Coherent mortality forecasting using functional time series models
Mortality rates

Australia: male mortality (1921)
Mortality rates

Australia: male death rates (1970)

Log death rate

Age
Mortality rates

Australia: male death rates (1990)
Mortality rates

Australia: male death rates (1921–2009)

Log death rate vs. Age

Coherent mortality forecasting
Mortality rates

Australia: male death rates (1921–2009)

Log death rate vs. Age
Mortality rates

Australia: mortality sex ratio (1921–2009)

Sex ratio of rates: M/F

Age

Coherent mortality forecasting
Outline

1. Functional forecasting
2. Forecasting groups
3. Coherent cohort life expectancy forecasts
4. Conclusions
Outline

1 Functional forecasting

2 Forecasting groups

3 Coherent cohort life expectancy forecasts

4 Conclusions
Some notation

Let $y_{t,x}$ be the observed (smoothed) data in period $t$ at age $x$, $t = 1, \ldots, n$.

$$ y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x} $$

$$ f_t(x) = \mu(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + e_t(x) $$

- Estimate $f_t(x)$ using penalized regression splines.
- Estimate $\mu(x)$ as mean $f_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using functional principal components.

$\varepsilon_{t,x} \sim \text{N}(0, 1)$ and $e_t(x) \sim \text{N}(0, \nu(x))$. 
Let $y_{t,x}$ be the observed (smoothed) data in period $t$ at age $x$, $t = 1, \ldots, n$.

$$y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x}$$

$$f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)$$

- Estimate $f_t(x)$ using penalized regression splines.
- Estimate $\mu(x)$ as mean $f_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using functional principal components.
- $\varepsilon_{t,x} \sim \text{iid } N(0,1)$ and $e_t(x) \sim \text{iid } N(0, v(x))$. 
Let \( y_{t,x} \) be the observed (smoothed) data in period \( t \) at age \( x \), \( t = 1, \ldots, n \).

\[
y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x}
\]

\[
f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

- Estimate \( f_t(x) \) using penalized regression splines.
- **Estimate** \( \mu(x) \) as mean \( f_t(x) \) across years.
- Estimate \( \beta_{t,k} \) and \( \phi_k(x) \) using functional principal components.
- \( \varepsilon_{t,x} \overset{iid}{\sim} N(0,1) \) and \( e_t(x) \overset{iid}{\sim} N(0, \nu(x)) \).
Some notation

Let \( y_{t,x} \) be the observed (smoothed) data in period \( t \) at age \( x \), \( t = 1, \ldots, n \).

\[
y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x}
\]

\[
f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

- Estimate \( f_t(x) \) using penalized regression splines.
- Estimate \( \mu(x) \) as mean \( f_t(x) \) across years.
- Estimate \( \beta_{t,k} \) and \( \phi_k(x) \) using functional principal components.
- \( \varepsilon_{t,x} \overset{iid}{\sim} N(0,1) \) and \( e_t(x) \overset{iid}{\sim} N(0, \nu(x)) \).
Let $y_{t,x}$ be the observed (smoothed) data in period $t$ at age $x$, $t = 1, \ldots, n$.

$$y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x}$$

$$f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)$$

- Estimate $f_t(x)$ using penalized regression splines.
- Estimate $\mu(x)$ as mean $f_t(x)$ across years.
- Estimate $\beta_{t,k}$ and $\phi_k(x)$ using functional principal components.
- $\varepsilon_{t,x} \overset{iid}{\sim} N(0, 1)$ and $e_t(x) \overset{iid}{\sim} N(0, v(x))$. 
Australian male mortality model

Coherent mortality forecasting

Functional forecasting
Australian male mortality model

Coherent mortality forecasting

Functional forecasting
Functional time series model

\[
y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x}
\]

\[
f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

- The eigenfunctions \( \phi_k(x) \) show the main regions of variation.
- The scores \( \{\beta_{t,k}\} \) are uncorrelated by construction. So we can forecast each \( \beta_{t,k} \) using a univariate time series model.
- Univariate ARIMA models can be used for forecasting.
The eigenfunctions \( \phi_k(x) \) show the main regions of variation.

The scores \( \{\beta_{t,k}\} \) are uncorrelated by construction. So we can forecast each \( \beta_{t,k} \) using a univariate time series model.

