Canonical Valuation of Mortality-linked Securities

Johnny S.H. Li

University of Waterloo

September 10, 2009

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Problem

- Mortality-linked Securities are becoming more popular.
- Pricing these securities is not straightforward.
- Reasons:
  - Incomplete market.
  - A lack of liquidly traded longevity indexes or securities.
  - A replicating hedge cannot be formed.
- Pricing mortality-linked securities is the focus of this study.

The Wang Transform

- The Wang transform (Wang, 1996, 2000, 2001).
- Let $F_P(x)$ be the d.f. for a future lifetime r.v. under P.
- Then, under Q, the distribution function for the r.v. is

$$F_Q(x) = \Phi(\Phi^{-1}(F_P(x)) + \lambda),$$

where Φ is the d.f. for a standard normal r.v., and λ is the market price of risk.

The Wang Transform

- The Wang transform (Wang, 1996, 2000, 2001).
- Let $F_P(x)$ be the d.f. for a future lifetime r.v. under P.
- Then, under Q, the distribution function for the r.v. is

$$F_Q(x) = \Phi(\Phi^{-1}(F_P(x)) + \lambda),$$

where Φ is the d.f. for a standard normal r.v., and λ is the market price of risk.

The Wang Transform

- The Wang transform (Wang, 1996, 2000, 2001).
- Let $F_P(x)$ be the d.f. for a future lifetime r.v. under P.
- Then, under Q, the distribution function for the r.v. is

$$F_Q(x) = \Phi(\Phi^{-1}(F_P(x)) + \lambda),$$

where Φ is the d.f. for a standard normal r.v., and λ is the market price of risk.

The Wang Transform

- Examples:
  1. Lin and Cox (2005): calibrate λ to market quotes of immediate annuities.
  2. Denuit, Devolder and Goderniaux (2007): integrate with the Lee-Carter model.
- Consistent with the classical capital asset pricing model (CAPM).
- Ruhm (2003) and Pelsser (2008) point out that it may not lead to a price consistent with the arbitrage-free price for general stochastic processes.

The Wang Transform

- Examples:
  1. Lin and Cox (2005): calibrate λ to market quotes of immediate annuities.
  2. Denuit, Devolder and Goderniaux (2007): integrate with the Lee-Carter model.
- Consistent with the classical capital asset pricing model (CAPM).
- Ruhm (2003) and Pelsser (2008) point out that it may not lead to a price consistent with the arbitrage-free price for general stochastic processes.

The Wang Transform

- Examples:
  1. Lin and Cox (2005): calibrate λ to market quotes of immediate annuities.
  2. Denuit, Devolder and Goderniaux (2007): integrate with the Lee-Carter model.
- Consistent with the classical capital asset pricing model (CAPM).
- Ruhm (2003) and Pelsser (2008) point out that it may not lead to a price consistent with the arbitrage-free price for general stochastic processes.

The Wang Transform

- Examples:
  1. Lin and Cox (2005): calibrate λ to market quotes of immediate annuities.
  2. Denuit, Devolder and Goderniaux (2007): integrate with the Lee-Carter model.

- Consistent with the classical capital asset pricing model (CAPM).

- Ruhm (2003) and Pelsser (2008) point out that it may not lead to a price consistent with the arbitrage-free price for general stochastic processes.

The Wang Transform

- Examples:
  1. Lin and Cox (2005): calibrate λ to market quotes of immediate annuities.
  2. Denuit, Devolder and Goderniaux (2007): integrate with the Lee-Carter model.

- Consistent with the classical capital asset pricing model (CAPM).

- Ruhm (2003) and Pelsser (2008) point out that it may not lead to a price consistent with the arbitrage-free price for general stochastic processes.

Risk-neutral dynamics of death/survival rates

- Step (1): Define a mortality model in P measure.
- E.g., Cairns, Blake and Dowd (2006) model:

$$\begin{aligned}\ln \frac{q_{x,t}}{1-q_{x,t}} &= A_1(t) + A_2(t)(x+t), \\ A(t+1) &= A(t) + \mu + CZ(t+1),\end{aligned}$$

where

- $A(t) = (A_1(t), A_2(t))'$,
- μ is a constant 2×1 vector,
- C is a constant 2×2 upper triangular matrix,
- $Z(t)$ is a 2-dimensional standard normal r.v.andom variable.

Risk-neutral dynamics of death/survival rates

- Step (2): Adjust the drift term to obtain a model in Q measure:

$$A(t+1) = A(t) + \tilde{\mu} + C\tilde{Z}(t+1),$$

  where
  - $\tilde{\mu} = \mu - C\lambda$,
  - $\tilde{Z}(t+1)$ is a standard 2-dim. normal r.v. under the Q-measure,
  - $\lambda = (\lambda_1, \lambda_2)'$ is a vector of market prices of risk.

- Cairns, Blake and Dowd (2006) obtain λ by calibrating to the price of the BNP/EIB longevity bond.

Risk-neutral dynamics of death/survival rates

- Problem (1): Parameter risk.
  - Even if the process is correct, parameters may be wrong.
  - Can be quantified by MCMC.
- Problem (2): Model risk.
  - The process itself may be incorrect.
  - May be reduced by considering a less stringent mortality model.

Risk-neutral dynamics of death/survival rates

- Problem (1): Parameter risk.
  - Even if the process is correct, parameters may be wrong.
  - Can be quantified by MCMC.
- Problem (2): Model risk.
  - The process itself may be incorrect.
  - May be reduced by considering a less stringent mortality model.

Risk-neutral dynamics of death/survival rates

- Problem (1): Parameter risk.
  - Even if the process is correct, parameters may be wrong.
  - Can be quantified by MCMC.
- Problem (2): Model risk.
  - The process itself may be incorrect.
  - May be reduced by considering a less stringent mortality model.

