Demography, Vol. 26, No. 4, November 1989

Forecasting Mortality: A Parameterized Time Series Approach

Robert McNown and Andrei Rogers
Population Program, University of Colorado,
Boulder, Colorado 80309

This article links parameterized model mortality schedules with time series methods to develop forecasts of U.S. mortality to the year 2000. The use of model mortality schedules permits a relatively concise representation of the history of mortality by age and sex from 1900 to 1985, and the use of modern time series methods to extend this history forward to the end of this century allows for a flexible modeling of trend and the accommodation of changes in long-run mortality patterns. This pilot study demonstrates that the proposed procedure produces medium-range forecasts of mortality that meet the standard tests of accuracy in forecast evaluation and that are sensible when evaluated against the comparable forecasts produced by the Social Security Administration.

The growing awareness of the aging of the U.S. population and its consequences has generated increased interest in forecasts of the elderly population and of the rates of mortality at all ages, which are needed as inputs for such forecasts. This shift in interest toward the future elderly population has elevated the importance of forecasts of mortality, because all of those who will be 65 years of age or older between now and 2050 have already been born. Clearly the size and composition of the elderly population depends on the age structure of mortality.

Despite the importance of forecasts in demography, systematic development of forecasting methodology has not been a major preoccupation of demographers, and forecasters have generally relied on simple extrapolative methods and demographic accounting procedures.

> Ask anyone outside the profession what he thinks demographers do and he will give a much bigger place to forecasting than the editors of our journals give it space on their pages. Our most distinguished demographers do not put their major efforts into forecasting . . . we assign the highest professional standing to those who derive relations among variables in application to past data, in short who can most convincingly explain the past. (Keyfitz, 1985:60)

The practical importance of mortality forecasts calls for basic research and innovation in the methods of demographic forecasting. This article outlines a forecasting strategy that involves the integration of two—and ultimately three—traditionally independent approaches to demographic forecasting: time series methods, demographic accounting, and explanatory models (Land, 1986; Long, 1984). It is likely that a wedding of these three traditions—combining methods and models of demography, statistics, economics, and sociology—would produce significant gains in forecasting accuracy.

The article sets out an approach to forecasting that combines parameterized model schedules and time series methods in generating forecasts of mortality. It demonstrates the feasibility and accuracy of the forecasting methodology and indicates how one can extend this approach to forecasts of mortality by cause. It describes mathematically the age- and sex-specific structure of mortality and the evolution of this structure over time, through the estimation of parameterized mortality schedules. A major component of this descriptive analysis is the fitting of parameterized mortality schedules to historical data, providing a structure to the data, permitting a relatively concise representation of mortality by age and sex, and enabling one to compare the structure of mortality at different points in time.

The estimation of the parameterized schedules for each year yields a set of observations on each parameter over time. Time series methods can then be applied to the data on each parameter to develop forecasting models of sex-specific mortality schedules and hence of probabilities of mortality by age and sex. We conjecture that improvements in accuracy will arise from the formal acknowledgment in the forecasting model of the pervasive regularity in such patterns of mortality.

On Forecasting Mortality With Time Series Models

Current approaches to mortality forecasting are based on the extrapolation of mortality rates by age, sex, race, and in more recent forecasts, cause of death. The Social Security Administration (1987), for example, uses average annual rates of change in death rates by cause, age, and sex (calculated over the period of 1968–1983) as the basis for extrapolation of mortality rates through 2010. Rates of decline in the cause-specific death rates are assumed to converge gradually to set subjectively the ultimate rates of change for the year 2011 and beyond.

An examination of the record of mortality forecasting indicates that projections have mirrored the pattern of mortality prior to the base period of the forecast and have consequently understated the decline in mortality that actually occurred in the forecast period (Olshansky, 1987). Extrapolation all too often imposes a *fixed* rate of change based on the history of the time series, so this method will clearly err in forecasting phenomena experiencing turning points or changes in rates of decline. Modern time series methodologies, on the other hand, allow for more flexible modeling of trend and consequently offer greater possibilities for accommodating changes in long-run growth rates. Time series methods also offer the possibility of formally incorporating behavioral information in forecasts, such as responses to projected or hypothesized changes in socioeconomic variables. They also permit the construction of confidence intervals in the forecasts, so probability statements can be attached to forecast ranges.

Several studies have demonstrated the power of time series methods in the analysis and forecasting of aggregate demographic phenomena. Saboia (1974) modeled directly the population of Sweden, producing forecasts and confidence limits by using simple univariate methods. Saboia (1977), McDonald (1979, 1981), and Carter and Lee (1986) applied time series methods to fertility data. McDonald (1979) supported his empirical analysis with a demonstration that time series equations can be derived from an underlying structural model. He employed both univariate and transfer function models in the analysis of total live births in Australia. Time series models of aggregate variables, such as population, total fertility, or total mortality, ignore information contained in the age distribution of demographic phenomena. Consequently, the forecasts from these aggregate time series models have tended to have extremely large confidence intervals (Land, 1986).

Lee (1974) and Carter and Lee (1986) applied time series methods to forecast the levels of fertility schedules, assuming a fixed age pattern of fertility over time. Their approach takes advantage of the age distribution of fertility, but the *constancy* of this distribution assumed in their studies yields inaccurate forecasts if the age pattern of fertility does change. Thompson et al. (1987) permitted the shape of the fertility profile to change over time by summarizing this age pattern with the four-parameter gamma function. Estimates of these parameters from each of 64 years of data on fertility by age were modeled as time series using univariate and multivariate methods.

A general category of models representative of a wide class of time series is the autoregressive integrated moving average (ARIMA) model. These ARIMA models have a number of features that make them attractive candidates for time series modeling. First, they can approximate most time series observed in practice, incorporating either local stochastic trends or universal fixed trends as the data indicate. Second, ARIMA models have been demonstrated to be successful forecasting models, even in comparisons with large simultaneous-equation models (Ascher, 1978; McNown, 1986). Third, unlike simple extrapolative models, ARIMA models are actually reduced forms of some underlying structural model (McDonald, 1979; Zellner and Palm, 1974); they therefore implicitly embody a theoretical structure that may or may not be expressed explicitly. Finally, there are generalizations of the univariate ARIMA models to multivariate forms, which can also be related to an underlying structural model (Tiao and Box, 1981), providing thereby instruments for examining directly the relations among variables using the time series methodology (Bohara, Bradley, and McNown, 1987; Sims, 1980) and for generating forecasts of potentially greater accuracy than those of univariate models (Kling and Bessler, 1985).

