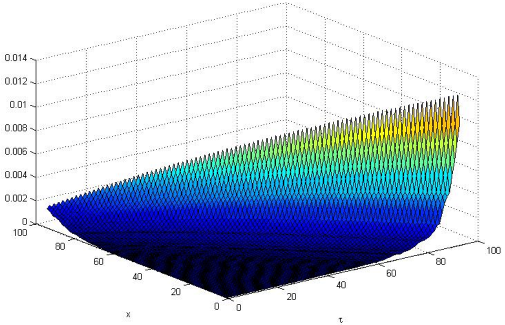


Coherent Modeling of the Risk in Mortality Projections: Theory and Applications

Georgia State University

Nan Zhu and Daniel Bauer

Department of RMI, Georgia State University

Motivation

- Consider a 60 year old individual who faces *systematic* mortality risk:
  1. Currently expected to die with probability of 0.7% next year (mortality rate). How uncertain is this appraisal? What are the chances that the rate is 0.65% or 0.75%?

  ⇒ Risk in mortality rates

  2. Currently expected to live 25.1 more years. How uncertain is this number? What are the chances that it changes to 23.7 or 24.5 next year?

  ⇒ Risk in mortality projections

- While the two questions are related, and the distinction does not matter in theory, it is relevant for the econometrical/statistical approach
- For personal financial planning/household finance, insurers' liability risk evaluation, and government/public economics, question 2 may be more suitable

Motivation

- Consider a 60 year old individual who faces *systematic* mortality risk:
  1. Currently expected to die with probability of 0.7% next year (mortality rate). How uncertain is this appraisal? What are the chances that the rate is 0.65% or 0.75%?

  ⇒ Risk in mortality rates

  2. Currently expected to live 25.1 more years. How uncertain is this number? What are the chances that it changes to 23.7 or 24.5 next year?

  ⇒ Risk in mortality projections

- While the two questions are related, and the distinction does not matter in theory, it is relevant for the econometrical/statistical approach
- For personal financial planning/household finance, insurers' liability risk evaluation, and government/public economics, question 2 may be more suitable

Motivation

- Consider a 60 year old individual who faces *systematic* mortality risk:
  1. Currently expected to die with probability of 0.7% next year (mortality rate). How uncertain is this appraisal? What are the chances that the rate is 0.65% or 0.75%?

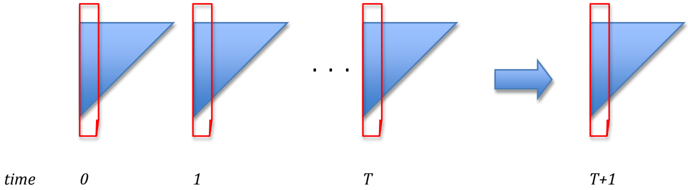
  $\Rightarrow$ Risk in mortality rates

  2. Currently expected to live 25.1 more years. How uncertain is this number? What are the chances that it changes to 23.7 or 24.5 next year?

  $\Rightarrow$ Risk in mortality projections

- While the two questions are related, and the distinction does not matter in theory, it is relevant for the econometrical/statistical approach
- For personal financial planning/household finance, insurers' liability risk evaluation, and government/public economics, question 2 may be more suitable

Motivation



- Current stochastic mortality models focus on stochastically forecasting mortality rates (question 1)

→ red in graphic

- This essay considers the risk in mortality projections (question 2)

→ blue in graphic

Comparison with Interest Rate Models/Yiled Curve Models

- Analogy in the motivation:
  - Want to forecast yield curve $p(T+1,\tau)$ based on $p(t,\tau)$, $0 \leq t \leq T$ (similar to question 2)
  - Know $p(t,\tau) = \mathbb{E}_t^{\mathbb{Q}}[\exp(-\int_t^{t+\tau} r_s\, ds)]$, so (theoretically) modeling risk in r_t should be sufficient (similar to question 1)
  - But to identify reasonable specification, we also need to consider cross-sectional data (persistence vs. transience of errors, etc.)
- But there are key differences:
  - Age: additional dimension
    - How does age enter into model equation?
    - What are appropriate models?
  - Data
    - Age/term panels
    - Where do we get "mortality forecasts"?
  - Risk adjustments ($\mathbb{P}$ vs. $\mathbb{Q}$) and consequences

- Analogy in the motivation:
  - Want to forecast yield curve $p(T+1,\tau)$ based on $p(t,\tau)$, $0 \leq t \leq T$ (similar to question 2)
  - Know $p(t,\tau) = \mathbb{E}_t^{\mathbb{Q}}[\exp(-\int_t^{t+\tau} r_s\, ds)]$, so (theoretically) modeling risk in r_t should be sufficient (similar to question 1)
  - But to identify reasonable specification, we also need to consider cross-sectional data (persistence vs. transience of errors, etc.)
- But there are key differences:
  - Age: additional dimension
    - How does age enter into model equation?
    - What are appropriate models?
  - Data
    - Age/term panels
    - Where do we get "mortality forecasts"?
  - Risk adjustments ($\mathbb{P}$ vs. $\mathbb{Q}$) and consequences

Literature Review: Mortality

Large literature with various methods on mortality forecasting:

- Lee-Carter approach (Lee and Carter (JASA, 1992)):

$$\log m(t,x) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$$

- CBD-Perks model (Cairns et al. (JRI, 2006)):

$$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_2^{(2)}(x - \bar{x})$$

- P-splines method (Currie et al. (Statistical Modeling, 2004)): non-parametric model

All these methodologies...

