Mortality and Longevity

Aging and Retirement

COVID-19 and Excess Mortality: an Actuarial Study

MAY | 2023

LIVING TO 100 SOA INTERNATIONAL SYMPOSIUM

SOA Research INSTITUTE

Long-Term Drivers of Future Mortality

AUTHORS Camille Delbrouck
Aon Benfield, Belgium

Jennifer Alonso-García
Department of Mathematics, Université Libre de Bruxelles, Belgium
RC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Sydney, Australia
Netspar, The Netherlands

SPONSORS Mortality and Longevity Strategic Research Program Steering Committee

Aging and Retirement Strategic Research Program Steering Committee

COVID-19 and excess mortality: an actuarial study

Camille Delbrouck[1], Jennifer Alonso-García[2,3,4*]

[1] Aon Benfield, Belgium
[2] Department of Mathematics, Université Libre de Bruxelles, Belgium
[3] ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Sydney, Australia
[4] Netspar, The Netherlands

October 17, 2022

Abstract

The study of mortality is an ever-active field of research and new methods or combinations of methods are constantly being developed. In the actuarial domain, the study of phenomena disrupting mortality and leading to excess mortality, as in the case of COVID-19, is of great interest. It is therefore relevant to dwelve on the investigation of the extent to which an epidemiological model can be integrated into an actuarial approach in the context of mortality. The aim of this project is to establish a method for the study of excess mortality due to an epidemic and to quantify these effects in the context of the insurance world and to anticipate certain possible financial instabilities. We consider a case study caused by the SARS-CoV-2, in Belgium, during the year 2020. We propose an approach that develops an epidemiological model simulating excess mortality and incorporates this model into a classical approach to pricing life insurance products.

Keywords: Actuarial Science; Epidemiology; Mortality; Excess Mortality; COVID-19; Life Insurance

*This project was done while Camille was a Msc student at ULB. The authors are responsible for any errors. Correspondence to: Jennifer Alonso García, Université Libre de Bruxelles, Department of Mathematics, Faculté des Sciences, Campus de la Plaine - CP 213, Boulevard du Triomphe ACC.2, 1050 Bruxelles, Belgium.

*Corresponding author: jennifer.alonso.garcia@ulb.be

1 Introduction

The recent SARS-CoV health crisis shows clear evidence that we are not immune to pandemics and its economic effects, despite medical advances and robust health systems. Indeed, Institute and Faculty of Actuaries (2015) indicate that the 1918 Spanish Flu, responsible for the death of more than 50 million people (CDC, 2021), had a cost of around £13 billion to insurance companies. Pandemics are also becoming more likely (Saunders-Hastings and Krewski, 2016; Institute and Faculty of Actuaries, 2015), see e.g. the H2N2 flu in 1957-1958, the H3N2 1968 flu, H1N1 2009-2010 flu (WHO, 2021a) or more recently the Ebola epidemic between 2014-2016 (WHO, 2021b).

The increasing frequency and severity call for indepth actuarial analysis of mortality and its effect on pricing and reserving. In particular, pandemics have an heterogenous effect on populations. Indeed, certain socio-economic groups are more susceptible than others to get sick and die, making basic life tables unsuitable for pricing and risk management. It becomes natural to consider a model that incorporates epidemiological insights and dependence (Feng and Garrido, 2011).[1]

The study of mortality is an ever-active field of research and new methods or combinations of methods are constantly being developed. In the actuarial domain, the study of phenomena disrupting mortality and leading to excess mortality, as in the case of COVID-19, is of great interest. Investigating the extent to which an epidemiological model can be integrated into an actuarial approach in the context of mortality is hence relevant. This paper proposes an approach that develops an epidemiological model simulating excess mortality and incorporates this model into a classical approach to pricing life insurance products. The aim of this project is to establish a method for the study of excess mortality due to an epidemic and to quantify these effects in the context of the insurance world and to anticipate certain possible financial instabilities.

In this paper, we consider the epidemic caused by the SARS-CoV-2, in Belgium, during the year 2020. For that, we use the daily COVID mortality data from national health data responsible Sciensano (2021). After capturing the COVID-19 specific mortality using an extended version of Franco (2021), we study the overall mortality, using data from the Human Mortality Database (2021) and the Cairns, Blake and Dowd model (Cairns et al., 2006). We focus on old ages in our study as these were those who were most affected by the pandemics from a mortality viewpoint.

Combining these two models allows us to perform an actuarial study on two life insurance products: whole life insurance and lifetime immediate annuities. We study these products in two settings: an old contract underwritten in 2000, and a recent contract underwritten in 2019, a year before the pandemic. We calculate their expected value and variance. We observe that the variance increases more importantly for recent contracts, despite the insured being younger. For older contracts, with older insured, the variance is reduced. We hypothesize that this could be due to a different distribution of mortalities in the COVID vs no COVID scenario. Indeed, the uncertainty would be greater for younger individuals who will still remain for a long period in the contract. However, for older individuals, the presence of COVID renders the event of death more "certain".

The remainder of this paper is structured as follows. Section 2 presents a review of the literature related to epidemiological models and aggregate mortality models. Section 3 presents

[1]The yearly seasonal flu does not fall within the scope of our study, being a recurrent sickness. Our aim is to model a one-off pandemic caused by a new virus as distinguished by the CDC.

the SIRD epidemiological and Cairns et al. (2006) model used for modeling mortality whereas section 4 provides a numerical estimation and illustration of the effect of COVID-19 in the valuation of insurance products. Section 5 concludes.

2 Literature review

In this section we will discuss various strands of literature related to mortality, in particular cause-of-death, epidemilogical and aggregate mortality models.

Since we are interested in COVID-19 specific mortality, a possible way of studying this phenomenon is by using cause-of-death mortality models. Insights about cause-of-death mortality and its long-term trend can be obtained, informing us about the contributors to aggregate mortality. A crucial choice has to be made with regards to the potential relationship between these causes as these models should also provide insights about aggregate mortality. The challenge is that this dependence is not observable. Indeed, upon death, it is impossible to know what would have been the potential cause-of-death had the individual stayed alive. Hence, a natural choice is to assume independence (see Rogers and Gard (1991); Wilmoth (1996); Tabeau et al. (1999) and recently Boumezoued et al. (2018) and Lyu et al. (2021) for France and The Netherlands respectively). A clear advantage of this approach is that aggregate mortality is simply obtained by adding the mortality per cause-of-death.

However, recent research on cause-of-death mortality incorporates dependence as these would yield better long-term forecasts of the aggregate mortality. Examples of this are Arnold and Sherris (2013) and Arnold and Sherris (2015) who use Vector Error Correlation Models; Zheng and Klein (1995); Li and Lu (2019) and Zittersteyn and Alonso-García (2021) who use copula theory and Li et al. (2019) who utilizes clustering methods to group different causes of death. In all cases, dependence between cause-specific death is considered and better cause-removal and aggregate mortality results are obtained. Despite their richness, these models are unable to capture the particular nature of infectious diseases in the context of a pandemic since transmission and mortality has a different behavior than natural causes of death.

To study COVID-19, we need to move beyond classical actuarial techniques and delve into compartmental models in epidemiology. Various models have been considered for COVID-19, going from SIR (Susceptible- Infectious-Recovered) (Abou-Ismail, 2020; Huang et al., 2020; Calafiore et al., 2020) to SUQC (Susceptible-Unquarantined-Quarantined-Confirmed) (Abou-Ismail, 2020). These models simplify the mathematical modeling of infectious diseases as they divide the population in different compartments, each of them with a different label such as S, I or R depending on whether they are Susceptible, Infectious or Recovered respectively. Individuals can move between compartments and the order of the labels generally represents the transitions between compartments. For instance, SIS means susceptible-infectious-susceptible.[2] These models originated at the beginning of the 20th century in the works of Kermack and McKendrick (1927) and Kermack and McKendrick (1932) and rely on Ordinary Differential Equations (ODE).

The susceptible-infectious-recovered (SIR) model is the basic building block of compartmental models (Hethcote, 2000; Tang et al., 2020). It has been widely used in the context of COVID-19, see e.g. Abou-Ismail (2020); Huang et al. (2020) and Tang et al. (2020). Calafiore et al. (2020), studying the case of Italy, extended this model to include the initial number of

[2]The SIS model is commonly used to model the common cold or the flu as infection does not provide long-term immunity.

susceptible individuals and relative factor between positive cases and real number of infected individuals as a model parameter. Another possible extension is to consider susceptible-exposed-infectious-recovered (SEIR) models. By including the compartment "exposed" we are able to include the incubation period of the sickness (Abou-Ismail, 2020; Tang et al., 2020). Both SIR and SEIR models disregard births and death, implicitely assuming that births and deaths have a negligeable impact in the models. More complex models can overcome this (Hethcote, 2000).

We are interested in analysing excess mortality, as well as the impact of the pandemic in the life table. Ultimately, the goal is to assess the effect of COVID-19 in pricing and reserving. Hence, there is a need to move beyond the SIR model to add at least a compartment related to Death. Fernández-Villaverde and Jones (2022) estimated the SIRD model for various countries, states and cities in the context of COVID-19. Our model will be based on the SIRD compartmental design with the addition of age-stratification as suggested by Balabdaoui and Mohr (2020). Note that more complex models could include quarantine and lockdown dynamics, such as the case of SUQC models described above (Abou-Ismail, 2020). However, we abstract from SEIR or SUQC models as our data is unable to sustain a model with the exposed (E), unquarantined (U) and quarantined (Q) compartments. Furthermore, SUQC models have been mostly used in China (Zhao and Chen, 2020), with strict quarantine rules, which vastly differ from those use in most Western countries. Besides, these models are of less interest within an insurance context as there are no cash-flow payments in case of exposure or quarantine.

Ultimately, after studying the COVID-19 mortality dynamics we are interested in incorporating it into a classical life table, using insights from single-population mortality models. The goal is to assess the effect of COVID-19 in products sold to individuals over 50 years old, as the effect of COVID-19 was the highest for this age segment. We hence choose to model aggregate mortality with a Cairns, Blake and Dowd (CBD) (Cairns et al., 2006) model which is fitted to the Belgian historical mortality using data from the Human Mortality Database (2021). We use this model as it is considered suitable for modelling higher ages and has a simple structure with few parameters. Indeed, empirical analysis show that the changes in mortality rate are imperfectly correlated, supporting the use of the CBD model.

3 Methodology

In this section we develop the theoretical framework for the epidemiological model, the aggregate mortality and their integration. The products and indeces studied to assess the effect of COVID-19 are also presented.

