Modeling and Forecasting U. S. Mortality

Ronald D. Lee, Lawrence R. Carter

Journal of the American Statistical Association, Volume 87, Issue 419 (Sep., 1992), 659-671.

Journal of the American Statistical Association

Modeling and Forecasting U.S. Mortality

RONALD D. LEE and LAWRENCE R. CARTER*

Time series methods are used to make long-run forecasts, with confidence intervals, of age-specific mortality in the United States from 1990 to 2065. First, the logs of the age-specific death rates are modeled as a linear function of an unobserved period-specific intensity index, with parameters depending on age. This model is fit to the matrix of U.S. death rates, 1933 to 1987, using the singular value decomposition (SVD) method; it accounts for almost all the variance over time in age-specific death rates as a group. Whereas e_0 has risen at a decreasing rate over the century and has decreasing variability, $\mathbf{k}(t)$ declines at a roughly constant rate and has roughly constant variability, facilitating forecasting. $\mathbf{k}(t)$, which indexes the intensity of mortality, is next modeled as a time series (specifically, a random walk with drift) and forecast. The method performs very well on within-sample forecasts, and the forecasts are insensitive to reductions in the length of the base period from 90 to 30 years; some instability appears for base periods of 10 or 20 years, however. Forecasts of age-specific rates are derived from the forecasts of $\mathbf{k}$, and other life table variables are derived and presented. These imply an increase of 10.5 years in life expectancy to 86.05 in 2065 (sexes combined), with a confidence band of plus 3.9 or minus 5.6 years, including uncertainty concerning the estimated trend. Whereas 46% now survive to age 80, by 2065 46% will survive to age 90. Of the gains forecast for person-years lived over the life cycle from now until 2065, 74% will occur at age 65 and over. These life expectancy forecasts are substantially lower than direct time series forecasts of e_0, and have far narrower confidence bands; however, they are substantially higher than the forecasts of the Social Security Administration's Office of the Actuary.

KEY WORDS: Demography; Forecast; Life expectancy; Mortality; Population; Projection.

From 1900 to 1988, life expectancy in the United States rose from 47 to 75 years. If it were to continue to rise at this same linear rate, life expectancy would reach 100 years in 2065, about seventy five years from now. The increase would be welcomed by most of us, but it would come as a nasty surprise to the Social Security Administration, which plans on the more modest life expectancy of 80.5 years predicted by its Office of the Actuary. We scarcely need dwell on the importance of the future course of mortality in our aging society. In contrast to the past, now mortality decline is a powerful cause of population aging.

There are many ways to forecast mortality (Land 1986; Olshansky 1988). The new method we propose here is extrapolative and makes no effort to incorporate knowledge about medical, behavioral, or social influences on mortality change. Its virtues are that it combines a rich yet parsimonious demographic model with statistical time series methods, it is based firmly on persistent long-term historical patterns and trends dating back to 1900, and it provides probabilistic confidence regions for its forecasts. While many methods assume an upper limit to the human life span or rationalize in some other way the deceleration of gains in life expectancy, our method allows age-specific death rates to decline exponentially without limit; the deceleration of life expectancy follows without any special additional assumptions. We believe that our method has important advantages over other extrapolative procedures, albeit with the usual shortcomings of its genre.

In this article we first consider the available data and their limitations. We then develop our demographic model of mortality, which represents mortality level by a single index. Next we fit the demographic model to U.S. data and evaluate its historical performance. Using standard time series methods, we then forecast the index of mortality and generate associated life table values at five-year intervals. Because we intend our forecasts to be more than illustrative, we present them in some detail and provide information to enable the reader to calculate life table functions and their confidence intervals for each year of the forecast.

1. THE HISTORICAL DATA

Annual age-specific death rates for the entire U.S. population are available for the years 1933 to 1987. For the years 1900 to 1932, these data are available annually only for the death registration states, which form a varying subset of the total U.S. population, and have a cruder age specificity (see Grove and Hetzel 1968, table 51, p. 309). While data generally are available by race and sex, here we restrict our analysis to the age-specific mortality of the total population. (We plan to extend the analysis to population subgroups in the future, but are concerned about extrapolating differentials.) Death rates are available for infants and standard five-year age groups up to age 85, and for age 85 and over. There is reason to be skeptical about measures of mortality at the older ages. With 46% of the population already surviving to age 80, and with future gains in life to be concentrated at old ages, it is particularly important to deal carefully with the older age groups. We will use a new method proposed by Coale and Kisker (1990) and Coale and Guo (1989).

Figure 1 plots life expectancy at birth from the years 1900 to 1989. (It also plots forecasts, which should be ignored for now.) Not surprisingly, the 28-year increase in life expectancy between 1900 and 1987 was accompanied by more dramatic declines in death rates at some ages than at others. The mortality rate for infants fell to .067 of its initial value and that for age group 1–4 fell to .026 of its initial value, but that for

* Ronald D. Lee is Professor, Departments of Demography and Economics, University of California, Berkeley, CA 94720. Lawrence R. Carter is Professor, Department of Sociology, University of Oregon, Eugene, OR 97403. This article was prepared as part of a project on "Modeling and Forecasting Demographic Time Series," supported by NICHHD Grant RO1-HD24982. An earlier draft was presented at the 1990 annual meeting of the Population Association of America in Toronto. The authors thank Kenneth W. Wachter, John Wilmoth, George Alter, Nathan Keyfitz, Jay Olshansky, William Bell, Gregory Spencer, Leo Goodman and our referees and editor for helpful comments.

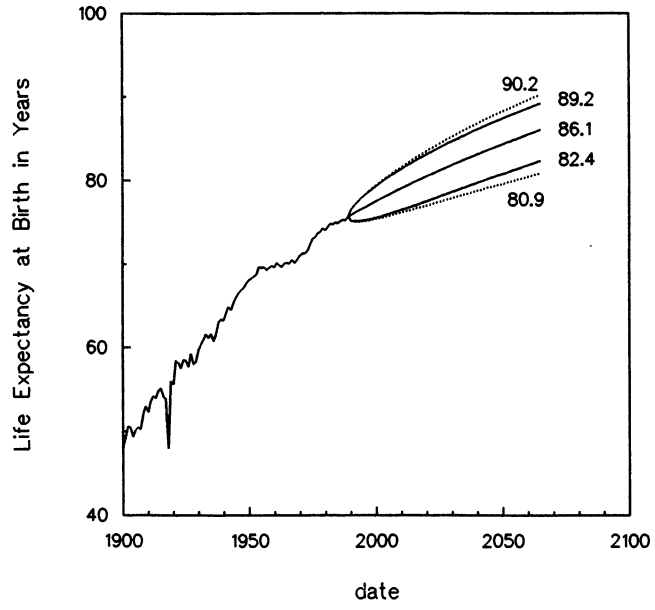

Figure 1. Actual U.S. Life Expectancy and Forecasts (95% Confidence Intervals With and Without Uncertainty From Trend Term). The forecasts use a (0, 1, 0) model with a flu dummy estimated on mortality data from 1900 to 1989. The 95% confidence intervals are shown with and without uncertainty from drift.

age group 85+ fell to only about .58 of its initial value. These proportions differ by a factor of 22 at the extremes!

We seek to develop a parsimonious model of the pattern of change over time in these death rates, so that variation in a single parameter can generate the main outlines of the observed pattern (see Keyfitz 1981). Of course we must accept from the outset that no model can (or should) reflect all age-time variation, some of which is highly irregular and arises from particular historical circumstances. To illustrate, the influenza epidemic of 1918 raised the death rate 34% above trend overall. But whereas for persons age 55 and above there was no increase, for persons age 25 to 34 the death rate rose 150% above trend; death rates at the younger ages also rose substantially (Grove and Hetzel 1968, p. 325). Such a pattern diverges from long-run trends and thus must remain outside the model. Likewise, from 1960 to the early 1980s the death rates for young adult males first rose substantially and then declined, dominating the behavior of the rates for sexes combined. This likewise must be viewed as an anomalous departure from the basic pattern of decline, requiring special study rather than absorption into our general model. We already have commented on the unreliability of the rates for age group 85+; there is no point in trying to match their erratic course. These exceptions notwithstanding, there is an overall pattern of fairly regular change, with coincident movement of death rates at all ages.

2. THE MODEL

Given an appropriate model, forecasts of the single parameter could then be used to generate forecasts of the level and age distribution of mortality for the next few decades. There are several familiar candidates for the model. A Brass logit model, with the beta parameter fixed, would fit the bill. Any one-parameter family of life tables, such as the Coale–Demeny Model West Female, likewise would do, with the "level" serving as the parameter—although such a model might not fit the U.S. experience particularly well. We could do a principal components analysis of the matrix of death rates, as Bozik and Bell (1989) did to forecast fertility. In the manner of Lederman (1969), we could construct a two-parameter model life table system using principal components analysis and could fix one parameter and forecast the other. We could fit a model to the logs of the death rates as Wilmoth (1990) and Wilmoth, Vallin, and Caselli (1989) did, including additive age and period effects and then an appropriate number of age-period interaction terms, although we would then have to forecast multiple parameters. The method we use here has most in common with the principal components method, with Wilmoth's method, and with the method of constructing the United Nations set of model life tables (Heligman 1984). For an altogether different approach to modeling and forecasting mortality, see Rogers (1986) and McNown and Rogers (1990), who fitted multiple parameter curves to mortality age schedules and then forecasted the time series of parameter estimates. It is difficult to obtain confidence intervals with this method. We believe that the method pursued here offers a number of important advantages, as discussed earlier.