... rely on past mortality data to project the *mortality experience* in some optimal sense

... pay little attention on the uncertainty (error estimates) associated with the *projections*

... may fail to identify the transiency of different random sources

Literature Review: Mortality

Large literature with various methods on mortality forecasting:

- Lee-Carter approach (Lee and Carter (JASA, 1992)):

$$\log m(t,x) = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)}$$

- CBD-Perks model (Cairns et al. (JRI, 2006)):

$$\text{logit } q(t,x) = \kappa_t^{(1)} + \kappa_2^{(2)}(x - \bar{x})$$

- P-splines method (Currie et al. (Statistical Modeling, 2004)): non-parametric model

All these methodologies...

... rely on past mortality data to project the *mortality experience* in some optimal sense

... pay little attention on the uncertainty (error estimates) associated with the *projections*

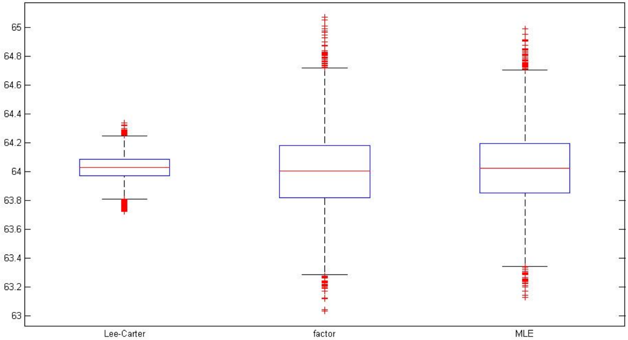
... may fail to identify the transiency of different random sources

Literature Review: Term Structure Models

- Factor models
  - **Factor analysis**: Litterman and Scheinkman (Journal of Fixed Income, 1991), Rebonato (Interest Rate Option Models, 1998)...
  - **VaR models**: Fama and Bliss (AER, 1987), Diebold and Li (JEconometrics, 2006), Duffee (JFin, 2002)...
- Forward rate models
  - **HJM model**: Heath et al. (Econometrica, 1992), Filipović (Lecture Notes in Mathematics, 2004)...
  - **In the case of mortality**: Cairns et al. (ASTIN Bull., 2006), Barbarin (IME, 2008), Bauer et al. (IME, 2010), Bauer et al. (2011)...
- Finite-dimensional realizations/Affine term structure models
  - **FDR**: Björk and Gombani (FinStoch, 1999), Björk and Svensson (MathFinanc, 2001)...
  - **Affine Models**: Duffie and Kan (MathFinanc, 1996), Duffie et al. (Econometrica, 2000), Piazzesi (Handbook of Financial Econometrics, 2010)...
- Improvement of Cross-sectional constraints
  - Joslin et al. (RFS, 2011), Duffee (2011)...

Introduction

Preview of Results



- Confidence intervals for life expectancies in one year for a now 20 year old female (USA)
- Comparison with conventional mortality forecasting approach – Lee-Carter
  - Generally underestimate the risk in mortality projections

Forward survival probabilities: $\{_\tau p_x(t) | (\tau, x) \in \mathcal{C}\}$

Forward force of mortality (easier to model/work with):

$$\mu_t(\tau, x) = -\frac{\partial}{\partial \tau} \log\{_\tau p_x(t)\}$$

Consider time-homogeneous, Gaussian models

$$d\mu_t = (A\,\mu_t + \Lambda)\,dt + \Sigma\,dW_t$$

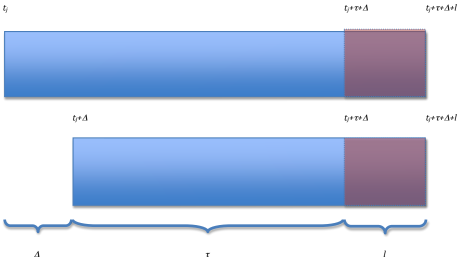
- $A = (\frac{\partial}{\partial \tau} - \frac{\partial}{\partial x})$ (infinitesimal generator of a strongly continuous semigroup)
- W_t: d-dimensional Brownian motion

Define

$$
\begin{aligned}
F_l\left(t_j, t_{j+1}, (\tau, x)\right) &= -\log\left\{ \frac{\dfrac{{}_{\tau+l}p_x(t_{j+1})}{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)}}{\dfrac{{}_{\tau}p_x(t_{j+1})}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)}} \right\} \\
&= -\log\left\{ \frac{{}_{\tau+l}p_x(t_{j+1})}{{}_{\tau}p_x(t_{j+1})} \Big/ \frac{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)} \right\}
\end{aligned}
$$

