Modeling and forecasting mortality with economic growth: a multi-population approach

Tim J. Boonen
Amsterdam School of Economic, University of Amsterdam

Hong Li
School of Finance, Nankai University

12th International Longevity Risk and Capital Markets Solutions Conference, September 29-30, 2016, Chicago, U.S.A.

Table of Contents

Background

- Human life expectancy has been increasing substantially over the past decades.
- Forecasting future life expectancy is crucial for various practices (demographics, finance, actuary, ...)
- Stochastic mortality modeling has been widely studied since 1990s.
- Modeling mortality in multiple populations has attracted increasing attention recently
  - Risk management, demographic analysis, capital market solution, ...;
  - Forecasts for individual populations may be improved by exploiting more data.

Motivation

- Most stochastic mortality models up to date are "extrapolative"
  - They summarise historical mortality trends by latent factor(s);
  - Forecasts are obtained by extrapolating these factors.
- Extrapolative models achieved great success
  - Single-population: Lee and Carter (1992), Cairns-Blake-Dowd (2006), Renshaw and Haberman (2006) ...
  - Multi-population: Li and Lee (2005), Dowd et al. (2011), Hyndman et al. (2013) ...
- However, there are potential limitations
  - The interpretation of these latent factors are sometimes unclear;
  - It is not sure whether the historical latent trends will continue in the future.

Motivation

- Can mortality trends be approximated by observable factors?
- Various candidates
  - Smoking pattern;
  - Temperature;
  - Unemployment rate;
  - ...
  - Economic growth;

Literature review

- Cross-sectional — economic level and mortality level.
  - Countries with a higher level of higher national income have on average a higher life expectancy (Preston 1975).
- Time-series — economic growth and mortality change. Mixed results.
- Positive:
  - Economic growth positively affects mortality decline (Brenner 2005 ; Bloom and Canning 2005);
  - Health and life expectancy positively affects economic growth (Bloom and Canning 2005; Bhargava et al. 2011).
- Negative:
  - Mortality declines termporarily accelerates during economic recessions and slows down during economic expansions (Ruhm 2005; Tapia Granados 2005);
  - No evidence that increase in life expectancy causes economic growth (Acemoglu and Johnson 2007).

Motivation

- However, few studies combined stochastic mortality modeling,
  - Hanewald (2011): correlation between k_t in the Lee-Carter model and macroeconomic fluctuations.
  - Niu and Melenberg (2014): include real GDP per capita in the Lee-Carter model.
- These studies are restricted to single-population contexts.

Research question

- Are there long-term relationships between aggregate economic growths and aggregate mortality changes in a group of closely related populations?
- If so, can mortality forecasts be improved by exploiting these relationships?
- We extend the Li-Lee model to include GDP data, and forecasts future mortality by extrapolating the observable trends of GDP.
- Satisfying results are obtained in empirical studies.

Table of Contents

The general specification

- I populations. For population i, age x, year t,

$$\log m_{i,x,t} = a_{i,x} + \sum_{\ell=1}^{L} \gamma_{\ell,x} g_{\ell,t} + B_x K_t + b_{i,x} k_{i,t} + \varepsilon_{i,x,t}$$

- $g_{\ell,t}$ the ℓ-th principal component of the logarithm of real GDP per capita of the I populations.
- K_t the common latent factor.
- $k_{i,t}$ the specific latent factor for population i.

Normalization

- To identify the model, we propose

$$
\begin{aligned}
\sum_{x=1}^{N} B_x &= 1, \\
\sum_{t=1}^{T} K_t &= 0, \\
K &= (K_1, \ldots, K_T) \neq 0, \\
\sum_{t=1}^{T} K_t g_{\ell,t} &= 0, \ \text{for} \ \ell = 1, \ldots, L,
\end{aligned}
$$

Table of Contents

Data

- Four groups
  - 14 low mortality countries (15-low in Li-Lee except West Germany), 1947 - 2010;
  - 6 Eastern European countries (as in Li-Lee), 1969 - 2009;
  - 6 former Soviet Union countries (Belarus, Estonia, Latvia, Lithuania, Russia, Ukraine), 1969 - 2010;
  - Female and male population of Sweden, 1947 - 2010.
- Data source
  - Uni-sex mortality rates from age 0 to 99;
    - Human mortality database.
  - Real GPD per capita (purchasing power parity based).
    - Maddison project on World Economy;
    - ERS International Macroeconomic Data Set.

