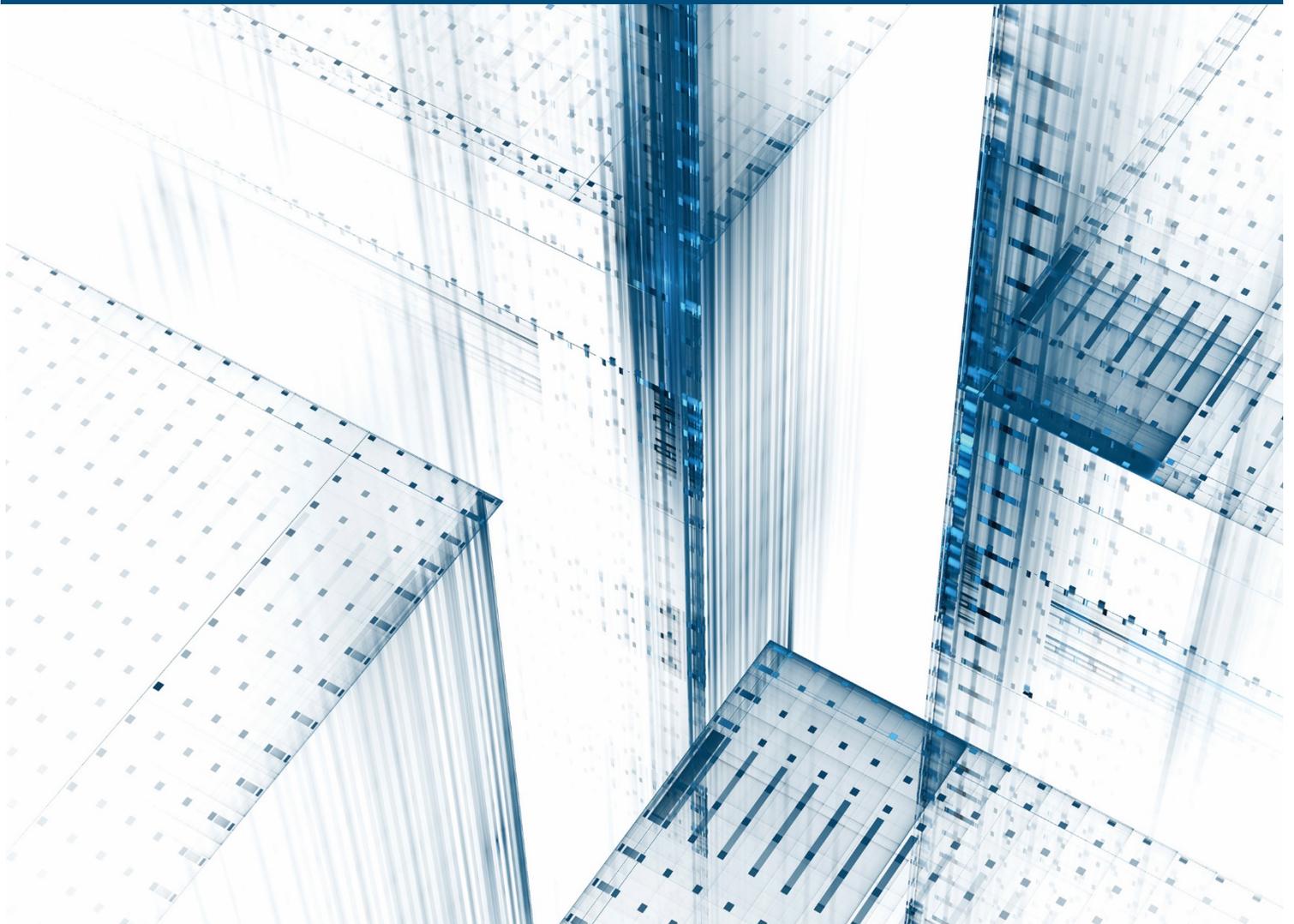


Components of Historical Mortality Improvement

Volume 2 – Mortality Rate Modeling and Conclusions





Components of Historical Mortality Improvement

Volume 2 – Mortality Rate Modeling and Conclusions

AUTHOR Johnny S.-H. Li, PhD, FSA
Rui Zhou, PhD, FSA
Yanxin Liu, PhD

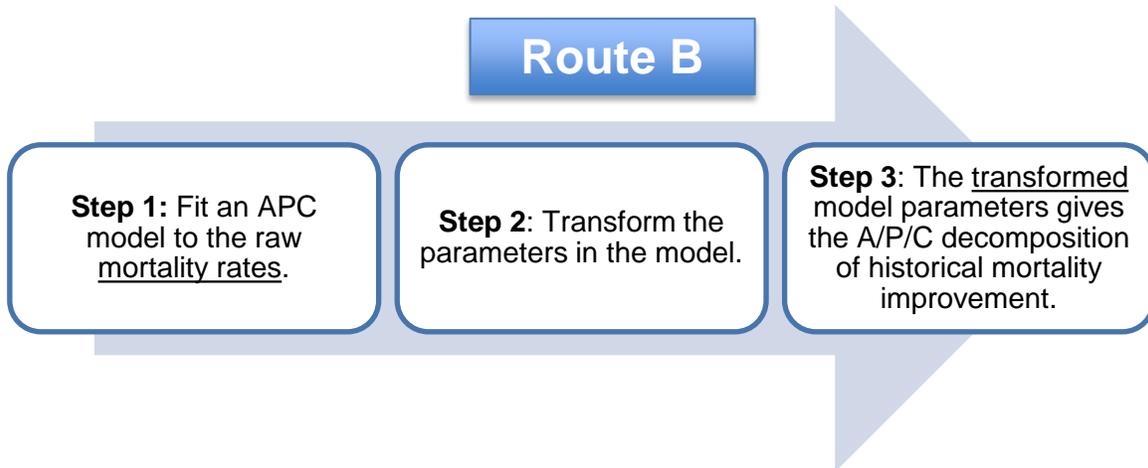
Caveat and Disclaimer

The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries or its members. The Society of Actuaries makes no representation or warranty to the accuracy of the information

Copyright © 2017 by the Society of Actuaries. All rights reserved.

Executive Summary

This volume of the report on the “Components of Historical Mortality Improvement” project documents the modeling work associated with Route B. It also compares the results from Routes A and B, and draw a final conclusion. To recapitulate, in Route B, APC models are fitted to mortality rates and the desired A/P/C decomposition of mortality improvement experience is obtained by transforming the parameters in the chosen APC model. The steps involved in Route B are shown in the flow chart below:



Route B encompasses the newer CMI method (the CMI-17 method), which is documented in the CMI Working Papers 97, 98 and 99 (CMI, 2017a,b,c).

Section 1 of this volume summarizes the CMI-17 method. Sections 2 and 3 implement and evaluate the CMI-17 method the U.S. data sets. It is found that the CMI-17 method is reasonably robust relative to (1) changes in the calibration window, (2) changes in the age range, (3) changes in the parameter constraints used, and (4) inclusion/exclusion of the oldest/newest cohorts. However, when applied to the U.S. male data sets, the CMI-17 method yields large residual clusters, which indicate that the APCI model (on which the CMI-17 method is based) is unable to pick up some features that are specific to the U.S. male mortality improvement experience.

Section 4 studies if the CMI-17 method may be improved by considering alternative, more sophisticated APC model structures for modeling mortality rates. Eight candidate model structures are examined: M2, M3, M6, M7, M8, the full Plat model, the simplified Plat model and the APCI model (on which the CMI-17 method is based).¹ We first perform a range of robustness tests to shortlist a smaller number of model structures that merit further consideration. We then examine the standardized residuals produced by the shortlisted models to identify the most effective model. Highlights of the model evaluation results are as follows:

- M2 and M8 show low robustness in some tests, and are therefore not given further consideration.

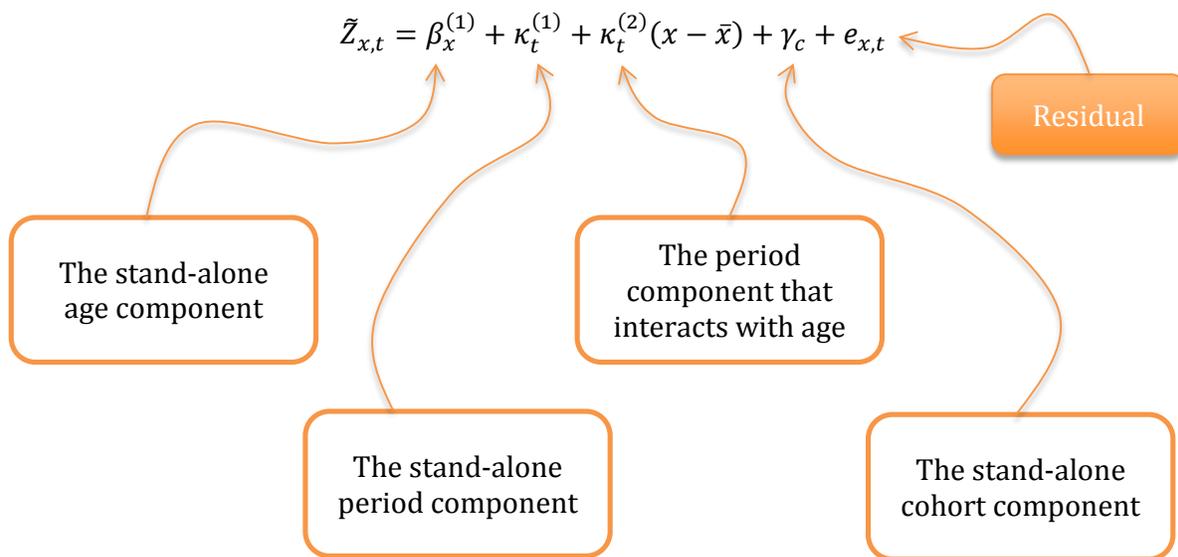
¹We exclude M1 (the Lee-Carter model) and M5 (the Cairns-Blake-Dowd model), because these models do not incorporate cohort effects. We also exclude M4 as it does not explicitly decompose historical mortality into APC components. Details concerning the models under consideration can be found in the papers by Cairns et al. (2009) and Plat (2009).

- M6 and M7 perform the worst in the residual analysis. Large horizontal clusters are found in the heat maps produced from these models, indicating that age effect is not adequately captured.
- Residuals clustering is significant in M3 and the APCI model for males. Large vertical clusters are observed between mid-80s and mid-90s. Note that a similar problem is found in the M3 for males when is Route A is used (see Volume 1 of the project report).
- The full Plat model performs the best in the residual analysis, closely followed by the simplified Plat model. However, the simplified Plat model outperforms the full Plat model in the robustness tests.

On the basis of the model evaluation results, the simplified Plat model appears to be the most effective A/P/C decomposition model when Route B is pursued.

Section 5 repeats the analyses using the data for ages 55 to 95 only. The results suggest that the simplified Plat model still performs well even when the data for younger ages are discarded.

Section 6 compares the optimal A/P/C decomposition results obtained from Routes A and B. The two optimal decomposition results are not too different. We choose **Route A simplified Plat model** as our final recommendation, on grounds that its smoothing parameters are optimized with a more statistically rigorous approach. The recommended model implies that mortality improvement $Z_{x,t}$ for age x and year t is driven by (1) a stand-alone age component, (2) a stand-alone period component, (3) a period component that interacts with (a linear function) of age, and (4) a stand-alone cohort component:



The numerical values of the identified A/P/C components are provided in the accompanying Excel workbook.

Acknowledgements

The SOA would like to extend their gratitude to the Project Oversight Group:

- Jennifer A. Haid, FSA (Chair)
- Jean-Marc Fix, FSA, MAAA
- Zachary Granovetter, FSA
- George A. Graziani, FSA, FCIA
- Robert C. W. Howard, FSA, FCIA
- Allen M. Klein, FSA, MAAA
- Alla Kleyner, FSA
- Laurence Pinzur, FSA
- Larry N. Stern, FSA, MAAA

The authors would like to thank the SOA staff for their leadership and coordination of the project:

- R. Dale Hall, FSA, MAAA, CERA
- Andrew J. Peterson, FSA, EA, MAAA, FCA
- Erika Schulty

Contents

1	A Summary of the CMI-17 Method	5
1.1	Model Structure	5
1.2	The Estimation Procedure	5
1.3	Choosing the Smoothing Parameters	6
2	Implementing the CMI-17 Method	7
2.1	Fitted vs. Crude Death Rates	7
2.2	The Age/Period/Cohort Components of the Mortality Improvement Rates	12
2.3	The Standardized Residuals	14
3	Testing the Robustness of the CMI-17 Method	16
3.1	Changes in the Calibration Window	16
3.2	Changes in the Age Range Used	18
3.3	Parameter Constraints Used	20
3.4	Inclusion/Exclusion of the Oldest/Newest Cohorts	22
4	Improving the CMI-17 Method	24
4.1	The Candidate APC Model Structures	24
4.1.1	Definitions and Overview	24
4.1.2	Mortality Rate Models	26
4.1.3	Parameter Estimation	27
4.1.4	A/P/C Components of Historical Mortality improvement	28
4.2	Selecting Penalty Parameters	31
4.3	Robustness Tests	33
4.3.1	Defining the Robustness Measure	33
4.3.2	Changes in the Tolerance Value Used in Optimizing Model Parameters	35
4.3.3	Changes in the Calibration Window	38
4.3.4	Changes in the Age Range Used	40
4.3.5	Choice of Parameter Constraints	43

4.3.6	Exclusion of the Oldest and Newest Cohorts	47
4.4	Analyzing the Standardized Residuals Produced by the Shortlisted Models	51
4.5	Concluding Remarks	52
5	Repeating the Analyses Using Data for Ages 55 to 95	56
6	Comparing Routes A and B	58
6.1	A Summary of the Modeling Work for Route A	58
6.2	A Summary of the Modeling Work for Route B	60
6.3	The Final Conclusion	62
7	Acknowledgments	65

1 A Summary of the CMI-17 Method

1.1 Model Structure

Let $m_{x,t}$ be the underlying central death rates at age x and in year t . The CMI-17 method is based on the APCI model, which describes the logarithm of $m_{x,t}$ as follows:

$$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)}(t - \bar{t}) + \kappa_t^{(1)} + \gamma_c,$$

where

- \bar{t} is the mid-point of the sample period used (for the full data set under consideration, $\bar{t} = (1968 + 2014)/2 = 1991$);
- $\beta_x^{(1)}$ is a stand-alone age-specific parameter;
- $\beta_x^{(2)}$ is an age-specific parameter that interacts with a linear function of time;
- $\kappa_t^{(1)}$ is a stand-alone time-varying parameter; and
- γ_c is a cohort-related parameter.

1.2 The Estimation Procedure

As with what was done in the CMI Working Paper No. 98, we estimate and smooth parameters in the APCI model simultaneously by minimizing a single objective function, which is formulated as follows:

$$\text{Objective} = \text{Deviance} + \text{Penalty}(\beta_x^{(1)}) + \text{Penalty}(\beta_x^{(2)}) + \text{Penalty}(\kappa_t^{(1)}) + \text{Penalty}(\gamma_c).$$

The deviance measures the goodness-of-fit. The smaller the deviance is, the better the goodness-of-fit is. Let $D_{x,t}$ be the number of deaths at age x and in year t , and $E_{x,t}$ be the corresponding number of exposures-to-risk. Assuming $D_{x,t}$ is a realization of a Poisson distribution, i.e., $D_{x,t} \sim \text{Poisson}(E_{x,t}m_{x,t})$, the deviance can be calculated with the following formula:

$$\text{Deviance} = 2 \sum_{x=x_0}^{x_1} \sum_{t=t_0}^{t_1} \left[D_{x,t} \ln \frac{D_{x,t}}{E_{x,t}m_{x,t}} - (D_{x,t} - E_{x,t}m_{x,t}) \right].$$

The other terms in the objective function are roughness penalty terms. The more jagged a parameter series is, the higher the penalty term for the parameter series is. The penalty terms can be calculated with the following formulas:

$$\text{Penalty}(\beta_x^{(1)}) = \lambda_{\beta^{(1)}} \sum_{x=x_0}^{x_1} \left(\beta_x^{(1)} - 3\beta_{x-1}^{(1)} + 3\beta_{x-2}^{(1)} - \beta_{x-3}^{(1)} \right)^2,$$

$$\text{Penalty}(\beta_x^{(1)}) = \lambda_{\beta^{(1)}} \sum_{x=x_0}^{x_1} \left(\beta_x^{(1)} - 3\beta_{x-1}^{(1)} + 3\beta_{x-2}^{(1)} - \beta_{x-3}^{(1)} \right)^2,$$

$$\text{Penalty}(\kappa_t^{(1)}) = \lambda_{\kappa^{(1)}} \sum_{t=t_0}^{t_1} \left(\kappa_t^{(1)} - 2\kappa_{t-1}^{(1)} + \kappa_{t-2}^{(1)} \right)^2,$$

and

$$\text{Penalty}(\gamma_c) = \lambda_{\gamma} \sum_c \left(\gamma_{c=t_0-x_1}^{t_1-x_0} - 3\gamma_{c-1} + 3\gamma_{c-2} - \gamma_{c-3} \right)^2,$$

where $\lambda_{\beta_x^{(1)}}$, $\lambda_{\beta_x^{(2)}}$, $\lambda_{\kappa_t^{(1)}}$ and λ_{γ_c} are penalty parameters which determine the degrees of smoothness of the parameters.

The objective function is optimized by an iterative Newton's method, in which parameters are updated one at a time. The following parameter constraints are applied at the end of each iteration to stipulate parameter uniqueness:

$$\sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0, \quad \sum_{c=t_0-x_1}^{t_1-x_0} c\gamma_c = 0, \quad \sum_{c=t_0-x_1}^{t_1-x_0} c^2\gamma_c = 0, \quad \sum_{t=t_0}^{t_1} \kappa_t = 0, \quad \sum_{t=t_0}^{t_1} t\kappa_t = 0.$$

1.3 Choosing the Smoothing Parameters

We select the smoothing parameters one at a time. The following procedure is used to select the penalty parameter for $\beta_x^{(1)}$:

1. Set all penalty parameters except $\lambda_{\beta_x^{(1)}}$ to zero.
2. Plot the estimated values of $\beta_x^{(1)}$ for $\lambda_{\beta_x^{(1)}} = 10^0, 10^1, 10^2, \dots, 10^{10}$.
3. The optimal value of $\lambda_{\beta_x^{(1)}}$ should be the one that removes the jaggedness and largely keeps the shape of the unsmoothed series.

Similar procedures are used to select the penalty parameters for other parameter series. Further details concerning the choice of smoothing parameters are provided in Section 4.2 where a general procedure for choosing smoothing parameters in Route B models is presented.

2 Implementing the CMI-17 Method

2.1 Fitted vs. Crude Death Rates

Figures 1 and 2 show the crude and fitted central death rates (in log scale) across ages for three specific years-of-birth: 1930, 1950 and 1970.² Two collections of fitted death rates are shown simultaneously:

- Fitted death rates that are not smoothed

They are calculated using the APCI model parameters that are optimized given

$$\lambda_{\beta_x^{(1)}} = \lambda_{\beta_x^{(2)}} = \lambda_{\kappa_t^{(1)}} = \lambda_{\gamma_c} = 0.$$

- Fitted death rates that are optimally smoothed

They are calculated using the APCI model parameters that are optimized given the chosen smoothing parameters

$$\lambda_{\beta_x^{(1)}} = 10^6, \lambda_{\beta_x^{(2)}} = 10^7, \lambda_{\kappa_t^{(1)}} = 10^6, \lambda_{\gamma_c} = 10^6.$$

The comparisons made in Figures 1 and 2 indicates that the APCI model gives a satisfactory fit. When the smoothing parameters are applied, the jaggedness is largely removed while the locations of the peaks and troughs are retained. The results from the HMD and SSA data sets are quite similar.

Figures 3 and 4 show the raw and smoothed central death rates (in log scale) across years-of-birth for three ages: 35, 65 and 95. As in Figures 1 and 2, we show both non-smoothed and optimally smoothed fitted values. The rates from the HMD and SSA data sets are broadly in line, but at age 95 the SSA rates are notably higher. This distinction was noted in the previous volume of the project report.

²Not all rates in the age range of 20 to 95 are available. For instance, as the data sets cover calendar years 1968 to 2014 only, for year-of-birth 1930 the available central death rates span age 38 (= 1968 – 1930) to age 84 (= 2014 – 1930) only.

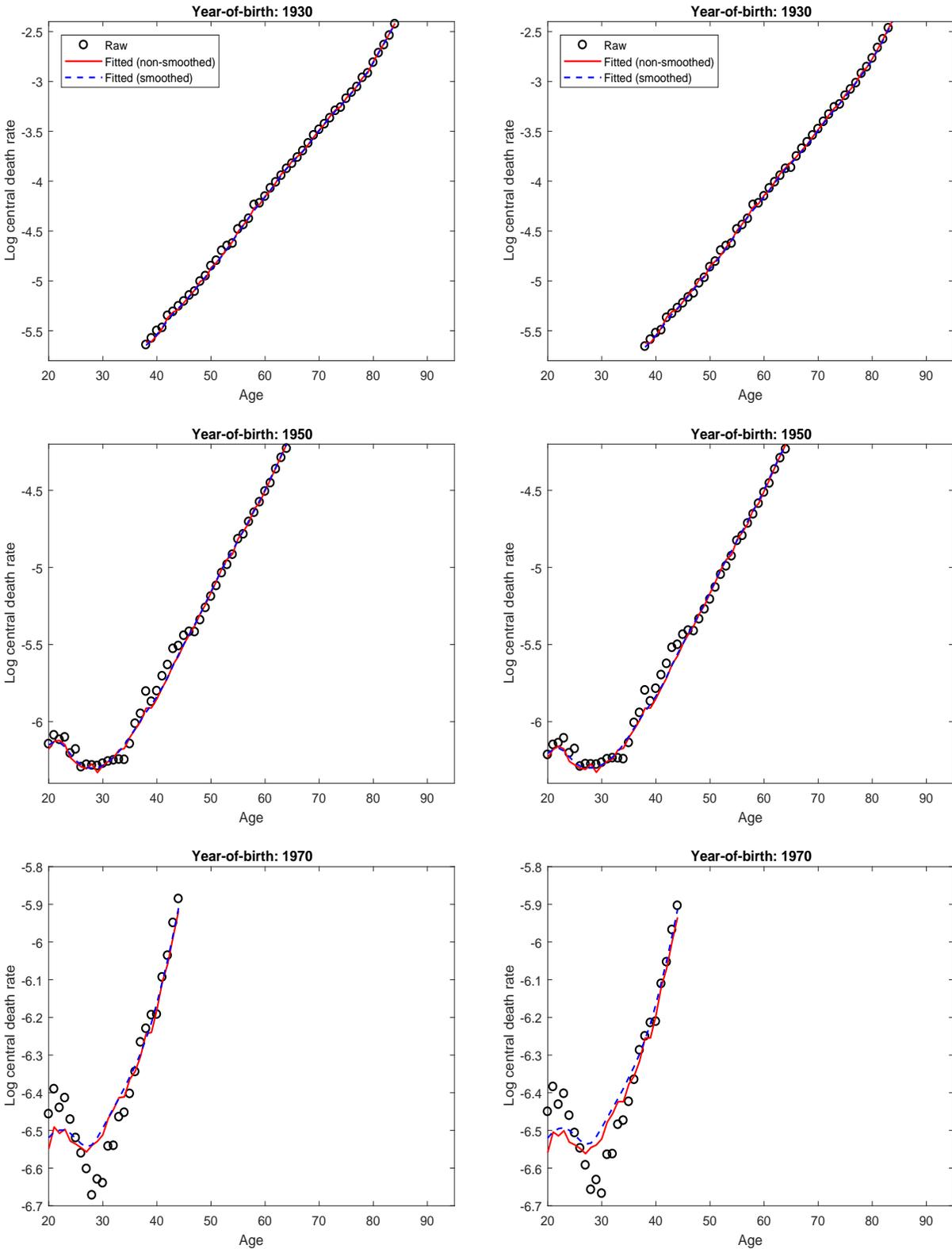


Figure 1: The raw and fitted mortality rates of U.S. **males** born in year 1930 (upper panels), year 1950 (middle panels) and year 1970 (lower panels).
 - Left panels: Data set (i) (HMD, ages 20-95, years 1968-2014).
 - Right panels: Data set (ii) (SSA, ages 20-95, years 1968-2014).