Demographers and statisticians have applied time series methods to forecast aggregate fertility rates and total populations (Lee, 1974; McDonald, 1979, 1981; Saboia, 1977). These procedures could also be used to model *age-specific* rates. Individual projections of age-specific rates are likely, however, to produce highly irregular age profiles that are not characteristic of known schedules:

> If one were to extrapolate the age-specific death rates, age by age using virtually any formula, one would obtain highly irregular rates within a very few cycles of projection. On a straight-line projection age by age, many ages would soon show negative death rates. One plainly ought to summarize the rates into some minimum parameter set. (Keyfitz, 1982:743)

A preferred approach is to model and forecast *the entire age pattern* of each schedule. Such patterns may be described by means of mathematical representations called *parameterized model schedules*, which are specified with the aid of a relatively small number of parameters. Forecasts of future demographic regimes may then be obtained by projecting the temporal evolution of each of these parameters (Rogers, 1986; Thompson et al., 1987).

Parameterized Model Mortality Schedules: The Heligman–Pollard Model

Demographers have often turned to model schedules to describe patterns of age-specific rates in terms of a relatively small number of parameters. A large number of functions have been developed and fitted to mortality, fertility, and migration data, for example. The results have been extensively applied to data interpolation, smoothing, and inference as well as to comparative analysis (e.g., see Brass, 1971; Coale and Demeny, 1966; Coale and Trussell, 1974; Rogers and Castro, 1981).

The age profile of mortality can be conveniently decomposed into three components describing mortality probabilities in childhood, middle life, and older ages:

$$f(x) = f_1(x) + f_2(x) + f_3(x),$$

with $f(x)$ denoting the mortality rate, probability, or intensity at age x.

Following the pioneering work of Thiele (1872), the first and third components are often represented as negative and positive exponentials, respectively:

$$f_1(x) = a_1 \exp(-b_1 x),$$
$$f_3(x) = a_3 \exp(b_3 x).$$

Recent work (Rogers and Planck, 1983; Rogers and Watkins, 1986) exploring the fit of a multiexponential model similar to Thiele's, with the middle-life mortality component specified by the double exponential distribution of Coale and McNeil (1972),

$$f_2(x) = a_2 \exp\{-\alpha_2(x - \mu_2) - \exp[\gamma_2(x - \mu_2)]\},$$

showed promising results. Alternatively, Heligman and Pollard (1980) demonstrated that a somewhat different set of component curves, representing the complete profile of mortality probabilities as

$$f(x) = A^{(x+B)^C} + D \exp[-E(\ln x - \ln F)^2] + GH^x/(1 + GH^x),$$

also provides good fits to observed data.

In each of these representations the parameters characterizing the mortality profiles have convenient demographic interpretations. Nonlinear curve-fitting techniques are applied to mortality rates at each age, or a coarser age disaggregation, to obtain estimates of the parameters. Although both the multiexponential and Heligman–Pollard models are complex and involve a large number of parameters, their flexibility is necessary to capture the full curvature of the mortality profile and the changes in this profile over time. Although other model schedules may nominally seem to employ fewer parameters, they in fact do so only by ignoring some feature of the age curve of mortality (e.g., Siler, 1983) or by adopting a "standard" that is then shaped to fit a particular application with the aid of a few parameters (e.g., Brass, 1971). For example, the nearly 100 probabilities of Brass's standard can be represented by the eight-parameter Heligman–Pollard model with the following values:

$$A = 0.06008, \quad B = 0.31087, \quad C = 0.34431, \quad D = 0.00698,$$
$$E = 1.98569, \quad F = 26.71071, \quad G = 0.00022, \quad H = 1.08800.$$

Figure 1 illustrates the Brass standard and contrasts its curve with those of U.S. male mortality schedules for 1900 and 1980. (In this and subsequent figures depicting the age pattern of mortality, the probability of dying is presented in logarithmic form to bring out the curvature of this profile, especially at the young adult ages. The estimated parameters for U.S. male mortality for these and the intermediate years, extending the time series to 1985, are available from the authors. The data were kindly provided by Alice Wade of the Social Security Administration and were used by them in the development of their mortality projections; see Social Security Administration, 1987.) The results describe in parametric form the history of U.S. male mortality during this century. Figure 2 gives a graphical display of this history, contrasting the evolution of male mortality with the corresponding pattern for females.

The U.S. crude death rate has declined dramatically during this century, exhibiting two periods of significant decline separated by a 14-year period (between 1954 and 1968) of very modest reductions in mortality levels (Crimmins, 1981). The early period of decline was largely a consequence of a reduction in deaths from infectious and parasitic diseases and from influenza and pneumonia. The recent decline, on the other hand, is mostly a result of reductions in deaths from cardiovascular diseases and from other causes at the older ages. This history of mortality levels can be identified in the temporal patterns of the eight Heligman–Pollard parameters illustrated in Figure 2.

The first three parameters show similar patterns for the history of male and female early-life mortality. Parameter A, reflecting mortality at age one, exhibited a dramatic decline for both sexes through 1945, followed by more moderate reductions. The *B* parameter for each sex also declined steadily until 1945, rising slightly thereafter. This parameter measures the difference between age 1 and age 0 mortality probabilities, with the increase in *B* values indicating some convergence between these two values. Parameter C, capturing the decline in mortality during childhood, fell throughout the period and at a slower rate since 1945.

The next three parameters create the pattern of young-adult mortality, with its char-

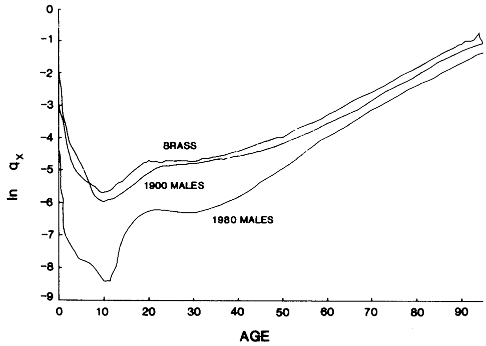

Figure 1. Probabilities of Dying at Each Age: U.S. Males, 1900 and 1980, and the Brass Standard. *Source:* Social Security Administration (1987) and Zaba (1979).

acteristic hump associated with deaths due to violence and maternal causes. The intensity of young-adult mortality is governed by the *D* parameters, which moved together for both sexes until the early 1940s. With the exception of the two spikes at 1918 (the influenza epidemic) and the early 1940s for males (World War II), both male and female *D* parameters declined persistently until 1950. This was followed by a slight divergence, with male mortality remaining high and even rising during the 1970s. Female young-adult mortality declined further until the mid-1950s, but it also showed a slight increase during the 1970s. The *E* parameter varies inversely with the spread of the young-adult mortality hump. The sharp increases in its values in 1940 for males and 1965 for females means the hump became more tightly concentrated around its modal age (approximated by *F*). The differences between the *F* parameter patterns for the sexes suggest that changing causes have played an important role in the history of young-adult mortality. For males the modal age of the young-adult mortality hump dropped gradually between 1905 and 1945, leveled off, and then began rising slightly during the 1970s. For females the modal age fluctuated about a relatively high age (between 27 and 32) until 1945, then dropped sharply to 21 years of age. This change most likely reflects a shift in importance away from maternal deaths to violent causes of mortality in the immediate post–World War II period.