- l: time lag, $t_{j+1} - t_j = \Delta$

→ Measures the log change of the l-year marginal survival probability



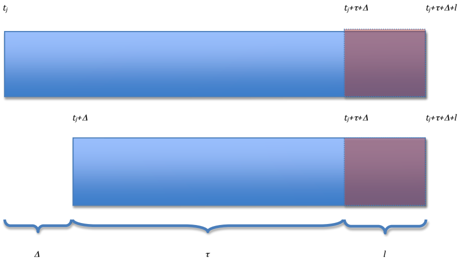
Define

$$F_l\left(t_j, t_{j+1}, (\tau, x)\right) = -\log\left\{ \frac{\dfrac{{}_{\tau+l}p_x(t_{j+1})}{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)}}{\dfrac{{}_{\tau}p_x(t_{j+1})}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)}} \right\}$$

$$= -\log\left\{ \frac{{}_{\tau+l}p_x(t_{j+1})}{{}_{\tau}p_x(t_{j+1})} \Big/ \frac{{}_{\tau+l+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)}{{}_{\tau+t_{j+1}-t_j}p_{x-t_{j+1}+t_j}(t_j)} \right\}$$

- l: time lag, $t_{j+1} - t_j = \Delta$

→ Measures the log change of the l-year marginal survival probability



Proposition I

If $t_{j+1} - t_j = \Delta$, the vectors

$$\bar{F}_l(t_j, t_{j+1}) = \left(\omega(\tau, x) \times \frac{F_l(t_j, t_{j+1}, (\tau, x))}{\sqrt{t_{j+1} - t_j}}\right)_{(\tau,x)\in\tilde{\mathcal{C}}}$$

$$= \left(\omega(\tau_1, x_1) \times \frac{F_l(t_j, t_{j+1}, (\tau_1, x_1))}{\sqrt{t_{j+1} - t_j}}, \ldots, \omega(\tau_K, x_K) \times \frac{F_l(t_j, t_{j+1}, (\tau_K, x_K))}{\sqrt{t_{j+1} - t_j}}\right),$$

$j = 1, 2, \ldots, N - 1$ are i.i.d. Gaussian distributed.

(the weights $\omega(\tau, x)$ allow for different weighting of future – e.g. $p(t, \tau)$)

"True" Mortality forecasts from market place or insurance prices → not available or at least not abundant/noisy...

Raw data: (deterministic) mortality forecasts generated from rolling windows of past mortality experience

- **Regions**: England/Wales (ENW), France (FRA), Japan (JPN), United States (USA), and West Germany (FRG)
- **Genders**: male and female
- **Years**: 1956-2006
  ⇒ 22 projections (each using mortality experience of past 30 years)
- **Methods**:
  Lee-Carter
    - Weighted-least-squares algorithm
    - Ages: 0-95
  CBD-Perks
    - Basic model w/o cohort effect
    - Ages: 25-95
  P-splines
    - Fixing the degree of freedom (*df*) at 20
    - Ages: 25-95

Factor Analysis of Mortality Forecasts

Data and Projection Methods

"True" Mortality forecasts from market place or insurance prices → not available or at least not abundant/noisy...

Raw data: (deterministic) mortality forecasts generated from rolling windows of past mortality experience

- **Regions**: England/Wales (ENW), France (FRA), Japan (JPN), United States (USA), and West Germany (FRG)
- **Genders**: male and female
- **Years**: 1956-2006
  - ⇒ 22 projections (each using mortality experience of past 30 years)
- **Methods**:

  Lee-Carter
  - Weighted-least-squares algorithm
  - Ages: 0-95

  CBD-Perks
  - Basic model w/o cohort effect
  - Ages: 25-95

  P-splines
  - Fixing the degree of freedom (*df*) at 20
  - Ages: 25-95

With $\Delta = t_{j+1} - t_j = 1$, $\bar{F}_l(t_j, t_{j+1})$ are i.i.d. Gaussian

$\Rightarrow \bar{F}_l(t_j, t_{j+1}) = a + bZ_j + \epsilon_j$

- $a \in \mathbb{R}^K$, $b \in \mathbb{R}^{K \times d}$, factors $Z_j \in \mathbb{R}^d$ with $\mathbb{E}(Z_j) = 0$ and $\text{Cov}(Z_j) = I_{d \times d}$

Estimates of a, b, and the number of factors, d, from principal component analysis:

- Decompose empirical covariance matrix of $\bar{F}_l(t_j, t_{j+1})$: $\hat{\Sigma}$

$$\hat{\Sigma} = U \times \text{diag}\{\lambda_1, \ldots, \lambda_K\} \times U' = \sum_{\nu=1}^{K} \lambda_\nu \, u_\nu \, u_\nu'$$

$$\approx \sum_{\nu=1}^{d} \lambda_\nu \, u_\nu \, u_\nu' = \text{Cov}\left(\sum_{\nu=1}^{d} u_\nu \sqrt{\lambda_\nu} \times Z_{\nu,j} \right)$$

- Determine the value of d
  - $\frac{\sum_{\nu=1}^{d} \lambda_\nu}{\sum_{\nu=1}^{K} \lambda_\nu} \geq \xi$
- Investigate the shape of the factors

With $\Delta = t_{j+1} - t_j = 1$, $\bar{F}_l(t_j, t_{j+1})$ are i.i.d. Gaussian

$\Rightarrow$ $\bar{F}_l(t_j, t_{j+1}) = a + bZ_j + \epsilon_j$

- $a \in \mathbb{R}^K$, $b \in \mathbb{R}^{K \times d}$, factors $Z_j \in \mathbb{R}^d$ with $\mathbb{E}(Z_j) = 0$ and $\text{Cov}(Z_j) = I_{d \times d}$

