An investigation of period and cohort effects in adult mortality across 31 countries

David G. McCarthy and Polin Wang

13th International Longevity Risk and Capital Markets Solutions Conference, Taiwan

September 22, 2017

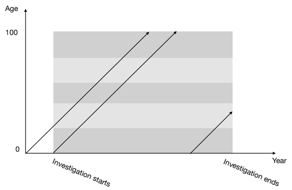
A quick summary

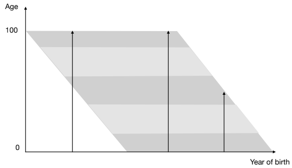
We fitted an age-cohort model with period effects using Bayesian maximum a posteriori estimation to mortality data of men and women in 31 countries. We find:

- Diminishing variance of period effects in almost all countries
- Very high correlation in estimated period effects between men and women in the same country
- Very high correlation in cohort mortality shocks between men and women in the same country
- Cohort mortality seems to be influenced by both cultural and geographical effects; period shocks seem to be largely geographical

Motivation

- Period analysis vs cohort analysis
- McCarthy (2017) argues that using cohort and age as primary variables, with period effects secondary, may have significant advantages
  - There is a natural link to survival distributions
  - Cohort effects are as good at capturing a constant time trend
  - Cohort effects are persistent across age
  - The variance of cohort effects does not appear to be diminishing
- He proposed an age-cohort model with random period effects estimated using BMAP that allows us to separate period and cohort effects





This paper

Here, we apply this model to male and female mortality data from 31 countries to:

- Identify international trends in period and cohort mortality
- Better understand common drivers of period and cohort mortality shocks
- Assist countries in understanding how best to diversify their mortality shocks (whether through capital markets or otherwise)

- (Will also serve to validate the model itself)

UNIVERSITY OF GEORGIA
Terry College of Business

Model

An age-cohort model of mortality (based on CBD, 2006)

$$log(m_{x,c}) = \alpha_c + \beta_c \frac{x - k_1}{k_2} + \gamma_t + \epsilon_{x,c}$$

where
$m_{x,c}$ = observed central rate of death;
x = age and k_1 and k_2 are chosen constants such that $\frac{x-k_1}{k_2} \in [-1, 1]$;
c = cohort;
t = period $= x + c$;
α_c = the average level of cohort effects for cohort c throughout their lives;
β_c = the slope of cohort effects for cohort c throughout their lives;
$\gamma_t \sim N(0, \sigma^2_\gamma)$, i.e. the period random effects;
$\epsilon_{x,c} \sim N(0, \sigma^2_\epsilon \omega_{x,c})$, where $\omega_{x,c} = \frac{1}{D_{x,c}}$ and $D_{x,c}$ is the number of deaths in cohort c and age x;

estimated using BMAP with prior distribution

$\begin{bmatrix} \alpha_c \\ \beta_c \end{bmatrix} = \theta_c$, where $(\theta_c - \theta_{c-1})$ follows a lag-1 VAR,

meaning $(\theta_c - \theta_{c-1}) = \mu + \Delta(\theta_{c-1} - \theta_{c-2}) + \nu_c$ and $\nu_c \sim MVN(0, \Sigma)$.

BMAP estimation steps

Step 1:
Estimate $\hat{\alpha}_c$ and $\hat{\beta}_c$ cohort by cohort using ML

Step 2:
Use $\hat{\alpha}_c$ and $\hat{\beta}_c$ to estimate VAR coefficients for prior distribution

Step 3:
Use Bayes' theorm, the estimated prior and the likelihood function of the data to generate the posterior distribution

Step 4:
Obtain point estimates $\hat{\alpha}_c$ and $\hat{\beta}_c$ as the joint mode of the posterior distribution

Data

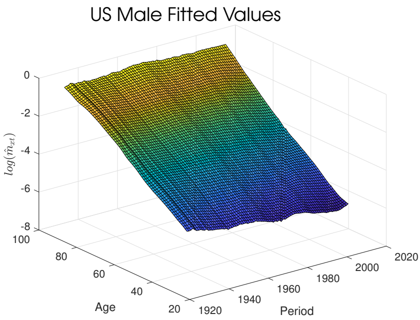
- The Human Mortality Database (HMD - mortality.org)