3.1 SIRD model

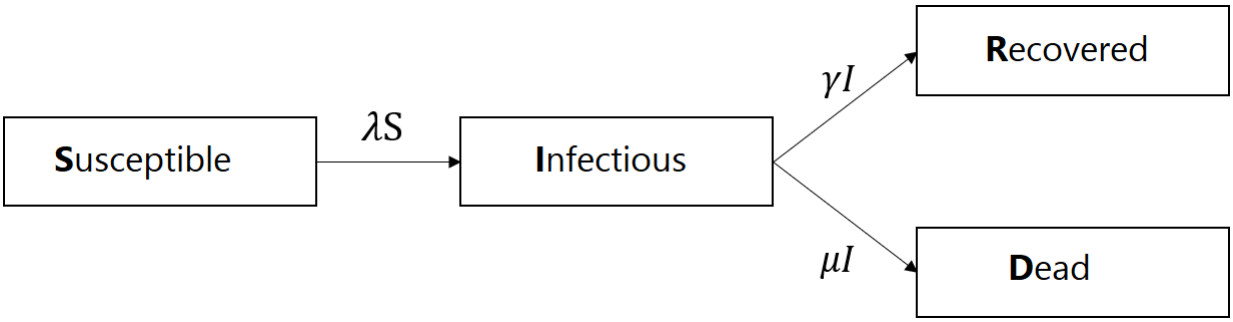
We study the age-stratified susceptible-infectious-recovered-death (SIRD) model as proposed by Balabdaoui and Mohr (2020). Given the differing mortality and recovery rate of infected individuals among different age groups, we incorporate age-stratification with the following age categories:

- Category 1 : Individuals aged 0–24
- Category 2 : Individuals aged 25–44
- Category 3 : Individuals aged 45–64

- Category 4 : Individuals aged 65–74
- Category 5 : Individuals aged 75–84
- Category 6 : Individuals aged 85+

Graphically, the SIRD model is depicted in Figure 1.

Figure 1: SIRD model



Source: the authors.

Integrating the age-component into the SIRD model, and denoting the age-group by "$_i$", we obtain the following EDO system (1) (with $i = 1, .., 6$):

$$\begin{aligned}
\frac{dS_i(t)}{dt} &= -S_i(t)\sum_j C_{i,j}\lambda_i \frac{I_j(t)}{N_j}, \\
\frac{dI_i(t)}{dt} &= S_i(t)\sum_j C_{i,j}\lambda_i \frac{I_j(t)}{N_j} - \gamma_i I_i(t) - \mu_i I_i(t), \\
\frac{dR_i(t)}{dt} &= \gamma_i I_i(t), \\
\frac{dD_i(t)}{dt} &= \mu_i I_i(t).
\end{aligned} \tag{1}$$

The rate λ_i represents the transmission rate of the infection for group i, whereas μ_i and γ_i represent the mortality and recovery rate of group i respectively. Clearly the system (1) indicates that you leave the susceptible compartment if infected, you enter the infectious state by infection and leave either when you recover or die. You enter the recovery state with intensity γ_i and do not leave this compartment.[3] Similarly, you enter the dead absorbant compartment with intensity μ_i. The total population in group i is given by

$$N_i = S_i(t) + I_i(t) + R_i(t) + D_i(t) \quad \forall i.$$

The matrix C represents social contacts between the different age categories. It is important to note that we study the epidemic in a short period, making transitions between age-categories redundant. The initial conditions of the EDO system (1) are given by:

$$\begin{aligned}
S_i(0) > 0, I_i(0) > 0, R_i(0) \geq 0, D_i(0) \geq 0 \\
S_i(0) + I_i(0) + R_i(0) + D_i(0) = N_i
\end{aligned} \tag{2}$$

[3]The *recovered* state is absorbent. This is a reasonable hypothesis in 2020, year on which the model is parametrized. COVID infection was assumed to provide long-term immunity. Models that allow for recovery could address this but are outside the scope of this study due to data limitations.

Alternatively, we can represent the EDO system (1) for all age-groups in matrix form as follows:

$$\begin{aligned}
\frac{dS}{dt} &= -\text{diag}\left(C\left[N_{Inv}I\right]S^T\right)\lambda, \\
\frac{dI}{dt} &= -\text{diag}\left(C\left[N_{Inv}I\right]S^T\right)\lambda - \gamma I - \mu I, \\
\frac{dR}{dt} &= \gamma I, \\
\frac{dD}{dt} &= \mu I,
\end{aligned} \tag{3}$$

where

$$\begin{aligned}
-\ S^T &= [S_1, ..., S_6] \\
-\ I^T &= [I_1, ..., I_6] \\
-\ R^T &= [R_1, ..., R_6] \\
-\ D^T &= [D_1, ..., D_6] \\
-\ C &= \begin{bmatrix} C_{1,1} & \cdots & C_{1,6} \\ \vdots & \ddots & \vdots \\ C_{6,1} & \cdots & C_{6,6} \end{bmatrix} \\
-\ N_{Inv} &= \begin{bmatrix} 1/N_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1/N_n \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
-\ \lambda^T &= [\lambda_1, ..., \lambda_6] \\
-\ \gamma &= \begin{bmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \gamma_6 \end{bmatrix} \\
-\ \mu &= \begin{bmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \mu_6 \end{bmatrix}.
\end{aligned}$$

The reproduction rate R_0, used to analyse the impulse of the pandemic, can be calculated as the maximum eigenvalue from the "next-generation" matrix (Franco, 2021), resulting in (4):

$$R_0 = \text{maxeigh}\left[\frac{\lambda}{\gamma + \mu} \cdot C_{i,j}\right]_{i,j} \tag{4}$$

Model identification The SIRD model defined by the EDO system (1) is solved using multistep methods as described in Hindmarsh (1983). Numerical multistep methods start with an initial point and start moving forward. These methods use past points and their derivatives in order to gain in efficiency (Hindmarsh, 1983).[4] Applying this method to (1) we obtain the number of people in each compartment, including the main compartment of interest $D_i(t)$ which represents the number of deaths. It does so for every time-step t which is equal to 1 day in our case.

The number of deaths according to the solution of the ODE system $D_i(t_{=2020})$ allows us to find the empirical mortality rate per age category using maximum likelihood arguments (5):

$$\widehat{m}_i(t_{=2020}) = \frac{D_i(t_{=2020})}{\widehat{ETR}_i(t_{=2020})} \tag{5}$$

[4]This is in contrast with one-period methodologies such as the Euler method which solely refers to a previous point and its derivative to determine the actual value, or the Runge-Kutta method which use a few intermediate points but reject all previous to obtain the higher order value.

where $m_i(t_{=2020})$, $D_i(t_{=2020})$ and $ETR_i(t_{=2020})$ are the central mortality rate, number of deaths and exposure to risk for year $t = 2020$ and age category $i = 1, ..., 6$ respectively. The exposure to risk for age-category i is estimated using $ETR_x(t)$ for the global population as follows:

$$ETR_i(t_{=2020}) = \sum_{x=i_1}^{i_N} ETR_x(t). \tag{6}$$

where i_1 and i_N correspond to the first and last age belonging to age-category i.

3.2 Cairns-Blake-Dowd model

We use Cairns et al. (2006) to model global mortality and its long-term trend for old-ages as these age categories have been most afected by COVID-19. The model is presented as a logistic regression (7):

$$\text{logit} q_{xt} = \kappa_t^{(1)} + (x - \overline{x})\kappa_t^{(2)} \tag{7}$$

where

- $\kappa_t^{(1)}$ (intercept): represents the global mortality trend and is generally a decreasing parameter since it improves over time.
- $\kappa_t^{(2)}$ (slope): represents the mortality improvements and has typically a positive slope, indicating that improvements have been greater at the first part of the age period considered.

We assume that the force of mortality is constant within each square of the Lexis diagram:

$$\mu_{x+\eta_1}(t + \eta_2) = \mu_x(t) \text{ for } 0 < \eta_1, \eta_2 < 1. \tag{8}$$

This allows to obtain the equivalence between the force of mortality $\mu_x(t)$ and the central mortality rate $m_x(t)$. We work within a Poisson framework with $D_x(t) \sim \text{Poi}(ETR_x(t)\mu_x(t))$ for $\mu_x(t)$ given by (9):

$$\mu_x(t) = -\ln(1 - q_x(t)) = \ln\left(1 + \exp\left(\kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})\right)\right). \tag{9}$$

This yields to the following log-likelihood:

$$\ln(\mathcal{L}(\kappa)) = \sum_{x=x_1}^{x_m} \sum_{t=t_1}^{t_n} \{D - d_{xt} \ln \mu_x(t) - ETR_x(t)\mu_x(t)\} + \text{constant} \tag{10}$$

where $\mu_x(t)$ should be replaced by (9). The model will be solved using the **R** package **StMoMo** (Villegas et al., 2018).[5] Goodness of fit is assessed through Pearson residuals, which should ideally have no trend. These are calculated as

$$r_{xt} = \frac{ln\frac{\hat{q}_x(t)}{\hat{p}_x(t)} - \kappa_t^{(1)} + \hat{\kappa}_t^{(2)}(x - \overline{x})}{\sqrt{\left((x_m - x_1)(t_n - t_1)\sum_{x=x_1}^{x_m}\sum_{t=t_1}^{t_n}(\hat{\epsilon}_x(t))^2\right)^{-1}}}. \tag{11}$$

[5]This package relies on generalized linear models and uses the package **gnm** to solve for numerous stochastic mortality models that can be expressed within a GLM framework. The algorithm follows two steps. Firstly, nonlinear parameters are updated, then the linear parameters are. Secondly, all parameters are updated jointly until convergence is attained.

Finally, projection is performed whereby the two stochastic processes $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are modeled through a bivariate random work with drift:

$$\begin{cases} \kappa_t^{(1)} &= \kappa_{t-1}^{(1)} + d_1 + \xi_t^{(1)} \\ \kappa_t^{(2)} &= \kappa_{t-1}^{(2)} + d_2 + \xi_t^{(2)} \end{cases} \tag{12}$$

where d_1 and d_2 correspond to the drift, and $\xi_t^{(1,2)}$ are independent bivariate and normally distributed parameters with zero mean and variance-covariance matrix $\begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 \\ \sigma_{21}^2 & \sigma_2^2 \end{bmatrix}$. We rely on closure of tables techniques to project to the ultimate age ω. We use the simple logistic regression of Thatcher (1999):

$$\mu_x(s) = \frac{\phi_{s,1} \exp(\phi_{s,2}.x)}{1 + \phi_{s,1} \exp(\phi_{s,2}.x)} \tag{13}$$

This model is calibrated as a linear logistic regression (14) for a carefully chosen age range $\tilde{x} = x_1, ..., x_{end}$:

$$\text{logit}(\mu_x(s)) = \ln(\phi_{s,1}) + \phi_{s,2} x. \tag{14}$$

with $\text{logit}(y) = \ln \frac{y}{1-y}$. Having $\phi_{s,1}$ and $\phi_{s,2}$, mortality for higher ages can be projected beyond x_{end} using (13).

3.3 Final model

Having the COVID-19 related deaths stemming from the epidemiological model and the general mortality model, we need to merge the results. We rely on cause-of-death mortality models techniques to aggregate *general* deaths to *COVID-19* deaths. We will assess the impact of our modeling through the study of the cohort life expectancy under the expression (8) given by:

$$\begin{aligned} e_x(t) &= E[T_x(t)] \\ &= \frac{1 - \exp(-\mu_x(t))}{\mu_x(t)} + \sum_{k \geq 1} \exp\left(- \sum_{j=0}^{k-1} \mu_{x+j}(t+j) \right) \frac{1 - \exp(-\mu_{x+k}(t+k))}{\mu_{x+k}(t+k)} \end{aligned} \tag{15}$$

where x and t indicate the age and year of study respectively. Comparing mortality rates that include or exclude COVID-19 mortality allows us to assess the impact of a part of the excess mortality due to the pandemic. Indeed, social isolation and health care system saturation has had an adverse impact on both COVID-19 and other patient's ability to receive quality and timely care. The total excess mortality is hence at this stage unknown.

