Statistical Inference For Mortality Models

Chen Ling (Patrick)
Risk Management and Insurance, Georgia State University
Southeastern Actuaries Conference 2019

November 22, 2019

Today's talk is based on . . .

- Liu, Q., **Ling, C.**, & Peng, L. (2019). Statistical Inference For Lee-Carter Mortality Model And Corresponding Forecasts. *North American Actuarial Journal*, to appear.
- Liu, Q., **Ling, C.**, Li, D., & Peng, L. (2019). Bias-Corrected Inference For A Modified Lee-Carter Mortality Model. *ASTIN Bulletin*, 49(2), 433-455.
- **Ling, C.**, Liu, Q., Peng, L. & Wu, X. (2019). Statistical Inference For A Modified Two-Population Lee-Carter Mortality Model. *Insurance: Mathematics and Economics*, submitted.

Today's Schedule

Introduction

- Underwriters of annuity products and administrators of pension funds are under financial obligation to their policyholder until the death of counterparty
  - They are subject to longevity risk when a population's average lifespan increases
- Things become even worse when U.S. interest rate is at its historic low level
  - When discount rate is low, present value of life-contingent cash flows can be hard to manage (subject to greatest amount of longevity risk)
  - During periods of low interest rate, large asset management firms have incentives to diversify their investment portfolio and seek alternative investment
  - The market for risk transfer of longevity risks are there (fact: longevity risk is uncorrelated with financial risk)
- This talk is not focused on interest rate risk or maturity mismatch; instead, we are modeling mortality rates quantitatively (so the next step of research would be development of pricing and hedging models)

Research Question

Can we develop a better parametric mortality model?

- Mortality rate (or, death rate): measure of number of deaths in a particular population, scaled to the size of that population

Research Significance

- To price insurance products
- To ensure the solvency through adequate reserves
  - Mortality models are developed for projection of death rates
- Increasing life expectancy of pensioners and policyholders can eventually translate to higher-than-expected pay-out-ratios (longevity risk); needs to be managed
  - Some firms match annuity policies with life insurance policies (natural hedge)
  - Often, finding a perfect match is difficult
- A potential way to manage such risk (pricing, risk transfer, regulation) is parametric mortality model [Li, Li and Balasooriya (2018)]

Research Significance

Research Significance

- As of now, longevity risk transfer market is underdeveloped but promising
- Bottleneck of research: lack of good data; we do theoretical modeling here

Insurance Economics

- Insurance choice of households is closely related to mortality risk
  - Households: maximize expected utility function which is homogeneous with respect to consumption at different times [Ando and Modigliani (1963)]
  - Attempts to explain for people's allocation of assets in annuities during or near retirement [Milevsky (1998)]
- This research is focused on institutions, so individual risk preferences are not considered, and pricing should be based on actuarial fairness

Lee-Carter's Method (1992)

- Lee-Carter model has become a benchmark in modeling mortality rates
- Central death rate: $m(x,t)$ is defined for age group x observed in year t as $\frac{d(x,t)}{L(x,t)}$
  - Number of deaths at age x, $d(x,t)$
  - Average number of living at the age x, $L(x,t) = \int_0^1 l(x+u,t)du$
  - Suppose there are M age groups and T years of mortality data
- Lee and Carter (1992) proposed the following simple linear regression model

$$\log m(x,t) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}, \quad \sum_{x=1}^{M} \beta_x = 1, \quad \sum_{t=1}^{T} k_t = 0 \qquad (1)$$

  - Only central death rate $m(x,t)$ is observed. k_t is unobserved and is called mortality index. $\epsilon_{x,t}$'s are random errors with mean zero and finite variance.

Lee-Carter's Method (1992)

Identification Constraints

Conditions $\sum_{x=1}^{M} \beta_x = 1$ and $\sum_{t=1}^{T} k_t = 0$ are imposed in finding the least squares estimators, which are obtained by the singular value decomposition (SVD) method.

Singular Value Decomposition (SVD)

- $M_{m\times n} = U_{m\times m}\Sigma_{m\times n}V_{n\times n}^T$, where U and V are orthogonal matrices and Σ is a diagonal matrix with nonnegative values on the diagonal. Write $U = (U_1, \cdots, U_m)$ and $V = (V_1, \cdots, V_n)$. Denote the positive values on the diagonal of Σ as $\lambda_1, \cdots, \lambda_r$. Then $M = \sum_{i=1}^{r} \lambda_i U_i V_i^T$
- SVD is used to obtain estimates of α_x, β_x and k_t

Lee-Carter's Method (1992)

Effective Prediction

- An important task of modeling mortality rates is to forecast future mortality pattern
  - Hence, to better forecast mortality risk and hedge longevity risk, Lee and Carter (1992) further proposed to model the estimated mortality index by a simple time series model
  - Assumed that $\{k_t\}$ follows from an ARIMA(p,d,q) model defined as

$$\left(1-\sum_{i=1}^{p}\phi_i B^i\right)(1-B)^d k_t = \mu + \left(1+\sum_{i=1}^{q}\theta_i B^i\right)e_t, \tag{2}$$

  - where e_t's are white noises

Lee-Carter's Method (1992)

Summary of LC Model

- The popular Lee-Carter mortality model is a combination of (1) and (2), and a proposed two-step inference procedure is to first estimate parameters in (1) by the singular value decomposition method and then to use the estimated k_t's to fit model (2).
- Many papers in actuarial science have claimed that an application of this model and this two-step inference procedure to mortality data prefers a unit root time series model, i.e., $d = 1$ in (2).

Extensions and Applications

- Since the seminal paper, many extensions and applications have appeared in the literature; also appeared is an open **R** package ('demography')
- For example,

Lee (2000)

Before proceeding directly to modeling the parameter $\hat{k}_t$ as a time series process, the $\hat{k}_t's$ are adjusted (taking $\hat{\alpha}_x$ and $\hat{\beta}_x$ estimates as given) to reproduce the observed number of deaths $\sum_x D_{xt}$, i.e., the $\tilde{k}_t's$ solve

$$\sum_x D_{xt} = \sum_x E_{xt} \exp\{\hat{\alpha}_x + \hat{\beta}_x \tilde{k}_t\},$$

where E_{xt} is the actual risk exposure. So for better prediction, a time series model is fitted to $\tilde{k}_t's$.