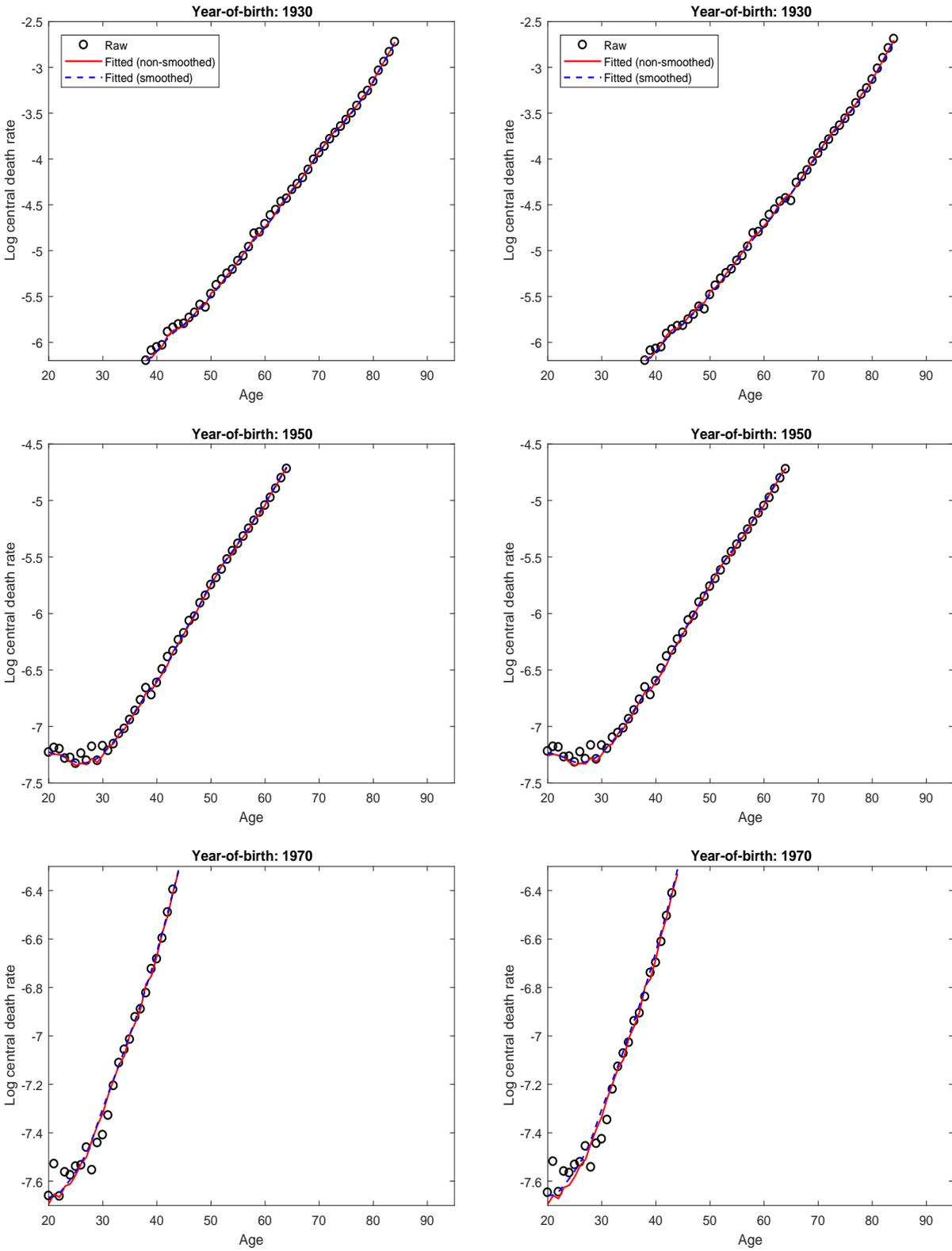


Figure 2: The raw and fitted mortality rates of U.S. **females** born in year 1930 (upper panels), year 1950 (middle panels) and year 1970 (lower panels).
 - Left panels: Data set (i) (HMD, ages 20-95, years 1968-2014).
 - Right panels: Data set (ii) (SSA, ages 20-95, years 1968-2014).

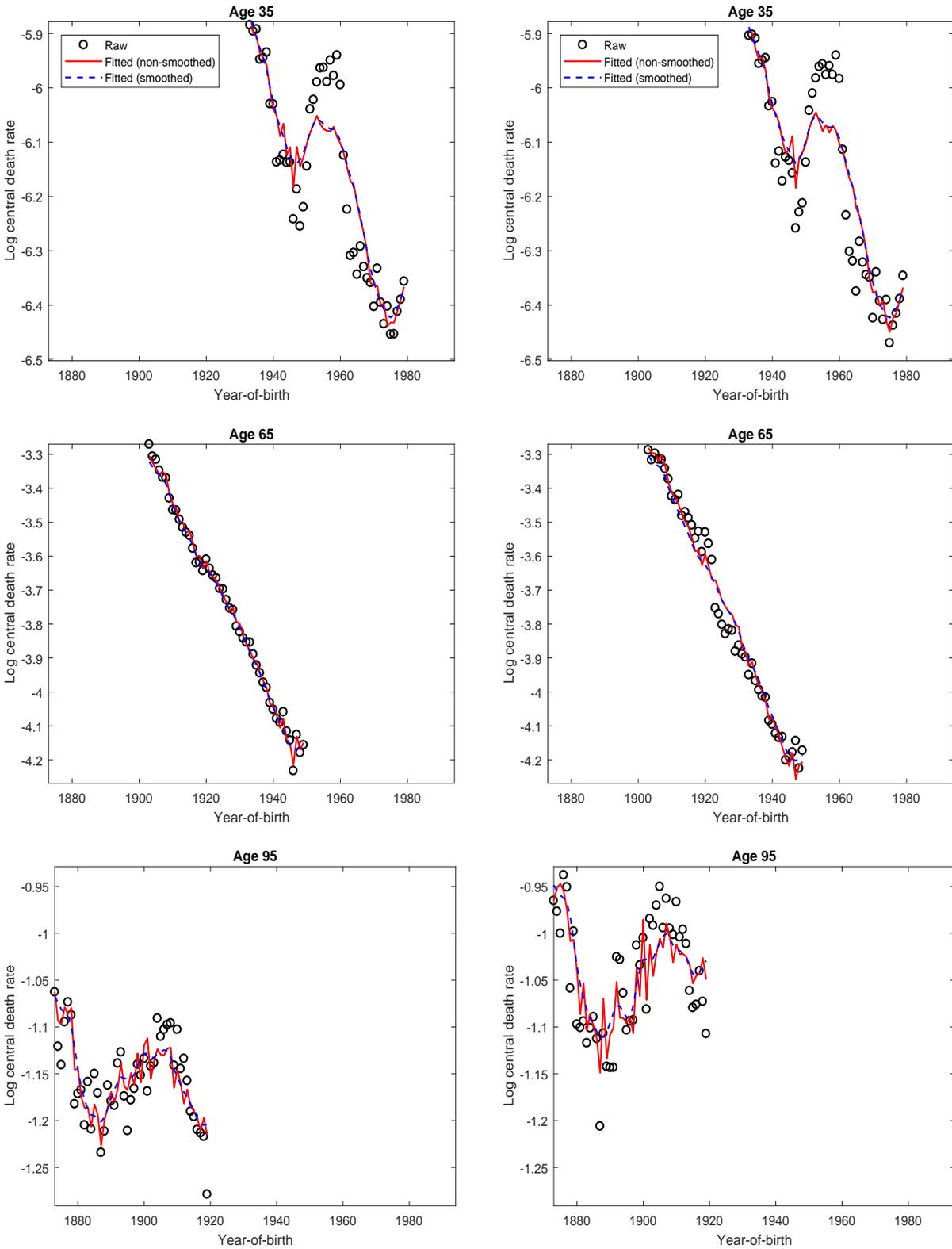


Figure 3: The raw and fitted mortality rates of U.S. **males** at age 35 (upper panels), age 65 (middle panels) and age 95 (lower panels).
 - Left panels: Data set (i) (HMD, ages 20-95, years 1968-2014).
 - Right panels: Data set (ii) (SSA, ages 20-95, years 1968-2014).

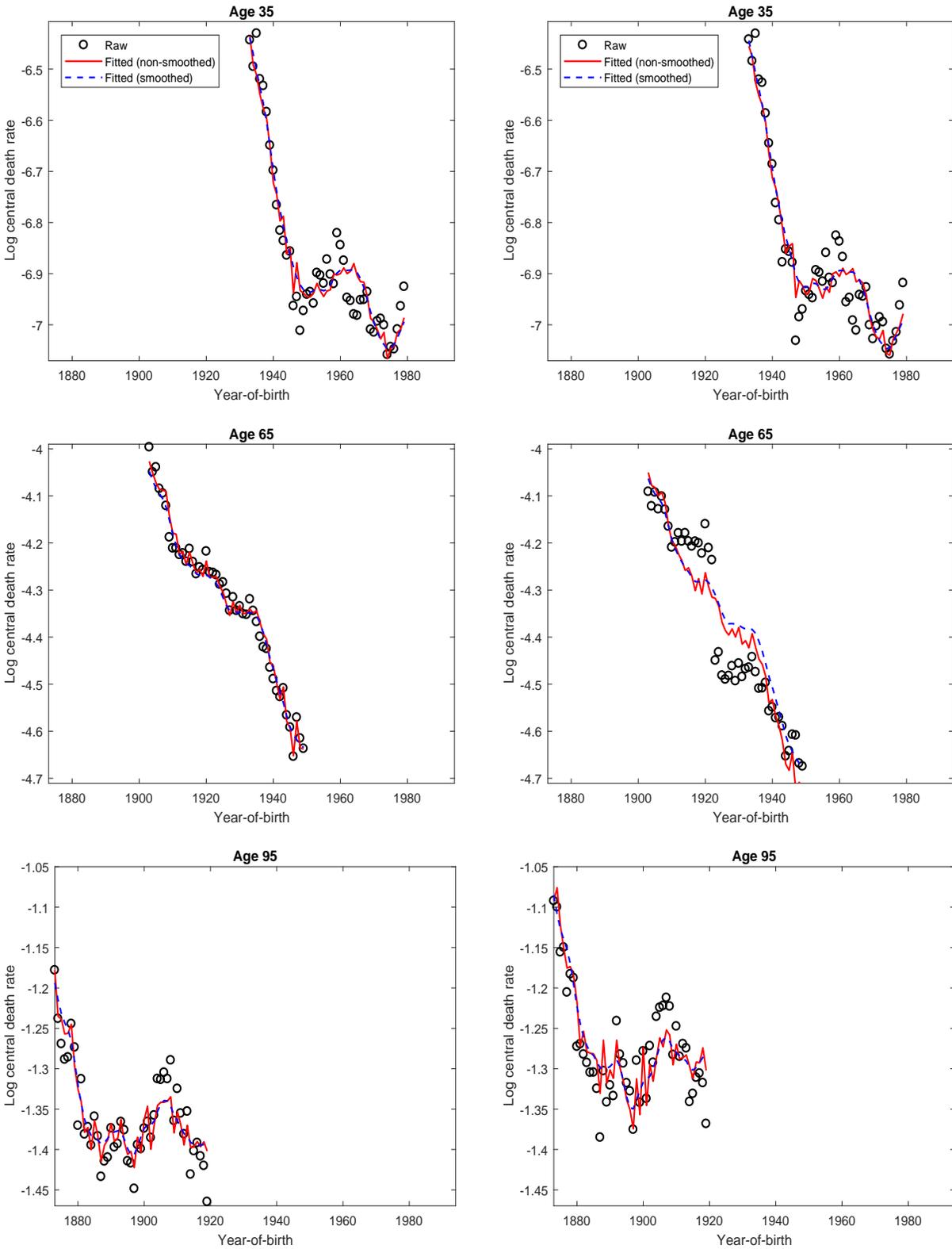


Figure 4: The raw and fitted mortality rates of U.S. **females** at age 35 (upper panels), age 65 (middle panels) and age 95 (lower panels).
 - Left panels: Data set (i) (HMD, ages 20-95, years 1968-2014).
 - Right panels: Data set (ii) (SSA, ages 20-95, years 1968-2014).

2.2 The Age/Period/Cohort Components of the Mortality Improvement Rates

The APCI model in the CMI-17 method implies that mortality improvement (in terms of the change in log central death rates) can be decomposed as follows:

$$MI_{x,t} = \ln(m_{x,t-1}) - \ln(m_{x,t}) = B_x^{(\cdot)} + K_t^{(1)} + G_c,$$

where

- $B_x^{(\cdot)} = -\beta_x^{(2)}$ is the age component;
- $K_t^{(1)} = \kappa_{t-1} - \kappa_t$ is the period component;
- $G_c = \gamma_{t-x-1} - \gamma_{t-x}$ is the cohort component.

A positive value of $MI_{x,t}$ means that mortality at age x is smaller in year t than in year $t - 1$. Figures 5 and 6 show the estimated age, period and cohort components of historical mortality improvement for U.S. males and females, respectively. As expected, the two data sets yield similar A/P/C decompositions.

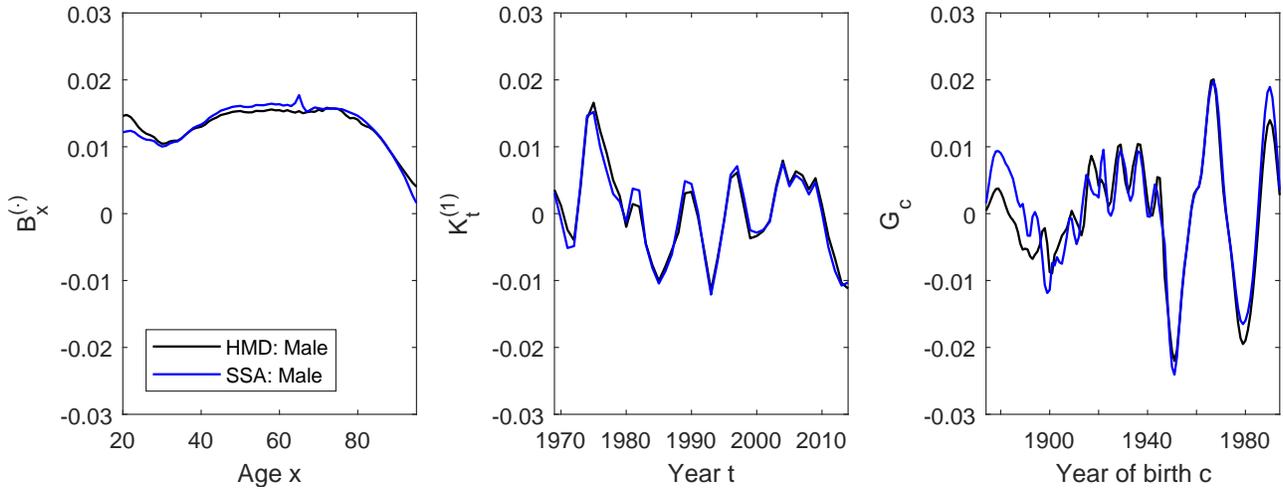


Figure 5: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model (U.S. **males**).

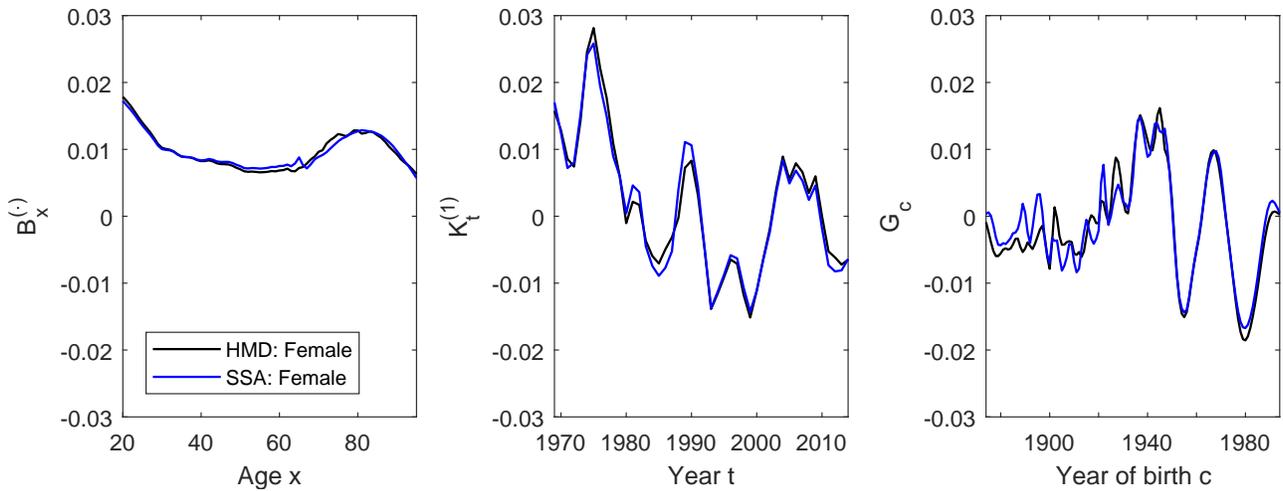


Figure 6: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model (U.S. **females**).

2.3 The Standardized Residuals

We now analyze the standardized residuals produced by the CMI-17 APCI models. The standardized residuals are calculated using the following formula:

$$\frac{D_{x,t} - E_{x,t}\hat{m}_{x,t}}{\sqrt{E_{x,t}\hat{m}_{x,t}}},$$

where $\hat{m}_{x,t}$ represents the fitted value of $m_{x,t}$ produced by the APCI model.

Because the APCI model is estimated using a Poisson death count assumption, the standardized residuals are not necessarily normally distributed even if the model is adequate. As such, we cannot apply normality tests and the ‘normal q-q plot’ here. However, we can still assess the adequacy of the APCI model by examining the heat map of its standardized residuals. If the model is adequate, then the pattern its standardized residuals should be random with little clustering.

Figure 7 and 8 show the standardized residuals from the estimated APCI models for U.S. males and females, respectively. For U.S. males, three large vertical clusters are observed in the age range of 20-50 from mid-80s and mid-00s, suggesting that the APCI model does not adequately the age-period interaction for this particular population. For U.S. females, the standardized residuals look reasonably random.

Note that the standardized residuals shown in Figure 7 and 8 cannot be directly compared and contrasted with those shown in the previous volume, which are computed in a different manner.

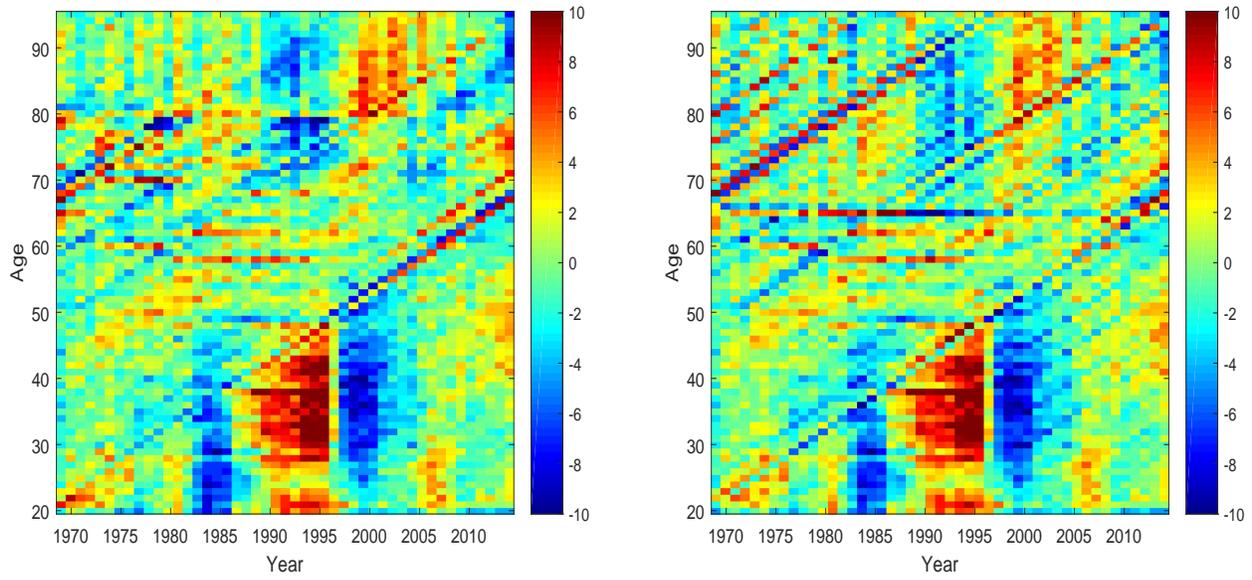


Figure 7: Heat maps of the standardize residuals calculated from the CMI-17 APCI model for U.S. **males**.
 - Left panel: Data set (i) (HMD, ages 20-95, years 1968-2014).
 - Right panel: Data set (ii) (SSA, ages 20-95, years 1968-2014).

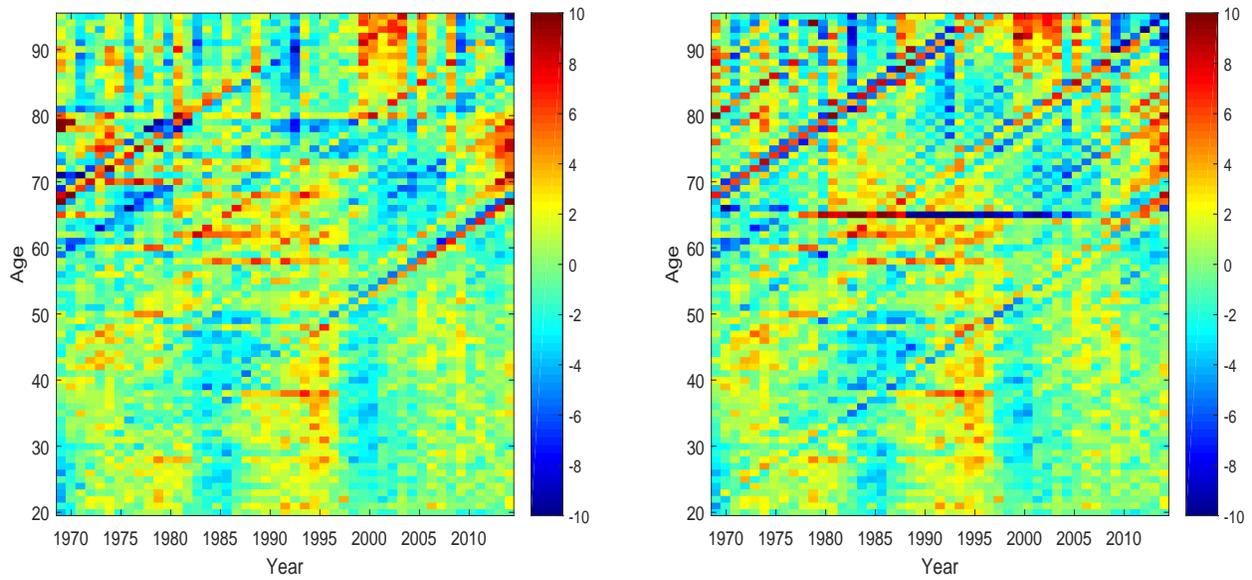


Figure 8: Heat maps of the standardized residuals calculated from the CMI-17 APCI model for U.S. **females**.
 - Left panel: Data set (i) (HMD, ages 20-95, years 1968-2014).
 - Right panel: Data set (ii) (SSA, ages 20-95, years 1968-2014).

3 Testing the Robustness of the CMI-17 Method

In this section, we perform several robustness tests on the CMI-17 APCI model. Specifically, we test the robustness of the APCI model to (i) changes in the calibration window, (ii) changes in the age range used, (iii) the parameter constraints used and (iv) the inclusion/exclusion of the oldest/newest cohorts. The results shown in this section are based on SSA data set. The conclusions drawn from HMD data set are very similar.