Univariate ARIMA models can be used for forecasting.
The eigenfunctions $\phi_k(x)$ show the main regions of variation.

The scores $\{\beta_{t,k}\}$ are uncorrelated by construction. So we can forecast each $\beta_{t,k}$ using a univariate time series model.

Univariate ARIMA models can be used for forecasting.
The eigenfunctions $\phi_k(x)$ show the main regions of variation.

The scores $\{\beta_{t,k}\}$ are uncorrelated by construction. So we can forecast each $\beta_{t,k}$ using a univariate time series model.

Univariate ARIMA models can be used for forecasting.
Forecasts

\[ y_{t,x} = f_t(x) + \sigma_t(x) \epsilon_{t,x} \]

\[ f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]
Forecasts

\[ y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x} \]

\[ f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

\[ \mathbb{E}[y_{n+h,x} \mid \mathbf{y}] = \hat{\mu}(x) + \sum_{k=1}^{K} \hat{\beta}_{n+h,k} \hat{\phi}_k(x) \]

\[ \text{Var}[y_{n+h,x} \mid \mathbf{y}] = \hat{\sigma}_\mu^2(x) + \sum_{k=1}^{K} v_{n+h,k} \hat{\phi}_k^2(x) + \sigma_t^2(x) + v(x) \]

where \( v_{n+h,k} = \text{Var}(\beta_{n+h,k} \mid \beta_{1,k}, \ldots, \beta_{n,k}) \)

and \( \mathbf{y} = [y_{1,x}, \ldots, y_{n,x}] \).
Forecasting the PC scores

Coherent mortality forecasting

Functional forecasting
Forecasts of \( f_t(x) \)

Australia: male death rates (1921–2009)
Forecasts of $f_t(x)$

Australia: male death rates (1921–2009)
Forecasts of $f_t(x)$

Australia: male death rates forecasts (2010–2059)
Forecasts of $f_t(x)$

Australia: male death rates forecasts (2010 and 2059)

Log death rate vs. Age

80% prediction intervals

Coherent mortality forecasting

Functional forecasting
Forecasts of mortality sex ratio

Australia: mortality sex ratio forecasts

Coherent mortality forecasting
Functional forecasting
Forecasts of mortality sex ratio

Australia: mortality sex ratio forecasts

Male and female mortality rate forecasts are diverging.
The problem

Let \( f_{t,j}(x) \) be the smoothed mortality rate for age \( x \) in group \( j \) in year \( t \).

- Groups may be males and females.
- Groups may be states within a country.
- Expected that groups will behave similarly.
- Coherent forecasts do not diverge over time.
- Existing functional models do not impose coherence.
The problem

Let \( f_{t,j}(x) \) be the smoothed mortality rate for age \( x \) in group \( j \) in year \( t \).

- Groups may be males and females.
- Groups may be states within a country.
- Expected that groups will behave similarly.
- Coherent forecasts do not diverge over time.
- Existing functional models do not impose coherence.
Let $f_{t,j}(x)$ be the smoothed mortality rate for age $x$ in group $j$ in year $t$.

- Groups may be males and females.
- Groups may be states within a country.
- Expected that groups will behave similarly.
- Coherent forecasts do not diverge over time.
- Existing functional models do not impose coherence.
The problem

Let \( f_{t,j}(x) \) be the smoothed mortality rate for age \( x \) in group \( j \) in year \( t \).

- Groups may be males and females.
- Groups may be states within a country.
- **Expected that groups will behave similarly.**
- Coherent forecasts do not diverge over time.
- Existing functional models do not impose coherence.
The problem

Let $f_{t,j}(x)$ be the smoothed mortality rate for age $x$ in group $j$ in year $t$.

- Groups may be males and females.
- Groups may be states within a country.
- Expected that groups will behave similarly.
- Coherent forecasts do not diverge over time.

- Existing functional models do not impose coherence.
Let $f_{t,j}(x)$ be the smoothed mortality rate for age $x$ in group $j$ in year $t$.

