The Locally-Linear Cairns-Blake-Dowd Model: A Note on Delta-Nuga Hedging of Longevity Risk

Yanxin Liu and Johnny Siu-Hang Li

Presenter: Yanxin Liu

September 4, 2014

Outline

Importance of Managing Longevity Risk

- As a result of various factors including:
  1. a low interest rate environment
  2. changes in regulatory regimes (e.g., Solvency II which is scheduled to come into effort in 2013).

$\rightarrow$ Longevity risk has become a high profile risk in recent years.

- Pension plans and annuities providers are paying more attention to managing longevity risk.
- Two main aspects in managing longevity risk:
  - modelling mortality patterns
  - hedging longevity risk

Issues Related to the Current Stochastic Mortality Models

Sensitivity of estimation and forecast to different sample periods:

- For example, the drift terms of the bivariate random walk encompassed in the original Cairns-Blake-Dowd (CBD) model are highly sensitive to the first and last data points.
- What length of calibration window should be used?
  - $\rightarrow$ a subjective judgement
    - Cairns(2013) regarded it as a source of "Knightian Uncertainty".
- main driven factor: the deterministic trend assumption encompassed in the stochastic model
- possible solution:
  - $\rightarrow$ allow the drift terms to be stochastic
    - The Locally-linear CBD model is introduced.

Outline

States Space Model

Suppose at each time t, we have an observation of N-dimensional multivariate time series $\vec{y}_t$, which is driven by some unobservable $m \times 1$ hidden states vector $\vec{\alpha}_t$.

Measurement equation:

$$y_t = B_t \alpha_t + \vec{\varepsilon}_t.$$

To model the unobservable states process, $\vec{\alpha}_t$ itself is assumed to follow a first-order Markov process.

Transition equation:

$$\alpha_t = A_t \alpha_{t-1} + \vec{\eta}_t.$$

$\vec{\varepsilon}_t \sim MVN(0, R)$ and $\vec{\eta}_t \sim MVN(0, Q)$ represent the error terms in measurement equations and transition equations, respectively.

Locally-Linear Cairns-Blake-Dowd model

- Denote by

$$y_{x,t} = \mathrm{logit}(q_{x,t}) = \log(\frac{q_{x,t}}{1 - q_{x,t}}).$$

- Suppose the dataset we use include in total K ages.

- For $i = 1, 2, \kappa_i(t)$ follows a random walk with drift:

$$\kappa_i(t) = C_i(t) + \kappa_i(t-1) + \eta_i(t),$$

  where $C_i(t)$ itself is stochastic.

- The underlying states vector becomes $(\kappa_1(t), \kappa_2(t), C_1(t), C_2(t))'$.

The corresponding measurement equation can be written as:

$$y_t = B_t \alpha_t + \vec{\epsilon}_t$$

and the transition equation can be shown as:

$$\alpha_t = A_t \alpha_{t-1} + \vec{\eta}_t,$$

where

$$y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{K,t} \end{pmatrix}, \alpha_t = \begin{pmatrix} \kappa_1(t) \\ \kappa_2(t) \\ C_1(t) \\ C_2(t) \end{pmatrix},$$

and both B_t and A_t can be expressed as time-invariant matrices:

$$B = \begin{pmatrix} 1 & x_1 - \bar{x} & 0 & 0 \\ 1 & x_2 - \bar{x} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_K - \bar{x} & 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

As a whole, the locally-linear Cairns-Blake-Dowd model (LLCBD) can be written as:

Measurement equations:

$$\begin{cases} \mathrm{logit}(q_{1,t}) & = & \kappa_1(t) + (x_1 - \bar{x})\kappa_2(t) + \varepsilon_{1,t} \\ \mathrm{logit}(q_{2,t}) & = & \kappa_1(t) + (x_2 - \bar{x})\kappa_2(t) + \varepsilon_{2,t} \\ & \vdots & \\ \mathrm{logit}(q_{K,t}) & = & \kappa_1(t) + (x_K - \bar{x})\kappa_2(t) + \varepsilon_{K,t} \end{cases}$$

Transition equations:

$$\begin{cases} \kappa_1(t) & = & \kappa_1(t-1) + C_1(t) + \eta_{1,t} \\ \kappa_2(t) & = & \kappa_2(t-1) + C_2(t) + \eta_{2,t} \\ C_1(t) & = & C_1(t-1) + \eta_{3,t} \\ C_2(t) & = & C_2(t-1) + \eta_{4,t} \end{cases}$$