Finally, late-life mortality, described by the Gompertz curve, is represented by the *G* (the intercept of the curve at age 0) and *H* (slope) parameters. Both of these parameters have shown wider fluctuations for males than for females, with *H* movements generally countering those of *G*. The combined impact of these two counteracting changes is difficult to gauge without numerical calculation, but old-age mortality in 1985 fell to considerably lower levels over mortality at the beginning of this century.

Univariate ARIMA Models of the Heligman–Pollard Parameters

Plots of the model schedule parameters for the entire 1900-1985 period show a sharp change in pattern for every parameter during the early 1940s (see Fig. 2). Although this shift raises some interesting questions concerning the behavior of mortality, it also poses major problems for modeling and forecasting the parameters. We experimented with time series models of the parameters, both including and excluding the observations before 1941, and discovered that models based on the more complete sample had inferior diagnostic statistics and out-of-sample forecasts. Using only post-1940 data reduces the number of observations below desired limits for time series modeling, because statistics used in model identification and model coefficients tend to be estimated with large standard errors. Despite

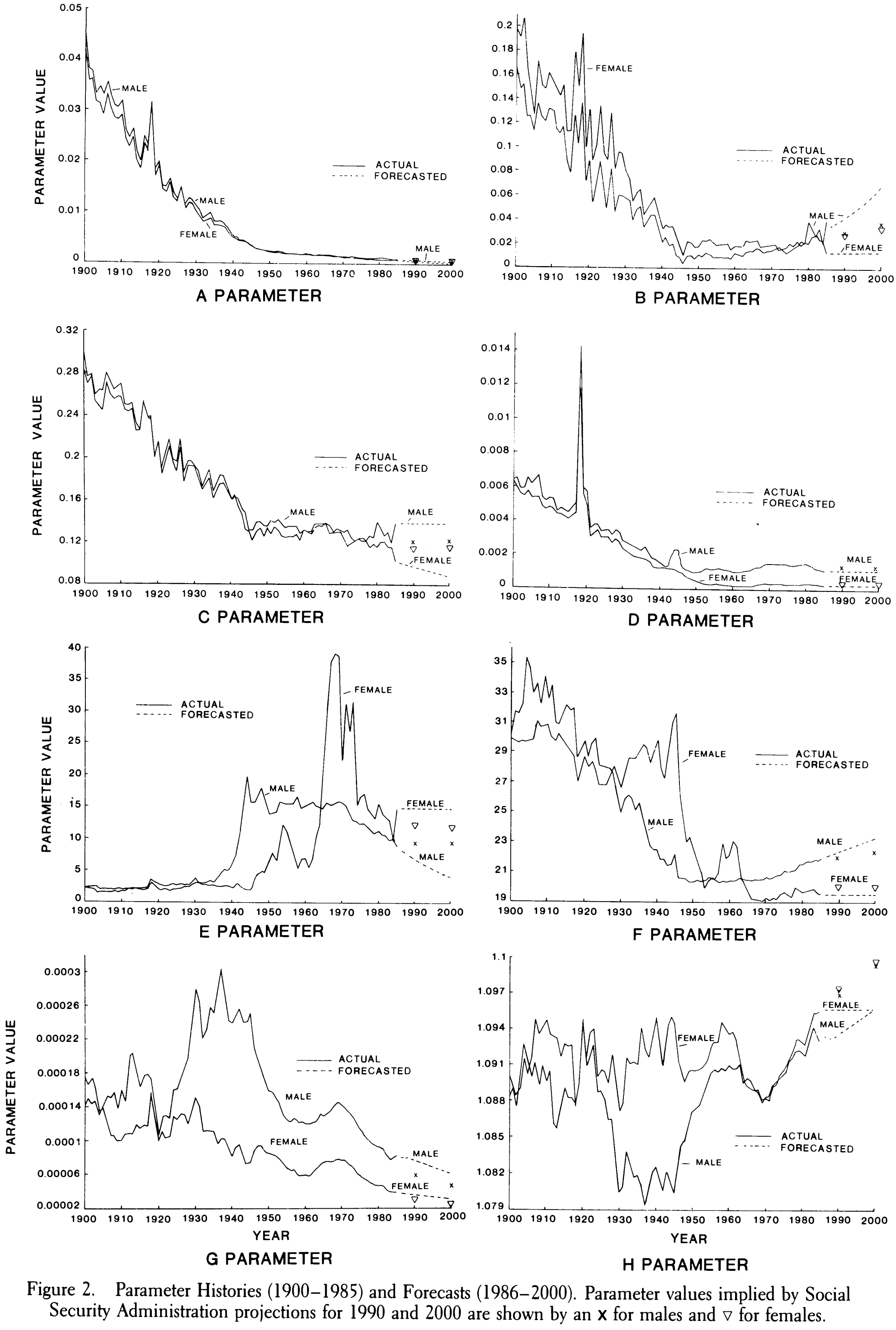

Figure 2. Parameter Histories (1900–1985) and Forecasts (1986–2000). Parameter values implied by Social Security Administration projections for 1990 and 2000 are shown by an **x** for males and ▿ for females.

these problems, we used the shorter sample in the modeling and forecasting analysis presented here, allowing forecast performance to override considerations of statistical significance in the decision to retain parameters.

Univariate time series models for each of the parameters were identified and estimated according to the methods of Box and Jenkins (1976), using the 35 observations from 1941–1975. The final models are reported in Table 1. Of the eight parameters of the male schedule, only parameter C was found to be most appropriately modeled as a random walk without drift, so the forecast for C is that of no change. For the female schedule the three parameters of the accident hump were identified as zero-drift random walk processes and were also forecasted to exhibit no change from the base-year values.

All parameters required a logarithmic transformation and a differencing of at least order 1 to achieve stationarity. The logarithmically transformed series tend to show less heteroscedasticity, and this transformation also sets a lower bound of 0 on parameter forecasts.