- Groups may be males and females.
- Groups may be states within a country.
- Expected that groups will behave similarly.
- **Coherent** forecasts do not diverge over time.
- Existing functional models do not impose coherence.
Forecasting the coefficients

\[
y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x}
\]

\[
f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

- We use ARIMA models for each coefficient \( \{\beta_{1,j,k}, \ldots, \beta_{n,j,k}\} \).
- The ARIMA models are non-stationary for the first few coefficients (\( k = 1, 2 \)).
- Non-stationary ARIMA forecasts will diverge. Hence the mortality forecasts are not coherent.
Forecasting the coefficients

\[ y_{t,x} = f_t(x) + \sigma_t(x)\varepsilon_{t,x} \]
\[ f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

- We use ARIMA models for each coefficient \( \{\beta_{1,j,k}, \ldots, \beta_{n,j,k}\} \).
- The ARIMA models are non-stationary for the first few coefficients \((k = 1, 2)\)
- Non-stationary ARIMA forecasts will diverge. Hence the mortality forecasts are not coherent.
Forecasting the coefficients

\[ y_{t,x} = f_t(x) + \sigma_t(x) \varepsilon_{t,x} \]

\[ f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

- We use ARIMA models for each coefficient \( \{\beta_{1,j,k}, \ldots, \beta_{n,j,k}\} \).
- The ARIMA models are non-stationary for the first few coefficients (\( k = 1, 2 \))
- Non-stationary ARIMA forecasts will diverge.
  Hence the mortality forecasts are not coherent.
Male fts model

Coherent mortality forecasting

Forecasting groups
Female fts model

Coherent mortality forecasting

Forecasting groups
Australian mortality forecasts

(a) Males

(b) Females

Coherent mortality forecasting
Forecasting groups
Key idea
Model the geometric mean and the mortality ratio instead of the individual rates for each sex separately.

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

- Product and ratio are approximately independent.
- Ratio should be stationary (for coherence) but product can be non-stationary.
Mortality product and ratios

Key idea

Model the geometric mean and the mortality ratio instead of the individual rates for each sex separately.

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{\frac{f_{t,M}(x)}{f_{t,F}(x)}}. \]

- Product and ratio are approximately independent.
- Ratio should be stationary (for coherence) but product can be non-stationary.
Mortality product and ratios

Key idea

Model the geometric mean and the mortality ratio instead of the individual rates for each sex separately.

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

- Product and ratio are approximately independent
- Ratio should be stationary (for coherence) but product can be non-stationary.
Mortality product and ratios

Key idea

Model the geometric mean and the mortality ratio instead of the individual rates for each sex separately.

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

- Product and ratio are approximately independent
- Ratio should be stationary (for coherence) but product can be non-stationary.
Australia: mortality sex ratio (1921–2009)
Model product and ratios

\[ p_t(x) = \sqrt{f_{t,M}(x) f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[
\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]
\[
\log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_\ell(x) + w_t(x).
\]

- \( \{\gamma_{t,\ell}\} \) restricted to be stationary processes: either ARMA\((p, q)\) or ARFIMA\((p, d, q)\).
- No restrictions for \( \beta_{t,1}, \ldots, \beta_{t,K}\).
- Forecasts:
  \[ f_{n+h|n,M}(x) = p_{n+h|n}(x)r_{n+h|n}(x) \]
  \[ f_{n+h|n,F}(x) = p_{n+h|n}(x)/r_{n+h|n}(x). \]
Model product and ratios

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[
\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + e_t(x)
\]

\[
\log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell}\psi_\ell(x) + w_t(x).
\]

- \( \{\gamma_{t,\ell}\} \) restricted to be stationary processes: either ARMA\((p, q)\) or ARFIMA\((p, d, q)\).
- No restrictions for \( \beta_{t,1}, \ldots, \beta_{t,K} \).
- Forecasts:
  \[
  f_{n+h|n,M}(x) = p_{n+h|n}(x)r_{n+h|n}(x)
  \]
  \[
  f_{n+h|n,F}(x) = p_{n+h|n}(x)/r_{n+h|n}(x).
  \]
Model product and ratios

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[ \log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

\[ \log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_{\ell}(x) + w_t(x). \]

