Quebec Pension Plan (QPP) multi-population data analysis

Jie Wen
supervised by Prof. Andrew Cairns and Dr. Torsten Kleinow

Heriot-Watt University, Edinburgh

PhD in Actuarial Science, School of Mathematical and Computer Sciences

Fourteenth International Longevity Risk and Capital Market Solutions Conference

Amsterdam, Netherlands

Content

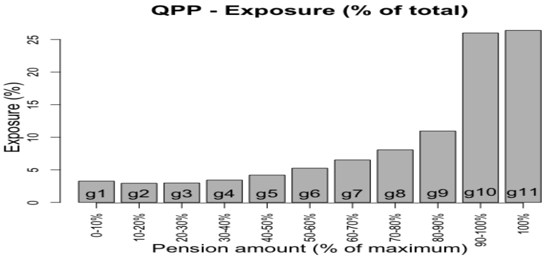
1. QPP Data Overview
2. Model Specification
3. Parameter Estimation and Model Selection
4. Fitting Diagnostics
5. Cluster Analysis
6. Summary
7. Q&A

QPP data overview

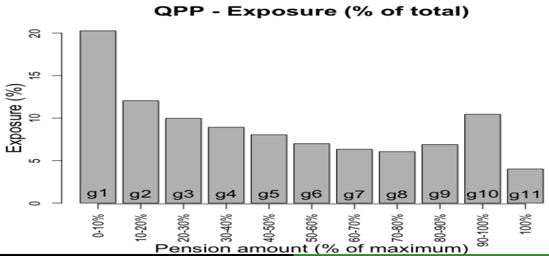
- 11 sub-populations ordered by increasing cohort pension amount in 10% bands.
- Only contains Quebec pensioners.
- Age over 65-89, and year over 1991-2015. $(11 \times 25 \times 25)$

QPP data overview (cont.)

- Males



- Females

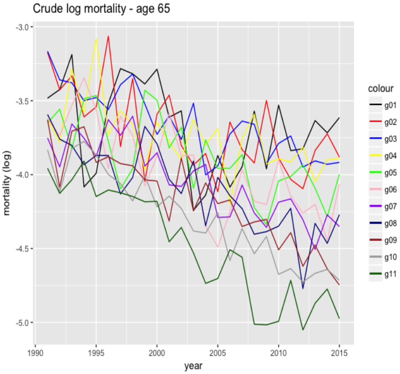


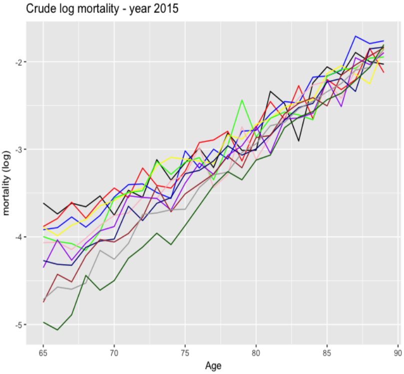
HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

QPP data overview (cont.)

- Group-wise crude death rates (log-scale):



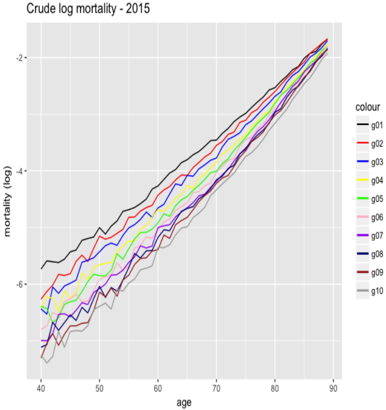


HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

QPP data overview (cont.)

- Comparison with **England IMD** (larger sample size, with groups evenly splited)



QPP data overview (cont.)

