test for stable central tendency. Thus, no need exists to explore other specifications. (See the appendix for estimates of the coefficients and values for the various test statistics.)

The 90 percent uncertainty bands for the projection of annual real wage growth cover a range of 4 percentage points in each direction. The range for average values

9. The uncertainty bands from bootstrap simulations are very similar. The next chapter presents the bootstrap technique as an alternative to Monte Carlo simulation when solving the entire model, because using actual errors is a direct means for testing the assumption that residuals are distributed normally.

FIGURE 12. UNCERTAINTY BANDS FOR THE OVERALL RATE OF MORTALITY IMPROVEMENT



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

narrows to only about 0.7 percentage points in each direction by 2075, which is the same order of magnitude as SSA's high/low variation of 0.5 percentage points.

Mortality Improvement

SSA projects rates of mortality improvement for both men and women in each of 21 separate age groups. Historical data suggest that the rates of improvement for each sex are somewhat correlated between age groups but that differences in central tendency exist within age/sex groups and should be accounted for. Thus, CBO estimated separate time-series equations for mortality improvement in the 21 age groups of each sex, but the equations were estimated such that correlations in annual random shocks could be accommodated (see the appendix for more details).¹⁰

Like the overall average of rates of mortality improvement, which can be aggregated over age and sex to generate a graph of how mortality is expected to change, uncertainty bands can also be aggregated and graphed (see Figure 12). The 90 per-

10. The basic concept is that each variable in the system of equations is unaffected by other variables but that the error terms are potentially correlated between the equations. Correlations between errors are measured after every equation in the system is estimated.

cent range of annual outcomes around the expected rate of improvement (0.7 percent per year) is quite large and is consistent with the significant historical variation. As expected, the range of average values is much smaller.

As noted earlier, although rates of mortality improvement pass the test for non-randomness in central tendency, the level of mortality (as represented by the central death rate) can be thought of as a variable with random central tendency. In other words, the input assumption being modeled is already the first-differenced version of a variable with expanding uncertainty ranges. Thus, although the uncertainty bands for the rate of mortality improvement are constant over time, a graph of the bands for central death rates would show increasing uncertainty.

Unemployment, Inflation, and the Real Interest Rate

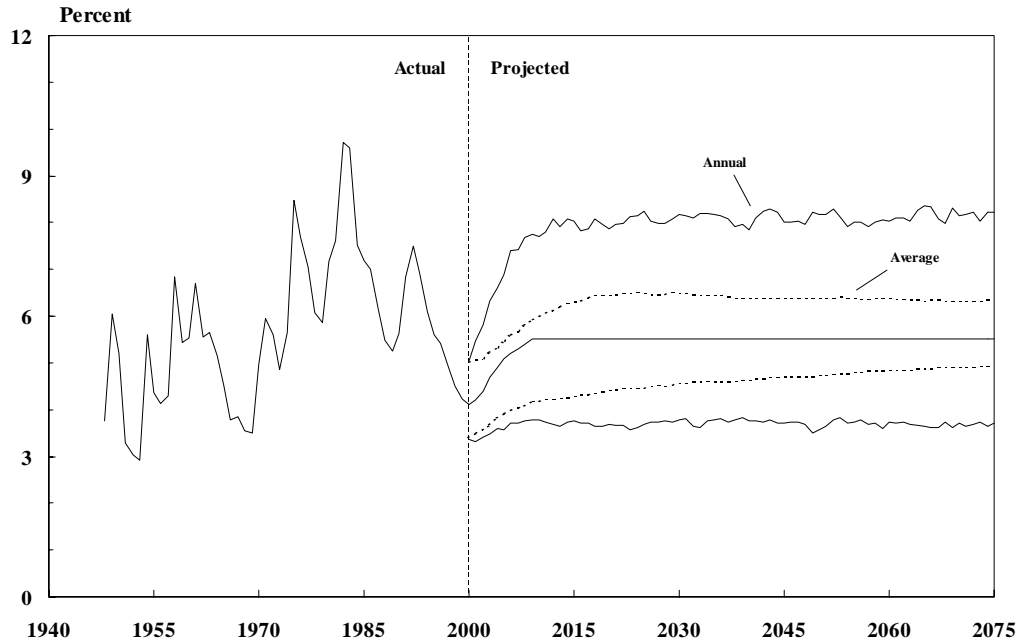
To understand why unemployment, inflation, and the real interest rate are estimated as a system of equations rather than independently, consider the effects of not doing so. In a model simulation, time-series equations for each variable would adequately generate the historical variability and correlation between annual outcomes for that variable. But the fact that outcomes among the three variables are correlated would not be captured. Thus, that technique could, in principle, violate a basic condition of model simulations and generate a *combination* of outcomes that would never have occurred historically.

As an example of such correlations, real interest rates were negative in the 1970s because inflation skyrocketed and nominal interest rates did not catch up for some time. Because no other good reason exists for real interest rates to become negative, no analyst would want to predict a negative real interest rate *unless* the underlying cause was a surge of inflation. Likewise, well-established correlations between inflation and unemployment have been drawn in the annual data; they are generally attributed to fluctuations in aggregate demand. An analyst would not want to accidentally simulate high-frequency positive correlations between inflation and unemployment when the data suggest a strong negative correlation.

The technique that CBO used to simultaneously model the three variables builds directly on the basic time-series approach. But rather than simply regressing a variable on its own lagged value, each equation includes lagged values for all of the variables under consideration.¹¹ Thus, for example, the equation for unemployment includes lagged values for unemployment, the real interest rate, and inflation over the previous two years (see the appendix for more details). The correlations between each variable and its own lagged values are generally positive. The effect of each var-

11. That technique is known as vector autoregression (VAR).

FIGURE 13. UNCERTAINTY BANDS FOR THE UNEMPLOYMENT RATE



SOURCES: Department of Labor, Bureau of Labor Statistics; Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

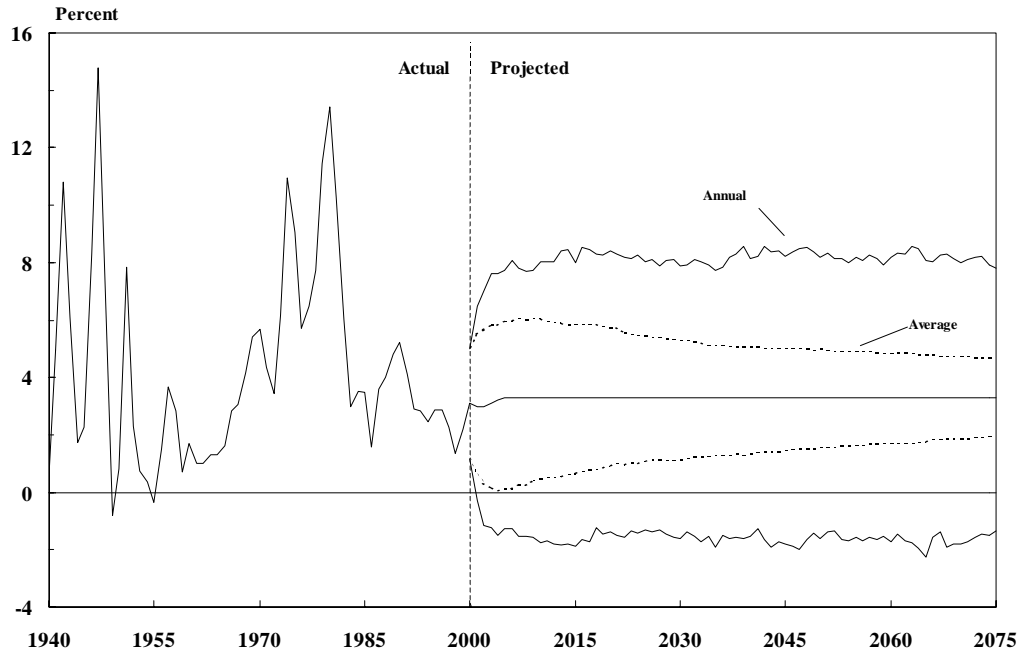
NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

iable on the other two differs over time but generally reflects well-known properties (such as the short-term negative relationship between inflation and real interest rates described above).

The uncertainty bands for annual values of unemployment, inflation, and the real interest rate are much larger than the bands for average values (see Figures 13, 14, and 15). For annual values, the range for the unemployment rate is 2 to 3 percentage points in each direction, and the ranges for inflation and the real interest rate are about 4 percentage points in each direction. For 75-year average values, the ranges for inflation and unemployment do not differ much from those in SSA's high- and low-cost scenarios (see Table 4). However, CBO's range for the real interest rate is nearly double that of SSA. (The 1999 Technical Panel of the Social Security Advisory Board suggested an even larger range—0.75 percentage points in each direction.)