3.4 Actuarial application

We assess the effect of COVID for two insurance products (whole life insurance and a monthly life annuity) and two scenarios: a contract underwritten in 2019, a year before the pandemic started, and 2000. For the sake of completeness, we provide the standard actuarial expressions for these products.

3.4.1 Whole life insurance

The Net Present Value (NPV) for a whole insurance paying C in case of death is given by

$$NPV = C \cdot A_n, \tag{16}$$

where A_n is calculated using the recurrence relationship

$$\begin{aligned} A_n &= vq_n + vp_n A_{n+1}, \\ A_{\omega-1} &= v, \end{aligned} \tag{17}$$

with

$$\begin{aligned} v &= \frac{1}{1+i}, \\ q_n(t) &= 1 - p_n(t), \\ p_n(t) &= e^{-\mu_n(t)}. \end{aligned} \tag{18}$$

The variance is given by

$$\text{Variance} = C^2 \cdot ({}^2A_n - A_n^2), \tag{19}$$

with

$$\begin{aligned} {}^2A_n &= v^2 q_n + v^2 p_n \, {}^2A_{n+1}, \\ {}^2A_{\omega-1} &= v^2. \end{aligned} \tag{20}$$

3.4.2 Life annuity

The NPV of a life annuity paying P per month is given by

$$NPV = 12 \cdot P \cdot a_n^{(12)}, \tag{21}$$

where $a_n^{(12)}$ can be expressed in terms of $A_n^{(12)}$ as follows:

$$a_n^{(12)} = \frac{1 - u^{(12)} A_n^{(12)}}{i^{(12)}}, \tag{22}$$

with

$$\begin{aligned} i^{(12)} &= ((e^{\delta})^{1/12} - 1) \cdot 12, \\ u^{(12)} &= 1 + i^{(12)}. \end{aligned}$$

We obtain the monthly whole life insurance equivalent $A_n^{(12)}$ using A_n by approximating:

$$A_n^{(12)} \approx \frac{i}{i^{(12)}} A_n. \tag{23}$$

yielding

$$NPV \approx 12P \frac{1 - u^{(12)} \frac{i}{i^{(12)}} A_n}{i^{(12)}}. \tag{24}$$

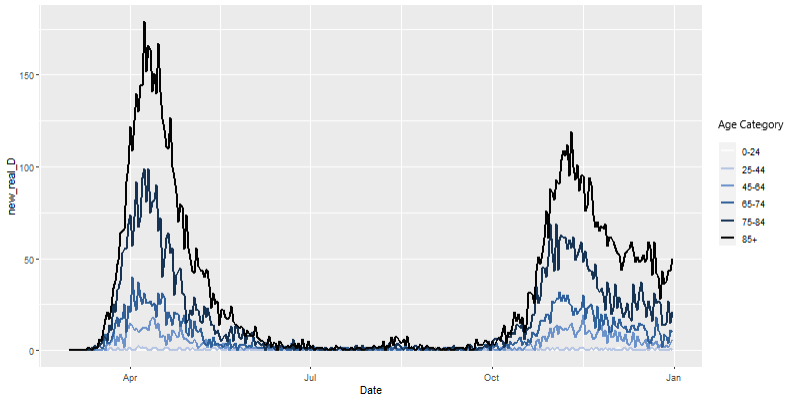
The variance is given in an analogous manner as

$$\text{Variance} = (12P)^2 \frac{{}^2A_n^{(12)} - (A_n^{(12)})^2}{(1 - v^{(12)})^2}, \tag{25}$$

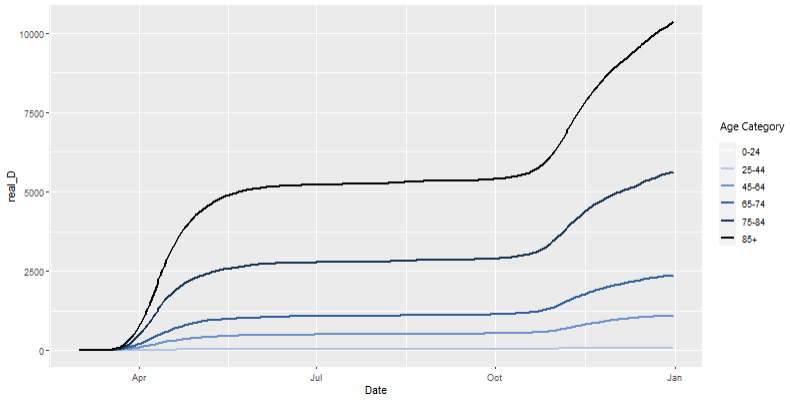
where

$${}^2A_n^{(12)} \approx \frac{i}{i^{(12)}} \, {}^2A_n. \tag{26}$$

Figure 2: Raw Belgian COVID-19 death data from Sciensano (2021) (01/03/2020-31/12/2020)



(a) Daily COVID-19 deaths



(b) Accumulated COVID-19 deaths

Source: the authors.

4 Numerical implementation

This section describes the database from Sciensano (2021) in detail and presents the epidemiological and global mortality model. The actuarial application and its analysis follows.

4.1 Database

Our model uses demographic and epidemiological data for the case of Belgium. The COVID-19 related data is given by the *Epistat* online platform from Sciensano (2021). Scienscano is the Belgian health institute and is responsible for following-up the epidemiological evolution of COVID-19. The database provides segmentation of deaths per age and sex.

Figure 2a and 2b depicts daily and cumulated COVID-19 deaths respectively. The two main waves appear clearly in the figures. Furthermore, the particular virulence to old ages is clear as most deaths belong to the 75-84 and 85+ age categories. We use Human Mortality Database (2021)[6] database for the global mortality. Our study is reduced to the year 2020.

[6]Human Mortality Database was created to provide detailed data about population and mortality to researchers, students, journalists, political analysist and individuals interested by the history of human longevity.

4.2 Epidemiological model

4.2.1 Parameters

To simulate the model, we need to identify the following parameters:

- the initial conditions: N, $I_i(0)$, $D_i(0)$;
- the social-contact matrix C;
- the COVID-19 related parameters: λ, γ, μ and R_0.

Some parameters are based on assumptions as they represent demographics or the biological and medical nature of the problem and are given by scientific articles. Others will be found by optimisation.

We use *Root Mean Square Error* (RMSE) to assess the goodness of our model. In particular we compare raw with estimated daily deaths :

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T}\sum_{i=1}^{6}(\text{daily\_}D_i(t) - \text{daily\_}D_i\text{\_real}(t))^2}{6 \cdot T}}, \tag{27}$$

with

- $t = 1$ corresponds to March 1st 2020 and T corresponds to either $T = 244$ to October 31st 2020 for the *preliminary model*[7] and $T = 306$ to December 31st 2020 for the *model with delay* and *final model*,
- daily_$D_i(t)$ is the daily number of deaths predicted by the model for day t and age-group i given by daily_$D_i(t) = D_i(t) - D_i(t-1)$
- daily_D_i_real(t) is the real number of datily deaths for time t and age-group i.

We assess our models graphically by comparing realised deaths, cumulated deaths as well as death surges with modeled ones. We study three models:

1. *Preliminary model*: this model relies on γ_i values that coincide between the two waves and R_0 values that are drawn from the literature. The parameters are based on the study of Franco (2021) who studies the pandemic until 31 october 2020,
2. *Model with delay*: this model relies on γ_i values that coincide between the two waves but uses R_0 that are estimated for our database,
3. *Final model*: this model relies on γ_i values that differ betwee two waves and uses estimated R_0 that are finetuned to the differences observed between the two waves.

The *model with delay* and *final model* rely on data for the whole pandemic period from 01-03-2020 until 31-12-2020. In what follows, we present the parameters that are used in our model and specify where needed whether these parameters are relevant for all models or not.

- $\boldsymbol{N}$: the belgian population as of January 1^{st} 2020 per age category is given by Table 1:

[7] The last weeks of the year are excluded in the *preliminary model* as the parameters are based on the study from Franco (2021) who studied the period until 31 October 2020.

Table 1: N_i per age category

	0-24	25-44	45-64	65-74	75-84	85+
N_i	3 237 498	2 968 631	3 082 034	1 170 399	698 940	335 139

Source: Statbel (2021).

- $\boldsymbol{I_i(0)}$: 19 total cases were reported March 1st 2020 according to Sciensano (2021). The age-decomposition of the initial infections are presented in Table 2:

Table 2: Number of COVID-19 infected individuals on 01/03/2020

	0-24	25-44	45-64	65-74	75-84	85+
$I_i(0)$	5	2	10	1	1	0

Source: Sciensano (2021).

- $\boldsymbol{D_i(0)}$: No deaths were reported until March 2nd 2020 (Sciensano, 2021).
- $\boldsymbol{C_{i,j}}$: the social contact matrix is based on the Socrates tool from Willem et al. (2020).
- $\boldsymbol{R_0}$: we use a time-dependent[8] R_0 from Franco (2021) as given in Table 3 for the *preliminary model*. For the *model with delay* and *final model* is calculated R_0 based on equation (4).

Table 3: R_0 values per period (format DD/MM) in 2020 and their confidence intervals ([ ])

	01/03 - 13/03	14/03 - 18/03	19/03 - 03/05	04/05 - 07/06
R_0	4.13 [3.89 ; 4.39]	2.24 [2.13;2.35]	0.65 [0.61 ; 0.72]	0.79 [0.75 ; 0.83]
	08/06 - 30/06	01/07 - 28/07	29/07 - 31/08	01/09 - 31/10
R_0	0.99 [0.91 ; 1.07]	1.40 [1.29 ; 1.53]	0.75 [0.63 ; 0.88]	1.73 [1.62 ; 1.85]

Source: Franco (2021).

- $\boldsymbol{\lambda}$: λ is determined using R_0 according to the following equation:

$$\lambda_i = \frac{R_0}{\text{maxeig}(\frac{C_{i,j}}{\gamma_i+\mu_i})} \quad \text{avec} \quad i = 1, ..., 6. \tag{28}$$

- $\boldsymbol{\gamma}$: it is a time-dependent parameter depending on the epidemiological stage. We use the values from Franco (2021):

[8] It is time-dependent as lockdown and quarantine rules have changed according to the evolution of the pandemic.

Table 4: γ values per age-category

	0-24	25-44	45-64	65-74	75-84	85+
γ_i	4.5294^{-1}	5.0786^{-1}	5.7858^{-1}	8.01^{-1}	9.0512^{-1}	17.76^{-1}

Source: Franco (2021).

- $\boldsymbol{\mu}$: we use the age and period dependent mortality rates from Franco (2021) as depicted in Table 5[9]

Table 5: μ values per age category (in %)

March-April	0-24	25-44	45-64	65-74	75+
μ_i	0.0	0.02	0.21	1.85	9.25
April-July	0-24	25-44	45-64	65-74	75+
μ_i	0.0	0.01	0.19	1.72	7.84
July-	0-24	25-44	45-64	65-74	75+
μ_i	0.0	0.01	0.08	0.86	1.89

Source: Franco (2021).

