A Multi-Parameter-Level Model for Simulating Future Mortality Scenarios with COVID-Alike Effects

Sixteenth International Longevity Risk and Capital Markets Solutions Conference

13 August 2021

Rui Zhou (Presenter) University of Melbourne
Johnny Siu-Hang Li University of Waterloo

Table of Contents

Intro

Challenges in Modelling Pandemic Impact

The COVID pandemic poses significant uncertainty on the future mortality experience of insurance policyholders and pension members

- Early days of the pandemic
  - Limited information on the disease
  - Relying on infectious disease studies
- Over a year into the pandemic
  - Virus mutation, vaccine rollout
  - When and how the pandemic will end is still uncertain
- Over the longer term
  - Difficult to quantify the long-term health effect of COVID, the potential positive impact induced by pandemic, and recurrence of pandemics

External Forecasts and Expert Opinions

We rely on external forecasts based on infectious disease models and expert opinions to understand the pandemic impact

- External forecasts
  - Up-to-date quantitative forecasts from infectious disease studies
- Expert opinions
  - Views drawing on the comprehensive knowledge of experts without quantitative modelling
  - Expert opinion has been shown useful in strengthening models where data is lacking and in modifying models to suit practical needs in various disciplines

Our Research

We develop a stochastic model for simulating future mortality scenarios with COVID-alike effects

- Incorporating expert opinions on and external forecasts of pandemic impact
- Encompassing three parameter levels
  - Level 1 captures the long-term pattern of mortality implied by the historical mortality data
  - Level 2 gauges the excess age-specific mortality due to COVID-19 based on external forecasts
  - Level 3 draws on expert opinions concerning difficult-to-quantify pandemic impact

Our Contribution

Our model utilizes the newest knowledge of the pandemic when predicting its impact on future mortality experience

- Previous studies (Chen and Cox, 2009; Zhou et al., 2013; Liu and Li, 2015) extrapolate the pattern of historical jump effects into the future
- Each pandemic has its distinct features
  - Half of the 1918 Spanish flu deaths are from the 20-40 age group
  - Majority of the COVID deaths are reported in people aged 70+
- Society changes, such as globalization, may also affect the transmission and severity of the disease

Our Contribution

We propose a single-stage estimation method

- Parameters are estimated by maximizing penalized quasi-likelihood (PQL)
- Commonly used two-stage estimation method for the Lee-Carter model cannot disentangle the pandemic disruption and the fluctuation of long-term mortality improvement

We investigate parameter uncertainty arising from

- Imputation
- Uncertainty in external forecasts

Data

Historical U.S. Mortality Data

Historical mortality data allow us to extrapolate the long-term mortality pattern into the future

- Obtained from the Human Mortality Database (2021)
- Including
  - the number of deaths and the number of exposures-at-risk in the period of 1970-2019 and the age range of 20-100
  - the U.S. population size in 2020

Estimated Infection Fatality Ratio (IFR)

Infection fatality ratio (IFR) is the number of deaths per infection

- Ferguson et al. (2020)
    - Conducted by the Imperial College COVID-19 Response Team in March 2020
    - Provides early estimates of COVID IFRs and predicts number of COVID deaths under no or various suppression strategies
- Centres for Disease Control and Prevention (2020)
    - Provides five scenarios representing different levels of disease transmission and severity to advance public health preparedness and planning
    - We use the set of scenarios published in September 2020
- O'Driscoll et al. (2021)
    - A systematic review and meta-analysis of published evidence on COVID-19 mortality until July 2020
    - Generates 15,000 simulated scenarios of IFRs

