Methods for Stochastic Trust Fund Projection*

Martin R. Holmer†

January 2003

Abstract

Current methods used for stochastic projection of the finances of the combined Old-Age and Survivors Insurance and Disability Insurance (OASDI) trust fund are described and found to have problems. Structural time series models and alternative Monte Carlo procedures are identified as new methods that may be able to solve these problems. The new methods are implemented using the productivity growth rate and total fertility rate as examples. Experience from the two examples provides some preliminary indications about how a switch to the new methods affects implementation effort and substantive results concerning the degree of uncertainty in trust fund finances.

*The research reported in this paper was supported by Social Security Administration contract J-9-P-7-0047 0440-02-51816. The views expressed in the paper are solely those of the author and do not necessarily represent the views of the Social Security Administration.

†Economist, Policy Simulation Group (holmer@polsim.com).

Over the past decade, stochastic projection of the financial condition of the combined OASDI trust fund has evolved from a suggestion by advisory panels to actual practice in government agencies. After describing the evolution of the stochastic projection methods currently in use, two problems with current methods are identified. Both problems are related to the representation of the long-run projected mean, or ultimate value, of key economic and demographic input variables. Current methods ignore the fact many input variables exhibit time-varying mean displacement and the fact that the long-run projection mean is rarely known with certainty. Statistical and simulation methods are proposed to solve these problems. These new methods are applied to two key input variables: the productivity growth rate and the total fertility rate. The distribution of the input variable and distribution of the trust fund actuarial balance (that is, actuarial surplus) that are generated by these new methods are compared to the distributions generated by current methods. The two examples are then combined permitting a comparison of the simulated distribution of the trust fund actuarial balance using current and new methods for both input variables. After summarizing the results of this exercise, suggestions are made for future work in this area.

Evolution of Current Methods

For many decades the Social Security Administration's Office of the Chief Actuary (OACT) has used non-stochastic methods to project trust fund finances in the annual Trustees Report.[1] Intermediate-cost time series projections for each of about a dozen key demographic and economic variables are used as input to a structural model of trust fund finances that produces a 75-year trust fund actuarial balance, the most common measure of trust fund solvency. Typically, each input variable is assumed to move gradually from its starting value to an ultimate value over the first few years of the projection, and is then assumed to remain at that ultimate value during all subsequent years of the projection. The intermediate-cost assumptions are characterized as the most likely projection. In addition to this intermediate-cost projection, two alternative projections are specified: a low-cost projection, in which each one of the key input variables is assumed to have an alternative ultimate value that increases the trust fund actuarial balance, and a high-cost

[1] See, for example, *The 2002 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.*

projection, in which each one of the key input variables is assumed to have an alternative ultimate value that decreases the trust fund actuarial balance (or surplus).

The 1991 Advisory Council report criticized these non-stochastic projection methods and recommended the use of methods that would permit the quantification of uncertainty in projections of input variables and trust fund finances.[2] The Advisory Council's critique of OACT's methods focused on the ad hoc nature of the time series assumed for the key input variables. The assumed time series have no cyclical fluctuations from year to year, and therefore, are often unrealistic. More problematic is the fact that the input variable time series used in the low-cost and high-cost projections are correlated in ways that are contrary to theoretical expectations and historical experience: either all the variables move simultaneously in ways that decrease the trust fund actuarial balance in the high-cost projection or all the variables move simultaneously in ways that increase the actuarial balance in the low-cost projection. And in addition, the low-cost and high-cost projections have no probabilistic foundations, leaving the probability of occurrence of the low-cost and high-cost projections unknown.

In order to implement the Advisory Council's recommendation, two new capabilities are needed: the capability of generating with Monte Carlo methods a large sample of realistic time series for each of the key demographic and economic input variables, and a structural model of trust fund finances that is integrated in the form of a single computer program that can process quickly thousands of projections.

OACT responded to this recommendation by experimenting with methods for generating more realistic time series for the input variables. Using quarterly time series data for a number of economic input variables, Foster estimated autoregressive integrated moving average (ARIMA) models[3] that would be suitable for use in the Trustees Report's short-run projections of the OASDI trust fund.[4] Following Foster's appointment as Chief Actuary at the Centers for Medicare & Medicaid Services, this line of research produced preliminary stochastic projections of Supplementary Medical Insurance (SMI)

[2] *Report of the 1991 Advisory Council on Social Security*, 1991.

[3] George P. Box and Gwilym M. Jenkins, *Times Series Analysis: Forecasting and Control*, New York, NY: Holden-Day, 1976.

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

costs in the Medicare Trustees Report.[5]

The 1994–96 Advisory Council report also recommended that "stochastic simulation modeling should be used as a tool for recognizing explicitly the uncertainty surrounding the Trustees' demographic and economic assumptions."[6] As part of its activities the Advisory Council supported the development of an integrated structural model of long-run trust fund finances, which has subsequently been used by various private organizations and government agencies interested in OASDI policy.[7] In addition to sponsoring the initial development of SSASIM, the Advisory Council supported the estimation of a vector autoregressive (VAR) model of three economic input variables using annual time series data on the unemployment rate, inflation rate, and nominal interest rate. A simple stochastic process representing annual returns on corporate equities was also estimated using historical data stretching back to the late 1920s. The estimated VAR model and equity return process were used in SSASIM to generate stochastic projections of the trust fund actuarial balance under alternative trust fund investment policies, including options for investing in corporate equities.[8]

The 1999 Social Security Advisory Board technical panel reiterated the recommendations of the earlier Advisory Councils. Stating that they "follow previous panels in strongly recommending efforts toward stochastic modeling or similar techniques that are better able to capture the interrelationships among assumptions," they emphasized that "what we seek is a method of displaying to policymakers and the public just how uncertain is some average cost outcome or date of exhaustion of the Trust Funds, and what are the probabilities that events will be close to or far away from that result."[9]

Incorporating ARIMA models of the mortality rate and the fertility rate that are estimated using annual data stretching back to the beginning of the

[5] *The 2002 Annual Report of the Board of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds*, Appendix IV.D: Supplementary Assessment of Uncertainty in SMI Cost Projections.

[6] *Report of the 1994–1996 Advisory Council on Social Security, Volume I*, page 22.

[7] Martin Holmer, *Introductory Guide to SSASIM* Washington, DC: Policy Simulation Group, January 2003. SSASIM is available at ⟨http://www.polsim.com/⟩.

[8] Martin Holmer and Christopher Bender, "Stochastic Simulation of Trust Fund Asset-Allocation," *Report of the 1994–1995 Advisory Council on Social Security, Volume II*, pages 431–450.

[9] Quoted on page 130 of *The 2002 Annual Report of the Board of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds*.

twentieth century,[10] an integrated reduced-form model of trust fund finances, called the Stochastic Social Security Simulator, or S^4, has been developed at Mountain View Research.[11]

During the past few years, the Congressional Budget Office (CBO) has developed an integrated structural model of OASDI trust fund finances, called CBOLT, and has used annual time series data to estimate ARIMA models for nine key input variables.[12] Three of the economic input variables are represented as a VAR model, while the mortality and fertility processes are similar to those in S^4.