Estimates of a, b, and the number of factors, d, from principal component analysis:

- Decompose empirical covariance matrix of $\bar{F}_l(t_j, t_{j+1})$: $\hat{\Sigma}$

$$\begin{aligned} \hat{\Sigma} &= U \times \text{diag}\{\lambda_1, \dots, \lambda_K\} \times U' = \sum_{\nu=1}^{K} \lambda_\nu \, u_\nu \, u'_\nu \\ &\approx \sum_{\nu=1}^{d} \lambda_\nu \, u_\nu \, u'_\nu = \text{Cov}\left(\sum_{\nu=1}^{d} u_\nu \sqrt{\lambda_\nu} \times Z_{\nu,j} \right) \end{aligned}$$

- Determine the value of d
  - $\frac{\sum_{\nu=1}^{d} \lambda_\nu}{\sum_{\nu=1}^{K} \lambda_\nu} \geq \xi$
- Investigate the shape of the factors

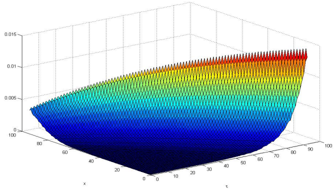
PCA Results: The Lee-Carter Approach

Country	Factor	Female Population Value	Female Population Percentage	Male Population Value	Male Population Percentage
United States					
	λ_1	4.50×10^{-2}	93.13%	7.51×10^{-2}	84.29%
	λ_2	1.80×10^{-3}	3.70%	8.30×10^{-3}	9.26%
	λ_3	6.15×10^{-4}	1.27%	2.60×10^{-3}	2.91%
	λ_4	4.79×10^{-4}	-	1.91×10^{-3}	-
	λ_5	2.29×10^{-4}	-	8.57×10^{-4}	-
	λ_6	2.07×10^{-4}	-	4.22×10^{-4}	-

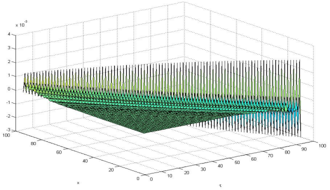
- First factor (PC1) dominates the rest
- Higher volatility for male population
  - Higher absolute value
  - Lower weight of first principle component

Factor Analysis of Mortality Forecasts

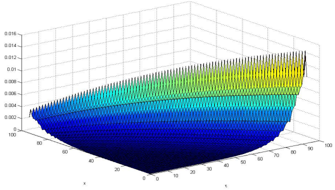
Surfaces of Eigenvectors: ENW/Lee-Carter



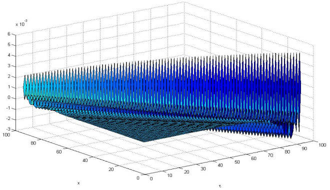
(a) female, PC1



(b) female, PC2



(c) male, PC1



(d) male, PC2

Factor Analysis of Mortality Forecasts

Summary

- Very similar shapes exhibited across different countries/genders/forecasting methods (at least for the first factor)

⇒ PC1

- Systematic, increasing in age/term
- Forward forces of mortality for high ages in the *far future* more volatile than in the *near future*

→ *slope factor*

⇒ PC2

- Male (consistently over time decreasing influence, even generate inverse relationship)
- Female (sometimes similar, sometimes rather unsystematic)

→ *twist factor*

- $d = 1$ for female population; $d = 2$ for male population (Lee-Carter)
- $d = 1$ for both genders (CBD & P-spline)

Factor Analysis of Mortality Forecasts

Summary

- Very similar shapes exhibited across different countries/genders/forecasting methods (at least for the first factor)

⇒ PC1

- Systematic, increasing in age/term
- Forward forces of mortality for high ages in the *far future* more volatile than in the *near future*

→ *slope factor*

⇒ PC2

- Male (consistently over time decreasing influence, even generate inverse relationship)
- Female (sometimes similar, sometimes rather unsystematic)

→ *twist factor*

- $d = 1$ for female population; $d = 2$ for male population (Lee-Carter)
- $d = 1$ for both genders (CBD & P-spline)

Factor Analysis of Mortality Forecasts

Summary

- Very similar shapes exhibited across different countries/genders/forecasting methods (at least for the first factor)

$\Rightarrow$ PC1

- Systematic, increasing in age/term
- Forward forces of mortality for high ages in the *far future* more volatile than in the *near future*

$\rightarrow$ *slope factor*

$\Rightarrow$ PC2

- Male (consistently over time decreasing influence, even generate inverse relationship)
- Female (sometimes similar, sometimes rather unsystematic)

$\rightarrow$ *twist factor*

- $d = 1$ for female population; $d = 2$ for male population (Lee-Carter)
- $d = 1$ for both genders (CBD & P-spline)

Factor Analysis of Mortality Forecasts

Summary

- Very similar shapes exhibited across different countries/genders/forecasting methods (at least for the first factor)

$\Rightarrow$ PC1

- Systematic, increasing in age/term
- Forward forces of mortality for high ages in the *far future* more volatile than in the *near future*