The unemployment rate is the first example in this analysis of a “bounded” input (one that is naturally restricted to a certain range). If the unemployment rate was modeled simply as a level variable, random shocks that led to negative unemployment

FIGURE 14. UNCERTAINTY BANDS FOR THE INFLATION RATE



SOURCES: Department of Labor, Bureau of Labor Statistics; Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

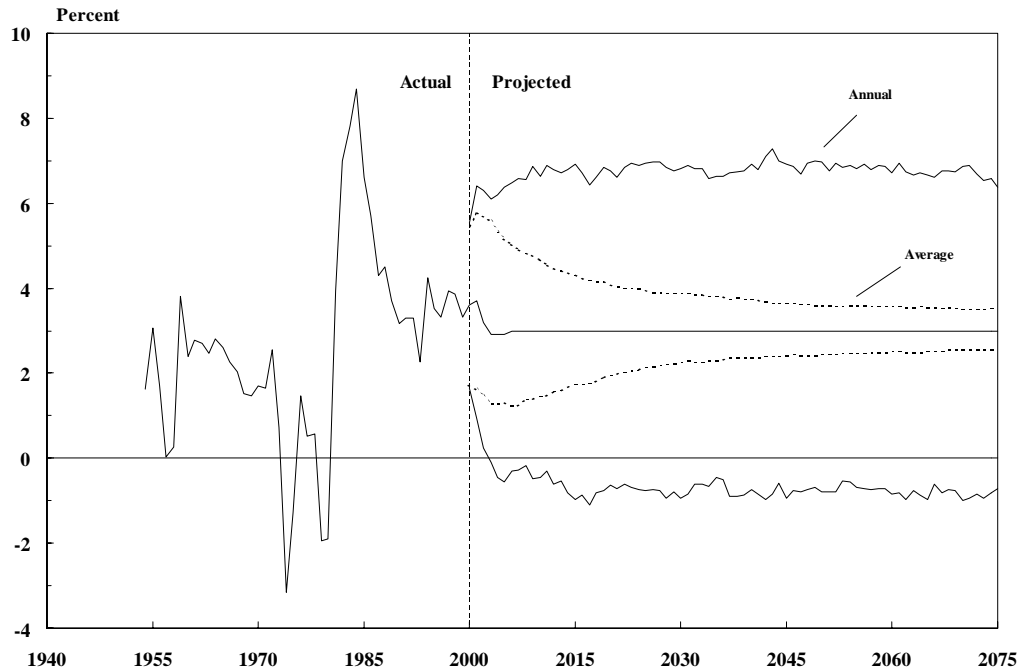
could be chosen—which is, of course, impossible in reality. Thus, the estimated equation coefficients are based on a “transformed” version of the unemployment rate that is restricted between zero and one (see the appendix).¹²

Immigration and Disability Incidence and Termination

Annual levels of legal immigration and rates at which people join and leave the Disability Insurance program are set directly in law or influenced strongly by policy changes. Historical data for each of those variables show clear indications of changing central tendencies. However, is it appropriate to think of those changes as random when they are determined to some extent by shifts in policy? The answer to that question determines which specification is appropriate for the three variables. The

12. The transformation involves taking the log-odds ratio: if u is the unemployment rate, the variable being modeled is $x = \log(u/(1-u))$. No matter what the shocks to x , the outcome of u is bounded between zero and one.

FIGURE 15. UNCERTAINTY BANDS FOR THE REAL INTEREST RATE ON ASSETS IN THE SOCIAL SECURITY TRUST FUNDS



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

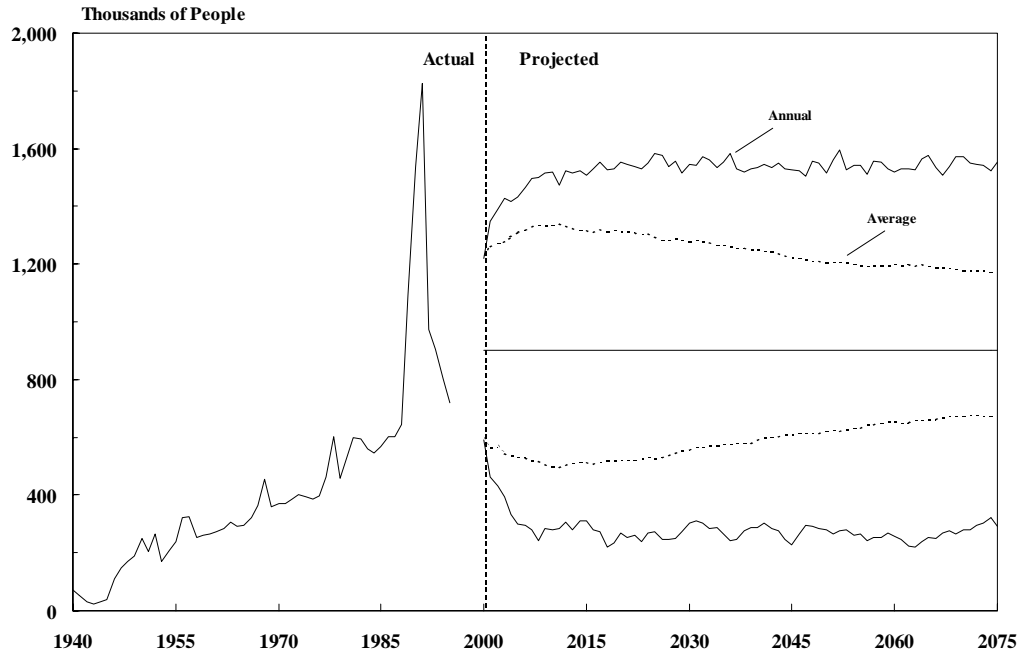
NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

approach that CBO used was to model the processes without random changes in central tendency, so that variation over time is attributed only to random shocks and correlation. To the extent that such variation results from changes in law, it will be overestimated.

Applying the standard time-series approach to those three variables produces significant uncertainty bands (see Figures 16, 17, and 18). The equation for immigration is somewhat more complicated than the standard (one-lag) time-series model because a clear trend in the level of immigration is apparent over time (see the appendix). The wide error bands for both annual and average values for immigration are consistent with the large autocorrelation, which magnifies shocks over time (although the initial shock eventually fades away).

Measuring uncertainty bands for disability incidence and termination is more difficult because of the limited data—only about 25 years' worth—available for each input. Both equations fail the test for a stable time-series variable; however, because that failure is driven by known policy changes and a short data series, CBO ignored

FIGURE 16. UNCERTAINTY BANDS FOR THE LEVEL OF IMMIGRATION



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTES: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

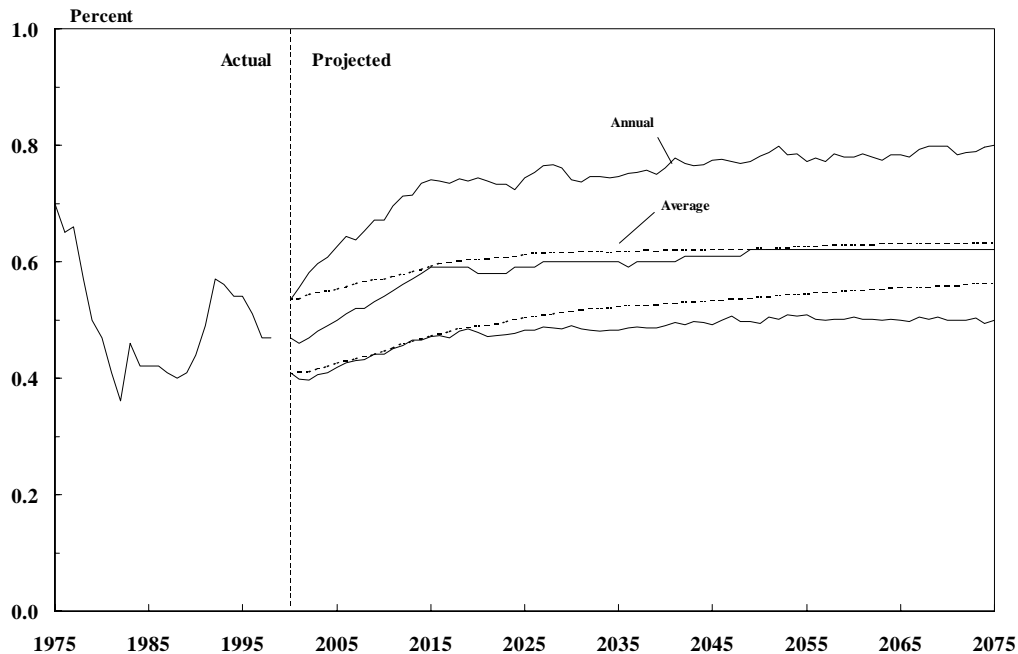
Recent historical data for the level of immigration are unavailable.

it for this analysis in order to generate fixed error bands. Those bands are quite large—for example, DI incidence roughly doubles between the high and low bands, which leads to a wide variation in estimates of Social Security's finances. (Figures 17 and 18 look different from the other figures because SSA's expected values vary. Thus, the uncertainty bands for average values do not always bracket the expected values.) In addition, the rates of both disability incidence and termination are naturally restricted between zero and one. Thus, like unemployment, those variables were estimated using a bounding transformation.