And finally, OACT initiated in 2002 a comprehensive effort to develop its own stochastic projection capabilities. This effort will require not only the estimation of stochastic processes for each of the key input variables, but also the construction of an integrated model of trust fund finances out of the component sub-models that have been used for decades to produce the low-cost, intermediate-cost, and high-cost projections. It is anticipated that the first report on this effort will appear in the 2003 Trustees Report.

Problems with Current Methods

The methods that have evolved for stochastic projection of trust fund finances are basically sound. The use of historical data to estimate stochastic processes for each of the key demographic and economic input variables is a sensible way to determine the range of variation in each variable and the degree of correlation between variables. The use of Monte Carlo simulation methods to realize a time series for each of the input variables, and the use of an integrated structural model of trust fund finances to translate these input time series into a trust fund actuarial balance (or other summary measure), are straightforward applications of methods that are commonly used in other areas of science and business. The resulting probability distribution of

[10] Ronald D. Lee and Lawrence Carter, "Modeling and Forecasting the Time Series of U.S. Mortality," *Journal of the American Statistical Association* 87:659–671, 1992. Ronald D. Lee, "Modeling and Forecasting the Time Series of U.S. Fertility: Age Patterns, Range, and Ultimate Level," *International Journal of Forecasting* 9:187–202, 1993.

[11] Ronald Lee and Shripad Tuljapurkar, "Uncertain Demographic Futures and Social Security Finances," *American Economic Review* 88:237–241, 1998. S^4 is available at ⟨http://simsoc.demog.berkeley.edu/⟩.

[12] Congressional Budget Office, "Uncertainty in Social Security's Long-Term Finances: A Stochastic Analysis," *A CBO Paper*, December 2001.

the trust fund actuarial balance (or other measure) summarizes expectations about its range of variation in the future.

There is little difference of opinion about what is involved in building an integrated structural model of trust fund finances that is adequate for stochastic projection. Both CBOLT and SSASIM exhibit about the same sensitivity to changes in input variables as does the model used by OACT to produce results in the Trustees Report. What is challenging, on the other hand, is the estimation of stochastic processes for the key input variables. Alternative specifications of the input variable stochastic process can lead to substantially different probability distributions of the trust fund actuarial balance (or other measure).

One problem with current methods is that neither ARIMA nor VAR models can explicitly represent a stochastic processes with a time-varying mean displacement. This is a potentially severe problem because a number of the key input variables are viewed by economists and demographers as exhibiting means that have differed across prior decades. For example, economists have produced a massive literature documenting the fact that the productivity growth rate has been either above or below the long-run mean for decades at a time. Any stochastic process that ignores this time-varying mean-displacement behavior of the productivity growth rate will understate the degree of uncertainty in simulated probability distributions of the trust fund actuarial balance (or other measure).

A second problem with current methods is that an input variable's long-run mean (or ultimate value) used in the stochastic projection is assumed to be certain. Historical experience is not always the most reliable guide for projecting the future, and therefore, it is often appropriate to assume that an input variable's projection mean is either higher or lower than its historical mean. In other words, the assumed long-run projection mean can be viewed as the sum of the long-run historical mean and a projection deviation from the long-run historical mean. Such adjustments to the long run historical mean are completely appropriate when there are reasons to believe that the future will be different than the past. But considering the long-run projection mean to be known with certainty is a potentially severe problem because this approach ignores not only uncertainty in measuring the long-run historical mean (measurement error), but also uncertainty in the projection deviation (prediction error). Any stochastic projection that ignores these two sources of uncertainty in the long-run projection mean assumed for each input variable will understate the degree of uncertainty in simulated probability distributions of the trust fund actuarial balance (or other measure).

Overview of New Methods

The first problem — that current methods ignore time-varying mean displacement in the input variables — can be addressed by using structural time series models.[13] These structural models have been developed as an alternative to ARIMA models, have the capability of explicitly representing time-varying mean displacements, and are easily estimated by applying maximum likelihood methods to the Kalman filter using available statistical software.[14] They can also be used in a multivariate setting as an alternative to VAR models. The exact specification of the structural time series model is likely to differ for each input variable. If an input variable has already been first-differenced (like the productivity growth rate), the specification of its model is likely be less complex than the specification of a model for an input variable that has not been first-differenced (like the total fertility rate). The exact specification of the structural time series models used in the two examples are shown below, both in their equation form and state space form.

The second problem — that current methods ignore uncertainty in the assumed long-run projection means — can be addressed by using standard Monte Carlo simulation methods.[15] Rather than assume a certain long-run projection mean, estimates of an input variable's measurement error (measuring the uncertainty in the variable's historical mean) and prediction error (measuring the uncertainty in the variable's projection deviation) can be used to construct a distribution for the long-run projection mean. This normal distribution has a mean equal to the assumed ultimate value and has a variance equal to the sum of the variance of the measurement error and the variance of the prediction error (assuming no correlation between the measurement and prediction errors). Before the estimated stochastic process for an input variable is used to realize a time series, the distribution of the

[13] Andrew C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge, UK: Cambridge University Press, 1989.

[14] All estimation results reported here have been produced using TSP, version 4.5, as documented in Bronwyn H. Hall and Clint Cummins, *User's Guide to Time Series Processor, Version 4.5*, Palo Alto, CA: TSP International, 1999.

[15] J.M. Hammersley and D.C. Handscomb, *Monte Carlo Methods*, London, UK: Chapman and Hall, 1964.

variable's long-run projection mean can be sampled to obtain the realized value of the long-run projection mean for this Monte Carlo replication. This straightforward generalization of the method used to generate a time series for the input variable requires a specification of the estimated stochastic process that permits use of different values for the long-run projection mean (or ultimate value) in each Monte Carlo replication. The two examples illustrate how the variances of the measurement and prediction errors can be estimated and used in the stochastic projection.[16]

Example: Productivity Growth Rate

The productivity growth rate is an important economic input variable because it largely determines the rate of increase in real wages, which influences the level of both OASDI benefits and taxes. In this first example, current methods are used to estimate an ARIMA model of the productivity growth rate, and this model is used to generate a stochastic projection in which only the productivity growth rate fluctuates in value and all other input variables assume their deterministic intermediate-cost values. Then the new methods are applied to the productivity growth rate, stochastic projections are generated, and the degree of uncertainty in the trust fund actuarial balance is compared between the stochastic projections.

Figure 1 shows how the annual productivity growth rate has varied around the sample mean of 1.85 percent during the period from 1960 through 2000 with no apparent long-run trend.

A specification search among the class of ARIMA models finds that the best model for the productivity growth rate is a constant plus white noise. Because the sample mean has been subtracted from each annual observation, the best ARIMA model is simply white noise, as shown in Equation 1 where y_t denotes the productivity growth rate in year t minus the sample mean.

$$y_t = \epsilon_t \qquad\qquad \epsilon_t \sim N(0, \sigma_\epsilon^2). \tag{1}$$

Using ordinary least squares, 1.19 is the estimated value of σ_ϵ in Equation 1.