$\rightarrow$ *slope factor*

$\Rightarrow$ PC2

- Male (consistently over time decreasing influence, even generate inverse relationship)
- Female (sometimes similar, sometimes rather unsystematic)

$\rightarrow$ *twist factor*

- $d = 1$ for female population; $d = 2$ for male population (Lee-Carter)
- $d = 1$ for both genders (CBD & P-spline)

Simple factor models from the factor analysis

- $Y_\nu(j) \triangleq (u_\nu\sqrt{\lambda_\nu})^T\bar{F}_l = (u_\nu\sqrt{\lambda_\nu})^T\mathbb{E}[\bar{F}_l] + \lambda_\nu Z_{\nu,j},\ \nu = 1,\ldots,d$
- Regression equation of $\bar{F}_l(t_j, t_{j+1})$ on $Y_\nu(j)$ (similar to Diebold and Li, (JEconometrics, 2006)):

$$\begin{aligned}\bar{F}_l(t_j,t_{j+1}) &= \mathbb{E}\left[\bar{F}_l(t_j,t_{j+1})\right] + \sum_{\nu=1}^{d}\frac{u_\nu}{\sqrt{\lambda_\nu}}[Y_\nu(j) - (u_\nu\sqrt{\lambda_\nu})^T\mathbb{E}[\bar{F}_l(t_j,t_{j+1})]] + \epsilon_j \\ &\triangleq \tilde{m} + \sum_{\nu=1}^{d}\tilde{s}_\nu \times Y_\nu(j) + \epsilon_j\end{aligned}$$

- Simple, easy-to-estimate mortality forecasting methodologies
  - Simulate $Y_\nu(N) \sim N(\mu^s_{Y,\nu}, \sigma^s_{Y,\nu})$ with $(\mu^s_{Y,\nu}, \sigma^s_{Y,\nu})$ as sample mean and standard error of $Y_\nu(j)$, $j = 1,\ldots,N-1$
  - Forecast: $\bar{F}_l(t_N, t_{N+1}) = \tilde{m} + \sum_{\nu=1}^{d}\tilde{s}_\nu \times Y_\nu(N)$ together w/ known ${}_\tau p_x(t_N)$

Simple factor models from the factor analysis

- $Y_\nu(j) \triangleq (u_\nu \sqrt{\lambda_\nu})^T \bar{F}_l = (u_\nu \sqrt{\lambda_\nu})^T \mathbb{E}[\bar{F}_l] + \lambda_\nu Z_{\nu,j},\ \nu = 1, \ldots, d$
- Regression equation of $\bar{F}_l(t_j, t_{j+1})$ on $Y_\nu(j)$ (similar to Diebold and Li, (JEconometrics, 2006)):

$$
\begin{aligned}
\bar{F}_l(t_j, t_{j+1}) &= \mathbb{E}\left[\bar{F}_l(t_j, t_{j+1})\right] + \sum_{\nu=1}^{d} \frac{u_\nu}{\sqrt{\lambda_\nu}} [Y_\nu(j) - (u_\nu \sqrt{\lambda_\nu})^T \mathbb{E}[\bar{F}_l(t_j, t_{j+1})]] + \epsilon_j \\
&\triangleq \tilde{m} + \sum_{\nu=1}^{d} \tilde{s}_\nu \times Y_\nu(j) + \epsilon_j
\end{aligned}
$$

- Simple, easy-to-estimate mortality forecasting methodologies
  - Simulate $Y_\nu(N) \sim N(\mu^s_{Y,\nu}, \sigma^s_{Y,\nu})$ with $(\mu^s_{Y,\nu}, \sigma^s_{Y,\nu})$ as sample mean and standard error of $Y_\nu(j)$, $j = 1, \ldots, N-1$
  - Forecast: $\bar{F}_l(t_N, t_{N+1}) = \tilde{m} + \sum_{\nu=1}^{d} \tilde{s}_\nu \times Y_\nu(N)$ together w/ known ${}_\tau p_x(t_N)$

Forward Mortality Factor Models

Self-consistency Condition

- Martingale property of $\left(\exp\left\{-\int_0^t \mu_s(0, x_0+s)\,ds\right\}\ {}_{T-t}p_{x_0+t}(t)\right)_{t\geq 0}$:

$$\mathbb{E}_t\left[\exp\left\{-\int_0^T \mu_s(0, x_0+s)\,ds\right\}\right] = \exp\left\{-\int_0^t \mu_s(0, x_0+s)\,ds\right\}\ {}_{T-t}p_{x_0+t}(t)$$

$\Rightarrow$ Drift condition (Cor. 3.1 in Bauer et al. (2011)):

$$\alpha(\tau, x) = \sigma(\tau, x) \times \int_0^\tau \sigma'(s, x)\,ds$$

- Bauer et al. (2011): μ_t allows for a Gaussian finite-dimensional realization (FDR) iff

$$\sigma(\tau, x) = C(x+\tau) \times \exp\{M\tau\} \times N$$

$$\Rightarrow \mu_t(\tau, x) = \mu_0(\tau+t, x-t) + \int_0^t \alpha(\tau+t-s, x-t+s)\,ds + C(x+\tau)\exp\{M\tau\}\underbrace{\int_0^t \exp\{M(t-s)\}\,N\,dW_s}_{=Z_t\text{(state process)}}$$