Risk-neutral dynamics of death/survival rates

- Problem (1): Parameter risk.
  - Even if the process is correct, parameters may be wrong.
  - Can be quantified by MCMC.
- Problem (2): Model risk.
  - The process itself may be incorrect.
  - May be reduced by considering a less stringent mortality model.

Risk-neutral dynamics of death/survival rates

- Problem (1): Parameter risk.
  - Even if the process is correct, parameters may be wrong.
  - Can be quantified by MCMC.
- Problem (2): Model risk.
  - The process itself may be incorrect.
  - May be reduced by considering a less stringent mortality model.

Risk-neutral dynamics of death/survival rates

- Problem (1): Parameter risk.
  - Even if the process is correct, parameters may be wrong.
  - Can be quantified by MCMC.
- Problem (2): Model risk.
  - The process itself may be incorrect.
  - May be reduced by considering a less stringent mortality model.

Our Idea...

- 'Canonical valuation' (Stutzer, 1996) as an alternative pricing method.
- Advantages:
  1. Largely non-parametric – reducing parameter and model risk.
  2. Useful even if only a few market prices are available.
- **Our objective:** to develop a framework for pricing mortality-linked securities using canonical valuation.

Our Idea...

- ‘Canonical valuation’ (Stutzer, 1996) as an alternative pricing method.
- Advantages:
  1. Largely non-parametric – reducing parameter and model risk.
  2. Useful even if only a few market prices are available.
- **Our objective:** to develop a framework for pricing mortality-linked securities using canonical valuation.

Our Idea...

- 'Canonical valuation' (Stutzer, 1996) as an alternative pricing method.
- Advantages:
  1. Largely non-parametric – reducing parameter and model risk.
  2. Useful even if only a few market prices are available.
- **Our objective:** to develop a framework for pricing mortality-linked securities using canonical valuation.

Our Idea...

- ‘Canonical valuation’ (Stutzer, 1996) as an alternative pricing method.
- Advantages:
  1. Largely non-parametric – reducing parameter and model risk.
  2. Useful even if only a few market prices are available.
- **Our objective:** to develop a framework for pricing mortality-linked securities using canonical valuation.

Our Idea...

- ‘Canonical valuation’ (Stutzer, 1996) as an alternative pricing method.
- Advantages:
  1. Largely non-parametric – reducing parameter and model risk.
  2. Useful even if only a few market prices are available.
- **Our objective:** to develop a framework for pricing mortality-linked securities using canonical valuation.

The Principle

- Assume there are m distinct primary securities.
- Each has a time-zero price of F_i and a random discounted payoff of $f_i(\omega)$.
- Let $\mathcal{Q}$ is the set of all equivalent martingale measures.
- We require, for any Q in $\mathcal{Q}$,

$$\mathbb{E}^Q[f_i(\omega)] = F_i, \quad i = 1, 2, \ldots, m. \qquad (1)$$

The Principle

- Assume there are m distinct primary securities.
- Each has a time-zero price of F_i and a random discounted payoff of $f_i(\omega)$.
- Let $\mathcal{Q}$ is the set of all equivalent martingale measures.
- We require, for any Q in $\mathcal{Q}$,

$$\mathbb{E}^{Q}[f_i(\omega)] = F_i, \quad i = 1, 2, \ldots, m. \tag{1}$$

The Principle

- Assume there are m distinct primary securities.
- Each has a time-zero price of F_i and a random discounted payoff of $f_i(\omega)$.
- Let $\mathcal{Q}$ is the set of all equivalent martingale measures.
- We require, for any Q in $\mathcal{Q}$,

$$\mathbb{E}^Q[f_i(\omega)] = F_i, \quad i = 1, 2, \ldots, m. \tag{1}$$

The Principle

- Assume there are m distinct primary securities.
- Each has a time-zero price of F_i and a random discounted payoff of $f_i(\omega)$.
- Let $\mathcal{Q}$ is the set of all equivalent martingale measures.
- We require, for any Q in $\mathcal{Q}$,

$$\mathbb{E}^Q[f_i(\omega)] = F_i, \quad i = 1, 2, \ldots, m. \tag{1}$$

The Principle

- The Kullback-Leibler (1951) information criterion (KLIC):

$$D(Q,P) = \mathbb{E}^P \left[\frac{\mathrm{d}Q}{\mathrm{d}P} \ln \frac{\mathrm{d}Q}{\mathrm{d}P} \right]$$

- We should choose an equivalent martingale measure Q_0 that minimizes the criterion, i.e.,

$$Q_0 = \arg\min_{Q \in \mathcal{Q}} D(Q,P),$$

subject to the constraints in equation (1).

The Principle

- The Kullback-Leibler (1951) information criterion (KLIC):

$$D(Q,P) = \mathbb{E}^P \left[\frac{\mathrm{d}Q}{\mathrm{d}P} \ln \frac{\mathrm{d}Q}{\mathrm{d}P} \right]$$

- We should choose an equivalent martingale measure Q_0 that minimizes the criterion, i.e.,

$$Q_0 = \underset{Q \in \mathcal{Q}}{\arg\min}\ D(Q,P),$$

subject to the constraints in equation (1).

Statistical Justifications

- $D(Q, P)$ represents the information gained by moving from P to Q.
- From a Bayesian viewpoint, we may regard P as the prior distribution.
- Given m market prices, we can update the prior by incorporating the information contained in equation (1).
- No information other than equation (1) should be incorporated.