3.1 Changes in the Calibration Window

In this sub-section, we apply the APCI model to three different calibration windows. The calibration windows have different starting and ending years, but have the same length of 37 years. This set-up is to mimic the situation when the model is updated every 5 years. The following table summarizes the calibration windows under consideration:

	Starting Age	Ending Age	Starting Year	Ending Year	Length of Calibration Window
Baseline			1968	2004	37 years
Alternative 1	20	95	1973	2009	37 years
Alternative 2			1978	2014	37 years

Figures 9 (males) and 10 (females) show the age, period and cohort components of the mortality improvement rates in the APCI model when different calibration windows are used. The following observations are made:

- Since we are shifting the starting and the ending year simultaneously, the lengths of the period and cohort components remain the same.
- The overall shapes of the age, period and cohort components are quite insensitive to changes in the calibration window.
- Among the three components, the age component appears to be the most sensitive to different calibration windows.

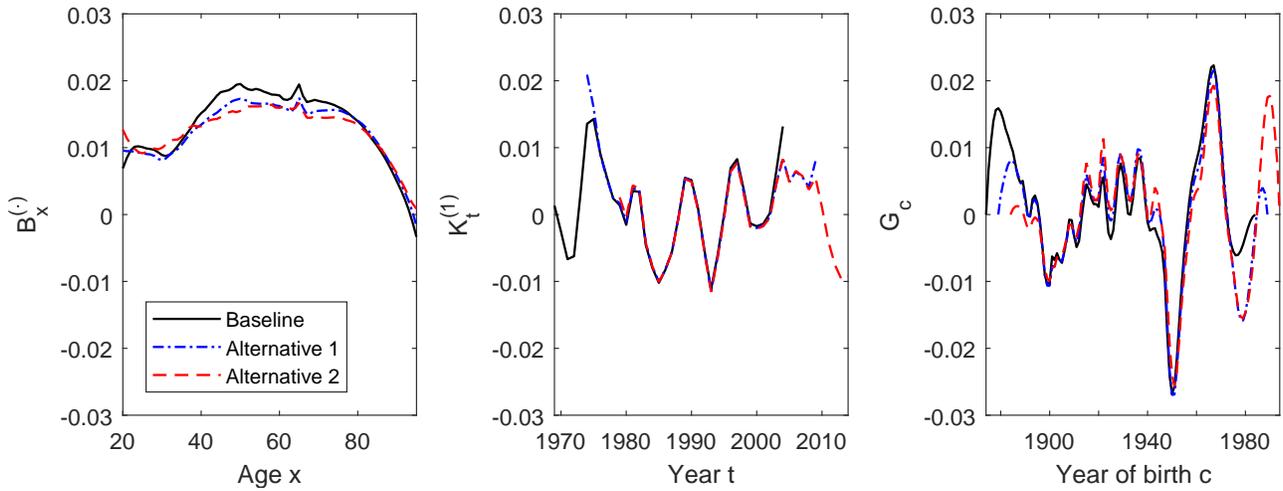


Figure 9: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model when three different calibration windows are used, U.S. **males**.

- Baseline: ages 20 to 95, years 1968 to 2004 (covering years-of-birth 1873 to 1984).
- Alternative 1: ages 20 to 95, years 1973 to 2009 (covering years-of-birth 1878 to 1989).
- Alternative 2: ages 20 to 95, years 1978 to 2014 (covering years-of-birth 1883 to 1994).

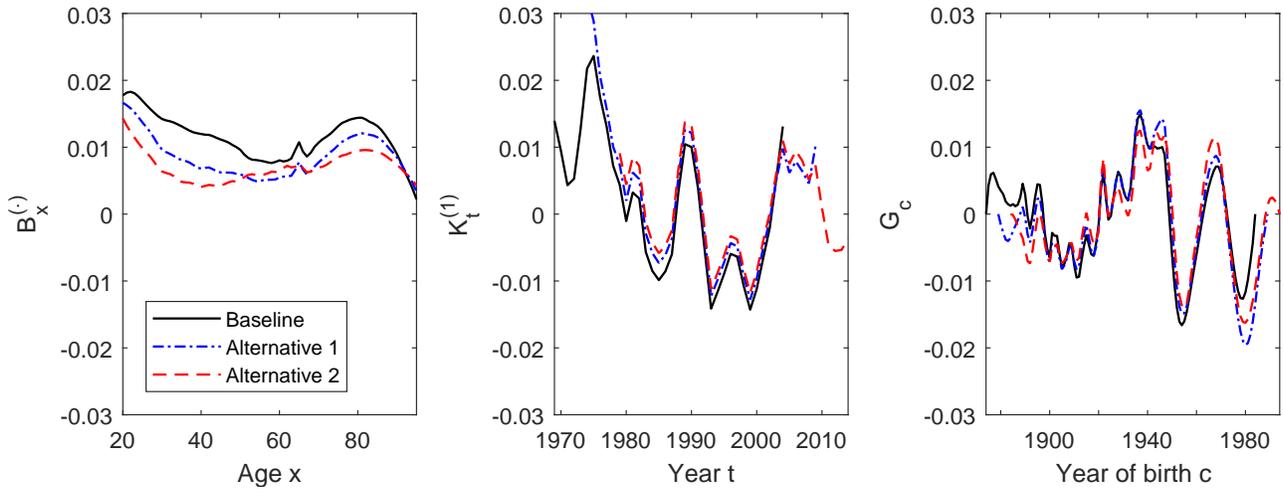


Figure 10: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model when three different calibration windows are used, U.S. **females**.

- Baseline: ages 20 to 95, years 1968 to 2004 (covering years-of-birth 1873 to 1984).
- Alternative 1: ages 20 to 95, years 1973 to 2009 (covering years-of-birth 1878 to 1989).
- Alternative 2: ages 20 to 95, years 1978 to 2014 (covering years-of-birth 1883 to 1994).

3.2 Changes in the Age Range Used

In this sub-section, we apply the APCI model to three different age ranges:

	Starting Age	Ending Age	Starting Year	Ending Year	Number of Ages
Baseline	20	95			76
Alternative 1	30	85	1968	2014	56
Alternative 2	40	75			36

Figures 11 (males) and 12 (females) show the estimated age, period and cohort components of historical mortality improvement obtained from the APCI model when different age ranges are used. The following observations can be made:

- The lengths of the age and cohort components become smaller as the age range is shortened.
- As the age range window is shortened, the values of the age, period and cohort components change, but the overall shapes of the components remain fairly the same.

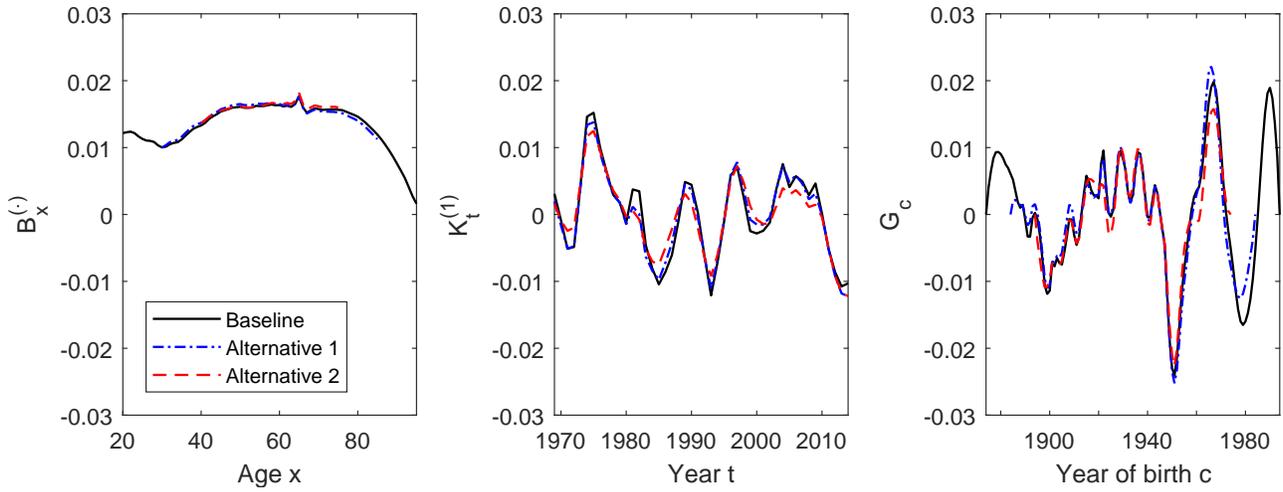


Figure 11: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model when different age ranges are used, U.S. **males**.
 - Baseline: ages 20 to 95, years 1968 to 2014 (covering years-of-birth 1873 to 1994).
 - Alternative 1: ages 30 to 85, years 1968 to 2014 (covering years-of-birth 1883 to 1984).
 - Alternative 2: ages 40 to 75, years 1968 to 2014 (covering years-of-birth 1893 to 1974).

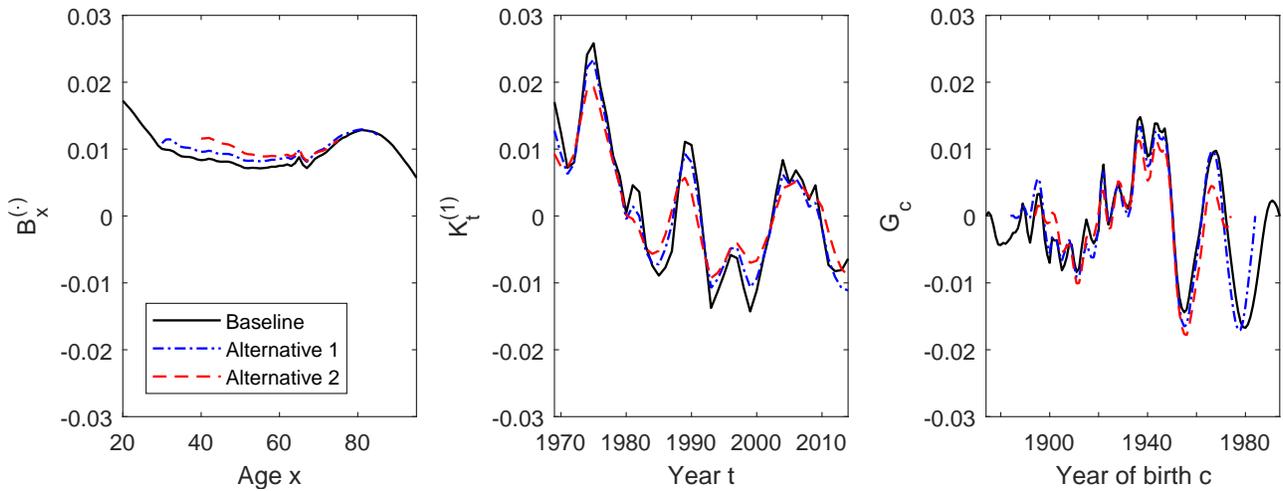


Figure 12: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model when different age ranges are used, U.S. **females**.
 - Baseline: ages 20 to 95, years 1968 to 2014 (covering years-of-birth 1873 to 1994).
 - Alternative 1: ages 30 to 85, years 1968 to 2014 (covering years-of-birth 1883 to 1984).
 - Alternative 2: ages 40 to 75, years 1968 to 2014 (covering years-of-birth 1893 to 1974).

3.3 Parameter Constraints Used

In this subsection, we study the robustness of the APCI model relative to the choice of identifiability constraints. The following sets of constraints are considered.

- **Baseline:**

$$\sum_t \kappa_t^{(1)} = 0, \quad \sum_t t\kappa_t^{(1)} = 0, \quad \sum_c \gamma_c = 0, \quad \sum_c c\gamma_c = 0, \quad \sum_c c^2\gamma_c = 0,$$

These constraints ensures that both the period and cohort parameters fluctuate around zero and exhibit no linear trend, and that the cohort parameter exhibits no quadratic trend.

- **Alternative:**

$$\sum_t \kappa_t^{(1)} = 0, \quad \sum_t t\kappa_t^{(1)} = 0, \quad \sum_c n_c\gamma_c = 0, \quad \sum_c n_c c\gamma_c = 0, \quad \sum_c n_c c^2\gamma_c = 0,$$

Compared to the baseline constraints, the alternative constraint include the number of data points related to year-of-birth n_c . By including n_c , the cohorts about which we have more information are weighted more heavily in the parameter constraints.

Figures 13 (males) and 14 (females) show the estimated age, period, and cohort components of the mortality improvement rates using the baseline and alternative constraints. The two sets of constraints yield very similar estimates, indicating that the weights (n_c) only have modest impact on the resulting estimates.

It should be noted that there are many other combinations of constraints that can be used to stipulate parameter uniqueness. When very different constraints, the resulting A/P/C components may be very different.

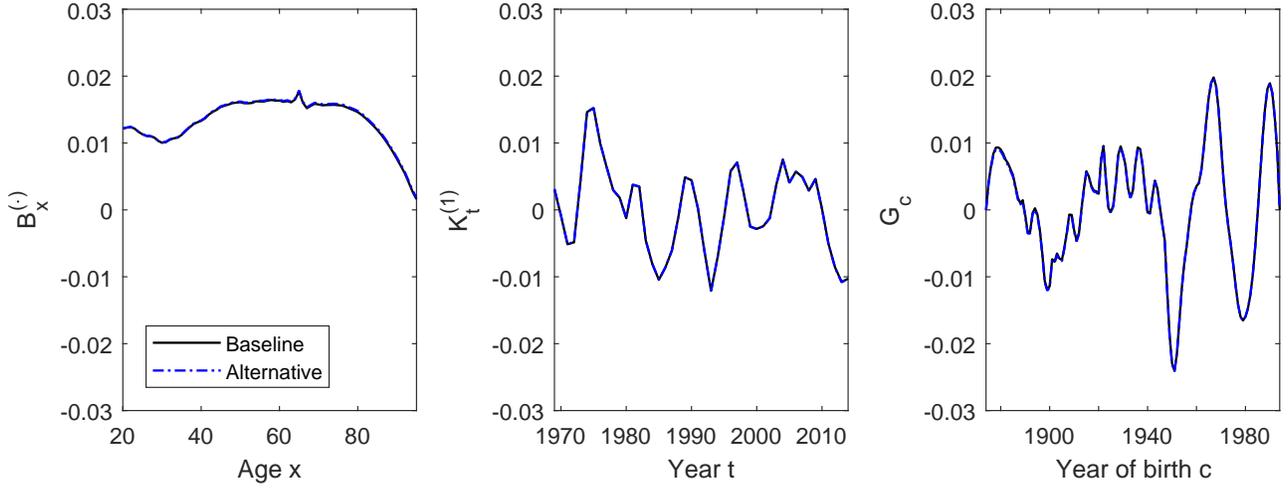


Figure 13: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model when different parameter constraints are used, U.S. **males**.

- Baseline: $\sum_t \kappa_t^{(1)} = \sum_t t\kappa_t^{(1)} = \sum_c \gamma_c = \sum_c c\gamma_c = \sum_c c^2\gamma_c = 0$
- Alternative: $\sum_t \kappa_t^{(1)} = \sum_t t\kappa_t^{(1)} = \sum_c n_c\gamma_c = \sum_c n_c c\gamma_c = \sum_c n_c c^2\gamma_c = 0$

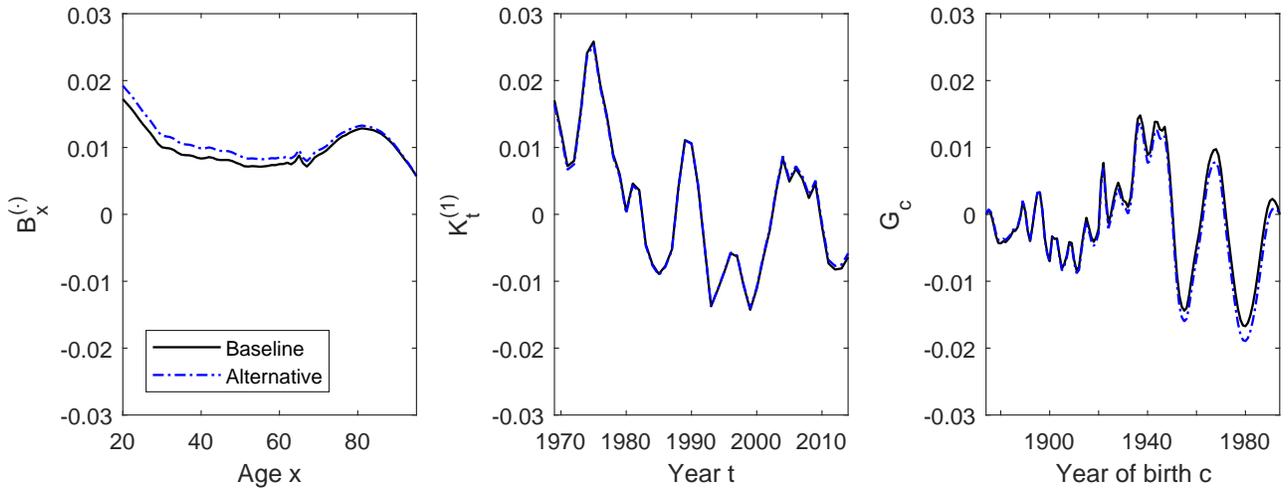


Figure 14: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model when different parameter constraints are used, U.S. **females**.

- Baseline: $\sum_t \kappa_t^{(1)} = \sum_t t\kappa_t^{(1)} = \sum_c \gamma_c = \sum_c c\gamma_c = \sum_c c^2\gamma_c = 0$
- Alternative: $\sum_t \kappa_t^{(1)} = \sum_t t\kappa_t^{(1)} = \sum_c n_c\gamma_c = \sum_c n_c c\gamma_c = \sum_c n_c c^2\gamma_c = 0$

3.4 Inclusion/Exclusion of the Oldest/Newest Cohorts

In this sub-section, we examine how the age, period and cohort components implied by the CMI-17 APCI model may change when the oldest and newest cohorts are excluded. The following three situations are considered:

- **Baseline:** All available data are used.
- **Alternative 1:** The oldest and youngest five cohorts in the data sample are excluded.
- **Alternative 2:** The oldest and youngest ten cohorts in the data sample are excluded.

The resulting age/period/cohort components are shown in Figures 15 (males) and 16 (females). It is found that the exclusion of the oldest and the newest cohorts has only a modest impact on the estimation results.

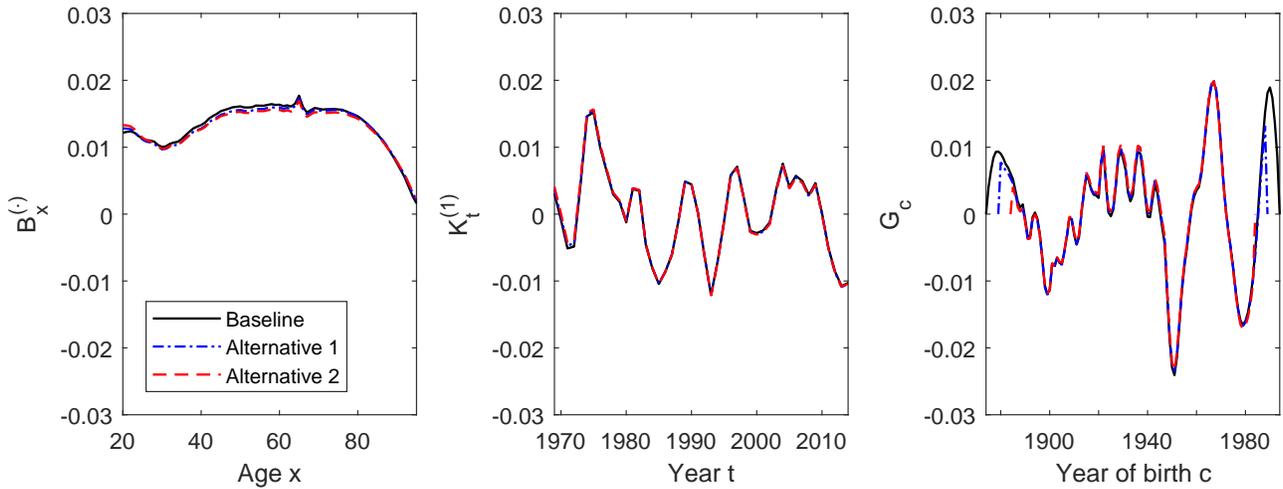


Figure 15: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model, with and without exclusion of the youngest/oldest cohorts, U.S. **males**.

- Baseline: All available data are used.
- Alternative 1: The oldest and youngest five cohorts in the data sample are excluded.
- Alternative 2: The oldest and youngest ten cohorts in the data sample are excluded.

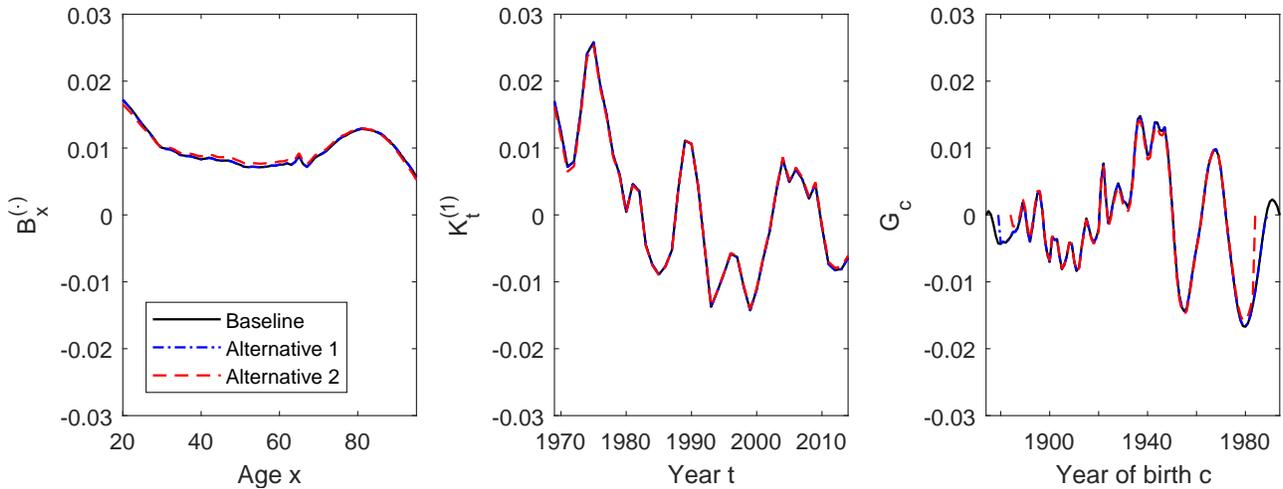


Figure 16: The estimated Age/Period/Cohort components of historical mortality improvement obtained from the CMI-17 APCI model, with and without exclusion of the youngest/oldest cohorts, U.S. **females**.

- Baseline: All available data are used.
- Alternative 1: The oldest and youngest five cohorts in the data sample are excluded.
- Alternative 2: The oldest and youngest ten cohorts in the data sample are excluded.

4 Improving the CMI-17 Method

It is found in the previous sections that the CMI-17 APC model is somewhat robust to small changes in different inputs. However, large vertical clusters are observed in the standardized residuals produced by the CMI-17 APC model that is fitted to U.S. male data, suggesting that the CMI-17 APC model does not adequately capture the age-period interaction for this particular population.

In this section, we attempt to improve the CMI-17 method by considering alternative APC models. These models are collectively known as Route B models, and are different from Route A models in that they are estimated to *mortality rates* instead of mortality improvement rates.

4.1 The Candidate APC Model Structures

4.1.1 Definitions and Overview

The following are some general definitions that are used throughout the rest of this section.

- $m_{x,t}$ is the central death rate at age x and in year t ;
- $q_{x,t}$ is the conditional probability of death at age x and in year t ;
- $[x_0, x_1]$ is the sample age range;
- $[t_0, t_1]$ is the sample period;
- $n_a = x_1 - x_0 + 1$ is the number of ages covered by the sample age range;
- $n_y = t_1 - t_0 + 1$ is the length of the sample period;
- $c = t - x$ is the year of birth; note that within the data sample c ranges from $t_0 - x_1$ to $t_1 - x_0$;
- n_c is the number of data points associated with year-of-birth c .
- $\bar{x} = (x_0 + x_1)/2$ is the mid-point of the sample age range;
- $\bar{t} = (t_0 + t_1)/2$ is the mid-point of the sample period;
- $(\bar{x} - x)_+ = \max(\bar{x} - x, 0)$ is the maximum of $\bar{x} - x$ and 0;

In Route B, we estimate APC models to $m_{x,t}$ or $q_{x,t}$. These models are composed of the following collection of elements:

- $\beta_x^{(i)}$, $i = 1, 2, 3$, are age-specific parameters:
 - $\beta_x^{(1)}$ is a stand-alone age-specific parameter,
 - $\beta_x^{(2)}$ is an age-specific parameter that interacts with a time-varying parameter,

- $\beta_x^{(3)}$ is an age-specific parameter that interacts with a cohort-related parameter,
- $\kappa_t^{(i)}$, $i = 1, 2, 3$, are time-varying parameters:
 - $\kappa_t^{(1)}$ is a stand-alone time-varying parameter;
 - $\kappa_t^{(2)}$ is a time-varying parameter that interacts with an age-specific parameter or a linear function of age;
 - $\kappa_t^{(3)}$ is a time-varying parameter that interacts with a non-linear function of age.
- γ_c is a cohort-related parameter.