Fertility

Among the nine inputs, the fertility rate stands out in terms of its complexity and the potential randomness of its central tendency. The historical data support two different methods for modeling fertility: one without random changes in central tendency and the other with random changes. CBO assumed a base-case scenario that includes no changes in central tendency, but it also analyzed an alternative scenario with such changes.

FIGURE 17. UNCERTAINTY BANDS FOR THE RATE OF DISABILITY INCIDENCE



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTES: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

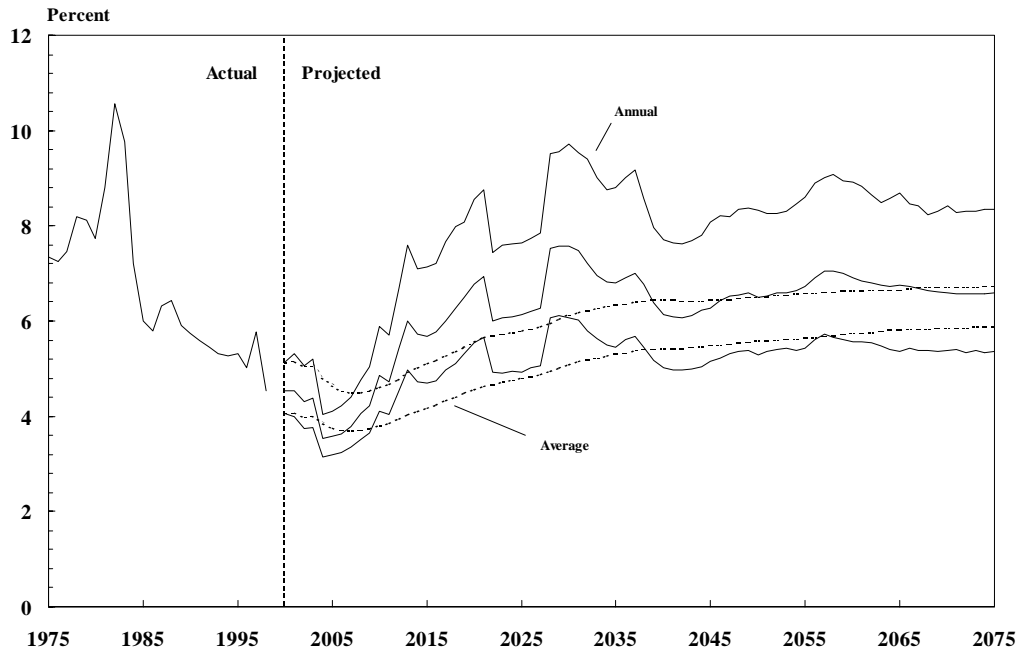
Recent historical data for the rate of disability incidence are unavailable.

Those two approaches provide different interpretations of the history of U.S. fertility rates since 1940. The base case explains the surge in fertility associated with the baby boom as a series of highly correlated shocks. An approach that assumes random changes in central tendency indicates that the baby-boom era and the post-1964 period have two very different central tendencies.

Fertility can also be modeled using a standard time-series approach that leads to stable error bands (see Figure 19).¹³ The estimated equation involves four lags for past fertility rates and a correlated error (moving-average) term (see the appendix for details). As suggested, the model explains the baby boom as a combination of annual shocks and highly correlated annual outcomes. Thus, the 90 percent range for fertility (roughly 1.0 to 3.0 children per woman) contains most of the data points associated with the baby boom.

13. Fertility is naturally bounded from below (the rate cannot drop below zero); however, using a bounding transformation requires setting limits in both directions. The uncertainty bands in Figure 19 are based on an (arbitrary) upward limit of 4.0 children per woman.

FIGURE 18. UNCERTAINTY BANDS FOR THE RATE OF DISABILITY TERMINATION



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTES: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

Recent historical data for the rate of disability termination are unavailable.

Allowing random changes in central tendency has a simple effect on the standard time-series disaggregation; however, the implications for the uncertainty bands are profound. In the random-change approach, the variable being modeled is first-differenced—a process akin to assuming perfect correlation in the levels of the variable. That means that every shock in the equation has a “permanent” effect until a shock in the other direction occurs.¹⁴ Thus, the variable tends to meander in one direction or another for long periods.

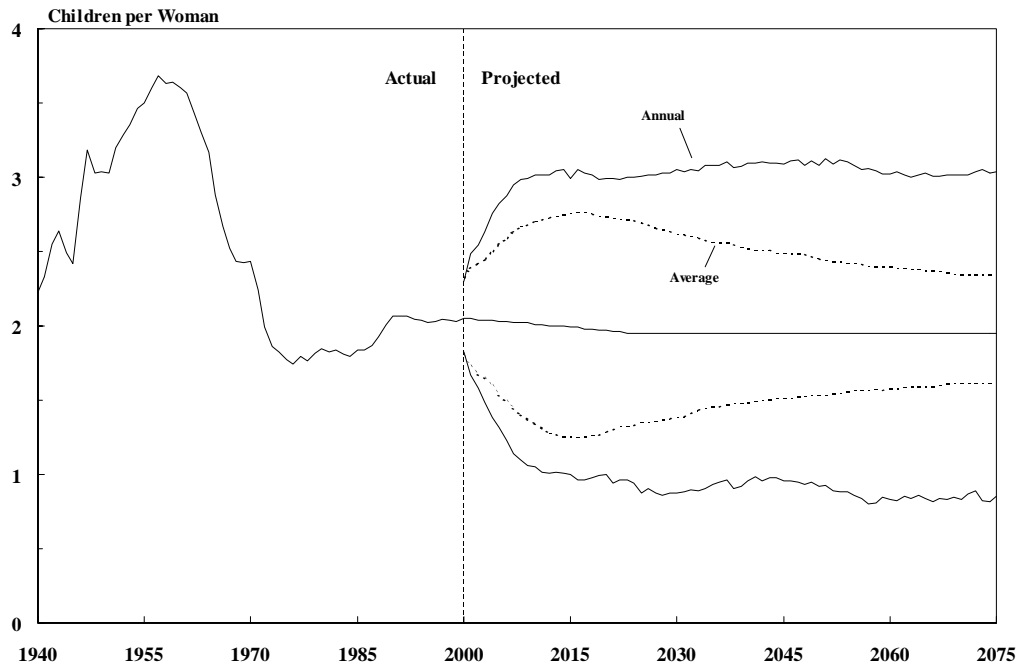
The error bands for *annual* values of fertility given a first-differenced specification show the possibility of meandering in either direction in the future (see Figure 20).

14. Building on footnote 4, the specification for a variable x_t is:

$$(x_t - x_{t-1}) = \alpha + \beta (x_{t-1} - x_{t-2}) + \varepsilon_t$$

where t denotes time, α and β are parameters to be estimated, and ε_t is the residual (unexplained error) at time t . That type of equation can also be estimated with standard regression techniques.

