MONASH University

Rob J Hyndman

# Functional time series

## with applications in demography

### 7. Common functional principal components

# Outline

1. **Product/ratio coherent forecasting**
2. Common functional principal components
3. Australian mortality
4. Testing for common functional principal components
5. References

# Product/ratio coherent forecasting

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

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

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

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

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

# Product/ratio coherent forecasting

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

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

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

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

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

# Product/ratio coherent forecasting

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

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

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

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

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

# Product/ratio coherent forecasting

$$
\begin{aligned}
\log[s_{t,j}(x)] &= \log[p_t(x)] + \log[r_{t,j}(x)] \\
&= \mu_p(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + e_t(x) \\
&\quad + \mu_{r,j}(x) + \sum_{\ell=1}^{L} \gamma_{t,j,\ell}\psi_{j,\ell}(x) + w_{t,j}(x) \\
&= \mu_j(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + \sum_{\ell=1}^{L} \gamma_{t,j,\ell}\psi_{j,\ell}(x) + \varepsilon_{t,j}(x)
\end{aligned}
$$

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

# Product/ratio coherent forecasting

- Long-term forecasts of ratio functions will converge to age-specific mean ratios, which depend on the fitting period.
- Using exponential weights helps overcome this problem.
- Convergence to constant ratios does not imply that mortality differences between groups tend to constants, or that life expectancies will not diverge.
- How to deal with coherence in more than one dimension: e.g., mortality by sex and state?

# Product/ratio coherent forecasting

- Long-term forecasts of ratio functions will converge to age-specific mean ratios, which depend on the fitting period.
- Using exponential weights helps overcome this problem.
- Convergence to constant ratios does not imply that mortality differences between groups tend to constants, or that life expectancies will not diverge.
- How to deal with coherence in more than one dimension: e.g., mortality by sex and state?

# Product/ratio coherent forecasting

- Long-term forecasts of ratio functions will converge to age-specific mean ratios, which depend on the fitting period.
- Using exponential weights helps overcome this problem.
- Convergence to constant ratios does not imply that mortality differences between groups tend to constants, or that life expectancies will not diverge.
- How to deal with coherence in more than one dimension: e.g., mortality by sex and state?

# Product/ratio coherent forecasting

- Long-term forecasts of ratio functions will converge to age-specific mean ratios, which depend on the fitting period.
- Using exponential weights helps overcome this problem.
- Convergence to constant ratios does not imply that mortality differences between groups tend to constants, or that life expectancies will not diverge.
- How to deal with coherence in more than one dimension: e.g., mortality by sex and state?

# Outline

1. Product/ratio coherent forecasting
2. **Common functional principal components**
3. Australian mortality
4. Testing for common functional principal components
5. References

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Each group has a different mean $\mu_j$
- A set of common principal components $\phi_1(x), \dots, \phi_K(x)$.
- Some uncommon principal components for each group, $\psi_{1,j}(x), \dots, \psi_{L,j}(x)$.
- Common features captured with the common principal components.
- Product-ratio model is special case when $\{\gamma_{t,j,\ell}\}$ constrained to be stationary.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Each group has a different mean $\mu_j$
- A set of common principal components $\phi_1(x), \dots, \phi_K(x)$.
- Some uncommon principal components for each group, $\psi_{1,j}(x), \dots, \psi_{L,j}(x)$.
- Common features captured with the common principal components.
- Product-ratio model is special case where $\{\gamma_{t,j,\ell}\}$ constrained to be stationary.