Model specification

- We compare two models.
- The GDP model without common latent factor

$$\log m_{i,x,t} = a_{i,x} + \sum_{\ell=1}^{L} \gamma_{\ell,x} g_{\ell,t} + b_{i,x} k_{i,t} + \varepsilon_{i,x,t}.$$

- The Li-Lee model

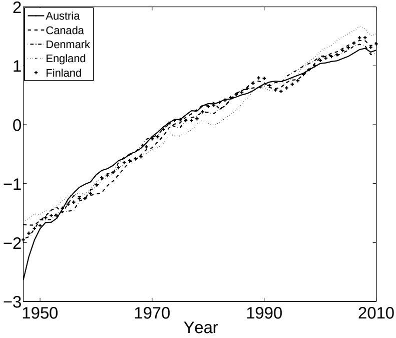
$$\log m_{i,x,t} = a_{i,x} + B_x K_t + b_{i,x} k_{i,t} + \varepsilon_{i,x,t}$$

Number of principal components

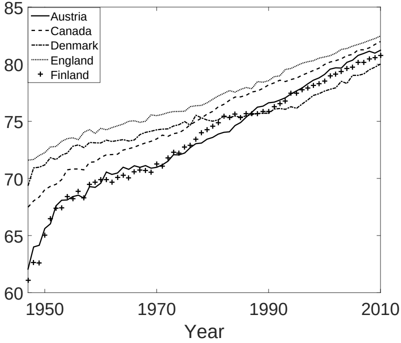
	1 PC	2 PCs
Group 1	98.70	99.57
Group 2	79.20	97.20
Group 3	77.81	98.01

- $L = 1$ for Group 1 and 4; $L = 2$ for Group 2 and 3.

Group 1 Life expectancy and GDP

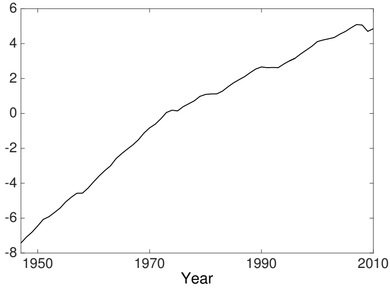


(a) GDP Group 1 (selected)

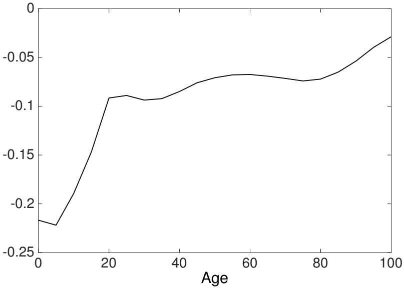


(b) Life Expectancy Group 1 (selected)