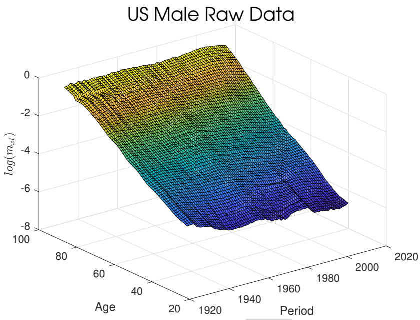
  Countries include:
  Australia, Japan, Slovakia, Austria, Finland, Slovenia, France, Spain, Belgium, Germany, Sweden, Bulgaria, Greece, Netherlands, Switzerland, Canada, Hungary, New Zealand, Taiwan, Chile, Norway, UK, Croatia, Ireland, Poland, USA, Czech Republic, Israel, Portugal, Denmark, Italy
- Male and female; ages from 35 to 100
  (This allows us to run the model 31 countries $\times$ 2 sexes times **independently**.)
- Periods (t): 1900 - 2015
- Cohorts (c): 1800 - 1975

(Countries have different available data length.)

Model performance

Fitted v. Raw - US Male

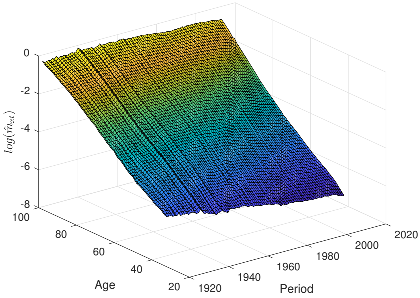




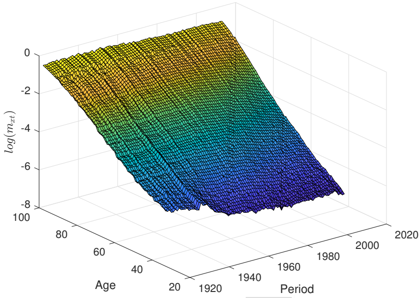
Model performance

Fitted v. Raw - UK Male

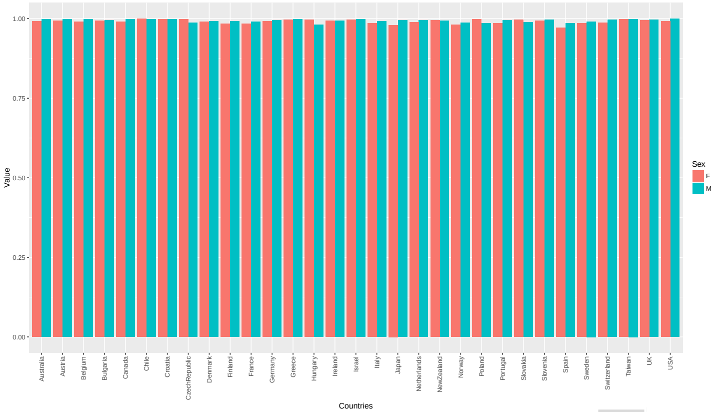
UK Male Fitted Values



UK Male Raw Data



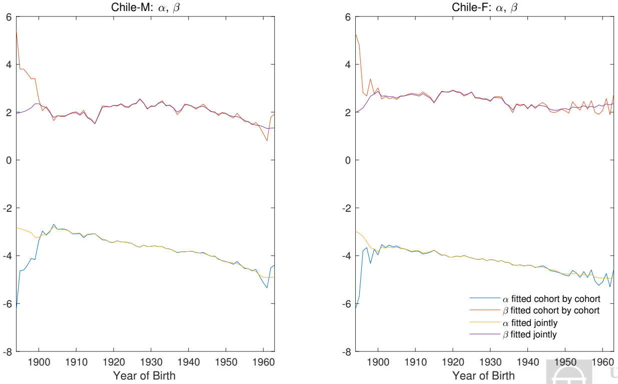
Model performance - R^2



- $R^2 = 1 - \frac{\text{Error sum of squares}}{\text{Total sum of squares}}$
- Mean: 99.25%