To the extent that forecast horizons required in demographic forecasting are rather long, proper modeling of trend is a crucial component of time series analysis. Forecasts from univariate time series models tend to converge fairly quickly to the long-run trends

Table 1. Estimated ARIMA Models, 1941–1975

Model	Q
Male Parameters	
$(1 - L)^2A(t) = e(t) - 1.11\,e(t-1) + 0.30\,e(t-2)$	6.42[a]
(0.17) (0.17)	(0.97)
$(1 - L)^2B(t) = e(t) - 1.22\,e(t-1) + 0.39\,e(t-2)$	3.52[a]
(0.16) (0.16)	(0.99)
$(1 - 0.27L + 0.11L^2 + 0.22L^3)(1 - L)D(t) = e(t)$	6.53[a]
(0.17) (0.18) (0.17)	(0.97)
$(1 - L)^2E(t) = e(t) - 0.42\,e(t-1) - 0.23\,e(t-2)$	13.69[a]
(0.17) (0.17)	(0.55)
$(1 - L)^2F(t) = e(t) - 1.26\,e(t-1) + 0.38\,e(t-2)$	4.63[a]
(0.16) (0.16)	(0.99)
$(1 - L)G(t) = -0.02 + e(t) + 0.20\,e(t-1) + 0.38\,e(t-3)$	8.58[b]
(0.01) (0.16) (0.16)	(0.80)
$(1 - L)H(t) = 0.0001 + e(t) + 0.12\,e(t-1) + 0.27\,e(t-3)$	8.74[b]
(0.0001) (0.17) (0.16)	(0.79)
Female Parameters	
$(1 - L)^2A(t) = e(t) - 0.82\,e(t-1)$	6.27[a]
(0.09)	(0.97)
$(1 - L)B(t) = e(t) - 0.15\,e(t-1)$	9.91[a]
(0.17)	(0.83)
$(1 - L)C(t) = -0.007 + e(t)$	9.88[a]
(0.006)	(0.83)
$(1 - L)G(t) = -0.006 + e(t) + 0.57\,e(t-1)$	9.86[c]
(0.017) (0.15)	(0.77)
$(1 - L)H(t) = e(t) + 0.38\,e(t-1)$	6.82[a]
(0.16)	(0.96)

Note: All models were identified and estimated using the 35 observations from 1941–1975. Each parameter has been transformed into logarithmic form. In each reported equation, L represents the lag operator, $LX(t) = X(t-1)$, and $e(t)$ is a white noise process. Numbers in parentheses below estimated coefficients are standard errors. Q is the Ljung–Box statistic, with the level of significance given in parentheses.

[a] df = 15.

[b] df = 13.

[c] df = 14.

incorporated in the models, which may be either deterministic and global or stochastic and local. In addition to the random walk processes, the models for parameter D from the male schedule and for parameters B and H from the female schedule contain no trend. Global trends are embodied in the remaining first-differenced models, although in several cases the intercepts, which measure these trends, are small and statistically insignificant. Local stochastic trends characterize the models of the second differenced series (A, B, E, and F from the male schedule and A from the female schedule). Forecasts of these parameters converge to a constant value determined by the last several observations in the sample.

All models reported in Table 1 show satisfactory diagnostics as summarized by the Q statistics of Ljung and Box (1978), which indicate no significant residual autocorrelation. Although the statistical insignificance of some coefficients in the reported models suggests possible overparameterization, we have retained these insignificant coefficients because they improve forecasting performance. We attribute their lack of statistical significance to the large standard errors produced by the shortened sample. The use of second differences for some series may appear surprising, raising the issue of possible overdifferencing. Although individual coefficients in some cases exceed unity, they will still satisfy the conditions for invertibility. A simple necessary condition is that the summation of the moving average coefficients be less than 1. According to the invertibility criterion, none of the series appears to be overdifferenced. Again the choice of second differencing is based on the superior forecasting performance of the resulting stochastic trend models relative to that of the corresponding first-differenced, deterministic trend models.

To generate reliable forecasts of age-specific mortality probabilities (the q_x), our methodology requires that the time series models be capable of providing satisfactory forecasts of the model schedule parameters. Nonupdated forecasts of both the parameters and the mortality probabilities are evaluated over the 10-year holdout period, 1976–1985. The accuracy of these out-of-sample forecasts is evaluated by using mean absolute percent error (MAPE) and Theil (1966) statistics. The Theil statistic reports the root mean squared error of the model forecast relative to that of the "naive" no-change forecast, with values less than unity indicating greater accuracy for the model forecast. Forecasted and actual mortality schedules are compared graphically in Figure 3, and summary statistics on forecast errors are displayed in Table 2.

The forecast errors observed for the mortality probabilities are related to underlying errors in the parameter forecasts. For the male mortality schedule the most serious absolute percentage errors are for infant mortality (MAPE = 20.9 percent). These large errors can be traced to the problems in forecasting the B parameter (MAPE = 41.9 percent), which experienced a change in trend at the end of the estimation period. After a decade of fairly flat values (1965–1975), this parameter began to rise sharply, reflecting some convergence in mortality probabilities at the ages of 0 and 1. The failure of the forecasts to capture this shift in trend highlights the progress against infant mortality that was achieved during the period of 1976–1985.

Relatively large forecast errors also appear for males at 20 years of age, and these are a reflection of the errors in forecasting the D parameter, which measures the intensity of young-adult mortality. From the late 1960s through 1980, mortality rates for young adults were rising against a tide of falling rates at all other ages. Easterlin (1980) suggested that the observed increase in mortality rates at these ages can be attributed to the stress experienced by this unusually large cohort from competitive pressures for jobs and education. With the maturation of this cohort, mortality probabilities for young adults turned down after 1980, a change that is not captured by the forecasts at these ages. Figure 3 shows vividly the failure of the forecasts to capture the downward shift in the accident hump between 1980 and 1985.

Despite these errors in male mortality forecasts, the overall forecast accuracy is high. The largest mean absolute error is 20.9 percent, and at only three other ages are the MAPEs