Let $\mathbf{m}(x, t)$ be the central death rate for age x in year t. We will fit this matrix of death rates by the model $\ln[\mathbf{m}(x, t)] = \mathbf{a}_x + \mathbf{b}_x\mathbf{k}_t + \varepsilon_{x,t}$, or $\mathbf{m}(x, t) = e^{\mathbf{a}_x+\mathbf{k}_t\mathbf{b}_x+\varepsilon_{x,t}}$, for appropriately chosen sets of age-specific constants, $\{\mathbf{a}_x\}$ and $\{\mathbf{b}_x\}$, and time-varying index $\mathbf{k}_t$. Because $\mathbf{k}$ is an index of the level of mortality, we sometimes will write $\mathbf{m}(x, k)$, by which we will mean $e^{\mathbf{a}_x+\mathbf{k}\mathbf{b}_x}$. $\mathbf{K}$ sometimes will be written $\mathbf{k}(t)$ or $\mathbf{k}_t$. $e^{\mathbf{a}_x}$ is the general shape across age of the mortality schedule. The $\mathbf{b}_x$ profile tells us which rates decline rapidly and which rates decline slowly in response to changes in $\mathbf{k}$ ($d\ln(\mathbf{m}_{x,t})/dt = \mathbf{b}_x dk/dt$). In principle $\mathbf{b}_x$ could be negative for some ages, indicating that mortality at those ages tends to rise when falling at other ages; in practice this does not seem to occur over the long run. When $\mathbf{k}$ is linear in time, mortality at each age changes at its own constant exponential rate. As $\mathbf{k}$ goes to negative infinity, each age-specific rate goes to 0; negative death rates cannot occur in this model, which is an advantage for forecasting.

The error term, $\varepsilon_{x,t}$, with mean 0 and variance σ_ε^2, reflects particular age-specific historical influences not captured by the model, as discussed earlier. Despite the fact that the age-specific death rates may vary by a factor of 1,000 in a given year, the variances over time of age-specific components of $\varepsilon_{x,t}$ should not differ greatly because they represent deviations from the logs of the rates.

An entire one-parameter family of life tables can be generated from two observed life tables using this model and viewing death rates as a function of $\mathbf{k}$ rather than of time. Choose a scale for $\mathbf{k}$ by letting $\mathbf{k} = 0$ for one life table and $\mathbf{k} = 1$ for the other. Then let $\mathbf{a}_x = \ln[\mathbf{m}(x, 0)]$ and $\mathbf{b}_x = \ln[\mathbf{m}(x, 0)] - \ln[\mathbf{m}(x, 1)]$. As $\mathbf{k}$ varies, a family of life tables will be generated that includes the two which form its basis. For $\mathbf{k}$ between 0 and 1, the model geometrically interpolates between the two life tables; for $\mathbf{k}$ less than 0 or greater than 1, it extrapolates from the two tables. This

method may be useful for some Third World countries with little data. But when more than two life tables are available, this procedure is inefficient and it is preferable to estimate the $\mathbf{a}_x$ and $\mathbf{b}_x$ schedules along with $\mathbf{k}(t)$ to minimize the squared deviations from a given matrix of age-specific rates; details are given in the section on "Fitting the Model." We will show that the period-specific life tables generated in this way fit actual U.S. life tables quite well over the years 1933–1987 and produce life expectancies that correspond quite closely to actual life expectancies. (See Lee 1974, 1977 for other applications.)

For any value of $\mathbf{k}$, the fitted model defines a set of central death rates that can be used to derive a life table. But the procedure also can be used in reverse, to solve analytically for the particular life tables in this family that would produce an observed number of deaths, $D(t)$, for a given population age distribution, $N(x, t)$. The task is to find $\mathbf{k}(t)$ such that:

$$D(t) = \sum [N(x, t)e^{\mathbf{a}_x + \mathbf{k}(t)\mathbf{b}_x}].$$

This can be done only by searching over a range of values of $\mathbf{k}$; no analytic solution is available.

This expression can be used for two purposes. First, the original least-squares estimates of $\mathbf{k}(t)$, $\mathbf{a}_x$, and $\mathbf{b}_x$ do not generate fitted life tables that imply the exact observed number of deaths for the actual historical population age distributions. If desired, the discrepancy can be removed by keeping $\mathbf{a}_x$ and $\mathbf{b}_x$ as estimated but calculating a new set of $\mathbf{k}(t)$'s as above. Second, there may be periods for which population age distributions and total deaths are known but age-specific death rates are not available. The model still can be fitted using this procedure. This is particularly useful for forecasting when, as in the United States, there is a lag of several years between publication of total deaths and publication of age-specific death rates. With this method, the base year for the forecast always can be taken to be the last year for which total deaths are available.

This method differs from forecasting each age-specific rate independently in many ways. First, if each rate were forecast independently, then we also would need to calculate $n(n - 1)/2$ different covariances of errors, where n is the number of age groups; these are necessary to find the confidence bounds for life expectancy. Our specification exploits the high degree of intertemporal correlation across the ages, by making all death rates functions of the same time-varying parameter. Their variances and covariances follow from the autoregressive integrated moving average (ARIMA) model of the time-varying parameter and the assumptions of the model. Second, each rate might be best modeled individually by a different order ARIMA process, requiring the estimation of many parameters. Only if each is well modeled by a random walk with drift will the number of parameters be similar to the number in our method. This is true whether or not $\mathbf{k}$ happens to be modeled by a random walk with drift. Third, with individual forecasts, rates in the distant future might combine to form highly implausible age profiles. Using our method, the individual age-specific rates are always constrained to belong to a life table system that fits the historical data. Fourth, this cohesion is obtained by forecasting the single parameter $\mathbf{k}$, which itself is a kind of compromise among the trends in all the individual age-specific rates. This leads to different forecasts of the individual rates than would be obtained by modeling them individually.

After we had chosen this representation of age-time variations in mortality rates, we came upon a working paper by Gomez (1990) that conducts an exploratory data analysis of the matrix of Norwegian mortality rates since the 19th century. Gomez considered a large number of possible models, involving different initial transformations of the death rates, inclusion of cohort effects, a pure period term, and so on. He found that the model we have chosen by an entirely different route was best and used it as the basis of a 10-year mortality forecast for Norway.

3. FITTING THE MODEL

To estimate the model for a given matrix of rates $\mathbf{m}_{x,t}$, we seek the least squares solution to the equation:

$$\ln(\mathbf{m}_{x,t}) = \mathbf{a}_x + \mathbf{b}_x\mathbf{k}_t + \varepsilon_{x,t}.$$

This model evidently is underdetermined, which can be seen as follows. Suppose that the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{k}$ are one solution. Then for any scalar c, $\mathbf{a} - \mathbf{b}c$, $\mathbf{b}$, $\mathbf{k} + c$ also must be a solution. It also is clear that if $\mathbf{a}$, $\mathbf{b}$, $\mathbf{k}$ is a solution, then $\mathbf{a}$, $\mathbf{b}c$, $\mathbf{k}/c$ also is a solution. Therefore, $\mathbf{k}$ is determined only up to a linear transformation, $\mathbf{b}$ is determined only up to a multiplicative constant, and $\mathbf{a}$ is determined only up to an additive constant. In what follows, we have normalized the $\mathbf{b}_x$ to sum to unity and the $\mathbf{k}_t$ to sum to 0, which implies that the $\mathbf{a}_x$ are simply the averages over time of the $\ln(\mathbf{m}_{x,t})$.

The model cannot be fit by ordinary regression methods, because there are no given regressors; on the right side of the equation we have only parameters to be estimated and the unknown index $\mathbf{k}(t)$. The singular value decomposition (SVD) method can be used to find a least squares solution when applied to the matrix of the logarithms of the rates after the averages over time of the (log) age-specific rates have been subtracted (Good 1969; Wilmoth et al. 1989). The first right and left vectors and leading value of the SVD, after the normalization described above, provide a unique solution. Appendix A provides further details and an alternate estimation method.

The fitted death rates derived in this way generally will not lead to the actual numbers of deaths when applied to given population age distributions. Furthermore, $\mathbf{k}$ is estimated to minimize errors in the logs of death rates rather than the death rates themselves. Using the equation given in Section 3 on this page, we have reestimated $\mathbf{k}(t)$ in a second step, taking the $\mathbf{a}_x$ and $\mathbf{b}_x$ estimates from the first step as given. We thereby find a new estimate of $\mathbf{k}$ such that for each year, given the actual population age distribution, the implied number of deaths will equal the actual number of deaths. These values of $\mathbf{k}(t)$ are found by an iterative search and differ somewhat from the direct SVD estimates with U.S. data after the early 1950s. This is because when fitting the log-transformed rates, the low death rates of youth receive the same weight as the high death rates of the older ages, yet they contribute far less to the total deaths. Additionally, differences in population age group sizes impart different weights in the second-stage estimation of $\mathbf{k}$.

Unfortunately, the annual time series of death rates extends only to the open age group 85 and over, whereas our interest extends to higher ages. Even in 1987, 30% of the population survived to age 85 (39% of the females); we forecast that by 2065 62% will do so. Under these circumstances, failing to take account of the population distribution within that open interval is at best uninformative and at worst may lead to serious distortions. Coale and Kisker (1990) showed that in populations with good data at old ages, mortality rates increase not at a constant rate with age, as the Gompertz curve assumes, but rather at a linearly decreasing rate. Their procedures were based on age patterns consistently observed in a number of populations with very low mortality: The Netherlands, Japan, France, West Germany, Austria, Sweden, and Norway. We apply the method suggested in Coale and Guo (1989, pp. 614–615) to extend our death rates up to age group 105–109, by expressing death rates at the older ages as functions of $\mathbf{k}$, $\mathbf{a}_x$, and $\mathbf{b}_x$ and then searching for the appropriate $\mathbf{k}$ as explained earlier. For our data, extending the life table in this way alters the calculated life expectancy over the sample period only negligibly. In the forecast period, however, the life expectancies calculated in this way came to exceed those based on fewer age groups; by 2065, e_0 was .7 years higher when death rates were extended up to age group 105–109 than it was when we used age group 85+.

4. THE FITTED MODEL

Figure 2 plots estimates of $\mathbf{k}$ (and also shows forecasts, which should be ignored until later). As shown, $\mathbf{k}$ declines roughly linearly from 1900–1989, which is striking because the pattern of change in life expectancy is definitely not linear. $\mathbf{k}$ declines at about the same pace during the first half of the period as it does during the second half: by 15.8 from 1900–1944 and by 16.7 from 1944–1988. In contrast, life expectancy at birth increased over these periods by 17.6 and 9.9 years, respectively. It also is striking that short-run fluctuations in $\mathbf{k}$ do not appear much greater in the first part of the period than they do in the second, with the exception of the influenza epidemic in 1918. Both these features of $\mathbf{k}$—its linear decline and its relatively constant variance—are very convenient for forecasting purposes. The linear decline of course is not an inherent feature of the index, and it would not hold if our series extended back into the 19th century.