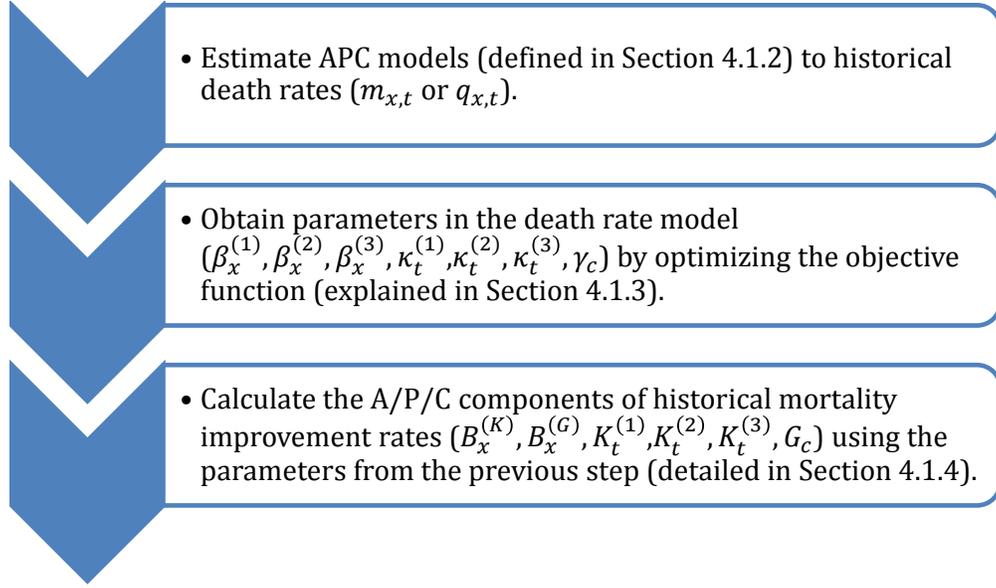
We may interpret parameters $\beta_x^{(i)}$, $\kappa_t^{(i)}$, and γ_c as the age, period and cohort components of historical *mortality rates*. These parameters are to be transformed to obtain A/P/C decompositions of historical *mortality improvement rates*.

For ease of exposition, the following symbols are defined to represent the A/P/C components of historical *mortality improvement rates*:

- $B_x^{(i)}$, $i = \cdot, K, G$ are the age components of the mortality improvement rates:
 - $B_x^{(\cdot)}$ is a stand-alone age component,
 - $B_x^{(K)}$ is an age component that interacts with a period component,
 - $B_x^{(G)}$ is an age component that interacts with a cohort component;
- $K_t^{(i)}$, $i = 1, 2, 3$, are period components of the mortality improvement rates:
 - $K_t^{(1)}$ is a stand-alone period component;
 - $K_t^{(2)}$ is a period component that interacts with an age component or a linear function of age;
 - $K_t^{(3)}$ is a period component that interacts with a non-linear function of age.
- G_c is the cohort component of the mortality improvement rates.

Section 4.1.4 details how $B_x^{(K)}$, $B_x^{(G)}$, $K_t^{(1)}$, $K_t^{(2)}$, $K_t^{(3)}$, and G_c are obtained from the estimated models.

The following flow chart summarizes the process undertaken in terms of the symbols defined above.



4.1.2 Mortality Rate Models

We consider eight candidate APC models in Route B, including M2, M3, M6, M7, M8, the full Plat model, the simplified Plat model and the APCI model (on which the CMI-17 method is based). These models are used to fit historical mortality rates ($m_{x,t}$'s or $q_{x,t}$'s).

The mathematical definitions of the eight candidate models are provided below. For all of the candidate structures, identifiability constraints are needed to stipulate parameter uniqueness. The identifiability constraints used for each candidate model are also given below.

- M2 – The Renshaw-Haberman Model (Renshaw and Haberman, 2006)

$$\ln m_{x,t} = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} + \beta_x^{(3)} \gamma_c$$

Identifiability constraints:

$$\sum_{x=x_0}^{x_1} \beta_x^{(1)} = 0, \sum_{x=x_0}^{x_1} \beta_x^{(2)} = 1, \sum_{x=x_0}^{x_1} \beta_x^{(3)} = 1, \text{ and } \sum_{c=t_0-x_1}^{t_1-x_0} n_c \gamma_c = 0.$$

- M3 – The Age-Period-Cohort Model (Osmond, 1985; used in the current CMI decomposition method)

$$\ln m_{x,t} = \beta_x^{(1)} + \kappa_t^{(1)} + \gamma_c$$

$$\text{Identifiability constraints: } \sum_{x=x_0}^{x_1} \beta_x^{(1)} = 0, \sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0, \text{ and } \sum_{c=t_0-x_1}^{t_1-x_0} c \gamma_c = 0.$$

- M6 – The CBD Model with a Cohort Effect (Cairns et al., 2009)

$$\ln \frac{q_{x,t}}{1 - q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)} (x - \bar{x}) + \gamma_c$$

$$\text{Identifiability constraints: } \sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0 \text{ and } \sum_{c=t_0-x_1}^{t_1-x_0} c \gamma_c = 0.$$

- M7 – The CBD Model with Quadratic Age and Cohort Effects (Cairns et al., 2009)

$$\ln \frac{q_{x,t}}{1 - q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_c$$

Identifiability constraints: $\sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0$, $\sum_{c=t_0-x_1}^{t_1-x_0} c\gamma_c = 0$, and $\sum_{c=t_0-x_1}^{t_1-x_0} c^2\gamma_c = 0$

- M8 – The CBD Model with an Age-Dependent Cohort Effect (Cairns et al., 2009)

$$\ln \frac{q_{x,t}}{1 - q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_c(x_c - x)$$

Identifiability constraints: $\sum_{c=t_0-x_1}^{t_1-x_0} n_c\gamma_c = 0$. Note: x_c is an estimated constant.

- The Full Plat model (Plat, 2009):

$$\ln m_{x,t} = \beta_x^{(1)} + \kappa_t^{(1)} + \kappa_t^{(2)}(\bar{x} - x) + \kappa_t^{(3)}(\bar{x} - x)_+ + \gamma_c$$

Identifiability constraints:

$$\sum_{x=x_0}^{x_1} \beta_x^{(1)} = 0, \sum_{t=t_0}^{t_1} \kappa_t^{(2)} = 0, \sum_{t=t_0}^{t_1} \kappa_t^{(3)} = 0, \sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0, \sum_{c=t_0-x_1}^{t_1-x_0} c\gamma_c = 0, \text{ and } \sum_{c=t_0-x_1}^{t_1-x_0} c^2\gamma_c = 0.$$

- The Simplified Plat model (Plat, 2009):

$$\ln m_{x,t} = \beta_x^{(1)} + \kappa_t^{(1)} + \kappa_t^{(2)}(\bar{x} - x) + \gamma_c$$

Identifiability constraints:

$$\sum_{x=x_0}^{x_1} \beta_x^{(1)} = 0, \sum_{t=t_0}^{t_1} \kappa_t^{(2)} = 0, \sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0, \sum_{c=t_0-x_1}^{t_1-x_0} c\gamma_c = 0, \text{ and } \sum_{c=t_0-x_1}^{t_1-x_0} c^2\gamma_c = 0.$$

- The APCI model (CMI, 2017a,b,c):

$$\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)}(t - \bar{t}) + \kappa_t^{(1)} + \gamma_c,$$

Identifiability constraints:

$$\sum_{c=t_0-x_1}^{t_1-x_0} \gamma_c = 0, \sum_{c=t_0-x_1}^{t_1-x_0} c\gamma_c = 0, \sum_{c=t_0-x_1}^{t_1-x_0} c^2\gamma_c = 0, \sum_{t=t_0}^{t_1} \kappa_t = 0 \text{ and } \sum_{t=t_0}^{t_1} t\kappa_t = 0.$$

4.1.3 Parameter Estimation

As with what was done in the CMI Working Paper No. 98, we estimate and smooth parameters simultaneously by minimizing a single objective function, which is formulated as follows:

$$\text{Objective} = \text{Deviance} + \sum_i \text{Penalty}(\beta_x^{(i)}) + \sum_j \text{Penalty}(\kappa_t^{(j)}) + \text{Penalty}(\gamma_c). \quad (1)$$

The deviance measures the goodness-of-fit. The smaller the deviance is, the better the goodness-of-fit is. Let $D_{x,t}$ be the number of deaths at age x and in year t , and $E_{x,t}$ be the corresponding

number of exposures-to-risk. Assuming $D_{x,t}$ is a realization of a Poisson distribution, i.e., $D_{x,t} \sim \text{Poisson}(E_{x,t}m_{x,t})$, the deviance can be calculated with the following formula:

$$\text{Deviance} = 2 \sum_{x=x_0}^{x_1} \sum_{t=t_0}^{t_1} \left[D_{x,t} \ln \frac{D_{x,t}}{E_{x,t}m_{x,t}} - (D_{x,t} - E_{x,t}m_{x,t}) \right].$$

The other three terms in the objective function are roughness penalty terms. The more jagged a parameter series is, the higher the penalty term for the parameter series is. The penalty terms can be calculated with the following formulas:

$$\begin{aligned} \text{Penalty}(\beta_x^{(i)}) &= \lambda_{\beta_x^{(i)}} \sum_{x=x_0}^{x_1} (\beta_x^{(i)} - 3\beta_{x-1}^{(i)} + 3\beta_{x-2}^{(i)} - \beta_{x-3}^{(i)})^2, \\ \text{Penalty}(\kappa_t^{(j)}) &= \lambda_{\kappa_t^{(j)}} \sum_{t=t_0}^{t_1} (\kappa_t^{(j)} - 2\kappa_{t-1}^{(j)} + \kappa_{t-2}^{(j)})^2, \\ \text{Penalty}(\gamma_c) &= \lambda_{\gamma_c} \sum_{c=t_0-x_1}^{t_1-x_0} (\gamma_c - 3\gamma_{c-1} + 3\gamma_{c-2} - \gamma_{c-3})^2, \end{aligned}$$

where $\lambda_{\beta_x^{(i)}}$, $\lambda_{\kappa_t^{(j)}}$ and λ_{γ_c} are penalty parameters which determine the degrees of smoothness of the parameters. The choice of penalty parameters is explained in Section 4.2.

4.1.4 A/P/C Components of Historical Mortality improvement

When using Route A (the previous volume), the following definition of mortality improvement was used:

$$Z_{x,t} = 1 - \frac{q_{x,t}}{q_{x,t-1}}.$$

When using Route B (this volume), alternative definitions of mortality improvement are needed in order to obtain neat and interpretable parameterizations. Two versions of mortality improvement are defined here.

The first definition is based on the change in log central death rates, and is given by

$$MI_{x,t} = \ln m_{x,t-1} - \ln m_{x,t}. \quad (2)$$

In the CMI Working Papers, this definition of mortality improvement is known as the ‘M-style’ mortality improvement. In Appendix 1 we show that on the basis of a first-order Taylor’s expansion, $MI_{x,t}$ defined in equation (2) is approximately equal to $Z_{x,t}$. We apply this definition of mortality improvement to M2, M3, the full Plat model, the simplified Plat models and the APCI model, as they are all created to model $\ln m_{x,t}$. The annual mortality improvement for these models can be expressed as follows:

- M2

$$MI_{x,t} = B_x^{(K)} K_t^{(2)} + B_x^{(G)} G_c,$$

where $B_x^{(K)} = \beta_x^{(2)}$, $K_t^{(2)} = \kappa_{t-1}^{(2)} - \kappa_t^{(2)}$, $B^{(G)} = \beta_x^{(3)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: M2 implies that mortality improvement is driven by a period component, a cohort component, and two age components which interact with the period and cohort components, respectively.

- M3

$$MI_{x,t} = K_t^{(1)} + G_c$$

where $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: M3 implies that mortality improvement is driven by the sum of a period component and a cohort component. There is no age component.

- The Full Plat model

$$MI_{x,t} = K_t^{(1)} + K_t^{(2)}(\bar{x} - x) + K_t^{(3)}(\bar{x} - x)_+ + G_c$$

where $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$, $K_t^{(2)} = \kappa_{t-1}^{(2)} - \kappa_t^{(2)}$, $K_t^{(3)} = \kappa_{t-1}^{(3)} - \kappa_t^{(3)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: The full Plat model implies that mortality improvement is driven by three period components and a cohort component. Among the period components, one is stand-alone, one interacts with a linear function of age, and one interacts with a non-linear function of age.

- The Simplified Plat model

$$MI_{x,t} = K_t^{(1)} + K_t^{(2)}(\bar{x} - x) + G_c$$

where $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$, $K_t^{(2)} = \kappa_{t-1}^{(2)} - \kappa_t^{(2)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: The simplified Plat model implies that mortality improvement is driven by two period components and a cohort component. Among the period components, one is stand-alone and one interacts with a linear function of age.

- The APCI model

$$MI_{x,t} = B_x^{(\cdot)} + K_t^{(1)} + G_c$$

where $B_x^{(\cdot)} = -\beta_x^{(2)}$, $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: The APCI model implies that mortality improvement is driven by a (stand-alone) age component, a period component and a cohort component. There is neither age-period nor age-cohort interaction.

The second definition of mortality improvement is based on the change in logit-transformed³ conditional death probabilities:

$$MI_{x,t} = \ln \frac{q_{x,t-1}}{1 - q_{x,t-1}} - \ln \frac{q_{x,t}}{1 - q_{x,t}}. \quad (3)$$

In Appendix 2 we prove that on the basis of a first-order Taylor's expansion, this specification of $MI_{x,t}$ is approximately equal to that specified in equation (2). We use this definition of mortality improvement for models M6, M7 and M8, because all of these models are created to model $\ln(q_{x,t}/(1 - q_{x,t}))$. The annual mortality improvement for these models can be expressed as follows:

³The logit transform of a quantity y is defined as $\ln(y/(1 - y))$.

- M6

$$MI_{x,t} = K_t^{(1)} + K_t^{(2)}(x - \bar{x}) + G_c$$

where $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$, $K_t^{(2)} = \kappa_{t-1}^{(2)} - \kappa_t^{(2)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: M6 implies that mortality improvement is driven by two period components and one cohort component. One period component is stand-alone, while the other period component interacts with a linear function of age.

- M7

$$MI_{x,t} = K_t^{(1)} + K_t^{(2)}(x - \bar{x}) + K_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + G_c$$

where $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$, $K_t^{(2)} = \kappa_{t-1}^{(2)} - \kappa_t^{(2)}$, $K_t^{(3)} = \kappa_{t-1}^{(3)} - \kappa_t^{(3)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: M7 implies that mortality improvement is driven by three period components and one cohort component. Among the three period components, one is stand-alone, one interacts with a linear function of age, and one interacts with a quadratic function of age.

- M8

$$MI_{x,t} = K_t^{(1)} + K_t^{(3)}(x - \bar{x}) + G_c(x_c - x)$$

where $K_t^{(1)} = \kappa_{t-1}^{(1)} - \kappa_t^{(1)}$, $K_t^{(2)} = \kappa_{t-1}^{(2)} - \kappa_t^{(2)}$, $K_t^{(3)} = \kappa_{t-1}^{(3)} - \kappa_t^{(3)}$ and $G_c = \gamma_{c-1} - \gamma_c$.

Comment: M8 implies that mortality improvement is driven by three period components and one cohort component. Among the three period components, one is stand-alone, one interacts with a linear function of age, and one interacts with a quadratic function of age. The cohort component interacts with a linear function of age.

4.2 Selecting Penalty Parameters

The penalty parameters control the smoothness of the parameter estimates and hence the resulting A/P/C components of the mortality improvement rates. Ideally, the selected penalty parameters should remove the unwanted jaggedness in the parameter series and at the same time retain the troughs and peaks in the unsmoothed parameter series. The penalty parameters selected for each Route B candidate model are shown in Table 1.

Penalty parameter	$\lambda_{\beta_x^{(1)}}$	$\lambda_{\beta_x^{(2)}}$	$\lambda_{\beta_x^{(3)}}$	$\lambda_{\kappa_t^{(1)}}$	$\lambda_{\kappa_t^{(2)}}$	$\lambda_{\kappa_t^{(3)}}$	λ_{γ_c}
M2	10^7	10^9	10^8	-	10^3	-	10^3
M3	10^6	-	-	10^6	-	-	10^7
M6	-	-	-	10^6	10^9	-	10^7
M7	-	-	-	10^6	10^9	10^{12}	10^7
M8	-	-	-	10^6	10^9	-	10^9
Full Plat	10^6	-	-	10^6	10^9	10^8	10^7
Simplified Plat	10^6	-	-	10^6	10^9	-	10^7
APCI	10^6	10^7	-	-	-	10^6	10^6

Table 1: Selected penalty parameters for each Route B candidate model.

The following procedure is used to select the penalty parameter for $\beta_x^{(1)}$:

1. Set all penalty parameters except $\lambda_{\beta_x^{(1)}}$ to zero.
2. Plot the estimated values of $\beta_x^{(1)}$ for $\lambda_{\beta_x^{(1)}} = 10^0, 10^1, 10^2, \dots, 10^{10}$.
3. The optimal value of $\lambda_{\beta_x^{(1)}}$ should be the one that removes the jaggedness and largely keeps the shape of the unsmoothed series.

Similar procedures are used to select the penalty parameters for other parameter series.

As an example, let us consider the choice of the smoothing parameter for $\kappa_t^{(1)}$ in M3. The following describes how this smoothing parameter is chosen.

- When $\lambda_{\kappa_t^{(1)}}$ is set to 10^5 (the left panel of Figure 17), the pattern of $\kappa_t^{(1)}$ still exhibit some jaggedness, indicating that the parameter series is under-smoothed.
- When $\lambda_{\kappa_t^{(1)}}$ is set to 10^7 (the right panel of Figure 17), some features in the unsmoothed series (represented by circles in the diagram) between 1980 and 2000 are completely smoothed out. The problem of over-smoothing can be more easily discerned in the pattern of $K_t^{(1)}$ (the period component in the mortality *improvement* rates), shown in Figure 18. When the penalty parameter is set to 10^7 , the peaks and troughs of $K_t^{(1)}$ are somewhat dislocated.
- Therefore, we choose 10^6 as the optimal value for $\lambda_{\kappa_t^{(1)}}$ (see the middle panel of Figure 17).

We have attempted to optimize the smoothing parameters using a cross validation (see Volume 1 of the project report) instead. Despite being more statistically rigorous, this method does not yield reasonable empirical results for Route B.⁴

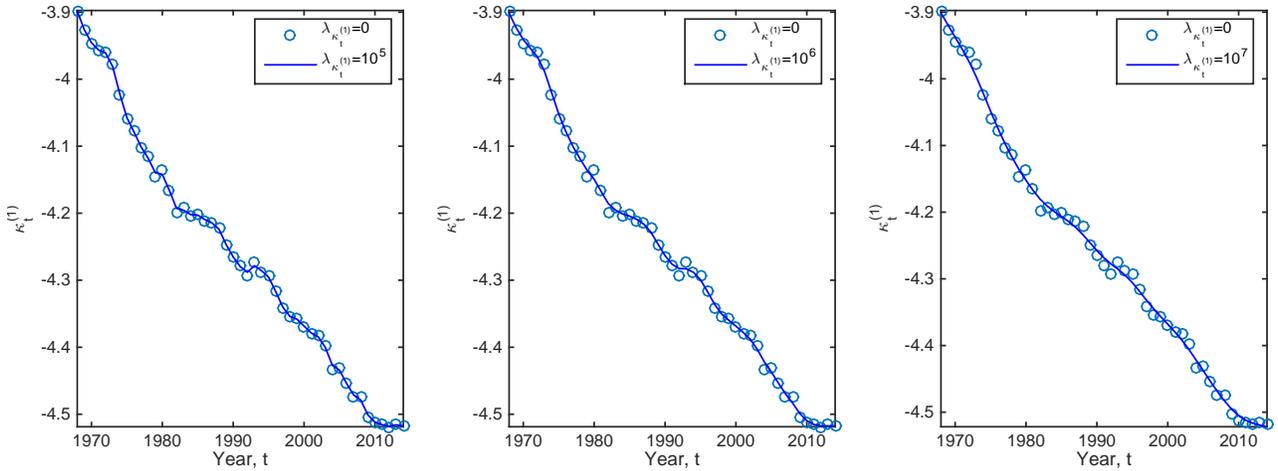


Figure 17: The estimates of $\kappa_t^{(1)}$ in the Route B M3 model for SSA males when different penalty parameters are used.

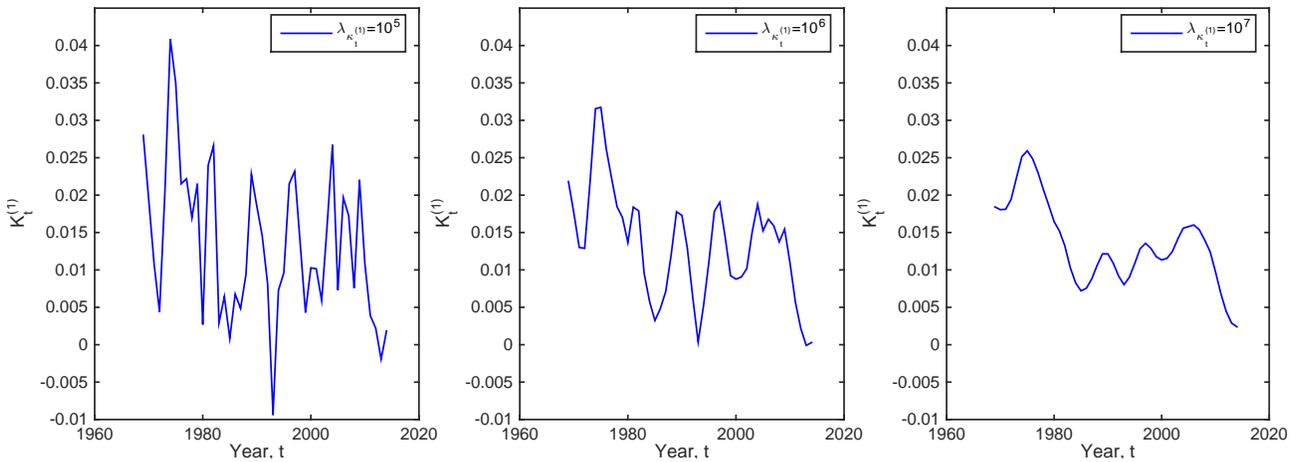


Figure 18: The values of $K_t^{(1)}$ (the period component of the mortality improvement rates) implied by the Route B M3 model fitted to the SSA male data set, when different penalty parameters are used.

⁴One possible explanation for this problem is that optimally smooth A/P/C parameters in the *mortality rate model* do not necessarily guarantee that the implied A/P/C components of historical *mortality improvements* are adequately smooth.

4.3 Robustness Tests

4.3.1 Defining the Robustness Measure

We define the following quantitative measure of robustness:

$$\text{robustness} = \frac{\max_i(\text{maximum absolute change in the } i\text{-th model term of } MI_{x,t})}{\max_{x,t}(MI_{x,t}) - \min_{x,t}(MI_{x,t})}, \quad (4)$$

where $\max_{x,t}(MI_{x,t})$ and $\min_{x,t}(MI_{x,t})$ represent the maximum and minimum values of the historical mortality improvement rates in the dataset⁵, respectively, and ‘the i -th model term’ refers to the i -th term on the right-hand-side of the equation specifying $MI_{x,t}$. The denominator $\max_{x,t}(MI_{x,t}) - \min_{x,t}(MI_{x,t})$ ‘standardizes’ the robustness measure by considering the variability of the data being fed into the model.