# Partial Common Functional Principal Components

## PCFPC($K,L$) model

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

- Each group has a different mean $\mu_j$
- A set of common principal components $\phi_1(x), \dots, \phi_K(x)$.
- Some uncommon principal components for each group, $\psi_{1,j}(x), \dots, \psi_{L,j}(x)$.
- Common features captured with the common principal components.
- Product-ratio model is special case where $\{\gamma_{t,j,\ell}\}$ constrained to be stationary.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Each group has a different mean $\mu_j$
- A set of common principal components $\phi_1(x), \dots, \phi_K(x)$.
- Some uncommon principal components for each group, $\psi_{1,j}(x), \dots, \psi_{L,j}(x)$.
- Common features captured with the common principal components.
- Product-ratio model is special case where $\{\gamma_{t,j,\ell}\}$ constrained to be stationary.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Each group has a different mean $\mu_j$
- A set of common principal components $\phi_1(x), \dots, \phi_K(x)$.
- Some uncommon principal components for each group, $\psi_{1,j}(x), \dots, \psi_{L,j}(x)$.
- Common features captured with the common principal components.
- Product-ratio model is special case where $\{\gamma_{t,j,\ell}\}$ constrained to be stationary.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Each group has a different mean $\mu_j$
- A set of common principal components $\phi_1(x), \dots, \phi_K(x)$.
- Some uncommon principal components for each group, $\psi_{1,j}(x), \dots, \psi_{L,j}(x)$.
- Common features captured with the common principal components.
- Product-ratio model is special case where $\{\gamma_{t,j,\ell}\}$ constrained to be stationary.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Coherence when $\{\gamma_{t,j,\ell}\psi_{j,\ell}(x) - \gamma_{t,i,\ell}\psi_{i,\ell}(x)\}$ is stationary for all $\ell$ and for each combination of $i$ and $j$:

$$\limsup_{t\to\infty} \mathrm{E}\|f_{t,j} - f_{t,i}\| < \infty \qquad \text{for all } i \text{ and } j.$$

- Can impose coherence by requiring either stationary scores or cointegrated scores with common eigenfunction $\psi_{i,\ell}(x) = \psi_{j,\ell}(x)$.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Coherence when $\{\gamma_{t,j,\ell}\psi_{j,\ell}(x) - \gamma_{t,i,\ell}\psi_{i,\ell}(x)\}$ is stationary for all $\ell$ and for each combination of $i$ and $j$:

$$\limsup_{t\to\infty} \mathrm{E}\|f_{t,j} - f_{t,i}\| < \infty \qquad \text{for all } i \text{ and } j.$$

- Can impose coherence by requiring either stationary scores or cointegrated scores with common eigenfunction $\psi_{i,\ell}(x) = \psi_{j,\ell}(x)$.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- Coherence when $\{\gamma_{t,j,\ell}\psi_{j,\ell}(x) - \gamma_{t,i,\ell}\psi_{i,\ell}(x)\}$ is stationary for all $\ell$ and for each combination of $i$ and $j$:

$$\limsup_{t\to\infty} \mathrm{E}\|f_{t,j} - f_{t,i}\| < \infty \qquad \text{for all } i \text{ and } j.$$

- Can impose coherence by requiring either stationary scores or cointegrated scores with common eigenfunction $\psi_{i,\ell}(x) = \psi_{j,\ell}(x)$.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- **Model 1: PCFPC($K, 0$)**. No idiosyncratic principal components in the model.
- **Model 2: PCFPC($K, L$)** with a coherence constraint. For each $\ell$, $\{\gamma_{t,i,\ell} - \gamma_{t,j,\ell}\}$ is stationary for all $i, j$, and $\psi_{j,\ell}(x) = \psi_{i,\ell}(x)$.
- **Model 3: PCFPC($K, L$)** with a coherence constraint. For each $\ell$ and $j$, $\{\gamma_{t,j,\ell}\}$ is stationary.
- **Model 4: PCFPC($0, L$)**. All principal components and scores are idiosyncratic.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- **Model 1: PCFPC($K, 0$).** No idiosyncratic principal components in the model.
- **Model 2: PCFPC($K, L$)** with a coherence constraint. For each $\ell$, $\{\gamma_{t,i,\ell} - \gamma_{t,j,\ell}\}$ is stationary for all $i, j$, and $\psi_{j,\ell}(x) = \psi_{i,\ell}(x)$.
- **Model 3: PCFPC($K, L$)** with a coherence constraint. For each $\ell$ and $j$, $\{\gamma_{t,j,\ell}\}$ is stationary.
- **Model 4: PCFPC($0, L$).** All principal components and scores are idiosyncratic.

# Partial Common Functional Principal Components

## PCFPC($K, L$) model

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

- **Model 1: PCFPC($K, 0$)**. No idiosyncratic principal components in the model.
- **Model 2: PCFPC($K, L$)** with a coherence constraint. For each $\ell$, $\{\gamma_{t,i,\ell} - \gamma_{t,j,\ell}\}$ is stationary for all $i, j$, and $\psi_{j,\ell}(x) = \psi_{i,\ell}(x)$.
- **Model 3: PCFPC($K, L$)** with a coherence constraint. For each $\ell$ and $j$, $\{\gamma_{t,j,\ell}\}$ is stationary.
- **Model 4: PCFPC($0, L$)**. All principal components and scores are idiosyncratic.