FIGURE 19. UNCERTAINTY BANDS FOR THE OVERALL RATE OF FERTILITY



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

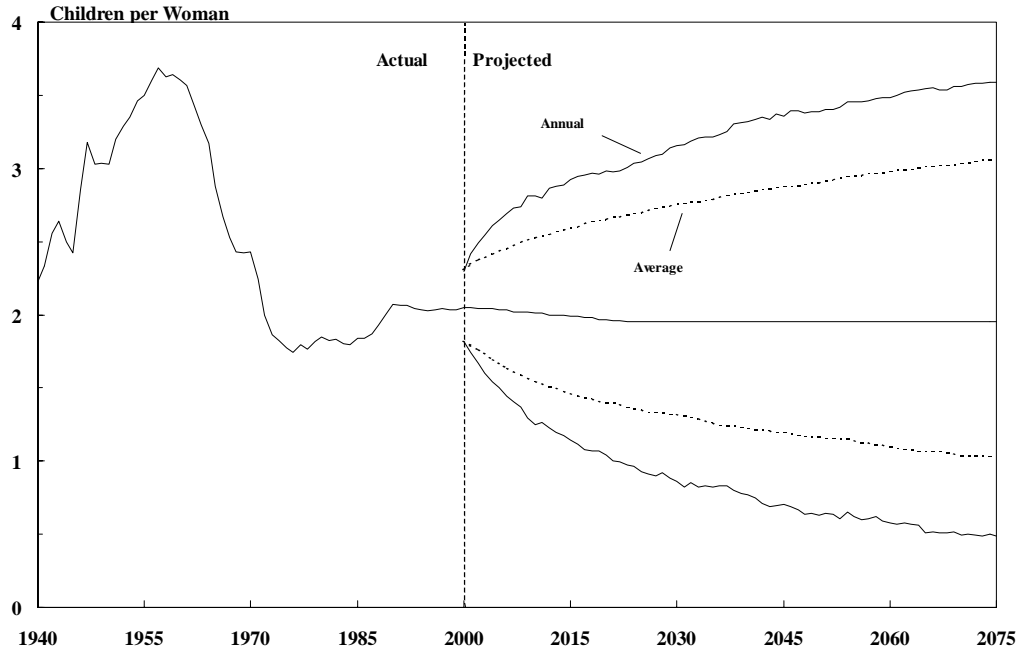
NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

The 90 percent range is much wider in 2075 than in the early years of the projection. Indeed, that band is narrower than the band of annual values for the standard specification (shown in Figure 19) in the first few decades, but it grows annually by a fixed amount, becoming wider by about 0.5 in either direction.

Even more striking is the difference in the shape of the error bands for *average* fertility values between the two specifications. For the standard specification, those bands narrow with time (see Figure 19). Thus, less uncertainty exists in the average value for 2000 through 2075 than in the average value for 2000 through 2020, for example. For the first-differenced specification, however, the bands for average values actually widen as shocks permanently affect the annual level of fertility and create scenarios with consistently higher or lower fertility than expected (see Figure 20).

After all of the estimation is complete, a sense of dissatisfaction about fertility (and possibly other inputs) remains. Using statistical techniques to forecast uncertainty in fertility implies the inclusion of a randomly changing central tendency. But how can the current annual fluctuation in the total fertility rate permanently affect all future outcomes? If that were the case, it would suggest that decisions about fertility

FIGURE 20. UNCERTAINTY BANDS FOR THE OVERALL RATE OF FERTILITY USING THE FIRST-DIFFERENCED SPECIFICATION



SOURCES: Social Security Administration, Office of the Chief Actuary; Congressional Budget Office.

NOTE: Annual uncertainty bands show the 90 percent confidence range for a given year. Average uncertainty bands show the 90 percent confidence range for the average of 2000 through a given year.

were determined by random changes in the fertility rate, which is not consistent with any logical or theoretical explanation of the level of fertility.

Another way to view the uncertainty about fertility is to look at other factors that may have caused fluctuations over time. The Depression, World War II, the great postwar economic expansion, the discovery of cheap and effective birth control—all of those events had unpredictable and dramatic effects on the fertility rate. By predicting uncertainty that is consistent with past variation, CBO is implicitly assuming that such events could happen again. Conversely, if the recurrence of those types of events is ruled out, the estimates of uncertainty about fertility presented here are truly an upper bound.

CHAPTER V
UNCERTAINTY ABOUT SOCIAL SECURITY'S
LONG-TERM FINANCES

Two different sets of statistics are commonly used to assess the long-term financial prospects for Social Security. The first set summarizes the expected adequacy of trust fund balances for a specific period of time. For example, the traditional present-value actuarial estimate basically indicates how far the system is expected to be from showing a positive balance at the end of the 75-year horizon. The second set of statistics focuses on the relationship between the program's costs and its income in any given future year. From those measures, inferences can be drawn about the path of the system's finances in that year.

Intermediate projections for both types of statistics are sobering. Actuaries at the Social Security Administration predict that the trust funds will run out of money around 2037 and that achieving solvency through 2075 would require an immediate tax increase or benefit cut equal to 1.89 percent of taxable payroll.¹ SSA's projections also imply that the program's costs will begin to exceed its income (excluding interest) around 2015 and that the gap between the two will widen to 6.2 percent of taxable payroll by 2075.

Given the length of time before the projected exhaustion of the trust funds and the dramatic widening of the cost/income gap, policymakers understandably want to know the likelihood that those dire predictions will come to pass. As explained in Chapter IV, the Congressional Budget Office uses time-series analysis to derive probability distributions for each of the nine major input assumptions that go into those projections. CBO then uses Monte Carlo or bootstrap simulations to generate random values for those inputs, and finally, it repeatedly solves its deterministic model of trust fund accumulation. That stochastic approach enables CBO to assign measures of likelihood for various ranges of trust fund outcomes, thus adding some probabilistic interpretations to SSA's basic actuarial forecasts.

The results of CBO's analysis are in some ways even more dramatic—and suggestive of impending problems—than SSA's conclusions based only on intermediate assumptions. For example, the base-case Monte Carlo simulations imply that the trust funds have only about a 1 percent chance of achieving a positive balance at the end

1. As noted in Chapter III, those estimates come from Social Security Administration, *The 2000 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds* (March 30, 2000).

of 75 years.² That conclusion is much starker than SSA's view of the range of outcomes, since even under SSA's low-cost projection, the Social Security system is expected to remain solvent in 2075.

Focusing on the gap between program costs and income over time, however, indicates a great deal of uncertainty, which suggests caution when considering the dire predictions of very long range forecasts. The 90 percent uncertainty band for the gap between costs and income in 2075 that was generated by CBO's Monte Carlo simulations is roughly the same size as the trust funds' expected deficit that year. In other words, the prediction that almost no gap will exist between costs and income and the prediction that the problem will be twice as large as SSA's intermediate value are both admissible conclusions within a 90 percent band of uncertainty.

Those differences in tone when describing the uncertainty of Social Security's finances are not inconsistent. The farther off an event, the more uncertain it generally is. Given that summary measures of the trust funds are weighted toward the near term (because of present-value discounting), events in the 75th year do not matter nearly as much as those in the next few years. Thus, it is possible to be much more certain about the present value of the system's finances than about their path 75 years from now.

CBO's stochastic approach is also useful for decomposing the uncertainty about those finances across various inputs. The contribution of any one input to overall uncertainty depends on its underlying variability (discerned from the time-series analysis) and its effect on system finances (which depends on the rules of the Social Security program). A great deal of uncertainty may exist about a particular variable, but if system finances are not significantly affected when the variable changes, its contribution to overall uncertainty will be modest. That observation may suggest ways in which Social Security can be altered to reduce financial uncertainty by modifying the rules so as to shift risk away from the trust funds and onto beneficiaries (such as by tying the retirement age to longevity).

BASE-CASE ESTIMATES OF TRUST FUND UNCERTAINTY

In assessing the health of the trust funds, analysts commonly look at three summary statistics: the income rate, the cost rate, and the actuarial balance. The income rate represents the percentage of taxable wages paid into the Social Security system (an amount basically fixed by law). The cost rate represents the percentage of taxable wages paid to beneficiaries; it is strongly affected by demographic changes, especially

2. As explained in the previous chapter, CBO's base case assumes that there is no change in central tendency for fertility.

changes in the ratio of Social Security beneficiaries to workers. The actuarial balance is the difference between the cost and income rates. Those three statistics can be computed for a particular year or used to create summary (present-value) measures of system finances.

Those statistics are often employed to describe Social Security's finances because they have convenient interpretations. The relationship between the income rate and the cost rate in a given year is a basic indicator of whether the trust funds are accumulating resources (although in the short run, they might be earning interest on past accumulations and still growing even if costs exceed income). The present-value measures indicate the size of the gap between income and costs over an entire period (including the accumulation of interest as the trust funds go through buildup and depletion phases). Expressing those measures relative to taxable payroll makes it possible to convert any gap into a statement about policy—the size of the tax increase or benefit cut needed to keep the Social Security system solvent for the next 75 years.

Differences Between Deterministic and Stochastic Expected Values

When CBO's Long-Term Actuarial Model is solved using stochastic Monte Carlo simulation, it yields an estimated long-term actuarial balance of -2.18 percent (see Table 5). That funding gap is noticeably greater than the SSA actuaries' deterministic estimate of -1.89 percent or LTAM's own deterministic estimate of -1.93 percent.³ The difference between deterministic and stochastic values occurs even though all of the input assumptions have been specified to vary *around* SSA's intermediate values.⁴ That difference shows up for all of the estimated measures—and indeed, it grows larger as one looks forward in time.