- Proposition II: Possible to consider each factor separately

Forward Mortality Factor Models

Analysis of $C(x)$, M, and N

- $$F_l(\tau,x) \stackrel{d}{=} \mathbb{E}[F_l(\tau,x)] + \underbrace{\int_\tau^{\tau+1} C(x+v)\, e^{Mv}\, dv}_{=O(\tau,x)} \times \underbrace{\int_0^\Delta e^{M(\Delta-s)}\, N\, dW_s}_{=Z_\Delta}$$

- $$u_\nu \sqrt{\lambda_\nu} \approx (O_\nu(\tau_i, x_i))_{1\leq i\leq K} \times \exp\{M_\nu(t_{j+1}-t_j)\}N_\nu$$

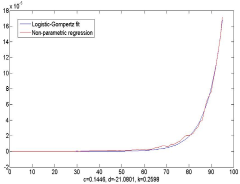
- Estimate $C_\nu(x)$, M_ν, and N_ν via **two-step identification**:

  1. Estimate M and N w/o functional assumptions on $C(x)$
     - ⋆ Rely on examples from interest rate modeling to find convenient shapes, in particular Björk and Gombani (FinStoch, 1999)
  2. Functional assumptions for $C(x)$

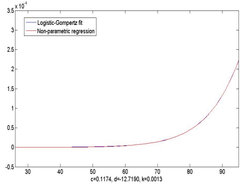
Forward Mortality Factor Models

The Slope Factor: U.S. Data

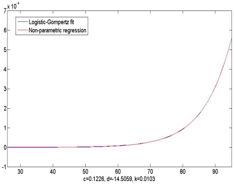
$$\sigma_1(\tau, x) = k \frac{\exp(c(x+\tau)+d)}{(1+\exp(c(x+\tau)+d))}(a+\tau)\exp(-b\tau)$$



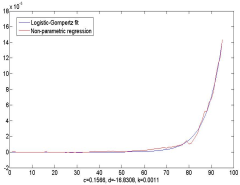
(e) female, Lee-Carter



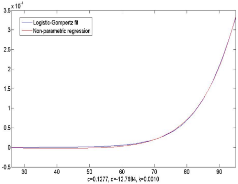
(f) female, CBD



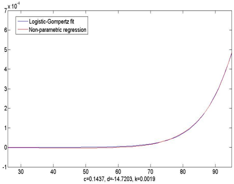
(g) female, *P*-spline



(h) male, Lee-Carter



(i) male, CBD

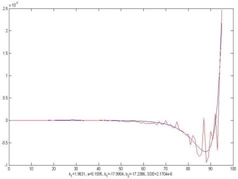


(j) male, *P*-spline

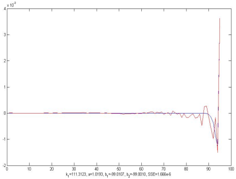
Forward Mortality Factor Models

The Twist Factor: Male/Lee-Carter

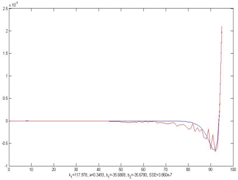
$$\sigma_2(\tau, x) = \left(k_1 \frac{\exp(c_1(x+\tau)+d_1)}{1+\exp(c_1(x+\tau)+d_1)} - k_2 \frac{\exp(c_2(x+\tau)+d_2)}{1+\exp(c_2(x+\tau)+d_2)} \right) \exp\{M_2\tau\} N_2$$



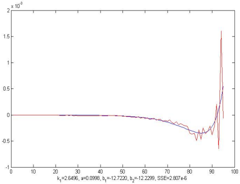
(k) ENW



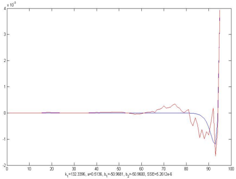
(l) FRA



(m) JPN



(n) USA



(o) FRG

Maximum Likelihood Estimation

From matching $\{C, M, N\}$ to $u\sqrt{\lambda}$ we can estimate parameter values, however

- Only consider the variance part of $\bar{F}(t_j, t_{j+1})$ while neglecting the moment: $\mathbb{E}(\bar{F}(t_j, t_{j+1}))$

$\Rightarrow$ Need to consider the *self-consistency condition*:

$$\alpha(\tau, x) = \sigma(\tau, x) \times \int_0^{\tau} \sigma'(s, x)\, ds$$

- In addition, allow for non-systematic deviations
  - $\bar{F}^{obs}(t_j, t_{j+1}) = \bar{F}^{mod}(t_j, t_{j+1}) + \underline{\epsilon}_j$
  - $\underline{\epsilon}_j \sim N(0, \alpha \cdot \text{diag}\{\Sigma\}), j = 1, \ldots, N-1$
  - α the weight (in the PCA) of all eigenvalues not considered in our model

$\Rightarrow$ $\bar{F}^{obs}(t_j, t_{j+1}) \sim N(\bar{\mu}, \tilde{\Sigma} = \Sigma + \alpha \cdot \text{diag}\{\Sigma\})$