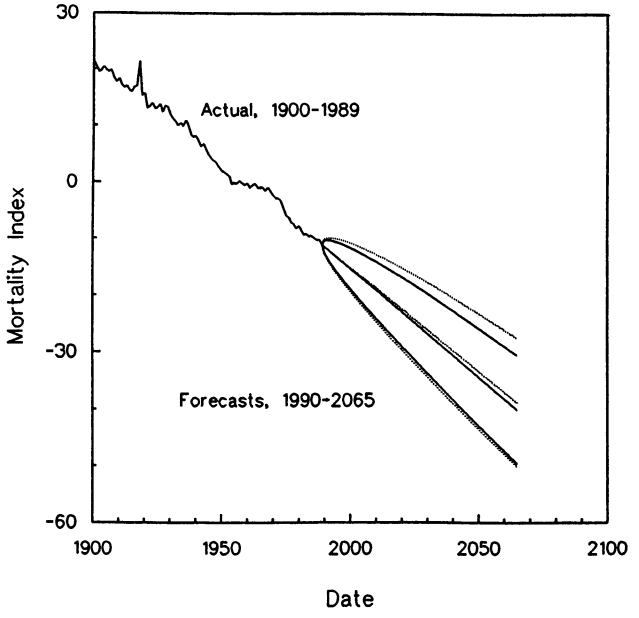

Figure 2. Comparison of Mortality Forecasts to 2065, From 1900–1989 (Dots) and From 1933–1989 (Solid), With 95% Confidence Band. Both forecasts use the (0, 1, 0) model; the forecast from 1900 has a dummy for the influenza epidemic.

Table 1. Fitted Values of $\mathbf{a}_x$ and $\mathbf{b}_x$ for 1933–1987 (SVD)

Age group	$\mathbf{a}_x$	$\mathbf{b}_x$
0	−3.64109	.09064
1–4	−6.70581	.11049
5–9	−7.51064	.09179
10–14	−7.55717	.08358
15–19	−6.76012	.04744
20–24	−6.44334	.05351
25–29	−6.40062	.05966
30–34	−6.22909	.06173
35–39	−5.91325	.05899
40–44	−5.51323	.05279
45–49	−5.09024	.04458
50–54	−4.65680	.03830
55–59	−4.25497	.03382
60–64	−3.85608	.02949
65–69	−3.47313	.02880
70–74	−3.06117	.02908
75–79	−2.63023	.03240
80–84	−2.20498	.03091
85–89	−1.79960	.03091
90–94	−1.40963	.03091
95–99	−1.03655	.03091
100–104	−.68035	.03091
105–109	−.34105	.03091

The estimates of $\mathbf{a}_x$ and $\mathbf{b}_x$ are given in Table 1. These can be used with forecasts of $\mathbf{k}$ (given later) to construct forecasts of age-specific death rates with confidence intervals (although see Appendix B for complications) and any other life table measures desired.

Figure 3 shows a sampling of fitted age-specific death rates compared to actual death rates (on a log scale) for 1933–1987. For most age groups, the fit is very good. The major exception is the rate for ages 20–24, which was discussed earlier. Note that this rate, although of interest in its own right, is very low relative to the others. Consequently, errors in modeling and forecasting this rate will have little impact on forecasts of life expectancy or the population age distribution.

Figure 4 shows the quality of the fit in a different way, by displaying the actual and fitted death rate schedules for all ages in 1933 and 1987. Once again we see that the fits are good, with the worst fits at the ages for which the rates are low (due in part to the reestimation of $\mathbf{k}$).

So far, we have treated the model of age-specific mortality change as if it were exact and deterministic. In fact it is approximate and does not account for all the variation in the rates. How close a fit does it provide? Viewing the fit for each age group separately, the lowest proportion of variance accounted for over the years 1933–1987 at any age is .830, for age group 15–19. (This is the ratio of the variance of

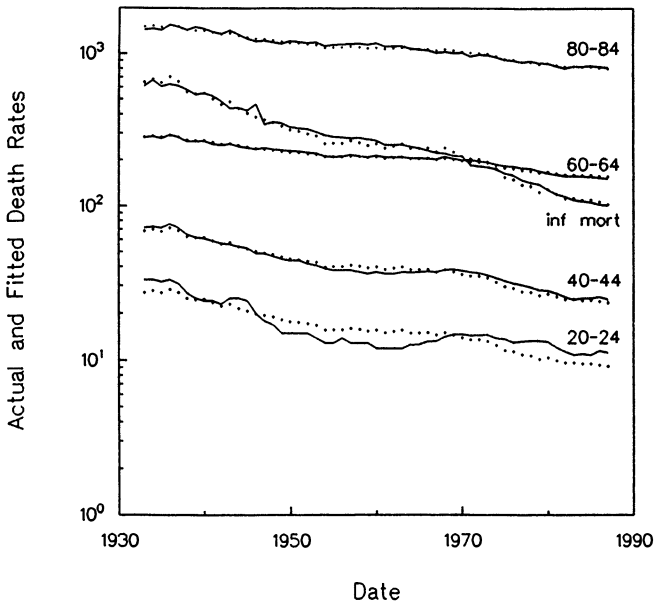

Figure 3. Actual and Fitted Log Death Rates, 1933–1987, Selected Ages. Death rates are fitted by model in text, with k recalculated.

differences between the actual and fitted rates to the variance of the actual rates.) For 13 of the 19 age groups, more than 95% of the variance over time is explained; for seven of these groups, more than 98% of this variance is explained. Unfortunately, for age group 85+, which is numerically very important, only 86% of the variance over time is explained. As noted earlier, the death rate in this age group fluctuates erratically, probably due more to measurement problems than to behavior of the true rate. In fact the mortality data above age 80 are highly suspect and fraught with measurement problems of various sorts (Coale and Kisker 1987). For this reason, and because it is important to have more age detail at the higher ages, we do not use the death rate for age group 85+, but instead impute death rates for age groups 85–90 to 105–109, as described earlier.

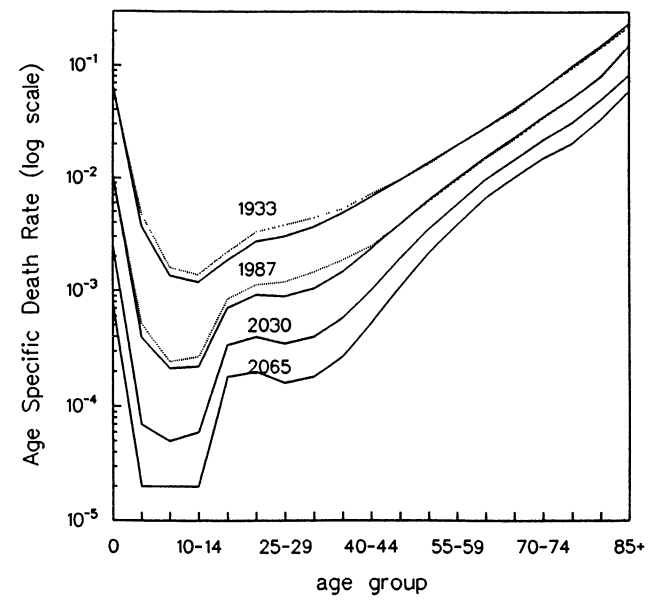

Figure 4. Comparison of Actual, Fitted, and Forecast Mortality by Age, for Sexes Combined, at Selected Dates. For 1933 and 1987, the last age group is 85+. For 2030 and 2065, it is 85–89 and, therefore, less than 85+. Actual schedule are dots. Fitted and forecasts are solid.

An overall measure of goodness of fit can be obtained by summing all the unexplained age group variances and taking their ratio to the sum of total variances over age groups. The advantage of this measure is that it reflects the differences in variability of death rates. This measure indicates that overall the fitted model accounts for .927 of the within-age group variance. Because age group 85+ is numerically important, and poorly fitted, it pulls down the overall figure considerably. If we ignore this age group, then the model explains 97.5% of the variance over time in all the other age groups, which is very good indeed.

One also might wonder whether the errors in a given year were correlated across age groups. We have calculated the correlation matrix for the differences between actual and fitted rates for the period 1933–1987. The correlations often are substantial and persist over very wide age gaps. This is unfortunate, because we would like to avoid incorporating this correlation matrix into the stochastic model. Appendix B provides an extended discussion of these errors and of the role of uncertainty in parameter estimates for $\mathbf{a}_x$ and $\mathbf{b}_x$.

For any value of $\mathbf{k}$ we can of course calculate the implied death rates $\mathbf{m}(x, \mathbf{k})$, and from these we can derive an entire life table and life expectancies. The function $\mathbf{m}(x, \mathbf{k})$ is nonlinear, and e_0 is in turn a nonlinear function of the $\mathbf{m}$'s. What then is the relation between $\mathbf{k}$ and e_0? Simulation reveals that it is nonlinear. Over the sample range, at lower levels of e_0 a decrease in $\mathbf{k}$ has an effect about twice as large as at higher levels of e_0, so a constant linear rate of decline in $\mathbf{k}$ will lead to a slowing rate of change in life expectancy. Their relationship may be understood in terms of the mortality entropy measure, H (Keyfitz, 1977, pp. 62–68) which declines as mortality falls. The plot of $\mathbf{k}$ against time (see Fig. 2) is fairly linear, whereas the plot of life expectancy against time (see Fig. 1) is not. This is an advantage of forecasting the death rates rather than the life expectancy directly. Direct forecasts of life expectancy must introduce some procedure for capturing the deceleration in gains, even when the rate of decline in death rates is undiminished. Some discussions of the purported deceleration in mortality decline fail to appreciate this point.

5. MODELING AND FORECASTING THE MORTALITY INDEX, K

Having developed and fitted the demographic model, we are now ready to move to the problem of forecasting. The first step is to find an appropriate ARIMA time series model for the mortality index $\mathbf{k}(t)$. After carrying out the standard model identification procedures (see Box and Jenkins 1970), we found that a random walk with drift describes $\mathbf{k}$ well (see Appendix A for details). The question arises of how to treat the influenza epidemic of 1918. One possibility is to view it as a highly unusual event whose inclusion in the series would inappropriately influence the results. In this case, we can fit an intervention model that uses a dummy variable for 1918, thereby effectively removing its influence. The other possibility is to view the epidemic as a kind of event that might

occur again in the future and thus conveys valuable information about the uncertainty of the forecast. In this case we can treat the observation like any other. Fortunately, it happens that only the confidence band of the mortality forecast, not the forecast itself, is affected by this decision. We have elected to treat the epidemic as an anomaly and use the intervention model. If we do not include a dummy for influenza in 1918, the parameter estimates and point forecasts are essentially unchanged but the confidence interval for **k** in 2065 (including uncertainty about drift) is 57% wider.