Group 1 GDP model

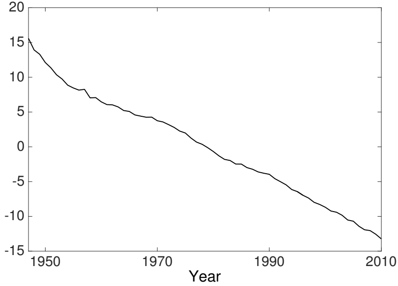


(a) g_t

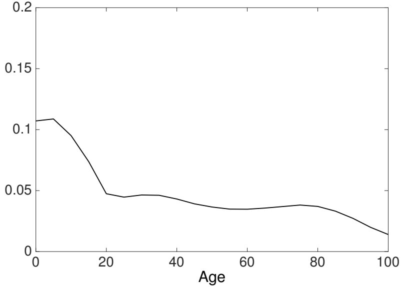


(b) γ_x

Group 1 Li-Lee model

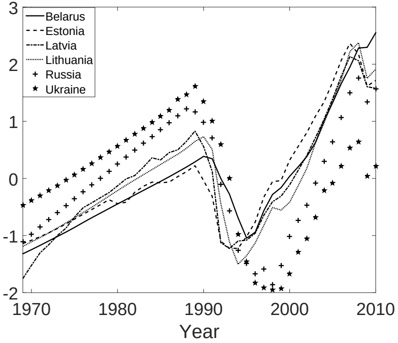


(a) K_t

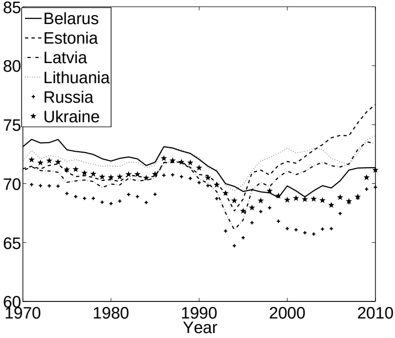


(b) B_x

Group 3 Life expectancy and GDP

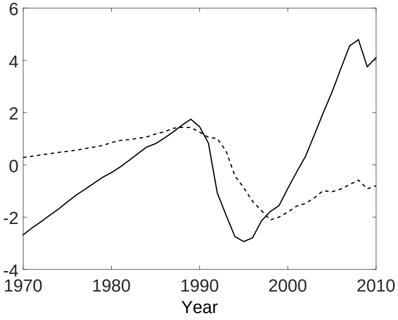


(a) GDP Group 3

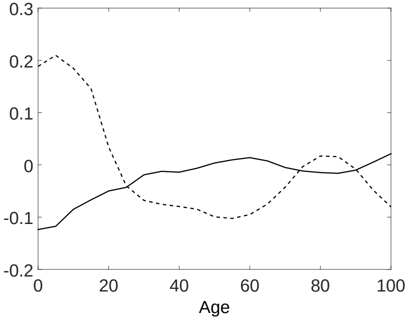


(b) Life Expectancy Group 3

Group 3 GDP model

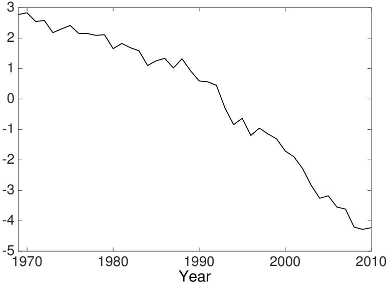


(a) $g_{1,t}$ (solid line) and $g_{2,t}$ (dashed line)

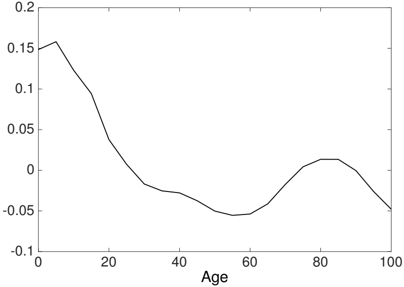


(b) $\gamma_{1,x}$ (solid line) and $\gamma_{2,x}$ (dashed line)

Group 3 GDP model



(a) K_t



(b) B_x

Variation explained by the common factors

- The ratio R_C used in Li and Lee (2015)

	Li-Lee	GDP
Group 1	0.90	0.85
Group 2	0.47	0.47
Group 3	0.53	0.55
Group 4	0.91	0.86

AIC and BIC ratios

- AIC ratio

$$2\log\hat{L} - 2m.$$

- BIC ratio

$$\log\hat{L} - \frac{1}{2}m\cdot\log M.$$

- $\hat{L}$ — likelihood;
- m — number of free parameters;
- M — sample size used for estimation.