- **Age-Standardized Mortality Rate (ASMR)**

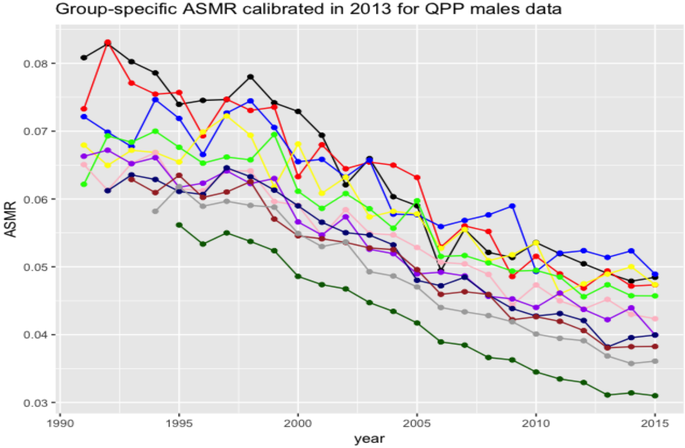
  - ASMR is a weighted average of the crude death rates over a defined age range, for certain specific calender year t.
  - E^s_x is the 'standard population' at age x (from European Standard Population, calibrated in 2013).
  - m_{tx} is crude death rate.
  - $$ASMR(t) = \frac{\sum_x m_{tx} E^s_x}{\sum_x E^s_x}$$
  - Use of ASMR:
    - comparison of mortality over different populations;
    - assessment of mortality term structure;
    - assessment of singal-to-noise ratio.

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

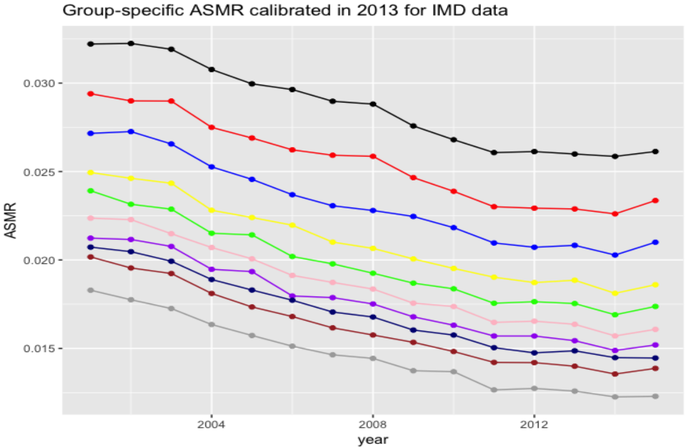
QPP data overview (cont.)

- ASMR of QPP males over age 65-89:



QPP data overview (cont.)

- (For comparison) ASMR of England IMD:



QPP data overview (cont.)

- ASMR is smoothier than the crude death rates, but still quite volatile for QPP males.
- Group 10 and 11 (larger size) are smoothier than others.
- Groups with higher pension tends to have lower mortality.
- QPP applies different grouping methodology (pension level) from England IMD (deprivation index) - less powerful predictor.

Model specification

m1 $\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti} + \beta^2_{xi}\kappa^2_{ti}$ (Renshaw and Haberman, 2003)

m2 $\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti} + \beta^2_{x}\kappa^2_{ti}$

m3 $\log m_{xti} = \alpha_{xi} + \beta^1_{x}\kappa^1_{t} + \beta^2_{xi}\kappa^2_{ti}$ (Li and Lee, 2005)

m4 $\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti}$ (Lee-Carter, 1992)

m5 $\log m_{xti} = \alpha_{xi} + \beta^1_{x}\kappa^1_{ti} + \beta^2_{x}\kappa^2_{ti}$ (CAE model by Kleinow, T, 2014)

m6 $\log m_{xti} = \alpha_{x} + \beta^1_{x}\kappa^1_{ti} + \beta^2_{x}\kappa^2_{ti}$ (CAE model with common α_x)

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Model specification

m1 $\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti} + \beta^2_{xi}\kappa^2_{ti}$ (Renshaw and Haberman, 2003)

m2 $\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti} + \beta^2_{x}\kappa^2_{ti}$

m3 $\log m_{xti} = \alpha_{xi} + \beta^1_{x}\kappa^1_{t} + \beta^2_{xi}\kappa^2_{ti}$ (Li and Lee, 2005)

m4 $\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti}$ (Lee-Carter, 1992)

m5 $\log m_{xti} = \alpha_{xi} + \beta^1_{x}\kappa^1_{ti} + \beta^2_{x}\kappa^2_{ti}$ (CAE model by Kleinow, T, 2014)

m6 $\log m_{xti} = \alpha_{x} + \beta^1_{x}\kappa^1_{ti} + \beta^2_{x}\kappa^2_{ti}$ (CAE model with common α_x)

- α, β and κ are stochastic parameters capturing age/period effect.
- α provides a form of base mortality table (while κ is zero).
- β determines the relative rates of mortality improvement at different ages.