Extensions and Applications

Brouhns, Denuit and Vermunt (2002)

First model

$$\begin{cases} D_{xt} = Poisson(E_{xt}\mu_x(t)) \\ \mu_x(t) = \exp\{\alpha_x + \beta_x k_t\} \end{cases}$$

Use the same constraints in Lee and Carter (1992) to estimate α_x, β_x, k_t by maximizing

$$\sum_{x,t}\{D_{xt}(\alpha_x + \beta_x k_t) - E_{xt}\exp(\alpha_x + \beta_x k_t)\}.$$

Denote the obtained estimators by $\hat{\alpha}_x, \hat{\beta}_x, \hat{k}_t$. Next step is to fit a time series to the mortality index $\hat{k}_t's$ as Lee and Carter (1992).

Extensions and Applications

Li and Lee (2005)

Extended the Lee-Carter model to model a group of population $m(x, t, i)$, where x, t, i denote the age group, time and ith population, respectively.

Girosi and King (2007)

Cited more than a dozen of papers to confirm the wide implementation of the Lee-Carter model by policy analysts around the world.

Extensions and Applications

Cairns et al. (2011)

Compared six different stochastic mortality models. The functions $\beta_x^{(i)}$, $k_t^{(i)}$, and $\gamma_{t-x}^{(i)}$ are age, period and cohort effects, respectively; $\bar{x}$ is the mean age over the range of ages being used in the analysis; $\hat{\sigma}_x^2$ is the mean value of $(x-\bar{x})^2$; n_a is the number of ages.

- $\log m(t,x) = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)}$;
- $\log m(t,x) = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)} + \beta_x^{(3)} \gamma_{t-x}^{(3)}$;
- $\log m(t,x) = \beta_x^{(1)} + n_a^{-1} k_t^{(2)} + n_a^{-1} \gamma_{t-x}^{(3)}$;
- $\mathrm{logit}\, q(t,x) = k_t^{(1)} + k_t^{(2)}(x-\bar{x})$;
- $\mathrm{logit}\, q(t,x) = k_t^{(1)} + k_t^{(2)}(x-\bar{x}) + k_t^{(3)}((x-\bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}^{4}$;
- $\mathrm{logit}\, q(t,x) = k_t^{(1)} + k_t^{(2)}(x-\bar{x}) + \gamma_{t-x}^{(3)}(x_c - x)$.

Extensions and Applications

D'amato et al (2014)

Employed the Lee-Carter model to detect common longevity trends. The model is

$$\log m_{xt,i} = \alpha_{x,i} + \beta_{x,i} k_{t,i} + \epsilon_{xt,i},$$

where i denotes the ith population.

Lin et al (2014)

Employed the extended Lee-Carter model in Li and Lee (2005) to study the risk management of a defined benefit plan.

Bisetti and Favero (2014)

Applied the Lee-Carter model to measure the impact of longevity risk on pension systems in Italy.

Extensions and Applications

Bootstrap Methods for Quantifying Uncertainty in Mortality Models

- For interval estimation and/or projection errors, bootstrap methods have been proposed
  - Haberman and Renshaw (2009) proposed three different bootstrap methods to construct confidence intervals for interesting quantities based on the Lee-Carter framework and a generalized linear Poisson model. Li (2010) used parametric bootstrap. D'Amato et al (2012) proposed sieve bootstrap method based on estimated errors in $\log m_{xt} = \alpha_x + \beta_x k_t + \epsilon_{xt}$, where ϵ_{xt} follows from an $AR(\infty)$ model.

Continuous Stochastic Differential Equations (SDE) for Modeling Mortality

Dahl (2004) selected an extended Cox-Ingersoll-Ross process; Biffis (2005) chose two different specifications for the intensity process; Schrager (2006) proposes an M-factor affine stochastic intensity; Elisa, Spreeuw and Vigna (2008) modeled stochastic mortality for dependent lives.

Our Discussion of Model Issues

Issues with Model Assumption

- Constraint on mortality index $\sum_{t=1}^{T} k_t = 0$
  - Since $\{k_t\}$ in (2) is random, the constraint $\sum_{t=1}^{T} k_t = 0$ in (1) becomes unrealistic and restrictive
  - For example, if one fits an AR(1) model to $\{k_t\}$, say $k_t = \mu + \phi k_{t-1} + e_t$, then we have $T^{-1} \sum_{t=1}^{T} k_t \xrightarrow{p} \mu/(1-\phi)$ as $T \to \infty$ when $|\phi| < 1$ independent of T
  - On the other hand, we have $k_T/T \xrightarrow{p} \mu\{\lim_{x \to \gamma} \frac{1-e^x}{-x}\}$ as $T \to \infty$ when $\mu \neq 0$ and $\phi = 1 + \gamma/T$ for some constant $\gamma \in \mathbb{R}$
- No trend term in time series; constraint too restrictive
  - Constraint $\sum_{t=1}^{T} k_t = 0$ in (1) basically says μ in (2) must be zero
- Hence, a modified model without any direct constraint on k_t's is more appropriate

Our Discussion of Model Issues

Issues with Inference Method

- Singular value decomposition
  - There is no way to quantify the inference uncertainty
  - When all β_x are the same (i.e., $\beta_1 = \cdots = \beta_M = 1/M$), Leng and Peng (2016) proved that the proposed two-step inference procedure in Lee and Carter (1992) is inconsistent when the model (2) is not an exact ARIMA(0,1,0) model.
- No asymptotic results available for the derived estimators

Misunderstandings of Lee-Carter's Method

- Some recent literature in actuarial science interpret the LC model in a wrong way
- Ignored the error terms $\epsilon_{x,t}$ in the LC model
  - For example, by defining $m_0(x,t)$ as the true central death rate for age x in year t, Dowd et al. (2010), Cairns et al (2011), Enchev, Kleinow and Cairns (2017) and others interpreted the Lee-Carter model as

$$\log m_0(x,t) = \alpha_x + \beta_x k_t, \ \ k_t = \mu + k_{t-1} + e_t, \ \sum_{x=1}^{M} \beta_x = 1, \ \sum_{t=1}^{T} k_t = 0. \tag{3}$$

  - Model (3) basically says the true mortality rate $m_0(x,t)$ is random due to the randomness of k_t's
  - Another misinterpretation appears in Li (2010) and Li, Chan and Zhou (2015), where the Lee-Carter model is treated as $\log m(x,t) = \alpha_x + \beta_x k_t$ without the random error $\varepsilon_{x,t}$ in (1)
  - This is quite problematic because it simply says that $\log m(x,t)$ and $\log m(y,t)$ are completely dependent as both are determined by the same random variable k_t.
  - That is, central death rates are completely dependent across ages.

