IA|BE PROSPECTIVE MORTALITY TABLES 2020 – COVID-19 impact analysis

NOVEMBER 2020

Assessing the impact of COVID-19 on the IA|BE 2020 mortality projections: a scenario analysis

Katrien Antonio [*1,2], Sander Devriendt [†], Jens Robben [‡], and Dominik Sznajder [§]

[1]Faculty of Economics and Business, KU Leuven, Belgium.
[2]Faculty of Economics and Business, University of Amsterdam, The Netherlands.

November 27, 2020

1 Introduction

In December 2019, the coronavirus disease (COVID-19) originated in the Chinese city Wuhan. In the months that followed, the virus spread across the world. At the time of writing, about 16.9 million positive cases and 383 213 deaths have been identified in Europe.[1] The United Kingdom has the highest absolute number of COVID-19 deaths in Europe (56 630), followed by Italy (52 028) and France (50 700). Belgium has reported 16 077 deaths.[2] While we are currently facing a second COVID-19 wave in Belgium, the recent announcements of two potential vaccines[3] bring hope for a (hopefully sharp) drop in the future number of COVID-19 deaths. But even in this bright scenario, the impact of COVID-19 on the frailty of survivors is unclear. In this note, we aim to outline the possible impact of the COVID-19 pandemic on the IA|BE 2020 mortality projection model and the scenarios generated for future mortality with this model. However, at this moment, statements and analysis regarding the impact of COVID-19 remain highly speculative in nature. The approach and results discussed in this note should be seen as *one* example of a COVID-19 impact analysis on the IA|BE 2020 model. Therefore, one should interpret the results with caution, in light of today's uncertainties regarding the further evolution of the disease.

[*]Corresponding author. E-mail address: katrien.antonio@kuleuven.be

[†]Voluntary research assistant on the mortality project at KU Leuven, AFI Department, Insurance Research Group. This paper reflects the personal views of the author and not the views of his employer.

[‡]Research assistant on the mortality project at KU Leuven, AFI Department, Insurance Research Group.

[§]Manager and actuarial consultant at Milliman. This paper reflects the personal views of the author and not the views of his employer.

[1]Numbers are retrieved from https://www.statista.com/statistics/1102209/coronavirus-cases-development-europe/ and https://www.statista.com/statistics/1102288/coronavirus-deaths-development-europe/ on 26 November 2020.

[2]These numbers of COVID-19 deaths come from the COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) on 26 November 2020, see https://www.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6.

[3]From Pfizer and BioNTech, see https://www.pfizer.com/news/press-release/press-release-detail/pfizer-and-biontech-announce-vaccine-candidate-against, and Moderna, see https://investors.modernatx.com/news-releases/news-release-details/modernas-covid-19-vaccine-candidate-meets-its-primary-efficacy.

The COVID-19 pandemic has impacted mortality in 2020 in multiple ways. The disease itself has led an still leads to an increase in the number of deaths, especially at the higher ages. However, measures taken by the governments worldwide also impacted mortality in a positive way, leading to less traffic or work-related accidents in 2020. While this note focuses on IA|BE 2020, we acknowledge some other, recent contributions that aim at assessing the impact of COVID-19 on mortality forecasts. Duyck et al. (2020) has actualised the demographical outlook of the Federaal Planbureau for 2019-2070. Koninklijk Actuarieel Genootschap (2020) performs a sensitivity analysis that shows the impact of the pandemic on the cohort life expectancies in 2021 by feeding virtual deaths and exposures in 2019-2020 to the KAG 2020 model. Further, van Delft and Huijzer (2020) use Koninklijk Actuarieel Genootschap (2018) as a starting point and investigate the impact of four different COVID-19 scenarios on the Dutch best estimate mortality table published in 2018. They create these COVID-19 scenarios by multiplying the mortality rates in 2020 (and later) with a shock factor. These shocks are defined in a particular age bucket as (a fraction of) the ratio of the observed death rate in the first 23 weeks of 2020 to the average of the observed death rates in the first 23 weeks in earlier years.