Our preferred model, estimated over 1900–1989, with standard errors in parentheses, is as follows:

$$\mathbf{k}_t = \mathbf{k}_{t-1} - .365 + 5.24\text{flu} + e_t, \qquad \text{see} = .651,$$
$$(.069) \quad (.461)$$
$$R^2 = .995.$$

The coefficient on flu indicates that **k** was 5.24 higher than otherwise expected in 1918. The constant term, −.365, indicates the average annual change in **k**, which drives the forecasts of long-run change in mortality. Over a 75-year horizon, we will forecast a decline in **k** of 75 times .365, or 27.4. The standard error of the estimate (see) indicates the uncertainty associated with a one-year forecast; as the forecast horizon increases, the standard error grows with the horizon's square root.

Figure 2 plots the past values of **k** along with the forecasts based on the time series model and the associated 95% confidence intervals. Two forecasts are shown: one based on a model fitted to the entire period 1900–1989, including a dummy for the influenza epidemic, and the other based on a model fitted solely to the years 1933–1989, when all states were included in the death registration area. It can be seen that the forecasts are very similar, both in expected values and in confidence bands. This also provides a rough test of the structural homogeneity of the period, because we can evaluate whether basing the forecast on the more recent data alters the results. Table 2 lists the forecasts based on the period 1900–1989, along with their standard errors. Later we will interpret these forecasts after first converting them into forecasts of death rates, survivorship, and life expectancy.

We also conducted a more thorough investigation of how base period length affects the forecasts by trying base periods starting in 1930, 1940, 1950, 1960, 1970, and 1980 and all ending in 1989. Table 3 shows the results. Evidently, the choice of starting dates from 1930 to 1960 makes little difference to either the point forecasts or the confidence intervals. A starting date of 1970, however, yields a much lower forecast for mortality; **k**(2065) is −47.0, versus −38.80 in our base run from 1900. This outcome reflects the exceptionally rapid mortality declines of the 1970s, when **k** fell on average by .548 per year versus a long-run average of .365. In the 1980s, **k** resumed decline at its long-run average rate (.363 for 1980–1989). Therefore, it is not clear why the estimates of drift and the forecast to 2065 both indicate slower decline (−32.2 in 2065 versus −38.80 in our base run). The fitted models and forecasts evidently exhibit some instability when the base period is reduced to 10 or 20 years. How severe are the implications for the derived forecasts of life

*Table 2. Forecasts of Mortality Index **k** With Standard Errors, (From (0, 1, 0) Model With Flu Dummy Estimated Over 1990–1989)*

Date	k	Standard deviation
1990	−11.41	.65
1991	−11.78	.92
1992	−12.14	1.13
1993	−12.51	1.30
1994	−12.87	1.46
1995	−13.24	1.60
1996	−13.60	1.72
1997	−13.97	1.84
1998	−14.33	1.95
1999	−14.70	2.06
2000	−15.06	2.16
2001	−15.43	2.26
2002	−15.79	2.35
2003	−16.16	2.44
2004	−16.52	2.52
2005	−16.89	2.61
2006	−17.25	2.69
2007	−17.62	2.76
2008	−17.98	2.84
2009	−18.35	2.91
2010	−18.71	2.98
2011	−19.08	3.05
2012	−19.44	3.12
2013	−19.81	3.19
2014	−20.18	3.26
2015	−20.54	3.32
2016	−20.91	3.38
2017	−21.27	3.45
2018	−21.64	3.51
2019	−22.00	3.57
2020	−22.37	3.63
2021	−22.73	3.68
2022	−23.10	3.74
2023	−23.46	3.80
2024	−23.83	3.85
2025	−24.19	3.91
2026	−24.56	3.96
2027	−24.92	4.01
2028	−25.29	4.07
2029	−25.65	4.12
2030	−26.02	4.17
2031	−26.38	4.22
2032	−26.75	4.27
2033	−27.11	4.32
2034	−27.48	4.37
2035	−27.84	4.42
2036	−28.21	4.46
2037	−28.57	4.51
2038	−28.94	4.56
2039	−29.30	4.61
2040	−29.67	4.65
2041	−30.03	4.70
2042	−30.40	4.74
2043	−30.76	4.79
2044	−31.13	4.83
2045	−31.49	4.87
2046	−31.86	4.92
2047	−32.22	4.96
2048	−32.59	5.00
2049	−32.95	5.04
2050	−33.32	5.09
2051	−33.68	5.13
2052	−34.05	5.17
2053	−34.41	5.21
2054	−34.78	5.25
2055	−35.14	5.29
2056	−35.51	5.33
2057	−35.87	5.37
2058	−36.24	5.41
2059	−36.61	5.45
2060	−36.97	5.49
2061	−37.34	5.53
2062	−37.70	5.56
2063	−38.07	5.60
2064	−38.43	5.64
2065	−38.80	5.68

Table 3. Forecasts of k From Various Base Periods Ending in 1989

Starting date k(2065)	Base per length	Constant (drift)	Standard error of constant	Standard error of equation	Forecast k(2065)	Standard error of forecast
1930	60	−.406	.071	.550	−41.9	4.80
1940	50	−.382	.071	.500	−40.1	4.36
1950	40	−.342	.079	.500	−37.1	4.36
1960	30	−.351	.094	.516	−37.7	4.50
1970	20	−.473	.104	.465	−47.0	4.06
1980	10	−.278	.137	.434	−32.2	3.78

NOTE: All base periods end in 1989. In each case, a random walk with drift was fit to the data.

expectancy at birth? Results to be discussed later indicate that in the neighborhood of $\mathbf{k} = -38.8$, a variation of unity in $\mathbf{k}$ implies a variation of .3 in life expectancy. Therefore, the forecast of $\mathbf{k}$ using data back to 1970 implies life expectancy greater by about 2.5 years; using data back to 1980 implies life expectancy less by about 2 years. From the point of view of forecasts based on relatively long time series, the results shown in the table are reassuring, because instability only emerges as the base period gets substantially shorter.

In addition, we tested our procedure by fitting models to earlier portions of the data set and forecasting over the later portions for which the outcome is known. Figure 5 shows forecasts from a model fitted to the period 1933–1962 and used to forecast over the 26-year period from 1963–1989. The forecast does very well over this horizon, which is 90% as long as the base period. In another test, we fitted a model to data for 1900–1944 and forecast from 1945–1989. This forecast (not shown) is right on target at the end of the period, and the actual $\mathbf{k}$ stays very well within the confidence band. Taken together, these experiments provide some basis for confidence in the approach taken. We must keep in mind, however, that no amount of analysis of the past can assure us that dramatic structural changes will not occur in the future.

As noted earlier, the estimation of the drift term is central to the forecasts, and of course its value cannot be known with certainty. The confidence intervals plotted in all the figures are based entirely on the estimated variance of the error term of the fitted model. For models without drift this is appropriate, because it has been shown that the effect of the standard error of parameter estimates is relatively small (Box and Jenkins 1970, pp. 267–269). In our model, taking account of parameter uncertainty increases the standard error of the forecast by less than 1% in the first year, by 6% after 10 years, by 25% after 50 years, and by 36% after 75 years (see Appendix B).

6. FORECASTS OF DEATH RATES AND LIFE EXPECTANCY

We now can use the mortality model to generate forecasts of the central death rates. Forecasts of the logs of representative rates are shown in Figure 6. Most of the forecasts match up well with the actual rates in 1989, but the rates for age groups 1–4 and 20–24 are forecast to drop quite suddenly. This is a consequence of the way the model was fit. A similar model easily could have been constrained to pass through the actual rates at the end of the base period, but with some deterioration in the goodness of fit for the rest of the base

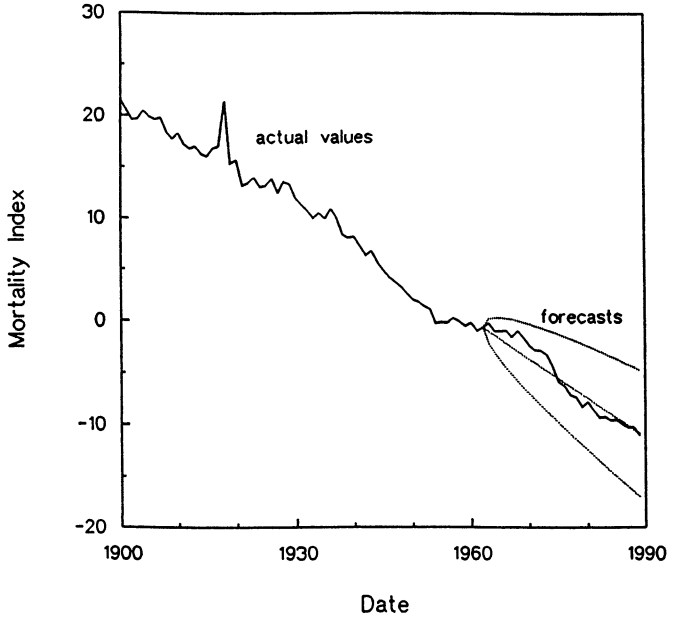

Figure 5. Within-Sample Forecast of Mortality Index, k Model Fitted on 1933–1962; Forecast for 1963–1989.