Our Discussion of Model Issues

Lessons Learned From Literature

- The random error term $\varepsilon_{x,t}$ in (1) is necessary in order to avoid the unrealistic implication that the central death rates are completely dependent across ages.
- Due to the presence of these random errors $\varepsilon_{x,t}$'s, the two-step inference procedure proposed by Lee and Carter (1992) may be inconsistent.
- Hence it is questionable for the existing claim on unit root mortality index and the use of bootstrap methods to quantify the forecast uncertainty of future mortality rates.

Proposed Model

First we propose to replace (1) by

$$\log m(x,t) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}, \quad \sum_{x=1}^{M} \beta_x = 1, \quad \sum_{x=1}^{M} \alpha_x = 0, \tag{4}$$

$$k_t = \mu + \phi k_{t-1} + e_t, \tag{5}$$

where e_t's and $\varepsilon_{x,t}$'s are random errors with zero mean and finite variance for each x. It is clear that we do not directly impose a constraint on the unobserved random mortality index k_t to ensure that the proposed model is identifiable. We also remark that the assumption of $\sum_{x=1}^{M} \alpha_x = 0$ is not restrictive at all as we can simply move the sum to k_t if $\sum_{x=1}^{M} \alpha_x \neq 0$.

Estimation

We need a different statistical inference without using the singular value decomposition method.
Put $Z_t = \sum_{x=1}^{M} \log m(x,t)$ and $\eta_t = \sum_{x=1}^{M} \varepsilon_{x,t}$ for $t = 1, \cdots, T$. Then, by noting that $\sum_{x=1}^{M} \alpha_x = 0$ and $\sum_{x=1}^{M} \beta_x = 1$, we have

$$Z_t = k_t + \eta_t \quad \text{for} \quad t = 1, \cdots, T. \tag{6}$$

When $\{k_t\}$ is nonstationary such as unit root (i.e., $\phi = 1$ in (5)) or near unit root (i.e., $\phi = 1 + \gamma/T$ for some constant $\gamma \neq 0$ in (5)), k_t dominates η_t as t large enough, and so Z_t behaves like k_t in this case.

Estimation

This motivates us to minimize the following least squares

$$\sum_{t=2}^{T} (Z_t - \mu - \phi Z_{t-1})^2 ,$$

and obtain the least squares estimator for μ and ϕ as

$$\begin{cases} \hat{\mu} = \dfrac{\sum_{s=2}^{T} Z_s \sum_{t=2}^{T} Z_{t-1}^2 - \sum_{s=2}^{T} Z_{s-1} \sum_{t=2}^{T} Z_t Z_{t-1}}{(T-1) \sum_{t=2}^{T} Z_{t-1}^2 - (\sum_{t=2}^{T} Z_{t-1})^2}, \\ \hat{\phi} = \dfrac{(T-1) \sum_{t=2}^{T} Z_t Z_{t-1} - \sum_{s=2}^{T} Z_s \sum_{t=2}^{T} Z_{t-1}}{(T-1) \sum_{t=2}^{T} Z_{t-1}^2 - (\sum_{t=2}^{T} Z_{t-1})^2}. \end{cases}$$

Estimation

Similarly, by minimizing the following least squares

$$\sum_{t=1}^{T} (\log m(x,t) - \alpha_x - \beta_x Z_t)^2 ,$$

we obtain the least squares estimator for α_x and β_x as

$$\begin{cases} \hat{\alpha}_x = \frac{\sum_{s=1}^{T} \log m(x,s) \sum_{t=1}^{T} Z_t^2 - \sum_{s=1}^{T} \log m(x,s) Z_s \sum_{t=1}^{T} Z_t}{T \sum_{t=1}^{T} Z_t^2 - (\sum_{t=1}^{T} Z_t)^2}, \\ \hat{\beta}_x = \frac{T \sum_{s=1}^{T} \log m(x,s) Z_s - \sum_{s=1}^{T} \log m(x,s) \sum_{t=1}^{T} Z_t}{T \sum_{t=1}^{T} Z_t^2 - (\sum_{t=1}^{T} Z_t)^2}. \end{cases}$$

Bias Correction

The above estimation solves the following score equations for $x = 1, \ldots, M$:

$$\begin{cases} \sum_{t=2}^{T}\{Z_t - \mu - \phi Z_{t-1}\} = 0, \\ \sum_{t=2}^{T}\{Z_t - \mu - \phi Z_{t-1}\}Z_{t-1} = 0, \\ \sum_{t=2}^{T}\{\log m(x,t) - \alpha_x - \beta_x Z_t\} = 0, \\ \sum_{t=2}^{T}\{\log m(x,t) - \alpha_x - \beta_x Z_t\}Z_t = 0. \end{cases} \tag{7}$$

Bias Correction

By noting that the inconsistency of the least squares estimators via solving (7) is due to the correlation between $Z_t - \mu_0 - \phi_0 Z_{t-1} = e_t + \eta_t - \phi_0 \eta_{t-1}$ and $Z_{t-1} = k_{t-1} + \eta_{t-1}$, we propose the simple bias corrected estimators via solving the following modified score equations:

$$\begin{cases} \sum_{t=3}^{T} \{Z_t - \mu - \phi Z_{t-1}\} = 0, \\ \sum_{t=3}^{T} \{Z_t - \mu - \phi Z_{t-1}\} Z_{t-2} = 0, \\ \sum_{t=3}^{T} \{\log m(x,t) - \alpha_x - \beta_x Z_t\} = 0, \\ \sum_{t=3}^{T} \{\log m(x,t) - \alpha_x - \beta_x Z_t\} Z_{t-1} = 0, \end{cases} \tag{8}$$

