Coherent functional forecasts of mortality rates and life expectancy

Rob J Hyndman

Business & Economic Forecasting Unit

MONASH University

Joint work with

Farah Yasmeen (Monash) and **Heather Booth** (ANU)

Mortality rates

Mortality rates

The problem

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

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.

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.

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.

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.
- We want to forecast **whole curve** $f_{t,j}(x)$ for future years.

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.
- We want to forecast **whole curve** $f_{t,j}(x)$ for future years.
- **Coherent** forecasts do not diverge over time.

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.
- We want to forecast **whole curve** $f_{t,j}(x)$ for future years.
- **Coherent** forecasts do not diverge over time.
- Existing functional models do not impose coherence.

Functional time series model

(Hyndman and Ullah, CSDA, 2007)

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\, \phi_{k,j}(x) + e_{t,j}(x)$$

where $e_{t,j}(x) \stackrel{\text{iid}}{\sim} \mathrm{N}(0, v(x))$.

Functional time series model

(Hyndman and Ullah, CSDA, 2007)

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

where $e_{t,j}(x) \overset{\text{iid}}{\sim} \mathrm{N}(0, v(x))$.

1. Estimate smooth functions $f_{t,j}(x)$ using nonparametric regression.

Functional time series model

(Hyndman and Ullah, CSDA, 2007)

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

where $e_{t,j}(x) \overset{\text{iid}}{\sim} \text{N}(0, v(x))$.

1. Estimate smooth functions $f_{t,j}(x)$ using nonparametric regression.
2. Estimate $\mu_j(x)$ as mean $\log[f_{t,j}(x)]$ across years.

Functional time series model

(Hyndman and Ullah, CSDA, 2007)

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

where $e_{t,j}(x) \overset{\text{iid}}{\sim} \mathrm{N}(0, v(x))$.

1. Estimate smooth functions $f_{t,j}(x)$ using nonparametric regression.
2. Estimate $\mu_j(x)$ as mean $\log[f_{t,j}(x)]$ across years.
3. Estimate $\beta_{t,j,k}$ and $\phi_{k,j}(x)$ using functional principal components.

Functional time series model

(Hyndman and Ullah, CSDA, 2007)

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\, \phi_{k,j}(x) + e_{t,j}(x)$$

where $e_{t,j}(x) \overset{\text{iid}}{\sim} \text{N}(0, v(x))$.

1. Estimate smooth functions $f_{t,j}(x)$ using nonparametric regression.
2. Estimate $\mu_j(x)$ as mean $\log[f_{t,j}(x)]$ across years.
3. Estimate $\beta_{t,j,k}$ and $\phi_{k,j}(x)$ using functional principal components.
4. Forecast $\beta_{t,j,k}$ using time series models.

Functional time series model

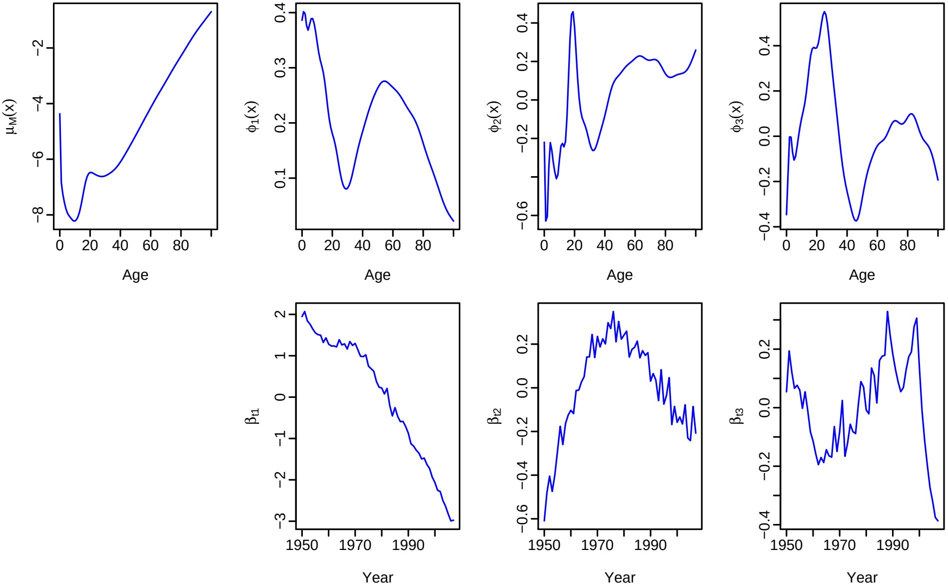
(Hyndman and Ullah, CSDA, 2007)