Statistical Justifications

- $D(Q, P)$ represents the information gained by moving from P to Q.
- From a Bayesian viewpoint, we may regard P as the prior distribution.
- Given m market prices, we can update the prior by incorporating the information contained in equation (1).
- No information other than equation (1) should be incorporated.

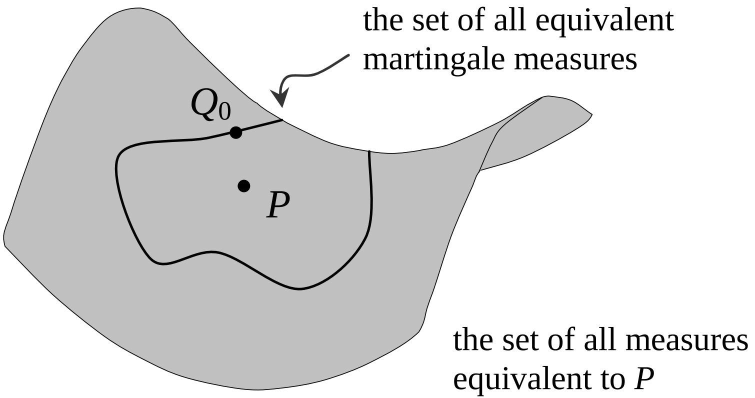
Statistical Justifications

- $D(Q, P)$ represents the information gained by moving from P to Q.
- From a Bayesian viewpoint, we may regard P as the prior distribution.
- Given m market prices, we can update the prior by incorporating the information contained in equation (1).
- No information other than equation (1) should be incorporated.

Statistical Justifications

- $D(Q, P)$ represents the information gained by moving from P to Q.
- From a Bayesian viewpoint, we may regard P as the prior distribution.
- Given m market prices, we can update the prior by incorporating the information contained in equation (1).
- No information other than equation (1) should be incorporated.

A Geometric Interpretation



Expected Utility Hypothesis

- Rittelli (2000) proved that maximizing the expected exponential utility is equivalent to minimizing the KLIC.
- The result also holds true in a multi-period setting.
- It implies linkages to the Esscher transform (Gerber and Shiu, 1994).

Expected Utility Hypothesis

- Rittelli (2000) proved that maximizing the expected exponential utility is equivalent to minimizing the KLIC.
- The result also holds true in a multi-period setting.
- It implies linkages to the Esscher transform (Gerber and Shiu, 1994).

Expected Utility Hypothesis

- Rittelli (2000) proved that maximizing the expected exponential utility is equivalent to minimizing the KLIC.
- The result also holds true in a multi-period setting.
- It implies linkages to the Esscher transform (Gerber and Shiu, 1994).

Implementation

- Generate N equally probable scenarios by the bootstrap.
- The p.f. for ω under P

$$\Pr(\omega = \omega_j) = \pi_j = \frac{1}{N}, \quad j = 1, 2, \ldots, N.$$

- Let π_j^*, $j = 1, 2, \ldots, N$, be the p.f. of ω under Q.
- We require

$$\sum_{j=1}^{N} f_i(\omega_j)\pi_j^* = F_i, \quad i = 1, 2, \ldots, m. \tag{2}$$

Implementation

- Generate N equally probable scenarios by the bootstrap.

- The p.f. for ω under P

$$\Pr(\omega = \omega_j) = \pi_j = \frac{1}{N}, \quad j = 1, 2, \ldots, N.$$

- Let π_j^*, $j = 1, 2, \ldots, N$, be the p.f. of ω under Q.

- We require

$$\sum_{j=1}^{N} f_i(\omega_j)\pi_j^* = F_i, \quad i = 1, 2, \ldots, m. \tag{2}$$

Implementation

- Generate N equally probable scenarios by the bootstrap.

- The p.f. for ω under P

$$\Pr(\omega = \omega_j) = \pi_j = \frac{1}{N}, \quad j = 1, 2, \ldots, N.$$

- Let π_j^*, $j = 1, 2, \ldots, N$, be the p.f. of ω under Q.

- We require

$$\sum_{j=1}^{N} f_i(\omega_j)\pi_j^* = F_i, \quad i = 1, 2, \ldots, m. \tag{2}$$

Implementation

- Generate N equally probable scenarios by the bootstrap.
- The p.f. for ω under P

$$\Pr(\omega = \omega_j) = \pi_j = \frac{1}{N}, \quad j = 1, 2, \ldots, N.$$

- Let π_j^*, $j = 1, 2, \ldots, N$, be the p.f. of ω under Q.
- We require

$$\sum_{j=1}^{N} f_i(\omega_j)\pi_j^* = F_i, \quad i = 1, 2, \ldots, m. \qquad (2)$$

Implementation

- We can rewrite KLIC as

$$\sum_{j=1}^{N} \pi_j^* \ln \frac{\pi_j^*}{\pi_j}.$$

- To find the canonical measure Q_0, we solve

$$Q_0 = \arg\min_{\pi_j^*} \sum_{j=1}^{N} \pi_j^* \ln \frac{\pi_j^*}{\pi_j},$$

subject to $\sum_{j=1}^{N} \pi_j^* = 1$ and equation (2).

Implementation

- We can rewrite KLIC as

$$\sum_{j=1}^{N} \pi_j^* \ln \frac{\pi_j^*}{\pi_j}.$$

- To find the canonical measure Q_0, we solve

$$Q_0 = \arg\min_{\pi_j^*} \sum_{j=1}^{N} \pi_j^* \ln \frac{\pi_j^*}{\pi_j},$$

subject to $\sum_{j=1}^{N} \pi_j^* = 1$ and equation (2).

The Challenge

- An empirical distribution of the mortality-linked security's payoff is needed.
- Generate from a time-series of past mortality rates or values of a longevity index.
- The data involve two dimensions: age and time.
- Potential dependency over both dimensions.