Bias Correction

which give

$$\hat{\mu} = \frac{\sum_{s=3}^{T} Z_s \sum_{t=3}^{T} Z_{t-1}Z_{t-2} - \sum_{s=3}^{T} Z_{s-1} \sum_{t=3}^{T} Z_t Z_{t-2}}{(T-2)\sum_{t=3}^{T} Z_{t-1}Z_{t-2} - \sum_{s=3}^{T} Z_{s-1} \sum_{t=3}^{T} Z_{t-2}},$$

$$\hat{\phi} = \frac{(T-2)\sum_{t=3}^{T} Z_t Z_{t-2} - \sum_{s=3}^{T} Z_s \sum_{t=3}^{T} Z_{t-2}}{(T-2)\sum_{t=3}^{T} Z_{t-1}Z_{t-2} - \sum_{s=3}^{T} Z_{s-1} \sum_{t=3}^{T} Z_{t-2}},$$

$$\hat{\alpha}_x = \frac{\sum_{s=3}^{T} \log m(x,s) \sum_{t=3}^{T} Z_t Z_{t-1} - \sum_{s=3}^{T} \log m(x,s) Z_{s-1} \sum_{t=3}^{T} Z_t}{(T-2)\sum_{t=3}^{T} Z_t Z_{t-1} - \sum_{s=3}^{T} Z_s \sum_{t=3}^{T} Z_{t-1}},$$

$$\hat{\beta}_x = \frac{(T-2)\sum_{t=3}^{T} \log m(x,t) Z_{t-1} - \sum_{s=3}^{T} \log m(x,s) \sum_{t=3}^{T} Z_{t-1}}{(T-2)\sum_{t=3}^{T} Z_t Z_{t-1} - \sum_{s=3}^{T} Z_s \sum_{t=3}^{T} Z_{t-1}}$$

for $x = 1, \ldots, M$. We remark that the estimator $\hat{\phi}$ is the same as the modified Yule-Walker estimator in Staudenmayer and Buonaccorsi (2005) for a time series model with measurement errors, and obviously we have $\sum_{x=1}^{M} \hat{\alpha}_x = 0$ and $\sum_{x=1}^{M} \hat{\beta}_x = 1$.

Regularity Conditions

In order to derive the asymptotic properties of the above proposed estimators, we assume the following regularity conditions for the stationary errors in (4) and (5).

- C1) $E(e_t) = 0$, $E(\varepsilon_{x,t}) = 0$ for $t = 1, \cdots, T$ and $x = 1, \cdots, M$;
- C2) there exist $\beta > 2$ and $\delta > 0$ such that $\sup_t E|e_t|^{\beta+\delta} < \infty$, $\sup_t E|\varepsilon_{x,t}|^{\beta+\delta} < \infty$ for $x = 1, \cdots, M$;
- C3) $\sigma_e^2 = \lim_{T\to\infty} E\{T^{-1}(\sum_{t=1}^T e_t)^2\} \in (0, \infty)$ and $\sigma_x^2 = \lim_{T\to\infty} E\{T^{-1}(\sum_{t=1}^T \varepsilon_{x,T})^2\} \in (0, \infty)$ for $x = 1, \cdots, M$;
- C4) sequence $\{(e_t, \varepsilon_{1,t}, \cdots, \varepsilon_{M,t})^\tau\}$ is strong mixing with mixing coefficients

$$\alpha_m = \sup_{k\geq 1} \sup_{A\in\mathcal{F}_1^k, B\in\mathcal{F}_{k+m}^\infty} |P(A\cap B) - P(A)P(B)|$$

such that $\sum_{m=1}^\infty \alpha_m^{1-2/\beta} < \infty$, where $\mathcal{F}_k^{k+m}$ denotes the σ-field generated by $\{(e_t, \varepsilon_{1,t}, \cdots, \varepsilon_{M,t})^\tau : k \leq t \leq k+m\}$ and A^τ denotes the transpose of the matrix or vector A.

Asymptotic Results

To present the asymptotic distribution of the proposed bias-corrected estimators, we need some notations. Put $\boldsymbol{\theta} = (\mu, \phi, \alpha_1, \beta_1, \cdots, \alpha_{M-1}, \beta_{M-1})^\tau$, $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\phi}, \hat{\alpha}_1, \hat{\beta}_1, \cdots, \hat{\alpha}_{M-1}, \hat{\beta}_{M-1})^\tau$ and let $\boldsymbol{\theta}_0 = (\mu_0, \phi_0, \alpha_{1,0}, \beta_{1,0}, \cdots, \alpha_{M-1,0}, \beta_{M-1,0})^\tau$ denote the true value of $\boldsymbol{\theta}$. Note that we exclude α_M and β_M in the above definitions due to the constraints $\sum_{x=1}^M \alpha_{x,0} = 0$ and $\sum_{x=1}^M \beta_{x,0} = 1$.