As outlined in Antonio et al. (2020), IA|BE 2020 calibrates mortality models collected on the data set with annual data on the observed numbers of deaths, $d_{x,t}$, and the corresponding exposures to risk, $E_{x,t}$. More specifically, IA|BE 2020 puts focus on a set of countries over the calibration period 1988-2018 (European trend) and 1988-2019 (Belgian trend) and over the age range 0-90. In the impact assessment strategy proposed in this note we stay close to the data set-up, model design and calibration strategy of IA|BE 2020. Hereto, we extend the data set underlying IA|BE 2020 with virtual exposure points $E_{x,t}$ and death counts $d_{x,t}$ until the year 2020. These extra data points are not available because the year 2020 is still ongoing and so the Human Mortality Database ([HMD]) and Eurostat only report mortality data up to the year 2019 (or earlier) at the time of writing.

This report is organised as follows. First, in Section 2 we discuss notation and list the data sources that provide us with weekly death counts and exposures in 2019 and part of 2020. In Section 3 we create the COVID-19 impacted data set of deaths and exposures till 2020. Next, we recalibrate the stochastic multi-population mortality model underlying the IA|BE 2020 methodology in Section 4 and assess the impact of the virtual data on the cohort life expectancy in 2020. We provide an outlook in Section 5. Technical details are deferred to the Appendix. We list the data sources in Appendix A. Appendix B describes the construction of the virtual exposure points $E_{x,t}$ for ages $x \in \{0, 1, 2, \ldots, 90\}$ and years 2019-2020 (Europe) and 2020 (Belgium). In Appendix C, we construct the three scenarios for the death counts $d_{x,t}$ for the same set of ages and years.

2 Data and notation

We have partial information for the years 2019-2020 available in the form of weekly death counts and death rates reported by the HMD, called the Short Term Mortality Fluctions ([STMF]),[4] and Eurostat.[5] These sources report mortality statistics for age buckets $[x_i, x_j]$ instead of individual ages x. Therefore, in order to be compliant with the design of IA|BE 2020, we need to transform the weekly mortality statistics in age buckets towards annual death counts and exposures at

[4]This information can be explored using the visualization toolkit on https://mpidr.shinyapps.io/stmortality/.

[5]Eurostat provides weekly death statistics at https://ec.europa.eu/eurostat/web/COVID-19/data.

individual ages. We use the following notations (for now, we leave out gender g in our notation):

$$d_{[x_i,x_j],t,w}, \qquad E_{[x_i,x_j],t,w} \quad \text{and} \quad m_{[x_i,x_j],t,w},$$

for the death counts, exposures and death rates respectively in age bucket $[x_i, x_j]$ in week w in year t. Further, $w \in \{1, 2, 3, \ldots, 52\}$.

Using the STMF reported weekly death counts $d_{[x_i,x_j],t,w}$ and death rates $m_{[x_i,x_j],t,w}$ for week w in year t and in age bucket $[x_i, x_j]$, we can derive the weekly exposures $E_{[x_i,x_j],t,w}$ using the following relation:

$$E_{[x_i,x_j],t,w} = \frac{d_{[x_i,x_j],t,w}}{m_{[x_i,x_j],t,w}}.$$

STMF reports mortality information in large age buckets:

$$[0, 14], \quad [15, 64], \quad [65, 74], \quad [75, 84], \quad 85+.$$

As mentioned in the STMF documentation,[6] the HMD team extrapolates annual death rates in order to estimate age-specific exposures and death counts under the assumption of zero migration. This strategy is needed since the exposure in 2020 is not known yet. They assume a constant weekly exposure per year, per age bucket and per gender. When going from weekly to yearly exposures, we simply multiply the weekly exposures with a factor 52, i.e. the number of weeks in a year:

$$E_{[x_i,x_j],t} = 52 E_{[x_i,x_j],t,w},$$

where $E_{[x_i,x_j],t}$ now denotes the total annual exposure in year t for age bucket $[x_i, x_j]$. STMF lists death counts and death rates for Northern Ireland, England and Wales and Scotland separately. A simple aggregation of their death counts leads to the death counts of the United Kingdom as a whole.