Properties of the LLCBD Model Specification

1. Stochastic nature in the drift terms
   $\Rightarrow$ More flexibility in capturing the longevity trend

2. State Space Form
   - Generalization to other time series structures
     - by adjusting A_t or B_t
   - Generalization to multi-population
     - by adjusting the structure in state vector α_t
   - Generalization to other mortality models
     - by adjusting the measurement equations
     - e.g. Hári et al.(2008) consider the Lee-Carter Model in SSM

3. Unconstrained Q matrix
   $\Rightarrow$ Availability to account for the correlation between different states

Outline

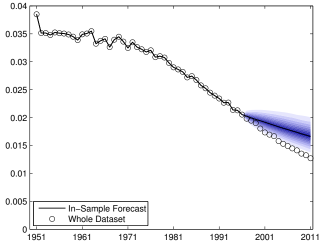
Robustness Study:

1. compare the in-sample forecasting performance between the original CBD model and the locally-linear CBD model;
2. investigate the robustness with respect to different look-back windows.

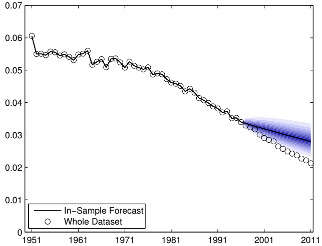
Data for the in-sample forecast:

- Population: English and Welsh (EW) male population
- Sample Period: 1951 to 1996 versus 1951 to 2011
- Age Range: 50 to 89

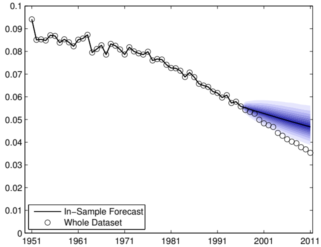
In-Sample Forecast Study in Terms of $q_{x,t}$



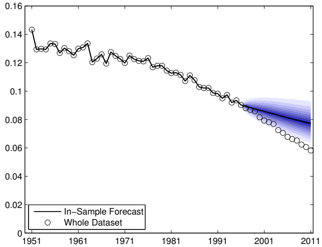
(a) Original CBD Age 65



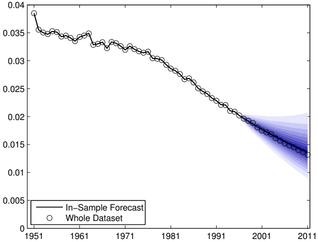
(b) Original CBD Age 70



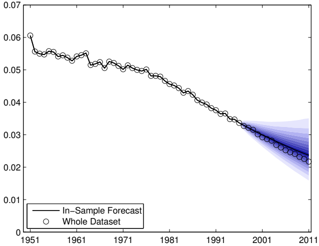
(c) Original CBD Age 75



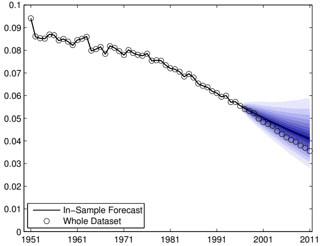
(d) Original CBD Age 80



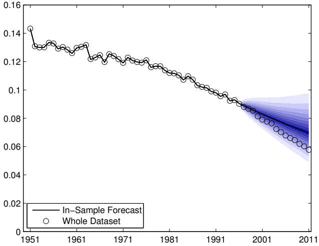
(a) LL CBD Model Age 65



(b) LL CBD Model Age 70



(c) LL CBD Model Age 75

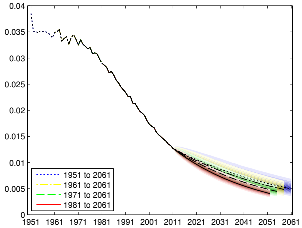


(d) LL CBD Model Age 80

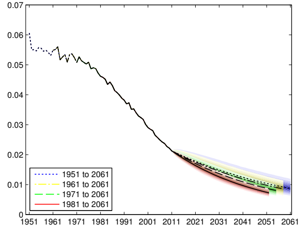
Robustness: Different Look-back Windows

Paradox in using different sample period

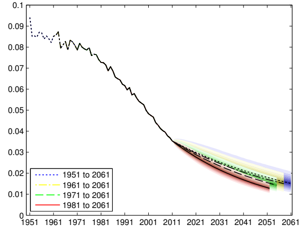
- It is reasonable to use the most up-to-date data;
- One paradox in using the original CBD model:
  longer sample period $\rightarrow$ less capacity in capturing the latest trend;
- The locally-linear CBD model provides a solution!
- Consider the following sample periods:
  - 1951 to 2011
  - 1961 to 2011
  - 1971 to 2011
  - 1981 to 2011