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

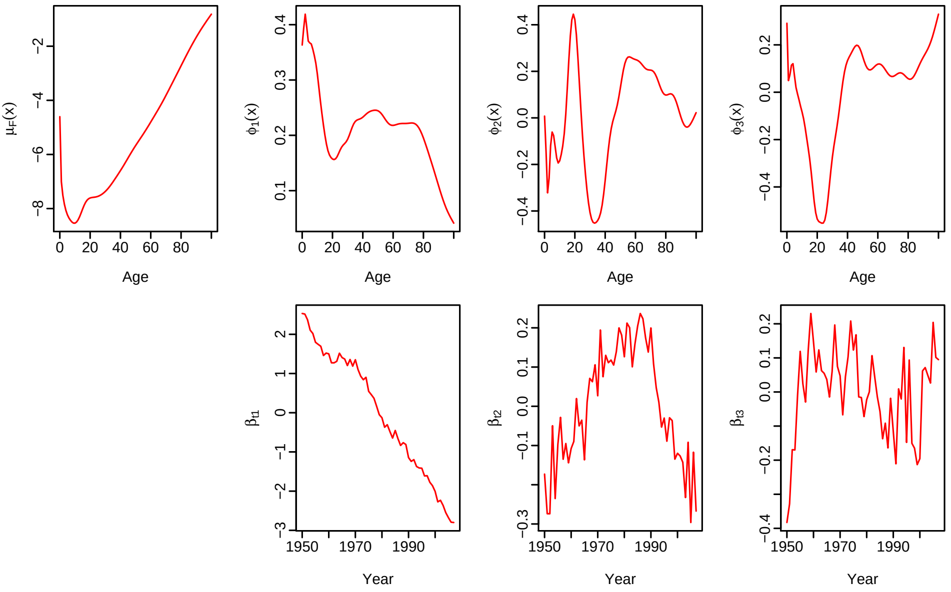
where $e_{t,j}(x) \overset{\text{iid}}{\sim} \mathrm{N}(0, v(x))$.

1. Estimate smooth functions $f_{t,j}(x)$ using nonparametric regression.
2. Estimate $\mu_j(x)$ as mean $\log[f_{t,j}(x)]$ across years.
3. Estimate $\beta_{t,j,k}$ and $\phi_{k,j}(x)$ using functional principal components.
4. Forecast $\beta_{t,j,k}$ using time series models.
5. Put it all together to get forecasts of $f_{t,j}(x)$.

Male fts model



Female fts model



Forecasting the coefficients

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

- We use ARIMA models for each coefficient $\{\beta_{1,j,k}, \dots, \beta_{n,j,k}\}$.

Forecasting the coefficients

$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

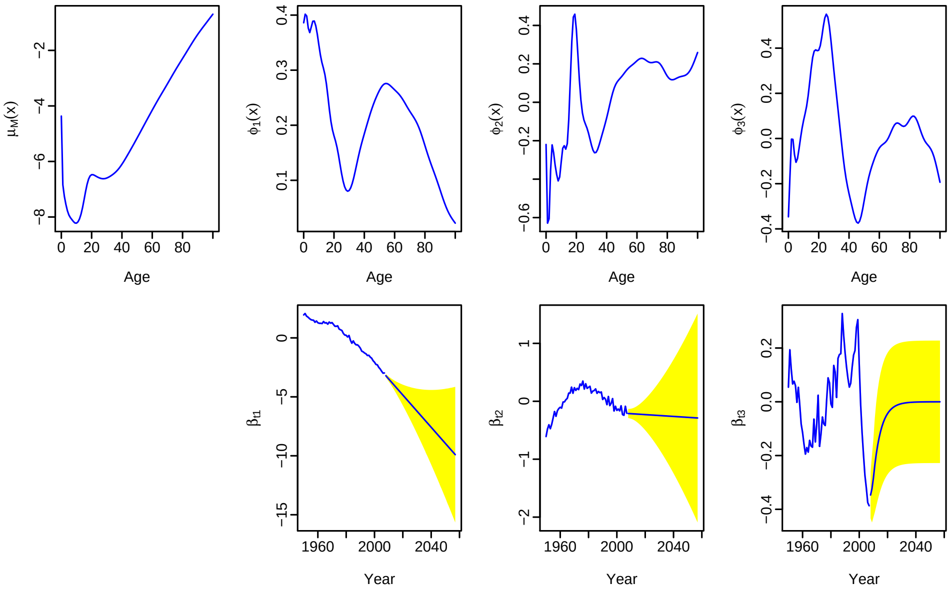
- We use ARIMA models for each coefficient $\{\beta_{1,j,k}, \dots, \beta_{n,j,k}\}$.
- The ARIMA models are non-stationary for the first few coefficients ($k = 1, 2$)

