WP 5.3
16 October 2013

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

Joint Eurostat/UNECE Work Session on Demographic Projections
organised in cooperation with Istat
(29-31 October 2013, Rome, Italy)

Item 5 – Assumptions on Future Mortality

Coherent forecasting of multiple-decrement life tables: Compositional models for French Cause of Death data, 1925-2008

Jim Oeppen, Max Planck Institute for Demographic Research (MPIDR)
Carlo Giovanni Camarda, Institut National d'Études Démographiques (INED)

Coherent forecasting of multiple-decrement life tables: compositional models for French Cause of Death data, 1925-2008.

Jim Oeppen* and Carlo Giovanni Camarda**

***Max Planck Institute for Demographic Research**
****Institut National d'Études Démographiques, Paris.**

Long Abstract.

There is both an attractive logic to, and a strong demand for, coherent mortality forecasting of sub-groups within populations that can be aggregated to a meaningful total. This has been difficult to achieve in practice, so that it is not currently recommended as a general approach. For example, separate sex, or spatial, forecasts using the Lee-Carter model have run into coherence problems (Lee and Nault, 1993). The components diverge in the long run in ways that are implausible when compared with the historical record. Similar problems apply to cause of death forecasting. Wilmoth (1995) quantified the way in which mortality forecasts disaggregated by cause of death tend to be more pessimistic than single decrement forecasts. Over time, the causes with the lower rates of decline come to dominate the projection. It is clear that the dependencies, or relative balances, between the decrements have not been adequately modeled.

This paper explores a different approach to disaggregation and combination. It recognises that there are two views of mortality change. One is that mortality rates are the fundamental entities to be modelled. The established practice in mortality forecasting is to model the log transform of the rates. This has two advantages. It exploits the approximately log-linear decline in mortality by age with respect to time. Secondly, it transforms the rates into the real space so that multivariate statistical methods for unbounded variables can be employed in the knowledge that the inverse transform ensures that all estimated rates exist in the positive real space.

An alternative view, explored by Vaupel and Yashin(1987), treats mortality improvement as saving lives and thus focuses on perturbation of the death distribution. The key feature for this paper is that the d(x) values are a density and obey a unit sum constraint for both single and multiple-decrement tables. Thus they are intrinsically relative rather than absolute values across decrements as well as ages. Since disaggregated models usually fail because they do not maintain relative values, it is hoped that this feature can be exploited to obtain coherent forecasts.

The d(x) values are converted to log-ratios with a common denominator for each period across all decrements. It can be shown that, because of the unit sum constraint, the results are independent of the value chosen for the common denominator and the original values can always be recovered by back-transformation. Once the death densities have been transformed into the real space, the full range of multivariate statistics can be applied. In this paper, both singular value decomposition and regression are employed to obtain forecasts. After forecasting, the estimated log-ratios are back-transformed to get positive values that conform to the unit sum constraint.

Although it may be unconventional to model the d(x) values it should be remembered that the original Lee-Carter (1992) model and the Booth-Maindonald-Smith (2002) variant adjust the time parameter to match D(x) and D(x,t) respectively.

The first step in evaluating this method was to apply the model in a single mortality-decrement context and contrast the results with those obtained from variants of the LC model. For this purpose we follow the approach of Booth et al. (2006) in the selection of countries and the time periods for fitting and forecasting. To make the comparison we take the conservative approach of comparing the models in terms of mortality rates and life expectancy at birth, rather than in d(x) terms. The results show that the new model consistently outperforms the original Lee-Carter model when using long fitting series from 1900 or 1920. Modeling d(x) change via perturbation seems to be more robust against large changes in mortality structure. When fitting over the shorter time periods of the Lee-Miller and Booth-Maindonald-Smith models, the d(x) model produces comparable average results. The variability of the results across countries makes it difficult to select a winner and suggests that a wider comparison is needed. It should be noted that the version of the log-ratio model used here is automatic and does not use the post-estimation iteration of the Lee-Carter model and its variants.

One of the features of this model for conventional life table forecasting is that the life expectancy trajectories tend to be linear for advanced economies, in line with the findings of Oeppen and Vaupel (2002) and White (2002). This is in contrast to models based on log mortality rates which tend to produce forecast life expectancies with a declining rate of increase. However, the d(x) model does produce such forecasts for some Eastern European countries.

The compositional d(x) model will be applied to time--series of mortality rates by cause of death, age, and sex for France from 1925 to 2008. The causes were classified into major headings to form seven coherent series across the revisions of the ICD. The age groups range over 0, 1--4, 5--9,…,95--99, 100+. The evaluation will include an assessment of whether the model reflects the usual pessimism of cause-specific projections identified by Wilmoth (1995).

Modeling log-ratios of demographic densities via regression models offers great flexibility. Covariates such as smoking or GDP per capita can be incorporated. Both mortality and fertility are frequently forecast by specifying a future series of e(0) or TFR, from which age-specific densities are derived. For life expectancy, L(x) / e(0) is a density, as is the completed family size distribution.