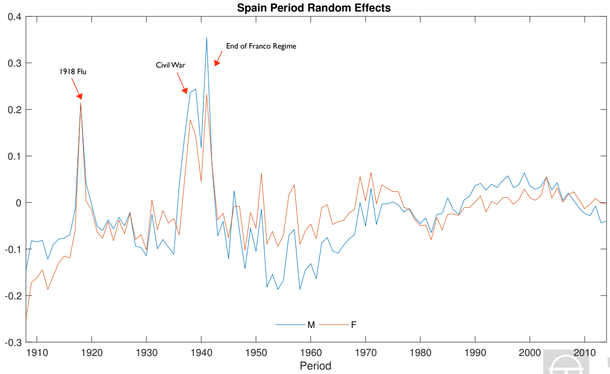
Estimates

Cohort Level and Slope (α and β)



Estimates

Period Effects (γ)



Which countries share similar mortality characteristics?

- Sets of cohort and period parameter estimates:

$$\{\hat{\alpha}_{i,s,c}, \hat{\beta}_{i,s,c}, \hat{\gamma}_{i,s,t}\}$$

$$i \in \{1, ..., 31\}$$

$$s \in \{M, F\}$$

$$c \in \{1800, ..., 1975\}$$

$$t \in \{1900, ..., 2015\}$$

- What are their statistical properties?
- Do they contain information about international mortality patterns?

Statistical properties of the cohort effects

- Strong evidence of unit roots in $\{\hat{\alpha}_{i,s,c}, \hat{\beta}_{i,s,c}\}$
- $\{\Delta\hat{\alpha}_{i,s,c}, \Delta\hat{\beta}_{i,s,c}\}$ generally follow AR(1) processes:

$$\Delta\hat{\alpha}_{i,s,c} = \mu^{\alpha}_{i,s} + \Phi^{\alpha}_{i,s}\Delta\hat{\alpha}_{i,s,c-1} + \epsilon^{\alpha}_{i,s,c}$$

$$\Delta\hat{\beta}_{i,s,c} = \mu^{\beta}_{i,s} + \Phi^{\beta}_{i,s}\Delta\hat{\beta}_{i,s,c-1} + \epsilon^{\beta}_{i,s,c}$$

- Define cohort mortality shocks as:

$$u^{\alpha}_{i,s,c} = \Delta\hat{\alpha}_{i,s,c} - \hat{\mu}^{\alpha}_{i,s} - \hat{\Phi}^{\alpha}_{i,s}\Delta\hat{\alpha}_{i,s,c-1}$$

$$u^{\beta}_{i,s,c} = \Delta\hat{\beta}_{i,s,c} - \hat{\mu}^{\beta}_{i,s} - \hat{\Phi}^{\beta}_{i,s}\Delta\hat{\beta}_{i,s,c-1}$$

UNIVERSITY OF GEORGIA
Terry College of Business

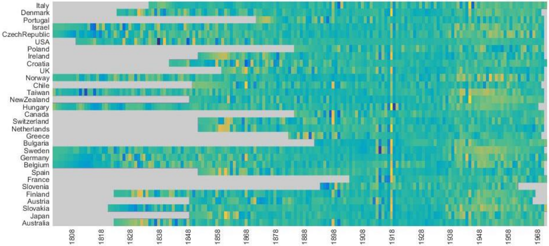
Statistical properties of the cohort mortality shocks within country

- Strong positive correlation between shocks for males and females
- Strong negative correlation between u^{α} and u^{β}: cohort mortality shocks affect mortality at younger ages more than at older ages

$\frac{1}{n}\sum_{c} Corr(u^{i}_{s_1,c}, u^{j}_{s_2,c})$	$i=\alpha, s_1=M$	$i=\alpha, s_1=F$	$i=\beta, s_1=M$	$i=\beta, s_1=F$
$j=\alpha, s_2=M$	1			
$j=\alpha, s_2=F$	0.585	1		
$j=\beta, s_2=M$	-0.408	-0.275	1	
$j=\beta, s_2=F$	-0.414	-0.284	0.938	1

Statistical properties of the cohort mortality shocks between countries

- Strong historical component in u^{α} but not so much in u^{β}
- 1918 flu epidemic, WWI, WWII all exerted permanent effects on cohort mortality in many countries



- How can we identify countries that are more and less similar?