To illustrate, let us use the metric to quantify the robustness of M3 to changes in the length of the calibration window (on the basis of the SSA male data set). We consider three calibration windows, which have the same length but begin in different years (1968, 1973, and 1978, respectively). The following table summarizes the calibration windows under consideration:

	Starting Age	Ending Age	Starting Year	Ending Year	Length of Calibration Window
Baseline			1968	2004	37 years
Alternative 1	20	95	1973	2009	37 years
Alternative 2			1978	2014	37 years

The maximum and the minimum values of SSA male historical improvement rates are 0.1963 and -0.1466 , respectively. Therefore, the denominator of equation (4) is $0.1963 - (-0.1466) = 0.3429$. The following steps are further taken to obtain the numerator of equation (4):

- (1) Fit M3 using the baseline setting and obtain estimates of the two terms in the equation for $MI_{x,t}$: $K_t^{(1)}$ and G_c .
- (2) Re-fit M3 using Alternative 1 setting and obtain new estimates of the two terms in the equation for $MI_{x,t}$: $\tilde{K}_t^{(1)}$ and \tilde{G}_c .
- (3) Compare the results from the baseline and Alternative 1 settings, and calculate the maximum absolute change in each of the two model terms:

$$\begin{aligned} \max_{t=1974,1975,\dots,2004} |K_t^{(1)} - \tilde{K}_t^{(1)}| &= 0.0107; \\ \max_{c=1879,1880,\dots,1984} |G_c - \tilde{G}_c| &= 0.0079. \end{aligned}$$

⁵For M2, M3, the full Plat model, the simplified Plat model and the APCI model, we calculate the historical mortality improvement rates as $\ln m_{x,t-1} - \ln m_{x,t}$; for M6, M7 and M8 models, the historical mortality improvement rates are calculated as $\ln \frac{q_{x,t-1}}{1-q_{x,t-1}} - \ln \frac{q_{x,t}}{1-q_{x,t}}$. This arrangement maximizes consistency with the way in which $MI_{x,t}$ is defined for the models.

- (4) Re-fit M3 using Alternative 2 setting and obtain new estimates of the two model terms in the equation for $MI_{x,t}$: $\hat{K}_t^{(1)}$ and \hat{G}_c .
- (5) Compare the results from the baseline and Alternative 2 settings, and calculate the maximum absolute change in each of the two model terms:

$$\begin{aligned} \max_{t=1979,1980,\dots,2004} |K_t^{(1)} - \hat{K}_t^{(1)}| &= 0.0151; \\ \max_{c=1884,1885,\dots,1984} |G_c - \hat{G}_c| &= 0.0043. \end{aligned}$$

- (6) Compare the results from Alternative 1 and Alternative 2 settings, and calculate the maximum absolute change in each of the two model terms:

$$\begin{aligned} \max_{t=1979,1980,\dots,2009} |\tilde{K}_t^{(1)} - \hat{K}_t^{(1)}| &= 0.0079 \\ \max_{c=1884,1885,\dots,1989} |\tilde{G}_c - \hat{G}_c| &= 0.0026. \end{aligned}$$

- (7) Calculate the overall maximum absolute change in the each of the two model term of $MI_{x,t}$:

$$\begin{aligned} \max(0.0107, 0.0151, 0.0079) &= 0.0151; \\ \max(0.0079, 0.0043, 0.0026) &= 0.0079. \end{aligned}$$

- (8) Set the maximum of the two values obtained in Step (7) as the numerator in equation (4), and calculate the robustness:

$$\frac{\max(0.0151, 0.0079)}{0.3429} = 4.4\%$$

In the following robustness tests, we rate the robustness of the candidate models using the following criteria:

- High robustness: $0 \leq \text{robustness measure} \leq 10\%$
- Medium robustness: $10\% < \text{robustness measure} \leq 20\%$
- Low robustness: $\text{robustness measure} > 20\%$

4.3.2 Changes in the Tolerance Value Used in Optimizing Model Parameters

To obtain parameter estimates, the objective function specified in equation (1) is maximized using an iterative Newton’s method. The iterations stop when the change in the value of the objective function is smaller than a pre-specified tolerance value. In this section, we test the robustness of the candidate models with respect to changes in the tolerance value. For all candidate models except M2 and M8, the following three tolerance values are considered:

- Baseline: tolerance value = 10^{-8} .
- Alternative 1: tolerance value = 10^{-6} .
- Alternative 2: tolerance value = 10^{-10} .

M2 and M8 require a special treatment because they do not converge when a tolerance value of 10^{-10} is used. We consider the following three tolerance values for these two models:

- Baseline: tolerance value = 10^{-6} .
- Alternative 1: tolerance value = 10^{-8} .
- Alternative 2: tolerance value = 10^{-4} .

The result of this robustness test is summarized in the following table:

Model	SSA male		SSA female	
	Robustness measure	Category	Robustness measure	Category
M2	29.3%	Low	25.1%	Low
M3	0.0%	High	0.0%	High
M6	0.0%	High	0.0%	High
M7	0.0%	High	0.0%	High
M8	0.0%	High	0.0%	High
Full Plat	0.0%	High	0.0%	High
Simplified Plat	0.0%	High	0.0%	High
APCI	0.0%	High	0.0%	High

Figures 19 and 20 illustrate high and low levels of robustness to changes in the tolerance value, respectively. In each diagram, the black, blue and red lines represent the resulting A/P/C components of mortality improvement rates that are obtained using different tolerance values. For the M3 model fitted to SSA male data (high robustness), the three lines overlap one another. For the M2 model fitted to SSA male data (low robustness), the three lines appear to be quite different.

The low robustness for M2 may be attributed to the convergence problem. For instance, when applying M2 to the SSA female data set, it takes 50,222 iterations to reach the tolerance value of 10^{-8} .⁶

⁶Typically, it takes less than 100 iterations to reach convergence.

In addition, we have to bound the absolute value of γ_c in this model, or otherwise the optimization may never converge.

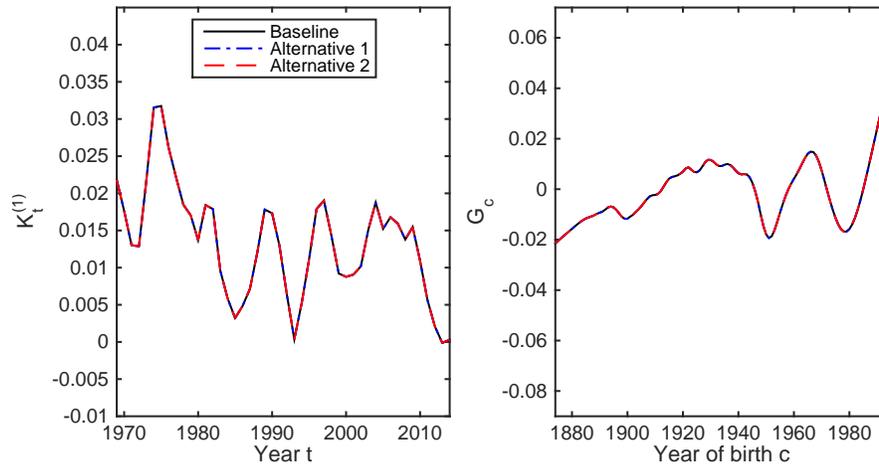


Figure 19: An example of **high robustness** to the tolerance value: The estimated A/P/C components of historical mortality improvement obtained from the Route B M3 model fitted to SSA female data.

- Baseline: tolerance value = 10^{-8} .
- Alternative 1: tolerance value = 10^{-6} .
- Alternative 2: tolerance value = 10^{-10} .

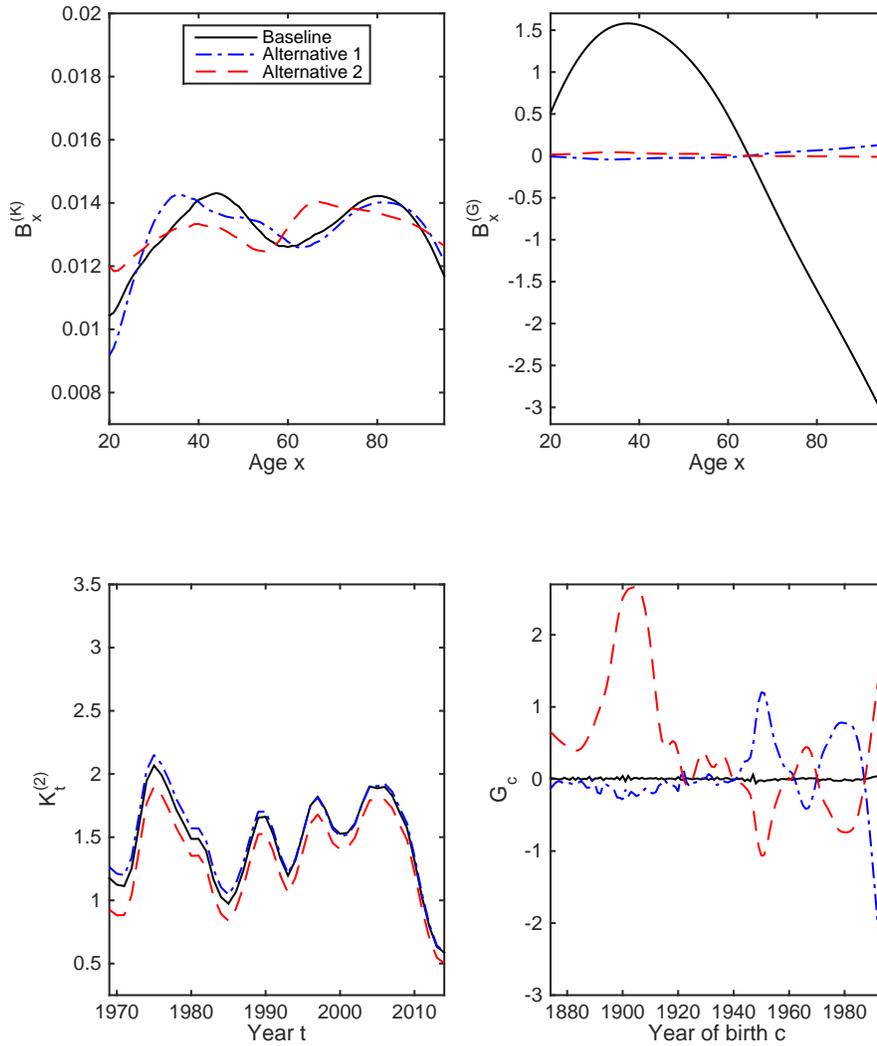


Figure 20: An example of **low robustness** to the tolerance value: The estimated A/P/C components of historical mortality improvement obtained from the Route B M2 model fitted to SSA male data.

- Baseline: tolerance value = 10^{-6} .
- Alternative 1: tolerance value = 10^{-8} .
- Alternative 2: tolerance value = 10^{-4} .

4.3.3 Changes in the Calibration Window

Here we test the robustness to changes in the calibration window. We consider three calibration windows, which have the same length but different starting years (1968, 1973, and 1978, respectively). This set-up is to mimic the situation when the models are updated every five years. The following table summarizes the calibration windows under consideration:

	Starting Age	Ending Age	Starting Year	Ending Year	Length of Calibration Window
Baseline			1968	2004	37 years
Alternative 1	20	95	1973	2009	37 years
Alternative 2			1978	2014	37 years

The result of this robustness test is summarized in the table below:

Model	SSA male		SSA female	
	Robustness measure	Category	Robustness measure	Category
M2	7.1%	High	8.8%	High
M3	4.4%	High	3.4%	High
M6	6.2%	High	11.3%	Medium
M7	9.1%	High	6.5%	High
M8	8.2%	High	8.5%	High
Full Plat	4.3%	High	12.9%	Medium
Simplified Plat	4.1%	High	5.2%	High
APCI	5.1%	High	3.8%	High

Figures 21 and 22 illustrate high and medium levels of robustness to changes in the length of the calibration window, respectively.

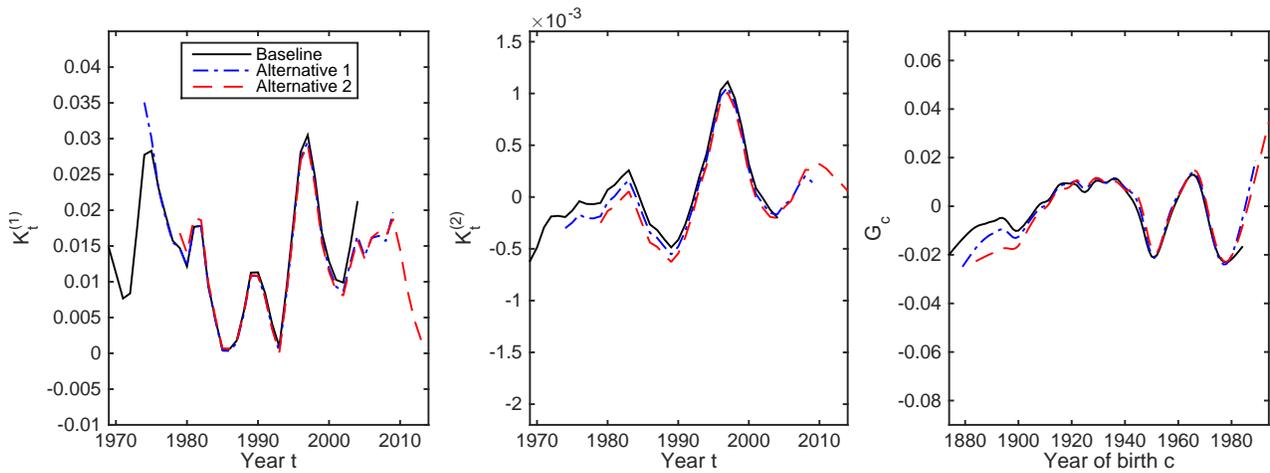


Figure 21: An example of **high robustness** to the calibration window: The estimated A/P/C components of historical mortality improvement obtained from the Route B simplified Plat model fitted to SSA male data.

- Baseline: ages 20 to 95, years 1968 to 2004.
- Alternative 1: ages 20 to 95, years 1973 to 2009.
- Alternative 2: ages 20 to 95, years 1978 to 2014.

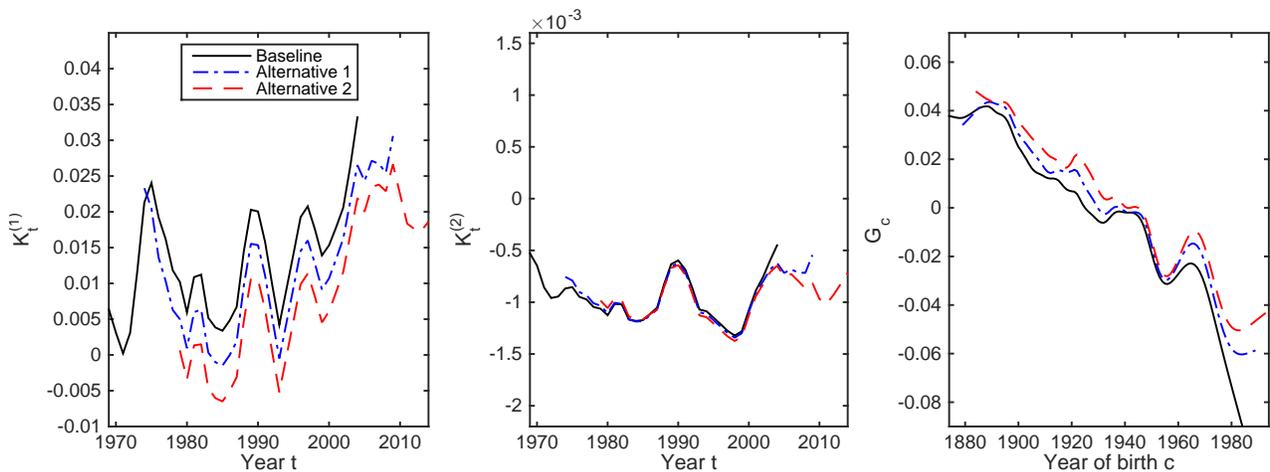


Figure 22: An example of **medium robustness** to the calibration window: The estimated A/P/C components of historical mortality improvement obtained from the Route B M6 model fitted to SSA female data.

- Baseline: ages 20 to 95, years 1968 to 2004.
- Alternative 1: ages 20 to 95, years 1973 to 2009.
- Alternative 2: ages 20 to 95, years 1978 to 2014.

4.3.4 Changes in the Age Range Used

Here we test the robustness to changes in age ranges. The following age ranges are considered:

	Starting Age	Ending Age	Starting Year	Ending Year	Number of Ages
Baseline	20	95			76
Alternative 1	30	85	1968	2014	56
Alternative 2	40	75			36

The result of this robustness test is summarized in the table below:

Model	SSA male		SSA female	
	Robustness measure	Category	Robustness measure	Category
M2	38.5%	Low	27.2%	Low
M3	5.2%	High	4.9%	High
M6	12.7%	Medium	10.2%	Medium
M7	7.8%	High	14.1%	Medium
M8	26.9%	Low	59.8%	Low
Full Plat	7.8%	High	10.0%	Medium
Simplified Plat	5.7%	High	5.6%	High
APCI	2.3%	High	3.2%	High

Figures 23, 24 and 25 illustrate high, low and medium levels of robustness to changes in the age ranges, respectively.

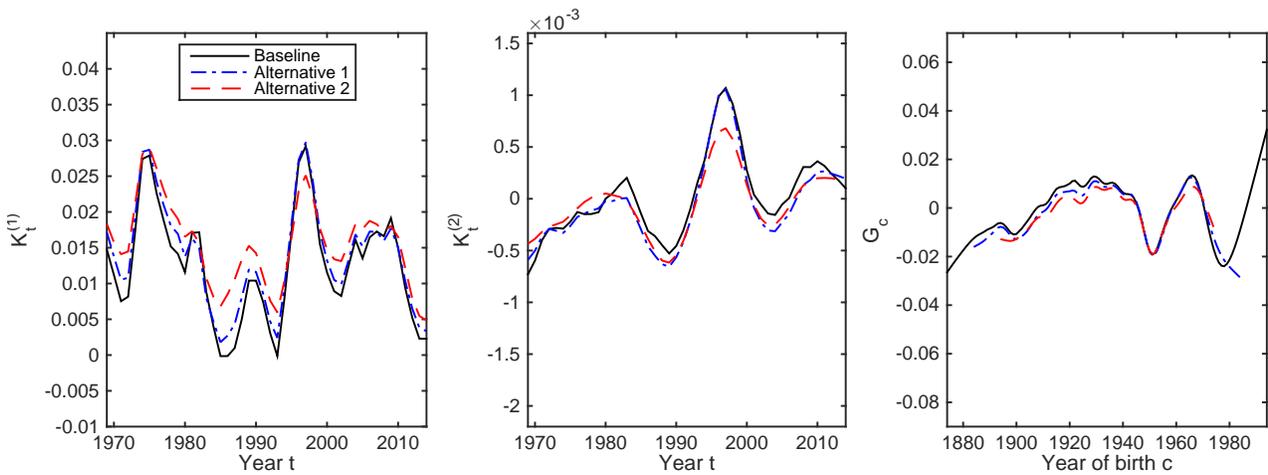


Figure 23: An example of **high robustness** to the age range used: The estimated A/P/C components of historical mortality improvement obtained from the Route B simplified Plat model fitted to SSA male data.

- Baseline: ages 20 to 95, years 1968 to 2014.
- Alternative 1: ages 30 to 85, years 1968 to 2014.
- Alternative 2: ages 40 to 75, years 1968 to 2014.

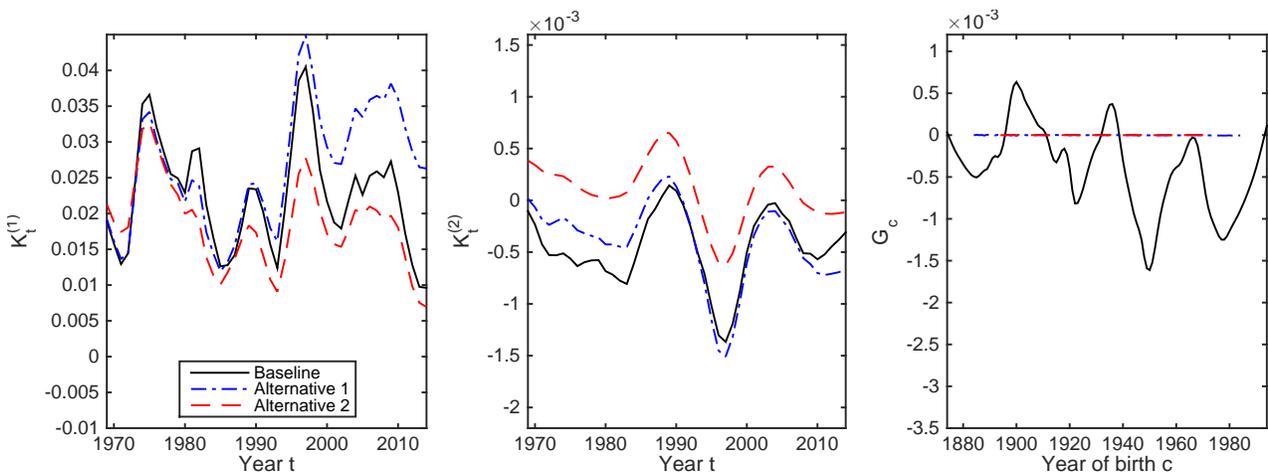


Figure 24: An example of **low robustness** to the age range used: The estimated A/P/C components of historical mortality improvement obtained from the Route B M8 model fitted to SSA male data.

- Baseline: ages 20 to 95, years 1968 to 2014.
- Alternative 1: ages 30 to 85, years 1968 to 2014.
- Alternative 2: ages 40 to 75, years 1968 to 2014.

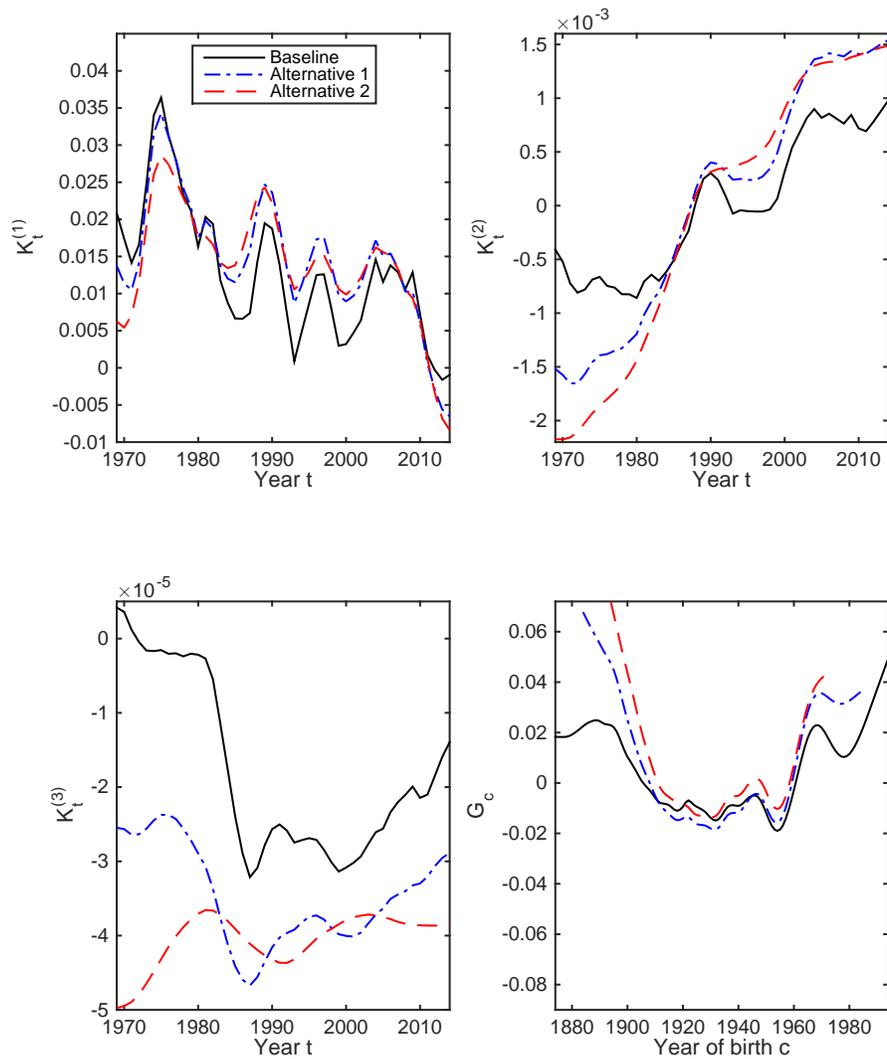


Figure 25: An example of **medium robustness** to the age range used: The estimated A/P/C components of historical mortality improvement obtained from the Route B M7 model fitted to SSA female data.

