Session 6A: projection and Statistical Modeling of Mortality at Late Age

This is a transcript of the discussion by W. Ward Kingkade, PhD

Presented at the Living to 100 and Beyond Symposium
Orlando, Fla.
January 7-9, 2008

Papers Presented:

On Simulation–Based Approaches to Risk Measurement in Mortality with Specific Reference to Binomial Lee–Carter Modelling
Steve Haberman and Arthur Renshaw

A Study of the Lee–Carter Model with Age–Shifts
Jack C. Yue, Sharon S. Yang and Hong–Chih Huang

Testing Deterministic versus Stochastic Trends in the Lee–Carter Mortality Indexes and Its Implications for Projecting Mortality Improvements at Advanced Ages
Wai–Sum Chan, Siu–Hang Li and Siu–Hung Cheung

The three papers in this session are all discussing statistical methodologies on projecting mortality, and each of them is presenting an alternative to the original Lee-Carter Methodology, which used time series models and has become sort of a standard.

Two of the papers provide a cross national perspective using the human mortality database developed at the Max Planck Institute in Rostock, Germany. Another common thread in these papers is how to deal with change over time in the fitted model. And the papers differ in various ways. The papers differ in various ways in the mortality measures they use, the type of age detail and the fitting procedures.

I'll review the last two papers first because they have a lot of overlap. The paper presented by Sharon Yang, deals with the question of parameter change over time as we've seen in the Lee-Carter model. It uses the similar fitting procedure to what Lee and Carter did which is the singular value decomposition, which amounts to the method of principal components, but they make an advance over Lee-Carter by adding one principle component. They note in their paper that the pace of reduction and mortality has been increasing over time, and the second principle component that they use shows a breakpoint very clearly. So they introduce a shift in the model where the parameters change after this cutoff point. That emerges as equivalent to a change in the pattern of mortality reduction. As a goodness of fit measure, they're using a mean average percent difference from the original formulation of the Lee-Carter model, which had only one principle component.

I assume that the measure for which goodness of fit is assessed, is the logarithm of the central death rate, which is an estimator of the force of mortality at the age.

In any case, they find that the modified model outperforms the Lee-Carter model in all four countries and they look into the impacts of this advance that they've introduced on life expectancy at later ages from 50 to 75. They find that this makes a big difference really only in Japan, but it does make a difference in the other countries too. What's very admirable about this paper is that it goes on to acknowledge its own limitations, which include the fact that there is only one principal component that's allowed to have jumps or shifts. They identify only two principle components, maybe there are more. There may be different ways of identifying where the optimal jump points are. Finally, having identified a jump point, their forecasting approach with this model does not include similar jumps in the future. I just want to complement the authors, and my reaction to this is that more work needs to be done and I assume it will be forthcoming from the authors.

The paper by Chan, et al. looks at the type of approach to modeling the trend in mortality reduction and it's within the framework of conventional econometrics. ARIMA models are auto-regressive, integrated moving average models. You have an auto-regressive part, a moving average part and what integrated means is that you keep adding in terms, either on a regressive or moving average until the model that you're using has a white noise type error pattern. They do this by differencing the model and a different series is called an integrated series and the order of integration is the number of differences taken. They distinguish between trend and different stationary models. A different stationary model is then the ARIMA model with the appropriate level of differencing and the right parameters.

A trend stationary model has a constant trend parameter over time, which behaves or looks like what you would get if you just did a regression on time. It's not exactly what we have. In this approach, you can model the error term and capture other components of it, but in any event, as far as the trend is concerned, you've got single parameter.

The authors look at Canada, England, Wales and the United States. Like Lee and Carter, they use the logarithm of the mortality rate. They use the singular value decomposition with only one principal component. Also, they employ a broken trend stationary model based on a certain statistical test that you can find in the paper. They find that the breakpoints in this trend—the places where the trend breaks—are remarkably similar in the countries that they are looking at, at 1974 and 1975, which has already been noted several times in this conference as a place where something happened to the trend of mortality.

Now the difference between the broken trend model and the unbroken trend makes a difference in Canada and England and Wales, but not in the United States. Now that doesn't mean that it's not useful. The authors go on to look at the annual percentage reduction and age-specific death rates at ages 45-95, and their evidence suggests an increase in the pace of mortality reduction. They go on to look at an impact measure of the total probability of surviving to age 100 from birth.

And to me the differences between their model and the unbroken model appear to be substantively significant, only in England and Wales, but that doesn't mean that they're not statistically significant.

The authors also address the problem that further jumps may occur in the future, and they suggest the method of adaptive forecasting to update the model. That has to be done as new data come in. So again, here's another paper where the conclusion is more work needs to be done.

Now on the Renshaw and Haberman paper, I've already got one comment, because the Census Bureau is cited as using the Lee-Carter model. That's not exactly correct. What happened was in the late 1990s, we produced a series of projections in which we interpolated between mortality at what the jump off time was and a projection made by Shripad Tuljapurkar, of Mountain View Associates, which was a Lee-Carter projection. Tuljapurkar did a Lee-Carter projection for the United States as a whole, and what we did, is we interpolated the mortality schedules by race to Shripad Tuljapurkar's projection of mortality for the U.S. as a whole at some point in the future, like 2025 (I forget exactly). So we were assuming convergence of mortality rates to a projection which was a Lee-Carter projection, we were not using Lee-Carter's method. We've gone on to try it—in my own personal experience—and found that as the other authors have, the original formulation of the Lee-Carter model doesn't work.

The paper by Renshaw & Haberman is mainly concerned with uncertainty and confidence intervals. They use the method of maximum likelihood, under the generalized linear modeling framework. They compare the original Lee-Carter model to a Linear Poisson model with two parameters. (Although, the second parameter—their gamma parameter—which is on page 2 of their paper is set to zero in the forecasts.) They use various measures including life expectancies over time and also fixed rate annuities. They construct their confidence intervals via bootstrap-related methods. What they come up with is that you get smaller confidence intervals. That is, you have less uncertainty with their Linear Poisson model than you would with Lee-Carter model. But, we've heard that there are a variety of things that can be introduced that widen the confidence intervals. Although, I think that those are constructive introductions.

Again, these authors also make the point that this rectangularization of mortality that's been commented upon in the literature and it goes back to a paper by Fries on the Compression of Mortality, is not really what's taking place at late age. That's been an observation also of the people who maintain this human mortality database at the Max Planck Institute.

And I fully agree with the authors on the fact that we need to compare different models as has been done in these analyses. Also that in the long run, judgment needs to be made. These types of things cannot be simply done mechanically. At some point or another, substantive judgment comes into play.

I will conclude the discussion by saying that there's another common theme that comes through in all of the papers, which is that more work needs to be done and I assume this is ongoing.