## PCFPC($K, L$) model

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

- **Model 1: PCFPC($K, 0$).** No idiosyncratic principal components in the model.
- **Model 2: PCFPC($K, L$)** with a coherence constraint. For each $\ell$, $\{\gamma_{t,i,\ell} - \gamma_{t,j,\ell}\}$ is stationary for all $i, j$, and $\psi_{j,\ell}(x) = \psi_{i,\ell}(x)$.
- **Model 3: PCFPC($K, L$)** with a coherence constraint. For each $\ell$ and $j$, $\{\gamma_{t,j,\ell}\}$ is stationary.
- **Model 4: PCFPC($0, L$).** All principal components and scores are idiosyncratic.

## PCFPC($K, L$) model

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

- **Model 1: PCFPC($K, 0$).** No idiosyncratic principal components in the model.
- **Model 2: PCFPC($K, L$)** with a coherence constraint. For each $\ell$, $\{\gamma_{t,i,\ell} - \gamma_{t,j,\ell}\}$ is stationary for all $i, j$, and $\psi_{j,\ell}(x) = \psi_{i,\ell}(x)$.
- **Model 3: PCFPC($K, L$)** with a coherence constraint. For each $\ell$ and $j$, $\{\gamma_{t,j,\ell}\}$ is stationary.
- **Model 4: PCFPC($0, L$).** All principal components and scores are idiosyncratic.

# Outline

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Experimental set up

- Use data for males/females in Australia from 1950 to avoid the outliers due to the World Wars.
- Data for 1950–2009 obtained from Human Mortality Database.
- All data smoothed (independently for each year) using penalized regression splines with monotonicity constraint above age 65.
- $K = L = 6$.
- ARIMA models for common PC scores.
- ARFIMA models for stationary PC scores with $0 < d < 0.5$.
- VECM using the Johansen procedure for cointegrated PC scores.
- Rolling forecast origin: 1969–2008, forecasting up to 20 years ahead.

# Out-of-sample MSE

| Forecast horizon | Groups | **Model 1** PCFPC(6,0) (All common) | **Model 2** PCFPC(6,6) (Cointegrated) | **Model 3** PCFPC(6,6) (Stationary) | **Model 4** PCFPC(0,6) (Divergent) |
|---|---|---|---|---|---|
| $h = 5$ | Combined (F & M) | 2.59 | 2.60 | **2.50** | 2.52 |
| | Female (F) | 2.81 | 2.75 | 2.70 | **2.63** |
| | Male (M) | 2.38 | 2.45 | **2.29** | 2.42 |
| $h = 10$ | Combined (F & M) | **4.57** | 4.66 | 4.60 | 4.65 |
| | Female(F) | 4.67 | 4.43 | 4.63 | **4.23** |
| | Male (M) | **4.48** | 4.89 | 4.57 | 5.06 |
| $h = 15$ | Combined (F & M) | **7.72** | 8.00 | 7.84 | 8.15 |
| | Female (F) | 7.31 | 6.64 | 7.23 | **6.47** |
| | Male(M) | **8.14** | 9.36 | 8.44 | 9.82 |
| $h = 20$ | Combined (F & M) | **12.97** | 13.56 | 13.35 | 14.10 |
| | Female (F) | 12.26 | 10.41 | 12.08 | **10.35** |
| | Male (M) | **13.69** | 16.70 | 14.63 | 17.86 |

# Common functional PC

- The independent (divergent) models work better for female data — due to the hump in male mortality being captured in common components?
- The best coherent model has all principal components and scores in common. So the models differ only in mean.
- PCFPC model more general, so poor performance a problem of model selection.
- Maybe PCFPC (cointegrated) would be better if we had a good automated VECM procedure.
- Perhaps we should consider models with some cointegrated scores and some stationary scores.
- PCFPC used $K = L = 6$. Too many? How to do order selection?