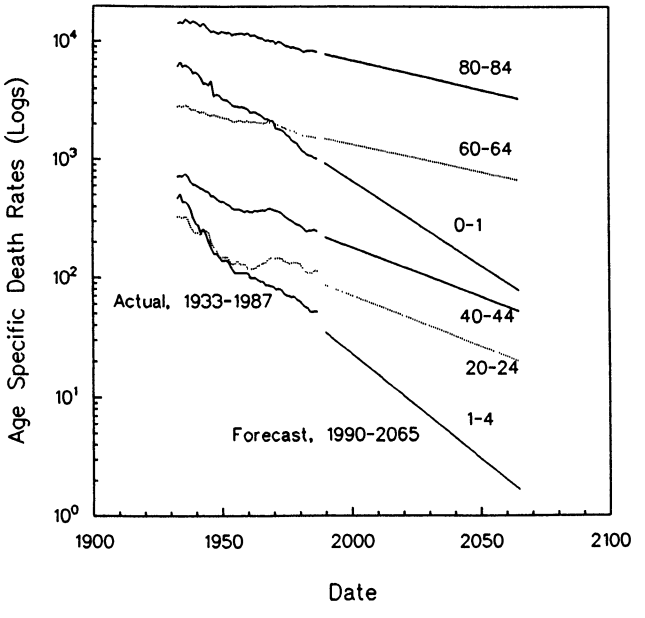

Figure 6. Selected Age-Specific Death Rates: Actual and Forecasts. Forecasts are based on 1900–1989, using model (0, 1, 0) with a flu dummy.

Table 4. Forecasts of Age-Specific Death Rates per 100,000 at Five-Year Intervals, 1990–2065 (Sexes Combined)

	Date								
Age group	1990	1995	2000	2010	2020	2030	2040	2050	2065
0–1	932	790	669	481	345	248	178	128	78
1–4	35	28	23	15	10	7	5	3	2
5–9	19	16	14	10	7	5	4	3	2
10–14	20	17	15	11	8	6	4	3	2
15–19	67	62	57	48	40	34	28	24	18
20–24	86	78	71	58	48	40	33	27	20
25–29	84	75	68	54	44	35	28	23	16
30–34	97	87	78	62	50	40	32	25	18
35–39	138	124	111	90	72	58	47	38	27
40–44	221	201	182	150	124	102	84	69	52
45–49	370	341	315	267	227	193	164	139	109
50–54	613	572	533	464	403	351	305	265	215
55–59	965	907	853	754	666	589	520	460	382
60–64	1,511	1,432	1,357	1,218	1,094	982	882	792	674
65–69	2,233	2,119	2,010	1,810	1,629	1,466	1,320	1,188	1,015
70–74	3,361	3,187	3,022	2,718	2,444	2,198	1,976	1,777	1,515
75–79	4,979	4,693	4,423	3,930	3,491	3,102	2,756	2,448	2,050
80–84	7,748	7,323	6,921	6,182	5,523	4,933	4,407	3,936	3,323
85–89	12,267	11,687	10,609	10,108	9,177	8,331	7,564	6,868	5,942
90–94	19,099	18,341	16,915	16,246	14,987	13,827	12,758	11,774	10,439
95–99	29,744	28,864	27,188	26,390	24,869	23,442	22,102	20,844	19,095
100–04	46,334	45,554	44,053	43,329	41,933	40,600	39,325	38,104	36,364
105–09	72,195	72,100	71,956	71,906	71,845	71,831	71,861	71,930	72,097

NOTE: Forecasts of the mortality index are based on the (0, 1, 0) model with a dummy for the flu year, fit to data from 1900–1989. See the text for the method of calculating the death rate forecasts from the forecasts of the mortality index. Confidence intervals for these forecasts may be calculated for each rate by using the procedure described in the text.

period. (To do this, set $\mathbf{a}_x$ equal to $\ln(\mathbf{m}_{x,T})$, where T is the terminal period, and apply SVD as before, disregarding the normalization rule for $\mathbf{k}$). We prefer to accept the discontinuity, which affects only rates that are absolutely very low in any event and have little influence on life expectancy.

Figure 4 shows the shapes of the age profiles that we forecast for 2030 and 2065. Two trends stand out in relation to the earlier profiles, which also are shown. First, the hump in mortality at young adult ages becomes more pronounced; second, by 2065 the death rates for age groups 1–4, 5–9, and 10–14 become virtually identical. The reader can assess the plausibility of these patterns.

Table 4 contains forecasts of all the five-year death rates for 1990, 1995, 2000, and every tenth year thereafter through 2065. Infant mortality rates are forecast to fall to less than one per thousand, which is perhaps implausibly low. Neither Table 4 nor Figure 6 gives confidence intervals for the death rate forecasts, but these can be derived using the information in Tables 1, 2, and 4 (see Appendix A). For shorter forecast horizons, these confidence intervals are too narrow; see the detailed discussion in Appendix B. The same information can be used to calculate death rate forecasts for every calendar year.

From the forecasts of death rates, it is straight forward to calculate life tables and life expectancy at birth. Table 5 contains forecasts of l_x (proportions surviving from birth to exact age x) for five-year age groups for periods as in Table 4. In 1987, about 46% of births survive to age 80 in a period life table. We forecast that by 2065, 46% will survive to age 90.

Figure 1 plots the forecasts of life expectancy along with their associated confidence intervals. There are several points to note. First, the forecast definitely is not a linear extrapolation of the trend in life expectancy, in contrast to the forecast for $\mathbf{k}$. Instead, life expectancy is forecast to increase slowly relative to past trend and at a rate that declines over the forecast horizon. Second, the confidence interval is surprisingly narrow; even by 2065, the 95% interval based only on the variance of the innovation terms is only plus 3.1 or minus 3.7 years; uncertainty from the estimated trend term raises these numbers to 4.1 and 5.2. These augmented intervals are shown by dotted lines in Figure 1. Third, of course, is the substance of the forecast, which foresees a 10.5-year increase in life expectancy over the next 75 years, in contrast to the Social Security Actuary and U.S. Census Bureau forecast of a 5.7-year increase (Wade 1989); reasons for the difference between the forecasts will be discussed later. Our forecast is low, however, compared to the recent extrapolative forecast by Guralnik, Yanagishita, and Schneider (1988), who anticipated life expectancy of 88.7 in 2040, whereas we expect only 83.1. Of the 10.5-year gain in life expectancy we forecast over the years 1989–2065, 2% of the gain in person-years lived occurs below age 15, 24% occurs in the working years 15–64, and 74% occurs at age 65 and older. Between 2064 and 2065, 86% of the gain in person-years lived occurs at age 65 and older. Table 6 shows forecasts of life expectancy at different ages and periods.

Why, as noted above previously, are the confidence bands on life expectancy so narrow? There are several reasons. One reason is that $\mathbf{k}$ hews so closely to a linear trend over the period, resulting in a narrow confidence band for its forecast. A second reason is that the entropy of the life table is decreasing, making life expectancy less responsive to errors in forecasting $\mathbf{k}$ and age-specific death rates. Thus, whereas the width of the 95% confidence interval for $\mathbf{k}$ in 2065 is 70% of its decline from 1900–1989, the comparable figure for e_0 is only 36%. A third reason is that we have chosen to remove the influence of the 1918 epidemic, and a fourth reason is that errors in forecasting individual death rates tend to be

Table 5. Forecasts of Numbers Surviving to Exact Ages Out of 100,000 Births at Five-Year Intervals, 1990–2065 (From Period Life Tables With Sexes Combined)

	Date									
Age	1990	1995	2000	2010	2020	2030	2040	2050	2060	2065
0	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000
1	99,120	99,253	99,366	99,544	99,672	99,764	99,830	99,878	99,912	99,926
5	98,985	99,143	99,276	99,483	99,631	99,737	99,812	99,866	99,904	99,919
10	98,890	99,062	99,208	99,434	99,596	99,712	99,794	99,853	99,895	99,911
15	98,791	98,977	99,134	99,380	99,556	99,682	99,773	99,837	99,883	99,901
20	98,458	98,671	98,853	99,143	99,357	99,514	99,631	99,718	99,783	99,809
25	98,033	98,285	98,503	98,854	99,118	99,318	99,469	99,585	99,673	99,710
30	97,622	97,915	98,170	98,586	98,902	99,143	99,329	99,472	99,582	99,628
35	97,148	97,490	97,789	98,280	98,657	98,947	99,172	99,346	99,482	99,539
40	96,480	96,888	97,247	97,840	98,301	98,659	98,939	99,158	99,330	99,402
45	95,420	95,921	96,365	97,108	97,694	98,157	98,523	98,814	99,046	99,144
50	93,669	94,297	94,860	95,818	96,590	97,214	97,718	98,128	98,461	98,604
55	90,837	91,636	92,362	93,620	94,661	95,523	96,239	96,835	97,333	97,550
60	86,550	87,565	88,500	90,153	91,556	92,749	93,765	94,632	95,373	95,702
65	80,235	81,500	82,681	84,813	86,673	88,296	89,714	90,953	92,037	92,528
70	71,723	73,275	74,745	77,451	79,872	82,036	83,969	85,694	87,234	87,940
75	60,561	62,419	64,204	67,561	70,643	73,464	76,039	78,384	80,514	81,504
80	47,098	49,256	51,364	55,423	59,256	62,850	66,201	69,311	72,183	73,532
85	31,780	33,972	36,166	40,531	44,821	48,992	53,007	56,840	60,471	62,206
90	16,953	18,681	20,470	24,202	28,091	32,078	36,108	40,131	44,101	46,055
95	6,290	7,220	8,226	10,462	12,983	15,763	18,771	21,968	25,314	27,030
100	1,301	1,569	1,874	2,608	3,520	4,621	5,920	7,418	9,113	10,031
105	104	131	163	248	363	515	710	955	1,254	1,427
110	0	0	0	0	0	0	0	0	0	0

NOTE: Forecasts of the mortality index are based on the (0, 1, 0) model with a dummy for the flu year, fit to data from 1900–1989. See the text for the method of calculating the death rate forecasts from the forecasts of the mortality index. Proportions surviving are derived from the death rates in Table 1 using standard procedures.

offsetting in the life expectancy calculation (see Appendix B).