Forecasting the coefficients

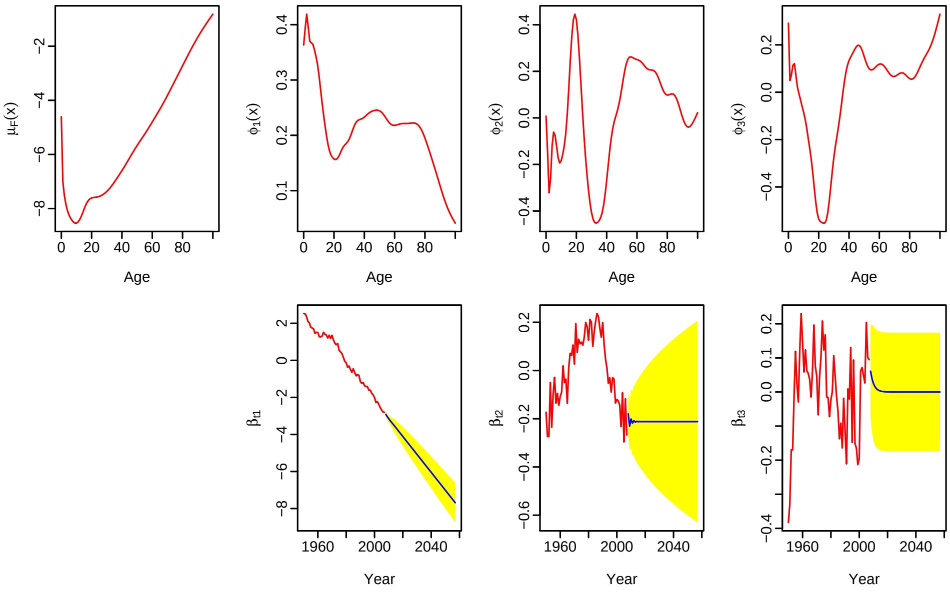
$$\log[f_{t,j}(x)] = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{k,j}(x) + e_{t,j}(x)$$

- We use ARIMA models for each coefficient $\{\beta_{1,j,k}, \dots, \beta_{n,j,k}\}$.
- 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.