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Model specification

m7 $\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$ (Plat, 2009)

m8 $\log m_{xti} = \alpha_{x} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{ti}^2$ (Plat model with common α_x)

m9 $\log m_{xti} = \alpha_{xi} + \kappa_{t}^1 + (x - \bar{x})\kappa_{ti}^2$ (Plat model with common κ_t^1)

m10 $\log m_{xti} = \alpha_{xi} + \kappa_{ti}^1 + (x - \bar{x})\kappa_{t}^2$ (Plat model with common κ_t^2)

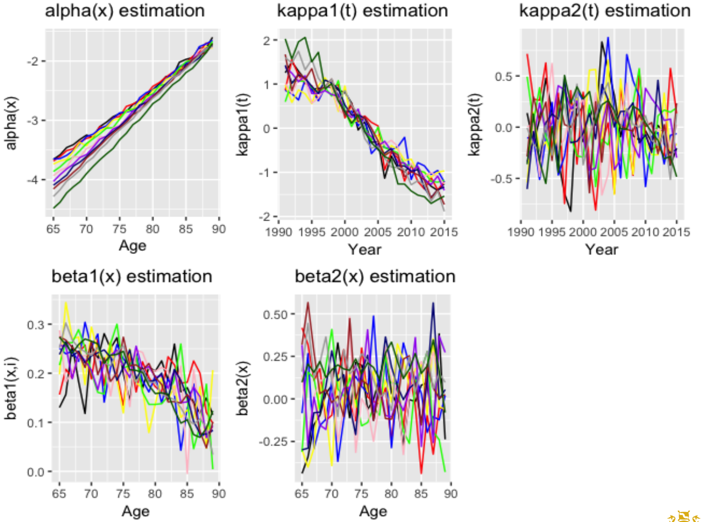
m11 $\log m_{xti} = \alpha_{xi} + \kappa_{t}^1 + (x - \bar{x})\kappa_{t}^2$ (Plat model with common κ_t^1 and κ_t^2)

Model specification (cont.)

- m1 is the 'basis' with most specified structure among all others.
- All other models are simplifications of m1.
- Parameters are estimated by Poisson assuption on number of deaths with Maximum Log-likelihood Estimation (MLE).

Parameter estimation and Model selection (cont.)

- Model m1 - estimated parameters (males)



HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Parameter estimation and Model selection (cont.)

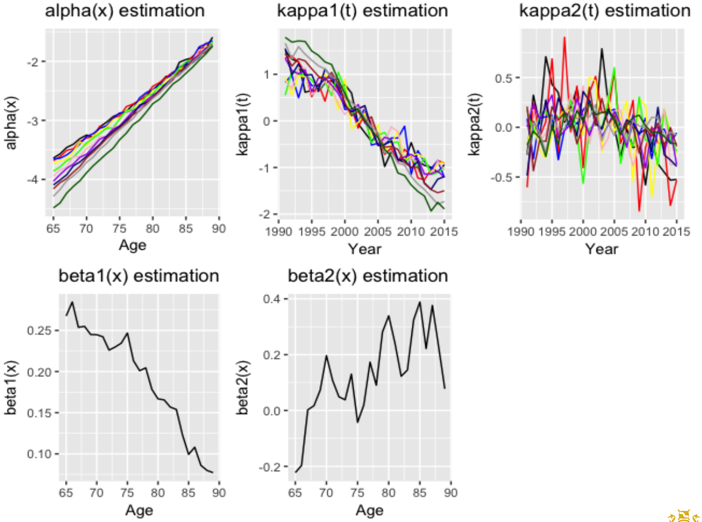
- Model m1:

$$\log m_{xti} = \alpha_{xi} + \beta^1_{xi}\kappa^1_{ti} + \beta^2_{xi}\kappa^2_{ti}$$

- Group specific α_{xi} gives observable group rankings.
- κ^1_{ti} and β^1_{xi} have decreasing pattern for all groups.
- κ^2_{ti} and β^2_{xi} are quite volatile.

Parameter estimation and Model selection (cont.)

- Model m5 - estimated parameters (males)



HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

- Model m5:

$$\log m_{xti} = \alpha_{xi} + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

- Group specific α_{xi} gives observable group rankings.
- κ_{ti}^1 has similar decreasing pattern for all groups.
- κ_{ti}^2 is quite volatile.
- β_x^1 decreases over age and is less volatile than β_x^2.