- \{\gamma_{t,\ell}\} restricted to be stationary processes: either ARMA\((p, q)\) or ARFIMA\((p, d, q)\).
- No restrictions for \(\beta_{t,1}, \ldots, \beta_{t,K}\).
- Forecasts:
  \[ f_{n+h|n,M}(x) = p_{n+h|n}(x)r_{n+h|n}(x) \]
  \[ f_{n+h|n,F}(x) = p_{n+h|n}(x)/r_{n+h|n}(x). \]
Model product and ratios

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[ \log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

\[ \log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_\ell(x) + w_t(x). \]

- \{\gamma_{t,\ell}\} restricted to be stationary processes: either ARMA(p, q) or ARFIMA(p, d, q).
- No restrictions for \(\beta_{t,1}, \ldots, \beta_{t,K}\).
- Forecasts: \(f_{n+h|n,M}(x) = p_{n+h|n}(x)r_{n+h|n}(x)\)
  \(f_{n+h|n,F}(x) = p_{n+h|n}(x)/r_{n+h|n}(x).\)
Product model
Ratio model

Coherent mortality forecasting

Forecasting groups

24
Product forecasts

Coherent mortality forecasting

Forecasting groups
Coherent forecasts

(a) Males

(b) Females

- Log death rate vs. Age for males and females.
Ratio forecasts

Independent forecasts

Coherent forecasts

Sex ratio of rates: M/F

Age
Life expectancy forecasts

Year
Age
1920 1960 2000 2040
70 75 80 85 90 95

Life expectancy difference: F−M

Year
Number of years
1960 1980 2000 2020
4 5 6 7
Coherent forecasts for $J$ groups

\[ p_t(x) = \left[ f_{t,1}(x)f_{t,2}(x)\cdots f_{t,J}(x) \right]^{1/J} \]

and

\[ r_{t,j}(x) = \frac{f_{t,j}(x)}{p_t(x)}, \]

\[ \log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \]

\[ \log[r_{t,j}(x)] = \mu_{r,j}(x) + \sum_{l=1}^{L} \gamma_{t,l,j} \psi_{l,j}(x) + w_{t,j}(x). \]

- $p_t(x)$ and all $r_{t,j}(x)$ are approximately independent.
- Ratios satisfy constraint:
  \[ r_{t,1}(x)r_{t,2}(x)\cdots r_{t,J}(x) = 1. \]
  \[ \log[r_t(x)] = \log[p_t(x)r_{t,1}(x)r_{t,2}(x)\cdots r_{t,J}(x)]. \]
Coherent forecasts for $J$ groups

$p_t(x) = \left[ f_{t,1}(x)f_{t,2}(x) \cdots f_{t,J}(x) \right]^{1/J}$

and

$r_{t,j}(x) = \frac{f_{t,j}(x)}{p_t(x)},$

$\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)$

$\log[r_{t,j}(x)] = \mu_{r,j}(x) + \sum_{l=1}^{L} \gamma_{t,l,j} \psi_{l,j}(x) + w_{t,j}(x).$

$p_t(x)$ and all $r_{t,j}(x)$ are approximately independent.

Ratios satisfy constraint $r_{t,1}(x)r_{t,2}(x) \cdots r_{t,J}(x) = 1.$
Coherent forecasts for $J$ groups

$p_t(x) = \left[ f_{t,1}(x)f_{t,2}(x) \cdots f_{t,J}(x) \right]^{1/J}$

and

$r_{t,j}(x) = f_{t,j}(x)/p_t(x),$

\[\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \epsilon_t(x)\]

\[\log[r_{t,j}(x)] = \mu_{r,j}(x) + \sum_{l=1}^{L} \gamma_{t,l,j} \psi_{l,j}(x) + w_{t,j}(x)\].