The Challenge

- An empirical distribution of the mortality-linked security's payoff is needed.
- Generate from a time-series of past mortality rates or values of a longevity index.
- The data involve two dimensions: age and time.
- Potential dependency over both dimensions.

The Challenge

- An empirical distribution of the mortality-linked security's payoff is needed.
- Generate from a time-series of past mortality rates or values of a longevity index.
- The data involve two dimensions: age and time.
- Potential dependency over both dimensions.

The Challenge

- An empirical distribution of the mortality-linked security's payoff is needed.
- Generate from a time-series of past mortality rates or values of a longevity index.
- The data involve two dimensions: age and time.
- Potential dependency over both dimensions.

Age Dependency

- Mortality rates at different ages are correlated with one another.
- Wills and Sherris (2008) point out that it is a critical factor in pricing mortality-linked securities.
- We consider mortality rates at different ages jointly by treating them as a vector.
- That is, we treat the data as a multivariate time-series of $\mathbf{m}_t = (m_{65,t}, m_{66,t}, \ldots, m_{90,t})'$.

Age Dependency

- Mortality rates at different ages are correlated with one another.
- Wills and Sherris (2008) point out that it is a critical factor in pricing mortality-linked securities.
- We consider mortality rates at different ages jointly by treating them as a vector.
- That is, we treat the data as a multivariate time-series of $\mathbf{m}_t = (m_{65,t}, m_{66,t}, \ldots, m_{90,t})'$.

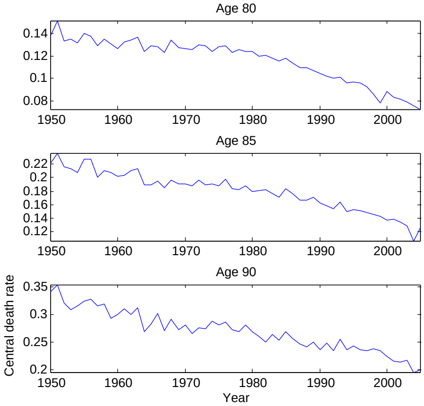
Age Dependency

- Mortality rates at different ages are correlated with one another.
- Wills and Sherris (2008) point out that it is a critical factor in pricing mortality-linked securities.
- We consider mortality rates at different ages jointly by treating them as a vector.
- That is, we treat the data as a multivariate time-series of $\mathbf{m}_t = (m_{65,t}, m_{66,t}, \ldots, m_{90,t})'$.

Age Dependency

- Mortality rates at different ages are correlated with one another.
- Wills and Sherris (2008) point out that it is a critical factor in pricing mortality-linked securities.
- We consider mortality rates at different ages jointly by treating them as a vector.
- That is, we treat the data as a multivariate time-series of $\mathbf{m}_t = (m_{65,t}, m_{66,t}, \ldots, m_{90,t})'$.

Time Dependency



Central death rates at representative ages.

Time Dependency

- We require the time-series to be weakly stationary.
- $m(x, t)$ has a clear downward trend, suggesting it is not weakly stationary.
- To solve this problem, we consider the transformation of $r_{x,t} = \frac{m_{x,t+1}}{m_{x,t}}$.
- This may be interpreted as a one-year mortality reduction factor.
- We observe no systematic change in $r_{x,t}$ over time.

Time Dependency

- We require the time-series to be weakly stationary.
- $m(x, t)$ has a clear downward trend, suggesting it is not weakly stationary.
- To solve this problem, we consider the transformation of $r_{x,t} = \frac{m_{x,t+1}}{m_{x,t}}$.
- This may be interpreted as a one-year mortality reduction factor.
- We observe no systematic change in $r_{x,t}$ over time.

Time Dependency

- We require the time-series to be weakly stationary.
- $m(x, t)$ has a clear downward trend, suggesting it is not weakly stationary.
- To solve this problem, we consider the transformation of $r_{x,t} = \frac{m_{x,t+1}}{m_{x,t}}$.
- This may be interpreted as a one-year mortality reduction factor.
- We observe no systematic change in $r_{x,t}$ over time.

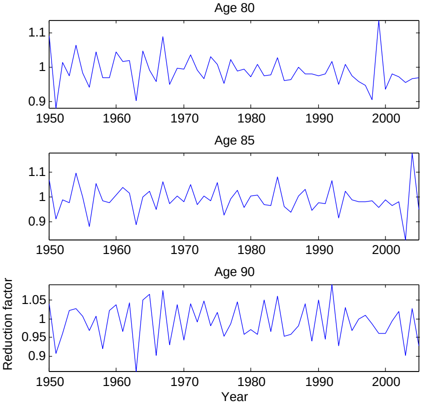
Time Dependency

- We require the time-series to be weakly stationary.
- $m(x,t)$ has a clear downward trend, suggesting it is not weakly stationary.
- To solve this problem, we consider the transformation of $r_{x,t} = \frac{m_{x,t+1}}{m_{x,t}}$.
- This may be interpreted as a one-year mortality reduction factor.
- We observe no systematic change in $r_{x,t}$ over time.

Time Dependency

- We require the time-series to be weakly stationary.
- $m(x, t)$ has a clear downward trend, suggesting it is not weakly stationary.
- To solve this problem, we consider the transformation of $r_{x,t} = \frac{m_{x,t+1}}{m_{x,t}}$.
- This may be interpreted as a one-year mortality reduction factor.
- We observe no systematic change in $r_{x,t}$ over time.

Time Dependency



Mortality reduction factors at representative ages.