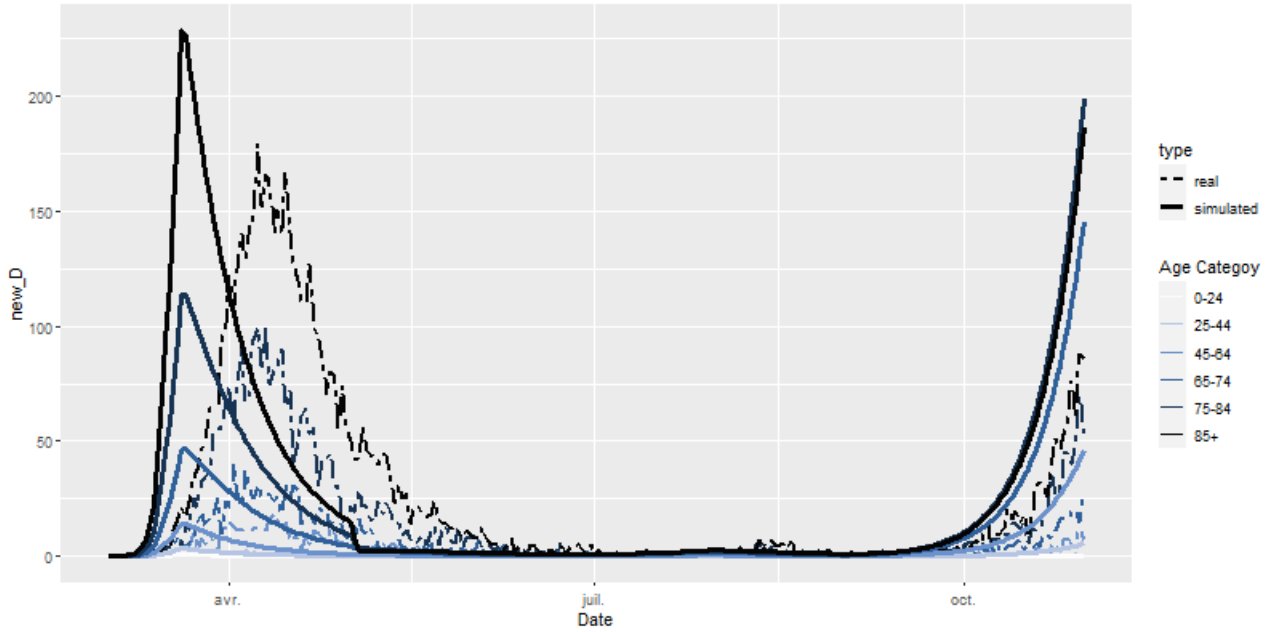
4.2.2 Preliminary model

As discussed in Section 3.1, we solve the EDO system (1) using a multi-step methodology. We use the **R** package **deSolve** for this. The model is used for the period between March 1st 2020 and October 31st 2020 for the reasons explained earlier in Section 4.2.1. Using the R_0 values from Table 3 from Franco (2021), an adjusted initial number of infected individuals I_0=6 for the 25-44 age category, we obtain an RMSE of 402.3329.

Figure 3 depicts the difference between the simulated and actual daily deaths. Two obvious trends appear: death is over-estimated in this model and there is a clear delay in the pandemic peaks. It becomes obvious that relying in pre-specified R_0 values is not providing satisfactory results, hence the need to calculate it ourselves.

[9]Mortality rates for Belgium were studied in various studies (Levin et al., 2020; Molenberghs et al., 2020). The meta-study from Levin et al. (2020) finds the relationship $\log_{10}(IFR) = -3.27 + 0.0524 \cdot age$. However, these results are not wave-dependent.

Figure 3: Daily observed vs simulated COVID-19 related deaths according to the *preliminary model* (01/03/2020-31/10/2020)



Source: the authors.

4.2.3 R_0 parameter calculation

The need to have a more granular study of R_0 appeared in the previous study. We calculate the value of R_0 value for shorter time intervals corresponding to changes in policy to fight the pandemic. For instance, we separate the period of 19/03-03/05 and 04/05-07/06 in 19/03-03/04 for the full lockdown and 04/04-07/06 for Phase 1 and 2 of the post-lockdown period, better reflecting restrictions. Table 6 provides a summary of the periods considered.

Table 6: Division of periods for R_0

Period	Level of restrictions
01/03/2020 - 13/03/2020	Pre-lockdown
14/03/2020 - 18/03/2020	Schools and leisure closed
19/03/2020 - 03/04/2020	Full lockdown
04/04/2020 - 07/06/2020	Phase 1-2
08/06/2020 - 30/06/2020	Phase 3
01/07/2020 - 28/07/2020	Phase 4
29/07/2020 - 31/08/2020	Phase 4 bis
01/09/2020 - 05/10/2020	Second wave
06/10/2020 - 18/10/2020	Limited social contacts
19/10/2020 - 01/11/2020	Curfew
02/11/2020 - 31/11/2020	(light) Lockdown
01/12/2020 - 23/12/2020	Reopening of shops
24/12/2020 - 31/12/2020	Public holiday period

Source: the authors.

Firstly, we calculate the R_0 parameter according to the period division proposed in Table 6. Given the time intensive procedure, this is done by parallel computing in **R**. The models obtained are compared with regards to their RMSE. We separate the first and second wave in the calculation of R_0. Table 7 highlights the computational time for an array of parameter sets and number of cores used in parallel computing.

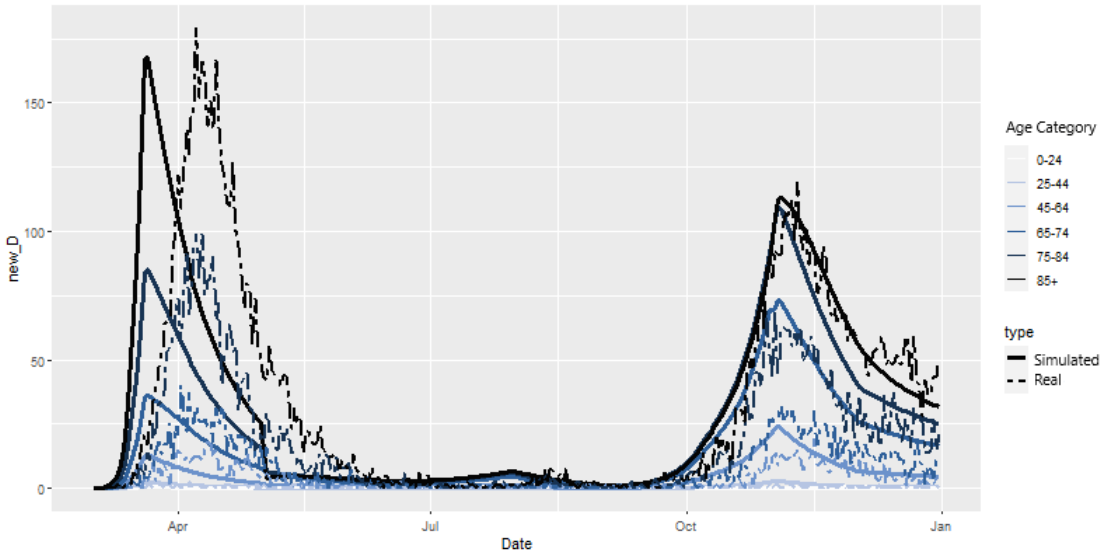
Table 7: R_0 identification benchmarking

		# parameter set			
		6	10	15360	94527
# de cores	1	14.00s	26.21s	NA	NA
	3	5.63s	15.26s	NA	NA
	8	2.92s	5.17s	1h39m50.09s	10h14m23.06s

Source: the authors.

The best model according to the RMSE calculation yields the model depicted in Figure 4. There is a clear delay between the peaks and valleys of the simulated model versus reality. Indeed, our simulated peak is in 2020-03-20 whereas the real peak is observed 2020-04-08, indicating a 18 day delay. We denote this model the *model with delay*.

Figure 4: Daily observed vs simulated COVID-19 related deaths according to the *model with delay* (01/03/2020-31/12/2020)



Source: the authors.

A priori, multiple reasons could be raised in order to understand this delay. A first explanation lies in the recovery rate γ. This empirically varies between the two waves, whereas our *preliminary model* considered the same γ for the whole period of 2020. We will hence use a new set of parameters for the recovery, assuming a longer recovery period during the

first wave compared to the second wave. The new parameters are given in Table 8. A second explanation lies in the calculation of R_0. This parameter changes at the beginning of the pandemic in a much quicker way than it does during the second wave. Utilizing the same time step for the calculation of R_0 may seem misleading. A weekly step instead is then adopted from 01-03 to 09-04.

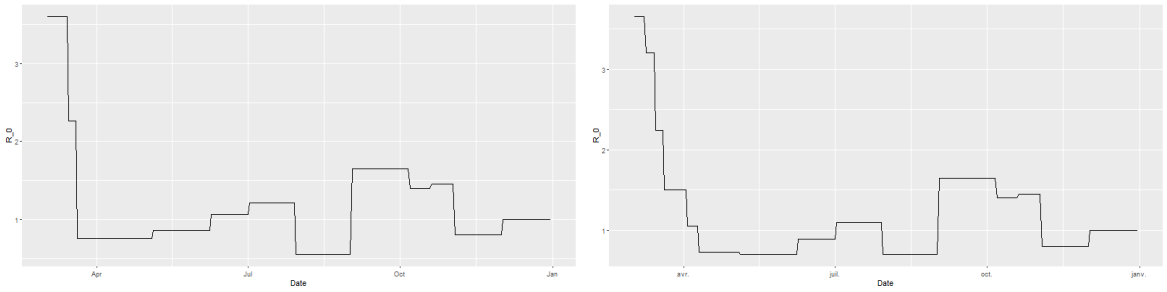
Table 8: γ values per age-category

First wave	0-24	25-44	45-64	65-74	75-84	85+
γ_i	5.163516^{-1}	5.789604^{-1}	6.595812^{-1}	8.52948^{-1}	10.152384^{-1}	20.2464^{-1}
Second wave	**0-24**	**25-44**	**45-64**	**65-74**	**75-84**	**85+**
γ_i	4.5294^{-1}	5.0786^{-1}	5.7858^{-1}	8.01^{-1}	9.0512^{-1}	17.76^{-1}

Source: The first wave values correspond to the second wave values increased by 14% based on the slower recovery during the first wave as seen in Franco (2021).

We observe in Figure 5a and 5b the R_0 parameter for the *model with delay* and *final model* respectively. We clearly see the need for a finer grid in order to replicate the rapid evolution of R_0 at the beginning of the pandemic[10]. The former grid yielded strong jumps.