HERIOT WATT UNIVERSITY

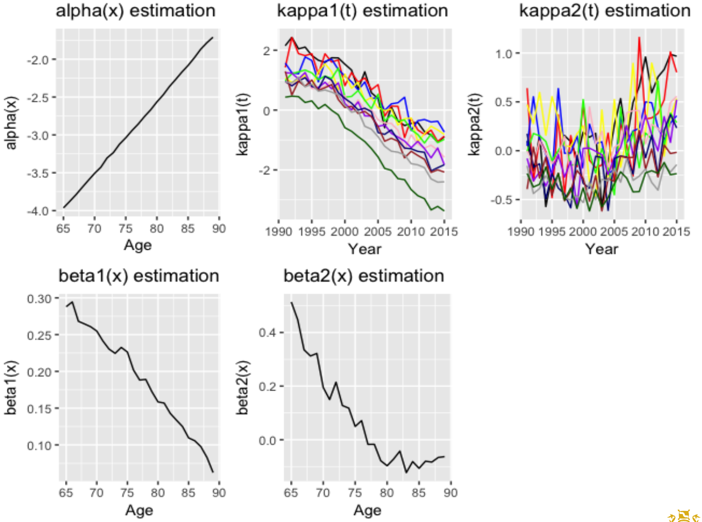
Actuarial Research Centre
Institute and Faculty of Actuaries

Parameter estimation and Model selection (cont.)

- Pattern of β^1 - model m5 (common - the grey fat solid line) and m1 (group-specific)

Parameter estimation and Model selection (cont.)

- Model m6 - estimated parameters (males)



Parameter estimation and Model selection (cont.)

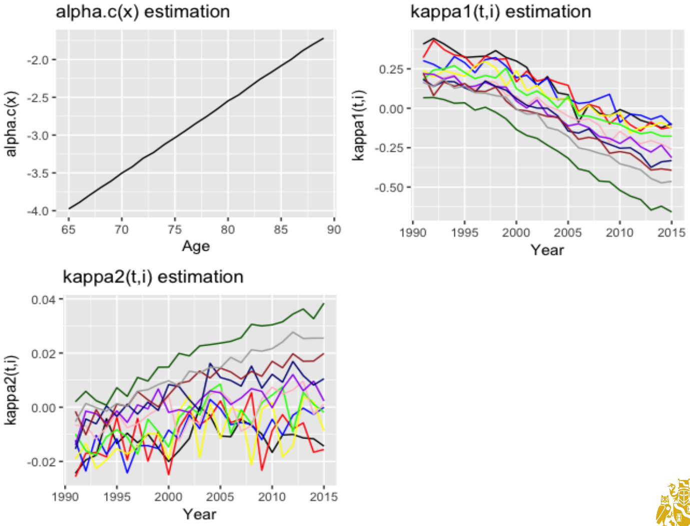
- Model m6:

$$\log m_{xti} = \alpha_x + \beta_x^1 \kappa_{ti}^1 + \beta_x^2 \kappa_{ti}^2$$

- As α_x is common, variations between subgroups are captured by κ_{ti}^1 and κ_{ti}^2.
- Group 11 stands clear of others in terms of κ_{ti}^1.
- β_x^1 and β_x^2 decreases over age, β_x^2 is smoothier than under m5.

Parameter estimation and Model selection (cont.)

- Model m8 - estimated parameters (males)



HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Parameter estimation and Model selection (cont.)

- Model m8:

$$\log m_{xti} = \alpha_x + \kappa_{ti}^1 + \kappa_{ti}^2(x - \bar{x})$$

- As α_x is common, variations between subgroups are captured by κ_{ti}^1 and κ_{ti}^2.
- Group 11 stands well below and above others for κ_{ti}^1 and κ_{ti}^2 respectively.

Parameter estimation and Model selection (cont.)

- Model selection criteria: log-likelihood and BIC: males

Bayes Information Criterion (BIC) is a statistic based on log-likelihood that penalises over-parameterized models and is used as a purely numerical criterion for selecting out the best model (m8).

Model	log-likelihood	# parameters	df	BIC
m1	-22,252.44	1375	1331	56,265.12
m5	-22,628.04	875	851	52,775.22
m6	-22,771.52	625	621	51,029.98
m8	-22,867.36	575	573	50,797.55

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Parameter estimation and Model selection (cont.)

- Model selection criteria: log-likelihood and BIC: males

Bayes Information Criterion (BIC) is a statistic based on log-likelihood that penalises over-parameterized models and is used as a purely numerical criterion for selecting out the best model (m8).