- Baseline: ages 20 to 95, years 1968 to 2014.
- Alternative 1: ages 30 to 85, years 1968 to 2014.
- Alternative 2: ages 40 to 75, years 1968 to 2014.

4.3.5 Choice of Parameter Constraints

Here we examine the robustness to the choice of identifiability constraints. The baseline and alternative constraints for each candidate model are listed below.

- **M2**
 Model structure: $\ln m_{x,t} = \beta_x^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} + \beta_x^{(3)} \gamma_{t-x}$
 Baseline constraints: $\sum_x \beta_x^{(1)} = 0, \sum_x \beta_x^{(2)} = 1, \sum_x \beta_x^{(3)} = 1, \sum_c n_c \gamma_c = 0$
 Alternative constraints: $\sum_x \beta_x^{(1)} = 0, \sum_x \beta_x^{(2)} = 1, \sum_x \beta_x^{(3)} = 1, \sum_c \gamma_c = 0$
- **M3**
 Model structure: $\ln m_{x,t} = \beta_x^{(1)} + \kappa_t^{(2)} + \gamma_{t-x}$
 Baseline constraints: $\sum_x \beta_x^{(1)} = 0, \sum_c \gamma_c = 0, \sum_c c \gamma_c = 0$
 Alternative constraints: $\sum_x \beta_x^{(1)} = 0, \sum_c n_c \gamma_c = 0, \sum_c n_c c \gamma_c = 0$
- **M6**
 Model structure: $\ln \frac{q_{x,t}}{1-q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}$
 Baseline constraints: $\sum_c \gamma_c = 0, \sum_c c \gamma_c = 0$
 Alternative constraints: $\sum_c n_c \gamma_c = 0, \sum_c n_c c \gamma_c = 0$
- **M7**
 Model structure: $\ln \frac{q_{x,t}}{1-q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \kappa_t^{(3)}((x - \bar{x})^2 - \hat{\sigma}_x^2) + \gamma_{t-x}$
 Baseline constraints: $\sum_c \gamma_c = 0, \sum_c c \gamma_c = 0, \sum_c c^2 \gamma_c = 0$
 Alternative constraints: $\sum_c n_c \gamma_c = 0, \sum_c n_c c \gamma_c = 0, \sum_c n_c c^2 \gamma_c = 0$
- **M8**
 Model structure: $\ln \frac{q_{x,t}}{1-q_{x,t}} = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}) + \gamma_{t-x}(x_c - x)$
 Baseline constraints: $\sum_c n_c \gamma_c = 0.$
 Alternative constraints: $\sum_c \gamma_c = 0$
- **The full Plat model**
 Model structure: $\ln m_{x,t} = \beta_x^{(1)} + \kappa_t^{(1)} + \kappa_t^{(2)}(\bar{x} - x) + \kappa_t^{(3)}(\bar{x} - x)_+ + \gamma_{t-x}$
 Baseline constraints: $\sum_x \beta_x^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_t \kappa_t^{(3)} = 0, \sum_c \gamma_c = 0, \sum_c c \gamma_c = 0, \sum_c c^2 \gamma_c = 0.$
 Alternative constraints: $\sum_x \beta_x^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_t \kappa_t^{(3)} = 0, \sum_c n_c \gamma_c = 0, \sum_c n_c c \gamma_c = 0, \sum_c n_c c^2 \gamma_c = 0$
- **The simplified Plat model**
 Model structure: $\ln m_{x,t} = \beta_x^{(1)} + \kappa_t^{(1)} + \kappa_t^{(2)}(\bar{x} - x) + \gamma_{t-x}$
 Baseline constraints: $\sum_x \beta_x^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_c \gamma_c = 0, \sum_c c \gamma_c = 0, \sum_c c^2 \gamma_c = 0.$
 Alternative constraints: $\sum_x \beta_x^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_c n_c \gamma_c = 0, \sum_c n_c c \gamma_c = 0, \sum_c n_c c^2 \gamma_c = 0$
- **The APCI model**
 Model structure: $\ln(m_{x,t}) = \beta_x^{(1)} + \beta_x^{(2)}(t - \bar{t}) + \kappa_t^{(1)} + \gamma_c$

Baseline constraints: $\sum_c \gamma_c = 0, \sum_c c\gamma_c = 0, \sum_c c^2\gamma_c = 0, \sum_t \kappa_t^{(1)} = 0$ and $\sum_t t\kappa_t^{(1)} = 0$.

Alternative constraints: $\sum_c n_c\gamma_c = 0, \sum_c n_c c\gamma_c = 0, \sum_c n_c c^2\gamma_c = 0, \sum_t \kappa_t^{(1)} = 0$ and $\sum_t t\kappa_t^{(1)} = 0$

The difference between the baseline and alternative constraints lies in the inclusion/exclusion of n_c , which represents the number of data points related to year-of-birth c . By including n_c , the cohorts about which we have more information are weighted more heavily in the parameter constraints.

We emphasize that there are many other combinations of constraints that can be used to stipulate parameter uniqueness. When very different constraints, the resulting A/P/C components may be very different. Our goal here is to examine the impact of small changes in the constraints used only.

The result of this robustness test is summarized in the table below:

Model	SSA male		SSA female	
	Robustness measure	Category	Robustness measure	Category
M2	10.5%	Medium	23.5%	Low
M3	0.4%	High	0.2%	High
M6	0.7%	High	0.0%	High
M7	1.1%	High	1.1%	High
M8	0.0%	High	18.5%	Medium
Full Plat	1.8%	High	1.4%	High
Simplified Plat	1.4%	High	0.7%	High
APCI	0.1%	High	0.7%	High

Figures 26, 27 and 28 illustrate high, medium and low levels of robustness to the choice of parameter constraints, respectively.

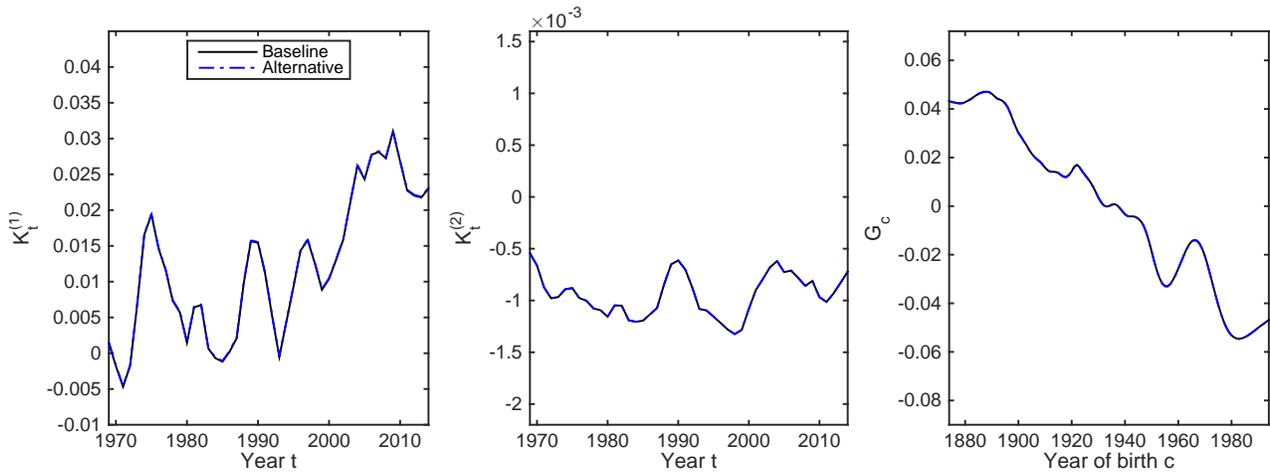


Figure 26: An example of **high robustness** to the parameter constraints used: The estimated A/P/C components of historical mortality improvement obtained from the Route B M6 model fitted to SSA female data.

- Baseline: $\sum_c \gamma_c = 0, \sum_c c\gamma_c = 0, \sum_c c^2\gamma_c = 0$
- Alternative: $\sum_c n_c\gamma_c = 0, \sum_c n_c c\gamma_c = 0, \sum_c n_c c^2\gamma_c = 0$

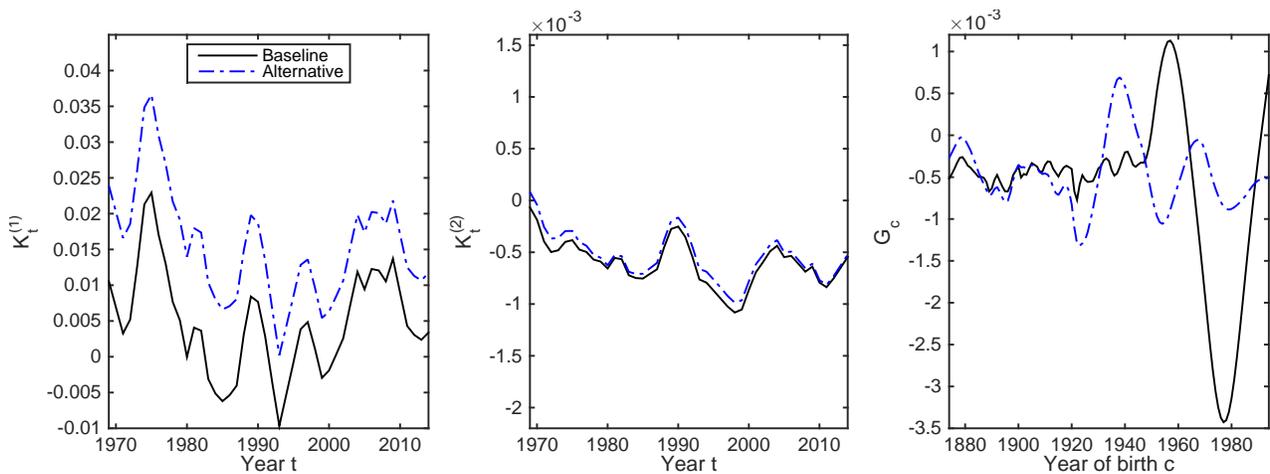


Figure 27: An example of **medium robustness** to the parameter constraints used: The estimated A/P/C components of historical mortality improvement obtained from the Route B M8 model fitted to SSA female data.

- Baseline: $\sum_c \gamma_c = 0$
- Alternative: $\sum_c n_c\gamma_c = 0$

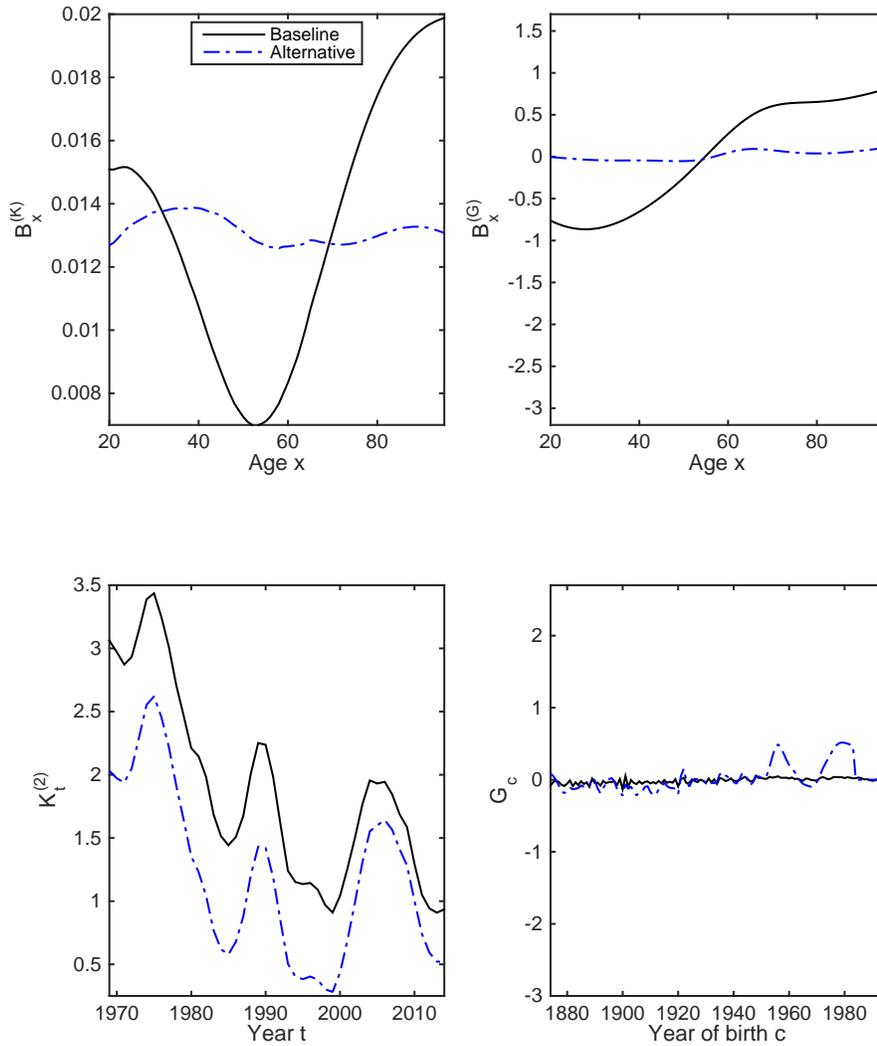


Figure 28: An example of **low robustness** to the parameter constraints used: The estimated A/P/C components of historical mortality improvement obtained from the Route B M2 model fitted to SSA female data.

- Baseline: $\sum_c \gamma_c = 0, \sum_c c\gamma_c = 0, \sum_c c^2\gamma_c = 0$
- Alternative: $\sum_c n_c\gamma_c = 0, \sum_c n_c c\gamma_c = 0, \sum_c n_c c^2\gamma_c = 0$

4.3.6 Exclusion of the Oldest and Newest Cohorts

Here we examine the robustness to the inclusion/exclusion of the oldest and newest cohorts. The following three situations are considered:

- Baseline: All available data are used.
- Alternative 1: The oldest and youngest five cohorts in the data sample are excluded.
- Alternative 2: The oldest and youngest ten cohorts in the data sample are excluded.

The result of this robustness test is summarized in the table below:

Model	SSA male		SSA female	
	Robustness measure	Category	Robustness measure	Category
M2	35.0%	Low	23.4%	Low
M3	2.9%	High	3.7%	High
M6	7.6%	High	6.3%	High
M7	8.9%	High	6.7%	High
M8	9.3%	High	11.7%	Medium
Full Plat	12.2%	Medium	11.0%	Medium
Simplified Plat	6.9%	High	6.8%	High
APCI	5.2%	High	3.2%	High

Figures 29, 30 and 31 demonstrate high, medium and low levels of robustness to the inclusion/exclusion of the youngest/oldest cohorts, respectively.

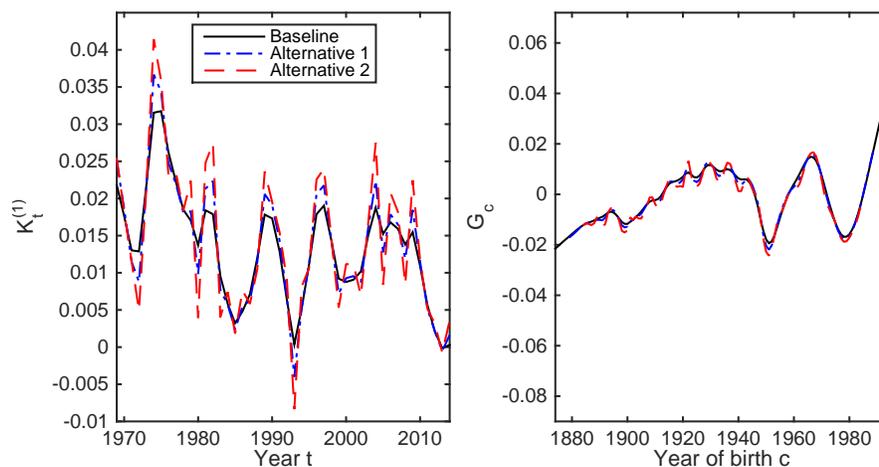


Figure 29: An example of **high robustness** to the inclusion/exclusion of the youngest/oldest cohorts: The estimated A/P/C components of historical mortality improvement obtained from the Route B M3 model fitted to SSA male data.

- Baseline: All available data are used.
- Alternative 1: The oldest and youngest five cohorts in the data sample are excluded.
- Alternative 2: The oldest and youngest ten cohorts in the data sample are excluded.

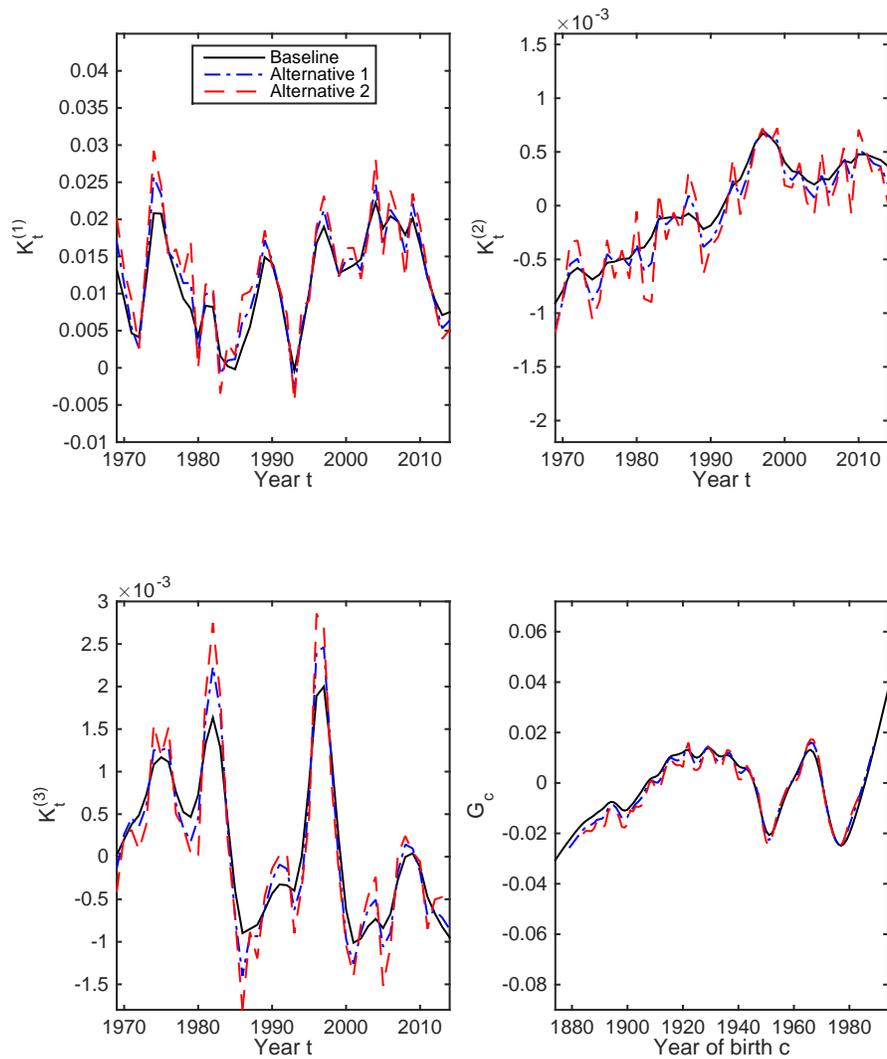


Figure 30: An example of **medium robustness** to the inclusion/exclusion of the youngest/oldest cohorts: The estimated A/P/C components of historical mortality improvement obtained from the Route B full Plat model fitted to SSA male data.

- Baseline: All available data are used.
- Alternative 1: The oldest and youngest five cohorts in the data sample are excluded.
- Alternative 2: The oldest and youngest ten cohorts in the data sample are excluded.

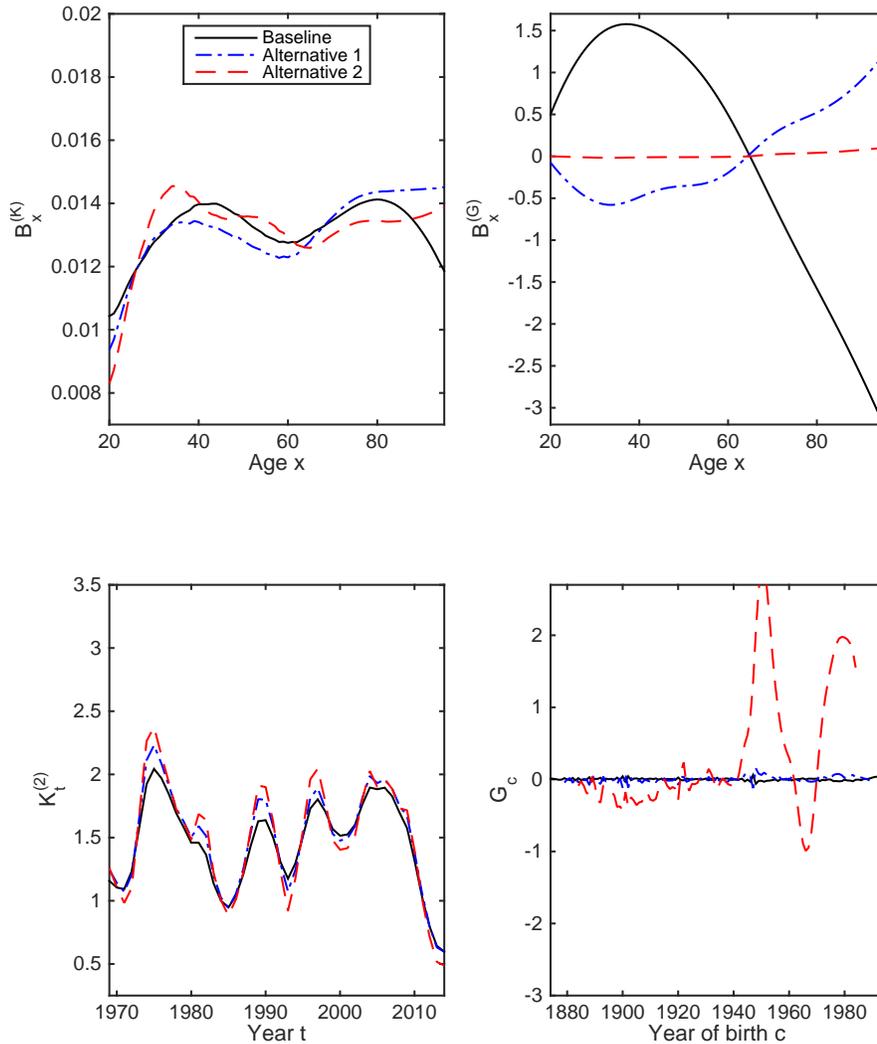


Figure 31: An example of **low robustness** to the inclusion/exclusion of the youngest/oldest cohorts: The estimated A/P/C components of historical mortality improvement obtained from the Route B M2 model fitted to SSA male data.

- Baseline: All available data are used.
- Alternative 1: The oldest and youngest five cohorts in the data sample are excluded.
- Alternative 2: The oldest and youngest ten cohorts in the data sample are excluded.

4.4 Analyzing the Standardized Residuals Produced by the Shortlisted Models

The table below summarizes the results of the robustness tests we performed for the Route B candidate models.