Asymptotic Distribution

Theorem

Assume $\mu_0 \neq 0$, *and* $\{(e_t, \varepsilon_{1,t}, \cdots, \varepsilon_{M,t})^\tau : t = 1, \ldots, T\}$ *is a sequence of independent and identically distributed random vectors with means zero and finite covariance matrix.*

i) When $|\phi_0| < 1$ *independent of* T *(i.e., stationary case), we have*

$$\sqrt{T}\Gamma\{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\} \xrightarrow{d} N(0, \Sigma).$$

ii) When $\phi_0 = 1 + \rho/T$ *for some constant* $\rho \in \mathbb{R}$ *(i.e., near unit root if* $\rho \neq 0$ *and unit root if* $\rho = 0$*), we have*

$$D_T\{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\} \xrightarrow{d} N(0, \widetilde{\Gamma}^{-1}\widetilde{\Sigma}\widetilde{\Gamma}^{-1}),$$

where D_T *is a diagonal matrix with* $T^{1/2}$ *in the odd diagonal elements and* $T^{3/2}$ *in the even diagonal element.*

Extension — Bivariate Lee-Carter Mortality Model

Bivariate mortality model

$$\log m^{(1)}(x,t) = \alpha_x^{(1)} + \beta_x^{(1)} k_t^{(1)} + \varepsilon_{x,t}^{(1)}, \tag{9}$$

$$\log m^{(2)}(x,t) = \alpha_x^{(2)} + \beta_x^{(2)} k_t^{(2)} + \varepsilon_{x,t}^{(2)}, \tag{10}$$

$$k_t^{(1)} = \mu^{(1)} + \phi^{(1)} k_{t-1}^{(1)} + e_t^{(1)}, \tag{11}$$

$$k_t^{(1)} - k_t^{(2)} = \mu^{(2)} + \phi^{(2)}(k_{t-1}^{(1)} - k_{t-1}^{(2)}) + e_t^{(2)}, \tag{12}$$

where $\{(\varepsilon_{x,t}^{(1)}, \varepsilon_{x,t}^{(2)})^\tau\}_{t=1}^T$ is a sequence of independent and identically distributed random vectors with zero means and finite variances for each x, $\{(e_t^{(1)}, e_t^{(2)})^\tau\}_{t=1}^T$ is a sequence of independent and identically distributed random vectors with zero means and finite variances.

Extension — Bivariate Lee-Carter Mortality Model

Still, we avoid adding constraints to the random mortality indexes by assuming that

$$\sum_{x=1}^{M} \alpha_x^{(i)} = 0, \quad \sum_{x=1}^{M} \beta_x^{(i)} = 1, \quad i = 1, 2, \tag{13}$$

Define $\eta_t^{(i)} = \sum_{x=1}^{M} \epsilon_{x,t}^{(i)}$ and $Z_t^{(i)} = \sum_{x=1}^{M} \log m^{(i)}(x, t)$, then we have

$$Z_t^{(i)} = k_t^{(i)} + \eta_t^{(i)}, \quad i = 1, 2.$$

When both $\{k_t^{(1)}\}$ and $\{k_t^{(1)} - k_t^{(2)}\}$ are near unit root or unit root, as t large enough, $Z_t^{(1)}$ and $k_t^{(1)}$ are approximately the same, and $Z_t^{(1)} - Z_t^{(2)}$ and $k_t^{(1)} - k_t^{(2)}$ are approximately the same.

Estimation

Therefore, we could employ the least squares estimators via solving the following score functions:

$$\begin{cases} \sum_{t=2}^{T}\{Z_t^{(1)} - \mu^{(1)} - \phi^{(1)} Z_{t-1}^{(1)}\} = 0 \\ \sum_{t=2}^{T}\{Z_t^{(1)} - \mu^{(1)} - \phi^{(1)} Z_{t-1}^{(1)}\} Z_{t-1}^{(1)} = 0 \\ \sum_{t=2}^{T}\{Z_t^{(1)} - Z_t^{(2)} - \mu^{(2)} - \phi^{(2)}(Z_{t-1}^{(1)} - Z_{t-1}^{(2)})\} = 0 \\ \sum_{t=2}^{T}\{Z_t^{(1)} - Z_t^{(2)} - \mu^{(2)} - \phi^{(2)}(Z_{t-1}^{(1)} - Z_{t-1}^{(2)})\}(Z_{t-1}^{(1)} - Z_{t-1}^{(2)}) = 0 \end{cases} \tag{14}$$

and

$$\begin{cases} \sum_{t=1}^{T}\{\log m^{(i)}(x,t) - \alpha_x^{(i)} - \beta_x^{(i)} Z_t^{(i)}\} = 0 \\ \sum_{t=1}^{T}\{\log m^{(i)}(x,t) - \alpha_x^{(i)} - \beta_x^{(i)} Z_t^{(i)}\} Z_t^{(i)} = 0 \end{cases} \tag{15}$$

Estimation

By noting that the inconsistency of the least squares estimators via solving (14) is due to the correlation between $Z_t^{(1)} - \mu_0^{(1)} - \phi_0^{(1)} Z_{t-1}^{(1)} = e_t^{(1)} + \eta_t^{(1)} - \phi_0^{(1)} \eta_{t-1}^{(1)}$ and $Z_{t-1}^{(1)} = k_{t-1}^{(1)} + \eta_{t-1}^{(1)}$, we follow the idea of shifting a lag and propose the unified bias corrected estimators via solving the following modified score equations:

$$\begin{cases} \sum_{t=3}^{T}\{Z_t^{(1)} - \mu^{(1)} - \phi^{(1)} Z_{t-1}^{(1)}\} = 0 \\ \sum_{t=3}^{T}\{Z_t^{(1)} - \mu^{(1)} - \phi^{(1)} Z_{t-1}^{(1)}\} Z_{t-2}^{(1)} = 0 \\ \sum_{t=3}^{T}\{Z_t^{(1)} - Z_t^{(2)} - \mu^{(2)} - \phi^{(2)}(Z_{t-1}^{(1)} - Z_{t-1}^{(2)})\} = 0 \\ \sum_{t=3}^{T}\{Z_t^{(1)} - Z_t^{(2)} - \mu^{(2)} - \phi^{(2)}(Z_{t-1}^{(1)} - Z_{t-1}^{(2)})\}(Z_{t-2}^{(1)} - Z_{t-2}^{(2)}) = 0 \end{cases} \tag{16}$$

and

$$\begin{cases} \sum_{t=2}^{T}\{\log m^{(i)}(x,t) - \alpha_x^{(i)} - \beta_x^{(i)} Z_t^{(i)}\} = 0 \\ \sum_{t=2}^{T}\{\log m^{(i)}(x,t) - \alpha_x^{(i)} - \beta_x^{(i)} Z_t^{(i)}\} Z_{t-1}^{(i)} = 0 \end{cases} \tag{17}$$