Time Dependency

	CCM		CCM
Lag 1	$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & - & \cdot & \cdot & - \\ \cdot & \cdot & - & \cdot & - \\ \cdot & \cdot & \cdot & - & - \\ \cdot & - & \cdot & - & - \end{pmatrix}$	Lag 2	$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & + \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & + \end{pmatrix}$
Lag 3	$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$	Lag 4	$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$

Simplified sample cross-correlation matrices constructed from $r_{x,t}$ at ages: 70, 75, 80, 85, and 90.

The Block Bootstrap

- The naïve bootstrap will lose the serial dependency in the data.
- We use the block bootstrap method (Carlstein, 1986; Künsch, 1989) to retain serial dependency.
- The sample CCMs indicate the cross-correlations taper off as the lag increases.
- Blocks of observations that are separated far enough will be (approximately) uncorrelated.

The Block Bootstrap

- The naïve bootstrap will lose the serial dependency in the data.
- We use the block bootstrap method (Carlstein, 1986; Künsch, 1989) to retain serial dependency.
- The sample CCMs indicate the cross-correlations taper off as the lag increases.
- Blocks of observations that are separated far enough will be (approximately) uncorrelated.

The Block Bootstrap

- The naïve bootstrap will lose the serial dependency in the data.
- We use the block bootstrap method (Carlstein, 1986; Künsch, 1989) to retain serial dependency.
- The sample CCMs indicate the cross-correlations taper off as the lag increases.
- Blocks of observations that are separated far enough will be (approximately) uncorrelated.

The Block Bootstrap

- The naïve bootstrap will lose the serial dependency in the data.
- We use the block bootstrap method (Carlstein, 1986; Künsch, 1989) to retain serial dependency.
- The sample CCMs indicate the cross-correlations taper off as the lag increases.
- Blocks of observations that are separated far enough will be (approximately) uncorrelated.

The Block Bootstrap

- We have 55 vectors of $\mathbf{r}_t$ (1950 – 2004).
- Assuming a block size of 5, we have 51 blocks: $(\mathbf{r}_{1950}, \mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954})$, $(\mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954}, \mathbf{r}_{1955})$, ..., $(\mathbf{r}_{2000}, \mathbf{r}_{2001}, \mathbf{r}_{2002}, \mathbf{r}_{2003}, \mathbf{r}_{2004})$.
- The optimal block size is not always evident.
- Hall et al. (1995) recommend a block size of $n^{1/5}$.
- We use a block size of 2 ($55^{1/5} = 2.23 \approx 2$).

The Block Bootstrap

- We have 55 vectors of $\mathbf{r}_t$ (1950 – 2004).
- Assuming a block size of 5, we have 51 blocks: $(\mathbf{r}_{1950}, \mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954})$, $(\mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954}, \mathbf{r}_{1955})$, …, $(\mathbf{r}_{2000}, \mathbf{r}_{2001}, \mathbf{r}_{2002}, \mathbf{r}_{2003}, \mathbf{r}_{2004})$.
- The optimal block size is not always evident.
- Hall et al. (1995) recommend a block size of $n^{1/5}$.
- We use a block size of 2 ($55^{1/5} = 2.23 \approx 2$).

The Block Bootstrap

- We have 55 vectors of $\mathbf{r}_t$ (1950 – 2004).
- Assuming a block size of 5, we have 51 blocks: $(\mathbf{r}_{1950}, \mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954})$, $(\mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954}, \mathbf{r}_{1955})$, ..., $(\mathbf{r}_{2000}, \mathbf{r}_{2001}, \mathbf{r}_{2002}, \mathbf{r}_{2003}, \mathbf{r}_{2004})$.
- The optimal block size is not always evident.
- Hall et al. (1995) recommend a block size of $n^{1/5}$.
- We use a block size of 2 ($55^{1/5} = 2.23 \approx 2$).

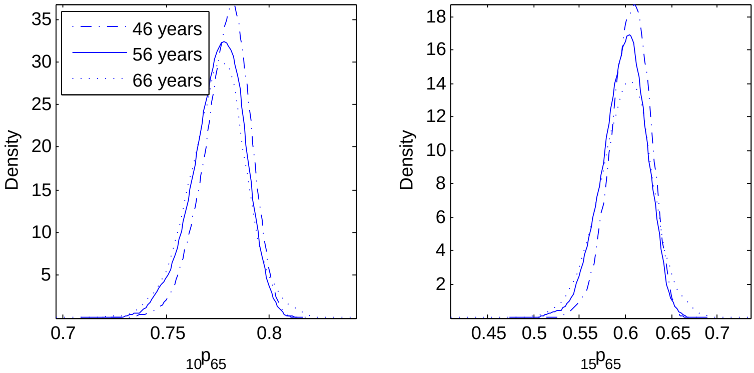
The Block Bootstrap

- We have 55 vectors of $\mathbf{r}_t$ (1950 – 2004).
- Assuming a block size of 5, we have 51 blocks: $(\mathbf{r}_{1950}, \mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954})$, $(\mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954}, \mathbf{r}_{1955})$, ..., $(\mathbf{r}_{2000}, \mathbf{r}_{2001}, \mathbf{r}_{2002}, \mathbf{r}_{2003}, \mathbf{r}_{2004})$.
- The optimal block size is not always evident.
- Hall et al. (1995) recommend a block size of $n^{1/5}$.
- We use a block size of 2 ($55^{1/5} = 2.23 \approx 2$).

The Block Bootstrap

- We have 55 vectors of $\mathbf{r}_t$ (1950 – 2004).
- Assuming a block size of 5, we have 51 blocks: $(\mathbf{r}_{1950}, \mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954})$, $(\mathbf{r}_{1951}, \mathbf{r}_{1952}, \mathbf{r}_{1953}, \mathbf{r}_{1954}, \mathbf{r}_{1955})$, ..., $(\mathbf{r}_{2000}, \mathbf{r}_{2001}, \mathbf{r}_{2002}, \mathbf{r}_{2003}, \mathbf{r}_{2004})$.
- The optimal block size is not always evident.
- Hall et al. (1995) recommend a block size of $n^{1/5}$.
- We use a block size of 2 ($55^{1/5} = 2.23 \approx 2$).