Robustness Test	M2	M3	M6	M7	M8	Full Plat	Simplified Plat	APCI
SSA males								
Tolerance value	Low	High	High	High	High	High	High	High
Calibration window	High	High	High	High	High	High	High	High
Age range	Low	High	Medium	High	Low	High	High	High
Parameter constraints	Medium	High	High	High	High	High	High	High
Exclusion of cohorts	Low	High	High	High	High	Medium	High	High
SSA females								
Tolerance value	Low	High	High	High	High	High	High	High
Calibration window	High	High	Medium	High	High	Medium	High	High
Age range	Low	High	Medium	Medium	Low	Medium	High	High
Parameter constraints	Low	High	High	High	Medium	High	High	High
Exclusion of cohorts	Low	High	High	High	Medium	Medium	High	High

The models that exhibit medium to high levels of robustness are given further consideration. The shortlisted models include M3, M6, M7, the full Plat, the simplified Plat model and the APCI model.

We now analyze the standardized residuals produced by the shortlisted models. The standardized residuals are calculated using the following formula:

$$\frac{D_{x,t} - E_{x,t}\hat{m}_{x,t}}{\sqrt{E_{x,t}\hat{m}_{x,t}}},$$

where $\hat{m}_{x,t}$ represents the fitted value of $m_{x,t}$ produced by the model being analyzed.

Because the models are estimated using a Poisson death count assumption, the standardized residuals are not necessarily normally distributed even if the model is adequate. As such, we cannot apply the ‘normal q-q plot’ and the normality test which we used when we pursue Route A. However, we can still assess the adequacy of a model by examining the heat map of its standardized residuals. If a model is adequate, then the pattern of its standardized residuals should be random with little clustering.

Figure 32 compares the heatmaps of standardized residuals obtained from M3, M6, M7, the full Plat, the simplified Plat model and the APCI model. The following observations are made:

- M6 and M7 perform the worst in the residual analysis. Large horizontal clusters are found in the heat maps produced from these models, indicating that age effect is not adequately captured.
- Residuals clustering is significant in M3 and the APCI model for males. Large vertical clusters are observed between mid-80s and mid-90s. Note that a similar problem is found in the M3 for males when Route A is used (see Volume 1 of the project report).

- The full Plat model performs the best in the residual analysis. The standardized residuals produced from the full Plat model appear to be the most random, although there still exist some gentle diagonal patterns.
- The simplified Plat model is the second best performing model in the residual analysis. Compared to those from full Plat model, the standardized residuals from the simplified Plat model exhibit slightly more clustering.

Note that the standardized residuals shown in Figure 32 cannot be directly compared and contrasted with those shown in Volume 1 of the project report, which are calculated in a different manner.

4.5 Concluding Remarks

Both the full Plat and simplified Plat models are reasonable choices for modeling the raw mortality rates of the US population. We recommend the simplified Plat model for use in Route B for the following reasons: (1) Although the full Plat model is the best performing model in the residual analysis, it only wins the simplified Plat model by a narrow margin. (2) The simplified Plat model outperforms the full Plat model in the robustness tests. The A/P/C decompositions resulting from the Route B simplified Plat model are shown in Figure 33 (males) and Figure 34 (females).

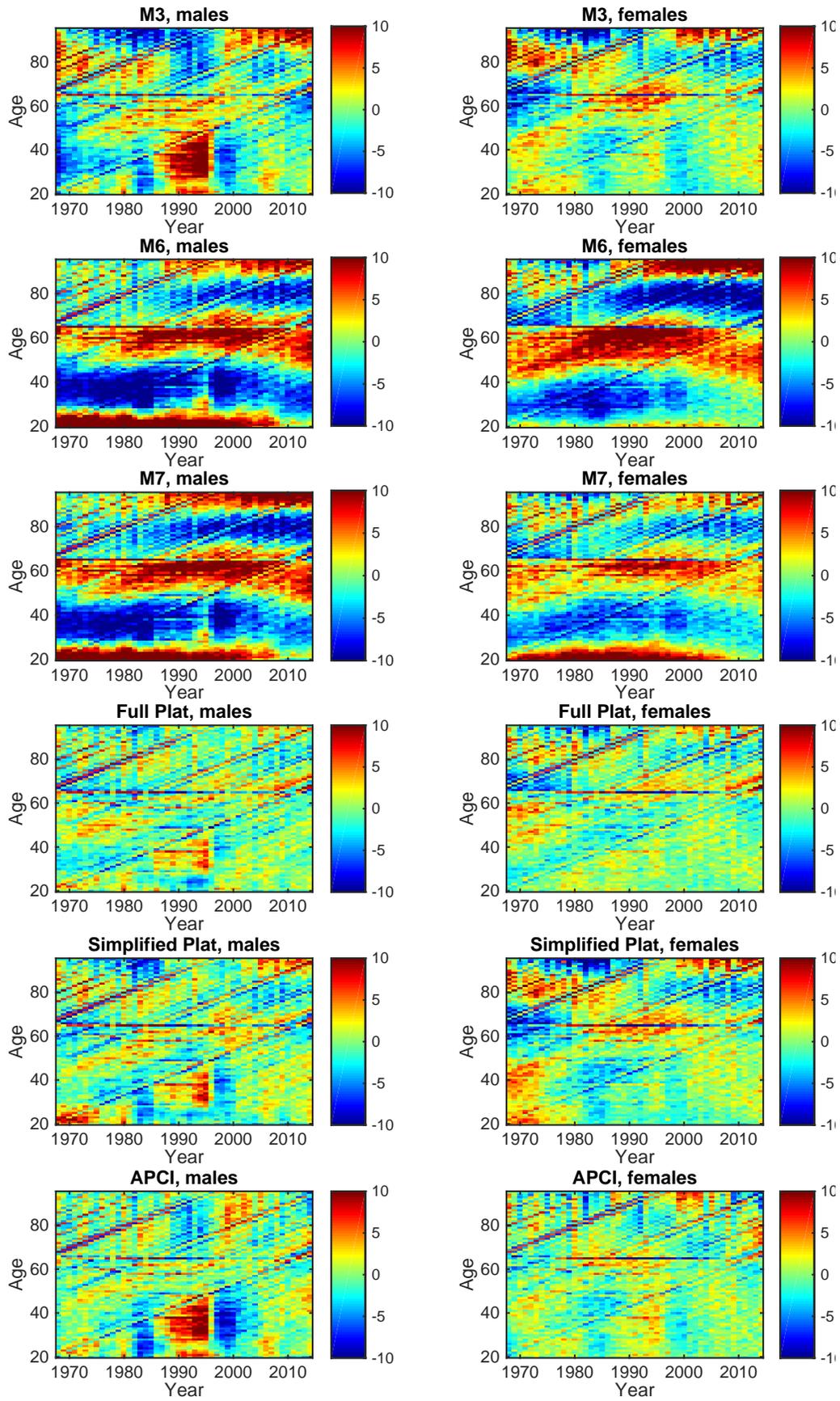


Figure 32: Heatmaps of the standardized residuals produced by the shortlisted Route B models.

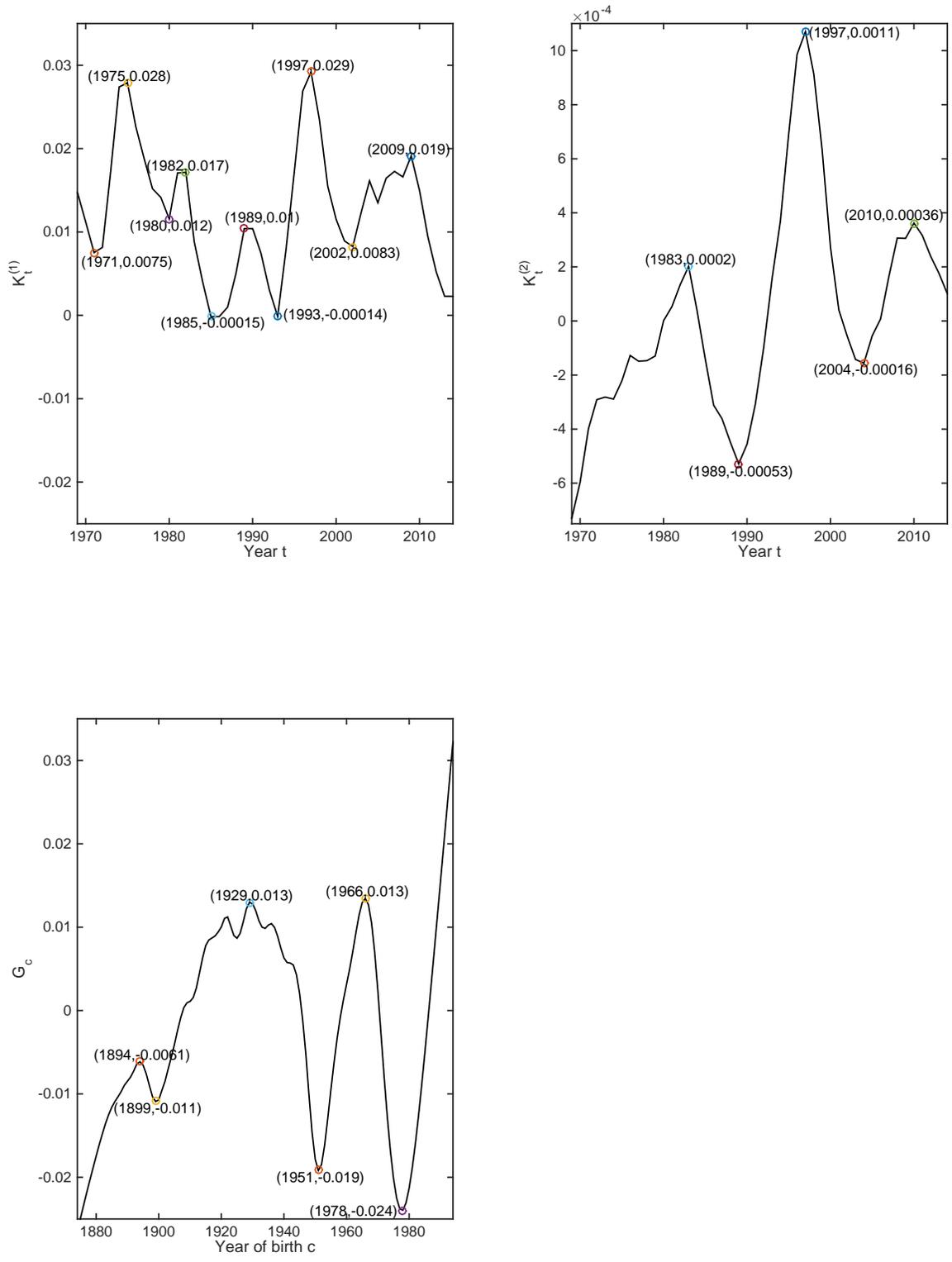


Figure 33: The estimated Age/Period/Cohort components obtained from the Route B simplified Plat model, U.S. males.

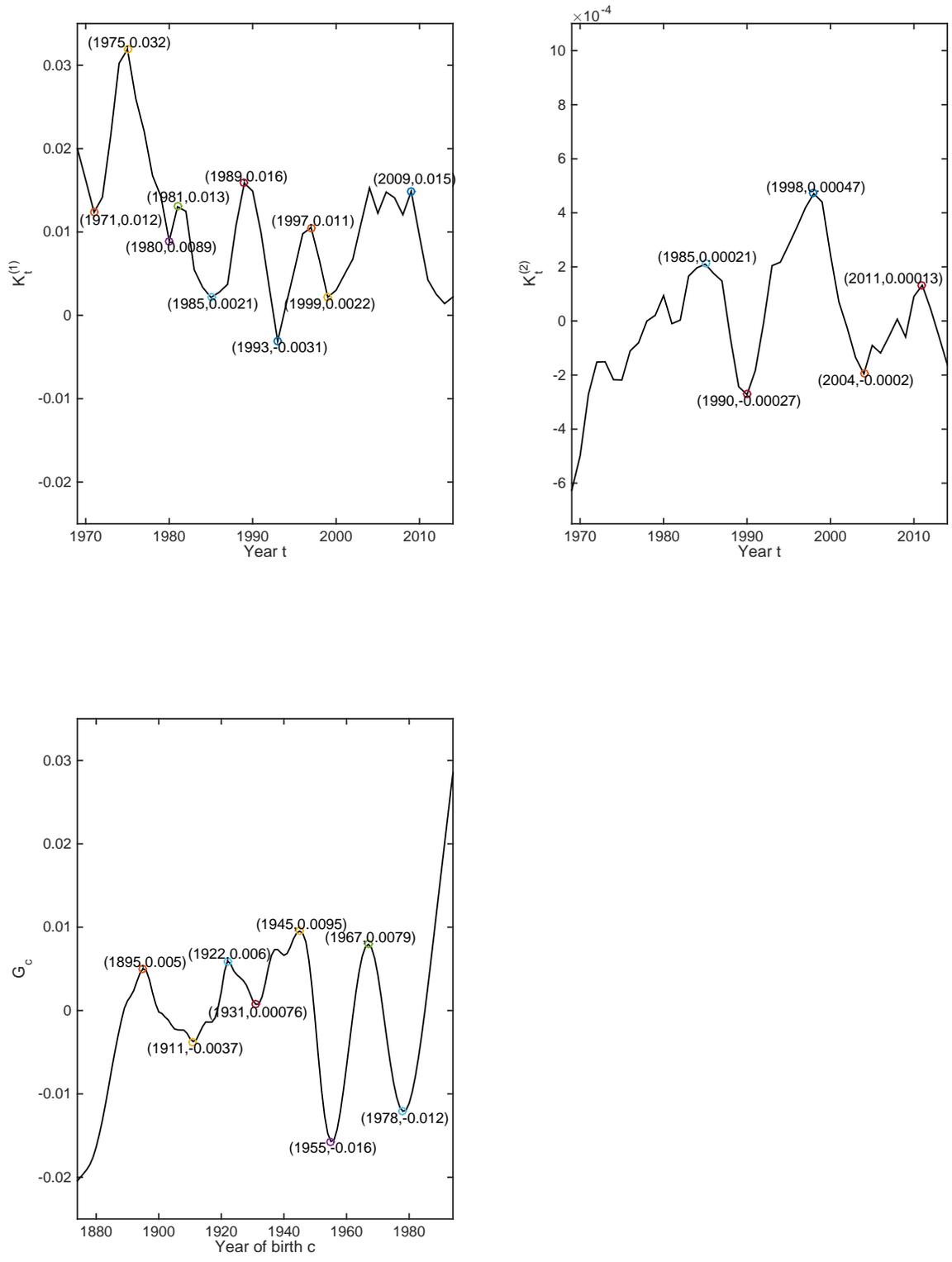


Figure 34: The estimated Age/Period/Cohort components obtained from the Route B simplified Plat model, U.S. females.

5 Repeating the Analyses Using Data for Ages 55 to 95

For the readers' information, in this section we repeat the analyses using the data for ages 55 to 95 only. The table below presents the results of the robustness tests for each of the Route B models. The baseline and alternative settings in these robustness tests are the same as before, except that when testing robustness changes in age ranges we consider ages 55-95 (baseline), ages 60-90 (alternative 1) and ages 65-85 (alternative 2). We also remark that M8 does not converge (for all of the tolerance values considered).

Robustness Test	M2	M3	M6	M7	M8	Full Plat	Simplified Plat	APCI
SSA males								
Tolerance value	High	High	High	High	High	High	High	High
Calibration window	High	High	High	High	Low	High	High	High
Age range	Low	High	High	High	Low	High	High	High
Parameter constraints	High	High	High	High	High	High	High	High
Exclusion of cohorts	Medium	High	High	High	Low	High	High	High
SSA females								
Tolerance value	Low	High	High	High	High	High	High	High
Calibration window	Medium	High	High	High	High	High	High	High
Age range	Medium	High	High	High	Low	High	High	High
Parameter constraints	Low	High	High	High	High	High	High	High
Exclusion of cohorts	Low	High	High	High	Low	High	High	High

M2 and M8 show low robustness in some robustness tests, and are therefore not given further consideration. Figure 35 compares the heatmaps of standardized residuals obtained from the short-listed models: M3, M6, M7, the full Plat, the simplified Plat model and the APCI model. In terms of goodness-of-fit (which is inversely related to the extent of residual clustering), the simplified Plat model clearly outperforms M3, M6 and M7, and perform similarly to the full Plat and APCI models.

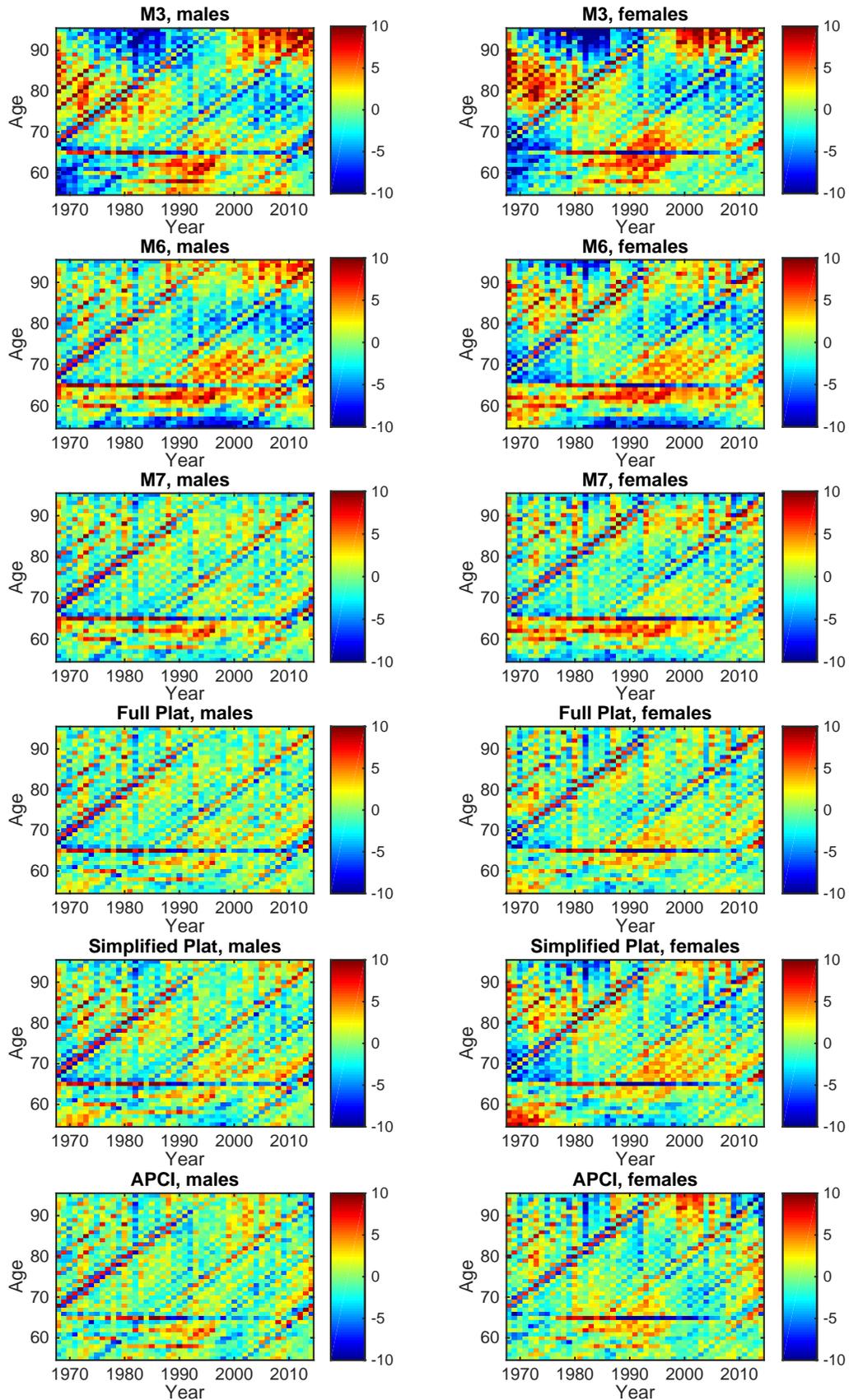
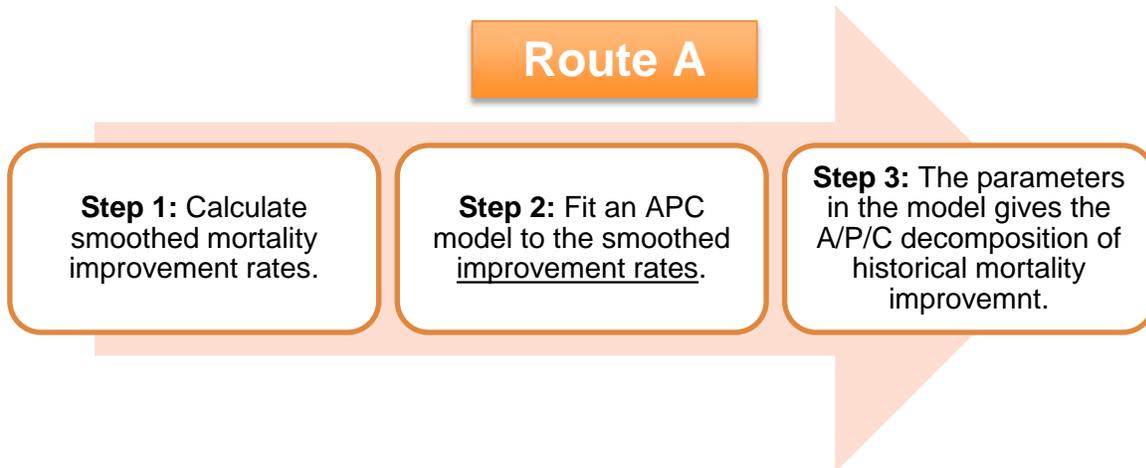


Figure 35: Heatmaps of the standardized residuals produced by the shortlisted Route B models for ages 55 to 95.

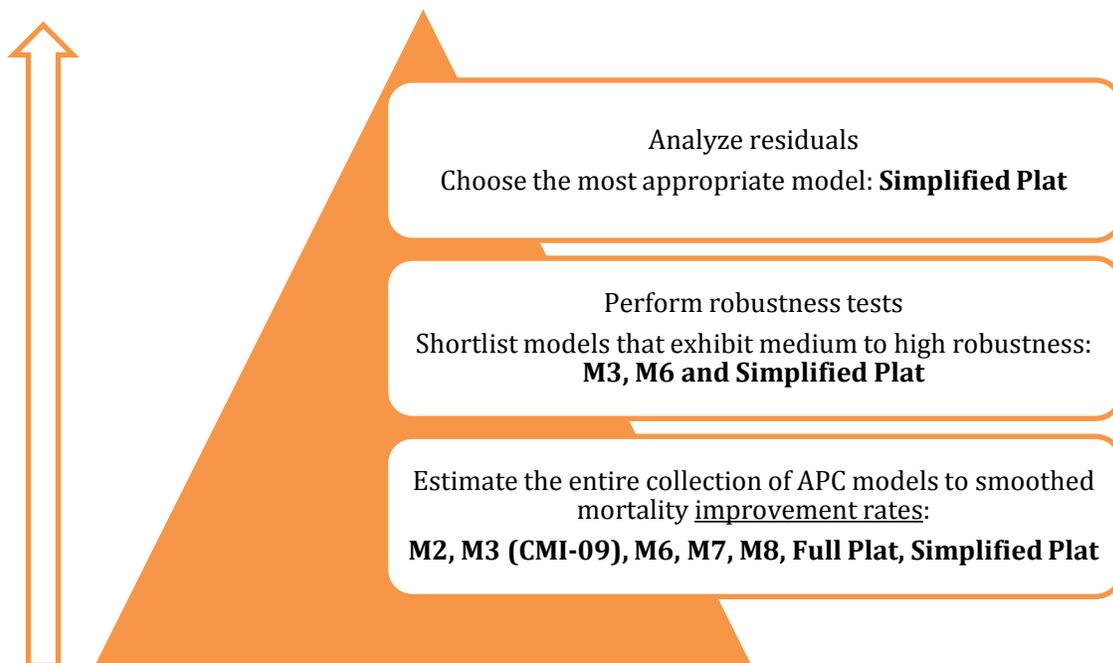
6 Comparing Routes A and B

6.1 A Summary of the Modeling Work for Route A

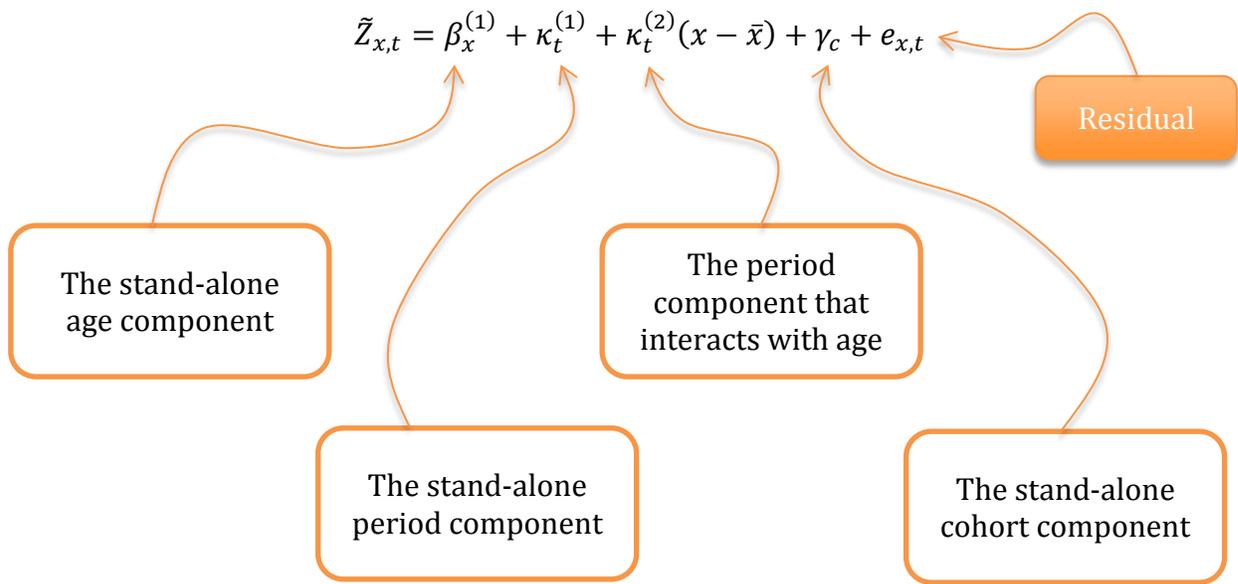
The distinction between the two routes lies in the quantity to which the APC models are fitted. In Route A, the APC models are fitted to *mortality improvement rates*. It entails the following steps:



Seven candidate APC structures are considered for this route. The model selection process is summarized in the flow chart below.