Estimated IFRs from the Three Studies

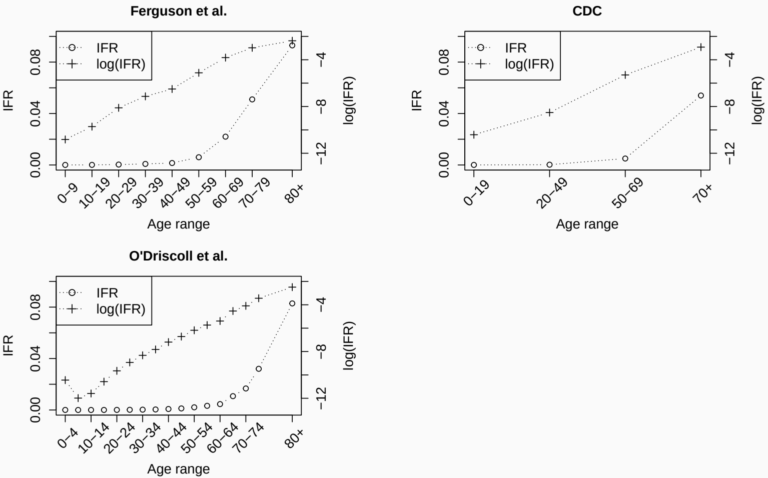

Figure 1: Best estimates of infection fatality ratio (IFR) from three different studies

Interpolated IFRs

To make the three sets of estimates comparable, we interpolate the log scaled IFR using cubic splines

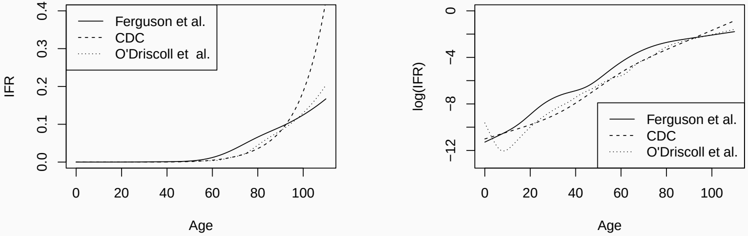

Figure 2: Interpolated IFRs using cubic splines

Predicted Number of COVID Deaths

Predicted number of COVID deaths = Population size $\times$ estimated IFR $\times$ predicted infection percentage

- Ferguson et al. (2020) predicts 81% infection percentage in the U.S. under the worst case scenario
- Total number of deaths can be significantly reduced under suppression strategies
- Expert opinion should be used to predict which suppression strategy is likely to be used and the resulting infection percentage
- We investigate various scenarios of infection percentage ranging from 10% to 80%

The Model

Level 1: The Long-Term Mortality Pattern

We decompose the long-term mortality pattern into age and period effects in the same way with the well known Lee-Carter model

$$\begin{aligned} D_{x,t} &\sim \text{Poisson}(E_{x,t} m_{x,t}) \\ \ln m_{x,t} &= a_x + b_x k_t \end{aligned}$$

- $x \in [x_0, x_1]$ and $t \in [t_0, t_1]$
- $D_{x,t}$, $E_{x,t}$, and $m_{x,t}$: number of deaths, number of exposures-at-risk, and central death rate at age x in year t
- a_x: average log mortality level at age x
- k_t: period effect driving the mortality change over time
- b_x: the age-specific sensitivity to the changes in k_t
- $\sum_x b_x = 1$, and $k_{t_0} = 0$ to ensure parameter uniqueness

Level 1: The Long-Term Mortality Pattern

k_t is often modelled by a random walk or an ARIMA model

$$k_t = k_{t-1} + \mu + \epsilon_t$$

- μ is the drift
- $\{\epsilon_t\}$ is a sequence of i.i.d. normal random variables and $\epsilon_t \sim N(0, \sigma^2)$

The first parameter level contains all the parameters related to the long-term mortality structure, including a_x, b_x, k_t, μ, and σ