We could have proceeded, as some have done, by modeling and forecasting life expectancy directly. The 50-year forecast (to 2038) from a direct ARIMA model of life expectancy is 92.3 years, with a confidence interval 13.9 years wide. The forecast from our preferred procedure, based on the mortality index, is 82.9 in 2038, with a confidence band only 6.4 years wide. Evidently the procedure suggested here leads to dramatically different forecasts that are both lower and (apparently) more certain. We believe our procedure incorporates prior understanding of the nature of the process of mortality

Table 6. Forecasts of Remaining Life Expectancy at Exact Ages at Five-Year Intervals, 1990–2065 (From Period Life Tables With Sexes Combined)

	Date									
Age	1990	1995	2000	2010	2020	2030	2040	2050	2060	2065
0	75.83	76.68	77.49	79.04	80.48	81.84	83.13	84.34	85.50	86.05
1	75.50	76.25	76.99	78.40	79.75	81.04	82.27	83.44	84.57	85.12
5	71.60	72.33	73.05	74.45	75.78	77.06	78.28	79.45	80.58	81.12
10	66.66	67.39	68.10	69.48	70.81	72.08	73.30	74.46	75.59	76.13
15	61.73	62.45	63.15	64.52	65.83	67.10	68.31	69.48	70.59	71.13
20	56.93	57.63	58.32	59.67	60.96	62.21	63.40	64.56	65.66	66.20
25	52.16	52.85	53.52	54.83	56.10	57.32	58.50	59.64	60.73	61.26
30	47.37	48.04	48.69	49.98	51.22	52.42	53.58	54.70	55.78	56.31
35	42.59	43.24	43.87	45.12	46.34	47.52	48.66	49.77	50.84	51.36
40	37.87	38.49	39.10	40.31	41.50	42.65	43.77	44.86	45.91	46.42
45	33.26	33.85	34.44	35.60	36.74	37.85	38.94	40.00	41.04	41.54
50	28.83	29.39	29.94	31.04	32.13	33.19	34.24	35.27	36.26	36.75
55	24.64	25.16	25.68	26.71	27.73	28.73	29.73	30.70	31.65	32.12
60	20.73	21.21	21.68	22.63	23.57	24.51	25.44	26.35	27.25	27.68
65	17.16	17.59	18.02	18.89	19.75	20.61	21.47	22.31	23.14	23.54
70	13.88	14.27	14.65	15.43	16.21	16.99	17.75	18.52	19.26	19.63
75	10.96	11.30	11.63	12.31	12.99	13.66	14.33	15.00	15.65	15.98
80	8.36	8.63	8.90	9.44	9.98	10.53	11.08	11.62	12.16	12.42
85	6.18	6.37	6.57	6.97	7.37	7.78	8.18	8.59	9.00	9.20
90	4.46	4.59	4.72	4.99	5.27	5.54	5.82	6.10	6.38	6.52
95	3.10	3.18	3.27	3.43	3.60	3.77	3.93	4.10	4.26	4.34
100	1.91	1.98	2.04	2.16	2.27	2.37	2.46	2.55	2.64	2.68
105	.00	.00	.11	.31	.48	.63	.76	.88	.99	1.04
110	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

NOTE: Forecasts of the mortality index are based on the (0, 1, 0) model with a dummy for the flu year, fit to data from 1900–1989. See the text for the method of calculating the death rate forecasts from the forecasts of the mortality index. Life expectancies are derived from the death rates in Table 1 using standard procedures.

decline, which after all is effected only through reductions in death rates at specific ages.

7. COMPARISON TO SOCIAL SECURITY FORECAST

Our forecasts differ significantly in both method and outcome from those prepared by the Actuary of the Social Security Administration (Wade 1989) and used by the U.S. Census Bureau as well. The Actuary's forecasts combine age-specific trend extrapolation with the views of medical experts on ultimate cause-specific (but not age-specific) rates of mortality decline, which are phased in gradually, with completion by 2010. The medical expert's views make the ultimate rate of decline substantially slower than the historic trends and imply that all age-specific rates ultimately decline at the same pace, which is sharply inconsistent with past age-specific trends. Confidence intervals (high-low intervals) are assigned subjectively. (See Alho and Spencer [1989] for a derivation of objective confidence intervals for these forecasts.) Alho (1990) found that the use of experts by the Actuary hindered rather than helped the forecasts in the past, in the sense that statistical time series models would have performed better.

We forecast a life expectancy of 86.05 by 2065; the Actuary projects a life expectancy of 80.45 (sexes averaged for the Alternative II forecast), or 5.6 years less. The age distribution of mortality in the Actuary's forecasts also is quite different than ours. The Actuary forecasts a female life expectancy of 84.7 in 2070; we forecast an identical sexes-combined level for 2054. Comparison of the age-specific death rates in the two forecasts reveals a clear pattern of differences. Up to age group 30–34, our death rate forecasts are much lower than the Actuary's; however, after this age, their forecasts are much lower than ours (Wade 1989). This pattern is to be expected, because their procedure eventually makes all age-specific death rates decline at the same rate and ours let each decline at its own specific historical rate. Because death rates have tended to decline more rapidly at younger ages and less rapidly at older ages, the observed pattern is generated.

To assess the implications of our mortality forecasts for population aging, we used them to prepare population forecasts based on assumptions that were otherwise identical to those of the "middle" scenarios of two official forecasts: Specifically, we replicated the fertility and migration assumptions of each, but combined these with our mortality forecasts. The Social Security Administration foresees an elderly dependency ratio (65+/20–64) of .411 in 2065, whereas we forecast .470; our forecast implies that payroll taxes will need to be 14% higher than the figure under their forecast, other things equal. The comparison with Census Bureau forecasts is their .436 versus our .480; our forecast implies a 10% higher payroll tax rate in 2065.

8. THE ROLE OF AIDS

The first reported cases of AIDS in the United States occurred in 1981; the numbers have grown each year since. The resulting AIDS deaths are of course reflected in the total registered deaths from which our mortality index, $\mathbf{k}$, is computed. By 1986 AIDS deaths raised adult male and female mortality in the United States by .7 and .07% (Curran 1988); AIDS is now the fifteenth leading cause of death overall. Calculating the rate of decline in $\mathbf{k}$ from 1980–1989, a period that spans rising mortality from AIDS, we find an average annual decline of .363, as compared to .365 in our fitted model for 1900–1989. This is essentially zero difference, suggesting that AIDS has not yet led to a structural discontinuity of the sort that would invalidate these forecasts. We must remember that the past 90 years have witnessed many important changes in diseases and their treatment, including the discovery of penicillin and other antibiotics.

One alternative to ignoring AIDS altogether would be to incorporate extraneous forecasts of AIDS mortality, including its age incidence, by simply adding them to the forecasts we have given, with a slight adjustment for competing risks. Given the current state of knowledge, however, the problems in forecasting U.S. AIDS mortality appear so considerable that nothing useful along these lines could be done (Curran 1988; Gail and Brookmeyer 1988; Lamp et al. 1990). For these reasons, we have not incorporated AIDS mortality into our forecasts in any specific way.

9. SUMMARY AND CONCLUSION

We have constructed a simple but powerful one-parameter family of life tables that closely fits the pattern of age-specific mortality in the United States from 1933 to the present. The time trend in $\mathbf{k}$, the parameter of the fitted life tables, is essentially linear from 1900 to the present; the parameter declines by the same amount in each of the two halves of the 90-year period. We use standard statistical methods to model and forecast the index of mortality as a random walk with drift, which implies that each age group's mortality continues to decline at its own age-specific exponential rate. From the forecasts of rates and their confidence intervals, we construct forecasts of life expectancy. We anticipate that it will rise by about 10 years to 86.05 in the year 2065, with a 95% confidence band of plus 3.9 or minus 5.6 years (including uncertainty about the trend term). Experiments indicate that the method performs well over intervals for which mortality is known, and that major variations in the length of the base period have relatively modest effects on the forecasts for the year 2065. The analysis also demonstrates that for life expectancy to rise to such a high value as 100 by 2065 would require a radical break in historical trends.

All purely extrapolative forecasts assume that the future will be in some sense like the past, and this forecast is no exception. Information from other sources may help us anticipate ways in which the future will be unlike the past and thereby lead to improved forecasts, although use of such information has its own problems. In any event, because this forecast spells out carefully what will happen if long-run trends continue, it can serve as a useful benchmark for assessing the assumptions and outcomes of other forecasts.

APPENDIX A: COMPUTATIONAL PROCEDURES

The SVD procedure is available in many statistical packages; we used Gauss. If SVD is not available, a close approximation to the SVD solution for this application can be found as follows. Choose the normalization used above; the $\mathbf{k}$'s sum to 0 and the $\mathbf{b}$'s to unity. Then $\mathbf{a}_x$ must equal the average over time of $\ln(\mathbf{m}_{x,t})$. (This follows

from setting the average value of $\mathbf{k}$ to 0.) Furthermore, $\mathbf{k}_t$ must (very nearly) equal the sum over age of $(\ln(\mathbf{m}_{x,t}) - \mathbf{a}_x)$, since the sum of the $\mathbf{b}_x$s has been chosen to be unity (this is not an exact relation, however, since the error terms will not in general sum to 0 for a given age). All that remains, then, is to estimate the $\mathbf{b}_x$s. Each $\mathbf{b}_x$ can be found by regressing, without a constant term, $(\ln(\mathbf{m}_{x,t}) - \mathbf{a}_x)$ on $\mathbf{k}_t$ separately for each age group x.

Time series estimation was done by Regression Analysis for Time Series (RATS) using a nonlinear least squares procedure. A similar model with an ar(1) term added was marginally superior, but we preferred the (0, 1, 0) model on grounds of parsimony. The preferred model had a DW statistic of 2.42 and a Ljung–Box Q statistic of 37, significant at the .090 level. We also experimented with numerous other specifications; all gave essentially the same point forecasts and differed only slightly on the confidence bands. We also used an automated nonparametric frequency domain method to forecast $\mathbf{k}$. This method, which was carried out without an intervention term for the influenza epidemic, generated a point mortality forecast nearly identical to the one based on time domain Box–Jenkins method (see Doan 1990, section 7-12).

To make the old age adjustments, we first estimated $\mathbf{k}$ in the way described, using observed death rates only up to 85+. We then used the implied fitted death rates for 75–79 and 80–84 to derive the death rates up to 105–109, using the Coale and Guo procedure. We did not adjust any death rates between 65 and 84.

If we ignore other sources of error (see Appendix B), then the confidence bounds on $\mathbf{k}$ can be used to calculate confidence bounds on life expectancy, because it is assumed that all age-specific rates vary together over time with $\mathbf{k}$ (correlations among logs of rates are assumed to be unity), so there is no cancellation of variations in rates. The approximate upper 95% confidence bound for the forecast of $\mathbf{m}_{x,t}$ is given by the point forecast in Table 4 times $\exp(2\mathbf{b}_x se_{kt})$; for the lower bound, the factor is $\exp(-2\mathbf{b}_x se_{kt})$. $\mathbf{b}_x$ is given in Table 1, and se_{kt} is given in Table 2.