Using a structural time series model to represent the possibility of a time-varying mean displacement and a long-run trend in the productivity growth

[16] All simulation results presented here have been produced using SSASIM (1/10/03 version), as documented in Martin R. Holmer, *Introductory Guide to SSASIM*, Washington, DC: Policy Simulation Group, January 2003.

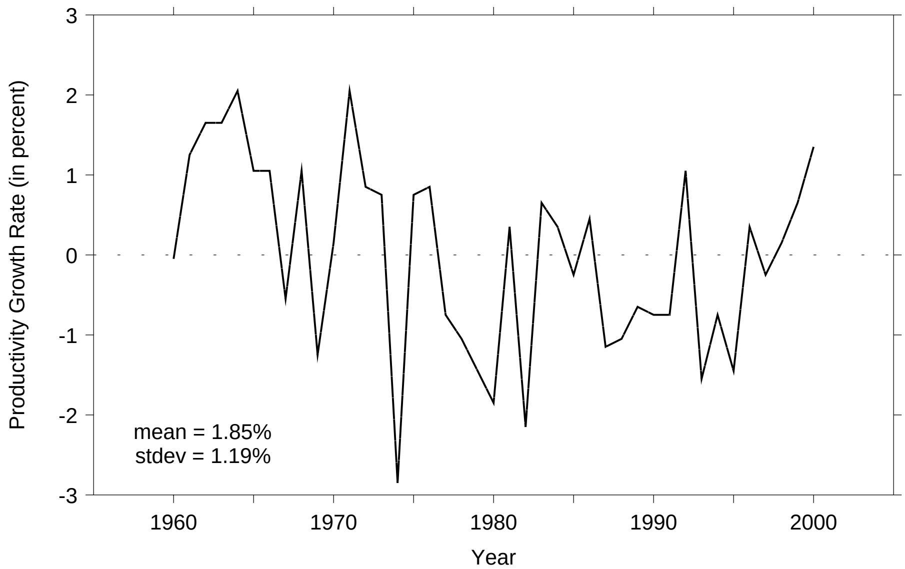

Figure 1: ***Productivity Growth Rate, 1960–2000, Measured Relative to Sample Mean***. *Data are from OACT web site.*

rate produces the model specified in Equations 2–3, where γ denotes the time-varying mean displacement and δ denotes the long-run trend.

$$y_t = \gamma_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{2}$$

where

$$\gamma_t = \lambda\gamma_{t-1} + \delta + \nu_t \qquad \nu_t \sim N(0, \pi_\nu \sigma_\epsilon^2). \tag{3}$$

The structural time series model specified in Equations 2–3 can be represented in state space form with the measurement equation as follows:

$$y_t = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ \gamma_t \\ \delta_t \end{bmatrix} \tag{4}$$

and the transition equation for the state variable as follows:

$$\begin{bmatrix} y_t \\ \gamma_t \\ \delta_t \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \gamma_{t-1} \\ \delta_{t-1} \end{bmatrix} + \omega_t \qquad \omega_t \sim N(0, \sigma_\epsilon^2 Q) \tag{5}$$

where

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \pi_\nu & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{6}$$

The hyperparameters (λ, σ_ϵ, π_ν) of the state space form in Equations 4–6 and the initial value of the state variable are estimated using maximum likelihood methods applied to the Kalman filter.[17] A likelihood ratio test shows that the estimated trend is not significantly different from zero. Assuming $\delta = 0$, the estimated value (standard error) of the hyperparameters are as follows: λ 0.875 (0.079), σ_ϵ 1.034 (—), and π_ν 0.033 (0.051). The last two estimates imply that 0.189 is the estimated value of σ_ν. A likelihood ratio test indicates that the null hypothesis of no time-varying mean displacement — that is, $\lambda = 0$ and $\pi_\nu = 0$ — is rejected at conventional significance levels ($p = 0.05$).

The estimated value of the time-varying mean displacement γ over the past forty years is shown in Figure 2. Its estimated movement, having a positive value during the 1960s and early 1970s and a negative value from the late 1970s through the late 1990s, is in accordance with the research literature on the productivity growth rate.

The results of using both the estimated models to simulate one thousand productivity growth rate time series are shown in Figure 3, which summarizes the variability of the simulated time series in a manner similar to that used by CBO.[18] The ARIMA model generates a more disperse annual distribution of the productivity growth rate, but a less disperse distribution of the cumulative average productivity growth rate.

These two sets of simulated time series for the productivity growth rate are used in SSASIM to generate two distributions of the combined OASDI trust fund actuarial balance. Each distribution consists of one thousand values of the actuarial balance. All input variables other than the productivity growth rate are assumed to have the deterministic values used in the intermediate-cost projection of the 2001 Trustees Report. Both sets of simulated time series are generated assuming a certain ultimate value (CUV) of

[17] For details see discussion of TSP's KALMAN procedure on pages 128–131 of Bronwyn H. Hall and Clint Cummins, *Reference Manual for Time Series Processor, Version 4.5*, Palo Alto, CA: TSP International, 1999.

[18] Congressional Budget Office, "Uncertainty in Social Security's Long-Term Finances: A Stochastic Analysis," *A CBO Paper*, December 2001, pages 44–54.

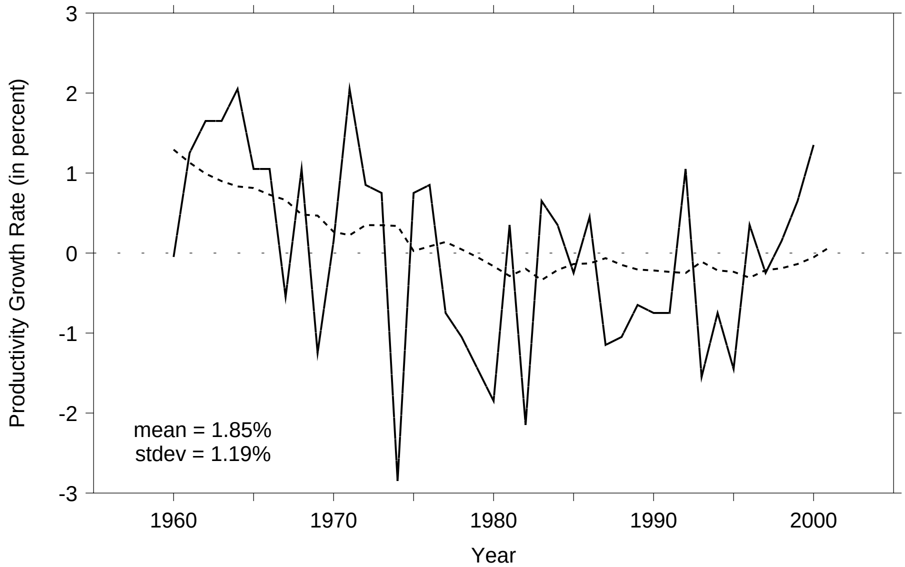

Figure 2: ***Productivity Growth Rate, 1960–2000, Measured Relative to Sample Mean, and Estimated Time-Varying Mean Displacement***. *Estimated mean displacement is shown as the dotted line.*