Model	log-likelihood	# parameters	df	BIC
m1	-22,252.44	1375	1331	56,265.12
m5	-22,628.04	875	851	52,775.22
m6	-22,771.52	625	621	51,029.98
m8	**-22,867.36**	**575**	**573**	**50,797.55**

HERIOT WATT UNIVERSITY

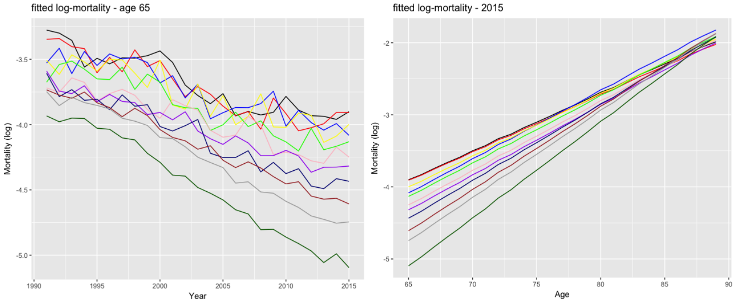
Actuarial Research Centre
Institute and Faculty of Actuaries

Diagnostic on fitting results

- m8 has fewest parameters and better BIC than other three.
- More parameters improves log-likelihood but is also penalized for over-parameterization.
- Greater complexity does not necessarily improve fitting significantly.
- Additional diagnostic is also required for selection.

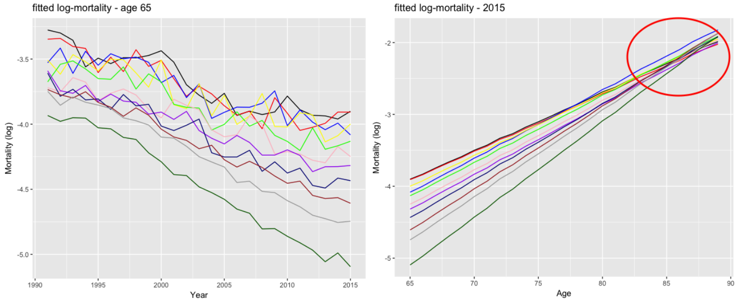
Diagnostic on fitting results

- Fitted mortalities (log-scale) from model m8 (males)



Diagnostic on fitting results

- Fitted mortalities (log-scale) from model m8 (males)

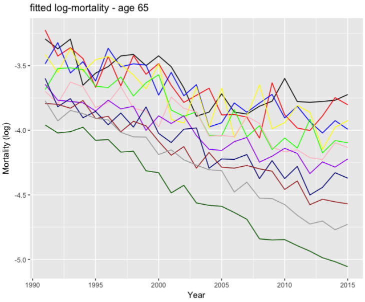


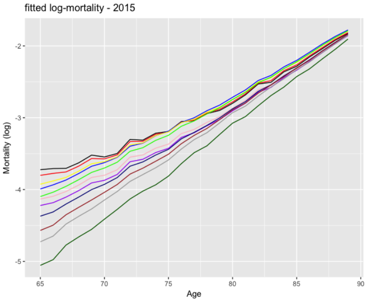
HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Diagnostic on fitting results

- Fitted mortalities (log-scale) from model m6 (males)





Diagnostic on fitting results

- Standardized Residuals

$$Z_{txi} = \frac{D_{txi} - E_{txi}\hat{m}_{txi}}{\sqrt{E_{txi}\hat{m}_{txi}}}$$

- Measures standardized difference between crude and estimated figures.
- Not affected by absolute scale of observations.
- Well-fitted model is expected to have random standardized residuals.

Diagnostic on fitting results

- Standardized residuals from m6: QPP males

Diagnostic on fitting results

- Standardized residuals from m8: QPP males

Diagnostic on fitting results

- Both m6 and m8 have quite random standardized residuals.
- There is no significant non-random cluster along x-axis (year), y-axis (age) or diagonal (cohort).
- m6 doesn't have significant crossover in fitted mortality curves. m8 has crossovers at high ages.
- m6 is selected as the most suitable model for QPP males.

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Cluster analysis on QPP data

- QPP has relatively small population size.
- Subpopulations are not evenly grouped.
- Crude mortalities are quite volatile.
- Some adjacent groups typically have quite similar levels of mortality.
- We consider to re-cluster the QPP dataset.