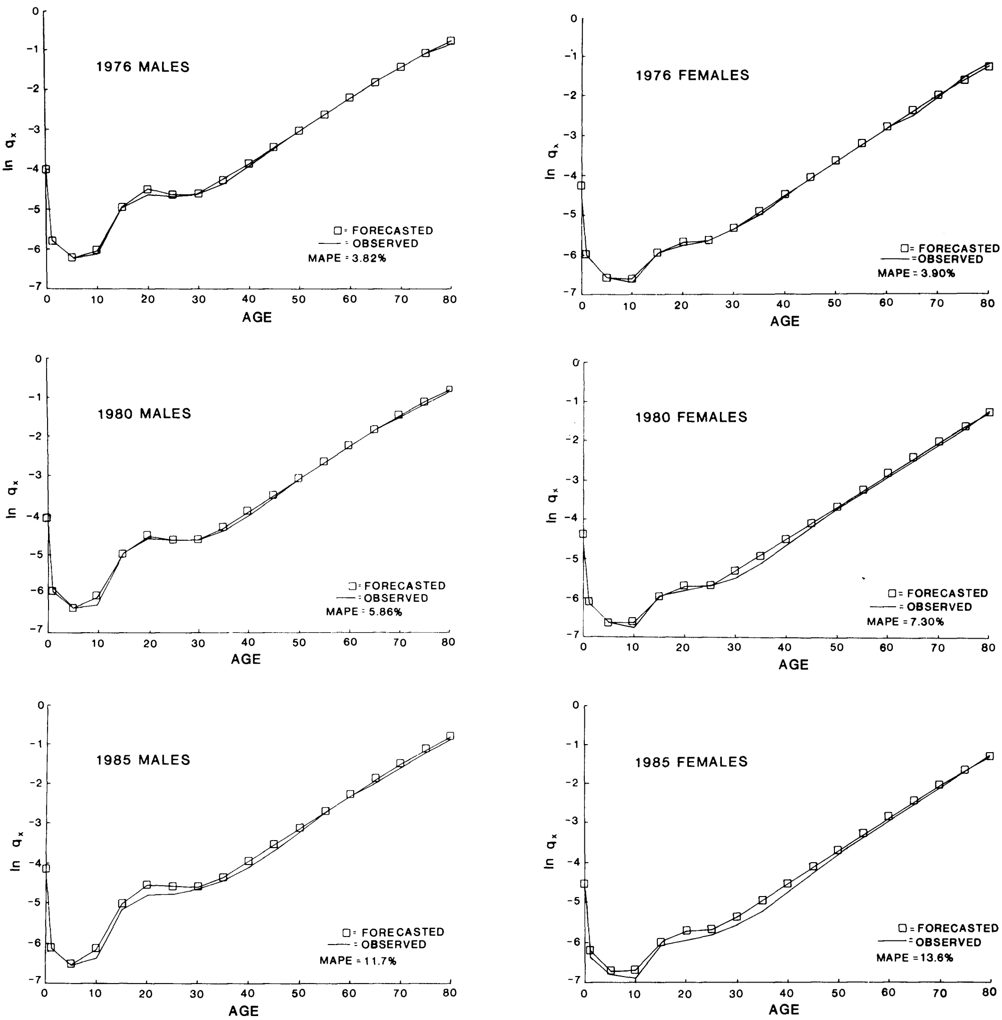

Figure 3. Forecasted Versus Observed Probabilities of Dying: U.S. Males and Females, 1976, 1980, 1985

greater than 10 percent. The model forecasts are superior to the naive no-change forecasts at all but two ages, and for one of these (80) the errors are small (MAPE = 5.8 percent). At many ages the improvement over the no-change forecast is substantial, with model root mean squared errors being less than one-third of those for the no-change forecast at the ages of 5 and 50–60.

The fit between the actual and forecasted schedules for each year is measured by the mean absolute percent error across the 18 age categories. These statistics increase from 3.8 percent to 11.7 percent as the horizon lengthens from 1 to 10 years, implying a reasonably strong association between actual and forecasted rate schedules at each year.

A similar degree of accuracy holds for the female mortality forecasts. Female infant mortality is forecasted with greater accuracy (MAPE = 9.0 percent) than for males. Females did not show the same degree of convergence of age-0 and age-1 mortality as did males over the forecast period. The largest MAPEs are at the peak age of the accident hump (20) and at 30–40 years of age, at which point adult mortality bottoms out before rising to its higher

Table 2. Forecast Error Statistics

	Male		Female	
Parameter	Mean absolute % error	Theil	Mean absolute % error	Theil
	Parameters			
$A(t)$	11.6	0.87	4.8	0.26
$B(t)$	41.9	1.09	18.0	1.03
$C(t)$			2.9	0.54
$D(t)$	12.9	1.05		
$E(t)$	5.6	0.37		
$F(t)$	0.7	0.31		
$G(t)$	10.9	0.41	26.7	0.86
$H(t)$	0.8	0.46	0.3	1.01
	Mortality Probabilities			
Age				
0	20.9	0.67	9.0	0.31
1	5.7	0.36	6.7	0.34
5	4.0	0.22	6.9	0.40
10	19.0	0.87	14.4	0.90
15	8.1	0.59	6.1	0.64
20	16.7	1.14	12.4	0.97
25	8.8	1.00	5.5	0.47
30	4.0	0.56	14.2	0.79
35	10.0	0.66	18.4	0.89
40	11.3	0.64	12.6	0.62
45	7.1	0.44	7.7	0.51
50	3.3	0.28	5.0	0.54
55	2.0	0.18	6.3	0.68
60	3.5	0.28	6.1	0.95
65	4.3	0.47	11.5	4.36
70	4.8	0.51	8.3	0.82
75	5.1	0.67	5.2	0.43
80	5.8	1.52	5.6	0.90

Note: All statistics are computed from nonupdated, out-of-sample forecasts for 1976–1985. Missing numbers for parameter forecasts correspond to random walk models.

levels due to senescence. As in the case for males, the time series models do not anticipate the decline in mortality rates in the accident hump for females during the last 5 years of the forecast. In fact, the 1985 accident hump is sufficiently diminished to eliminate the trough in mortality probabilities around the ages of 30–40 observed in prior years. The forecasted mortality profile also shows a leveling of this trough, but not to the degree exhibited by the actual schedule.

A final anomaly in the female forecasts is the large Theil statistic at the age of 65. The plot of actual mortality probabilities for 1976 shows an unusual deflection in the actual schedule below the almost linear relation between the logarithm of q_x and age, which is characteristic of senescent mortality. Since the Heligman–Pollard schedule imposes a smooth Gompertz curve on senescent mortality, the forecasts of mortality from the schedule cannot capture this irregularity. The errors are not large by the MAPE criterion, but the coincidental accuracy of the no-change forecast leads to the extremely large Theil statistic.

The fits between forecasted and actual mortality schedules are close across all ages, again showing a looser correspondence as the forecast horizon widens. The MAPE statistic calculated for the 18 observed age intervals grows from 3.9 percent in 1976 to 13.6 percent in 1985. Considering both the overall fits at each year and the summary statistics over the 10 years, the forecasts of female mortality appear acceptable in terms of accuracy.

Forecasts of U.S. Mortality

The analysis of holdout sample forecasts demonstrates the viability of the forecasting methodology for horizons of moderate length. In this section, the ARIMA models are reestimated through 1985 to produce true forecasts of mortality through the year 2000. Although it is not possible to evaluate the accuracy of these forecasts until the end of this century, some perspective can be gained from a comparison of the model's forecasts with those of the Social Security Administration for the years 1990 and 2000 (Social Security Administration, 1987). To enhance this comparison the Heligman–Pollard model was fitted to the mortality probabilities implied by Social Security Administration mortality rate projections. The comparison is illustrated in Figure 4. Divergences in forecasted mortality rates can therefore be traced to different implicit assumptions about the future of the Heligman–Pollard parameters.