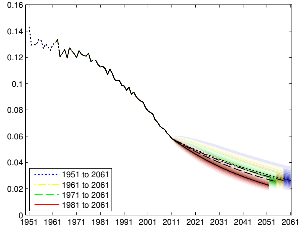
(a) Original CBD Age 65



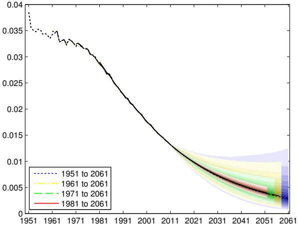
(b) Original CBD Age 70



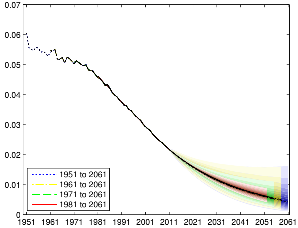
(c) Original CBD Age 75



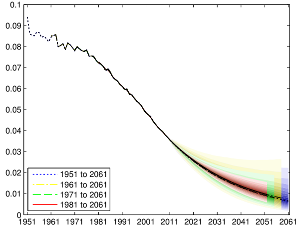
(d) Original CBD Age 80



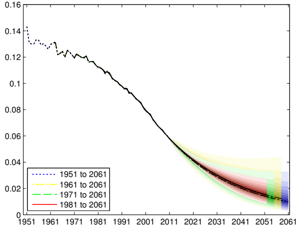
(a) LL CBD Model Age 65



(b) LL CBD Model Age 70



(c) LL CBD Model Age 75



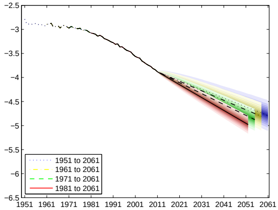
(d) LL CBD Model Age 80

Comments:

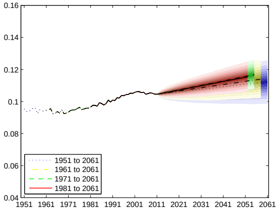
- Original CBD Model:
  - The use of different sample periods lead to very different median forecasts;
  - Relatively small forecast errors.
- Locally-linear CBD Model:
  - The median forecasts are substantially more robust;
  - Relatively large forecast errors as it incorporates the stochastic nature of drift terms.

What about the unobservable states?

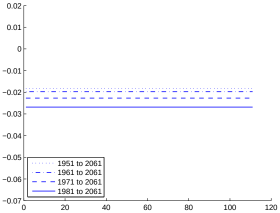
Different Look-Back Windows in Terms of Unobservable States



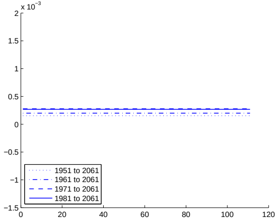
(a) Random Walk κ_1



(b) Random Walk κ_2

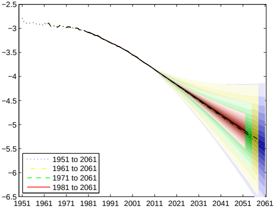


(c) Random Walk C_1

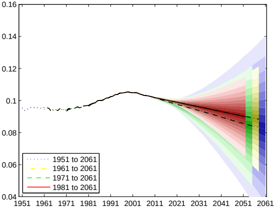


(d) Random Walk C_2

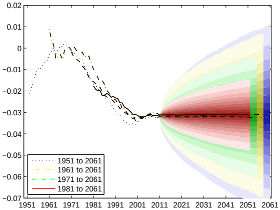
Different Look-Back Windows in Terms of Unobservable States



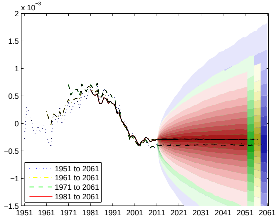
(a) State Space Model κ_1



(b) State Space Model κ_2



(c) State Space Model C_1



(d) State Space Model C_2

Comments:

- Key states: $C_1(t)$ and $C_2(t)$.
- LL CBD model:
  - the difference in the median forecasts of $C_1(t)$ over different sample periods is much smaller than that from the original CBD model;
  - great similarity in the patterns of drift processes $C_1(t)$ and $C_2(t)$ under different look-back windows;

  $\rightarrow$ the median forecast is substantially more robust than the CBD model.