- $p_t(x)$ and all $r_{t,j}(x)$ are approximately independent.
- Ratios satisfy constraint $r_{t,1}(x)r_{t,2}(x) \cdots r_{t,J}(x) = 1.$
- $\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)]$
Coherent forecasts for $J$ groups

$$p_t(x) = [f_{t,1}(x)f_{t,2}(x)\cdots f_{t,J}(x)]^{1/J}$$

and

$$r_{t,j}(x) = f_{t,j}(x)/p_t(x),$$

$$\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)$$

$$\log[r_{t,j}(x)] = \mu_{r,j}(x) + \sum_{l=1}^{L} \gamma_{t,l,j} \psi_{l,j}(x) + w_{t,j}(x).$$

- $p_t(x)$ and all $r_{t,j}(x)$ are approximately independent.
- Ratios satisfy constraint $r_{t,1}(x)r_{t,2}(x)\cdots r_{t,J}(x) = 1.$
- $\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)].$
Coherent forecasts for $J$ groups

$$p_t(x) = [f_{t,1}(x)f_{t,2}(x) \cdots f_{t,J}(x)]^{1/J}$$

and

$$r_{t,j}(x) = f_{t,j}(x)/p_t(x),$$

$$\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)$$

$$\log[r_{t,j}(x)] = \mu_{r,j}(x) + \sum_{l=1}^{L} \gamma_{t,l,j} \psi_{l,j}(x) + w_{t,j}(x).$$

- $p_t(x)$ and all $r_{t,j}(x)$ are approximately independent.
- Ratios satisfy constraint $r_{t,1}(x)r_{t,2}(x) \cdots r_{t,J}(x) = 1$.
- $\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)]$
Coherent forecasts for $J$ groups

$$
\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}]
$$

$$
= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j} \psi_{\ell,j}(x) + z_{t,j}(x)
$$

- $\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)$ is group mean.
- $z_{t,j}(x) = e_t(x) + w_{t,j}(x)$ is error term.
- $\{\gamma_{t,\ell}\}$ restricted to be stationary processes: either ARMA($p, q$) or ARFIMA($p, d, q$).

No restrictions for $\beta_{t,1}, \ldots, \beta_{t,K}$.
Coherent forecasts for $J$ groups

$log[f_{t,j}(x)] = log[p_t(x)r_{t,j}(x)] = log[p_t(x)] + log[r_{t,j}]$

$$= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j} \psi_{\ell,j}(x) + z_{t,j}(x)$$

- $\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)$ is group mean
- $z_{t,j}(x) = e_t(x) + w_{t,j}(x)$ is error term.
- $\{\gamma_{t,\ell}\}$ restricted to be stationary processes: either ARMA($p, q$) or ARFIMA($p, d, q$).
- No restrictions for $\beta_{t,1}, \ldots, \beta_{t,K}$. 
Coherent forecasts for $J$ groups

\[
\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}]
\]

\[
= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j} \psi_{\ell,j}(x) + z_{t,j}(x)
\]

- $\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)$ is group mean
- $z_{t,j}(x) = e_t(x) + w_{t,j}(x)$ is error term.
- \{\gamma_{t,\ell}\} restricted to be stationary processes: either ARMA($p$, $q$) or ARFIMA($p$, $d$, $q$).
- No restrictions for $\beta_{t,1}, \ldots, \beta_{t,K}$. 

Coherent forecasts for $J$ groups

$$\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}]$$

$$= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j} \psi_{\ell,j}(x) + z_{t,j}(x)$$

- $\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)$ is group mean
- $z_{t,j}(x) = e_t(x) + w_{t,j}(x)$ is error term.
- $\{\gamma_{t,\ell}\}$ restricted to be stationary processes: either ARMA($p, q$) or ARFIMA($p, d, q$).
- No restrictions for $\beta_{t,1}, \ldots, \beta_{t,K}$.
Coherent forecasts for J groups

\[
\log[f_{t,j}(x)] = \log[p_t(x)r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}]
\]

\[
= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j} \psi_{\ell,j}(x) + z_{t,j}(x)
\]

- \(\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)\) is group mean
- \(z_{t,j}(x) = e_t(x) + w_{t,j}(x)\) is error term.
- \(\{\gamma_{t,\ell}\}\) restricted to be stationary processes: either ARMA\((p, q)\) or ARFIMA\((p, d, q)\).
- No restrictions for \(\beta_{t,1}, \ldots, \beta_{t,K}\).
Li & Lee (Demography, 2005) method is a special case of our approach.

\[ f_{t,j}(x) = \mu_j(x) + \beta_t \phi(x) + \gamma_{t,j} \psi_j(x) + e_{t,j}(x) \]

where \( f \) is unsmoothed log mortality rate, \( \beta_t \) is a random walk with drift and \( \gamma_{t,j} \) is AR(1) process.