Level 2: Excess Age-Specific Mortality due to Pandemic

The impact of pandemic on mortality varies significantly by age

- Our model (Level 1 + Level 2)

$$\ln m_{x,t} = a_x + b_x k_t + c_{x,t}\pi_t \mathbb{1}_{\{T \le t \le T+n\}}$$

  - Pandemic begins in year T and its impact lasts for n years
  - π_t: overall severity of the pandemic in year t
  - $c_{x,t}$: the age pattern of the pandemic impact in year t
  - COVID increases mortality by the multiplicative factor $e^{c_{x,t}\pi_t \mathbb{1}_{\{T \le t \le T+n\}}}$
- Previous studies

$$\ln m_{x,t} = a_x + b_x(k_t + c_t \mathbb{1}_{\{T \le t \le T+n\}})$$

  - Mortality shocks share the same age pattern with the long-term mortality improvement

Level 3: Expert Opinions on the Long-Term Pandemic Impact

The parameters in the third level gauge the long-term impact of a pandemic which is often difficult to quantify

- The complete model

$$\ln m_{x,t} \quad = \quad a_x + b_x k_t + \sum_i \left(c_{x,t}^{(i)} \pi_t^{(i)} \mathbb{1}_{\{T_i \leq t \leq T_i + n_i\}} \right) + d_{x,t}^{(i)} \psi_t^{(i)} \mathbb{1}_{\{t \geq T_i + 1\}} \Big)$$

  - T_i: arrival year of the ith pandemic
  - The arrival of pandemics follows a Poisson process

$$T_i - T_{i-1} \sim Exp(1/\lambda),$$

    where λ is chosen based on expert opinions
  - n_i: number of years with excess mortality due to the ith pandemic
  - $c_{x,t}^{(i)}$ and $\pi_t^{(i)}$: age pattern and size of excess mortality in year t
  - $d_{x,t}^{(i)}$ and $\psi_t^{(i)}$: age pattern and size of positive impact

Estimation of Parameters in the First Two Levels

The Two-Stage Model Estimation

The Lee-Carter model and its variants are often estimated using a two-stage method

- The two stages
  1. Estimate a_x, b_x, and k_t by maximising the loglikelihood function

$$\sum_{x,t}(D_{x,t} \ln m_{x,t} - E_{x,t} m_{x,t})$$

  2. Estimate μ and σ in the random walk model for $\{k_t\}$
- Missing data
  - The total number of deaths $D_{x,2020}$ is required for estimation
  - The number of COVID deaths in 2020 is calculated from external forecasts
  - The number of non-COVID deaths is unknown until end of 2020
  - We impute the number of non-COVID deaths in 2020 with simulated number of deaths from the long-term mortality pattern

The Baseline Case

We assume COVID does not affect the number of deaths by other causes, i.e. COVID deaths = Excess deaths

- 80% of infection percentage
- Best estimates of IFRs from Ferguson et al. (2020)
- Excess mortality due to COVID only occurs in 2020

Estimated Parameters using the Two-Stage Estimation

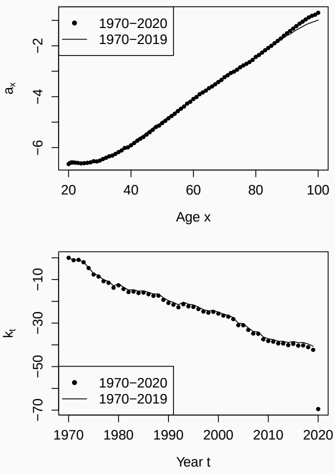

Figure 3: Estimated parameters using the two-stage procedure and a set of imputed non-COVID deaths

A Single-Stage Model Estimation

A single-stage method simultaneously estimates the first- and second-level parameters

- View the model as a general linear mixed model (GLMM)
- Parameters in the GLMM can be estimated by maximizing penalized quasi-likelihood (PQL) (Breslow and Clayton, 1993)

$$\sum_{x,t}(D_{x,t}\ln m_{x,t} - E_{x,t}m_{x,t}) - \frac{1}{2}(\vec{k}-\vec{\mu})'V^{-1}(\vec{k}-\vec{\mu})$$