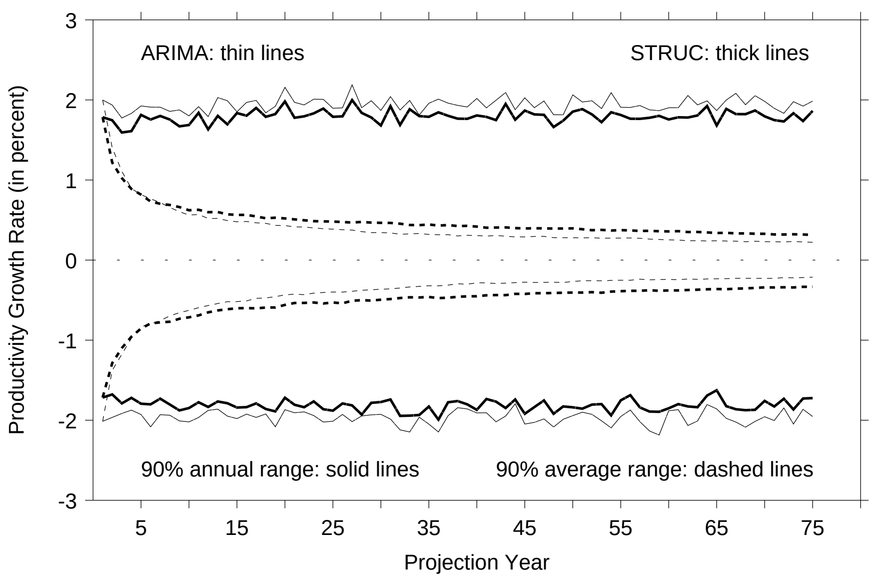

Figure 3: ***Range of Projected Productivity Growth Rates, Measured Relative to Long-Run Projection Mean***.

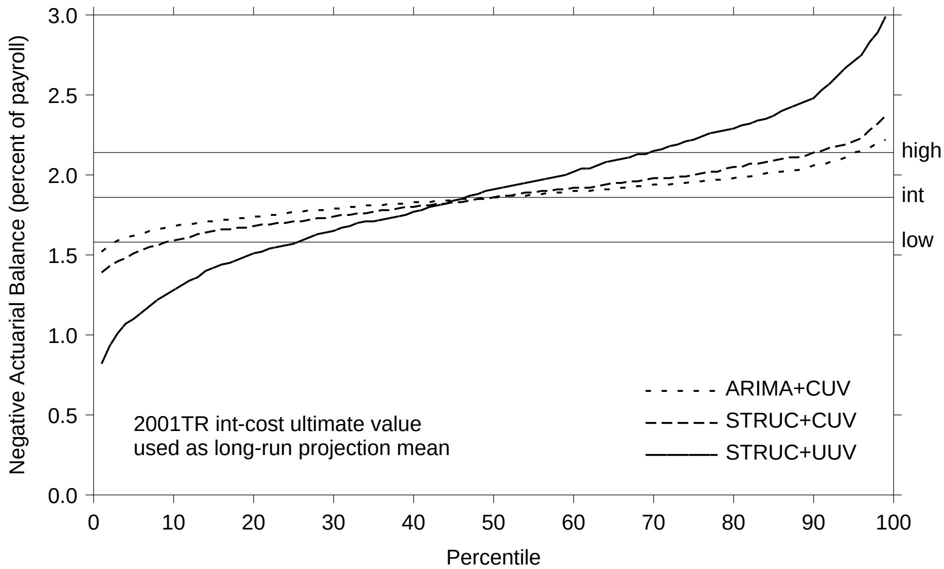

Figure 4: ***Negative OASDI Actuarial Balance Distribution for Three Models of Projected Productivity Growth Rate.***

1.5 percent for the productivity growth rate, which is the intermediate-cost ultimate value in the 2001 Trustees Report. The two resulting distributions of the actuarial balance are shown in Figure 4, where the actuarial balance associated the low-cost, intermediate-cost, and high-cost values of the productivity growth rate in the 2001 Trustees Report are shown as horizontal lines for comparison. The actuarial balance distribution generated by the structural time series model of the productivity growth rate is somewhat more disperse than the distribution generated by the ARIMA model.

The third simulated distribution of the trust fund actuarial balance shown as the solid line in Figure 4 is generated by combining the structural time series model of the productivity growth rate with an uncertain ultimate value (UUV) assumption.

Sometimes there are good reasons to believe that the long-run mean in the future will differ from the long-run mean in the past. Discussions about ultimate values most often take the form of establishing the historical mean and then the reasons for a projection deviation from that historical mean. Given an estimate of the historical mean and the projection deviation, the assumed ultimate value is calculated as follows:

$$u = h + d \tag{7}$$

where the assumed ultimate value is denoted by u, the historical long-run mean value by h, and the projection deviation by d. This procedure is sensible, but ignores the uncertainty associated with both the estimate of the historical mean and the estimate of the projection deviation. The long-run historical mean is uncertain because it is estimated with limited sample data and the projection deviation is uncertain because the extent to which the future will differ from the past is not known with certainty. If it is assumed that the measurement error associated with the historical mean, which is distributed with a normal distribution whose standard deviation is denoted by σ_h, is uncorrelated with the prediction error associated with the projection deviation, which can be assumed to be represented by a normal distribution whose standard deviation is denoted by σ_d, then the uncertainty associated with the ultimate value can be expressed as follows:

$$\sigma_u^2 = \sigma_h^2 + \sigma_d^2 \tag{8}$$

where σ_u^2 denotes the normal variance of the assumed ultimate value.

Applying Equation 8 to the productivity growth rate, the measurement error σ_h is 0.186 percent using the 1960–2000 sample observations. The use of a 1.50 percent ultimate value in the 2001 intermediate-cost projection implies a projection deviation of –0.35 percent relative to the historical mean of 1.85 percent. If it is assumed that there is only a twenty percent chance that the long-run mean in the future will equal or exceed the long-run historical mean, then the prediction error σ_d is 0.416 percent. This is a conservative estimate of the prediction error because it assumes that there is an eighty percent chance that the long-run mean in the future will be less than the long-run mean in the past. Under the plausible assumption of independence, the variances of the two errors are added to produce an estimate of the variance of the ultimate value. This implies a 0.455 percent standard deviation (σ_u) for the productivity growth rate ultimate value, whose mean value is assumed to be 1.50 percent. Notice that about 83 percent of the variance is associated with the prediction error, which itself is probably being under estimated in these calculations.

Using the structural time series model and the assumption that the ultimate value of the productivity growth rate is a normal distribution with a mean of 1.50 percent and a standard deviation of 0.455 percent produces the simulated distribution of the trust fund actuarial balance shown as the solid line in Figure 4. This distribution is substantially more disperse than the other two distributions which are both generated using the assumption of a certain ultimate value. Combining the estimated structural time series model with a conservative assumption about the size of the prediction error associated with the assumed ultimate value of the productivity growth rate produces an actuarial balance distribution with about 30 percent of the observations above the high-cost projection and about 25 percent below the low-cost projection. In other words, more than half the actuarial balance distribution lies beyond the low-cost/high-cost range shown in the 2001 Trustees report for the productivity growth rate.