# Common functional PC

- The independent (divergent) models work better for female data — due to the hump in male mortality being captured in common components?
- The best coherent model has all principal components and scores in common. So the models differ only in mean.
- PCFPC model more general, so poor performance a problem of model selection.
- Maybe PCFPC (cointegrated) would be better if we had a good automated VECM procedure.
- Perhaps we should consider models with some cointegrated scores and some stationary scores.
- PCFPC used $K = L = 6$. Too many? How to do order selection?

# Common functional PC

- The independent (divergent) models work better for female data — due to the hump in male mortality being captured in common components?
- The best coherent model has all principal components and scores in common. So the models differ only in mean.
- PCFPC model more general, so poor performance a problem of model selection.
- Maybe PCFPC (cointegrated) would be better if we had a good automated VECM procedure.
- Perhaps we should consider models with some cointegrated scores and some stationary scores.
- PCFPC used $K = L = 6$. Too many? How to do order selection?

# Common functional PC

- The independent (divergent) models work better for female data — due to the hump in male mortality being captured in common components?
- The best coherent model has all principal components and scores in common. So the models differ only in mean.
- PCFPC model more general, so poor performance a problem of model selection.
- Maybe PCFPC (cointegrated) would be better if we had a good automated VECM procedure.
- Perhaps we should consider models with some cointegrated scores and some stationary scores.
- PCFPC used $K = L = 6$. Too many? How to do order selection?

# Common functional PC

- The independent (divergent) models work better for female data — due to the hump in male mortality being captured in common components?
- The best coherent model has all principal components and scores in common. So the models differ only in mean.
- PCFPC model more general, so poor performance a problem of model selection.
- Maybe PCFPC (cointegrated) would be better if we had a good automated VECM procedure.
- Perhaps we should consider models with some cointegrated scores and some stationary scores.
- PCFPC used $K = L = 6$. Too many? How to do order selection?

# Common functional PC

- The independent (divergent) models work better for female data — due to the hump in male mortality being captured in common components?
- The best coherent model has all principal components and scores in common. So the models differ only in mean.
- PCFPC model more general, so poor performance a problem of model selection.
- Maybe PCFPC (cointegrated) would be better if we had a good automated VECM procedure.
- Perhaps we should consider models with some cointegrated scores and some stationary scores.
- PCFPC used $K = L = 6$. Too many? How to do order selection?

# Outline

# Implied volatiliity

## Benko, Härdle & Kneip (2009)

### Independent FPC models

$$f_{t,j}(x) = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{j,k}(x) + \varepsilon_{t,j}(x), \quad j = 1, 2$$

- Each group has a different mean $\mu_j$
- Inference to test if $\phi_{1,k}(x) = \phi_{2,k}(x)$
- Weaker hypothesis: equality of eigenspaces spanned by first $K$ PCs.
- Application to implied volatility where $f_{t,j}(x)$ denotes log-return for option at price $x$ on maturity $j$.

# Implied volatiliity

## Benko, Härdle & Kneip (2009)

### Independent FPC models

$$f_{t,j}(x) = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\, \phi_{j,k}(x) + \varepsilon_{t,j}(x), \quad j = 1, 2$$

- Each group has a different mean $\mu_j$
- Inference to test if $\phi_{1,k}(x) = \phi_{2,k}(x)$
- Weaker hypothesis: equality of eigenspaces spanned by first $K$ PCs.
- Application to implied volatility where $f_{t,j}(x)$ denotes log-return for option at price $x$ on maturity $j$.

# Implied volatiliity

## Benko, Härdle & Kneip (2009)

### Independent FPC models

$$f_{t,j}(x) = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{j,k}(x) + \varepsilon_{t,j}(x), \quad j = 1, 2$$

- Each group has a different mean $\mu_j$
- Inference to test if $\phi_{1,k}(x) = \phi_{2,k}(x)$
- Weaker hypothesis: equality of eigenspaces spanned by first $K$ PCs.
- Application to implied volatility where $f_{t,j}(x)$ denotes log-return for option at price $x$ on maturity $j$.

# Implied volatiliity

## Benko, Härdle & Kneip (2009)

### Independent FPC models

$$f_{t,j}(x) = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\, \phi_{j,k}(x) + \varepsilon_{t,j}(x), \quad j = 1, 2$$

- Each group has a different mean $\mu_j$
- Inference to test if $\phi_{1,k}(x) = \phi_{2,k}(x)$
- Weaker hypothesis: equality of eigenspaces spanned by first $K$ PCs.
- Application to implied volatility where $f_{t,j}(x)$ denotes log-return for option at price $x$ on maturity $j$.