Figure 5: R_0 values



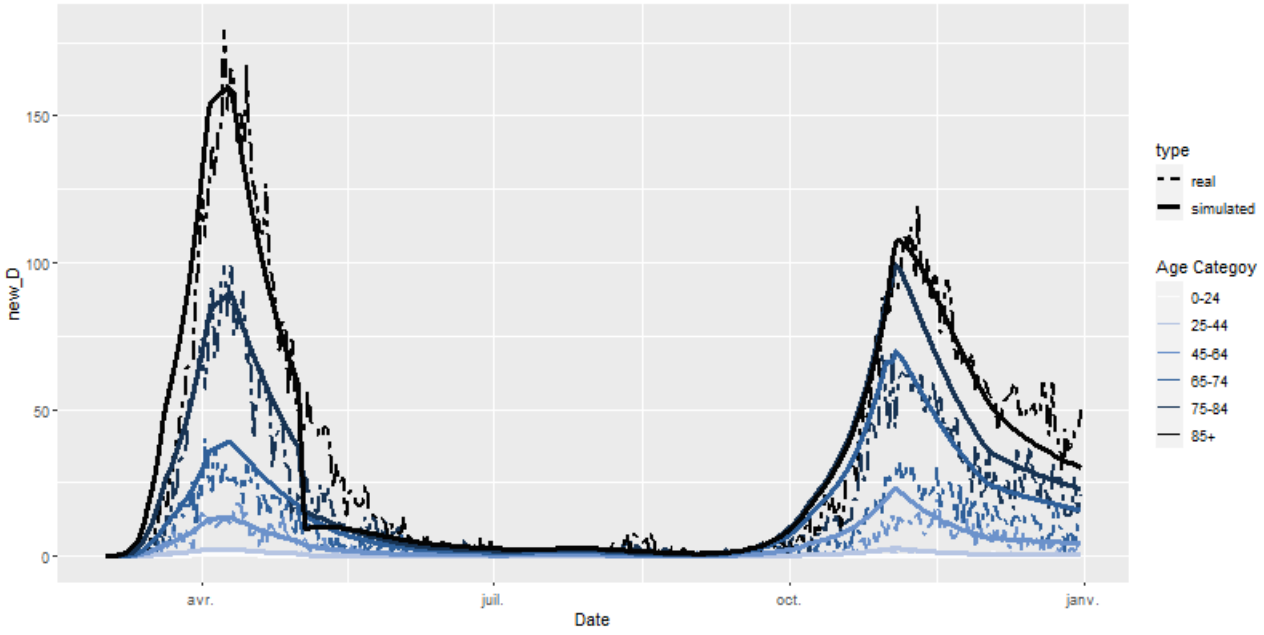
(a) R_0 in the *model with delay* (b) R_0 in the *final model*

Source: the authors.

Of course, we could consider working on a finer grid for the whole duration of the study. However, computational time would be greatly affected. Using a weekly time step for the whole study would yield to a number of parameter sets equal to 2377728 which would take 10 days and 17 hours to fit. Furthermore, we observe in Figure 6 that the enhanced model fits the raw data reasonably well. Indeed, incorporating the new R_0 calculation and recovery rates yields the first wave peak to 08/04/2020, coinciding with the raw data. Hence, we reduce the timestep to a weekly basis for the beginning of the first wave whilst keeping the longer timestep for the remainder of the analysis. We denote this model as *final model*.

[10]This corresponds to the following intervals in the *final model*: 1/3/2020, 8/3/2020, 14/03/2020, 19/03/2020, 26/03/2020, 2/4/2020, 9/4/2020, 4/5/2020, 8/6/2020, 1/7/2020, 29/07/2020, 1/9/2020, 6/10/2020, 19/10/2020, 2/11/2020, 1/12/2020, 24/12/2020 and 31/12/2020.

Figure 6: Daily observed vs simulated COVID-19 related deaths according to the *final model* (01/03/2020-31/12/2020)



Source: the authors.

Finally, we compare the three models, *preliminary model*, *model with delay* and *final model*, with regards to their RMSE (Table 9). We distinguish two time periods, Period 1 going until 31/10/2020 and period 2 going until 31/12/2020. Of course, the *preliminary model* does not have a value for period 2 as the R_0 values of Franco (2021) were only available until 31/10/2020 limiting our preliminary study to that period. It is clear that our *final model* outperforms the others in terms of RMSE.

Table 9: RMSE for the different models

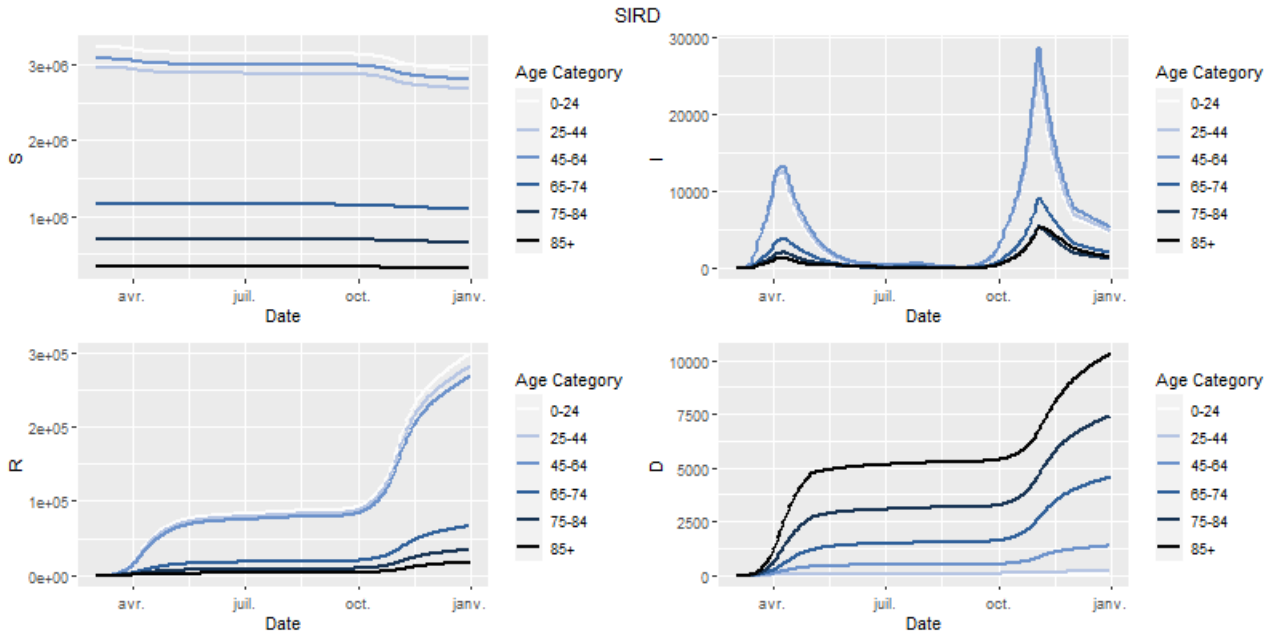
Model	Period 1	Period 2
Preliminary model	402.3329	NA
Model with delay	393.6614	543.0819
Final model	239.5903	525.7816

Source: the authors.

4.2.4 Epidemiological model: results

Figure 7 depicts the evolution of **S**usceptible, **I**nfectious, **R**ecovered et **D**ead as a solution of the *final model*.

Figure 7: Final model compartments



Source: the authors.

The **I**nfectious graph shows clearly the two waves of the epidemic. Despite the greatest number of sick individuals during the second waves, deaths were more numerous during the first wave. Our solution also allows us to calculate the empirical mortality rate. These empirical mortality rates, merged with the global mortality, will allow us to assess the effect of excess of mortality within a life insurance product context.

The empirical mortality rates are calculated using the number of deaths per age-category given by the model together with the estimated $ETR_i(t_{=2020})$ given in equation (6). Exposure data being not available for 2020, we approximate it through $ETR_{xt_{=2018}}$. We obtain the results in Table 10.

Table 10: Empirical mortality rate per age-category (en %)

	0-24	25-44	45-64	65-74	75-84	85+
μ_i	0	0.005	0.039	0.4	1.081	3.39

Source: the authors.

4.3 Cairns-Blake-Dowd model

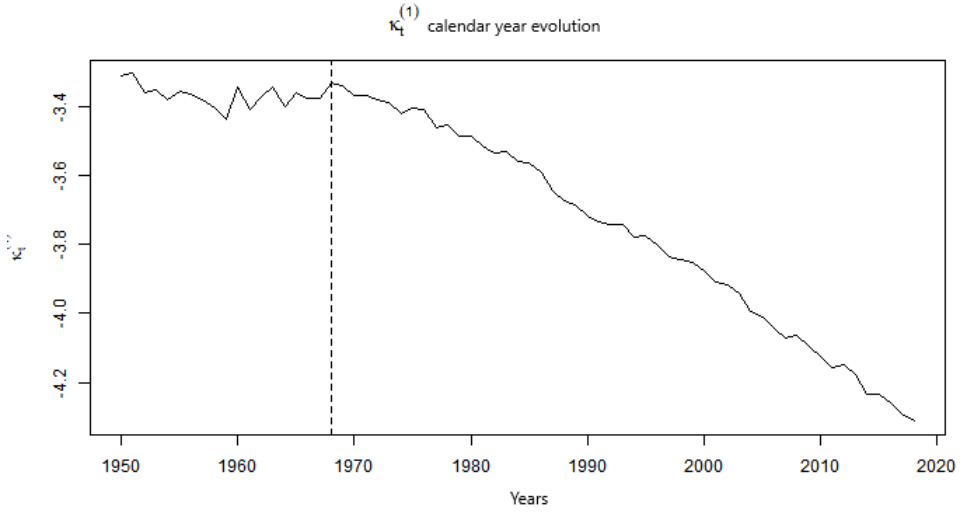
We need the base mortality in order to be able to assess excess mortality. Here we focus on the fit of the Cairns et al. (2006) model for Belgium using Human Mortality Database (2021).

4.3.1 Estimation

The theoretical methodological procedure is detailed in Section 3.2. We calibrate the model from 1968 to 2018, last year available under Human Mortality Database (2021), using the

StMoMo package in **R**.[11] This period is chosen as a linear trend becomes clear from 1968 as shown in Figure 8. We consider the age period 45 to 100, excluding the accident hump that would be otherwise poorly fitted by this model. Furthermore, we are solely interested in the effect of COVID-19 and its insurance products for adults and old ages since COVID-19 mortality rates are negligible under the age of 44 (Table 10).

Figure 8: Choice of time period

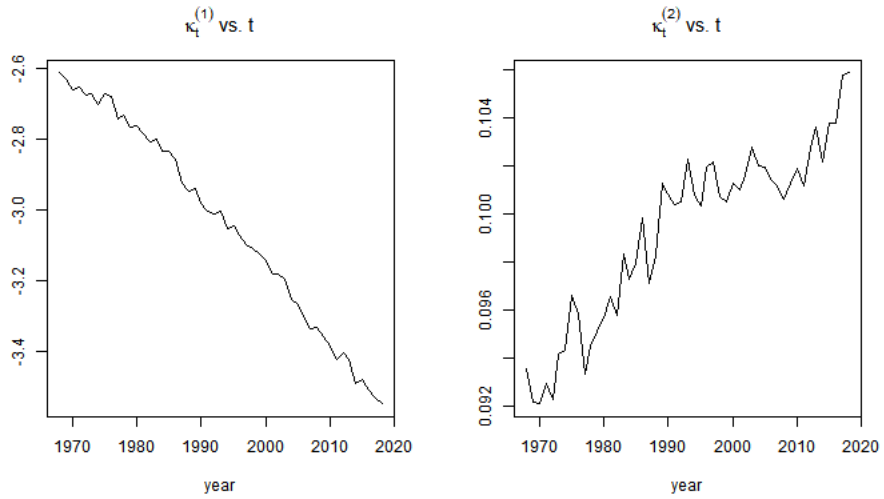


Source: the authors.