The low-cost, intermediate-cost, and high-cost ultimate values for the productivity growth rate were all increased by 0.10 percentage points in the 2002 Trustees Report, which offers this explanation for the change: "This increase reflects ongoing assessment of historical data, including the period of rapid productivity growth between 1995 and 2000."[19] The fact that a few years of above average values for the productivity growth rate would lead to a change in the ultimate value used as the mean in a 75-year projection suggests a need to recognize the fact that the productivity growth rate exhibits time-varying mean displacements. It also suggests that the prediction error associated with the assumed ultimate value for the productivity growth rate is substantial in the minds of the Trustees, perhaps much larger than the prediction error assumed here.

Example: Total Fertility Rate

The total fertility rate is an important demographic input variable because it influences the number of workers paying taxes several decades later and influences the number of beneficiaries many decades later. In this second example, current methods are used to estimate an ARIMA model of the total fertility rate, and this model is used to generate a stochastic projection in which only the total fertility rate fluctuates in value and all other input variables assume their deterministic intermediate-cost values. Then the new methods are applied to the total fertility rate, stochastic projections are

[19] *The 2002 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds*, page 85.

generated, and the degree of uncertainty in the trust fund actuarial balance is compared between the stochastic projections.

Figure 5 shows how the total fertility rate has varied around the sample mean of 2.55 children per woman during the period from 1917 through 1999. Clearly there have been substantial swings up and down in fertility: a sharp fall from the late 1910s through the 1930s, the high rates associated with the baby boom during the late 1940s through the early 1960s, and another major decline during the late 1960s and early 1970s. In addition to substantiating these major fluctuations, the figure suggests the possibility that there has been a long-run trend towards lower total fertility rates.

The natural logarithm of the total fertility rate is shown for the same years in Figure 6. The log rate is used here for two reasons. First, the rate cannot be negative and a log transformation ensures that positive rates are always produced in the simulations. Second, because the possibility of a long-run trend in the total fertility rate is explored, the log transformation implies that a constant decline corresponds to a constant percentage decline in the untransformed rate, which is analogous to the OACT representation of the mortality rate input variable. Comparing Figure 5 and Figure 6 reveals that the log transformation introduces very little distortion into the time series, unlike a log-odds transformation where the rate is constrained to be between zero and four, which has been used by others.[20]

A specification search among the class of ARIMA models finds that the best model for log total fertility rate is an ARIMA(4,0,1) model, which is the specification used in earlier versions of S^4 and in CBOLT. Because the sample mean has been subtracted from each annual observation, the ARIMA(4,0,1) model shown in Equation 9 has no constant term.

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \\ &\quad \phi_4 y_{t-4} + \epsilon_t - \theta_1 \epsilon_{t-1} \qquad\qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \end{aligned} \tag{9}$$

where y_t denotes the log total fertility rate in year t minus the sample mean.

The same maximum-likelihood/Kalman-filter methods used to estimate the structural time series models are used to estimate the ARIMA(4,0,1)

[20]For CBO's use of the log-odds transformation, see Congressional Budget Office, "Uncertainty in Social Security's Long-Term Finances: A Stochastic Analysis," *A CBO Paper*, December 2001, page 73. For the reasons Lee has abandoned this transformation, see footnote 9 in Ronald D. Lee and Shripad Tuljapurkar, "Population Forecasting for Fiscal Planning: Issues and Innovations" in Alan J. Auerbach and Ronald D. Lee, eds., *Demographic Change and Fiscal Policy*, Cambridge, UK: Cambridge University Press, 2001.

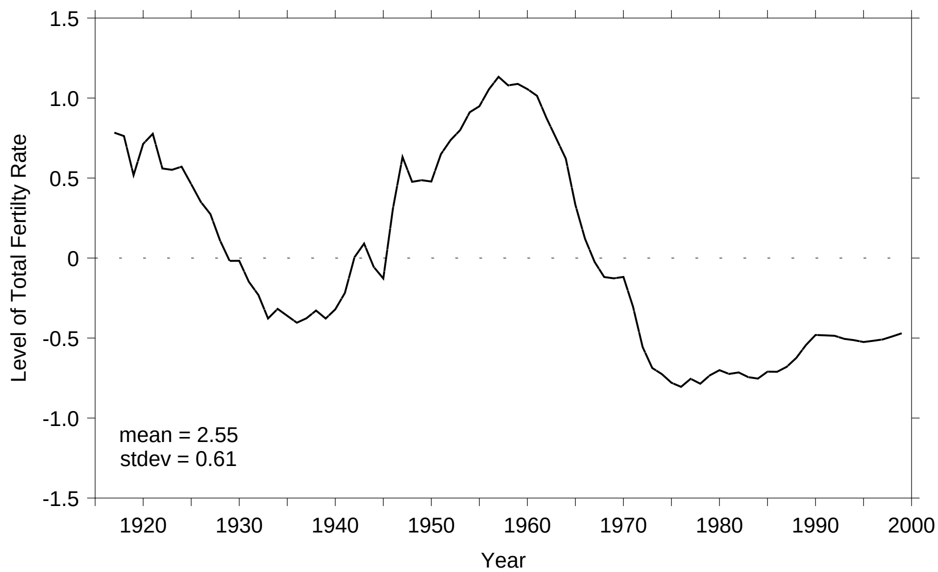

Figure 5: ***Level of Total Fertility Rate, 1917–1999, Measured Relative to Sample Mean****. Data are from an unpublished OACT fertility file provided complements of CBO.*

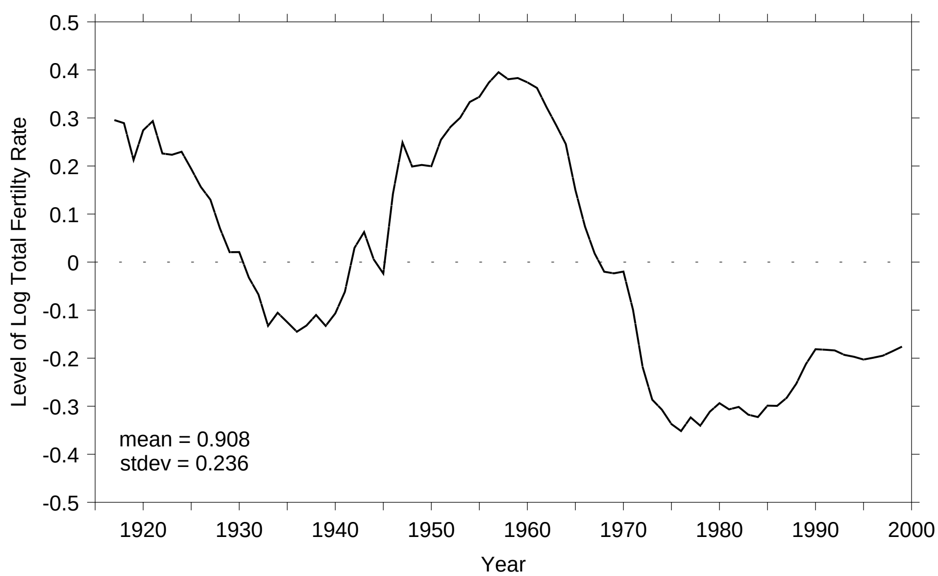

Figure 6: ***Level of Log Total Fertility Rate, 1917–1999, Measured Relative to Sample Mean****.*

model. These methods are used to facilitate statistical comparisons with the structural time series model and they produce estimated coefficients that are statistically indistinguishable from the those produced using Box-Jenkins estimation methods. The estimated values (standard errors) of the coefficients of the ARIMA(4,0,1) model of log total fertility rate are as follows: ϕ_1 1.971 (0.139), ϕ_2 –1.480 (0.246), ϕ_3 0.903 (0.207), ϕ_4 –0.416 (0.095), θ_1 0.615 (0.147), and σ_ϵ 0.033 (—). The log likelihood is 164.461.