- No smoothing.
- Only one basis function for each part.
- Random walk with drift very limiting.
- AR(1) very limiting.
- The \( \gamma_{t,j} \) coefficients will be highly correlated with each other, and so independent models are not appropriate.
Li-Lee method

Li & Lee (Demography, 2005) method is a special case of our approach.

\[ f_{t,j}(x) = \mu_j(x) + \beta_t \phi(x) + \gamma_{t,j} \psi_j(x) + e_{t,j}(x) \]

where \( f \) is unsmoothed log mortality rate, \( \beta_t \) is a random walk with drift and \( \gamma_{t,j} \) is AR(1) process.

- No smoothing.
- Only one basis function for each part,
  - Random walk with drift very limiting.
  - AR(1) very limiting.
- The \( \gamma_{t,j} \) coefficients will be highly correlated with each other, and so independent models are not appropriate.
Li-Lee method

Li & Lee (Demography, 2005) method is a special case of our approach.

\[ f_{t,j}(x) = \mu_j(x) + \beta_t \phi(x) + \gamma_{t,j} \psi_j(x) + e_{t,j}(x) \]

where \( f \) is unsmoothed log mortality rate, \( \beta_t \) is a random walk with drift and \( \gamma_{t,j} \) is AR(1) process.

- No smoothing.
- Only one basis function for each part,
- Random walk with drift very limiting.
- AR(1) very limiting.
- The \( \gamma_{t,j} \) coefficients will be highly correlated with each other, and so independent models are not appropriate.
Li & Lee (Demography, 2005) method is a special case of our approach.

\[ f_{t,j}(x) = \mu_j(x) + \beta_t \phi(x) + \gamma_{t,j} \psi_j(x) + e_{t,j}(x) \]

where \( f \) is unsmoothed log mortality rate, \( \beta_t \) is a random walk with drift and \( \gamma_{t,j} \) is AR(1) process.

- No smoothing.
- Only one basis function for each part,
- Random walk with drift very limiting.
- AR(1) very limiting.
- The \( \gamma_{t,j} \) coefficients will be highly correlated with each other, and so independent models are not appropriate.
Li-Lee method

Li & Lee (Demography, 2005) method is a special case of our approach.

\[ f_{t,j}(x) = \mu_j(x) + \beta_t \phi(x) + \gamma_{t,j} \psi_j(x) + e_{t,j}(x) \]

where \( f \) is unsmoothed log mortality rate, \( \beta_t \) is a random walk with drift and \( \gamma_{t,j} \) is AR(1) process.

- No smoothing.
- Only one basis function for each part,
- Random walk with drift very limiting.
- AR(1) very limiting.
- The \( \gamma_{t,j} \) coefficients will be highly correlated with each other, and so independent models are not appropriate.
1 Functional forecasting

2 Forecasting groups

3 Coherent cohort life expectancy forecasts

4 Conclusions
Life expectancy calculation

Using standard life table calculations:

For $x = 0, 1, \ldots, \omega - 1$:

$$q_x = \frac{m_x}{1 + (1 - a_x)m_x}$$

$$\ell_{x+1} = \ell_x (1 - q_x)$$

$$L_x = \ell_x [1 - q_x (1 - a_x)]$$

$$T_x = L_x + L_{x+1} + \cdots + L_{\omega-1} + L_{\omega+}$$

$$e_x = \frac{T_x}{L_x}$$

where $a_x = 0.5$ for $x \geq 1$ and $a_0$ taken from Coale et al (1983).

$q_{\omega+} = 1$, $L_{\omega+} = l_x/m_x$, and $T_{\omega+} = L_{\omega+}$.

- Period life expectancy: let $m_x = m_{x,t}$ for some year $t$.
- Cohort life expectancy: let $m_x = m_{x,t|x}$ for birth cohort in year $t$. 

Coherent mortality forecasting

Coherent cohort life expectancy forecasts
Life expectancy calculation

Using standard life table calculations:

For \( x = 0, 1, \ldots, \omega - 1 \):

\[
q_x = \frac{m_x}{(1 + (1 - a_x)m_x)}
\]

\[
\ell_{x+1} = \ell_x (1 - q_x)
\]

\[
L_x = \ell_x [1 - q_x (1 - a_x)]
\]

\[
T_x = L_x + L_{x+1} + \cdots + L_{\omega-1} + L_{\omega+}
\]

\[
e_x = \frac{T_x}{L_x}
\]

where \( a_x = 0.5 \) for \( x \geq 1 \) and \( a_0 \) taken from Coale et al (1983).