APPENDIX B: THE STRUCTURE AND MAGNITUDE OF ERRORS IN THE MORTALITY FORECAST

Our analysis of forecast error so far has focused on the error in forecasting the mortality index, $\mathbf{k}$, while ignoring errors in fitting the original matrix of mortality rates. This appendix explores the forecast errors more completely.

The log of each age-specific mortality rate is forecast s periods ahead from base period t using the following equation:

$$\text{forecast}[\ln(\mathbf{m}_{x,t+s})] = \hat{\mathbf{a}}_x + \hat{\mathbf{k}}_{t+s}\hat{\mathbf{b}}_x,$$

where the $\hat{}$ indicates an estimate (for $\mathbf{a}_x$ and $\mathbf{b}_x$) of a forecast (for $\mathbf{k}_t$). The true value of $\ln(\mathbf{m}_{x,t+s})$, assuming the model specification and data are correct, is given by:

$$\ln(\mathbf{m}_{x,t+s}) = (\hat{\mathbf{a}}_x + \alpha_x) + (\hat{\mathbf{k}}_{t+s} + u_{t+s})(\hat{\mathbf{b}}_x + \beta_x) + \varepsilon_{x,t+s},$$

where α_x and β_x are the errors in estimating $\mathbf{a}_x$ and $\mathbf{b}_x$ and u_{t+s}, is, as before, the error in forecasting $\mathbf{k}$ ahead s periods from base period t.

The forecast error, $E_{x,t+s}$, is their difference:

$$E_{x,t+s} = \alpha_x + \varepsilon_{x,t+s} + (\hat{\mathbf{b}}_x + \beta_x)u_{t+s} + \beta_x\hat{\mathbf{k}}_{t+s}.$$

There are four error terms in this expression, but actually u is itself the sum of two errors due to innovations and error in estimating drift. To incorporate forecast uncertainty in $\mathbf{k}$ due to the estimate of the drift term, we reasoned as follows. Innovation terms after 1989 must be independent of the error in estimating the drift term; therefore, their influence on the forecast error is additive with that of the error terms. If the forecast interval is s, then the associated forecast error variance is the sum of the cumulated error variance, $s*\text{see}^2$, and the squared error associated with estimate of the drift, $(s*\text{sec})^2$, where sec is the standard error of the estimate of the constant. Thus the standard error of the forecast is the square root of $[s*\text{see}^2 + (s*\text{sec})^2]$. In our model, see = .653 and sec = .0696.

It is not clear exactly how these different sources of error may be correlated with one another. Informal experiments and use of the bootstrap indicate no correlation of errors in estimates of $\mathbf{a}_x$ and $\mathbf{b}_x$ and modest negative correlations across ages in the $\mathbf{b}_x$ estimates. To get an idea of their relative importance, we will assume them to be independent. Under this assumption the variance of $E_{x,t+s}$ is given by

$$\sigma^2_{Ex,t+s} = \sigma^2_{\varepsilon x,t+s} + \sigma^2_{\alpha x} + \hat{\mathbf{b}}^2_x\sigma^2_{ut+s} + \sigma^2_{\beta x}(\hat{\mathbf{k}}^2_{t+s} + \sigma^2_{ut+s}).$$

The question is whether the term $\hat{\mathbf{b}}^2_x\sigma^2_u$ dominates the others, for this is the only term included in the forecast intervals shown in the tables and graphs.

The variance of the term $\varepsilon_{x,t+s}$ is estimated by the variance of the error in fitting age group x within the sample period, which is readily calculated. $\hat{\mathbf{a}}_x$ is simply the average over time of the log of the death rate for age x, so its error variance is the variance of $\ln(\mathbf{m}_{x,t})$ divided by T, the number of observations of $\mathbf{m}_x$.

The error variance in the estimate of $\mathbf{b}_x$ is obtained less easily. An analytic expression is available for errors in SVD estimates in the usual case when these arise from sampling error or measurement error in the data. In our application such errors are minimal, and errors arise instead from genuine departures of the rates from the simple model posited due to omitted historical influences. For this reason we used a bootstrap procedure to estimate the variance of the estimator of $\mathbf{b}_x$. We constructured 400 pseudo-data matrices by adding errors to the matrix $\{\ln(\mathbf{m}_{x,t}) - \ln(\mathbf{m}_{x,\cdot})\}$. The errors were obtained by sampling with replacement from the corresponding age column of the matrix of residuals from the initial fit of the model to the matrix of observed rates (that is, based on the first elements of the SVD analysis). We then used our SVD procedure to reestimate $\{\mathbf{b}_x\}$ for each pseudo-data matrix. From the results we calculated variances and covariances for $\{\mathbf{b}_x\}$. Following Freedman and Peters (1984, p. 151), we did not inflate the errors to compensate for their reduction in fitting.

Table B1 evaluates the share of the forecast error for the logs of a few representative death rates in 2065 arising from each source. It shows that in 2065 the error in forecasting the mortality index dominates the errors in fitting the mortality matrix. Even in the

Table B1. Sources of Error in Log Age-Specific Death Rate Forecasts for 2065

Age group	Total σ^2_E	fit $\mathbf{m}$ σ^2_ε	$\mathbf{a}_x$ σ^2_α	$\mathbf{b}_x$ $(\sigma_\beta\hat{\mathbf{k}}_{t+s})^2$	$\mathbf{k}$, $\mathbf{b}$ errs $\sigma^2_{\beta x}\sigma^2_u$	$\mathbf{k}$ forc $\beta^2\sigma^2_u$	$\mathbf{k}$ forc share σ^2	$\mathbf{k}$ forc share σ
0	.6682	.0156	.0054	.0097	.00039	.6371	.953	.976
20–24	.2449	.0115	.0020	.0089	.00036	.2221	.907	.952
40–44	.2199	.0010	.0018	.0008	.00003	.2163	.984	.992
60–64	.0714	.0018	.0006	.0014	.00005	.0675	.946	.973
80–84	.0796	.0028	.0006	.0019	.00008	.0742	.932	.966

NOTE: The forecasts of $\mathbf{k}$ for 2065 is −38.8, which was used to calculate $(\sigma^2_\beta\hat{\mathbf{k}}_{t+s})^2$. The forecast variance for $\mathbf{k}$ in 2065, σ^2_u, is 60.39.

Table B2. k's Share of the Standard Errors of Age-Specific Death Rate Forecasts and Life Expectancy

	Date (forecast horizon)			
Age group	1990 (1)	1999 (10)	2014 (25)	2065 (76)
0	.37	.80	.93	.98
20–24	.28	.69	.88	.95
40–44	.54	.90	.97	.99
60–64	.36	.78	.93	.97
80–84	.32	.74	.91	.97
e_0	.81	.98	.99	1.00

worst of the five age groups examined, the standard error based solely on the forecast of **k** is 95% of the standard error based on all sources.

We also examined the sources of error for shorter run forecasts, for which the outcome is not so favorable. For the five selected age groups, Table B2 gives the size of the standard error derived from **k** alone as a proportion of the standard error calculated for all sources (corresponding to the last column of Table B1). From these figures we must conclude that the confidence intervals based on **k** alone seriously understate the errors in forecasting individual age-specific death rates over shorter horizons, say less than 15 years.

What about life expectancy? When passing from errors in forecasting the age-specific death rates to errors in forecasting life expectancy there is bound to be very substantial cancellation of age-specific errors, whereas the error in forecasting **k** applies to all age groups and does not cancel at all. Therefore, error in **k** should be relatively more important in this case. We will consider this point in some detail.

First, note that errors in the log of mortality differ from those in mortality itself. If $\ln(\mathbf{m}_{x,t+s}) = \hat{\ln}(\mathbf{m}_{x,t+s}) + E_{x,t+s}$, then the error in $\mathbf{m}_{x,t+s}$ is approximately $E_{x,t+s}\hat{\mathbf{m}}_{x,t+s}$ (since $\exp(E_{x,t+s})$ is approximately $1 + E_{x,t+s}$ for small values of $E_{x,t+s}$), which we will call $\theta_{x,t+s}$.

Second, note that if $\mathbf{m}_{x,t+s}$ is overestimated by an amount $\theta_{x,t+s}$, then $\hat{e}_0$ for time $t + s$ will be reduced by roughly $\theta_{x,t+s}T_{x,t+s}$, where $T_{x,t+s}$ is the total years lived above the middle point of the age group x and the life table radix, l_0, is unity. This is a linear approximation that evaluates $T_{x,t+s}$, a random variable, at its expected value. The error $\theta_{x,t+s}$ consists of a component $\eta_{x,t+s}$, which is perfectly correlated across x, and a component $\phi_{x,t+s}$, which is assumed independent across x. Letting $\mathcal{E}$ denote an expectation, we can then express the approximate error variance of $\hat{e}_0$ as a function of the error components as

$$\mathrm{Var}(\hat{e}_0) = \mathcal{E}(\sum \eta_{x,t+s}T_{x,t+s})^2 + \mathcal{E}\{\sum (\phi_{x,t+s}T_{x,t+s})^2\}.$$

Replacing $\eta_{x,t+s}$ by $\hat{\mathbf{m}}_{x,t+s}u_{x,t+s}\mathbf{b}_x$ and $\phi_{x,t+s}$ by $\hat{\mathbf{m}}_{x,t+s}(\alpha_x + \varepsilon_{x,t+s} + \beta_x\hat{\mathbf{k}}_{x,t+s} + \beta_x u_{t+s})$, and substituting variances, we get

$$\mathrm{Var}(\hat{e}_0) = \sigma^2_{ut+s}(\sum \mathbf{b}_x\hat{\mathbf{m}}_{x,t+s}T_{x,t+s})^2 + \sum (\hat{\mathbf{m}}_{x,t+s}T_{x,t+s})2 \times (\sigma^2_{\alpha x} + \sigma^2_{\varepsilon x,t+s} + \sigma^2_{\beta x}\hat{\mathbf{k}}2_{t+s} + \sigma^2_{\beta x}\sigma^2_{ut+s}).$$

This expression can be evaluated to find the approximate error variance of life expectancy forecasts at each horizon, under the assumption that different sources of error at a given age are uncorrelated and that the error sources other than **k** are uncorrelated across age. These assumptions probably overstate the contribution of these non-**k** errors to uncertainty in life expectancy forecasts, because although the age group residuals $\varepsilon_{x,t+s}$ are indeed correlated across age, by construction of **k** these errors must exactly cancel in the sense that the fitted death rates must produce exactly the correct number of deaths in each sample calendar year, given the actual population age distribution. Any tendency for the forecasts to be wrong in the same direction for all age groups is in principle absorbed by **k**. The calculations of the standard error of e_0 ignore this constraint and thus probably overstate the role of the non-**k** errors.