Outline

Basic Set Up for Hedging Longevity Risk

Notation:

- L_0: the time-0 value of the liability being hedge
- H_j: the time-0 value of the jth hedging instrument
- N_j: the number of the jth hedging instrument
- P_0: the time-0 value of the constructed hedging portfolio

Setting:

- Suppose at time 0 we have a liability of 15 years term life annuity which is sold to age 65.
- The present value of this liability can be expressed as:

$$L_0 = \sum_{u=1}^{15} e^{-ru} \prod_{t=1}^{u} p_{x_0+t-1,t},$$

where

$$p_{x,t} = 1 - q_{x,t} = \frac{1}{1 + e^{\kappa_1(t)+(x-\bar{x})\kappa_2(t)+\varepsilon_{x,t}}}.$$

- To hedge against this liability, we use q-forward contracts as the hedging instruments.

- The present value of the jth q-forward contract can be expressed as:

$$e^{-rt_j}(q_{x_j,t_j} - q^f_{x_j,t_j}),$$

where $q^f_{x_j,t_j}$ is the forward mortality rate and r is the interest rate. In particular, we set $r = 0.01$.

Hedging Longevity Risk

Objective:

- To stabilize the present value of the unexpected cash flow.

Main Idea:

- To use the hedging instrument to construct a hedging portfolio s.t. the variance of the portfolio is minimized.

Evaluation:

- Define Hedge Effectiveness (HE) in terms of the amount of risk reduction:

$$HE = 1 - \frac{Var(P_0)}{Var(L_0)}.$$

A higher value of HE represents a greater amount of risk reduction.

- At time 0, we use the best estimate of $q_{x,t}$ at time 0 to approximate the value for different contracts. The hedging strategy is based on this approximation.

- The value of the liability and the jth hedging instrument becomes:

$$\hat{L}_0 = \sum_{u=1}^{15} e^{-ru} \prod_{t=1}^{u} \hat{p}_{x_0+t-1,t}$$

and

$$\hat{H}_j = e^{-rt_j}(\hat{q}_{x_j,t_j} - q^f_{x_j,t_j}),$$

where

$$\hat{p}_{x,t} = 1 - \hat{q}_{x,t} = \frac{1}{1 + e^{\kappa_1(0)+C_1\times t+(x-\bar{x})(\kappa_2(0)+C_2\times t)}}.$$

Delta-Hedging

- Using the delta-hedging approach, the hedging portfolio would be constructed with at least two different q-forwards and should satisfy the following system of equations:

$$\begin{cases} \frac{\partial \hat{L}_0}{\partial \kappa_1(0)} & = & N_1 \times \frac{\partial \hat{H}_1}{\partial \kappa_1(0)} + N_2 \times \frac{\partial \hat{H}_2}{\partial \kappa_1(0)} \\ \frac{\partial \hat{L}_0}{\partial \kappa_2(0)} & = & N_1 \times \frac{\partial \hat{H}_1}{\partial \kappa_2(0)} + N_2 \times \frac{\partial \hat{H}_2}{\partial \kappa_2(0)} \end{cases}.$$

- In matrix form, we have

$$\begin{pmatrix} \frac{\partial \hat{L}_0}{\partial \kappa_1(0)} \\ \frac{\partial \hat{L}_0}{\partial \kappa_2(0)} \end{pmatrix} = \begin{pmatrix} \frac{\partial \hat{H}_1}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_2}{\partial \kappa_1(0)} \\ \frac{\partial \hat{H}_1}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_2}{\partial \kappa_2(0)} \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$$