Figure 9 shows the estimated parameters for the age and period interval chosen. The trends observed coincide with expectation, $\kappa_t^{(1)}$ decreases over time reflecting mortality improvements whereas $\kappa_t^{(2)}$ increases indicating improvements have been more important under $\bar{x}$ than beyond the mean age considered in our study.

[11] Missing values, as wel as NA, are associated zero weight and are hence not included in the fit.

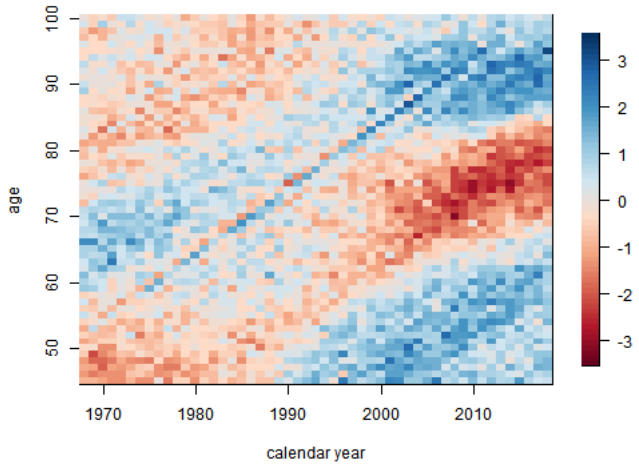
Figure 9: CBD fit



Source: the authors.

Residuals are calculated as in equation 11 and are depicted in figure 10. We observe no clear trends beyond a reduced cohort effect for the individiuals born in 1920 after the first world war.

Figure 10: Residual heatmap



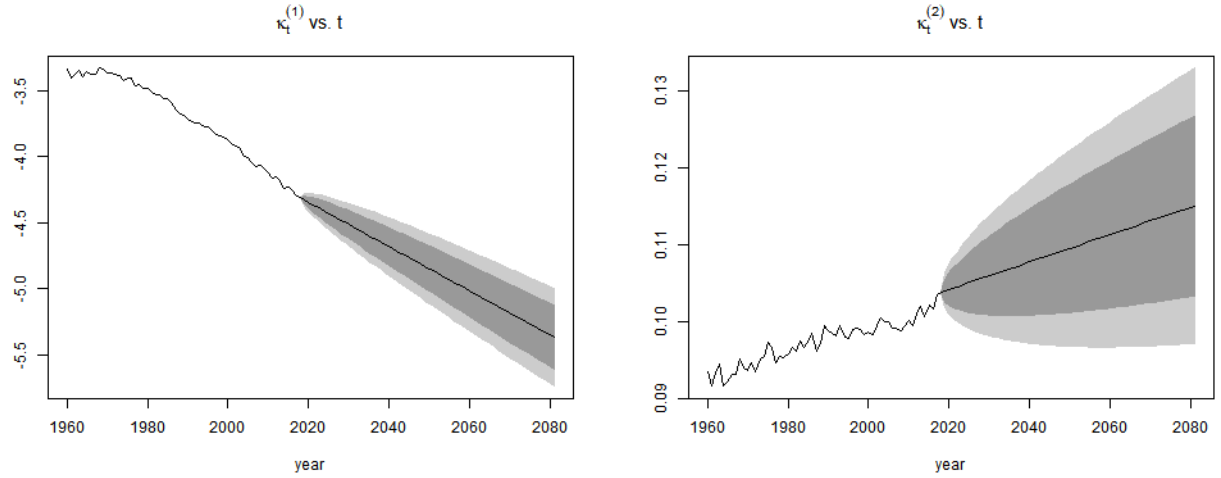
Source: the authors.

4.3.2 Mortality projection and extrapolation

We project κ_t as indicated in Section 3.2 and obtain the trends depicted in Figure 11. Having unsatisfactory data beyond the age of 100, we perform the logistic regression indicated in equation 14 using $\tilde{x} = 90, ..., 100$ allowing us to project mortality rates $\mu_x(t)$ until ω=120, the ultimate age. Results are given in Figure 12a and 12b. The vertical dotted line indicates the start of the logistic extrapolation. The discontinued line depicts the modeled rates whereas

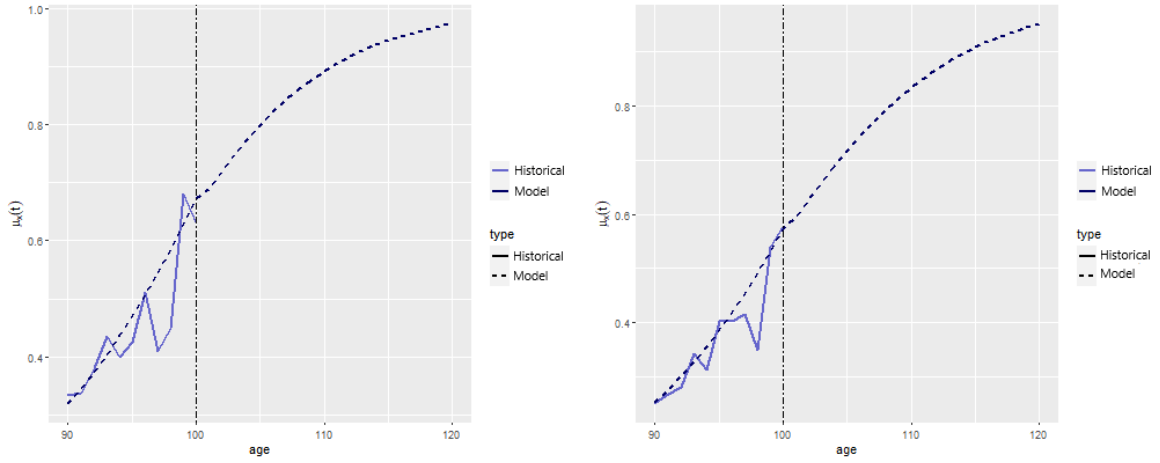
the blue lines correspond to the historical information.

Figure 11: Projection of $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$



Source: the authors.

Figure 12: Old age logistic regression



(a) 1968 (b) 1990

Source: the authors.

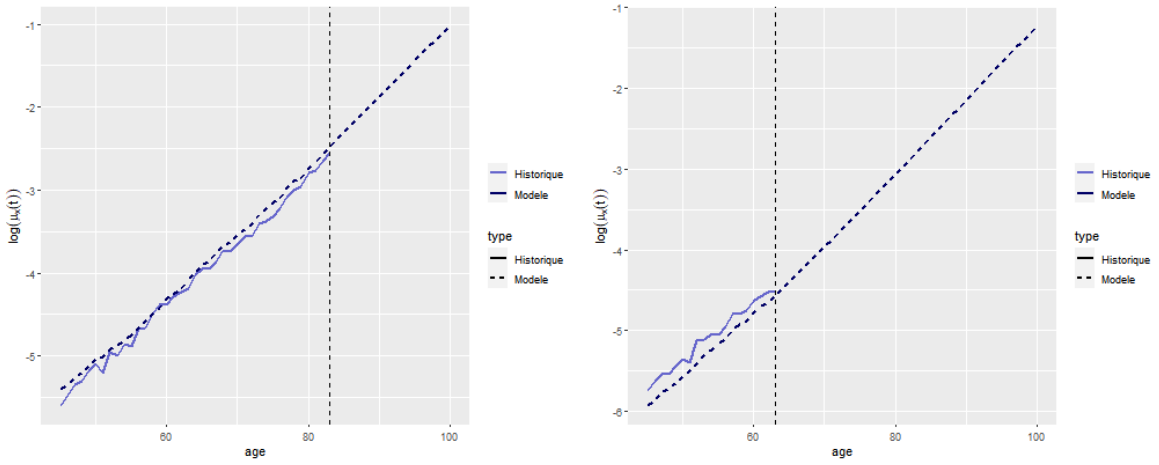
4.3.3 Global mortality: results

We analyse the effect of COVID-19 mortality in life insurance products. This is done for various cohorts to assess the relative impact of the pandemic for recent contracts versus contracts underwritten in 2000. In particular, we are interested in studying the male cohort born in 1935 and a second male cohort born in 1955.[12] Figures 13a and 13b show the fitted model (discontinued line) versus empirical mortality rates (blue continuous line) given by the maximum likelihood estimator for the two cohorts of interest. It is clear that the mortality

[12]Details about the characteristics of the contract are given in Section 4.5.

rate for the 1935 cohort is greater than the 1955, which clearly follows from the fact that the former was born during the world war II. The vertical dotted line represents the age from which no empirical data is available.[13]

Figure 13: Simulated versus historical log mortality rates



(a) 1935 cohort

(b) 1955 cohort

Source: the authors.

4.4 Model reconciliation

Agggregating the mortality rates obtained in Sections 4.2.4 and 4.3 we obtain the COVID and no COVID scenarios. Of course, as long as $\mu^{\text{COVID}} > 0$ we will have some excess mortality since $\mu^{\text{total}} = \mu^{\text{CBD}} + \mu^{\text{COVID}} > \mu^{\text{CBD}}$. However, the extend of this excess of mortality, and whether it greatly impacts life expectancy and product valuation is unknown. We investigate this here.

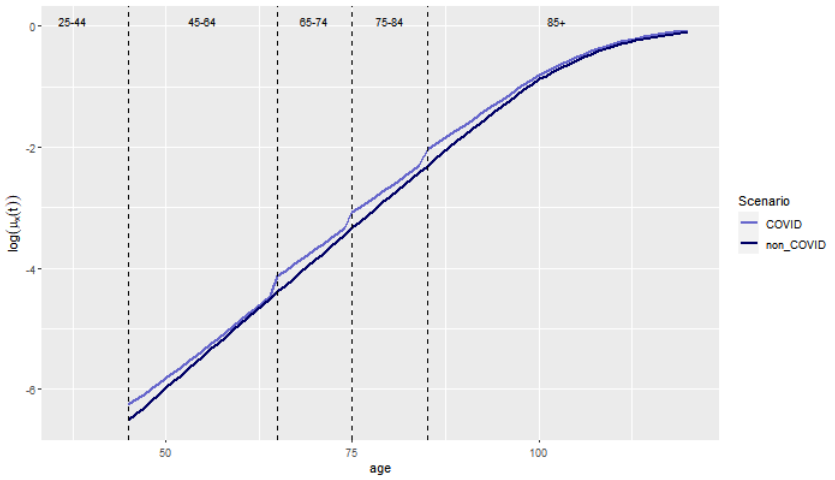
Log mortality rates $log(\mu_x(t))$ for year 2020, for the two scenarios, are presented in Figure 14. Vertical discontinuous lines represent the age categories, the blue and black line represent the COVID and non COVID scenarios respectively. We observe a jump upon each age-category change. These are explained considering differences between *natural* mortality and excess of mortality COVID for the first part of the age interval. Indeed, let us focus on 75-84 to clarify this. The excess of mortality is, in relative, much more important for a person aged 75 with respect to someone aged 84 since μ^{COVID} is constant for a particular age-category.[14]

We also calculate the cohort life expectancies using equation (15) for cohorts born in 1935, 1945 and 1970 for the COVID and non-COVID scenario (Figures 15a, 15b and 15c) aged 75, 85 and 50 in 2020 respectively. We observe that the relative big differences in the log mortality rate do not translate in extreme differences in cohort life expectancy, especially for old ages. We observe that the older the cohort, the greater the effect of COVID-19. However, this effect is at most three months. We deduce that the effect of COVID-19 in lifetime annuity pricing might be hence also reduced.