  - $\vec{k} = [k_{t_0+1}, k_{t_0+2}, \ldots, k_{t_1}]'$ and $\vec{k} \sim MVN(\vec{\mu}, V)$
  - The joint density function of $\vec{k}$ is

$$f(\vec{k}) = 2\pi^{-\frac{t_1-t_0}{2}}|V|^{-\frac{1}{2}}e^{-\frac{1}{2}(\vec{k}-\vec{\mu})'V^{-1}(\vec{k}-\vec{\mu})}$$

  - The PQL estimation makes a trade-off between likelihood and how closely $\vec{k}$ follows a multivariate normal distribution

Estimated Parameters using the Single-Stage Estimation

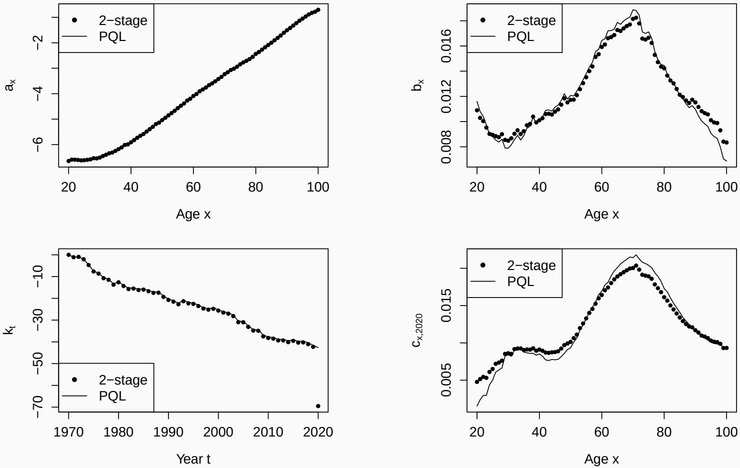

Figure 4: Estimated parameters by maximizing PQL

$\pi_{2020} = 60.74$ using the two-stage procedure

$\pi_{2020} = 45.62$ using the PQL method

Parameter Uncertainty – Imputations

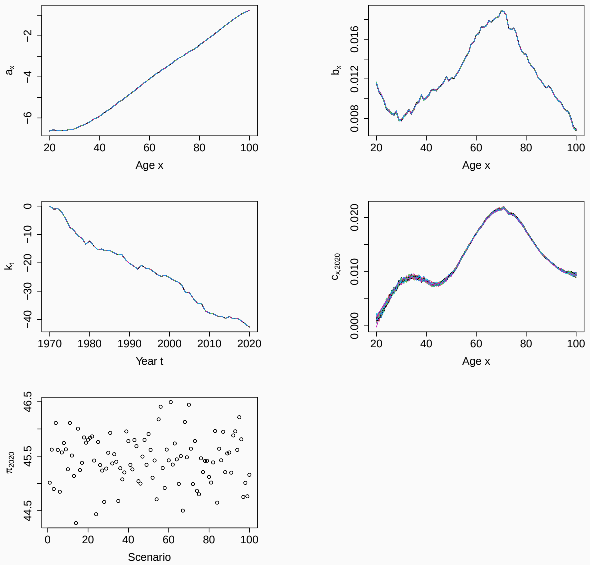

Figure 5: Estimated parameters using 100 sets of imputed values

Parameter Uncertainty – Choice of External Forecasts

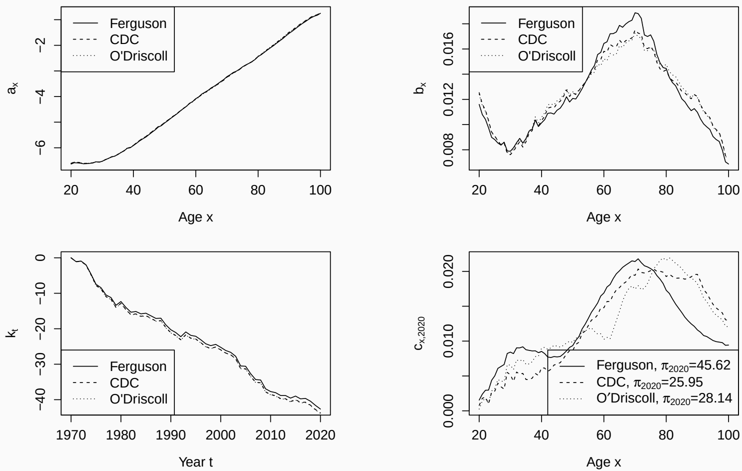

Figure 6: Estimated parameters using different sets of COVID deaths

Using Expert Opinions to Choose the Third Level Parameters

Infection Percentage

- Infection percentage may change significantly with the suppression strategy used
- We consider infection percentages ranging from 10% to 80%
- Expert opinion should be used to predict suppression strategy and the resulting infection percentage

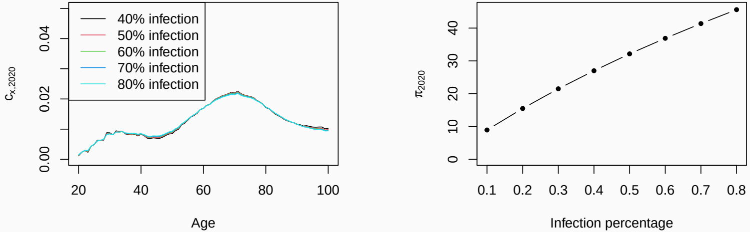

Figure 7: Estimated $c_{x,2020}$ and π_{2020} using various infection percentages