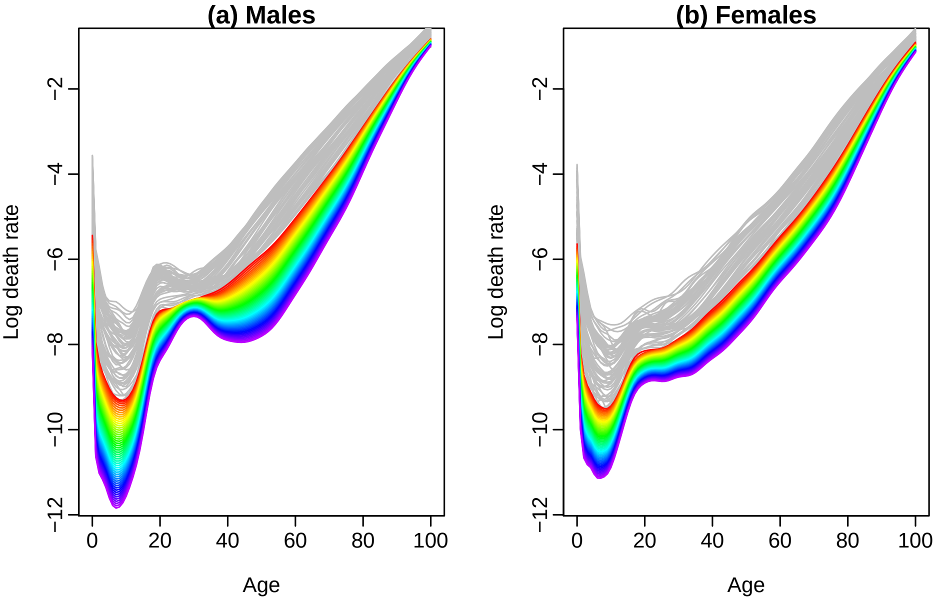
Male fts model



Female fts model



Australian mortality forecasts



Mortality product and ratios

Key idea

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

Mortality product and ratios

Key idea

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

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

Mortality product and ratios

Key idea

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

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

- Product and ratio are approximately independent

Mortality product and ratios

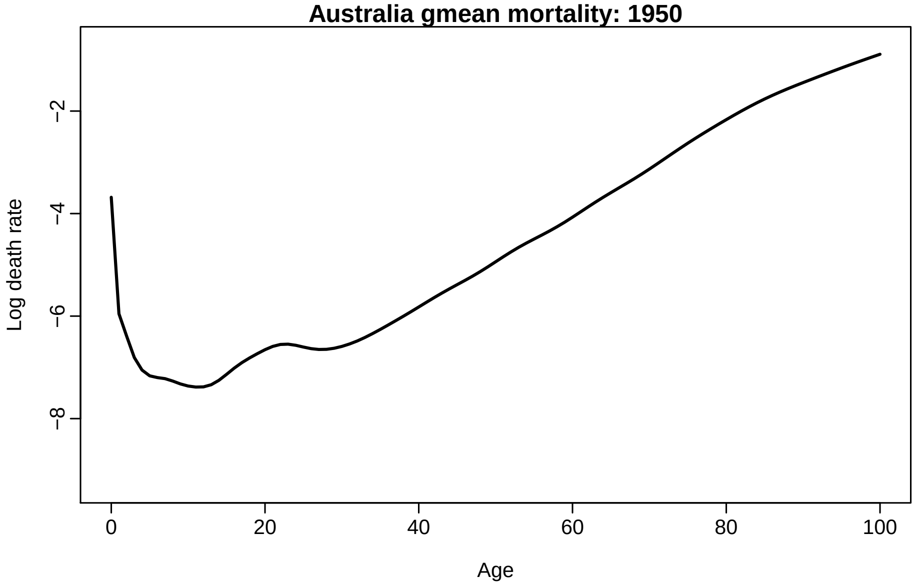
Key idea

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

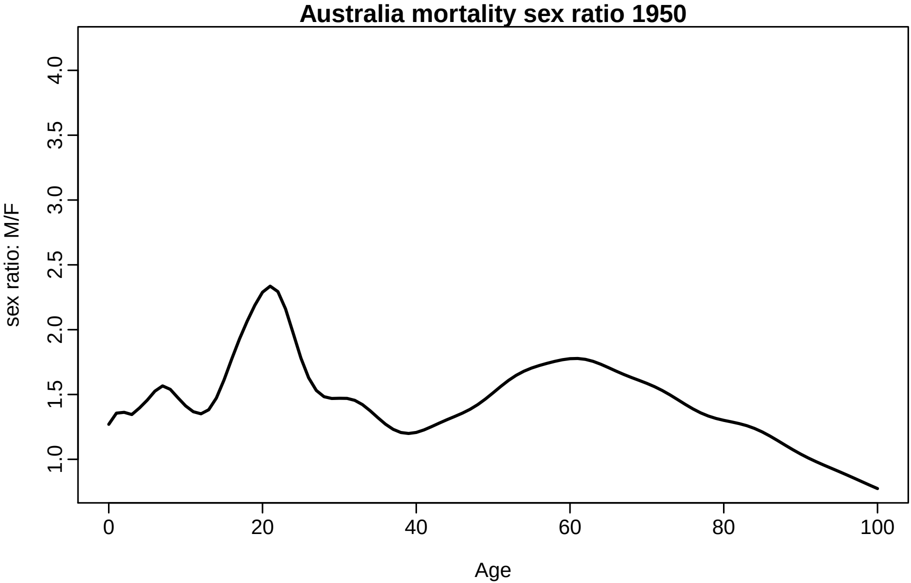
$$p_t(x) = \sqrt{f_{t,M}(x) f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x) \big/ f_{t,F}(x)}.$$