The primary reason for the difference between deterministic and stochastic estimates is asymmetry in LTAM's response to changes in inputs.⁵ For instance, when the average fertility rate is increased, system finances change by less (in absolute

3. The deterministic value from LTAM (-1.93 percent) differs from SSA's estimate (-1.89 percent) because of different approaches to rounding in the two models.

4. In the time-series equations described in the previous chapter, the constant terms are set to zero. Thus, the equation basically produces deviations around a zero mean, which are then added to the deterministic intermediate projections. An alternative strategy would be to use the estimated time-series equations to produce forecasts directly, but that would change the nature of the exercise to some extent.

5. Noah Meyerson and John Sabelhaus, "Uncertainty in Social Security Trust Fund Projections," *National Tax Journal*, vol. 53, no. 3, part 1 (September 2000), pp. 515-529.

TABLE 5. ESTIMATED UNCERTAINTY ABOUT TRUST FUND OUTCOMES
IN CBO'S BASE CASE

	Deterministic Expected Value	Stochastic Expected Value	Standard Deviation	90 Percent Range of Uncertainty
75-Year Summary Measures				
Actuarial Balance	-1.93	-2.18	1.16	-4.17 to -0.34
Cost Rate	15.40	15.67	1.29	13.62 to 17.86
Income Rate	13.47	13.49	0.16	13.24 to 13.79
Annual Measures for 2030				
Actuarial Balance	-4.30	-4.78	2.54	-9.49 to -1.00
Cost Rate	17.35	17.85	2.65	13.90 to 22.86
Income Rate	13.05	13.07	0.11	12.91 to 13.27
Annual Measures for 2075				
Actuarial Balance	-6.20	-8.20	5.26	-18.38 to -0.97
Cost Rate	19.51	21.62	5.55	14.02 to 32.43
Income Rate	13.31	13.42	0.28	13.04 to 13.98
Aged Dependency Ratio				
In 2030	35.3	35.5	2.0	32.5 to 38.9
In 2075	42.0	44.8	9.5	31.8 to 62.3

SOURCES: Social Security Administration; Congressional Budget Office.

NOTE: The stochastic expected values, standard deviations, and 90 percent ranges of uncertainty are based on 1,000 Monte Carlo simulations using CBO's Long-Term Actuarial Model.

terms) than they do when the average fertility rate is decreased by the same amount.⁶ Such asymmetries in the model's response, when combined with symmetric variation in the inputs, cause average statistics in the stochastic simulations to diverge from

6. Those asymmetries are verified by looking at the sensitivity analysis in Table 2. For example, the improvement in the 75-year actuarial balance is 2.27 percent of taxable payroll when all inputs are set to the Social Security Administration's low-cost values, but the deterioration is -3.11 percent when the input values are set to SSA's high-cost levels.

those in the deterministic simulations even though the average values for the inputs are the same.

Uncertainty About the 75-Year Summary Actuarial Balance

This analysis measures the uncertainty associated with 75-year summary statistics in two ways: using the standard deviation and the 90 percent range of uncertainty (or confidence range). The standard deviation is a common measure used to describe the variability in a particular estimate; it is interpreted as being the plus-or-minus figure within which the outcome will lie about two-thirds of the time. The 90 percent range of uncertainty marks the values that 5 percent of the outcomes are below and 5 percent are above.

In CBO's stochastic simulations, 90 percent of the 75-year actuarial balances lie within a range of -0.34 percent to -4.17 percent (see Table 5). Therefore, the trust funds face a high probability of being exhausted before the end of the 75-year period. The chief source of uncertainty about the actuarial balance is the 75-year cost rate; with a probability of 90 percent, it ranges between 13.62 percent and 17.86 percent. The 75-year income rate, by contrast, has a fairly narrow range (13.24 percent to 13.79 percent). Because the tax rates for Social Security are fixed, the only variation in the income rate arises from the taxation of Social Security benefits.

Increasing Variability of Annual Actuarial Balances

Currently, the Social Security program's cost rate is below its income rate, which means that the trust funds are accumulating resources. By 2030, however, the cost rate is expected to exceed the income rate by 4.3 percent of taxable payroll in the intermediate deterministic case. That gap is predicted to widen to 6.2 percent by 2075. The stochastically generated cost and income rates show an even wider gap: 8.2 percent of taxable payroll by 2075. (Annual stochastic estimates diverge from their equivalent deterministic estimates for the same reason that 75-year summary measures do: the model's asymmetric response to symmetrically distributed inputs. But the difference is even more noticeable with annual values than with summary ones.)

That widening gap results from expected growth in the cost rate over time, which in turn reflects the increasing burden that Social Security is expected to bear as the baby-boom generation retires. It also reflects the combination of increased longevity and lower fertility that is expected to permanently increase the ratio of beneficiaries to workers, even after the baby boomers die off.

Besides rising, the annual cost rate is expected to grow increasingly variable over time (thus boosting the variability of the actuarial balance). That change is projected to be dramatic. As measured by the standard deviation, the variability of the annual cost rate more than doubles between 2030 and 2075, while the expected value rises by only about one-fifth (see Table 5).

One simple factor drives that dramatic growth in the amount of variation in the forecast: eventually, the random outcomes of the nine input processes will become the foundation for generating future random inputs. In 75 years, the cohort born in 2001 will have worked, had children who might also be working, and possibly retired or even died. Randomly changing the fertility rate in 2001 produces great variation in the size of the cost rate 75 years later, because in that case, the total amount of Social Security benefits will include a randomly generated number of beneficiaries as well as the randomly drawn inputs that were used at the beginning of the projection (mortality, real wage growth, inflation, and so forth).

That principle—that early variation results in even higher future variation—is most clearly seen by looking at how variation in the aged dependency ratio grows over time (see Table 5).⁷ Although the expected ratio increases from 35.5 percent to 44.8 percent between 2030 and 2075, the standard deviation of the ratio rises much more dramatically, from 2.0 percent to 9.5 percent. (The aged dependency ratio in 2030 can be predicted with a fair amount of certainty—after all, most of the people who make up that ratio are already alive, and mortality is the only real source of uncertainty.)

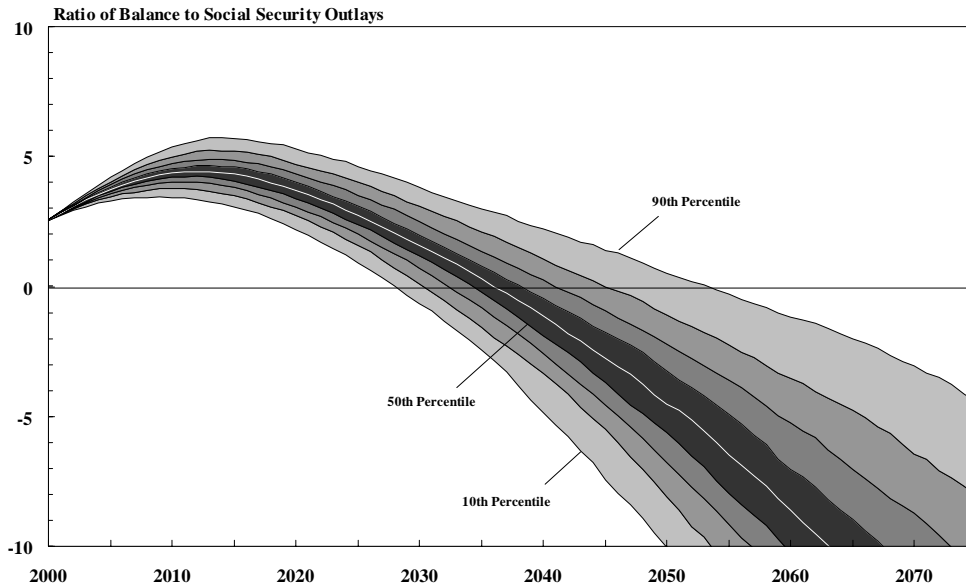
The Likelihood of Trust Fund Exhaustion Over Time

In addition to the standard statistics (cost rates, income rates, and actuarial balances), a policymaker may want to know how likely it is that the trust funds will remain solvent through a particular period. The size of the trust funds can be looked at in a number of ways—a preferred method is to relate the balance in any given year to Social Security's spending that year (the trust fund ratio). In both SSA's intermediate deterministic projections and CBO's stochastic projections, the trust fund ratio is expected to rise for the next decade or so and then begin to fall as the trust funds are depleted (see Figure 21).