Forecasts of Survival Probabilities



Empirical distributions of the survival probabilities for the cohort aged 65 in year 2005, on the basis of 46, 56, and 66 years of data.

Comparing with Model-Based Methods

	Non-parametric	Lee-Carter	Cairns, Blake and Dowd
${}_{10}p_{65}$	0.7790	0.7755	0.7814
${}_{15}p_{65}$	0.6048	0.6011	0.6135
${}_{20}p_{65}$	0.3999	0.3995	0.4132
${}_{25}p_{65}$	0.2080	0.2039	0.2146

Central estimates of the survival probabilities for the cohort aged 65 in year 2005, on the basis of the non-parametric bootstrap, the Lee-Carter model and the Cairns, Blake and Dowd model.

The BNP/EIB Longevity Bond

- We use the BNP/EIB bond for the price constraint.
- It is 25-year amortising bond, which pays \$50$I(t)$, for $t = 1, \ldots, 25$.
- $I(t)$ is defined as:

$$I(t) = I(t-1)(1 - m_{64+t,2002+t}), \qquad t = 1, 2, \ldots, 25,$$

where
  - $I(0) = 1$,
  - $m_{x,t}$ is the crude central death rate for the E&W male population at age x and in year t.

The BNP/EIB Longevity Bond

- We use the BNP/EIB bond for the price constraint.
- It is 25-year amortising bond, which pays $\$50I(t)$, for $t = 1, \ldots, 25$.
- $I(t)$ is defined as:

$$I(t) = I(t-1)(1 - m_{64+t,2002+t}), \qquad t = 1, 2, \ldots, 25,$$

where
- $I(0) = 1$,
- $m_{x,t}$ is the crude central death rate for the E&W male population at age x and in year t.

The BNP/EIB Longevity Bond

- We use the BNP/EIB bond for the price constraint.
- It is 25-year amortising bond, which pays $\$50I(t)$, for $t = 1, \ldots, 25$.
- $I(t)$ is defined as:

$$I(t) = I(t-1)(1 - m_{64+t,2002+t}), \qquad t = 1, 2, \ldots, 25,$$

where

  - $I(0) = 1$,
  - $m_{x,t}$ is the crude central death rate for the E&W male population at age x and in year t.

The BNP/EIB Longevity Bond

- We use the BNP/EIB bond for the price constraint.
- It is 25-year amortising bond, which pays $\$50I(t)$, for $t = 1, \dots, 25$.
- $I(t)$ is defined as:

$$I(t) = I(t-1)(1 - m_{64+t,2002+t}), \qquad t = 1, 2, \dots, 25,$$

where

- $I(0) = 1$,
- $m_{x,t}$ is the crude central death rate for the E&W male population at age x and in year t.

The BNP/EIB Longevity Bond

- We use the BNP/EIB bond for the price constraint.
- It is 25-year amortising bond, which pays $\$50I(t)$, for $t = 1, \ldots, 25$.
- $I(t)$ is defined as:

$$I(t) = I(t-1)(1 - m_{64+t,2002+t}), \qquad t = 1, 2, \ldots, 25,$$

  where
  - $I(0) = 1$,
  - $m_{x,t}$ is the crude central death rate for the E&W male population at age x and in year t.

The BNP/EIB Longevity Bond

- The issue price was determined by discounting at LIBOR minus 35 basis points the anticipated coupon payments.
- The time-0 value of the bond is £561.

The BNP/EIB Longevity Bond

- The issue price was determined by discounting at LIBOR minus 35 basis points the anticipated coupon payments.
- The time-0 value of the bond is £561.

Step (1)

- Generate a number, say N, of equally probable mortality scenarios.

- From each scenario, calculate the longevity index $I(t)$ at $t = 1, 2, \ldots, 25$.

- The time-0 value of the BNP/EIB bond in the jth scenario is

$$v(\omega_j) = 50 \times \sum_{t=1}^{25} B(0, t) I(t, \omega_j),$$

where $I(t, \omega_j)$ be the index value at time t in the jth scenario, and $B(0, t)$ is the time-0 price of a risk-free zero-coupon bond that pays £1 at time t.

Step (1)

- Generate a number, say N, of equally probable mortality scenarios.
- From each scenario, calculate the longevity index $I(t)$ at $t = 1, 2, \ldots, 25$.
- The time-0 value of the BNP/EIB bond in the jth scenario is

$$v(\omega_j) = 50 \times \sum_{t=1}^{25} B(0,t) I(t, \omega_j),$$

where $I(t, \omega_j)$ be the index value at time t in the jth scenario, and $B(0,t)$ is the time-0 price of a risk-free zero-coupon bond that pays £1 at time t.

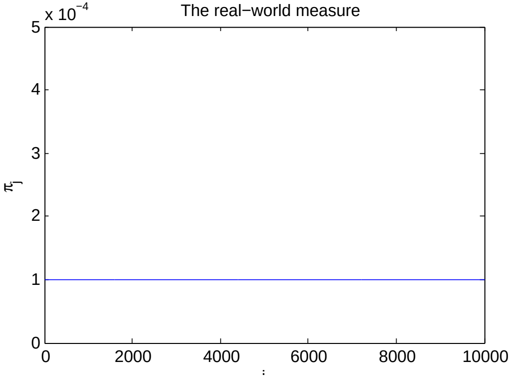
Step (1)

- Generate a number, say N, of equally probable mortality scenarios.
- From each scenario, calculate the longevity index $I(t)$ at $t = 1, 2, \ldots, 25$.
- The time-0 value of the BNP/EIB bond in the jth scenario is

$$v(\omega_j) = 50 \times \sum_{t=1}^{25} B(0,t) I(t, \omega_j),$$

where $I(t, \omega_j)$ be the index value at time t in the jth scenario, and $B(0,t)$ is the time-0 price of a risk-free zero-coupon bond that pays £1 at time t.