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

Mortality rates



Mortality rates



Model product and ratios

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

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

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

Model product and ratios

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

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

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

- $\{\gamma_{t,\ell}\}$ restricted to be stationary processes: either ARMA(p,q) or ARFIMA(p,d,q).

Model product and ratios

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

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

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

- $\{\gamma_{t,\ell}\}$ restricted to be stationary processes: either ARMA(p,q) or ARFIMA(p,d,q).
- No restrictions for $\beta_{t,1}, \dots, \beta_{t,K}$.

Model product and ratios

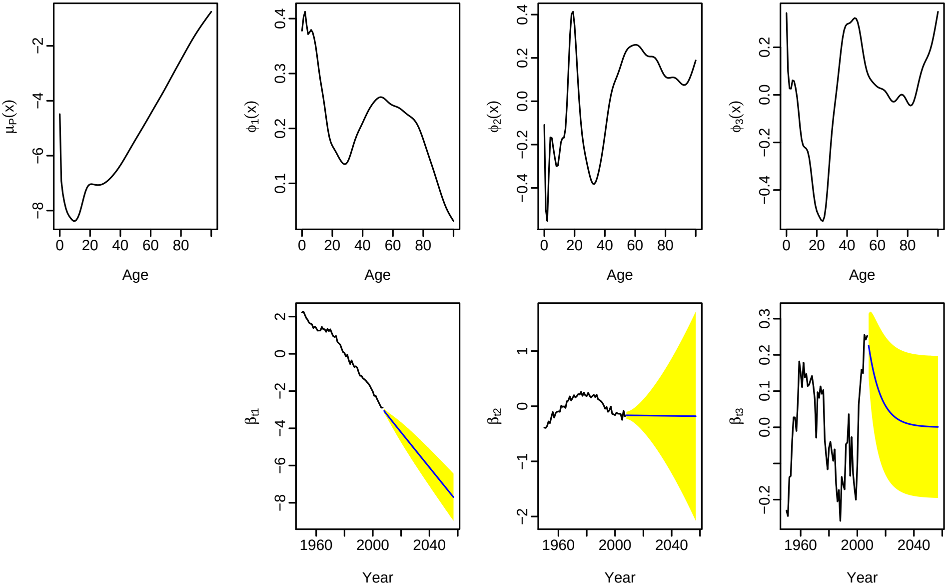
$$p_t(x) = \sqrt{f_{t,M}(x) f_{t,F}(x)} \quad \text{and} \quad r_t(x) = \sqrt{f_{t,M}(x) \big/ f_{t,F}(x)}.$$

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

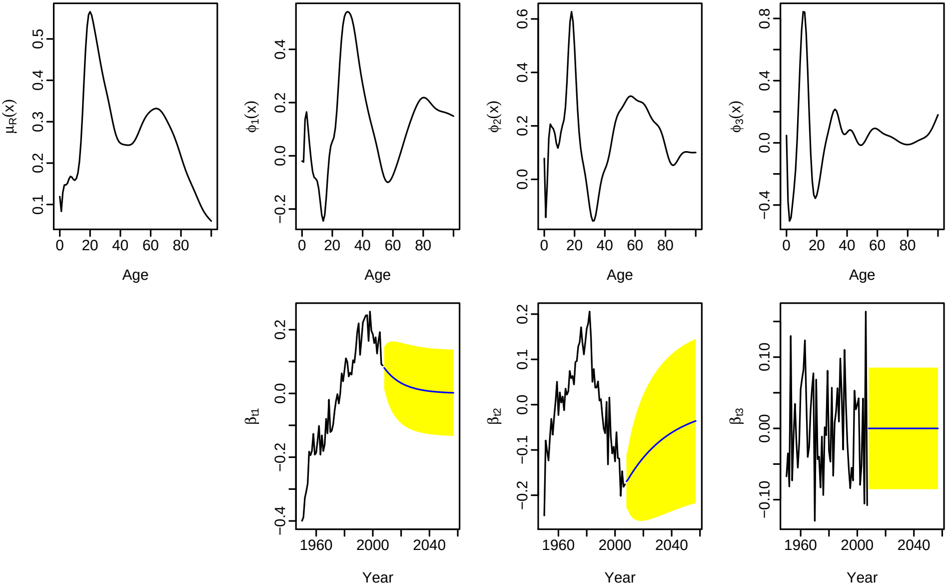
$$\log[r_t(x)] = \mu_r(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell}\psi_\ell(x) + w_t(x).$$