Adjusting $c_{x,t}$ and π_t to Infection Percentage Change

Observations from Figure 7

- $c_{x,2020}$ change slightly with the increase of infection percentage
- π_{2020} increases with the infection percentage almost linearly

When infection percentage changes but IFRs remain the same, we do not need to re-estimate the model

- $c_{x,t}$ remains the same
- π_t can be approximated by a linear interpolation

Pandemic Impact over Multiple Years

Denote the infection rate in year t by F_t

- **Scenario 1**: Excess mortality gradually wears off and mortality reverts back to the long-term mortality pattern

$$F_t = F_{2020}\delta^{t-2020}, 0 < \delta < 1$$

- **Scenario 2**: Excess mortality reduces for several years and then remains at a constant level

$$F_t = F_{2020}\delta^{\min(t-2020,n)}, 0 < \delta < 1$$

- **Scenario 3**: Excess mortality gradually reduces and a permanent downward shift occurs due to positive behavioural changes and improved healthcare practice

$$F_t = F_{2020}0.5^{t-2020}$$
$$d_{x,t}\psi_t = -0.05$$

Pandemic Impact over Multiple Years

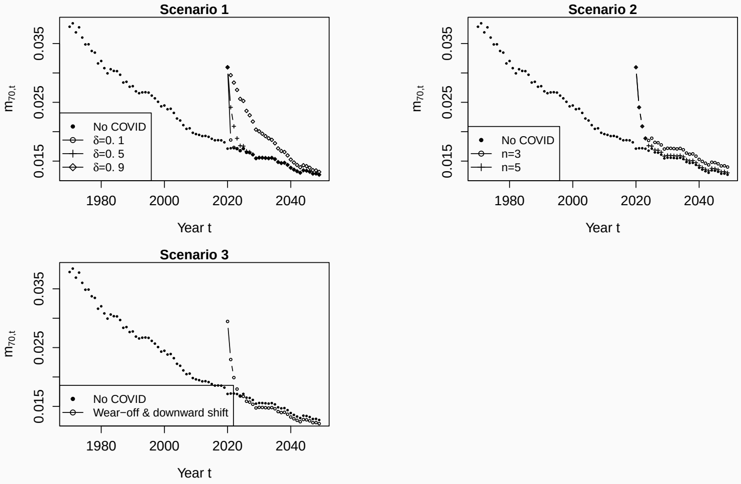

Figure 8: Simulated paths of $m_{70,t}$ using a range of expert opinions

Recurrence of Pandemics in the Future

We use $\lambda = 1/100$ to illustrate the impact of pandemic recurrence on future mortality