As it does with other statistics, the stochastic approach allows analysts to assign probabilistic interpretations to the possible path for trust fund ratios. In CBO's base-case Monte Carlo simulations, a 90 percent chance exists that the funds will be sol-

7. The aged dependency ratio measures the number of people age 65 or older relative to the population between ages 20 and 64.

FIGURE 21. PROBABILITY DISTRIBUTION OF THE BALANCE OF THE SOCIAL SECURITY TRUST FUNDS



SOURCE: Congressional Budget Office.

NOTE: In each year, there is a 10 percent chance that the balance will fall within each shaded band, a 10 percent chance that it will fall below the lowest band, and a 10 percent chance that it will fall above the highest band.

vent at least through 2029 (in Figure 21, that is the year in which the 10th percentile path crosses the zero line). The expected year of exhaustion—based on the 50th percentile trajectory—is 2036, basically the same as in SSA's deterministic case. At the other extreme, the trust funds have only a 10 percent chance of remaining solvent through 2054 (where the 90th percentile trajectory hits zero) or beyond.

Note that in Figure 21, the distribution of trust fund exhaustion dates is skewed to the right. That skew is attributable to the nonlinear combinations of the inputs that will have a postponed effect on the trust funds. In particular, the variation in the average level of fertility will affect the size of the trust funds over time. High average fertility in the early years of the projection prolongs solvency by adding more workers to the system after 20 years. The response is asymmetric, in that an equal reduction in fertility will not make the trust funds go broke much faster than is expected in the worst case—the reason is that the workforce in 2029 will be composed mostly of people who have already been born.

TABLE 6. SOURCES OF UNCERTAINTY ABOUT SOCIAL SECURITY'S FINANCES

	<u>75-Year Actuarial Balance</u>		<u>Cost Rate in 2075</u>	
	Expected Value for Actuarial Balance ^a	Standard Deviation	Expected Value for Cost Rate ^b	Standard Deviation
Effect of Changing Individual Inputs				
Fertility	-1.94	0.63	20.11	3.17
Mortality improvement	-2.01	0.38	19.77	1.00
Immigration	-1.93	0.10	19.52	0.38
Real wage growth	-1.92	0.64	19.93	2.69
Inflation	-2.00	0.39	20.17	2.24
Unemployment	-1.94	0.04	19.53	0.32
Real interest rate	-1.94	0.24	19.51	n.a.
Disability incidence	-1.99	0.17	19.60	0.46
Disability termination	-1.91	0.06	19.47	0.21
Effect of Changing All Inputs Together	-2.18	1.16	21.62	5.55

SOURCES: Social Security Administration; Congressional Budget Office.

NOTES: Based on 1,000 Monte Carlo simulations using CBO's Long-Term Actuarial Model.

n.a. = not applicable.

a. The 75-year actuarial balance in SSA's intermediate projection published in 2000 is -1.89 percent of taxable payroll.

b. The 2075 cost rate in SSA's intermediate projection published in 2000 is 19.51 percent.

SOURCES OF TRUST FUND VARIABILITY

The Monte Carlo simulation technique can be applied to any combination of the nine inputs, including one at a time. That exercise is useful because it indicates how much of the overall variability is attributable to each input. Breaking down uncertainty by source shows whether the specification of any given time-series equation significantly affects the results. That process is also helpful in considering policy changes that are designed to lessen the financial risk in the Social Security system.

How Much Trust Fund Uncertainty Is Attributable to Each Input?

Holding all other variables fixed at their expected values, the rates of real wage growth and fertility cause the most variation in the 75-year actuarial balance and the 2075 cost rate (see Table 6). Indeed, together those two inputs account for about

half of the total uncertainty. Variability in the rates of mortality improvement and inflation also have large effects, but each effect is still only about half as large as the impacts of real wage growth and fertility.

When only real wage growth or fertility is allowed to vary randomly, the 75-year actuarial balance has a standard deviation of about 0.64 percent. But when every input varies, the standard deviation is 1.16 percent, which is smaller even than the sum of those two individual changes. That relationship indicates an important feature of the Monte Carlo simulation process—it is highly unlikely to generate especially bad or good draws for all of the inputs at the same time (the assumption used in SSA's approach of high- and low-cost scenarios). Rather, the simulations tend to produce some high and some low draws, which cancel each other out. Thus, the overall standard deviation is much less than the sum of the individual standard deviations when each input is run separately.

The effect of uncertainty about fertility and real wage growth increases dramatically looking forward in time; uncertainty about inflation also becomes important. Tying that finding back to the discussion in the previous chapter, real wage growth and inflation represent bounded changes in other variables (wage levels and price indexes, respectively) that are really meandering. Because it is the levels of wages and prices that actually affect the system, bounded uncertainty in the changes is consistent with expanding ranges of variability in the levels over time. Except for the real interest rate (which does not affect annual cost rates), the variables follow the same order of magnitude in their effect on the 2075 cost rate as they do in their effect on the 75-year actuarial balance.

Comparing Time-Series and Scenario-Based Sources of Uncertainty

The stochastic and deterministic models yield strikingly different effects on the uncertainty of the 75-year actuarial balance for three inputs: fertility, mortality, and the real interest rate. Those different results arise from the difference between allowing each input to move annually according to its historical distribution and projecting inputs without annual fluctuations.

- *Fertility.* In the stochastic model, fertility has nearly twice as great an impact on the uncertainty of the 75-year actuarial balance as it does in SSA's high- and low-cost analysis. That difference occurs because the stochastic projection of fertility is based on data from 1917 to 1998, whereas the deterministic model (at least implicitly) bases the uncertainty of fertility on a much shorter range.
- *Mortality.* Variation in the rate of mortality improvement causes only half as much variation in the 75-year actuarial balance in the stochastic model as

it does in the high- and low-cost approach. That disparity is probably attributable to the difference in how the two approaches project the reduction in mortality rates. The deterministic model assumes that the reduction will be perfectly correlated across age/sex groups. The stochastic model assumes that there will be some correlation across age/sex groups, but because that correlation is not perfect, it also assumes that there will be some offsetting differences among the groups. For example, even when a high overall rate of mortality improvement is projected, some age/sex groups may experience a smaller reduction, or even an increase, in mortality rates. Thus, the overall variation in rates for different scenarios will be smaller in the stochastic model than in the deterministic model.

- *Real Interest Rate.* This variable produces much less variation in the 75-year actuarial balance under the stochastic model than under the high- and low-cost scenarios.

The rest of the inputs have almost the same effect on the uncertainty of the 75-year actuarial balance in both approaches. However, for inflation and real wage growth, the differences are small but noticeable.

THE EFFECT OF RANDOM CHANGES IN THE CENTRAL TENDENCY FOR FERTILITY

When the assumption of a fixed central tendency for fertility is modified to allow for random changes in central tendency, the effect on the uncertainty of the 75-year summary statistics is modest. The uncertainty of the 2075 cost rate, by contrast, increases by nearly three-quarters (see Table 7).

Allowing random changes in the central tendency for fertility boosts the variability of the 75-year summary cost rate by 12 percent (the standard deviation rises from 1.29 percent of taxable payroll to 1.45 percent). When all of the other inputs are held fixed at the values expected by SSA, that change also increases the standard deviation of the 75-year actuarial balance by 40 percent (from 0.63 percent to 0.88 percent)—making fertility the most important source of uncertainty. That increase in variability suggests that the 75-year funding gap could become even wider (the new lower bound of the 90 percent uncertainty band for the actuarial balance is -4.26 percent) or could disappear entirely (the new upper bound is 0.00).