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

Product model

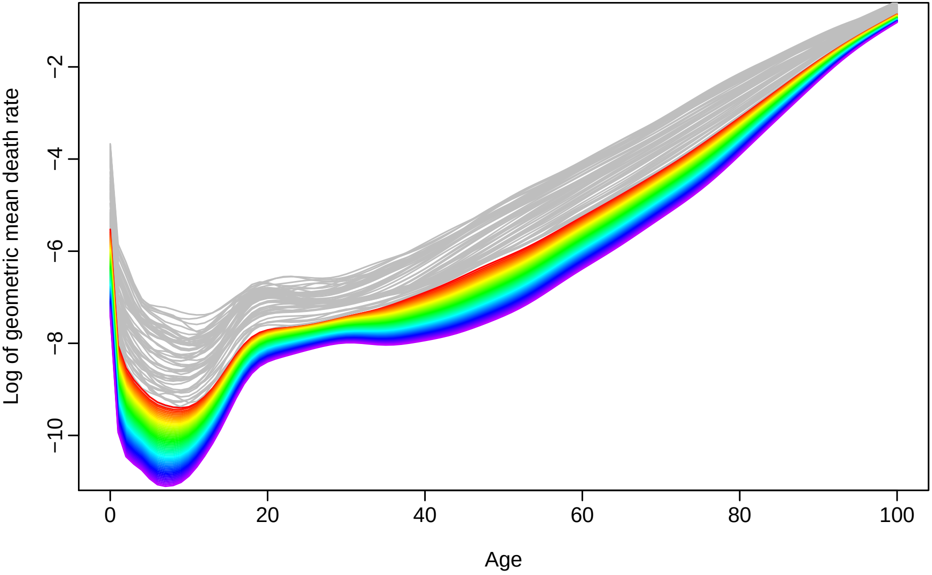


Ratio model

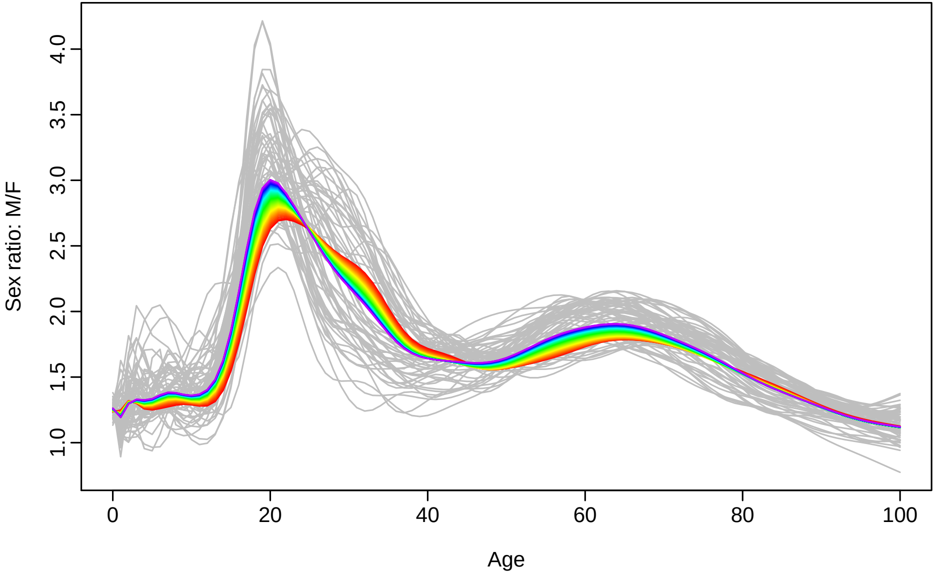


$\mu_R(x)$
Age
$\phi_1(x)$
Age
$\phi_2(x)$
Age
$\phi_3(x)$
Age
β_{t1}
Year
β_{t2}
Year
β_{t3}
Year