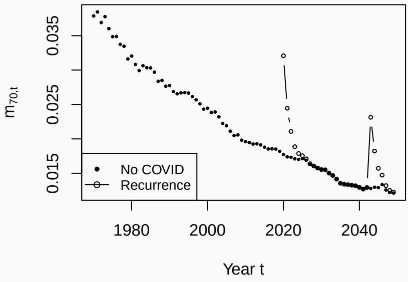

Figure 9: A path of simulated $m_{70,t}$ with recurrence of COVID-alike events

Recurrence of Pandemics in the Future

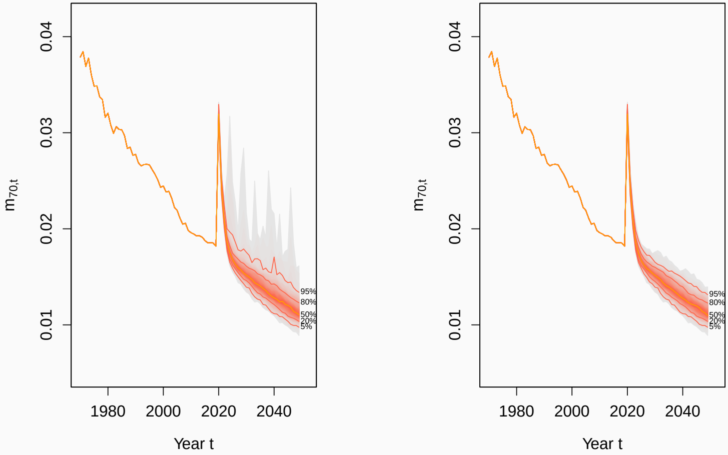

Figure 10: Simulated paths with and without the recurrence of COVID-alike events

Incorporating Parameter Uncertainty in Simulated Mortality Scenarios

Revisiting Parameter Uncertainty

We have observed parameter uncertainty arising from

- Imputations
- Choice of external forecasts (best estimates of IFRs)

The IFRs from the three studies are estimated using COVID death data collected from various sources and also suffer from uncertainty

- Centres for Disease Control and Prevention (2020) and Ferguson et al. (2020) provide confidence intervals for the estimated IFRs
- O'Driscoll et al. (2021) generate 15,000 simulated values for the IFR at each age range

Uncertainty in the Estimated IFRs

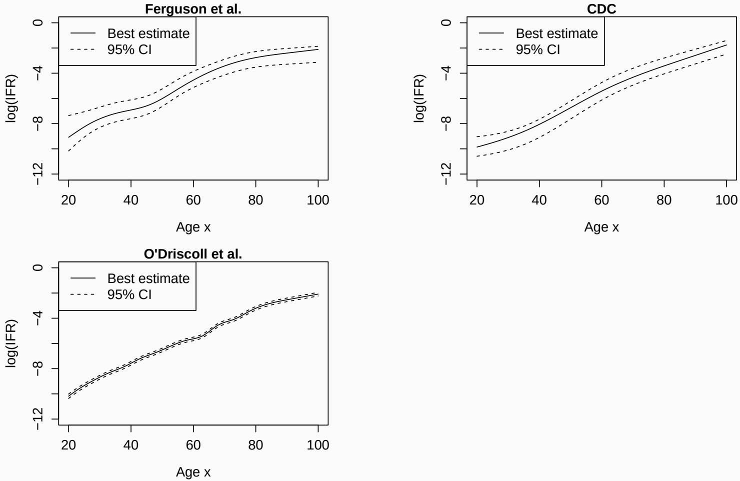

Figure 11: Best estimates and 95% confidence intervals of IFRs from the three studies