Using a structural time series model to represent the possibility of a time-varying mean displacement (whose level and slope are both subject to random shocks) and a long-run trend in the total fertility rate produces the model specified in Equations 10–12, where μ denotes the level of the time-varying mean displacement, γ represents the slope of the displacement, and δ denotes the long-run trend in the log total fertility rate. Autoregressive terms with one-year and two-year lags are also included in the model to test the hypothesis that short-run dynamics of fertility differ qualitatively from long-run movements in fertility, as suggested by earlier work.[21]

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \mu_{t-1} + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2) \tag{10}$$

where

$$\mu_t = \rho\mu_{t-1} + \gamma_{t-1} + \eta_t \qquad \eta_t \sim N(0, \pi_\eta\sigma_\epsilon^2) \tag{11}$$

and where

$$\gamma_t = \lambda\gamma_{t-1} + \delta + \nu_t \qquad \nu_t \sim N(0, \pi_\nu\sigma_\epsilon^2). \tag{12}$$

The structural time series model specified in Equations 10–12 can be represented in state space form with the measurement equation as follows:

$$y_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ y_{t-1} \\ \mu_t \\ \gamma_t \\ \delta_t \end{bmatrix} \tag{13}$$

[21]Shripad Tuljapurkar and Carl Boe, "Validation, Bayesian Methods, and Information in Stochastic Forecasts," *Mountain View Research working paper*, 1997.

and the transition equation for the state variable as follows:

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \mu_t \\ \gamma_t \\ \delta_t \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \mu_{t-1} \\ \gamma_{t-1} \\ \delta_{t-1} \end{bmatrix} + \omega_t \tag{14}$$

where

$$\omega_t \sim N(0, \sigma_\epsilon^2 Q) \tag{15}$$

and

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \pi_\eta & 0 & 0 \\ 0 & 0 & 0 & \pi_\nu & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{16}$$

The hyperparameters (ϕ_1, ϕ_2, ρ, λ, δ, σ_ϵ, π_η, π_ν) of the state space form in Equations 13–16 and the initial value of the state variable are estimated using maximum likelihood methods applied to the Kalman filter. A likelihood ratio test shows that the variance of the shock to the displacement level is not significantly different from zero. Assuming $\pi_\eta = 0$, the estimated value (standard error) of the hyperparameters are as follows: ϕ_1 –0.148 (0.078), ϕ_2 –0.418 (0.049), ρ 0.900 (0.099), λ 0.780 (0.146), δ –0.0035 (0.0038), σ_ϵ 0.0001 (—), and π_ν 110398 (109767). The last two estimates imply that 0.0343 is the estimated value of σ_ν. The log likelihood is 173.715. A likelihood ratio test indicates that the null hypothesis of no long-term trend in the total fertility rate — that is, $\delta = 0$ — is rejected at the $p = 0.10$ significance level, but not at the $p = 0.05$ significance level ($\chi^2(1) = 3.34$). The structural time series model clearly fits the data better than the ARIMA(4,0,1) model even when adjusting for the different number of parameters estimated in the two models: the Bayesian information criterion (BIC) for the structural model is much smaller (–151.621) than the BIC for the ARIMA(4,0,1) model (–142.367).

The negative ϕ coefficients in the structural time series model of the total fertility rate suggest that the short-run dynamics of fertility are influenced by the fact that women who have given birth within the past year are less likely to conceive than women without young babies. The structural model's

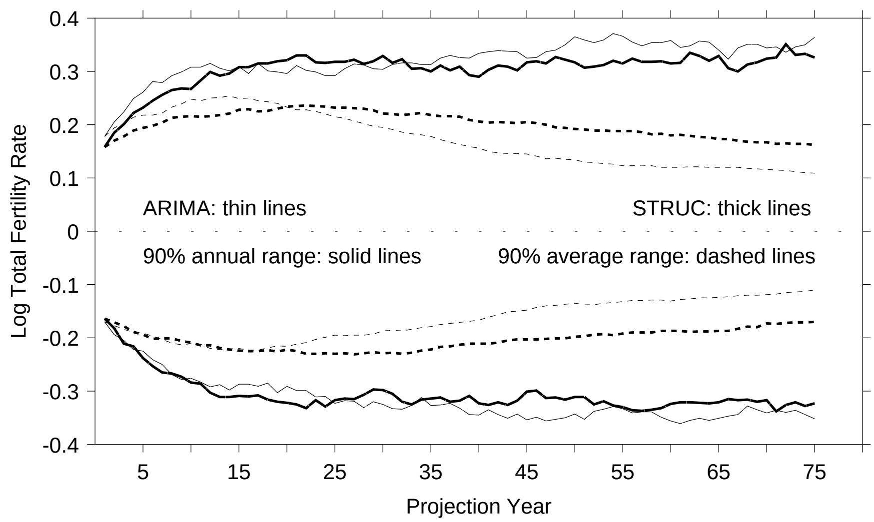

Figure 7: ***Range of Projected Log Total Fertility Rates, Measured Relative to Long-Run Projection Mean.***

explicit specification of a time-varying mean displacement appears to be the reason why these two autoregressive parameters have significantly different values than they do in the ARIMA(4,0,1) model.

The results of using both the estimated ARIMA model and the estimated structural model (with the trend parameter δ set to zero) to simulate one thousand total fertility rate time series are shown in Figure 7. The ARIMA model generates distributions of the annual total fertility rate that have roughly the same dispersion as the annual distributions generated by the structural model, but the ARIMA model generates a cumulative average total fertility rate distribution that is much less disperse than the cumulative average distribution generated by the structural model of fertility.