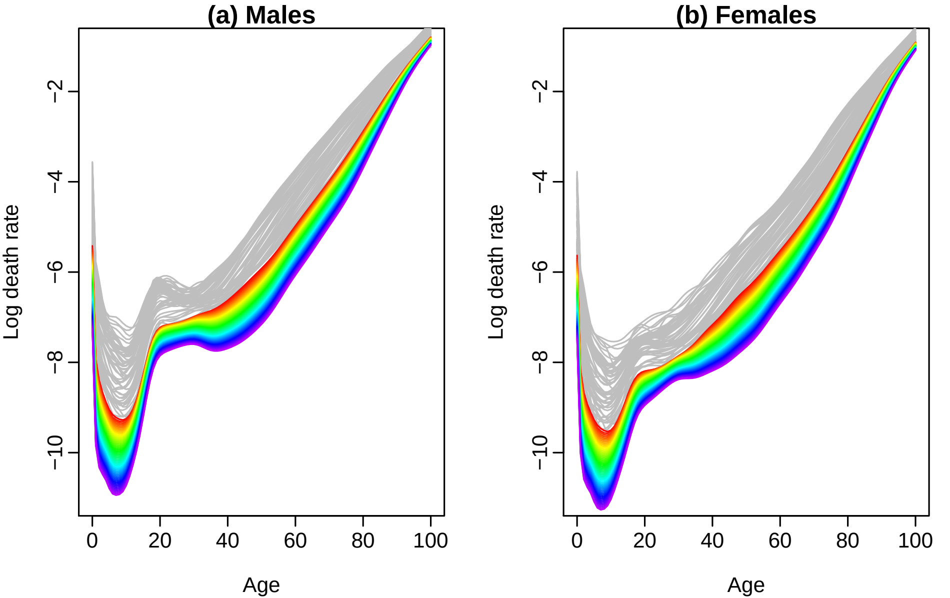
Product forecasts



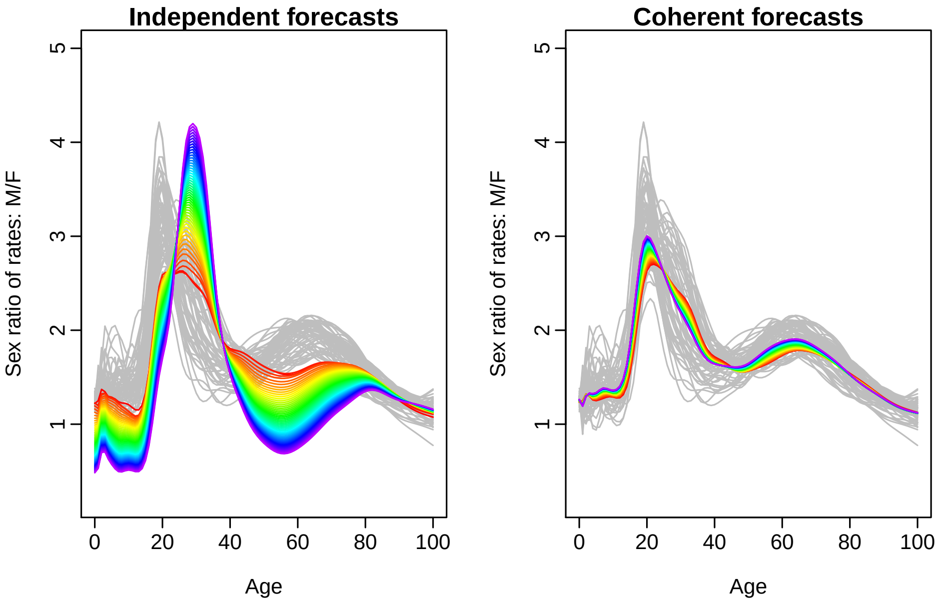
Ratio forecasts



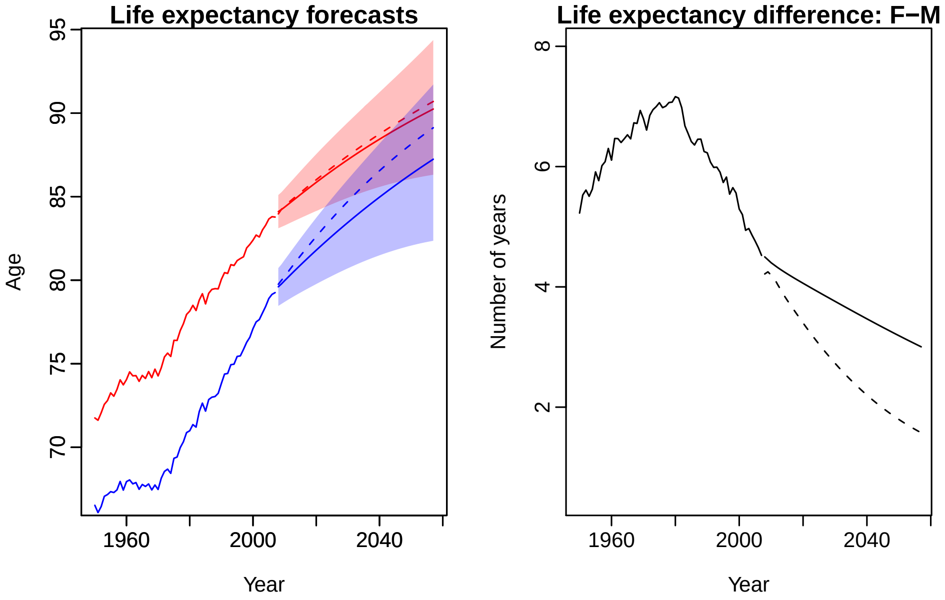
Coherent forecasts



Ratio forecasts



Life expectancy forecasts



Coherent forecasts for J groups

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

$$\text{and} \qquad r_{t,j}(x) = f_{t,j}(x) \big/ p_t(x),$$

Coherent forecasts for J groups

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

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

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

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

Coherent forecasts for J groups

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

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

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

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

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

Coherent forecasts for J groups

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

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

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

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

- $p_t(x)$ 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

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

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

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

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

- $p_t(x)$ 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

$$\begin{aligned}\log[f_{t,j}(x)] &= \log[p_t(x) r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}] \\ &= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j}\psi_{\ell,j}(x) + z_{t,j}(x)\end{aligned}$$

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

Coherent forecasts for J groups

$$
\begin{aligned}
\log[f_{t,j}(x)] &= \log[p_t(x) r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}] \\
&= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j}\psi_{\ell,j}(x) + z_{t,j}(x)
