Forecasting death rates using exogenous determinants

Declan French (Queens University, Belfast, UK)
Colin O'Hare (Monash University, Australia)

Motivation

- 3 major risks faced by a pension provider :
  interest rate risk, inflation risk and longevity risk.
- Deal with longevity risk by selling the liability via an insurance or reinsurance contract.
- Pay a fixed amount based on expected mortality rates in return for a payment based on actual realized mortality rates (a 'q-forward').
- Pricing relies on accurate mortality forecasts.

Forecasting death rates using exogenous determinants

Background

- Lee-Carter (1992): $\ln m_{x,t} = a_x + b_x k_t$
- Renshaw et al.(2006) : $\ln m_{x,t} = a_x + b_x^1 k_t + b_x^2 \gamma_{t-x}$
- Currie (2006) : $\ln m_{x,t} = a_x + k_t + \gamma_{t-x}$
- Cairns et al. (2006): $\ln\left(q_{x,t} \big/ 1 - q_{x,t} \right) = k_t^{1} + k_t^{2}\left(x - \bar{x} \right)$
- Hyndman and Ullah (2007) $f_t(x) = \mu(x) + \Sigma_{k=1}^{K} \beta_{tk} \varphi_k(x)$
- Plat (2009) : $\ln m_{x,t} = a_x + k_t^{1} + k_t^{2}\left(\bar{x} - x \right) + k_t^{3}\left(\bar{x} - x \right)^{+} + \gamma_{t-x}$

Forecasting death rates using exogenous determinants

Research questions

- can the factors found in stochastic mortality forecasting models be associated with real-world trends in health-related variables;
- does inclusion of health-related factors in models improve forecasts?
- do resulting models give better forecasts than existing stochastic mortality models?

Forecasting death rates using exogenous determinants

QUEEN'S UNIVERSITY MANAGEMENT SCHOOL

Queen's University Belfast

Methodology

$$m_{xt} = \lambda_x' \mathrm{F_t} + \mathrm{e_{xt}} \quad \text{for r factors}$$

- Determine r (Bai and Ng, 2002).
- Are observed variables $\mathrm{G_t}$ a linear combination of latent variables $\mathrm{F_t}$? (Bai and Ng, 2006).
  - Exact case $\mathrm{G_{jt}} = \delta_j{}'\mathrm{F_t} \; \forall \, t$ then $\hat{\tau}_t(j) = \frac{\hat{\mathrm{G}}_{jt} - \mathrm{G}_{jt}}{(var\hat{\mathrm{G}}_{jt})^{1/2}}$
    1. $A(j) = \frac{1}{T}\sum_1^T 1(|\hat{\tau}_t(j)| > \phi_\alpha)$
    2. $M(j) = \max_{1 \le t \le T} |\hat{\tau}_t(j)|$
  - Approximate factor $\mathrm{G_{jt}} = \delta_j{}'\mathrm{F_t} + \varepsilon_{jt} \; \forall \, t$.
    3. $NS(j) = \frac{\widehat{\mathrm{var}}(\hat{\varepsilon}(j))}{\widehat{\mathrm{var}}(\hat{\mathrm{G}}(j))}$
    4. $R^2(j) = \frac{\widehat{\mathrm{var}}(\hat{\mathrm{G}}(j))}{\widehat{\mathrm{var}}(\mathrm{G}(j))}$

Forecasting death rates using exogenous determinants

QUEEN'S UNIVERSITY MANAGEMENT SCHOOL

Queen's University Belfast

Methodology

$$m_{xt} = \lambda_x' \mathrm{F_t} + \mathrm{e_{xt}} \quad \text{for r factors}$$

- Are observed variables $\mathrm{G_t}$ a linear combination of latent variables $\mathrm{F_t}$? (Bai and Ng, 2006)
  - Testing as a group
    5. $(\rho_k^{2-}, \rho_k^{2+}) = \left(\tilde{\rho}_k^2 - 2\phi_\alpha \frac{\tilde{\rho}_k(1-\tilde{\rho}_k^2)}{\sqrt{T}}, \tilde{\rho}_k^2 + 2\phi_\alpha \frac{\tilde{\rho}_k(1-\tilde{\rho}_k^2)}{\sqrt{T}}\right)$

    where $\tilde{\rho}_k$ is the kth canonical correlation between $\mathrm{G_t}$ and $\tilde{\mathrm{F}}_\mathrm{t}$.

Forecasting death rates using exogenous determinants

QUEEN'S UNIVERSITY MANAGEMENT SCHOOL

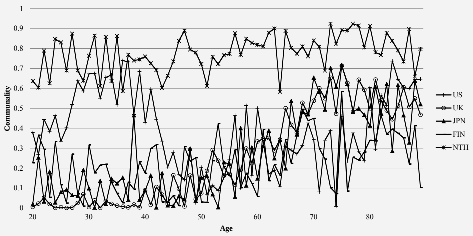
Queen's University Belfast

Data

- Mortality
  - Human Mortality Database (Berkeley/Max Planck)
  - males in the US, UK, Japan, Finland, Netherlands and Sweden
  - Estimates 1970-2000 & forecasts, 2001-2009
- Health
  - OECD Health data 2009
  - Alcohol, Tobacco, Fat, Fruit and vegetable, GDP, Health expenditure

Forecasting death rates using exogenous determinants

Proportion of variance explained by principal component extraction (communality) by age for male age-specific mortality rates, 1970-2000.