The variability of the cost rate in 2075 rises by a much larger amount: 71 percent (the standard deviation grows from 5.55 percent to 9.48 percent). Moreover, when all of the other inputs are held fixed at their deterministic expected values, the variability of the 2075 cost rate jumps by 235 percent (the standard deviation is 7.46 per-

TABLE 7. ESTIMATED UNCERTAINTY ABOUT TRUST FUND OUTCOMES ASSUMING RANDOM CHANGES IN THE CENTRAL TENDENCY FOR FERTILITY

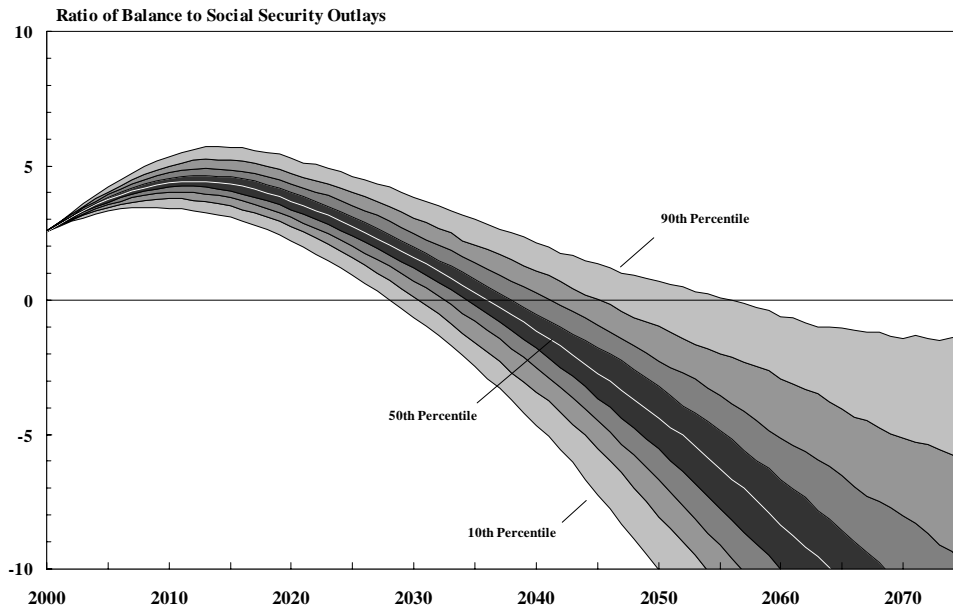
	Base Case (Fixed central tendency)		Random Central Tendency	
	Standard Deviation	90 Percent Range of Uncertainty	Standard Deviation	90 Percent Range of Uncertainty
75-Year Summary Measures				
Actuarial Balance	1.16	-4.17 to -0.34	1.31	-4.26 to 0.00
Cost Rate	1.29	13.62 to 17.86	1.45	13.28 to 17.97
Income Rate	0.16	13.24 to 13.79	0.17	13.22 to 13.79
Annual Measures for 2030				
Actuarial Balance	2.54	-9.49 to -1.00	2.43	-9.31 to -1.28
Cost Rate	2.65	13.90 to 22.86	2.53	14.20 to 22.61
Income Rate	0.11	12.91 to 13.27	0.10	12.92 to 13.26
Annual Measures for 2075				
Actuarial Balance	5.26	-18.38 to -0.97	9.00	-26.11 to 1.50
Cost Rate	5.55	14.02 to 32.43	9.48	11.42 to 40.57
Income Rate	0.28	13.04 to 13.98	0.49	12.90 to 14.38
Aged Dependency Ratio				
In 2030	2.0	32.5 to 38.9	1.7	32.8 to 38.4
In 2075	9.5	31.8 to 62.3	19.5	24.5 to 84.1

SOURCE: Congressional Budget Office.

NOTE: Based on 1,000 Monte Carlo simulations using CBO's Long-Term Actuarial Model.

cent, compared with 3.17 percent in the base case). Allowing the central tendency for fertility to move randomly produces larger variability in the annual cost rates and actuarial balances for later years of the projection period than for earlier years. The distribution of trust fund ratios looks very similar to that of the base case, but a slight outward shift occurs in the 90th percentile, from exhaustion in 2054 to exhaustion in 2056 (see Figure 22).

FIGURE 22. PROBABILITY DISTRIBUTION OF THE BALANCE OF THE SOCIAL SECURITY TRUST FUNDS ASSUMING RANDOM CHANGES IN THE CENTRAL TENDENCY FOR FERTILITY



SOURCE: Congressional Budget Office.

NOTE: In each year, there is a 10 percent chance that the balance will fall within each shaded band, a 10 percent chance that it will fall below the lowest band, and a 10 percent chance that it will fall above the highest band.

THE EFFECT OF USING BOOTSTRAPPED RANDOM SHOCKS

A final sensitivity analysis is based on bootstrap simulations. The idea behind performing this analysis is that, in the Monte Carlo simulations, the distribution of the random shocks might not be normal, and thus the assumption of normality pushes the results in one direction or another. The problem with bootstrap simulations, however, is that they require making random picks from the list of historical shocks—that is, from the actual error terms of the original estimation. If the time period used for that estimation was short, the sample of error terms to choose from will be small.

The impact of shifting to bootstrapped random shocks is fairly modest (see Table 8). The standard deviation of the 75-year actuarial balance rises from 1.16 to 1.32 percent, and the 90 percent uncertainty range expands by about 0.3 percent of taxable payroll in either direction. The standard deviation of the 75-year cost rate does not change disproportionately.

TABLE 8. ESTIMATED UNCERTAINTY ABOUT TRUST FUND OUTCOMES
USING BOOTSTRAPPED RANDOM SHOCKS

	Base Case (Normal random shocks)		Bootstrapped Random Shocks	
	Standard Deviation	90 Percent Range of Uncertainty	Standard Deviation	90 Percent Range of Uncertainty
75-Year Summary Measures				
Actuarial Balance	1.16	-4.17 to -0.34	1.32	-4.47 to -0.10
Cost Rate	1.29	13.62 to 17.86	1.47	13.36 to 18.21
Income Rate	0.16	13.24 to 13.79	0.17	13.23 to 13.80
Annual Measures for 2030				
Actuarial Balance	2.54	-9.49 to -1.00	2.60	-9.37 to -1.06
Cost Rate	2.65	13.90 to 22.86	2.71	13.98 to 22.64
Income Rate	0.11	12.91 to 13.27	0.11	12.91 to 13.27
Annual Measures for 2075				
Actuarial Balance	5.26	-18.38 to -0.97	5.47	-18.88 to -1.05
Cost Rate	5.55	14.02 to 32.43	5.76	14.13 to 33.03
Income Rate	0.28	13.04 to 13.98	0.30	13.03 to 14.01
Aged Dependency Ratio				
In 2030	2.0	32.5 to 38.9	2.3	31.8 to 39.3
In 2075	9.5	31.8 to 62.3	9.9	30.3 to 62.0

SOURCE: Congressional Budget Office.

NOTE: Based on 1,000 Monte Carlo and bootstrapped simulations using CBO's Long-Term Actuarial Model.

APPENDIX

ESTIMATES OF TIME-SERIES EQUATIONS

FOR INPUT ASSUMPTIONS

In the Congressional Budget Office’s analysis, the equations for each input are estimated according to basic techniques of time-series econometrics.¹ For each series, the goal is to find an equation that sufficiently captures the properties of the historical time series and employs a limited number of variables to yield a plausible fit for the input variable. Two sets of tests are performed to guarantee that the historical time series is stationary and that the residuals of the fitted equation are white noise. In this analysis, the only inputs that fail to pass tests of stationarity are fertility and the variables controlled primarily by law (disability incidence, disability termination, and immigration). For a few age groups, the residuals of the time-series equations for mortality improvement fail to pass the test for white noise.

Currently, the possibility of nonstationarity is recognized only in CBO’s estimates of uncertainty in the fertility projection. The other inputs that may contain nonstationarities are estimated as stationary processes, because random changes in their central tendencies are strongly influenced by changes in law. CBO plans to pursue improvements in how the system of equations for mortality measures that variable; for the present, however, it felt that a “corrected” model would produce results similar to the equations now in place.

Each variable has been estimated using either an AR(1) process, a vector autoregression model, or an ARMA model involving three or four autoregressive variables along with a moving-average representation of the annual fluctuations.²

- *Real Wage Growth.* This input was estimated according to an AR(1) process, such that

$$x_t = 0.006 + 0.539 \cdot x_{t-1} + \varepsilon_t$$

where x_t represents real wage growth and ε_t represents the random variable that describes the annual random shocks to real wage growth and has a standard deviation of 0.018. The p-value for the Dickey-Fuller test is 0.003,

1. See, for example, Walter Enders, *Applied Econometric Time Series* (New York: John Wiley & Sons, 1995).

2. To preserve the series with an expected value set to the Social Security Administration’s intermediate assumption, each result is expressed in deviation form. The deviation is forecast and then added to the deterministic expected value.