[13]Obviously, the 1935 and 1955 cohorts are aged 83 and 63 in 2018 the last observed year according to Human Mortality Database (2021), making it impossible to compare with empirical data beyond these ages.

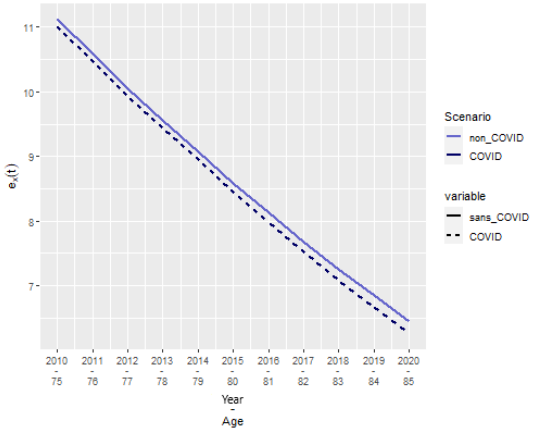
[14]In reality COVID related mortality will most likely vary within the age category. However, we are unable to extract this trend due to data limitations.

Figure 14: $log(\mu_x(t))$, in 2020, two scenarios

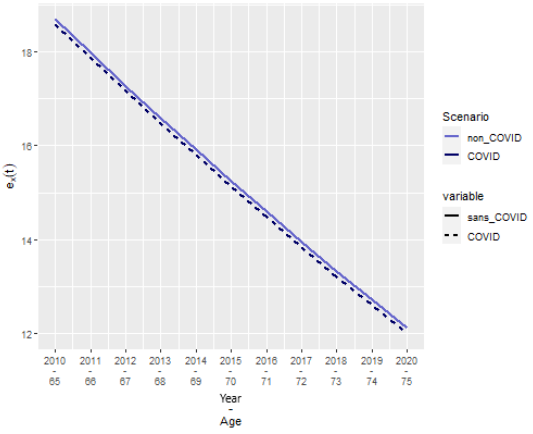


Source: the authors.

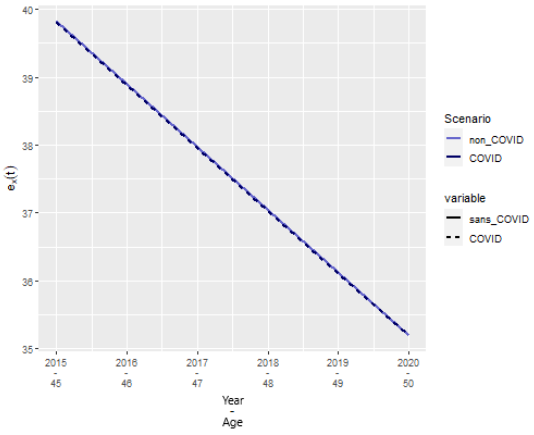
Figure 15: $e_x(t)$: two scenarios



(a) Cohort born in 1935



(b) Cohort born in 1945



(c) Cohort born in 1970

Source: the authors.

4.5 Actuarial application

We study a whole life insurance, paying capital €20,000 $(= C)$ at the end of the year of death to the insured's beneficiary. The second product is a lifetime annuity immediate, paying €2,000 $(= P)$ per month while the insured is alive. We price at the technical rate of $\delta = 3\%$. These product have an inverse relationship with mortality. Indeed, an adverse mortality shock will increase the price of a whole life insurance whereas it will decrease the price of a lifetime annuity as death becomes more likely.

We consider two cohorts per underwriting year. We study cohorts born in 1970 and 1955 for the contracts underwritten in 2019, a year before the pandemic. These are 49 and 64 at underwriting and 65 and 50 during the pandemic. We consider cohorts born in 1945 and 1935 for contracts underwritten in 2000. These are aged 55 and 65 at underwriting and 85 and 75 during the COVID-19 year of 2020. Furthermore, we consider the effect of COVID-19 excess mortality during the year 2020 by studying a *with* and *without* scenario. The NPV, its variance, presented in Section 3.4 and the relative difference between COVID and non-COVID scenarios given by equation (29) are analysed:

$$\Delta NPV = \frac{NPV_{\text{COVID}}}{NPV_{\text{non COVID}}} - 1. \tag{29}$$

We also study the difference between the no COVID case and the *catastrophic* COVID scenario, whereby

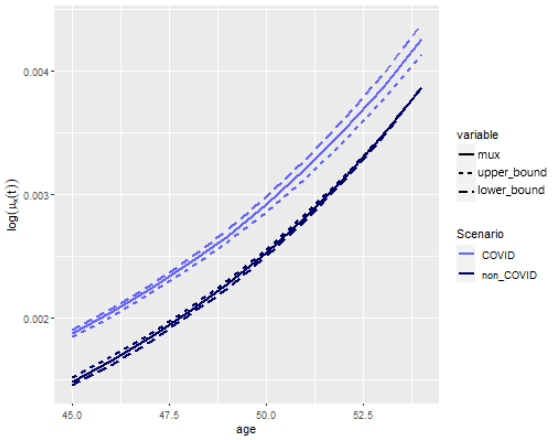
$$\mu^{\text{COVID CAT}} = 10 \cdot \mu^{\text{COVID}}.$$

Table 11 shows the NPV and standard deviation for the *base* COVID case, corresponding to the estimated mortality presented in the previous sections for the whole life insurance (left block) and lifetime annuity (right block). For the whole insurance case, we observe that the NPV increases in presence of COVID as expected. Indeed, higher mortality increases the likelihood of paying the capital, increasing accordingly the NPV. On the contrary, the NPV of the annuity decreases in presence of COVID. However, as in the case of the whole life insurance, the variation is negligible.

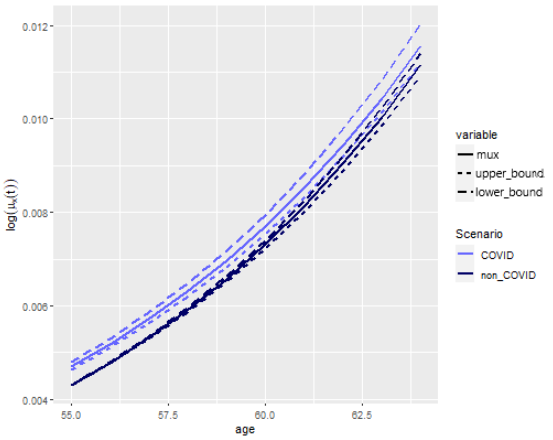
The standard deviation, on the other hand, has a more interesting trend. Indeed, for both insurance products we observe that the the variance increases for the contracts underwritten soon before the pandemic to younger individuals whereas it decreases for old contracts. Table 12 presents the effect of a *catastrophic* COVID scenario where mortality rate is multiplied by 10. We see that the relative change in NPV corresponds to roughly 10 times as well. Similarly decreases or increases in variance are almost 10-fold.

We found that the variance changed, depending on whether the contract was underwritten soon or long before the start of the pandemic. The worst possible outcome is to have a variance increase, which is the case for contracts underwritten in 2019 to individuals aged 49 and 64. On the contrary, old contracts see their variance decrease. This is linked to the behavior of the confidence intervals (CI) of the mortality rates which we depict in Figure 16. This graph shows the CI for all age categories with an adjusted scale in order to see the differences between the two scenarios.

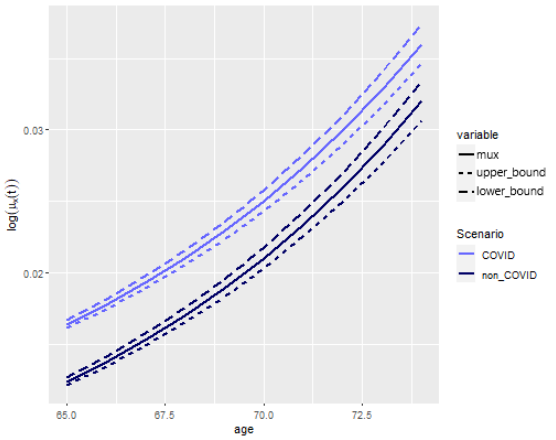
Figure 16: Confidence intervals for 2020: two scenarios



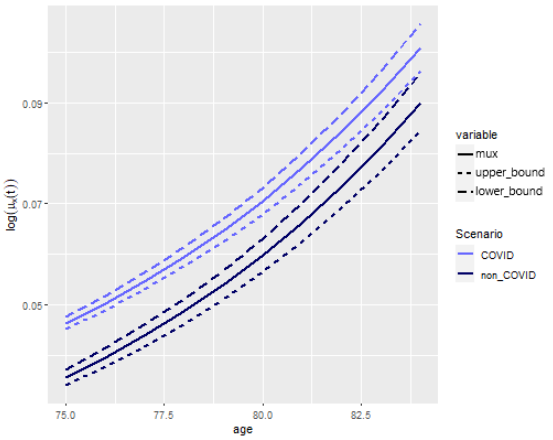
(a) ages 45-54



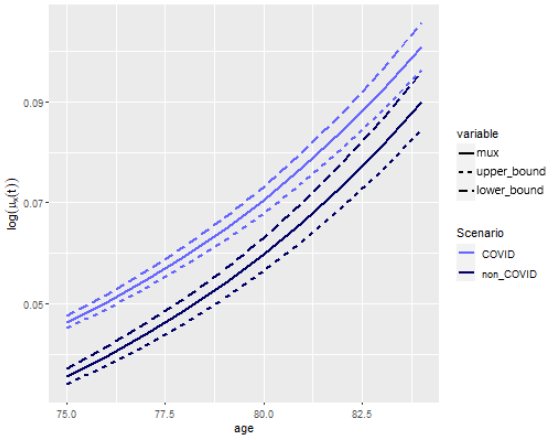
(b) ages 55-64 (zoom)



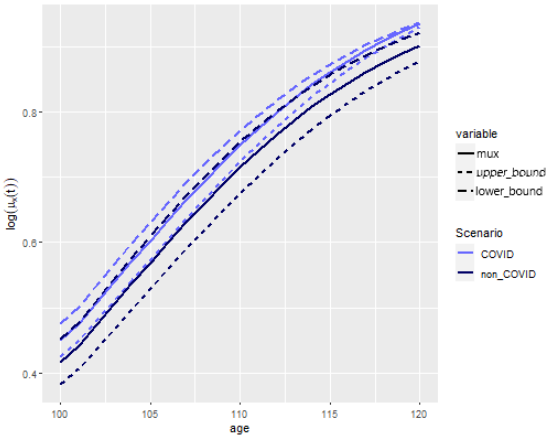
(c) ages 65-74 (zoom)



(d) ages 75-84



(e) ages 85+ (zoom)



(f) ages 100+ (zoom)

Table 11: NPV and variance for the COVID vs no COVID scenario: whole life insurance (left) and lifetime immediate annuity (right)

		Whole life insurance						Lifetime immediate annuity					
		NPV			Standard deviation (= σ)			NPV			Standard deviation (= σ)		
		Scenario		Δ^1 (%)	Scenario		Δ^1 (%)	Scenario		Δ^1 (%)	Scenario		Δ^1 (%)
		COVID	no COVID		COVID	no COVID		COVID	no COVID		COVID	no COVID	
		Underwriting year: 2000											
Age	75	9,492	9,469	0.241	3,374	3,379	-0.148	402,980	403,932	-0.235	131,879	132,138	-0.196
	85	12,225	12,203	0.180	3,289	3,308	-0.567	288,976	289,892	-0.316	122,519	123,434	-0.741
		Underwriting year: 2019											
Age	50	7,156	7,151	0.064	2,951	2,943	0.288	500,451	500,641	-0.038	117,177	116,817	0.308
	65	11,023	10,992	0.285	3,309	3,279	0.935	339,090	340,397	-0.384	126,118	124,801	1.055

Table 12: Difference between *catastrophic* COVID and no COVID: whole life insurance (left) and lifetime immediate annuity (right)

		Whole life insurance		Lifetime immediate annuity	
		Δ VAP^1 (%)	$\Delta\sigma^1$ (%)	Δ VAP^1 (%)	$\Delta\sigma^1$ (%)
		Underwriting year: 2000			
Age	75	2.294	-1.611	-2.244	-2.099
	85	1.551	-5.150	-2.724	-6.781
		Underwriting year: 2019			
Age	50	0.637	2.826	-0.379	3.019
	65	2.799	8.474	-3.770	9.491

[1] given by equation (29). We define the *catastrophic* COVID as $10 \cdot \mu^{\text{COVID}}$. *Source*: the authors.