Reestimation of the ARIMA models with the 10 additional observations yielded only small differences from the results reported in Table 1. The additional data resulted in estimated parameters with smaller standard errors, but otherwise the estimates and the model diagnostics remain qualitatively unchanged.

In Figure 2 the historical time series for the male and female parameters are extended from 1986 through 2000 with the parameter forecasts from the updated models. The implied

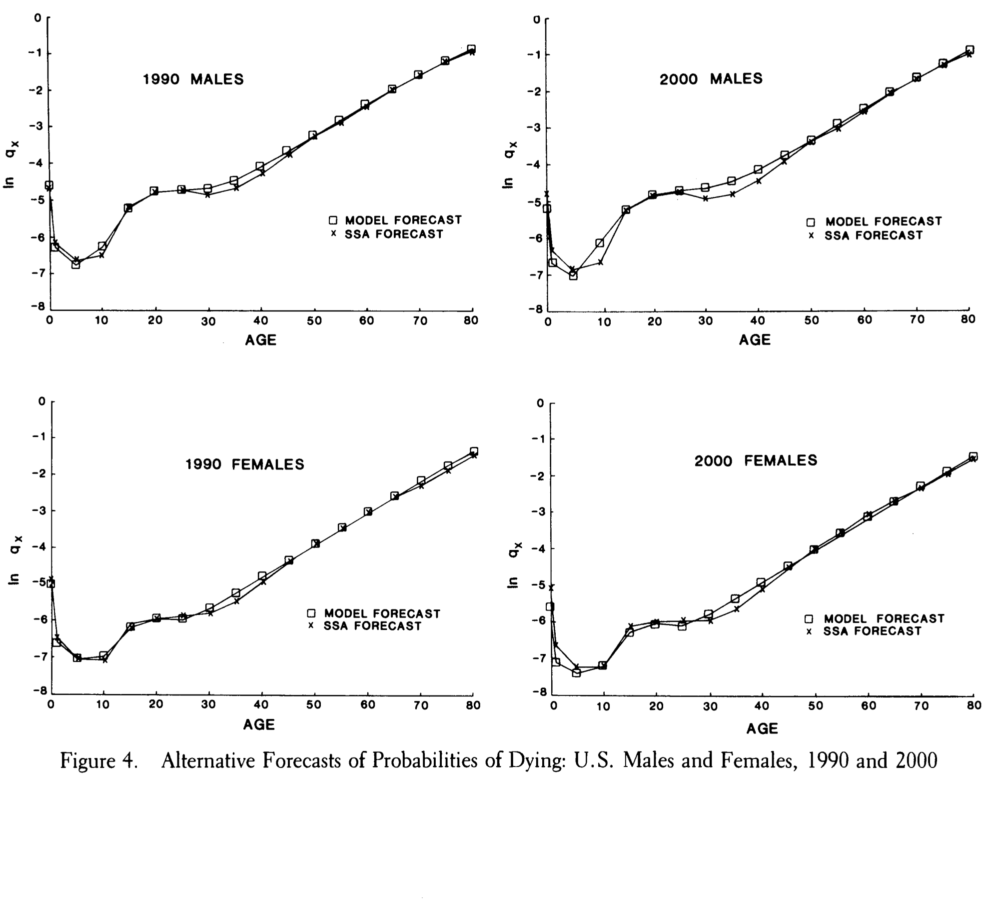

Figure 4. Alternative Forecasts of Probabilities of Dying: U.S. Males and Females, 1990 and 2000

Table 3. Life Expectancies in 2000

Age	Males: Model	Males: Social Security Administration	Females: Model	Females: Social Security Administration
0	72.79	73.58	80.77	81.26
20	53.77	54.77	61.34	62.15
40	35.45	36.23	41.95	42.75
60	18.66	19.08	23.91	24.58
70	11.97	12.24	16.13	16.77

parameter values for 1990 and 2000 from the Social Security Administration projections are also shown in these plots. Figure 4 presents comparisons of the entire mortality profiles from the two sets of forecasts, and Table 3 shows the two sets of projected life expectancies at various ages for the year 2000.

For males, the model forecasts call for lower infant and childhood mortality probabilities than shown by the Social Security Administration projections. This reflects the more rapid decline in the A parameter (measuring mortality at the age of 1) and a greater convergence of age-0 and age-1 mortality probabilities produced by the rise in B generated by the ARIMA model. Both of these parameter forecasts maintain trends established during the 1970s and early 1980s.

Both male mortality profiles show minimum mortality probabilities occurring at a younger age than has been exhibited historically, and mortality under the Social Security Administration projection rises to a well-defined accident peak at the age of 23. The ARIMA model predicts a tempering of the concentration in young adult mortality. The forecasts of D from the ARIMA model are flat throughout the projection period and at a value slightly lower than implied by the Social Security Administration projections. More dramatic is the decline in E from the model, which implies an increase in the dispersion around the midpoint of the accident hump. This midpoint, defined by F, rises slowly in both forecasts, but reaches a higher age (23.3 in 2000) in the model forecast. The general tendency for young male adult mortality, according to the model, is an attenuation of the accident hump as the large post–World War II birth cohort passes through the violence-prone years. (A reviewer suggested the possible influences of changes in fuel costs and in the legal age of drinking.)

The more optimistic picture for youth and young-adult mortality described by the model forecasts is reversed at the higher ages. The model forecasts for senescent mortality are largely driven by the G parameter, which is projected to decline by about 25 percent over the projection period. This rate of decline is slower than that implied by the Social Security Administration projections (43 percent), which envision a pattern that is comparable to the rate of decline over the previous 15 years.

The lower senescent mortality probabilities under the Social Security Administration forecasts also account for the correspondingly greater life expectancies projected at every age. In the year 2000 the largest gap in projected life expectancies is 1 year at the age of 20. Model projections imply the following gains over 1985 life expectancies: 1.64 years at 0, 1.19 years at 40, and 0.87 year at 60. These may be compared with gains of 4.16, 2.75, and 1.86 years at these ages over the 1970–1985 period.

The differences between the two projected mortality schedules for females are similar to those for males. Infant and youth mortality probabilities are lower under the model projections, again reflecting a rapid decline in the A parameter from the ARIMA model forecasts. Young-adult mortality is also slightly lower in the model projection, and both forecasts still show a distinct accident peak at 20 years of age. The shape of the accident peak under the model forecast is static in this case, since the three relevant parameters were modeled as random walk processes. Senescent mortality is projected to fall more rapidly according to the Social Security Administration projection than is indicated by the ARIMA model. Again, the model forecasts reflect a more moderate decline in the G parameter (20 percent vs. 42 percent). The lower senescent mortality of the Social Security Administration projections translates into greater life expectancies at all ages, although the gap is always less than a year. The model forecasts of life expectancy gains for females are also more conservative than the record of the previous 15 years.