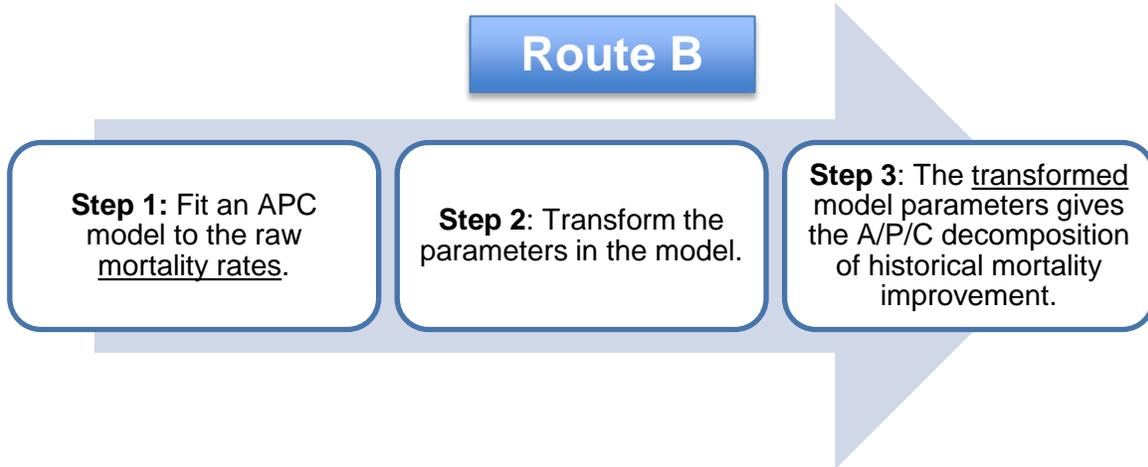
In Route A, the most effective model (the simplified Plat model) implies that mortality improvement is driven by (1) a stand-alone age component, (2) a stand-alone period component, (3) a period component that interacts with (a linear function) of age, and (4) a stand-alone cohort component:



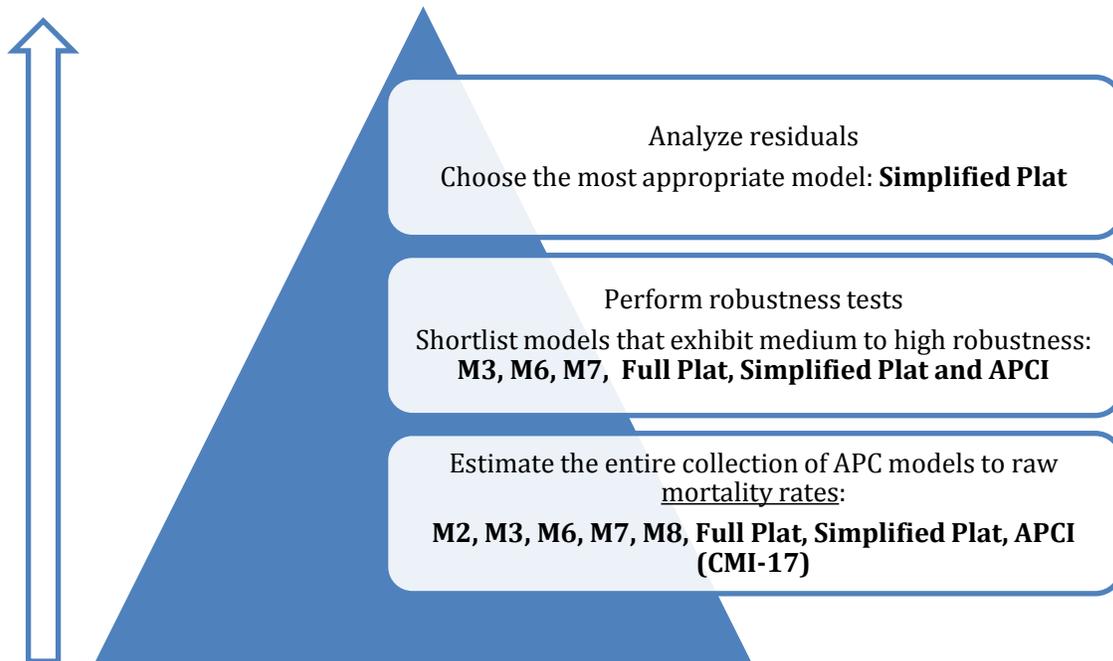
Note that in the Route A simplified Plat model, mortality improvement is defined as the percentage reduction in (smoothed) death probabilities: $\tilde{Z}_{x,t} = 1 - \frac{\tilde{q}_{x,t}}{\tilde{q}_{x,t-1}}$, where \tilde{q} denotes a smoothed one-year death probability.

6.2 A Summary of the Modeling Work for Route B

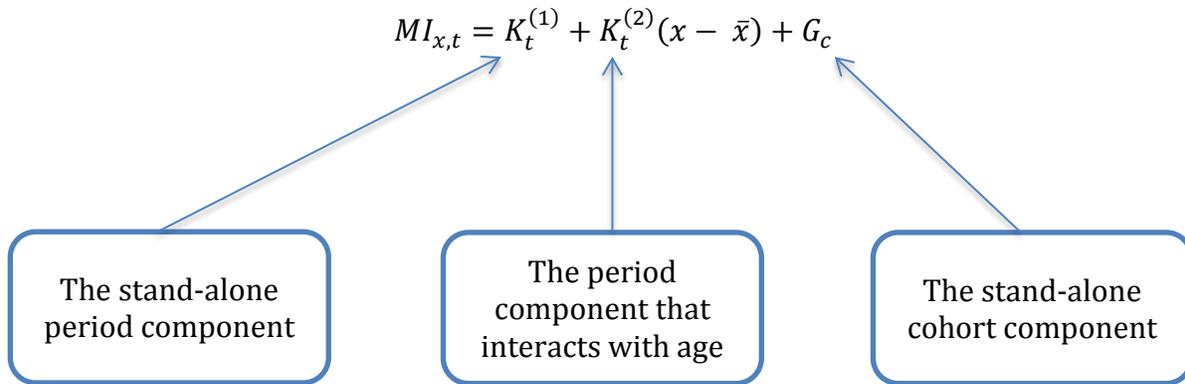
In Route B, the APC models are fitted to *mortality rates*. It entails the following steps:



Eight candidate APC structures are considered for this route. The model selection process is summarized in the flow chart below.



In Route B, the most effective model (the simplified Plat model) implies that mortality improvement is driven by (1) a stand-alone period component, (2) a period component that interacts with (a linear function) of age, and (3) a stand-alone cohort component:



Note that in the Route B simplified Plat model, mortality improvement is defined as the change in log central death rates (known as M-style mortality improvement by the CMI): $MI_{x,t} = \ln(m_{x,t-1}) - \ln(m_{x,t})$, where m denotes a central death rate.

6.3 The Final Conclusion

We conclude that the Route A simplified Plat model is the most appropriate model for identifying the A/P/C components of the US gender-specific historical mortality improvements. The conclusion is drawn on the following bases.

- We have argued previously that the simplified Plat model performs the best in both Routes A and B. Figures 36 and 37 provide a side-by-side comparison between the A/P/C components of historical mortality improvement derived from the two routes.⁷ The A/P/C components derived from the two routes are not too different. However, the A/P/C components from Route B are noticeably less smooth. The lack of smoothness is a concern, in part because the underlying A/P/C components should in principle be free of noise (the noise should be captured by the residual component) and in part because it would be more challenging (for the researchers of the follow-up project) to link the A/P/C components to intrinsic factors if they are jagged.
- Also, one technical limitation of Route B is that the smoothing parameters for Route B models are chosen with a rather subjective approach instead of a more statistically rigorous cross validation. Therefore, from a technical viewpoint, the Route A simplified Plat model is more preferred than the Route B counterpart.

The numerical values of the A/P/C components identified from this model are provided in the accompanying Excel workbook.

We emphasize that the conclusion drawn is data-dependent, applicable only to the data set under consideration (US gender-specific, ages 20-95, years 1968-2014). Although the model selection procedures (Routes A and B) stands for any data set, the conclusion drawn can be (very) different when a different population is considered.

We also emphasize that the conclusion might be different if we focus on a particular age range (say, pensionable ages only). This point is manifested in the the analyses performed in Section 7 of Volume 1: when applied to ages above 55 only, the advantage of the simplified Plat model over the CMI-09 model (M3) becomes less apparent.

Furthermore, the selected model (Route A simplified Plat) is only the best currently. While the robustness test results indicate that the conclusion is unlikely to change in the next few years, it is entirely possible that the conclusion will change when a substantial volume of new data is added to the analyses, for reasons such as structural changes in the underlying A/P/C components.

Finally, we fully acknowledge that this study focuses only on the overall US population, without considering the possible differences in mortality improvements due to, for example, geographical differences and differences in socioeconomic status. It is warranted to revisit this study using more granular data when such data becomes available in the future.

⁷Note that the ‘stand-alone age component’ does not apply to Route B.

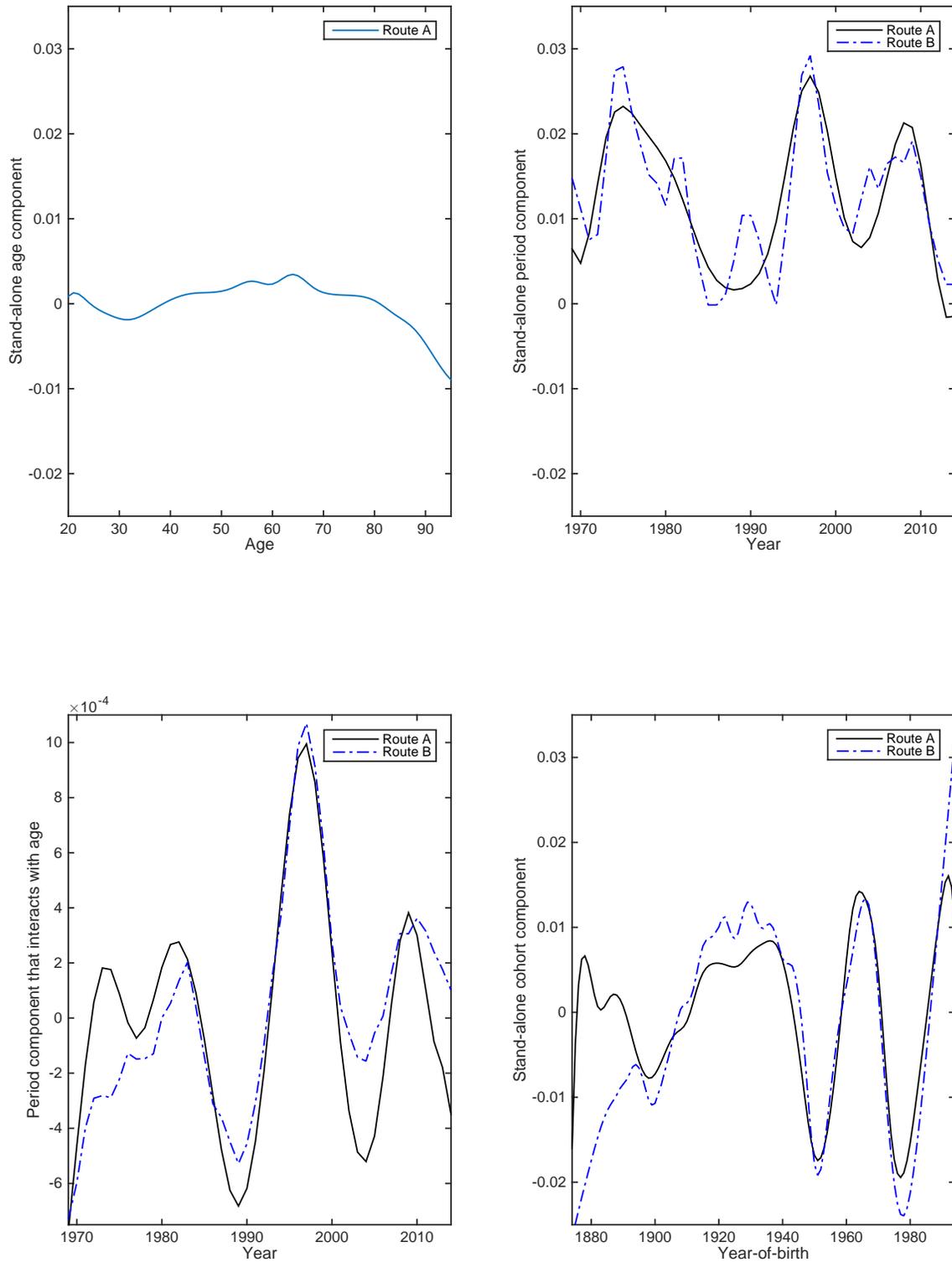


Figure 36: A comparison between the A/P/C components obtained from the most effective Route A and Route B models, U.S. **males**.

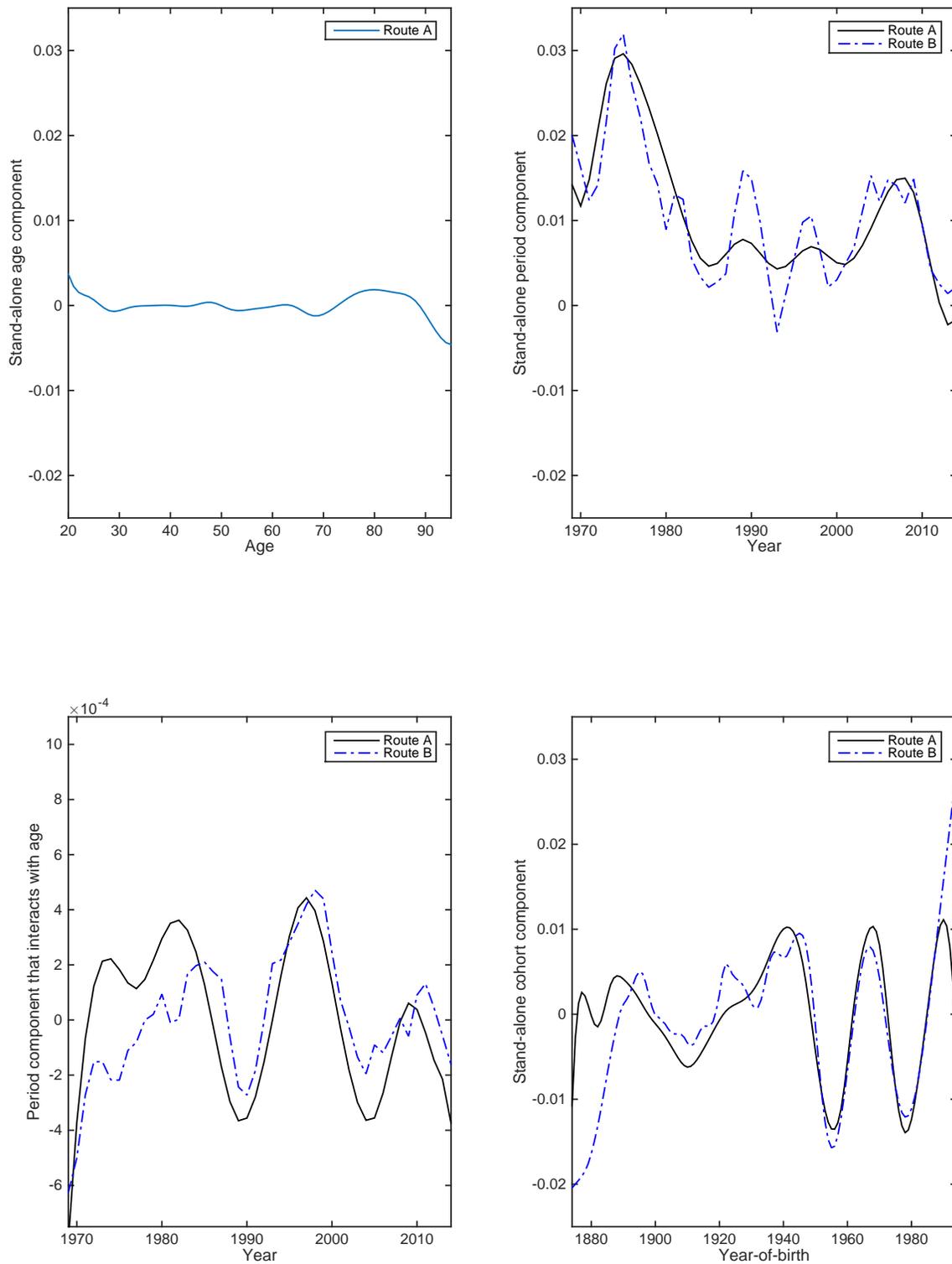


Figure 37: A comparison between the A/P/C components obtained from the most effective Route A and Route B models, U.S. **females**.

7 Acknowledgments

The project team is grateful to all members of the Project Oversight Group, formed by Jennifer Haid (Chair), Jean-Marc Fix, Zach Granovetter, George Graziani, Alla Kleyner, Dale Hall, Bob Howard, Al Klein, Andy Peterson, Larry Pinzur, Erika Schulty, and Larry Stern, for their guidance and insightful comments.

Appendix 1

In this appendix, we show that $MI_{x,t} = \ln m_{x,t-1} - \ln m_{x,t} \approx Z_{x,t} = 1 - \frac{q_{x,t}}{q_{x,t-1}}$ under the assumption that the force of mortality between two consecutive integer ages is a constant.

Proof. First, we rewrite $MI_{x,t}$ as follows:

$$\begin{aligned} MI_{x,t} &= \ln m_{x,t-1} - \ln m_{x,t} \\ &= -\ln \frac{m_{x,t}}{m_{x,t-1}} \\ &= -\ln \left(1 - \frac{m_{x,t-1} - m_{x,t}}{m_{x,t-1}} \right). \end{aligned}$$

Let $f(y) = \ln(y)$. Using the first order Taylor expansion, $f(y + dy) \approx f(y) + \frac{d}{dy}f(y)dy$ at $y = 1$, we have

$$\begin{aligned} -\ln \left(1 - \frac{m_{x,t-1} - m_{x,t}}{m_{x,t-1}} \right) &\approx -\ln(1) + \frac{m_{x,t-1} - m_{x,t}}{m_{x,t-1}} \\ &= \frac{m_{x,t-1} - m_{x,t}}{m_{x,t-1}}. \end{aligned}$$

Assuming that the force of mortality between two consecutive integer ages is a constant, we have $q_{x,t} = 1 - e^{m_{x,t}}$. Let $f(y) = e^y$. Using the first order Taylor expansion at $y = 0$, we have

$$\begin{aligned} e^{-m_{x,t}} &\approx e^0 + e^0(-m_{x,t}) \\ q_{x,t} &= 1 - e^{m_{x,t}} \\ &\approx 1 - (1 - m_{x,t}) \\ &= m_{x,t} \end{aligned}$$

Therefore,

$$\begin{aligned} MI_{x,t} &\approx \frac{m_{x,t-1} - m_{x,t}}{m_{x,t-1}} \\ &\approx \frac{q_{x,t-1} - q_{x,t}}{q_{x,t-1}} \\ &= 1 - \frac{q_{x,t}}{q_{x,t-1}} = Z_{x,t}. \end{aligned}$$

□

Appendix 2

In this appendix, we show that $\ln m_{x,t-1} - \ln m_{x,t} \approx \ln \frac{q_{x,t-1}}{1-q_{x,t-1}} - \ln \frac{q_{x,t}}{1-q_{x,t}}$ when the force of mortality between two consecutive integer ages is a constant.

Proof. Let $y_{x,t} = \ln \frac{q_{x,t}}{1-q_{x,t}}$. We have

$$q_{x,t} = \frac{e^{y_{x,t}}}{1 + e^{y_{x,t}}}.$$

Assuming the force of mortality between two consecutive integer ages is a constant, we have $q_{x,t} = 1 - e^{-m_{x,t}}$, which gives

$$\begin{aligned} m_{x,t} &= -\ln(1 - q_{x,t}) \\ &= -\ln \frac{1}{1 + e^{y_{x,t}}} \\ &= \ln(1 + e^{y_{x,t}}). \end{aligned}$$

Using the first order Taylor expansion at 1, we have

$$\begin{aligned} m_{x,t} &\approx \ln(1) + e^{y_{x,t}} \\ &= e^{y_{x,t}}, \end{aligned}$$

so that $\ln m_{x,t} \approx y_{x,t}$. Therefore,

$$\ln m_{x,t-1} - \ln m_{x,t} \approx \ln \frac{q_{x,t-1}}{1 - q_{x,t-1}} - \ln \frac{q_{x,t}}{1 - q_{x,t}}.$$

□

References

- Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., and Balevich, I. (2009). A Quantitative Comparison of Stochastic Mortality Models Using Data from England and Wales and The United States. *North American Actuarial Journal*, **13**, 1-35.
- Continuous Mortality Investigation Bureau (2017a). *CMI Mortality Projections Model: CMI_2016*. Mortality Projections Committee Working Paper 97.
- Continuous Mortality Investigation Bureau (2017b). *CMI Mortality Projections Model: Methods*. Mortality Projections Committee Working Paper 98.
- Continuous Mortality Investigation Bureau (2017c). *CMI Mortality Projections Model: Software user guide*. Mortality Projections Committee Working Paper 99.
- Osmond, C. (1985). Using Age, Period and Cohort Models to Estimate Future Mortality Rates. *International Journal of Epidemiology*. **14**, 124-129.
- Plat, R. (2009). On stochastic mortality modeling. *Insurance: Mathematics and Economics*, **45**, 393-404.
- Renshaw, A.E., and Haberman, S. (2006). A Cohort-Based Extension to the Lee-Carter Model for Mortality Reduction Factors. *Insurance: Mathematics and Economics*, **38**, 556-570.

About The Society of Actuaries

The Society of Actuaries (SOA), formed in 1949, is one of the largest actuarial professional organizations in the world dedicated to serving more than 27,000 actuarial members and the public in the United States, Canada and worldwide. In line with the SOA Vision Statement, actuaries act as business leaders who develop and use mathematical models to measure and manage risk in support of financial security for individuals, organizations and the public.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and nonactuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.

Society of Actuaries
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org