Next to the HMD and its STMF project, Eurostat also lists valuable datasets related to death counts, useful to assess the impact of COVID-19 on mortality rates. From Eurostat we obtain the deaths by week, sex and 5-year age bucket.[7] For many countries, the STMF reported death counts correspond to the aggregated death counts reported by Eurostat. If this correspondence holds true, the data from Eurostat is more preferable due to the smaller age buckets, which will eventually lead to a more accurate transition towards death counts at individual ages, as used in the IA|BE 2020 mortality model. The 19 age buckets, reported by Eurostat, are

$$[0, 4], \quad [5, 9], \quad [10, 14], \quad [15, 19], \quad \ldots, \quad [85, 89], \quad 90+.$$

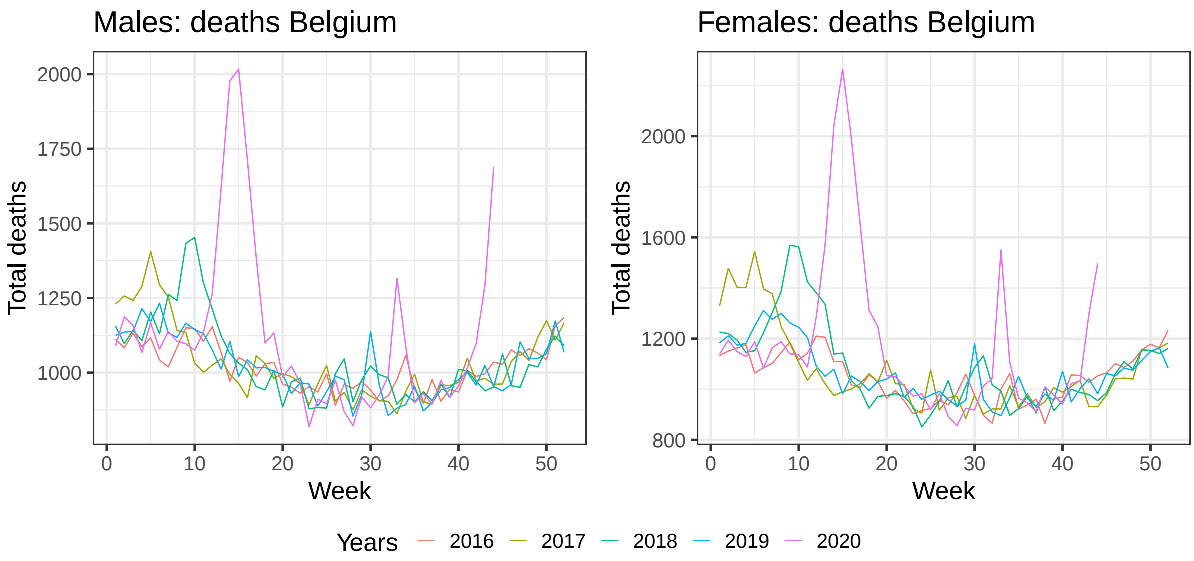
Eurostat does not report any information about weekly exposures. For data quality reasons we only use the Eurostat reported weekly death counts in the small age buckets whenever their aggregated death counts correspond to the ones reported by STMF. This is the case for all countries, except Germany[8], France and the United Kingdom.

In Antonio et al. (2020), we use a collection of 14 European countries as a basis for a European trend, namely Belgium, The Netherlands, Luxembourg, Norway, Switzerland, Austria, Ireland, Sweden, Denmark, Germany, Finland, Iceland, United Kingdom and France. Neither STMF

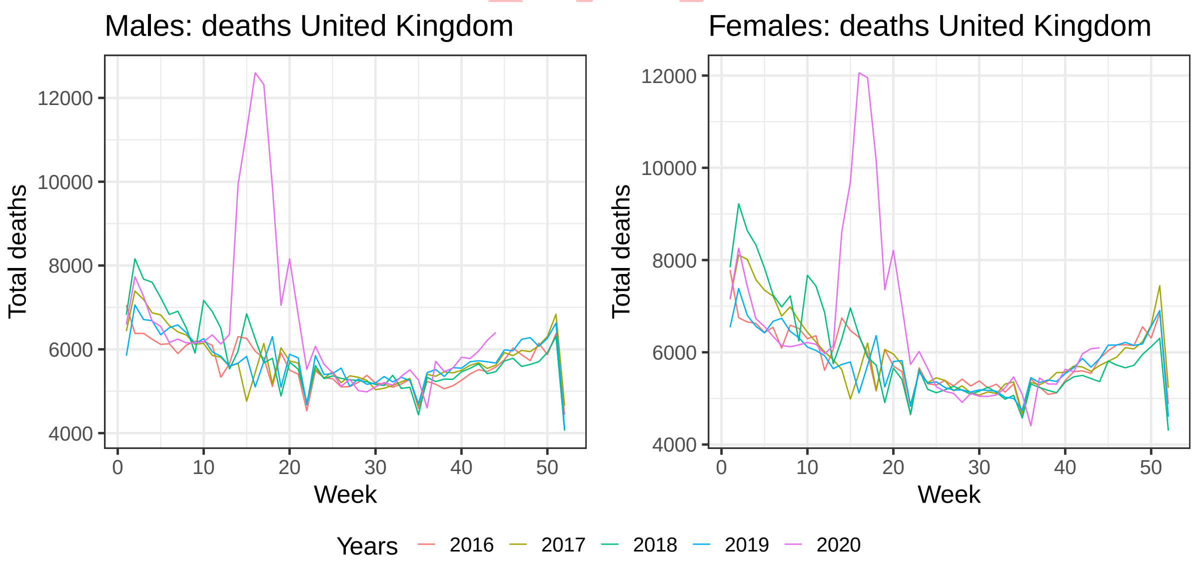
[6] This note can be consulted on https://www.mortality.org/Public/STMF_DOC/STMFNote.pdf.
[7] See https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_mwk_05&lang=en.
[8] Eurostat does not provide weekly death counts for Germany for age buckets of length 5.