- The present value of the constructed portfolio P can be expressed as

$$P = L - N_1 \times H_1 - N_2 \times H_2.$$

Nuga-Hedging

- The treatment of the drift terms as random variables naturally calls for Nuga-hedging, a technique proposed by Cairns (2013) to hedge the risk associated with changes in drifts.
- Using the Delta-Nuga-hedging approach, we need to match all four derivatives:

$$\begin{pmatrix} \frac{\partial \hat{L}_0}{\partial \kappa_1(0)} \\ \frac{\partial \hat{L}_0}{\partial \kappa_2(0)} \\ \frac{\partial \hat{L}_0}{\partial C_1(0)} \\ \frac{\partial \hat{L}_0}{\partial C_2(0)} \end{pmatrix} = \begin{pmatrix} \frac{\partial \hat{H}_1}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_2}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_3}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_4}{\partial \kappa_1(0)} \\ \frac{\partial \hat{H}_1}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_2}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_3}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_4}{\partial \kappa_2(0)} \\ \frac{\partial \hat{H}_1}{\partial C_1(0)} & \frac{\partial \hat{H}_2}{\partial C_1(0)} & \frac{\partial \hat{H}_3}{\partial C_1(0)} & \frac{\partial \hat{H}_4}{\partial C_1(0)} \\ \frac{\partial \hat{H}_1}{\partial C_2(0)} & \frac{\partial \hat{H}_2}{\partial C_2(0)} & \frac{\partial \hat{H}_3}{\partial C_2(0)} & \frac{\partial \hat{H}_4}{\partial C_2(0)} \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \end{pmatrix}$$

- And the corresponding present value of the constructed portfolio is

$$P = L - N_1 \times H_1 - N_2 \times H_2 - N_3 \times H_3 - N_4 \times H_4.$$

Issues Related to the Current Hedging Strategy:

- The information underlying the correlation between different states is not fully utilized.
- In the previous literature, main focus is the correlation of the states across different populations (basis risk).
- The correlation of the states within the same population is not the main concern, if there are sufficient instruments to hedge all states. For example
  - Cairns (2011) considers the probit transform of the survival probabilities and derives the dynamic hedging strategies where both states κ_1 and κ_2 are hedged;
  - Cairns (2013) investigates the robust hedging strategy where, in addition to the hedging of κ_t, the trend state ν_κ would be hedged.

Issues Related to the Current Hedging Strategy: Con't

- The hedging strategies based on highly correlated states may reduce the hedge effectiveness

- More importantly, for a market when only a handful of hedging instruments can be used, the utilization of each instrument becomes one of the decisive factors when deriving the hedging strategies.

⋆ We consider a new hedging technique that is able to

1. account for the correlation between different states within the same population and across different populations simultaneously;
2. obtain higher hedge effectiveness when only a few instruments can be used.

Transformed Hedging Technique

- Recall that in the setting of LLCBD model, the Q matrix which represents the covariance matrix of different states is unconstrained.
  $\Rightarrow$ In other words, the states vector is correlated.
- To further improve the delta- and nuga-hedging techniques, we consider a LDL transform to the Q matrix.
- The objective of this method is to decompose Q into the product of three matrix, denoted as KQ^*K', where Q^* is a diagonal matrix with diagonal elements in decreasing order.
- Similar transformation techniques include the singular value decomposition.

Under the LDL transformation, the transition equation then follows

$$\begin{array}{rcl} \alpha_t & = & A\alpha_{t-1} + \eta_t \\ \Rightarrow \quad K^{-1}\alpha_t & = & K^{-1}AKK^{-1}\alpha_{t-1} + K^{-1}\eta_t \\ \Rightarrow \quad \alpha_t^* & = & A^*\alpha_{t-1}^* + \eta_t^*, \end{array}$$

where

$$\alpha_t^* = K^{-1}\alpha_t, \quad A^* = K^{-1}AK \quad \text{and} \quad \eta_t^* \sim \mathrm{MVN}(0, Q^*).$$

The partial derivatives under the transformed states are then calculated as:

$$\frac{\partial \hat{L}}{\partial \alpha_t^*} = K' \frac{\partial \hat{L}}{\partial \alpha_t},$$

where $\partial \hat{L}/\partial \alpha_t = (\partial \hat{L}/\partial \kappa_1(t), \partial \hat{L}/\partial \kappa_2(t), \partial \hat{L}/\partial C_1(t), \partial \hat{L}/\partial C_2(t))'$.