Real World Probability Measure, π_j



Step (2)

- Let π_j^* be the probability associated with $v(\omega_j)$ under Q.
- We require $\sum_{j=1}^{N} v(\omega_j)\pi_j^* = 561$ and $\sum_{j=1}^{N} \pi_j^* = 1$.
- We minimize the KLIC as follows:

$$L = \sum_{j=1}^{N} \pi_j^* \ln \pi_j^* - \lambda_0 \left(\sum_{j=1}^{N} \pi_j^* - 1 \right) - \lambda_1 \sum_{j=1}^{N} \left(v(\omega_j)\pi_j^* - 561 \right).$$

Step (2)

- Let π_j^* be the probability associated with $v(\omega_j)$ under Q.
- We require $\sum_{j=1}^{N} v(\omega_j)\pi_j^* = 561$ and $\sum_{j=1}^{N} \pi_j^* = 1$.
- We minimize the KLIC as follows:

$$L = \sum_{j=1}^{N} \pi_j^* \ln \pi_j^* - \lambda_0 \left(\sum_{j=1}^{N} \pi_j^* - 1 \right) - \lambda_1 \sum_{j=1}^{N} \left(v(\omega_j)\pi_j^* - 561 \right).$$

Step (2)

- Let π_j^* be the probability associated with $v(\omega_j)$ under Q.
- We require $\sum_{j=1}^{N} v(\omega_j)\pi_j^* = 561$ and $\sum_{j=1}^{N} \pi_j^* = 1$.
- We minimize the KLIC as follows:

$$L = \sum_{j=1}^{N} \pi_j^* \ln \pi_j^* - \lambda_0 \left(\sum_{j=1}^{N} \pi_j^* - 1 \right) - \lambda_1 \sum_{j=1}^{N} \left(v(\omega_j)\pi_j^* - 561 \right).$$

Step (2), Continued

- Let $\tilde{\pi}_j^*$, $j = 1, 2, \ldots, N$, be the solution.
- We have

$$\tilde{\pi}_j^* = \frac{\exp(\lambda_1 v(\omega_j))}{\sum_{j=1}^N \exp(\lambda_1 v(\omega_j))}, \quad j = 1, 2, \ldots, N.$$

$$\lambda_1 = \arg\min_{\gamma} \sum_{j=1}^N \exp(\gamma(v(\omega_j) - 561)).$$

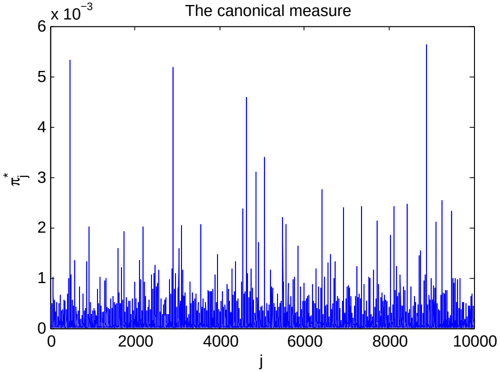
Step (2), Continued

- Let $\tilde{\pi}_j^*$, $j = 1, 2, \ldots, N$, be the solution.
- We have

$$\tilde{\pi}_j^* = \frac{\exp(\lambda_1 v(\omega_j))}{\sum_{j=1}^{N} \exp(\lambda_1 v(\omega_j))}, \quad j = 1, 2, \ldots, N.$$

$$\lambda_1 = \arg\min_{\gamma} \sum_{j=1}^{N} \exp(\gamma(v(\omega_j) - 561)).$$

The Canonical Measure, π_j^*



Incorporating More Prices

- What if more market prices are available?
- The method can be extended to incorporate additional primary securities.
- Assume the ith securitiy has a time-0 price of V_i and a discounted payoff of $v_i(\omega_j)$ in the jth scenario.
- To price m securities correctly, we require

$$\sum_{j=1}^{N} v_i(\omega_j)\pi_j^* = V_i, \quad i = 1, 2, \ldots, m. \tag{3}$$

Incorporating More Prices

- We minimize the KLIC subject to the m constraints and $\sum_{j=1}^{N} \pi_j^* = 1$.

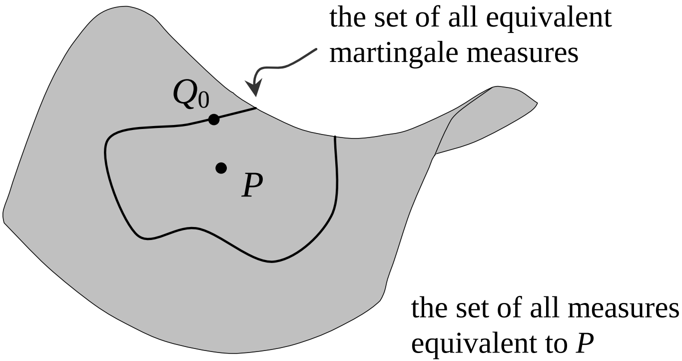
- It can be shown that the resulting canonical measure $\tilde{\pi}_j^*$, $j = 1, 2, \ldots, N$ is

$$\tilde{\pi}_j^* = \frac{\exp(\sum_{i=1}^{m} \lambda_i v(\omega_j))}{\sum_{j=1}^{N} \exp(\sum_{i=1}^{m} \lambda_i v(\omega_j))}, \quad j = 1, 2, \ldots, N,$$

where $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)'$ can be expressed as