Results of evaluating the expression are shown in the bottom row of Table B2. In 1990 **k** accounts for 81% of the standard error in the forecast of life expectancy; by 1994 (not shown) this rises to 95%. The approximation for $\mathrm{Var}(\hat{e}_0)$ developed above is evidently good, because when restricted to the term in σ^2_{ut}, it matches the exact result very closely.

We conclude that for life expectancy forecasts, it is reasonable to restrict attention to the errors in forecasting the mortality index and to ignore those in fitting the mortality matrix, even for shortrun forecasts. For individual death rate forecasts, however, confidence intervals based on **k** alone are a reasonable approximation only for forecast horizons greater than 10 to 25 years. If there is particular interest in forecasting these individual rates over the shorter term, then the other sources of error can be incorporated in the confidence interval as was done previously.

[*Received September 1990. Revised July 1991.*]

REFERENCES

Alho, J. M. (1990), "Stochastic Methods in Population Forecasting," *International Journal of Forecasting,* 6, 521–530.

Alho, J. M., and Spencer, B. D. (1990), "Error Models for Official Mortality Forecasts," *Journal of the American Statistical Association,* 85, 609–616.

Alter, G. (1990), "Old Age Mortality and Age Misreporting In the United States, 1900–1940," paper presented at the 1990 Meetings of the Population Association of America, Toronto.

Bozik, J., and Bell, W. (1989), "Time Series Modeling for the Principal Components Approach to Forecasting Age-Specific Fertility," paper presented at the 1989 Meetings of the Population Association of America, Baltimore.

Coale, A., and Guo, G. (1989), "Revised Regional Model Life Tables At Very Low Levels of Mortality," *Population Index,* 55, 613–643.

Coale, A., and Kisker, E. E. (1990), "Defects in Data on Old Age Mortality in the United States: New Procedures for Calculating Approximately Accurate Mortality Schedules and Life Tables at the Highest Ages," *Asian and Pacific Population Forum,* 4(1), 1–31.

Curran, J. W., Jaffe, H. W., Hardy, A. M., Morgan, W. M., Selik, R. M., and Dondero, T. J. (1988), "Epidemiology of HIV Infection and AIDS in the United States," *Science,* 239, 610.

Doan, Thomas, A. (1990), *RATS User's Manual: Version 3.10,* VAR Econometrics, Inc., Evanston, IL.

Freedman, D. A., and Peters, S. C. (1984), "Bootstrapping an Econometric Model: Some Empirical Results," *Journal of Business and Economic Statistics,* 2, 150–158.

Gail, M., and Brookmeyer, R. (1988), "Methods for Projecting Course of Acquired Immunodeficiency Syndrome Epidemic," *Journal of the National Cancer Institute,* 80, 900–911.

Good, I. J. (1969), "Some Applications of the Singular Decomposition of a Matrix," *Technometrics,* 11, 823–831.

Gomez de Leon, J. (1990), "Empirical DEA Models to Fit and Project Time Series of Age-Specific Mortality Rates," unpublished manuscript, Central Bureau of Statistics, Norway.

Grove, R. D., and Hetzel, A. M. (1968), *Vital Statistics Rates in the United States 1940–1960,* U.S. Department of Health, Education and Welfare, Public Health Service, National Center for Health Statistics, Washington, DC: U.S. Government Printing Office.

Guralnik, J. M., Yanagishita, M., and Schneider, E. L. (1988), "Projecting the Older Population of the United States: Lessons From the Past and Prospects for the Future," *The Millbank Quarterly,* 66, 283–308.

Heligman, L. (1984), "Construction of the New United Nations Model Life Table System" in *Methodologies for the Collection and Analysis of Mortality Data,* eds. J. Vallin, J. H. Pollard, and L. Heligman, Ordina Liege, pp. 179–202.

Keyfitz, N. (1977), *Applied Mathematical Demography,* New York: John Wiley.

——— (1981), "Choice of Function for Mortality Analysis: Effective Forecasting Depends on a Minimum Parameter Representation," *Theoretical Population Biology,* 21, 329–352.

Lamp, G. F., Payne, S. F., Rutherford, G. W., Hessol, N. A., Winkelstein,

W., Jr., Wiley, J. A., Moss, A. R., Chaisson, R. E., Chen, R. T., Feigel, D. W., Jr., Thomas, P. A., and Werdegar, D. (1990), "Projections of AIDS Morbidity and Mortality in San Francisco," *Journal of the American Medical Association,* 263, 1497–1501.

Land, K. (1986), "Methods for National Population Forecasts: A Review," *Journal of the American Statistical Association,* 81, 888–901.

Lederman, S. (1969), "Nouvelles Tables-Type de Mortalite," *Travaux and Documents,* n. 53, Paris: Institut National d'Etudes Demographiques.

Lee, R. (1974), "Estimating Series of Vital Rates and Age Structures From Baptisms and Burials: A New Technique," *Population Studies,* 28, 495–512.

——— (1977), "Methods and Models for Analyzing Historical Time Series of Births, Deaths and Marriages," in *Population Patterns in the Past,* ed. R. Lee, New York: Academic Press, pp. 337–370.

McNown, R., and Rogers, A. (1990), "Forecasting Cause-Specific Mortality Using Time Series Methods," paper presented at the 1990 Meetings of the Population Association of America, Toronto.

Olshansky, S. J. (1988), "On Forecasting Mortality," *Millbank Quarterly,* 66, 482–530.

Rogers, A. (1986), "Parameterized Multistate Population Dynamics and Projections," *Journal of the American Statistical Association,* 81, 48–61.

Wade, A. (1989), *Social Security Area Population Projections: 1989,* Actuarial Study No. 105, Washington DC: Office of the Actuary, Social Security Administration.

Wilmoth, J. R. (1990), "Variation in Vital Rates by Age, Period, and Cohort," *Sociological Methodology: 1990,* Basil Blackwell Ltd. for The American Sociological Association, pp. 295–336.

Wilmoth, J. R., Vallin, J., Caselli, G. (1989), "Quand Certaines Generations Ont Une Mortalite Differente de Celle Que l'on Pourrait Attendre," *Population,* 44, 335–376.

Comment

ROBERT McNOWN*

A common criterion for evaluating forecasting methodologies is the accuracy with which the forecasts match the eventual realizations of the actual data. Users of forecasts must choose among alternative forecasting models before such facts become available, however. Unwilling to wait for several decades to pass before writing this comment, I analyze the methodology of Lee and Carter using criteria suggested by Keyfitz (1981), Long (1984), and Rogers and Woodward (1991): (a) the transparency of assumptions used to generate the forecasts, (b) the ability of the model to generate measures of forecast uncertainty, (c) the extent of disaggregation (by age, sex, and race) permitted by the model, and (d) the quality of the data on which the forecasts are based.

Lee and Carter have presented a methodology for forecasting mortality that falls squarely within the extrapolative tradition of demographic forecasting. They intend their forecasts to be more than illustrative, presenting projections that clearly trace the implications of the continuation of historical trends. These projections can serve as a benchmark for the comparison and evaluation of official forecasts. As a set of benchmark forecasts, these will be most useful if we clearly understand the process generating the forecasts.

The singular value decomposition (SVD) is applied to the matrix of age-specific central death rates for each calendar year in this century. This decomposition yields an index of mortality, $\mathbf{k}_t$, and a set of age-specific constants, $\mathbf{b}_x$, which relate the central death rates, $\mathbf{m}_{x,t}$, to the index of mortality. In particular, $d(\ln \mathbf{m}_{x,t})/dt = \mathbf{b}_x dk_t/dt$. Because $\mathbf{k}_t$ is modeled as a random walk with drift, it is projected to decline at a constant linear rate. Therefore, each age-specific mortality rate is predicted to decline at its own constant exponential rate, as determined by the individual $\mathbf{b}_x$ parameters. In fact, as stated at several points in the article and as shown in their Figure 6, each central death rate is forecasted to decline "at its own specific historical rate" (p. 665). *From this analysis it becomes clear that forecasts of mortality identical to Lee and Carter's will be produced by directly projecting each age-specific mortality rate at its own historical rate of exponential decline.* Because the forecasts produced by straight extrapolation of individual mortality rates are the same as those generated indirectly from Lee and Carter's mortality index, the two methods are equivalent, despite the authors' statements to the contrary (p. 661).

Viewed in these terms, the Lee and Carter forecasts can be given high marks in terms of transparency of assumptions. Their projections capture the implications of a continuation of past exponential trends in age-specific mortality rates, uncomplicated by expert opinion or assumptions about medical advances, delay of deaths by cause, or ultimate levels of life expectancy.

Lee and Carter claim that "variation in a single parameter can generate the main outlines of the observed pattern" (p. 660), meeting Keyfitz's (1981) criterion of parsimony in representing the mortality profile. Actually their methodology involves 23 parameters—the $\mathbf{b}_x$ coefficients that capture the rates of change in each age-specific mortality rate, relative to changes in the mortality index $\mathbf{k}_t$.

There are disadvantages to mortality forecasts that are straight projections of individual age-specific central death rates. One concern is that if each age-specific rate is allowed to change at its own individual rate, the projected age profile of mortality may depart from plausible, historically observed patterns (Keyfitz 1981). Evidence of this outcome is seen in Lee and Carter's Figure 4, in which the projected profiles for 2030 and 2065 are characterized by several irregularities: mortality rates that are constant across ages 1–14, a very sharp rise to a prominent accident peak for young adults, and a strong upward curvature of senescent mortality rates

* Robert McNown is Professor, Department of Economics, University of Colorado, Boulder, CO 80309.

Journal of the American Statistical Association
September 1992, Vol. 87, No. 419, Applications & Case Studies