Obviously mortality is greater in presence of COVID for all ages. However, the variability varies. Indeed, we observe wider bounds in presence of COVID prior to retirement (Figures 16a and 16b), slightly greater level of uncertainty for ages 65-74 (Figure 16c) and narrower bounds with COVID beyond the age of 75 (Figures 16d, 16e and 16f). We hypothesize that this is due to deaths being more unlikely in normal circumstances before retirement, so every additional death due to COVID yields a more uncertain outcome and affects the CI and variability of the insurance products accordingly. After retirement, it has the opposite effect as *natural* deaths are more common, and considering COVID renders them even more *likely*, narrowing down the CIs of the mortality rates and contracts for insured these ages accordingly.

Hence, the CI widens for younger ages whereas they narrow for older ages in presence of COVID. The narrowing (widening) of the CIs translates to a reduction (increase) in the standard deviation, respectively. Consequences of such a pandemic scenario show clear risks: for products that pay in the event of death we have a risk of underpricing, and overall we can observe a volatility increase that would put our reserves and capital levels at risk.

5 Conclusion

This paper provides an actuarial analysis based on excess of mortality in the context of COVID-19. We develop an epidemiological model to estimate the COVID-19 related mortality in Belgium in 2020 and reconciliate it with an aggregate mortality model. We contribute in this way to actuarial recent work on epidemiological models of e.g. Feng and Garrido (2011), Chen et al. (2021) and Hall et al. (2020). We extend Feng and Garrido (2011) considering a SIRD component with a death compartment and both Feng and Garrido (2011) and Chen et al. (2021) by adding age categories inspired by recent epidemiological work of Balabdaoui and Mohr (2020) for Switzerland and Franco (2021) for Belgium.

We present a SIRD epidemiological model that relies on data-specific R_0. It results in a model that very accurately reflects number of deaths for different age categories as well as accurate timing with regards to peaks and valleys. From an empirical epidemiological perspective, we extend the work of Franco (2021) to the full year of 2020, whereas their study is limited to 31/10/2020 due to data constraints.

We find that considering COVID-19 increases the Net Present Value (NPV) of a whole life insurance and decreases the NPV of a lifetime immediate annuity as expected. We find that variability of the contract increases for recently underwritten contracts whereas variability decreases for old contracts, and this for both insurance products. We hypothesize this is due to COVID-19 rendering death at older ages more *certain*, decreasing the variability of such products. In all cases, the level of standard deviation changes remains limited.

Our work is comprehensive but has various venues to improve. The study is limited to one year of COVID-19 related mortality data. Adding more years, combined with the various age categories, would raise the need to create a transition mechanisms between age categories. That could be added within the EDO system. Furthermore, our model reflects the main compartments of interest in the context of an insurance application, be it recovery or death as payments are contingent on either survival or death. It does not reflect other aspects of the pandemic such as quarantine, incubation or hospitalization. Finally, we have abstracted, due to data limitations, from considering the effect of COVID-19 in the global health landscape. Indeed, one of the challenges of the pandemic has been managing the sudden inflow of sick people in the intensive care units. This obviously has had a (negative) effect in the treatment of other diseases for which the short and long-term effects are still unknown.

6 Bibliography

A. Abou-Ismail. Compartmental models of the covid-19 pandemic for physicians and physician-scientists. *SN comprehensive clinical medicine*, 2(7):852–858, 2020.

S. Arnold and M. Sherris. Forecasting mortality trends allowing for cause-of-death mortality dependence. *North American Actuarial Journal*, 17(4):273–282, 2013.

S. Arnold and M. Sherris. Causes-of-death mortality: What do we know on their dependence? *North American Actuarial Journal*, 19(2):116–128, 2015.

F. Balabdaoui and D. Mohr. Age-stratified discrete compartment model of the covid-19 epidemic with application to switzerland. *Scientific reports*, 10(1):1–12, 2020.

A. Boumezoued, H. L. Hardy, N. El Karoui, and S. Arnold. Cause-of-death mortality: What can be learned from population dynamics? *Insurance: Mathematics and Economics*, 78: 301–315, 2018.

A. J. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73(4): 687–718, 2006.

G. C. Calafiore, C. Novara, and C. Possieri. A modified sir model for the covid-19 contagion in italy. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 3889–3894. IEEE, 2020.

CDC. 1918 pandemic (h1n1 virus). retrieved from `https://www.cdc.gov/flu/pandemic-resources/1918-pandemic-h1n1.html`, 2021. Accessed: 24-05-2021.

X. Chen, W. F. Chong, R. Feng, and L. Zhang. Pandemic risk management: resources contingency planning and allocation. *Insurance: Mathematics and Economics*, 101:359–383, 2021.

R. Feng and J. Garrido. Actuarial applications of epidemiological models. *North American Actuarial Journal*, 15(1):112–136, 2011.

J. Fernández-Villaverde and C. I. Jones. Estimating and simulating a sird model of covid-19 for many countries, states, and cities. *Journal of Economic Dynamics and Control*, page 104318, 2022.

N. Franco. Covid-19 belgium: Extended seir-qd model with nursing homes and long-term scenarios-based forecasts. *Epidemics*, 37:100490, 2021.

R. D. Hall, C. S. MacDonald, P. J. Miller, A. N. Natsis, L. A. Schilling, S. C. Siegel, and J. P. Wiese. Society of actuaries research brief : Impact of covid-19, june 12, 2020. Technical report, Society of Actuaries, 2020.

H. W. Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.

A. C. Hindmarsh. Odepack, a systematized collection of ode solvers. *Scientific computing*, pages 55–64, 1983.

Y. Huang, L. Yang, H. Dai, F. Tian, and K. Chen. Epidemic situation and forecasting of covid-19 in and outside china. *Bull World Health Organ*, 10, 2020.

Human Mortality Database. Life tables belgium 1968-2018 - total (both sexes). retrieved from `https://www.mortality.org/`, 2021. Accessed: 24-05-2021.

Institute and Faculty of Actuaries. Longevity bulletin 6: The pandemic edition. Technical report, Institute and Faculty of Actuaries, 2015.

W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. *Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character*, 138(834):55–83, 1932.

A. T. Levin, W. P. Hanage, N. Owusu-Boaitey, K. B. Cochran, S. P. Walsh, and G. Meyerowitz-Katz. Assessing the age specificity of infection fatality rates for covid-19: systematic review, meta-analysis, and public policy implications. *European journal of epidemiology*, 35(12): 1123–1138, 2020.

H. Li and Y. Lu. Modeling cause-of-death mortality using hierarchical archimedean copula. *Scandinavian Actuarial Journal*, 2019(3):247–272, 2019.

H. Li, H. Li, Y. Lu, and A. Panagiotelis. A forecast reconciliation approach to cause-of-death mortality modeling. *Insurance: Mathematics and Economics*, 86:122–133, 2019.

P. Lyu, A. De Waegenaere, and B. Melenberg. A multi-population approach to forecasting all-cause mortality using cause-of-death mortality data. *North American Actuarial Journal*, 25(sup1):S421–S456, 2021.

G. Molenberghs, C. Faes, J. Verbeeck, P. Deboosere, S. Abrams, L. Willem, J. Aerts, H. Theeten, B. Devleesschauwer, N. B. Sierra, et al. Belgian covid-19 mortality, excess deaths, number of deaths per million, and infection fatality rates (9 march—28 june 2020). *medRxiv*, 2020.

A. Rogers and K. Gard. Applications of the heligman/pollard model mortality schedule. *Population Bulletin of the United Nations*, (30):79–105, 1991.

P. R. Saunders-Hastings and D. Krewski. Reviewing the history of pandemic influenza: understanding patterns of emergence and transmission. *Pathogens*, 5(4):66, 2016.

Sciensano. Covid-19 database. retrieved from `https://epistat.wiv-isp.be/covid/`, 2021. Accessed: 24-05-2021.

Statbel. Population structure. retrieved from `https://statbel.fgov.be/fr/themes/population/structure-de-la-population`, 2021. Accessed: 24-05-2021.

E. Tabeau, P. Ekamper, C. Huisman, and A. Bosch. Improving overall mortality forecasts by analysing cause-of-death, period and cohort effects in trends. *European Journal of Population/Revue Européenne de Démographie*, 15(2):153–183, 1999.

L. Tang, Y. Zhou, L. Wang, S. Purkayastha, L. Zhang, J. He, F. Wang, and P. X.-K. Song. A review of multi-compartment infectious disease models. *International Statistical Review*, 88 (2):462–513, 2020.

A. R. Thatcher. The long-term pattern of adult mortality and the highest attained age. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 162(1):5–43, 1999.

A. M. Villegas, V. K. Kaishev, and P. Millossovich. Stmomo: An r package for stochastic mortality modeling. *Journal of Statistical Software*, 84:1–38, 2018.

WHO. Past pandemics. retrieved from `https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/pandemic-influenza/past-pandemics`, 2021a. Accessed: 24-05-2021.

WHO. Ebola virus disease. retrieved from `https://www.who.int/csr/disease/ebola/en/`, 2021b. Accessed: 24-05-2021.

L. Willem, T. Van Hoang, S. Funk, P. Coletti, P. Beutels, and N. Hens. Socrates: an online tool leveraging a social contact data sharing initiative to assess mitigation strategies for covid-19. *BMC research notes*, 13(1):1–8, 2020.

J. R. Wilmoth. 13 mortality projections for japan. *Health and mortality among elderly populations*, page 266, 1996.

S. Zhao and H. Chen. Modeling the epidemic dynamics and control of covid-19 outbreak in china. *Quantitative biology*, 8(1):11–19, 2020.

M. Zheng and J. P. Klein. Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika*, 82(1):127–138, 1995.

G. Zittersteyn and J. Alonso-García. Common factor cause-specific mortality model. *Risks*, 9 (12):221, 2021.

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its strategic research programs: aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of topical research available, including an expanding collection of international and market-specific research, experience studies, models and timely research.

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