Cluster analysis on QPP data

- Algorithm:

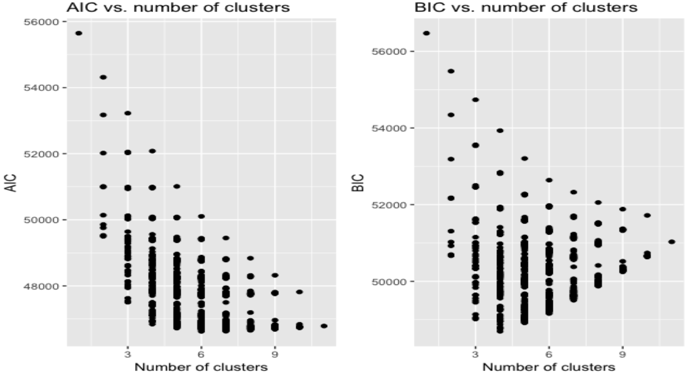
1. Restructure the data by combining neighbouring groups into clusters. Each cluster could contain $1, 2, \ldots, 11$ groups.
2. We obtain new restructured datasets with ≤ 11 groups.
3. There are $1,024$ different combinations in total. $(\sum_{i=0}^{11-1} C_{11-1}^{i})$
4. Fit underlying models to each reclustered dataset.

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Cluster analysis on QPP data

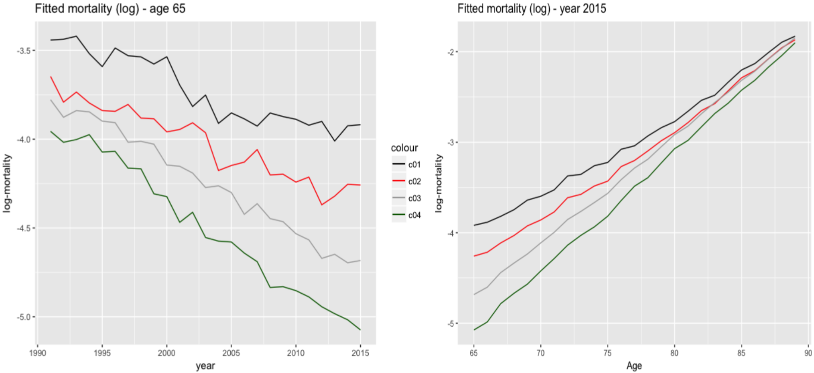
- AIC and BIC for all $1,024$ cluster combinations fitted for model m6:



BIC is $48,694.69$ under the optimized scenario (used to be $51,029.98$).

Cluster analysis on QPP data

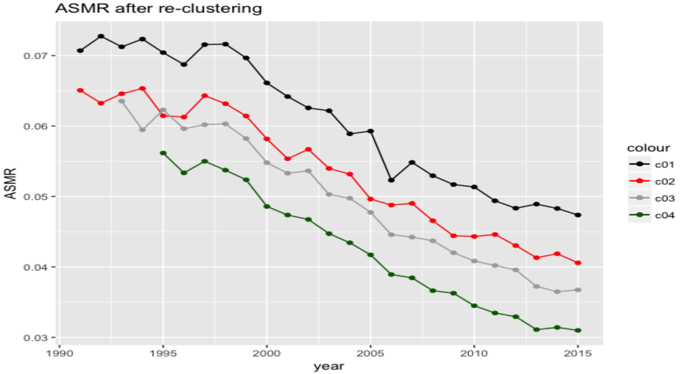
- Fitted mortalities (log-scale) from m6 after re-clustered into 4 groups.



HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

- ASMR of QPP males after re-clustering:



Cluster analysis on QPP data

- Conclusion from cluster analysis:

  - All models suggest the same optimal clustering by BIC - with 4 clusters:
    - Cluster 1: group 1-5;
    - Cluster 2: group 6-8;
    - Cluster 3: group 9 and 10;
    - Cluster 4: group 11.
  - Volatilities are reduced significantly.
  - It enables us to see more clearly the different trends of clusters.

HERIOT WATT UNIVERSITY

Actuarial Research Centre
Institute and Faculty of Actuaries

Summary

- For volatile population, models with simpler structure fits better, i.e. model m6 and m8 over m1.
- Besides quantitative criteria, qualitative criteria like graphical diagnostics are the same important.
- Clustering improves fitting quality and signal-to-noise ratio.
- Future researches: Smoothing of modelling results; More detailed cluster analysis; Long-term mortality projection.

ANY QUESTIONS ?