\end{aligned}
$$

- $\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)$ is group mean
- $z_{t,j}(x) = e_t(x) + w_{t,j}(x)$ is error term.

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).

Coherent forecasts for J groups

$$
\begin{aligned}
\log[f_{t,j}(x)] &= \log[p_t(x) r_{t,j}(x)] = \log[p_t(x)] + \log[r_{t,j}] \\
&= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,\ell,j}\psi_{\ell,j}(x) + z_{t,j}(x)
\end{aligned}
$$

- $\mu_j(x) = \mu_p(x) + \mu_{r,j}(x)$ 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 $\beta_{t,1},\dots,\beta_{t,K}$.

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, β_t is a random walk with drift and $\gamma_{t,j}$ is AR(1) process.

- No smoothing.

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, β_t is a random walk with drift and $\gamma_{t,j}$ is AR(1) process.

- No smoothing.
- Only one basis function for each part,

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, β_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.

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, β_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.

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, β_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.

Some conclusions

- New, automatic, flexible method for coherent forecasting of groups of functional time series.

Some conclusions

- New, automatic, flexible method for coherent forecasting of groups of functional time series.
- Based on geometric means and ratios, so interpretable results.

Some conclusions

- New, automatic, flexible method for coherent forecasting of groups of functional time series.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.

Some conclusions

- New, automatic, flexible method for coherent forecasting of groups of functional time series.
- Based on geometric means and ratios, so interpretable results.
- More general and flexible than existing methods.
- Easy to compute prediction intervals.