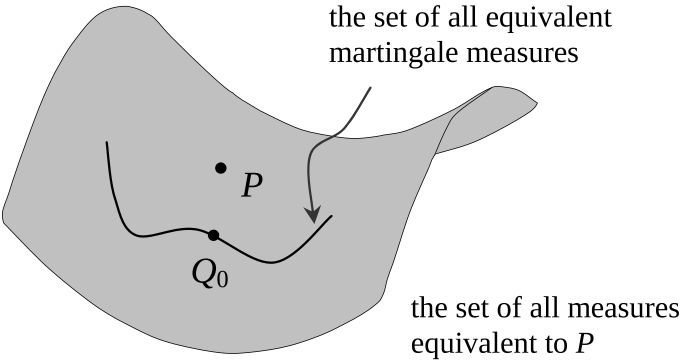
$$\vec{\lambda} = \underset{\gamma_1, \ldots, \gamma_m}{\arg\min} \sum_{j=1}^{N} \exp\left(\sum_{i=1}^{m} \gamma_i (v_i(\omega_j) - V_i) \right).$$

With One Primary Security, $m = 1$



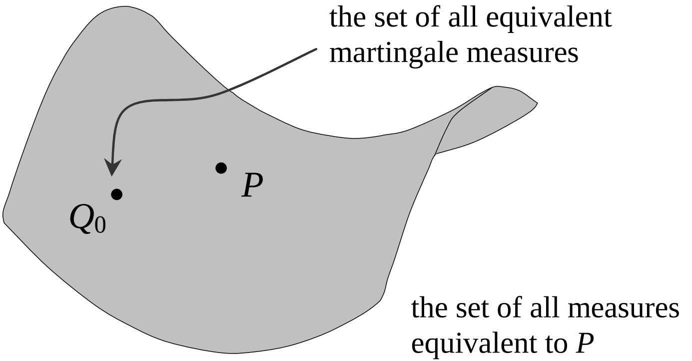
The canonical measure Q_0 when $m = 1$.

With Two Primary Securities, $m = 2$



The canonical measure Q_0 when $m = 2$.

With Infinitely Many Primary Securities, $m \to \infty$



The canonical measure Q_0 when $m \to \infty$.

Pricing Vanilla Survivor Swaps

- We consider vanilla survivor swaps with a fixed proportional premium θ and a fixed maturity T.
- At $t = 1, 2, \ldots, T$, the fixed-payer pays a preset amount of $(1 + \theta)K(t)$.
- The fixed-reciever pays a random amount of $S(t)$, which is linked to the realized survival function of the reference population.
- The reference population is the same as that of the BNP/EIB longevity bond.

Pricing Vanilla Survivor Swaps

- We set

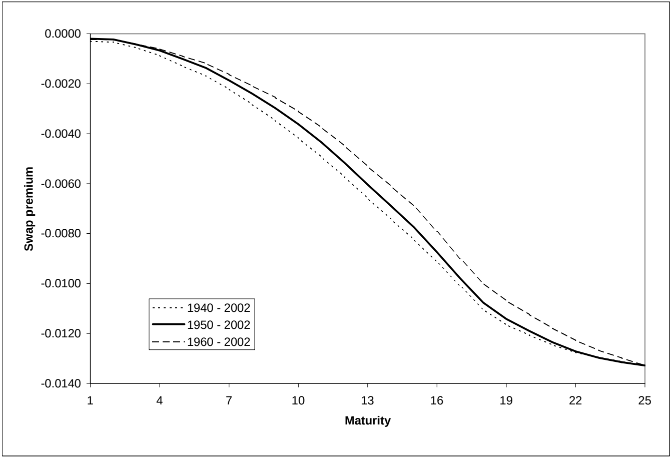
$$S(t) = S(t-1)(1 - q_{64+t,2002+t}), \quad t = 1, 2, \ldots, T,$$

where $S(0) = 1$, and $q_{x,t}$ is the realized death probability.

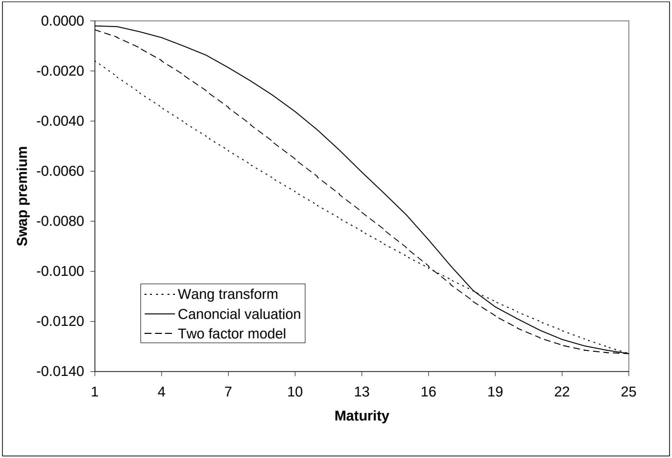
- We set $K(t)$ to the projected survival function for the reference population, on the basis of GAD's projection.

- $K(t)$ for declines over time.

The Calculated Swap Premium



Comparing with Other Pricing Methods



Conclusions

- The pricing framework is reasonably robust relative to the amount of data used.
- It avoids model risk and parameter risk.
- Additional prices can be incorporated into the canonical measure easily.
- Due to its non-parametric nature, our framework can be applied to reference populations with limited volume of data available.

Q&A