Asymptotic Results

To present the asymptotic distribution of the proposed bias-corrected estimators, we need some notations.
Put

$$\boldsymbol{\theta} = (\mu^{(1)}, \phi^{(1)}, \mu^{(2)}, \phi^{(2)}, \alpha_1^{(1)}, \beta_1^{(1)}, \alpha_1^{(2)}, \beta_1^{(2)}, \cdots, \alpha_{M-1}^{(1)}, \beta_{M-1}^{(1)}, \alpha_{M-1}^{(2)}, \beta_{M-1}^{(2)})^\tau,$$

and let

$$\hat{\boldsymbol{\theta}} = (\hat{\mu}^{(1)}, \hat{\phi}^{(1)}, \hat{\mu}^{(2)}, \hat{\phi}^{(2)}, \hat{\alpha}_1^{(1)}, \hat{\beta}_1^{(1)}, \hat{\alpha}_1^{(2)}, \hat{\beta}_1^{(2)}, \cdots, \hat{\alpha}_{M-1}^{(1)}, \hat{\beta}_{M-1}^{(1)}, \hat{\alpha}_{M-1}^{(2)}, \hat{\beta}_{M-1}^{(2)})^\tau$$

and

$$\boldsymbol{\theta}_0 = (\mu_0^{(1)}, \phi_0^{(1)}, \mu_0^{(2)}, \phi_0^{(2)}, \alpha_{1,0}^{(1)}, \beta_{1,0}^{(1)}, \alpha_{1,0}^{(2)}, \beta_{1,0}^{(2)}, \cdots, \alpha_{M-1,0}^{(1)}, \beta_{M-1,0}^{(1)}, \alpha_{M-1,0}^{(2)}, \beta_{M-1,0}^{(2)})^\tau$$

denote the above bias-corrected estimators and the true value of $\boldsymbol{\theta}$, respectively. Note that we exclude $\alpha_M^{(i)}$ and $\beta_M^{(i)}$ in the above definitions due to the constraints $\sum_{x=1}^M \alpha_{x,0}^{(i)} = 0$ and $\sum_{x=1}^M \beta_{x,0}^{(i)} = 1$ for $i = 1, 2$.

Asymptotic Results — Unit Root Case

Theorem

Assume models (9) to (13). Further assume $\mu_0^{(1)} \neq 0$, $\mu_0^{(2)} \neq 0$, $\phi_0^{(1)} = 1 + \frac{\rho_1}{T}$ *and* $\phi_0^{(2)} = 1 + \frac{\rho_2}{T}$ *for some* $\rho_1, \rho_2 \in \mathbb{R}$. *Then*

$$D_T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Gamma}_1^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Gamma}_1^{-1}) \text{ as } T \to \infty,$$

where $\boldsymbol{\Sigma}_1$ *and* $\boldsymbol{\Gamma}_1$ *are respectively defined in the appendix, and*

$$D_T = diag(\sqrt{T}, T^{3/2}, \cdots, \sqrt{T}, T^{3/2}).$$

Asymptotic Results — Stationary Case

Theorem

Suppose models (9) and (11) hold with (13). Further assume $\mu_0^{(1)} \neq 0$, $\mu_0^{(2)} \neq 0$, $|\phi_0^{(1)}| < 1$ and $|\phi_0^{(2)}| < 1$. Then

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Gamma}_2^{-1}\boldsymbol{\Sigma}_2\boldsymbol{\Gamma}_2^{-1}) \text{ as } T \to \infty,$$

where $\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Gamma}_2$ are respectively defined in the appendix.

We also have asymptotic results for mixed cases (one of the two series $\{k_t^{(1)}\}$ and $\{k_t^{(1)} - k_t^{(2)}\}$ is near unit root or unit root; the other is stationary).

Data Analysis

We investigate the US mortality data of male and female cohorts from year 1933 to year 2017, which are available from the Human Mortality Database (https://www.mortality.org).
We fit the Lee-Carter model based on singular value decomposition, as well as the proposed method in this paper. (For comparison purposes, we fit the proposed least squares estimate in this paper without and with bias correction.)

Data Analysis — Lee-Carter Estimates

Figure: Estimates of Bivariate Model Parameters Based on U.S. Cohorts

Lee-Carter Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-6.265	-6.128	-5.850	-5.482	-5.060	-4.624	-4.215	-3.810	-3.429	-3.028
$\beta_x^{(1)}$	0.085	0.091	0.104	0.109	0.108	0.109	0.104	0.101	0.098	0.093
$\alpha_x^{(2)}$	-7.018	-6.741	-6.385	-5.994	-5.582	-5.169	-4.780	-4.359	-3.941	-3.477
$\beta_x^{(2)}$	0.134	0.126	0.119	0.106	0.096	0.091	0.083	0.081	0.082	0.084

Lee-Carter Estimates of $\mu^{(i)}$ and $\phi^{(i)}$

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.113	0.994	0.049	0.958

Data Analysis — Without Bias Correction

Figure: Estimates of Bivariate Model Parameters Based on U.S. Cohorts

Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$ by the Proposed Model (Without Bias Correction)

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-2.187	-1.755	-0.877	-0.276	0.129	0.569	0.741	0.994	1.245	1.415
$\beta_x^{(1)}$	0.085	0.091	0.104	0.109	0.108	0.108	0.103	0.100	0.098	0.093
$\alpha_x^{(2)}$	0.094	-0.030	-0.067	-0.344	-0.459	-0.317	-0.318	-0.020	0.444	1.017
$\beta_x^{(2)}$	0.133	0.126	0.118	0.106	0.096	0.091	0.083	0.081	0.082	0.084

Estimates of $\mu^{(i)}$ and $\phi^{(i)}$ by the Proposed Model (Without Bias Correction)

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.622	0.989	0.274	0.958

Data Analysis — With Bias Correction

Figure: Estimates of Bivariate Model Parameters Based on U.S. Cohorts

Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$ by the Proposed Model (With Bias Correction)

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-2.294	-1.867	-0.954	-0.289	0.139	0.610	0.809	1.053	1.324	1.469
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.101	0.099	0.094
$\alpha_x^{(2)}$	-0.012	-0.105	-0.106	-0.330	-0.447	-0.285	-0.278	7.739e-3	0.488	1.068
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085