nor Eurostat report granular mortality information on Ireland. For this reason we exclude this country in the COVID-19 impact assessment covered in this note. However, given the exposure size of the Irish population, we do not expect that this has a major impact on the results obtained with the multi-population mortality model. To provide a clear overview, Table 4 in Appendix A lists the data sources we use for each country.

Figure 1 illustrates the number of deaths per week for the last couple of years for Belgium, United Kingdom and Germany. In the first two cases we observe the peaks corresponding to the COVID-19 waves, whereas such pattern is missing for the German data.

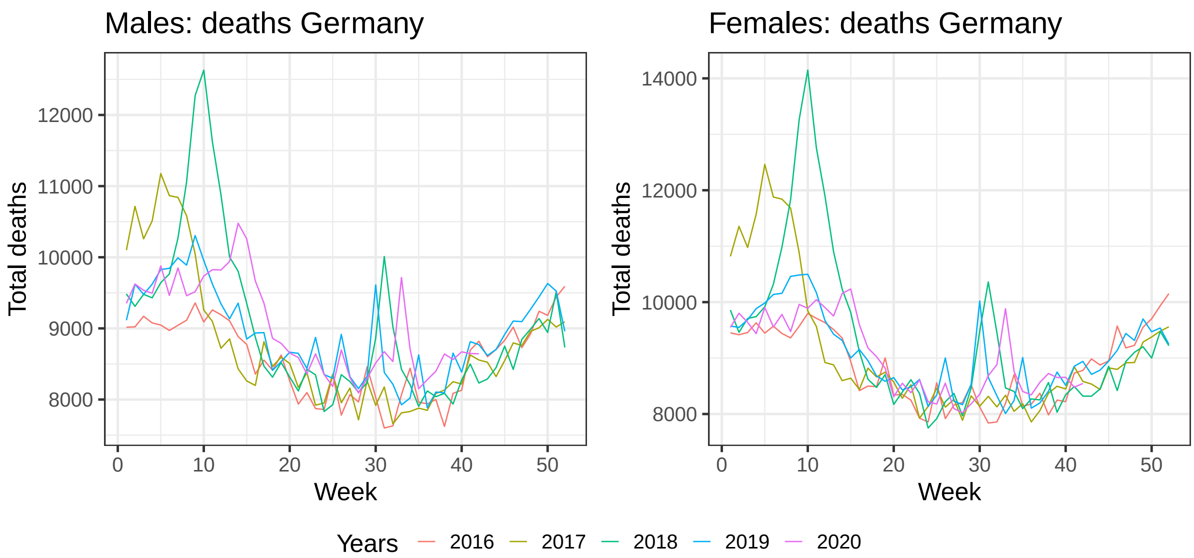


(a)



(b)

Figure 1: Total weekly deaths in Belgium (top) and the United Kingdom (bottom) in the years 2016-2020 for males and females. Eurostat data, first 44 weeks of 2020 are displayed.



(c)

Figure 1: Total weekly deaths in Germany in the years 2016-2020 for males and females. STMF data, first 42 weeks of 2020 are displayed.