Distance measure

Correlation distance between countries i and j

$$d(x, y)_{i,j} = 1 - corr(u^x_{i,y}, u^x_{j,y})$$

where $x \in \{\alpha, \beta\}$;
$y \in \{M, F\}$

Distance measures

Distance Matrix

The pairwise correlation gives us a distance matrix:

	Australia	Japan	Slovakia	Austria	...
Australia	0	0.922	0.778	0.729	...
Japan	0.922	0	1.178	0.702	...
Slovakia	0.777	1.177	0	0.800	...
Austria	0.729	0.702	0.800	0	...
Finland	0.910	0.995	0.711	0.778	...
Slovenia	0.830	0.788	0.600	0.616	...
France	0.833	0.810	0.657	0.653	...
⋮	⋮	⋮	⋮	⋮	⋱

But how to interpret this in an accessible way?

Cluster analysis

Methods of grouping data given a distance matrix

- Principal component analysis allows us to infer a map given a distance matrix
  - The first k eigenvalues of the Gram matrix give the best k-dimensional approximation to the map (we use two)
- Hierarchical clustering
  - Cluster by “cluster distances”
  - Start with each point as a cluster
  - Group the closest pairs → calculate the pairwise cluster distances
  - Group again... → cluster dendrogram

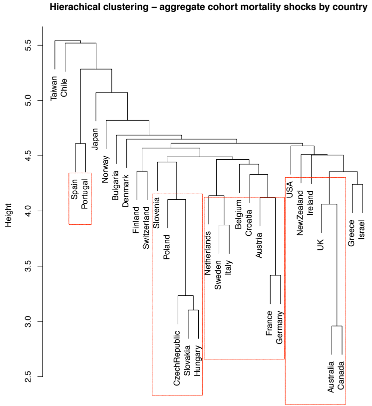
Clustering - cohort mortality shocks (u^α and u^β)

We calculate an aggregate cohort mortality shock distance measure between countries i and j as:

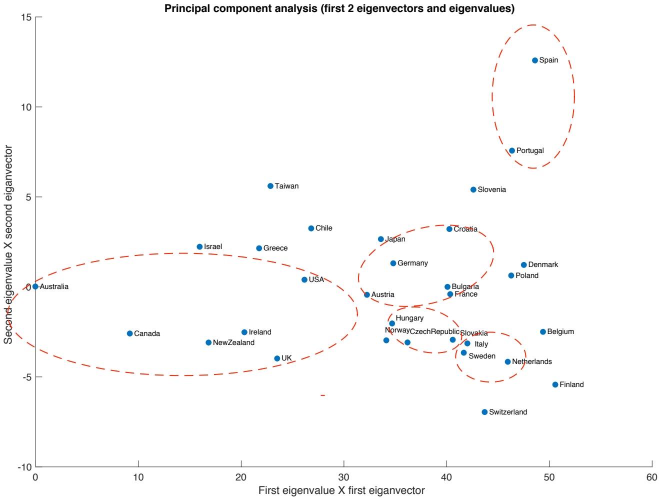
$$S_{i,j} = d(\alpha, M)_{i,j} + d(\alpha, F)_{i,j} + d(\beta, M)_{i,j} + d(\beta, F)_{i,j}$$

Seem to be both cultural and regional components

- Anglosphere falls into a single group
- The other groupings appear to be more geographical
  - Western Europe
  - Eastern Europe
  - Scandinavia (plus Bulgaria and Switzerland)
  - Iberian peninsula



Clustering - cohort mortality shocks (u^{α} and u^{β})