Estimates of $\mu^{(i)}$ and $\phi^{(i)}$ by the Proposed Model (With Bias Correction)

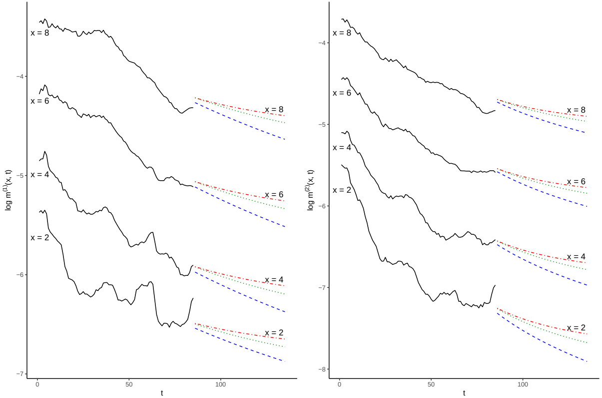
$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.856	0.985	0.285	0.956

Data Analysis

- It would be cautious to use the Lee-Carter inference and the proposed least squares estimation without bias correction.
    - The estimate for $\phi^{(1)}$ based on the Lee-Carter inference is much closer to one (i.e., unit root) than the proposed least squares estimation, and the estimates for $\mu^{(i)}$ and $\phi^{(i)}$ are different for the proposed least squares estimation with or without bias correction. As estimates for $\phi^{(2)}$ based on these three inferences are 'significantly' smaller than one, it suggests that $\{k_t^{(1)} - k_t^{(2)}\}$ is a stationary sequence.

Out-of-Sample Forecast

Figure: Solid lines represent true historic values; dashed lines represent forecasts by Lee-Carter model; dotted lines represent forecasts according to the proposed least squares estimators without bias correction; dash-dotted lines are forecasts by proposed bias corrected estimators



Simulation Study — Setup

- Simulation setting:
  - We simulate 10,000 random samples with sample size $T = 80$, 150 and 300 with parameters being the proposed bias-corrected estimates in fitting the real dataset above except $\phi^{(1)} = 1$ and $\phi^{(2)} = 0.95$. We choose $\varepsilon_{x,t}$ and e_t to be independent and identically distributed random variables with $N(0, 0.1^2)$.
- Many papers in the study of longevity risk simply assume $\phi^{(1)} = 1$
  - This section uses simulated data to show that the Lee-Carter inference and the proposed least squares estimation without bias correction lead to an inconsistent inference when $\phi^{(1)} = 1$ and $|\phi^{(2)}| < 1$, but the proposed bias-corrected inference performs well.

Simulation Study

Figure: Estimates with standard deviations in brackets for $T = 80$

Lee-Carter Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-5.171	-4.955	-4.500	-4.049	-3.625	-3.178	-2.827	-2.466	-2.115	-1.786
	(0.044)	(0.047)	(0.053)	(0.057)	(0.057)	(0.057)	(0.055)	(0.053)	(0.052)	(0.049)
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.102	0.099	0.094
	(5.472e-4)	(5.351e-4)	(5.367e-4)	(5.353e-4)	(5.388e-4)	(5.274e-4)	(5.418e-4)	(5.379e-4)	(5.384e-4)	(5.337e-4)
$\alpha_x^{(2)}$	-5.131	-4.953	-4.693	-4.466	-4.198	-3.852	-3.565	-3.181	-2.746	-2.251
	(0.072)	(0.068)	(0.065)	(0.059)	(0.054)	(0.051)	(0.047)	(0.046)	(0.047)	(0.048)
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085
	(5.060e-4)	(5.097e-4)	(5.089e-4)	(5.047e-4)	(5.037e-4)	(5.055e-4)	(5.030e-4)	(5.022e-4)	(5.068e-4)	(5.082e-4)

Lee-Carter Estimates of $\mu^{(i)}$ and $\phi^{(i)}$

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.847	1.001	0.070	0.941
(0.012)	(6.724e-4)	(7.744e-3)	(0.013)

Simulation Study

Figure: Estimates with standard deviations in brackets for $T = 80$

Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$ by the Proposed Model (Without Bias Correction)

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-2.294	-1.867	-0.954	-0.289	0.138	0.610	0.808	1.053	1.324	1.470
	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)	(0.022)	(0.021)
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.102	0.099	0.094
	(5.470e-4)	(5.350e-4)	(5.366e-4)	(5.352e-4)	(5.386e-4)	(5.272e-4)	(5.417e-4)	(5.378e-4)	(5.382e-4)	(5.335e-4)
$\alpha_x^{(2)}$	-0.012	-0.106	-0.106	-0.330	-0.447	-0.286	-0.278	8.205e-3	0.489	1.068
	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085
	(5.059e-4)	(5.096e-4)	(5.088e-4)	(5.046e-4)	(5.036e-4)	(5.054e-4)	(5.029e-4)	(5.021e-4)	(5.067e-4)	(5.081e-4)

Estimates of $\mu^{(i)}$ and $\phi^{(i)}$ by the Proposed Model (Without Bias Correction)

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.866	1.000	0.669	0.861
(0.029)	(7.541e-4)	(0.119)	(0.028)

Simulation Study

Figure: Estimates with standard deviations in brackets for $T = 80$

Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$ by the Proposed Model (With Bias Correction)

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-2.294	-1.867	-0.954	-0.289	0.138	0.610	0.808	1.053	1.324	1.470
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.102	0.099	0.094
	(5.567e-4)	(5.438e-4)	(5.462e-4)	(5.440e-4)	(5.484e-4)	(5.373e-4)	(5.520e-4)	(5.472e-4)	(5.479e-4)	(5.433e-4)
$\alpha_x^{(2)}$	-0.012	-0.106	-0.106	-0.330	-0.447	-0.286	-0.278	8.087e-3	0.488	1.068
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085
	(5.167e-4)	(5.215e-4)	(5.194e-4)	(5.152e-4)	(5.140e-4)	(5.164e-4)	(5.131e-4)	(5.132e-4)	(5.176e-4)	(5.183e-4)