These two sets of simulated time series for the total fertility rate are used in SSASIM to generate two distributions of the combined OASDI trust fund actuarial balance. Each distribution consists of one thousand values of the actuarial balance. All input variables other than the total fertility rate are assumed to have the deterministic values used in the intermediate-cost projection of the 2001 Trustees Report. Both sets of simulated time series are generated assuming a certain ultimate value of 1.95 children for the total fertility rate, which is the intermediate-cost ultimate value in the 2001 and 2002 Trustees Reports. The two resulting distributions of the actuarial

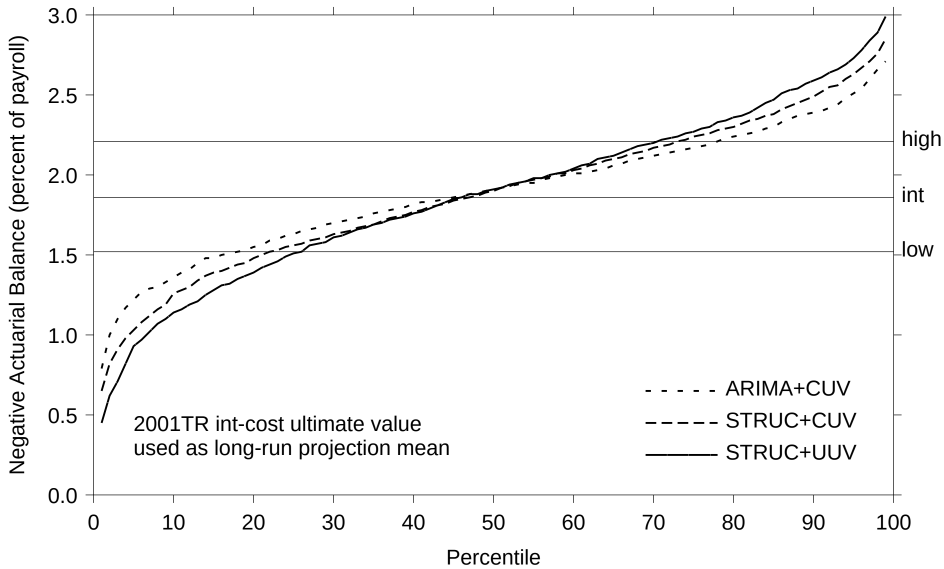

Figure 8: ***Negative OASDI Actuarial Balance Distribution for Three Models of Projected Total Fertility Rate.***

balance are shown in Figure 8, where the actuarial balance associated the low-cost, intermediate-cost, and high-cost values of the total fertility rate in the 2001 Trustees Report are shown as horizontal lines for comparison. The actuarial balance distribution generated by the structural model of the total fertility rate (with its trend coefficient δ set to zero) is noticeably more disperse than the distribution generated by the ARIMA model.

The third simulated distribution of the trust fund actuarial balance shown as the solid line in Figure 8 is generated by combining the structural time series model of the total fertility rate (assuming no long-run trend) with an uncertain ultimate value assumption.

Calculating the standard error of the ultimate value for the total fertility rate involves several steps.

The results of using the estimated structural model (including the negative estimated trend parameter δ) to simulate one thousand total fertility rate time series indicates that the mean value of the cumulative 75-year average total fertility rate is about 1.79 children, which is about eight percent below the 1.95 ultimate value assumed in recent Trustees Reports. Assuming that the historical mean is interpreted as the long-run mean produced by the continuation of the historical trend, applying Equation 7 produces $log(1.95) = log(1.79) + 0.0856$. In other words, the intermediate-cost ultimate value for total fertility rate assumes a slowdown in the long run historical trend towards lower fertility rates. This may turn out to be an accurate prediction in the distant future, but the uncertainty of this projection must be recognized in the present.

The measurement error associated with the $log(1.79)$ term can be estimated as $\sqrt{\sigma^2/T}$ where σ^2 denotes the variance of estimated residuals in an ordinary least squares regression of the log total fertility rate on a constant term and a time trend using $T = 83$ annual observations. These calculations produce an estimate of 0.021 for the measurement error, the σ_h in Equation 8.

The prediction error associated with the 0.0856 projection deviation, the σ_d in Equation 8, is estimated to be 0.102 using the conservative assumption that there is only a twenty percent chance that the long-run mean in the future will be less than the mean associated with a continuation of the historical fertility trend. This is a conservative estimate of the prediction error because it assumes that there is an eighty percent chance that the long-run mean in the future will be more than the long-run mean implied by projecting the historical fertility trend into the future.

Under the plausible assumption of independence, the variances of the measurement and prediction errors are added to produce an estimate of the variance of the ultimate value. This implies a 0.104 standard error (σ_u) for the total fertility rate ultimate value, whose mean value is assumed to be 0.668 ($= log(1.95)$). Notice that 96 percent of the variance is associated with the prediction error, which itself is probably being under estimated in these calculations.

Using the structural time series model (with its trend parameter set to zero) and the assumption that the ultimate value of the log total fertility rate is a normal distribution with a mean of 0.668 and a standard deviation of 0.104 produces the simulated distribution of the trust fund actuarial balance shown as the solid line in Figure 8. This distribution is somewhat more disperse than the other two distributions which are both generated using the assumption of a certain ultimate value. Combining the estimated structural time series model (with its trend parameter set to zero) with a conservative assumption about the size of the prediction error associated with the assumed ultimate value of the total fertility rate, produces an actuarial balance distribution with about 30 percent of the observations above the high-cost projection and about 25 percent below the low-cost projection. In other words, more than half the actuarial balance distribution lies beyond the low-cost/high-cost range shown in the 2001 Trustees report for the total fertility rate.

It could be argued that the baby boom was a unique historical experience that will never be repeated. This is a debatable demographic assumption, analogous to an economic assumption that no major wars or depressions will occur in the future. Excluding historical data for such "unique" demographic and economic experiences from the samples used to estimate time series models for the key input variables is likely to reduce substantially the projected uncertainty in trust fund finances. To get an idea of the magnitude of the reduction, the above analysis of the total fertility rate has been redone based on a sample that excludes the twenty baby-boom years from 1946 to 1965. The post-1965 observations are placed immediately following the 1945 observation in this edited sample. The coefficients of the ARIMA and structural models, and the degree of uncertainty in the assumed ultimate value, are estimated using the same methods with this edited sample. The resulting simulated distributions of the trust fund actuarial balance are shown in Figure 9, which has exactly the same scale as Figure 8 where the distributions resulting from use of the complete 1917–1999 sample are shown. While the distributions produced by the edited sample clearly show less dispersion than those produced by the complete sample, about 40 percent of the distribution generated by the structural model and uncertain ultimate value falls outside the low-cost/high-cost range.

An alternative approach to editing the sample would be to estimate the time series models using only post-1965 fertility data. This more drastic approach ignores not only the baby boom experience, but also major fluctuations in the total fertility rate that occurred before 1946. The rationale for such an approach is unclear. What is clear is that such a severe sample-editing approach would produce estimated time series models that generate total fertility rate distributions that are much less disperse than the ones shown here. Whether or not the simulated distribution of the trust fund actuarial balance would be much less disperse than the ones shown here depends on whether or not uncertainty in the ultimate value was recognized and estimated to be larger because of the reduced reliability of the severely edited sample.