\( q_{\omega+} = 1, L_{\omega+} = l_x/m_x \), and \( T_{\omega+} = L_{\omega+} \).

- Period life expectancy: let \( m_x = m_{x,t} \) for some year \( t \).
- Cohort life expectancy: let \( m_x = m_{x,t+x} \) for birth cohort in year \( t \).
Life expectancy calculation

Using standard life table calculations:

For $x = 0, 1, \ldots, \omega - 1$:

$$q_x = \frac{m_x}{1 + (1 - a_x)m_x}$$

$$\ell_{x+1} = \ell_x (1 - q_x)$$

$$L_x = \ell_x [1 - q_x (1 - a_x)]$$

$$T_x = L_x + L_{x+1} + \cdots + L_{\omega-1} + L_{\omega+}$$

$$e_x = \frac{T_x}{L_x}$$

where $a_x = 0.5$ for $x \geq 1$ and $a_0$ taken from Coale et al (1983).

$q_{\omega+} = 1$, $L_{\omega+} = l_x/m_x$, and $T_{\omega+} = L_{\omega+}$.

- **Period life expectancy**: let $m_x = m_{x,t}$ for some year $t$.
- **Cohort life expectancy**: let $m_x = m_{x,t+x}$ for birth cohort in year $t$.
Life expectancy calculation

Using standard life table calculations:

For \( x = 0, 1, \ldots, \omega - 1 \):

\[
q_x = \frac{m_x}{1 + (1 - a_x)m_x}
\]

\[
\ell_{x+1} = \ell_x (1 - q_x)
\]

\[
L_x = \ell_x [1 - q_x (1 - a_x)]
\]

\[
T_x = L_x + L_{x+1} + \cdots + L_{\omega-1} + L_{\omega+}
\]

\[
e_x = \frac{T_x}{L_x}
\]

where \( a_x = 0.5 \) for \( x \geq 1 \) and \( a_0 \) taken from Coale et al (1983).

\( q_{\omega+} = 1, L_{\omega+} = l_x/m_x, \) and \( T_{\omega+} = L_{\omega+} \).

- **Period life expectancy**: let \( m_x = m_{x,t} \) for some year \( t \).
- **Cohort life expectancy**: let \( m_x = m_{x,t+x} \) for birth cohort in year \( t \).
Because we can forecast $m_{x,t}$ we can estimate the mortality rates for each birth cohort (using actual values when they are available).

We can simulate future $m_{x,t}$ in order to estimate the uncertainty associated with $e_x$. 
Because we can forecast $m_{x,t}$ we can estimate the mortality rates for each birth cohort (using actual values when they are available).

We can simulate future $m_{x,t}$ in order to estimate the uncertainty associated with $e_x$. 
Cohort life expectancy

Coherent mortality forecasting

Coherent cohort life expectancy forecasts
Simulate future mortality rates

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[
\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

\[
\log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_\ell(x) + w_t(x).
\]

- \{\gamma_{t,\ell}\} and \{\beta_{t,k}\} simulated.
- \{e_t(x)\} and \{w_t(x)\} bootstrapped.
- Generate many future sample paths for \(f_{t,M}(x)\) and \(f_{t,F}(x)\) to estimate uncertainty in \(e_x\).
Simulate future mortality rates

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)} \]

\[
\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

\[
\log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_\ell(x) + w_t(x).
\]

- \(\{\gamma_{t,\ell}\}\) and \(\{\beta_{t,k}\}\) simulated.
- \(\{e_t(x)\}\) and \(\{w_t(x)\}\) bootstrapped.
- Generate many future sample paths for \(f_{t,M}(x)\) and \(f_{t,F}(x)\) to estimate uncertainty in \(e_x\).