# Implied volatiliity

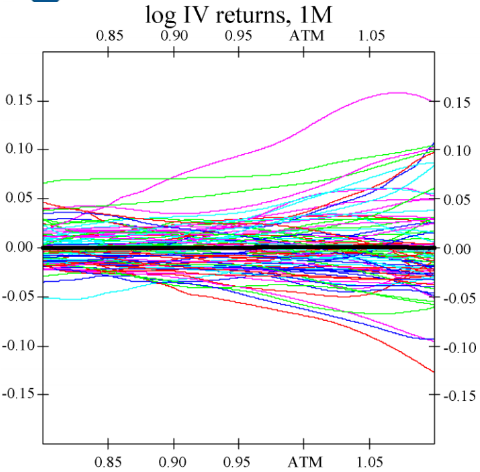
## Benko, Härdle & Kneip (2009)

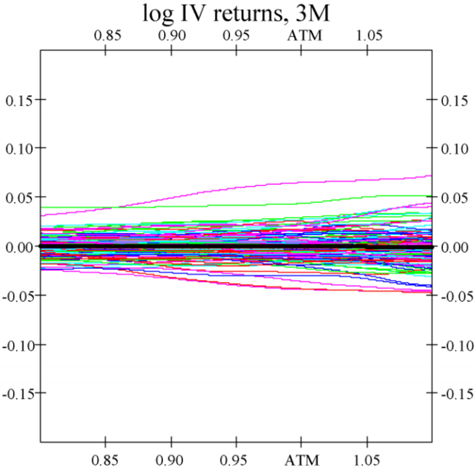
### Independent FPC models

$$f_{t,j}(x) = \mu_j(x) + \sum_{k=1}^{K} \beta_{t,j,k}\,\phi_{j,k}(x) + \varepsilon_{t,j}(x), \quad j = 1, 2$$

- Each group has a different mean $\mu_j$
- Inference to test if $\phi_{1,k}(x) = \phi_{2,k}(x)$
- Weaker hypothesis: equality of eigenspaces spanned by first $K$ PCs.
- Application to implied volatility where $f_{t,j}(x)$ denotes log-return for option at price $x$ on maturity $j$.