★ Maximum likelihood estimation

Maximum Likelihood Estimation

From matching $\{C, M, N\}$ to $u\sqrt{\lambda}$ we can estimate parameter values, however

- Only consider the variance part of $\bar{F}(t_j, t_{j+1})$ while neglecting the moment: $\mathbb{E}(\bar{F}(t_j, t_{j+1}))$

$\Rightarrow$ Need to consider the *self-consistency condition*:

$$\alpha(\tau, x) = \sigma(\tau, x) \times \int_0^\tau \sigma'(s, x)\, ds$$

- In addition, allow for non-systematic deviations
  - $\bar{F}^{obs}(t_j, t_{j+1}) = \bar{F}^{mod}(t_j, t_{j+1}) + \underline{\epsilon}_j$
  - $\underline{\epsilon}_j \sim N(0, \alpha \cdot \text{diag}\{\Sigma\}), j = 1, \ldots, N-1$
  - α the weight (in the PCA) of all eigenvalues not considered in our model

$\Rightarrow$ $\bar{F}^{obs}(t_j, t_{j+1}) \sim N(\bar{\mu}, \tilde{\Sigma} = \Sigma + \alpha \cdot \text{diag}\{\Sigma\})$

★ Maximum likelihood estimation

Should the Self-consistency Condition be Imposed?

Debate on the necessity of imposing cross-sectional constraints:

- Unlike term structure modeling, increase efficiency of estimates ("$\mathbb{P} = \mathbb{Q}$")
- BUT: Models not satisfying cross-sectional constraints should produce similar forecasts. In particular, imposing cross-sectional constraints should not invalidate estimates in their absence (Duffee (2011))

$\Rightarrow$ Test self-consistency of forecasting approaches

Able to compare estimates for μ_Y and σ_Y in three cases:

1. Directly calculated as the sample mean and standard error from the data set: (μ_Y^s, σ_Y^s)
2. Estimated based on the specific functional assumption on $\sigma(\tau, x)$ but without the self-consistency condition in place: (μ_Y^u, σ_Y^u)
3. Estimated via the MLE from the previous subsection under the specific functional assumption on $\sigma(\tau, x)$ with the self-consistency condition in place: (μ_Y^c, σ_Y^c)

Debate on the necessity of imposing cross-sectional constraints:

- Unlike term structure modeling, increase efficiency of estimates ("$\mathbb{P} = \mathbb{Q}$")
- BUT: Models not satisfying cross-sectional constraints should produce similar forecasts. In particular, imposing cross-sectional constraints should not invalidate estimates in their absence (Duffee (2011))

$\Rightarrow$ Test self-consistency of forecasting approaches

Able to compare estimates for μ_Y and σ_Y in three cases:

1. Directly calculated as the sample mean and standard error from the data set: (μ_Y^s, σ_Y^s)
2. Estimated based on the specific functional assumption on $\sigma(\tau, x)$ but without the self-consistency condition in place: (μ_Y^u, σ_Y^u)
3. Estimated via the MLE from the previous subsection under the specific functional assumption on $\sigma(\tau, x)$ with the self-consistency condition in place: (μ_Y^c, σ_Y^c)

Forward Mortality Factor Models

Should the Self-consistency Condition be Imposed?

Methodology	μ_Y^s	σ_Y^s	μ_Y^u	σ_Y^u	μ_Y^c	σ_Y^c
Lee-Carter	0.0157	0.0235	0.0156	0.0235	0.0170	0.0282
	(0.0047, 0.0266)	(0.0184, 0.0347)				
CBD-Perks	0.0016	0.0204	0.0016	0.0204	0.0043	0.1418
	(−0.0079, 0.0111)	(0.0160, 0.0302)				
P-splines	−0.0249	0.1698	−0.0249	0.1698	−0.0575	0.9968
	(−0.1041, 0.0543)	(0.1331, 0.2512)				

- (μ_Y^u, σ_Y^u) and (μ_Y^s, σ_Y^s) are very close $\rightarrow$ indicates good parametric fit
- (μ_Y^c, σ_Y^c) are only close to the (μ_Y^s, σ_Y^s) for the Lee-Carter method
- For the other forecasting approaches, σ_Y^c considerably exceeds the upper bound of the corresponding confidence interval
- Endorse the use of the Lee-Carter method for producing (deterministic) mortality forecasts

Forward Mortality Factor Models

Should the Self-consistency Condition be Imposed?

Methodology	μ_Y^s	σ_Y^s	μ_Y^u	σ_Y^u	μ_Y^c	σ_Y^c
Lee-Carter	0.0157	0.0235	0.0156	0.0235	0.0170	0.0282
	(0.0047, 0.0266)	(0.0184, 0.0347)				
CBD-Perks	0.0016	0.0204	0.0016	0.0204	0.0043	0.1418
	(−0.0079, 0.0111)	(0.0160, 0.0302)				
P-splines	−0.0249	0.1698	−0.0249	0.1698	−0.0575	0.9968
	(−0.1041, 0.0543)	(0.1331, 0.2512)				