Forecasting death rates using exogenous determinants

Testing the factors in US age-specific mortality rates by single age 20-89, 1970-2000

G_j	A(j)	M(j)	R^2(j)	NS(j)	$\hat{\rho}(k)^2$
G_1, Alcohol	0.968	95.5	0.025 (-0.084,0.133)	39.0	0.253 (-0.011, 0.518)
G_2, Tobacco	0.968	58.6	0.034 (-0.091,0.158)	28.7	0.076 (-0.103, 0.256)
G_3, Fat	0.903	58.2	0.043 (-0.097,0.183)	22.1	-
G_4, Fruit & Veg	0.935	30.4	0.151 (-0.081,0.383)	5.6	-
G_5, GDP	0.935	98.1	0.034 (-0.091,0.158)	28.8	-
G_6, Health exp	0.935	143.6	0.046 (-0.098,0.191)	20.6	-

Forecasting death rates using exogenous determinants

Testing the factors in UK age-specific mortality rates by single age 20-89, 1970-2000

G_j	A(j)	M(j)	R^2(j)	NS(j)	$\hat{\rho}(k)^2$
G_1, Alcohol	1.000	97.8	0.020 (-0.078,0.119)	48.0	0.180 (-0.065,0.425)
G_2, Tobacco	0.903	72.6	0.093 (-0.102,0.287)	9.8	-
G_3, Fat	1.000	138.5	0.016 (-0.072,0.104)	61.0	-
G_4, Fruit & Veg	0.839	55.0	0.080 (-0.103,0.263)	11.5	-
G_5, GDP	0.968	464.4	0.002 (-0.026,0.029)	651.9	-
G_6, Health exp	1.000	458.9	0.003 (-0.036,0.043)	314.2	-

Forecasting death rates using exogenous determinants

Forecasting (King and Soneji, 2011)

- *Standard linear regression formulation*

$$m_{x,t} \sim N\left(\mu_{x,t}, \frac{\sigma_x^2}{b_{x,t}}\right), \qquad with \ \mu_{x,t} = \mathbf{Z}_{x,t}\beta_x$$

- *with smoothing across ages* $P(\beta|\theta) \propto exp\left(-\frac{1}{2}H^{\mu}[\beta,\theta]\right)$

$$H^{\mu}[\beta,\theta] \equiv \frac{1}{2}\sum s_{i,j}\left\|\beta_i - \beta_j\right\|_{\theta}^2$$

Forecasting death rates using exogenous determinants

Root mean square error for forecasting results, 2001-2009

	US	UK	Japan	Finland	Nld
Lee Carter	0.0043	**0.0051**	0.0022	0.0056	0.0079
Hyndman Ullah	0.0046	0.0058	0.0033	0.0076	0.0087
Girosi and King	0.0042	0.0075	**0.0018**	0.0058	0.0081
King and Soneji	**0.0041**	0.0070	0.0021	**0.0053**	**0.0075**

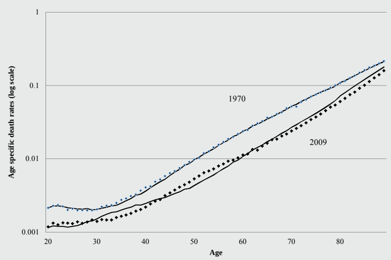
Forecasting death rates using exogenous determinants

Root mean square error for forecasting results, 2001 and 2009

	US		UK		Japan		Finland		Netherlands	
	2001	2009	2001	2009	2001	2009	2001	2009	2001	2009
Lee Carter	0.0011	0.0071	0.0029	**0.0058**	0.0036	**0.0011**	0.0028	0.0057	**0.0030**	0.0125
Hyndman Ullah	**0.0010**	0.0076	**0.0023**	0.0071	**0.0020**	0.0045	0.0045	0.0076	0.0031	0.0135
Girosi and King	0.0011	0.0070	0.0044	0.0089	0.0020	0.0012	**0.0024**	0.0061	0.0033	0.0127
King and Soneji	0.0011	**0.0069**	0.0037	0.0083	0.0026	0.0013	0.0027	**0.0053**	0.0032	**0.0116**

Forecasting death rates using exogenous determinants

US actual, fitted and forecast mortality by age at 1970 and 2009. Actual schedule are dots. Fitted and forecasts are solid.



Forecasting death rates using exogenous determinants

Conclusions

- Identified statistical factors with exogenous factors
- Exogenous factors improve forecasts
- Results comparable or better than conventional models.
- Fruit & veg consumption and smoking best describes the factor structure
- Age-specific variables would improve forecasts.

Further work :
(i) Lagged variables (ii) the semi-parametric estimation Connor et al. (2012)

Forecasting death rates using exogenous determinants