# Implied volatiliity

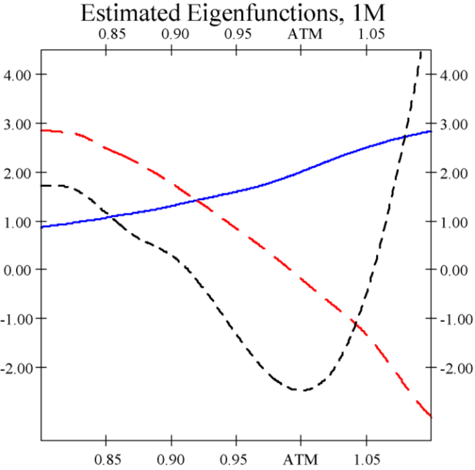


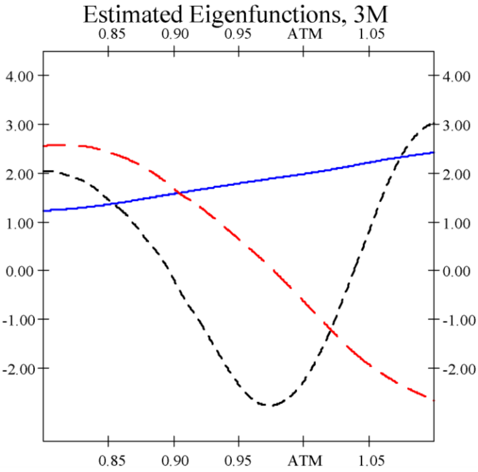




Relative price
One month maturity

Relative price
3 month maturity

# Implied volatiliity

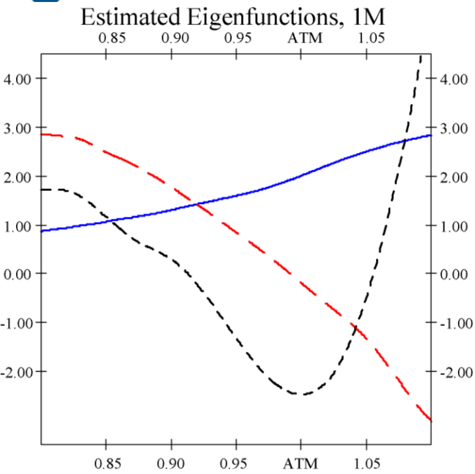


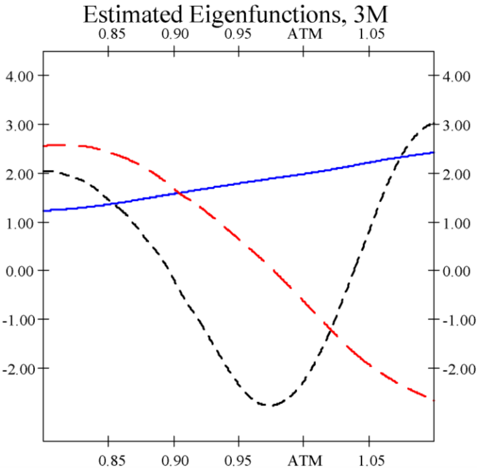




Relative price
One month maturity

Relative price
3 month maturity

# Implied volatiliity




Relative price
One month maturity

Relative price
3 month maturity

Var explained: 89.9 7.7 1.7 0.6

93.0 4.2 1.0 0.4

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ ($K = 3$)
- Practical implications:

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ $(K = 3)$
- Practical implications:

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ ($K = 3$)
- Practical implications:
  - Not enough data to see the differences between the maturities
  - Better estimates obtained by combining the two data sets.

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ $(K = 3)$
- Practical implications:
  - Not enough data to see the differences between the maturities
  - Better estimates obtained by combining the two data sets.

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ ($K = 3$)
- Practical implications:
  - Not enough data to see the differences between the maturities
  - Better estimates obtained by combining the two data sets.

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ $(K = 3)$
- Practical implications:
  - Not enough data to see the differences between the maturities
  - Better estimates obtained by combining the two data sets.

# Inference on FPC

- Regularity and stationarity assumptions (not applicable to the 2-sex mortality problem)
- Restricted to Nadaraya-Watson or local linear smoothing (not splines?)
- Test of equivalent eigenfunctions: $p = 0.01$
- Test of equivalent eigenspaces: $p = 0.09$ $(K = 3)$
- Practical implications:
  - Not enough data to see the differences between the maturities
  - Better estimates obtained by combining the two data sets.

# To test or predict?

- Testing equivalent of eigenfunctions or eigenspaces is not the same as determining if common estimates give better predictions.
- Testing observational data is almost always about whether there is enough data to detect differences. We know that differences exist.
- If prediction is the aim, then compare predictive accuracy, not model equivalence.
- When is eigenfunction/eigenspace equivalence useful?

# To test or predict?

- Testing equivalent of eigenfunctions or eigenspaces is not the same as determining if common estimates give better predictions.
- Testing observational data is almost always about whether there is enough data to detect differences. We know that differences exist.
- If prediction is the aim, then compare predictive accuracy, not model equivalence.
- When is eigenfunction/eigenspace equivalence useful?

# To test or predict?

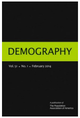
- Testing equivalent of eigenfunctions or eigenspaces is not the same as determining if common estimates give better predictions.
- Testing observational data is almost always about whether there is enough data to detect differences. We know that differences exist.
- If prediction is the aim, then compare predictive accuracy, not model equivalence.
- When is eigenfunction/eigenspace equivalence useful?

# To test or predict?

- Testing equivalent of eigenfunctions or eigenspaces is not the same as determining if common estimates give better predictions.
- Testing observational data is almost always about whether there is enough data to detect differences. We know that differences exist.
- If prediction is the aim, then compare predictive accuracy, not model equivalence.
- When is eigenfunction/eigenspace equivalence useful?

# Outline

# Selected references


Hyndman, Booth, Yasmeen (2013). "Coherent mortality forecasting: the product-ratio method with functional time series models". *Demography* **50**(1), 261–283.


Benko, Härdle, Kneip (2009). "Common functional principal components". *Annals of Statistics* **37**(1), 1–34.

Hyndman (2014). *demography: Forecasting mortality, fertility, migration and population data*. `cran.r-project.org/package=demography`