Statistical properties of the period effects

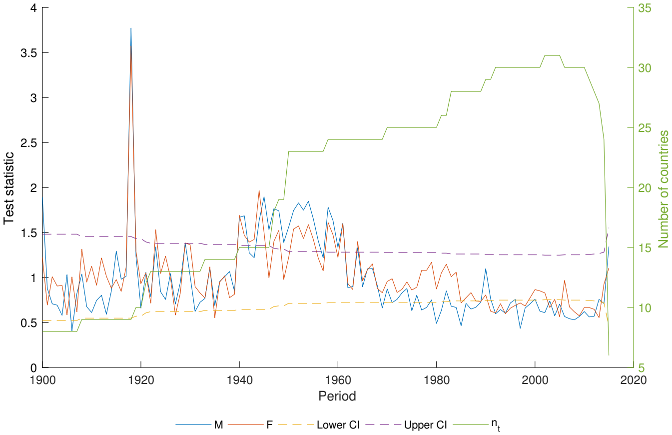
- These were estimated as random effects $\sim N(0, \sigma_\gamma^2)$
- We find no evidence of unit roots, but we do find significant first-order autocorrelation (on average $\Phi \approx 0.4$, but with significant variation from country to country)
- Rather than estimating residuals as we did in the cohort analysis, to maintain consistency with our estimation assumption, we set:

$$u^{\gamma}_{i,s,t} = \hat{\gamma}_{i,s,t}$$

Statistical properties of the period mortality shocks within country

- Very high correlation between males and females (average ≈ 0.767)
- But, strong evidence that the variance of these shocks has declined
- Hypothesis test:
  - Null: The variance of period effects before and after 1960 is the same.
  - Test statistic: $F_{i,s} = \frac{var(\hat{\gamma}_{i,s,t \leq 1960})}{var(\hat{\gamma}_{i,s,t > 1960})} \sim F(n_1, n_2)$, where n_1 is the number of periods before 1960 and n_2 the number after
  - Of the 24 countries with enough data, 21 reject for men and 17 for women, often with negligable p-values

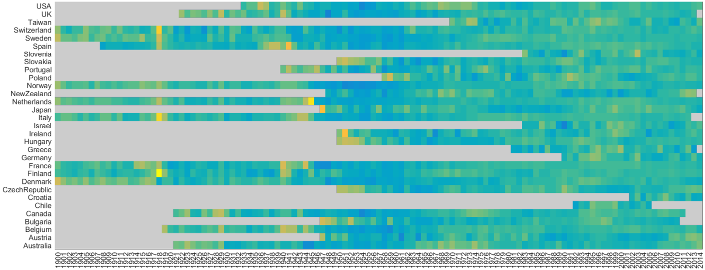
Testing for diminishing period effects across countries



- Confidence interval ↓ as number of countries (degree of freedom) ↑
- Picks up 1918 spanish flu
- Exceeding upper CI, then lower CI

Statistical properties of the period mortality shocks across countries

Various historical events can be identified, as well as broad patterns

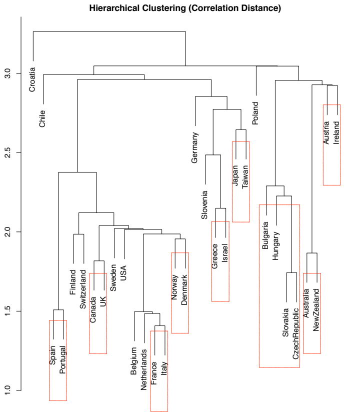


- 1918 flu epidemic
- World War II
- Dutch Hongerwinter
- Decline in importance of period effects

Clustering - period mortality shocks (γ)

Cultural effects seem less important, geographic poximity seems to matter more

- Anglosphere no longer falls into a single group
- Most groupings are geographical



Conclusion

We fitted an age-cohort model with period effects using Bayesian maximum a posteriori estimation to mortality data of men and women in 31 countries. We find:

- Diminishing variance of period effects in almost all countries
- Very high correlation in estimated period effects between men and women in the same country
- Very high correlation in cohort mortality shocks between men and women in the same country
- Cohort mortality seems to be influenced by both cultural and geographical effects; period shocks seem to be largely geographical

Q & A

Thank you!