Estimates of $\mu^{(i)}$ and $\phi^{(i)}$ by the Proposed Model (With Bias Correction)

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.857	1.000	0.293	0.948
(0.030)	(7.743e-4)	(0.106)	(0.024)

Simulation Study

Figure: Estimates with standard deviations in brackets for $T = 150$

Lee-Carter Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-7.658	-7.623	-7.565	-7.298	-6.878	-6.452	-5.969	-5.507	-5.088	-4.599
	(0.060)	(0.064)	(0.073)	(0.078)	(0.078)	(0.078)	(0.075)	(0.073)	(0.071)	(0.067)
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.102	0.099	0.094
	(2.094e-4)	(2.067e-4)	(2.082e-4)	(2.127e-4)	(2.082e-4)	(2.073e-4)	(2.084e-4)	(2.086e-4)	(2.076e-4)	(2.108e-4)
$\alpha_x^{(2)}$	-9.140	-8.751	-8.286	-7.706	-7.135	-6.645	-6.140	-5.679	-5.279	-4.850
	(0.096)	(0.091)	(0.086)	(0.078)	(0.071)	(0.067)	(0.062)	(0.060)	(0.061)	(0.063)
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085
	(2.047e-4)	(2.043e-4)	(2.037e-4)	(2.035e-4)	(2.034e-4)	(2.028e-4)	(2.031e-4)	(2.065e-4)	(2.062e-4)	(2.033e-4)

Lee-Carter Estimates of $\mu^{(i)}$ and $\phi^{(i)}$

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.851	1.000	0.041	0.937
(8.505e-3)	(2.434e-4)	(4.727e-3)	(0.011)

Simulation Study

Figure: Estimates with standard deviations in brackets for $T = 150$

Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$ by the Proposed Model (Without Bias Correction)

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-2.293	-1.867	-0.954	-0.289	0.138	0.610	0.809	1.053	1.324	1.470
	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.016)
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.102	0.099	0.094
	(2.094e-4)	(2.066e-4)	(2.082e-4)	(2.127e-4)	(2.082e-4)	(2.073e-4)	(2.084e-4)	(2.086e-4)	(2.076e-4)	(2.108e-4)
$\alpha_x^{(2)}$	-0.012	-0.106	-0.106	-0.330	-0.447	-0.285	-0.278	7.924e-3	0.488	1.068
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085
	(2.047e-4)	(2.043e-4)	(2.037e-4)	(2.035e-4)	(2.034e-4)	(2.028e-4)	(2.031e-4)	(2.065e-4)	(2.062e-4)	(2.033e-4)

Estimates of $\mu^{(i)}$ and $\phi^{(i)}$ by the Proposed Model (Without Bias Correction)

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.862	1.000	0.836	0.839
(0.019)	(2.620e-4)	(0.127)	(0.026)

Simulation Study

Figure: Estimates with standard deviations in brackets for $T = 150$

Estimates of $\alpha_x^{(i)}$ and $\beta_x^{(i)}$ by the Proposed Model (With Bias Correction)

x	1	2	3	4	5	6	7	8	9	10
$\alpha_x^{(1)}$	-2.294	-1.867	-0.954	-0.289	0.138	0.610	0.809	1.053	1.324	1.470
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
$\beta_x^{(1)}$	0.083	0.089	0.102	0.108	0.109	0.109	0.105	0.102	0.099	0.094
	(2.111e-4)	(2.095e-4)	(2.097e-4)	(2.148e-4)	(2.108e-4)	(2.093e-4)	(2.106e-4)	(2.114e-4)	(2.101e-4)	(2.124e-4)
$\alpha_x^{(2)}$	-0.012	-0.106	-0.106	-0.330	-0.447	-0.285	-0.278	7.866e-3	0.488	1.068
	(0.017)	(0.016)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)
$\beta_x^{(2)}$	0.131	0.124	0.118	0.106	0.096	0.091	0.084	0.082	0.083	0.085
	(2.070e-4)	(2.066e-4)	(2.058e-4)	(2.063e-4)	(2.051e-4)	(2.056e-4)	(2.049e-4)	(2.086e-4)	(2.082e-4)	(2.051e-4)

Estimates of $\mu^{(i)}$ and $\phi^{(i)}$ by the Proposed Model (With Bias Correction)

$\mu^{(1)}$	$\phi^{(1)}$	$\mu^{(2)}$	$\phi^{(2)}$
-0.857	1.000	0.291	0.949
(0.019)	(2.667e-4)	(0.102)	(0.020)

Simulation Study

- The proposed least squares estimates without bias correction for $\mu^{(2)}$ and $\phi^{(2)}$ are inconsistent;
- The Lee-Carter inference clearly gives an inconsistent estimation for $\phi^{(2)}$ when $T = 150$;
- The standard errors for estimating $\alpha_x^{(i)}$ based on the Lee-Carter inference are much larger than those based on the proposed least squares estimation with or without bias correction;
- The bias-corrected inference performs quite well.

Conclusions

- How mortality rates are modeled have quite a few implications in annuity and pension fund practices because parametric mortality models can be relied on to manage longevity risk
- Our proposed model is based on our understanding of pitfalls of LC model
    - Different asymptotic results under different cases whether AR(1) process is stationary or unit root
    - Unit root test rejects null hypothesis for female and combined mortality rates; fails to reject for male rates
- Our bias corrected estimators overcome the issue of estimator inconsistency
    - Confirmed by simulation study that shows BC estimators display smaller bias and smaller mean squared error
    - Another dedicated paper that argues choice of AR(1) is adequate

Conclusions

- Methodology is now applied to a two population case
  - Since underwriters have policyholders from multiple populations, modeling multiple populations allows for quantifying the risk of whole portfolio
  - Asymptotic results; confirmed by nice simulation study results
- Implications for longevity hedging:
  - Can use derivatives to hedge against time-t values of longevity deltas of (unpaid) annuity liability
  - Risk transfer of longevity risk from annuity underwriter / pension fund to another financial institution

Thank you!