- (μ_Y^u, σ_Y^u) and (μ_Y^s, σ_Y^s) are very close $\rightarrow$ indicates good parametric fit
- (μ_Y^c, σ_Y^c) are only close to the (μ_Y^s, σ_Y^s) for the Lee-Carter method
- For the other forecasting approaches, σ_Y^c considerably exceeds the upper bound of the corresponding confidence interval
- Endorse the use of the Lee-Carter method for producing (deterministic) mortality forecasts

A Non-negative Model Variant

Framework

Based on our specific one-factor assumption, the spot force of mortality is

$$\begin{aligned} \mu_t(0,x) &= \mu_0(t,x-t) + \int_0^t \alpha(t-s,x-t+s)\,ds + f(x) \times Z_t^{(2)} \\ &= \mu_0(t,x-t) + \int_0^t \alpha(t-s,x-t+s)\,ds - \xi_2 f(x) + f(x) \times \underbrace{(Z_t^{(2)} + \xi_2)}_{\tilde{Z}_t^{(2)}} \end{aligned}$$

$$d\tilde{Z}_t = d\begin{pmatrix} Z_t^{(1)} \\ \tilde{Z}_t^{(2)} \end{pmatrix} = \left\{ \begin{pmatrix} 2\xi_1 b + \xi_2 b^2 \\ -\xi_1 \end{pmatrix} + \begin{pmatrix} -2b & -b^2 \\ 1 & 0 \end{pmatrix} \tilde{Z}_t \right\} dt + \begin{pmatrix} 1-ab \\ a \end{pmatrix} dW_t$$

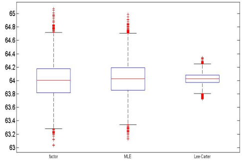
since $\mu_t(0,x)$ only depends on $\tilde{Z}_t^{(2)}$, we thus change $\tilde{Z}_t^{(2)}$ to a square-root process (analogy: Vasicek model to CIR model)

$$d\tilde{Z}_t = \left\{ \begin{pmatrix} 2\xi_1 b + \xi_2 b^2 \\ -\xi_1 \end{pmatrix} + \begin{pmatrix} -2b & -b^2 \\ 1 & 0 \end{pmatrix} \tilde{Z}_t \right\} dt + \begin{pmatrix} (1-ab)\sqrt{\tilde{Z}_t^{(2)}} \\ a\sqrt{\tilde{Z}_t^{(2)}} \end{pmatrix} dW_t$$

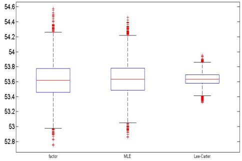
Preserves affine structure ⇒ Calibration w/ generalized Kalman filter

Application

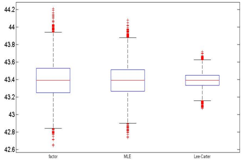
Confidence Intervals for Future LE: USA Female (After 1 yr)



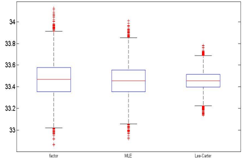
(p) age 20



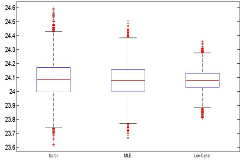
(q) age 30



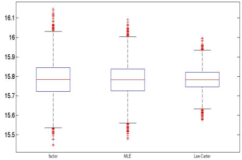
(r) age 40



(s) age 50



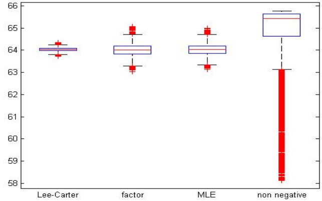
(t) age 60



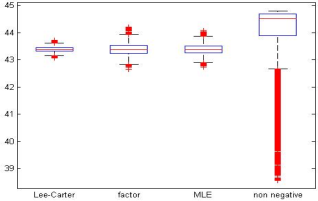
(u) age 70

Application

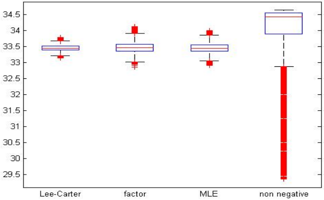
Confidence Intervals for Future LE: USA Female (After 1 yr)



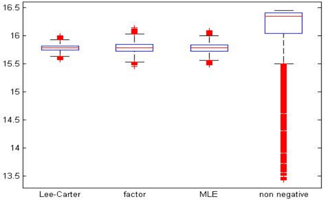
(v) age 20



(w) age 40



(x) age 50



(y) age 70

Conclusion

Conclusion

Conclusion

- Having appropriate estimates for the risk in mortality projections is important
- Common approach may not be suitable to appraise risk within medium or long-term projections
- We provide a parsimonious and tractable alternative

$\Rightarrow$ Mortality Surface Models / Mortality Term Structure Models

- Applications in the life insurance context: "Applications of Forward Mortality Factor Models in Life Insurance Practice", *Geneva Papers*, 2011, **36**: 567-594.

Future Work

- Multiple populations
- Application in household finance: Annuitization decision in portfolio context/influence of systematic mortality risk

Contact

Nan Zhu & Daniel Bauer
nzhu1@student.gsu.edu & dbauer@gsu.edu
Georgia State University

https://sites.google.com/site/nanzhugsu/ & **www.rmi.gsu.edu**

Thank you!