AIC and BIC ratios

	Li-Lee		GDP	
	BIC	AIC	BIC	AIC
Group 1	-15283	-18505	-14987	-18391
Group 2	-4152	-4735	-4064	-4676
Group 3	-4266	-4903	-4205	-4905
Group 4	-2250	-2789	-2015	-2680

Model specification

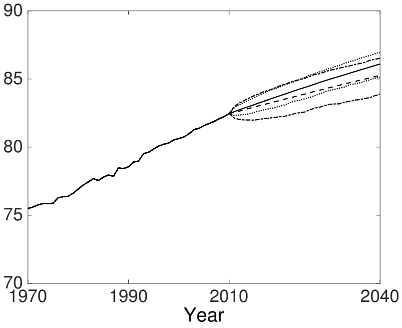
- Common factors — random walk with drift

$$y_t = c + y_{t-1} + \eta_t,\ \ \eta_t \overset{i.i.d.}{\sim} N(0, \sigma_y^2),$$

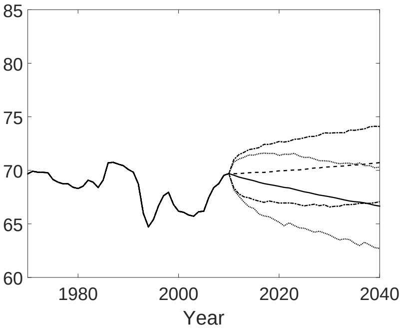
- Population-specific factors — AR(1)

$$k_{i,t} = c_{i,0} + c_{i,1} k_{i,t-1} + \omega_{i,t},\ \ \omega_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma_{\omega,i}^2).$$

Model specification



(a) US



(b) Russia (Group 3)

Out-of-sample forecast: RMSFE

- Using data up to 2000, and forecast from 2001 to the end of the sample.
- Relative root mean squared forecast error (RMSFE)

$$\sqrt{\frac{1}{N(U-\hat{u})}\sum_{u=\hat{u}}^{U}\sum_{x=1}^{N}\frac{(\log m_{i,x,u}-\widehat{\log m}_{i,x,u})^2}{|\log m_{i,x,u}|}}$$

- $\hat{u}$ — 2001;
- U — end of sample.

Choice of sample size

- The common factors do not seem linear in sample.
- Test for multiple-structural break (Van Berkum et al. 2016).
- Starting date for estimating the random walk with drift

	Li-Lee	GDP
Group 1	1956	1971
Group 2	1991	1993
Group 3	1989	1993
Group 4	1956	1949

Performance

	Li-Lee	GDP
Group 1	0.07	0.07
Group 2	0.08	0.06
Group 3	0.08	0.07
Group 4	0.07	0.06

- Similar for Group 1.
- Better for other groups.

Table of Contents

A more general framework

- The product-ratio model in Hyndman et al. (2013).
- It can be interpreted as a extension of the Li-Lee model.

$$\log f_{i,x,t} = \tilde{\mu}_{i,x} + \sum_{p=1}^{P} b_{p,x}\phi_{p,t} + \sum_{q=1}^{Q} b_{i,q,x}\Phi_{i,q,t} + \tilde{\varepsilon}_{i,x,t},$$

- $\log f_{i,x,t}$ — smoothed $\log m_{i,x,t}$;
- P common factors;
- Q specific factors for each population.
- More general time series processes
  - ARIMA for $\phi_{p,t}$;
  - ARFIMA for $\Phi_{i,q,t}$.

GDP product-ratio

- Replace the first common latent factor by the mean of the logarithm of GDP

$$\phi_{1,t} = \frac{\sum_{i=1}^{I} \log G_{i,t}}{I}$$

- Replace the first population-specific factor for population i by the difference between its GDP and the common factor

$$\Phi_{1,t} = \log G_{i,t} - \phi_{1,t}$$

Out-of-sample forecast

	PR	GDP
Group 1	0.05	0.05
Group 2	0.09	0.07
Group 3	0.10	0.08
Group 4	0.07	0.06

- $P = 3, \; Q = 3.$

Conclusion

- We study the long-term relationship between the aggregate economic growth and the aggregate mortality change in a group of closely related populations.
- We extend the Li-Lee model by including the observable trend of real GDP per capita.
- The proposed model has better in-sample fit (AIC, BIC), and out-of-sample forecast performance than the Li-Lee model.
- The relationship seems to hold in a more general framework.

Conclusion

Thank You!