Illustration I: Simulation Study

Illustration I:

- EW male population
- Sample period: 1951 to 2011
- Assume no basis risk
- Data source: HMD (2014)

Hedging Instrument:

j_{th} Hedging Instrument	x_j	t_j		
H_1	73	8	$\leftarrow$	approx. the median age of the cohort in the liability
H_2	68	3	$\leftarrow$	approx. the 1/4 age of the same cohort
H_3	78	13	$\leftarrow$	approx. the 3/4 age of the same cohort
H_4	65	5	$\leftarrow$	selected to avoid singular problem

Why the singularity problem occurs when four q-forwards linked to the same cohort are used?

$$\begin{cases} \frac{\partial \hat{H}_j}{\partial \kappa_1(0)} & = & \hat{p}_{x_j,t_j}\hat{q}_{x_j,t_j} \\ \frac{\partial \hat{H}_j}{\partial \kappa_2(0)} & = & (x_j - \bar{x})\hat{p}_{x_j,t_j}\hat{q}_{x_j,t_j} \\ \frac{\partial \hat{H}_j}{\partial C_1(0)} & = & t_j\hat{p}_{x_j,t_j}\hat{q}_{x_j,t_j} \\ \frac{\partial \hat{H}_j}{\partial C_2(0)} & = & t_j(x_j - \bar{x})\hat{p}_{x_j,t_j}\hat{q}_{x_j,t_j} \end{cases} \Rightarrow \begin{pmatrix} \frac{\partial \hat{H}_1}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_2}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_3}{\partial \kappa_1(0)} & \frac{\partial \hat{H}_4}{\partial \kappa_1(0)} \\ \frac{\partial \hat{H}_1}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_2}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_3}{\partial \kappa_2(0)} & \frac{\partial \hat{H}_4}{\partial \kappa_2(0)} \\ \frac{\partial \hat{H}_1}{\partial C_1(0)} & \frac{\partial \hat{H}_2}{\partial C_1(0)} & \frac{\partial \hat{H}_3}{\partial C_1(0)} & \frac{\partial \hat{H}_4}{\partial C_1(0)} \\ \frac{\partial \hat{H}_1}{\partial C_2(0)} & \frac{\partial \hat{H}_2}{\partial C_2(0)} & \frac{\partial \hat{H}_3}{\partial C_2(0)} & \frac{\partial \hat{H}_4}{\partial C_2(0)} \end{pmatrix}$$

Illustration I: Empirical Result

Hedge Effectiveness				
	1 State	2 States	3 States	4 States
CBD	70.4416%	91.8240%	89.2252%	97.5299%
LLCBD	70.8834%	91.8779%	90.1934%	98.1765%
LLCBD*	77.4641%	92.3658%	95.3269% (97.1192%)	98.1765%

- In LLCBD*, the hedging strategies are obtained by matching the transformed states.
- (97.1192%) represents the H.E. from hedging the first three transformed states.

Illustration II: Real Data Analysis

Illustration II:

- EW male population
- Assume no basis risk
- Data source: HMD (2014)
- Same hedging instruments as Illustration I
- Sample period: 1951 to 1996
- Whole sample: 1951 to 2011

Evaluation of the Actual Hedging Performance (AHP)

- Define AHP as the absolute value of the actual cash flow discounted to time 0:

$$AHP = |L_0^{(actual)} - \hat{L}_0 - \sum_{j=1}^{m} N_j H_j^{(actual)}|,$$

  where m is the number of states being hedged.
- A smaller value in AHP represents a better hedging performance.

Actual Hedging Performance

Actual Hedging Performance				
	1 State	2 States	3 States	4 States
CBD	0.0812	0.0782	0.1049	0.0213
LLCBD	0.0660	0.0584	0.0775	0.0162
LLCBD*	0.0424	0.0072	0.0156 (0.0403)	0.0162

Outline

Conclusions

- By allowing the drift term to be random, Locally-linear CBD is capable in capturing the latest trend in the sample period.
- The median forecast of Locally-linear CBD model is substantially more robust than that from the original CBD model.
- Without considering basis risk, the hedging strategies derived from the transformed states of LLCBD model has the ability to retrieve more hedge effectiveness.
- In the real data analysis, the hedging performance from the transformed technique also shows to be more efficient.

Future Extensions

- By adapting the measurement equation, Locally-linear CBD model can be easily generalized to model multiple populations.
- Under the multi-population LLCBD model, we are able to account for the basis risk in longevity hedging.
- When deriving the hedging strategies, the LDL transform can also be applied to the multi-population LLCBD model to achieve higher hedge effectiveness.