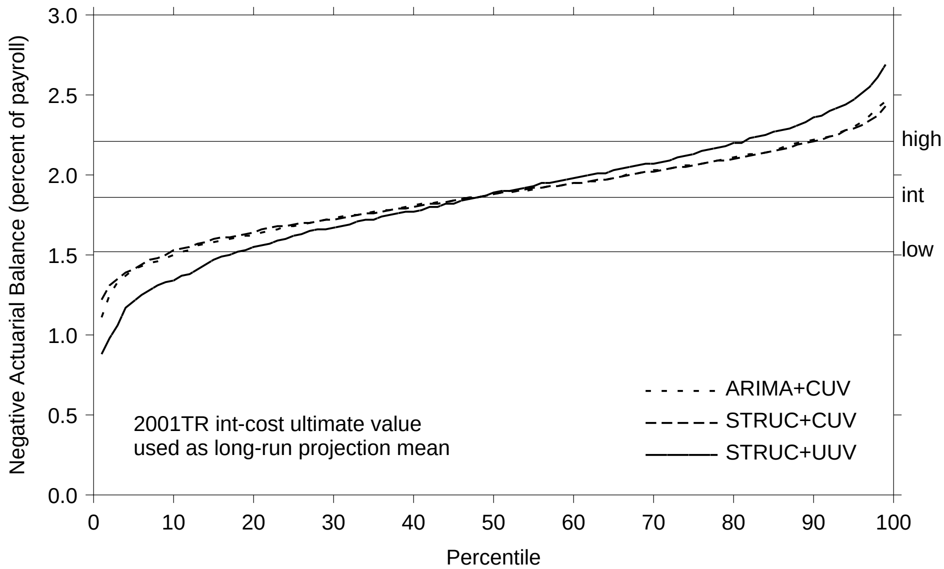

Figure 9: ***Negative OASDI Actuarial Balance Distribution for Three Models of Projected Total Fertility Rate Estimated Excluding Baby-Boom Years in Sample***. *The excluded years are 1946–1965.*

Projections with Examples Combined

As a final exercise, the time series model for productivity growth rate and the time series model for total fertility rate (estimated with the complete sample) are used together to produce a stochastic projection of the trust fund actuarial balance. All other input variables are assumed to have the deterministic values used in the intermediate-cost projection of the 2001 Trustees Report. As in the two examples, three distributions of the actuarial balance are compared in Figure 10, which has exactly the same scale as other figures that show actuarial balance distributions. The actuarial balance generated using the low-cost, intermediate-cost, and high-cost ultimate values of the productivity growth rate and total fertility rate in the 2001 Trustees Report are shown as horizontal lines for comparison. These horizontal lines are more widely spaced than in the figures above because the low-cost (high-cost) actuarial balance assumes, as do the Trustees Reports, that both input variables simultaneously take on their low-cost (high-cost) ultimate value. This extreme assumption is equivalent to assuming that the ultimate values of the productivity growth rate and total fertility rate are perfectly positively

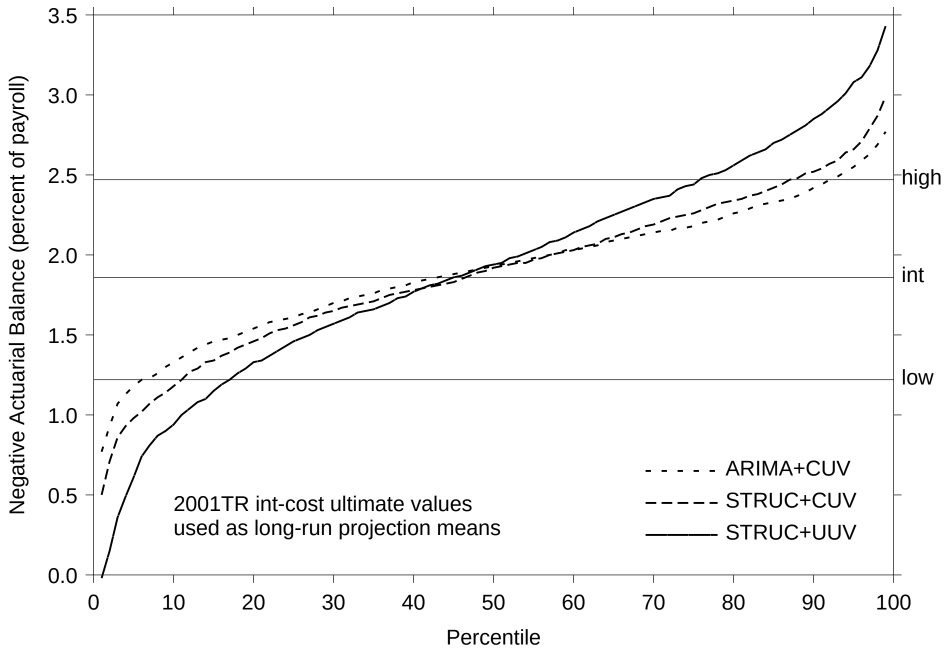

Figure 10: ***Negative OASDI Actuarial Balance Distribution for Three Combined Models of Projected Productivity Growth Rate and Projected Total Fertility Rate****. Total fertility rate models are all estimated with the complete 1917–1999 sample.*

correlated (that is, their correlation coefficient equals plus one).

One actuarial balance distribution is produced using the ARIMA models for productivity and fertility combined with the assumption of certain ultimate values for the productivity growth rate and total fertility rate. This distribution has a standard deviation of 0.43 percent with only about 12 percent of its observations located beyond the low-cost/high-cost range. A second actuarial balance distribution is generated using the structural models for productivity and fertility combined with the assumption of certain ultimate values for the productivity growth rate and total fertility rate. This distribution has a standard deviation of 0.52 percent with about 22 percent of its observations located beyond the low-cost/high-cost range. The third actuarial balance distribution is produced by combining the structural models for productivity and fertility with the assumption of uncertain ultimate values for the productivity growth rate and total fertility rate. This distribution has a standard deviation of 0.74 percent with about 40 percent of its observations located beyond the low-cost/high-cost range. The difference between 40 percent beyond the low-cost/high-cost range, when recognizing both time-varying mean displacements and uncertain ultimate values, and 12 percent beyond the low-cost/high-cost range, when using current methods that ignore both these issues, illustrates how much current methods under estimate the uncertainty in the trust fund actuarial balance.

Conclusion and Future Work

After reviewing the evolution of methods currently used to produce stochastic projections of the combined OASDI trust fund, two problems with these methods are identified. First, the ARIMA models currently used to estimate stochastic processes for the key demographic and economic input variables do not permit time-varying mean displacements. And second, the Monte Carlo simulation methods currently used to generate a probability distribution for the trust fund actuarial balance (or other financial statistic) do not recognize uncertainty in the ultimate values (or long-run projection means) of the key input variables. Both of these problems with current methods cause the dispersion of the actuarial balance distribution to be under estimated.

After describing a statistical estimation method and a stochastic simulation method that have the potential to solve these problems, the new methods are tested with one economic input variable, the productivity growth rate, and one demographic input variable, the total fertility rate. These two examples indicate that the new methods can be implemented without difficulty, that the use of structural time series models permits the representation of time-varying mean displacements that lead to better fitting models, and that a straightforward application of Monte Carlo methods can be used to simulate ultimate-value uncertainty arising primarily from prediction errors.

Comparing the actuarial balance distributions generated using current methods with distributions generated using these new methods indicates, at least in the two examples considered here, that current methods under estimate the dispersion in the actuarial balance distribution by a substantial amount. These preliminary results suggest a need to apply these new methods to all the key variables used as model inputs. Only a comprehensive implementation of the new methods can provide an accurate indication of how much uncertainty in the trust fund actuarial balance is being missed by the current methods.