Uncertainty in the Estimated IFRs

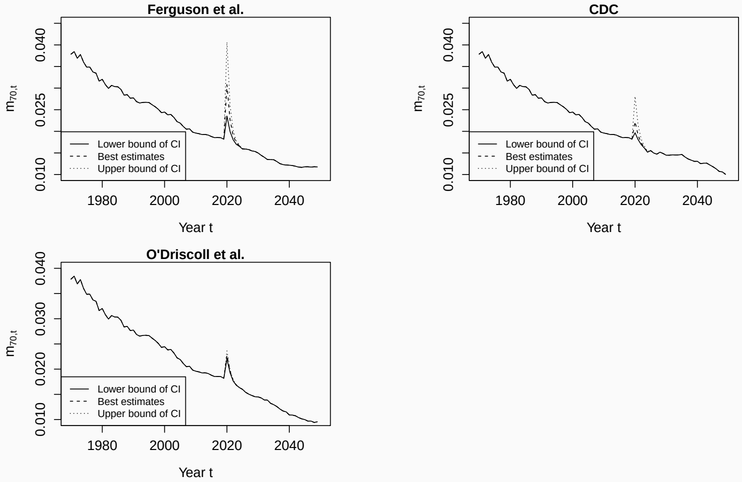

Figure 12: Impact of the uncertainty in the estimated IFRs on the simulated mortality paths

Incorporating Uncertainty into Simulated Mortality Scenarios

Uncertainty arises from k_t, recurrence of pandemic, imputation, and IFR estimates

1. Select a set of IFR estimates and obtain projections of the number of COVID deaths
2. Obtain a set of imputed values for the total number of deaths
3. Estimate the first and second level of parameters in the mortality model by maximizing PQL
4. Choose additional parameters based on expert opinions
5. Simulate a path of k_t and a path for the recurrences of COVID-alike events
6. Determine the corresponding mortality path

Conclusion

Conclusion

- The proposed model uses three levels of parameters to incorporate
  - Long-term mortality improvement pattern extracted from historical mortality data
  - External forecasts of pandemic impact based on quantitative modelling
  - Expert opinions on difficult-to-quantify pandemic impact
- We use single-stage estimation to separate the pandemic impact from changes in the long-term mortality improvement
- Significant uncertainty exhibits in parameter estimates
  - Due to uncertainty in the external forecasts
  - Parameter uncertainty can be included into mortality simulation
- The difference in the three sets of external forecasts highlights the importance of using the most recent information when forecasting mortality during an evolving pandemic

Questions?

References

N. E. Breslow and D. G. Clayton. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88:9 – 25, 1993.

Centres for Disease Control and Prevention. Covid-19 pandemic planning scenarios. 2020. *https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios-archive/planning-scenarios-2020-09-10.pdf* Accessed on 22nd March, 2021.

Hua Chen and Samuel H. Cox. Modeling mortality with jumps: Applications to mortality securitization. *The Journal of Risk and Insurance*, 76(3):727–751, 2009.

N.M. Ferguson, D. Laydon, G Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, and A. Dighe. Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. 2020. URL `https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf`.

Human Mortality Database, 2021. data retrieved from `https://www.mortality.org/` on 22nd March, 2021.

Yanxin Liu and Johnny Siu-Hang Li. The age pattern of transitory mortality jumps and its impact on the pricing of catastrophic mortality bonds. *Insurance: Mathematics and Economics*, 64: 135–150, 2015.

Megan O'Driscoll, Gabriel Ribeiro Dos Santos, Lin Wang, Derek A. T. Cummings, Andrew S. Azman, Juliette Paireau, Arnaud Fontanet, Simon Cauchemez, and Henrik Salje. Age-specific mortality and immunity patterns of sars-cov-2. *Nature*, 590(7844):140–145, 2021.

Rui Zhou, Johnny Siu-Hang Li, and Ken Seng Tan. Pricing standardized mortality securitizations: A two-population model with transitory jump effects. *Journal of Risk and Insurance*, 80(3):733–774, 2013.