These comparisons show the sources of divergences in the two projections and can provide the basis for informed debates and revisions. Comparison with previous trends suggests that the model forecasts may be too conservative in predicting senescent mortality and that the Social Security Administration may be underpredicting the gains in infant and youth mortality. Such revisions or similar sensitivity analyses could be conducted by introducing changes in sets of future parameter values.

Future Research: Models for Mortality by Cause and Multivariate Models

A natural extension of this methodology is to mortality by cause. Demographers have long been interested in causes of death, how the cause structure of mortality has changed over time, and how it differs across populations. Parameterized model schedules offer a powerful instrument for representing mortality by cause and for allowing comparisons across time or subgroups. Preston (1976), for example, presented graphically the age profiles of mortality for several important causes. With these charts, one can observe the incidence of a particular cause of death at each age and how this pattern varies with overall levels of life expectancy. Compact representation of mortality by cause at each age across a number of years is facilitated by the estimation of cause-specific mortality functions. Since the parameters of these functions capture the complete age and cause structure of mortality, they can be used as variables in behavioral analysis and time series forecasting.

Several issues must be addressed in the extension of the methodology to mortality by cause. It is clear from the profiles displayed by Preston (1976) that the age structure of mortality varies by cause. Variations of the Heligman–Pollard model or the multiexponential model (Rogers, 1986; Rogers and Watkins, 1986), found suitable for modeling a wide range of demographic phenomena, may have to be modified to fit the data on cause-specific mortality. There is no research of which we are aware that provides guidance on the choice of model schedules for cause-specific mortality, but preliminary tests of the Heligman–Pollard model and the flexibility and wide range of applications of the multiexponential model give us confidence that both can also be applied here.

A *multiple* time series model would permit an examination of linkages among different causes of death and the introduction of socioeconomic variables as predictors of the cause-specific model parameters. Boulier and Paqueo (1988) rightly observed that despite "an extremely rich body of empirical work describing the influences of biological, demographic, and environmental factors on mortality, there is yet no formal theoretical model of mortality that can be used to organize and synthesize the vast quantitative literature" (p. 249). Nevertheless plausible linkages between aggregate socioeconomic variables and mortality have been identified. Land and McMillan (1980), for example, established the existence of strong linkages between cause-specific mortality rates and variables measuring exposure to risk, general living standards, specific social and economic conditions, and public health practices. The variables they examined in their regression analysis would also be suitable candidates for a multiple time series analysis. Of course, for forecasting purposes the socioeconomic variables must also be predicted within the time series framework in the generation of mortality forecasts.

Alternatively, the multiple time series model can be used to make conditional predictions of mortality based on alternative projections of demographic and other socioeconomic determinants (Sims, 1982). For example, one could analyze the effect of an exogenous reduction in the parameters of a particular cause of death, such as cardiovascular disease, on the probabilities of death from other causes. The validity of these conditional projections, as in any other model-based projection, would depend on the constancy of the relations among the variables examined—in this case among the parameters of the different causes of death.

Two immediate extensions of our work are suggested by the time series tradition in demographic forecasting. The first is the development of confidence intervals for the mortality probability forecasts. Standard errors of the *parameter* forecasts are provided directly in the ARIMA modeling methodology. However, since the mortality probabilities are nonlinear functions of the parameters, computation of their forecast standard errors requires either an asymptotic approximation or a bootstrap simulation (Efron and Gong, 1983; Peters and Freedman, 1985). We are currently investigating the application of the bootstrap procedure to our forecasting methodology.

The second direct extension of our work is the application of multiple time series methods to forecast the Heligman–Pollard parameters (Thompson et al., 1987). The univariate models yield independent projections of each parameter, ignoring potential gains in accuracy arising from taking into account the covariation among the eight parameters. Future work will apply the vector autoregression methodology (Sims, 1980) to the simultaneous modeling and forecasting of the eight parameters.

Summary

Long (1984) and Land (1986) discussed three modeling traditions in population forecasting: demographic accounting, time series analysis, and explanatory modeling. Demographic accounting provides a structure in the form of accounting identities that permits the generation of long-term projections for detailed population subgroups from assumptions concerning age-specific mortality, fertility, and migration rates. Because of the need for such long-range, detailed predictions, the Census Bureau has relied on the demographic accounting procedure, employing "informed judgment" and "historical continuity" in the assumed age-specific rates of demographic change. The use of an essentially nonstatistical method of projection, particularly the use of informed judgment in establishing future rates of births and deaths, has made difficult any inferences concerning the precision of the Census Bureau's projections. Stoto (1988) found that the range of population values encompassed in past Census Bureau projections does not correspond to meaningful confidence intervals. Land (1986) called for a "statistical perspective" in demographic forecasting, which, at a minimum, would yield confidence intervals about population forecasts.

We have demonstrated how the estimation of a model mortality schedule for each year over a period of time yields a set of observations on each parameter that can be modeled and forecasted with time series methods. Inserting the forecasted parameter values into the mathematical expression for the model schedule produces forecasts of mortality probabilities at each age and hence of the entire mortality profile. Our application of this procedure to data on male and female mortality probabilities for the first 86 years of this century yielded forecasts of age-specific mortality of reasonable accuracy, confirming the viability of this methodology.

Our modeling strategy contains two sources of forecasting error that must be weighed against the gains obtained from imposing a flexible structure on the mortality age profile. First, there are errors in approximating mortality by age with an algebraic function and in estimating the parameters of this function with a small number of observations on mortality by age. The approximation errors are quite small, with mean absolute percentage errors of 5.5 percent or less in each year. Second, the time pattern of each parameter of the model schedule is represented by a parsimonious ARIMA model, whose coefficients are also estimated with error. This appears to be a larger source of error than the use of the model schedule, as indicated by the large errors in forecasting certain parameters.

Our pilot study illustrates the use of the time series methodology in a real mortality forecasting situation and analyzes the accuracy of out-of-sample forecasts in a simulated forecasting environment. The procedure produces medium-range forecasts that are sensible when evaluated against those of the Social Security Administration and that meet standards of accuracy in the holdout sample evaluation. The introduction of cohort effects would no doubt further improve the proposed methodology.

To some extent, the strength of the methodology with the proposed extensions goes beyond forecast accuracy. Forecasts of the parameters produce a schedule that captures the fundamental characteristics of the age distribution of mortality. Smooth curves showing mortality probabilities at single-year age intervals can be inferred from the forecasted parameters. The parameterized schedules and their forecasts provide a convenient summary of entire mortality distributions, and their representation by parameters provides a useful instrument for comparing mortality at different points in time or across populations. Furthermore, the forecasts of individual mortality probabilities can be traced to assumptions or projections of individual parameters, providing a convenient basis for the comparison of mortality forecasts obtained from different sources.