TABLE A-1. ESTIMATED COEFFICIENTS FOR MORTALITY REDUCTIONS

Age Range	Intercept	Beta	Sigma	P-Value for Test of Unit Root	P-Value for Test of White Noise
Male Mortality					
Under 1	-0.036	-0.244	0.042	0.0001	0.996
1 to 4	-0.050	-0.419	0.083	0.0001	0.342
5 to 9	-0.035	-0.227	0.074	0.0001	0.326
10 to 14	-0.025	-0.234	0.085	0.0001	0.785
15 to 19	-0.008	-0.198	0.137	0.0001	0.404
20 to 24	-0.008	-0.176	0.146	0.0001	0.332
25 to 29	-0.005	-0.219	0.180	0.0001	0.461
30 to 34	-0.006	-0.229	0.158	0.0001	0.524
35 to 39	-0.011	-0.240	0.099	0.0001	0.303
40 to 44	-0.012	-0.162	0.061	0.0001	0.411
45 to 49	-0.012	-0.161	0.044	0.0001	0.263
50 to 54	-0.010	-0.099	0.037	0.0001	0.042*
55 to 59	-0.008	-0.040	0.034	0.0009	0.592
60 to 64	-0.008	-0.130	0.032	0.0001	0.028*
65 to 69	-0.007	-0.235	0.031	0.0001	0.168
70 to 74	-0.007	-0.282	0.031	0.0001	0.019*
75 to 79	-0.006	-0.260	0.033	0.0001	0.229
80 to 84	-0.007	-0.329	0.037	0.0001	0.001*
85 to 89	-0.005	-0.229	0.041	0.0001	0.010*
90 to 94	-0.003	-0.256	0.042	0.0001	0.018*
95 and Over	-0.002	-0.251	0.045	0.0001	0.064

(Continued)

which indicates the rejection of the presence of a unit root. The p-value for the Ljung-Box test for white noise of the residuals is 0.849.

- *Mortality Improvement.* Over a range of 42 age and sex groups, the ratio of the percentage reduction between the current year's central death rate (by age and sex) and the previous year's central death rate is estimated according to an AR(1) process. (See Table A-1 for the coefficients as well as for the p-values for the tests of unit root and white noise for the residuals.) The covariance of the annual random shocks is also calculated and used during the simulation process. For instance, the correlation between mortality reductions for males ages 10 to 14 and males ages 15 to 19 is 0.93. Because the deterministic

TABLE A-1. CONTINUED

Age Range	Intercept	Beta	Sigma	P-Value for Test of Unit Root	P-Value for Test of White Noise
Female Mortality					
Under 1	-0.036	-0.272	0.043	0.0001	0.997
1 to 4	-0.051	-0.429	0.090	0.0001	0.222
5 to 9	-0.038	-0.266	0.090	0.0001	0.502
10 to 14	-0.027	-0.188	0.110	0.0001	0.685
15 to 19	-0.021	-0.166	0.116	0.0001	0.851
20 to 24	-0.023	-0.199	0.147	0.0001	0.874
25 to 29	-0.018	-0.186	0.179	0.0001	0.750
30 to 34	-0.021	-0.222	0.135	0.0001	0.687
35 to 39	-0.022	-0.239	0.080	0.0001	0.772
40 to 44	-0.022	-0.230	0.049	0.0001	0.929
45 to 49	-0.019	-0.261	0.036	0.0001	0.777
50 to 54	-0.017	-0.252	0.032	0.0001	0.455
55 to 59	-0.015	-0.220	0.031	0.0001	0.700
60 to 64	-0.015	-0.345	0.027	0.0001	0.063
65 to 69	-0.012	-0.222	0.029	0.0001	0.125
70 to 74	-0.013	-0.274	0.031	0.0001	0.003*
75 to 79	-0.012	-0.283	0.036	0.0001	0.029*
80 to 84	-0.011	-0.299	0.039	0.0001	0.000*
85 to 89	-0.008	-0.185	0.045	0.0001	0.006*
90 to 94	-0.007	-0.229	0.044	0.0001	0.014*
95 and Over	-0.005	-0.237	0.042	0.0001	0.115

SOURCES: Social Security Administration; Congressional Budget Office.

NOTE: * = fails to pass test for white noise.

model does not include that estimated covariation, mortality varies far more in that model than in the stochastic model. To simulate the annual random shocks and the covariance among them, a random vector of 21 normal random errors is generated for each sex from a random-number generator. The 21 normal random errors are then transformed according to the variance-covariance structure of the errors by multiplying the vector of errors by the Cholesky vector, which is the triangular decomposition of the variance-covariance matrix of the random shocks.

TABLE A-2. ESTIMATED COEFFICIENTS FOR CBO'S ECONOMIC MODEL

Variable	Dependent Variable		
	Unemployment	Inflation	Real Interest Rate
Intercept	-1.074	-0.004	0.029
Unemployment _{t-1}	0.885	-0.079	0.058
Unemployment _{t-2}	-0.204	0.074	-0.048
Inflation _{t-1}	6.727	0.571	0.246
Inflation _{t-2}	-3.589	0.354	-0.191
Real Interest _{t-1}	3.383	-0.683	1.164
Real Interest _{t-2}	-1.433	0.424	-0.263
Sigma	0.134	0.013	0.013
P-Value for Test of Unit Root	0.025	0.001	0.043
P-Value for Test of White Noise	0.281	0.222	0.313

SOURCES: Social Security Administration; Congressional Budget Office.

- Economic Variables.* The economic variables of unemployment, inflation, and the real interest rate are estimated together in a VAR model, such that each variable is a function of its own previous values as well as the previous values of the other two variables.³ In each series, the residuals appear to be white noise, and the series is stationary. As in the mortality projection, the variance and covariance of the random shocks of the three variables are estimated in order to have plausible comovements between the economic variables (see Table A-2).
- Immigration.* Immigration fails its test for stationarity. It is estimated without a correction for that property as an ARMA(4,1) model:

$$x_t = 487,552 + 0.956 \cdot x_{t-1} - 0.278 \cdot x_{t-2} + 0.316 \cdot x_{t-3} - 0.121 \cdot x_{t-4} + \varepsilon_t - 0.036 \cdot \varepsilon_{t-1}$$

3. Unemployment rates are expressed as a log-odds ratio in order to bound them between 0 and 1.

where the standard deviation of the annual random shock is 187,869 people. The tests for stationarity reveal that the series fails to reject the presence of a unit root at the 5 percent level but rejects it at the 10 percent level (p-value = 0.097). The residuals of the model are white noise, based on a p-value of 0.092.

- *Disability Incidence and Termination.* Both of these variables fail their tests for stationarity. They are estimated without a correction for that property as an AR(1) model. For the model of disability incidence,

$$x_t = -0.796 + 0.732 \cdot x_{t-1} + \varepsilon_t$$

where the standard deviation is equal to 0.128, the p-value for the unit-root test is 0.322, and the p-value for white noise of the residuals is 0.592. For the model of disability termination,

$$x_t = 0.49 + 0.747 \cdot x_{t-1} + \varepsilon_t$$

where the standard deviation of the annual random shock is 0.146, the p-value for the unit-root test is 0.96, and the p-value for white noise of the residuals is 0.67.⁴

- *Fertility.* Two models of fertility are estimated. In both, the annual level of fertility is transformed logarithmically so the projected values of the annual total fertility level will lie between zero and four, using the formula $x_t = \log(\text{TFR}_t / (4 - \text{TFR}_t))$. After the transformed projected values are calculated, they are converted into annual total fertility rates according to the formula $\text{TFR}_t = 4 \cdot \exp(x_t) / (1 + \exp(x_t))$.⁵ The first model does not allow for the possibility that the time series is nonstationary—an approach similar to what has been used in other projections of fertility, where it has been argued that allowing for nonstationarity may not be appropriate for creating forecasts of the level of fertility. On the basis of that argument, the first model is used in the baseline projection of the trust funds. Fertility is estimated according to an ARMA (4,1) process,

$$x_t = 0.687 + 1.802 \cdot x_{t-1} - 1.189 \cdot x_{t-2} + 0.787 \cdot x_{t-3} - 0.429 \cdot x_{t-4} + \varepsilon_t - 0.567 \cdot \varepsilon_{t-1}$$

4. The Long-Term Actuarial Model uses a disability retention rate, which is simply (1 - termination rate).

5. Ronald D. Lee, "Modeling and Forecasting the Time Series of U.S. Fertility: Age Distribution, Range, and Ultimate Level," *International Journal of Forecasting*, vol. 9 (1993), pp. 187-202.

where the standard deviation of the random shock is 0.136, the p-value for the unit-root test is 0.530, and the p-value for white noise of the residuals is 0.196. The second model explicitly acknowledges the presence of stochastic trends, and the change in the transformed fertility level is estimated as an AR(3) process,

$$\Delta x_t = 0.375 \cdot \Delta x_{t-1} - 0.155 \Delta x_{t-2} + 0.33 \cdot \Delta x_{t-3} + \varepsilon_t$$

where the standard deviation of the annual random shock is 0.001, the p-value for the unit-root test is 0.96, and the p-value for white noise of the residuals is 0.121.



This paper and other CBO publications are
available at CBO's Web site:
www.cbo.gov