Modeling and Forecasting U.S. Mortality: Rejoinder
Author(s): Ronald D. Lee and Lawrence R. Carter
Reviewed work(s):
Source: *Journal of the American Statistical Association*, Vol. 87, No. 419 (Sep., 1992), pp. 674-675
Published by: American Statistical Association
Stable URL: http://www.jstor.org/stable/2290204
Accessed: 16/11/2012 17:03

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Statistical Association is collaborating with JSTOR to digitize, preserve and extend access to *Journal of the American Statistical Association.*

http://www.jstor.org

This content downloaded by the authorized user from 192.168.82.209 on Fri, 16 Nov 2012 17:03:17 PM
All use subject to JSTOR Terms and Conditions

reserve this term for interval estimators of parameters. Interval estimators of random variables (which is what we are discussing here) might be better called "prediction intervals" or "forecast intervals" as, say, in Kendall and Buckland (1971).

ADDITIONAL REFERENCES

Alho, J. M. (1992), "Estimating the Strength of Expert Judgment: The Case of U.S. Mortality Rates," *Journal of Forecasting*, 11, 157-167.
Kendall, M. G., and Buckland, W. R. (1976), *A Dictionary of Statistical Terms*, London: Longman Group.

Rejoinder

RONALD D. LEE and LAWRENCE R. CARTER

We are grateful for the thoughtful comments on our article provided by two scholars who have themselves made pioneering contributions to the problems we discuss. Because we disagree with some of McNown's points, we will devote the most space to them.

McNown asserts that our method is "equivalent" to "directly projecting each age-specific mortality rate at its own historical rate of exponential decline" . . . "despite their statements to the contrary." This is an important point, because our method is rather complicated, whereas straight extrapolation is very simple. We have addressed this point in our article and also will respond at some length here: First, in our model each death rate declines at its own exponential rate only when **k** declines linearly. This is not an assumption of the model, and in other applications **k** might follow some other sort of process. Second, if each age-specific rate is forecast separately, then deriving confidence intervals for forecasts of period life table functions such as life expectancy, that depend on many death rates, requires taking into account the covariance matrix of errors.

Third, in response to this comment, we have tried two versions of directly extrapolating individual age specific rates. We forecasted to 2065 using the endpoint-to-endpoint exponential rates of decline from 1933–1987 to extrapolate to 2065. The resulting rate forecasts were lower than ours for ages below 10, higher from 10–45, lower from 45–75, and higher thereafter. The percentage differences were often appreciable, ranging from plus 65 to minus 11. We also forecasted using regressions of the logs of the death rates on a constant and time. Such forecasts did indeed often come close to ours for 2065, although individual age group differences are as large as 25%. For example, for age group 30–34 our own forecast is .000180. An endpoint-to-endpoint extrapolation yields .00298, and the regression-based extrapolation yields .000225. Furthermore, comparing the regression estimates of rates of decline to our $\mathbf{b}_x$s shows that they differ by up to 8% after equivalent normalization. Comparison of the 95% probability interval for this age group in 2065 shows wider discrepancies: Our range is .00009 to .00036; the regression interval is .00016–.00031, or about half as wide (these figures do not reflect parameter uncertainty). In sum, the methods we tried for directly forecasting the individual rates led to point forecasts which, although somewhat similar to ours, differed in both level and age pattern, contrary to McNown's assertion that they would be "identical." They also led to very different confidence intervals.

Fourth, our method incorporates procedures for indirect estimation of mortality in periods when age-specific mortality data are unavailable. In our article we extended the time series back from 1933 to 1900 in this way, and forward from 1987 to 1989. This aspect of the method is helpful in applications for populations of developed countries and will be absolutely essential in many applications for populations of less developed countries such as China, where estimates of age-specific mortality may be available for only one or two years.

McNown also suggests that we actually have a 24-parameter model of mortality change, consisting of the 23$\mathbf{b}_x$s plus **k**. Perhaps our difference on this point is just semantic. The $\mathbf{b}_x$s are fixed by age and so do not change over time. Only **k** changes over time, and so only a single parameter, **k**, needs to be forecast. In the common language of demography we have a one-parameter family of life tables in exactly the same sense that the Coale–Demeny model life tables for a given region and sex are one-parameter life tables, even though the construction of the Coale–Demeny life tables involved two regression coefficients at each age, corresponding to our $\mathbf{a}_x$ and $\mathbf{b}_x$ coefficients. In our case a value of **k** allows us to identify uniquely a corresponding life table from the family. Of course one could vary the underlying coefficients ($\mathbf{a}_x$ and $\mathbf{b}_x$), but then one would be providing a basis for a new family of life tables. To forecast from a two-parameter family of life tables, such as those of Ledermann or Brass, one would have to forecast two parameters. The model used by McNown and Rogers can describe a single life table very efficiently, using only nine parameters versus the 47 required for our model. But in their 1990 forecasting application, even though six of these parameters are held constant over the forecast range, it still is necessary to forecast three of them. With six parameters held constant, this could be identified as a three-parameter life table system.

The two commentators suggest that our out-of-sample forecasts of the age pattern of mortality may not be "rea-

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

This content downloaded by the authorized user from 192.168.82.209 on Fri, 16 Nov 2012 17:03:17 PM
All use subject to JSTOR Terms and Conditions

sonable" (Alho) and "may depart from plausible, historically observed patterns" (McNown). This certainly is a real concern. But because we are forecasting outside the historically observed range, historically observed patterns can be compared to our forecasts only in a very general sense. We agree that the pattern of forecasted rates for the teen years and early twenties looks somewhat implausible, as we have noted in the article. However, we see nothing implausible in our forecasted pace of increase in mortality from ages 50–70. It might be useful to compare our forecasted age patterns to those in some contemporary populations that have both high quality data and substantially superior life expectancies, such as Sweden or some regions of Japan. Data from such populations could be included in the matrix of death rates from which $\mathbf{a}_x$ and $\mathbf{b}_x$ are computed, or the information could be incorporated in some other way. In any event, we fail to see why using a highly nonlinear nine-parameter system to fit within-sample age patterns of mortality, as McNown and Rogers (1990) did, would lead to age patterns that "adhere to standard age profiles" when some subset of these parameters has been forecast over a 75-year horizon.

Alho makes the interesting suggestion that extrapolative forecasts such as ours be combined with forecasts based on opinions of medical experts, such as those of the Social Security Administration's Office of the Actuary. Such combinations could be useful for anticipating progress against specific causes of death, but we doubt they would help produce forecasts with plausible age patterns. Alho also suggests that the forecast intervals would be more credible had we not used a dummy for the influenza epidemic. We believe that there are good arguments on both sides of this issue. Perhaps we should add the disclaimer that our intervals do not reflect the possibilities of nuclear war or global environmental catastrophe, for indeed they do not.

This content downloaded by the authorized user from 192.168.82.209 on Fri, 16 Nov 2012 17:03:17 PM
All use subject to JSTOR Terms and Conditions