Acknowledgments

A number of individuals contributed in various ways to the preparation of this paper. Foremost among them was Kathy Gard, who, as research assistant to the project, carried out a prodigious amount of data processing work with good cheer, for which we are most grateful. Thanks also go to Alice Wade of the Social Security Administration for supplying the data and to Larry Heligman of the United Nations Population Division for providing the software (included in MORTPAK) for the Heligman–Pollard model. Finally, we also wish to acknowledge with thanks the accurate typing of Carrie Andree and the useful comments of three anonymous reviewers. This research is being supported by National Science Foundation Grant SES-8705591 to Andrei Rogers, Principal Investigator.

References

Ascher, W. 1978. *Forecasting: An Appraisal for Policy Makers*. Baltimore, Md.: Johns Hopkins University Press.

Bohara, A. K., M. Bradley, and R. F. McNown. 1987. New evidence on targets for monetary policy. *Southern Economic Journal* 53:591–604.

Boulier, B. L., and V. B. Paqueo. 1988. On the theory and measurement of the determinants of mortality. *Demography* 25:249–263.

Box, G. E. P., and G. M. Jenkins. 1976. *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.

Brass, W. 1971. On the scale of mortality. Pp. 86–110 in W. Brass (ed.), *Biological Aspects of Demography*. London: Taylor and Francis.

Carter, L. R., and R. D. Lee. 1986. Joint forecasts of U.S. marital fertility, nuptiality, births and marriages using time series models. *Journal of the American Statistical Association* 81:902–911.

Coale, A. J., and P. Demeny. 1966. *Regional Model Life Tables and Stable Populations*. Princeton, N.J.: Princeton University Press.

Coale, A. J., and D. R. McNeil. 1972. The distribution by age of the frequency of first marriage in a female cohort. *Journal of the American Statistical Association* 67:743–749.

Coale, A. J., and T. J. Trussell. 1974. Model fertility schedules: Variations in the age structure of childbearing in human populations. *Population Index* 40:185–206.

Crimmins, E. M. 1981. The changing pattern of American mortality decline, 1940–77, and its implications for the future. *Population and Development Review* 7:229–254.

Easterlin, R. A. 1980. *Birth and Fortune*. New York: Basic Books.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross validation. *American Statistician* 37:36–48.

Heligman, L., and J. H. Pollard. 1980. The age pattern of mortality. *Journal of the Institute of Actuaries* 107: 49–80.

Keyfitz, N. 1982. Choice of function for mortality analysis: Effective forecasting depends on a minimum parameter representation. *Theoretical Population Biology* 21:329–352.

———. 1985. Demography in the twenty-first century: The uses of forecasting. Pp. 59–81 in *Proceedings of the IUSSP International Population Conference* (Vol. 1). Liege: International Union for the Scientific Study of Population.

Kling, J. L., and D. A. Bessler. 1985. A comparison of multivariate forecasting procedures for economic time series. *International Journal of Forecasting* 1:5–24.

Land, K. C. 1986. Methods for population forecasts: A review. *Journal of the American Statistical Association* 81:888–901.

Land, K. C., and M. M. McMillan. 1980. A macrodynamic analysis of changes in mortality indexes in the United States, 1940–1975: Some preliminary results. *Social Indicators Research* 7:1–46.

Lee, R. D. 1974. Forecasting births in post-transition populations: Stochastic renewal with serially correlated fertility. *Journal of the American Statistical Association* 69:607–617.

Ljung, G. M., and G. E. P. Box. 1978. On a measure of lack of fit in time series models. *Biometrika* 65:297–303.

Long, J. F. 1984. U.S. national population projection methods: A view from four forecasting traditions. *Insurance: Mathematics and Economics* 3:231–239.

McDonald, J. 1979. A time series approach to forecasting Australian total live births. *Demography* 16:575–601.

———. 1981. Modeling demographic relationships: An analysis of forecast functions for Australian births. *Journal of the American Statistical Association* 76:782–792.

McNown, R. F. 1986. On the uses of econometric models: A guide for policy makers. *Policy Sciences* 19:359–380.

Olshansky, S. J. 1987. *On Forecasting Mortality.* Unpublished manuscript (mimeo), National Opinion Research Center, University of Chicago.

Peters, S. C., and D. A. Freedman. 1985. Using the bootstrap to evaluate forecasting equations. *Journal of Forecasting* 4:251–262.

Preston, S. H. 1976. *Mortality Patterns in National Populations.* New York: Academic Press.

Rogers, A. 1986. Parameterized multistate population dynamics and projections. *Journal of the American Statistical Association* 81:48–61.

Rogers, A., and L. Castro. 1981. *Model Migration Schedules,* Research Report 81-30. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., and F. Planck. 1983. *Model: A General Program for Estimating Parameterized Model Schedules of Fertility, Mortality, Migration, and Marital and Labor Force Status Transitions,* Working Paper 83-102. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., and J. Watkins. 1986. *Parameterized Projections of Regional Populations in the United States,* Working Paper 86-5. Boulder: University of Colorado, Institute of Behavioral Science, Population Program.

Saboia, J. L. M. 1974. Modeling and forecasting populations by time series—The Swedish case. *Demography* 11:483–492.

———. 1977. Autoregressive integrated moving average (ARIMA) models for birth forecasting. *Journal of the American Statistical Association* 72:264–270.

Siler, W. 1983. Parameters of mortality in human populations with widely varying life spans. *Statistics in Medicine* 2:373–380.

Sims, C. A. 1980. Macroeconomics and reality. *Econometrica* 48:1–48.

———. 1982. Policy analysis with econometric models. *Brookings Papers on Economic Activity,* 107–152.

Social Security Administration. 1987. *Social Security Area Population Projections, 1987,* Actuarial Study No. 99. Washington, D.C.: U.S. Office of the Actuary.

Stoto, M. A. 1988. Dealing with uncertainty: Statistics for an aging population. *American Statistician* 42:103–110.

Theil, H. 1966. *Applied Economic Forecasting.* Amsterdam: North Holland.

Thiele, T. N. 1872. On a mathematical formula to express the rate of mortality throughout the whole of life. *Journal of the Institute of Actuaries* 16:313–329.

Thompson, P. A., W. Bell, J. Long, and R. B. Miller. 1987. *Multivariate Time Series Projections of Parameterized Age-Specific Fertility Rates,* Working Paper 87-3. Columbus: Ohio State University, College of Business.

Tiao, G. C., and G. E. P. Box. 1981. Modeling multiple time series with applications. *Journal of the American Statistical Association* 76:802–816.

Zaba, B. 1979. The four parameter logit life table system. *Population Studies* 33:79–100.

Zellner, A., and F. Palm. 1974. Time series analysis and simultaneous equations models. *Journal of Econometrics* 2:17–54.