Coherent mortality forecasting
Simulate future mortality rates

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[
\begin{align*}
\log[p_t(x)] &= \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x) \\
\log[r_t(x)] &= \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_{\ell}(x) + w_t(x).
\end{align*}
\]

- \{\gamma_{t,\ell}\} and \{\beta_{t,k}\} simulated.
- \{e_t(x)\} and \{w_t(x)\} bootstrapped.
- Generate many future sample paths for \(f_{t,M}(x)\) and \(f_{t,F}(x)\) to estimate uncertainty in \(e_x\).
Simulate future mortality rates

\[ p_t(x) = \sqrt{f_{t,M}(x)f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x)/f_{t,F}(x)}. \]

\[
\log[p_t(x)] = \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k} \phi_k(x) + e_t(x)
\]

\[
\log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell} \psi_\ell(x) + w_t(x).
\]

- \( \{\gamma_{t,\ell}\} \) and \( \{\beta_{t,k}\} \) simulated.
- \( \{e_t(x)\} \) and \( \{w_t(x)\} \) bootstrapped.
- Generate many future sample paths for \( f_{t,M}(x) \) and \( f_{t,F}(x) \) to estimate uncertainty in \( e_x \).
Cohort life expectancy

Australia: cohort life expectancy at age 50

<table>
<thead>
<tr>
<th>Year</th>
<th>Remaining life expectancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920</td>
<td></td>
</tr>
<tr>
<td>1940</td>
<td></td>
</tr>
<tr>
<td>1960</td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td></td>
</tr>
<tr>
<td>2040</td>
<td></td>
</tr>
<tr>
<td>2060</td>
<td></td>
</tr>
</tbody>
</table>

Remaining life expectancy
library(demography)

# Read data
aus <- hmd.mx("AUS","username","password","Australia")

# Smooth data
aus.sm <- smooth.demogdata(aus)

# Fit model
aus.pr <- coherentfdm(aus.sm)

# Forecast
aus.pr.fc <- forecast(aus.pr, h=100)

# Compute life expectancies
e50.m.aus.fc <- flife.expectancy(aus.pr.fc, series="male",
                             age=50, PI=TRUE, nsim=1000, type="cohort")
e50.f.aus.fc <- flife.expectancy(aus.pr.fc, series="female",
                             age=50, PI=TRUE, nsim=1000, type="cohort")
Forecast accuracy evaluation

Australian female cohort e50: Data to 1955

Year

Remaining years

Coherent mortality forecasting
Coherent cohort life expectancy forecasts
Forecast accuracy evaluation

- Compute age 50 remaining cohort life expectancy with a rolling forecast origin beginning in 1921.
- Compare against actual cohort life expectancy where available.
- Compute 80% prediction interval actual coverage.
Forecast accuracy evaluation

- Compute age 50 remaining cohort life expectancy with a rolling forecast origin beginning in 1921.
- Compare against actual cohort life expectancy where available.
- Compute 80% prediction interval actual coverage.
Forecast accuracy evaluation

- Compute age 50 remaining cohort life expectancy with a rolling forecast origin beginning in 1921.
- Compare against actual cohort life expectancy where available.
- Compute 80% prediction interval actual coverage.
Forecast accuracy evaluation

Mean Absolute Forecast Errors

- Female
- Male

Forecast horizon

Years

0.0 0.2 0.4 0.6 0.8 1.0 1.2

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
Forecast accuracy evaluation

80% prediction interval coverage

Percentage coverage vs. Forecast horizon

Coherent mortality forecasting

Coherent cohort life expectancy forecasts
Outline

1. Functional forecasting
2. Forecasting groups
3. Coherent cohort life expectancy forecasts
4. Conclusions
Some conclusions

- Suitable for age-specific mortality.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.
- Easy to compute prediction intervals for any computable statistics.
Some conclusions

- Suitable for age-specific mortality.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.
- Easy to compute prediction intervals for any computable statistics.
Some conclusions

- Suitable for age-specific mortality.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.
- Easy to compute prediction intervals for any computable statistics.
Some conclusions

- Suitable for age-specific mortality.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.
- Easy to compute prediction intervals for any computable statistics.
Some conclusions

- Suitable for age-specific mortality.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.
- Easy to compute prediction intervals for any computable statistics.
Selected references


Selected references


---

Papers and R code: robjhyndman.com

Email: Rob.Hyndman@monash.edu