3 Creating virtual data points for 2019 and 2020

Our strategy to evaluate the impact of COVID-19 on the IA|BE 2020 model is to focus on the input data and to follow the same design principles of this stochastic mortality model. Because IA|BE 2020 starts from a multi-population mortality data set with annual deaths $d_{x,t}$ and exposures $E_{x,t}$ for years 1988-2018 (Europe) and 1988-2019 (Belgium), we have to construct the corresponding virtual exposures and death counts for the European countries in 2019 (if necessary) and 2020. As Table 4 in Appendix A indicates, for some European countries the 2019 data is already available on the HMD while for others this information is not yet available. In order to assess the impact of the pandemic on the mortality forecasts obtained from IA|BE 2020, we will recalibrate the model using these virtually extended data sets.

We start from the weekly mortality data as discussed in Section 2. Exposures for 2019-2020 in age buckets are retrieved from STMF. In principle we also have weekly death counts for 2019-2020 in certain age buckets on STMF and Eurostat (see Appendix A). However, the death counts for 2020 are only partially observed. Therefore, in order to obtain the total death counts for 2020, we consider three scenarios that extrapolate the number of deaths in the first 26 weeks to the remaining weeks of 2020. At the time of writing, only the first 26 weeks are stable and reliable enough to conduct the analysis. Moreover, this choice allows to clearly distinguish between the first and second COVID-19 wave. The death counts in the first half year are by now fully observed, whereas an assumption has to be made for the second half year. Therefore, we first define an unrealistic base scenario S_0 where in each age bucket the death rate in the second half of 2020 equals the death rate in the second half of 2019, i.e. there is no second COVID-19 wave in this base scenario. In addition, we define a worst case scenario S_∞ where the total number of deaths in the second half and in the first half of 2020 are equal, i.e. we assume that

COVID-19 causes as much deaths in the second wave compared to the deaths in the first wave.[9] Using these two scenarios S_0 and S_∞, we set up three different, realistic scenarios in Table 1, namely starting from the base scenario, we add 50%, 75% or 100% of the excess of deaths in scenario S_∞ compared to scenario S_0. This roughly corresponds to a second COVID-19 wave in the last 26 weeks of 2020 causing 50%, 75% or 100% of the total number of deaths observed in the first COVID-19 wave in the first half of 2020. At the moment of writing 9 635 COVID-19 deaths were observed in the first wave (first half of 2020), versus 6 442 deaths in the second wave.[10]

Death scenario	Excess of deaths in second half of 2020
Scenario 1	50% of (deaths in S_∞ minus deaths in S_0)
Scenario 2	75% of (deaths in S_∞ minus deaths in S_0)
Scenario 3	100% of (deaths in S_∞ minus deaths in S_0)

Table 1: Scenarios for the number of deaths in the last 26 weeks of 2020.

In Appendix B we explain how we go from weekly exposures in age buckets $E_{[x_i,x_j],t,w}$ to annual exposures at individual ages $E_{x,t}$. Figure 2 visualizes the stacked exposures at ages 0-90 for the year 2019 (left panel) and 2020 (right panel) for all 13 European countries. Figure 3 shows the evolution of the (virtual) annual exposures for Belgium, the United Kingdom and Germany. The exposures in year 2020 (and 2019) are created using our approach.

Appendix C then puts focus on the scenarios for the number of deaths in the second half of 2020 and again describes the transformation of these weekly death counts in age buckets $d_{[x_i,x_j],t,w}$ to annual death counts at individual ages $d_{x,t}$. Figure 4 shows the stacked European number of deaths for 2019, and for the second death scenario of 2020. Figure 5 shows the evolution of deaths through time for three European countries and shows the impact on the deaths for each scenario.

In addition, we look at the estimated excess of mortality in Belgium in 2020 in each of the three scenarios. Figure 6 compares the estimated number of deaths in 2020 under the original IA|BE 2020 model with the virtual deaths in 2020 in each of the three proposed scenarios. Figure 7 visualizes the same situation for the mortality rates. We observe a clear excess of mortality at the higher ages. The opposite is true for the younger ages.

[9] In this scenario S_∞ we even over-estimate the number of COVID-19 deaths in the second half compared to the deaths in first half of 2020 because the number of deaths in the first half of any year typically exceeds the number of deaths in the second half of that same year.

[10] Numbers are retrieved from https://www.statista.com/statistics/1111078/cumulative-coronavirus-deaths-in-belgium/ on 26 November 2020.

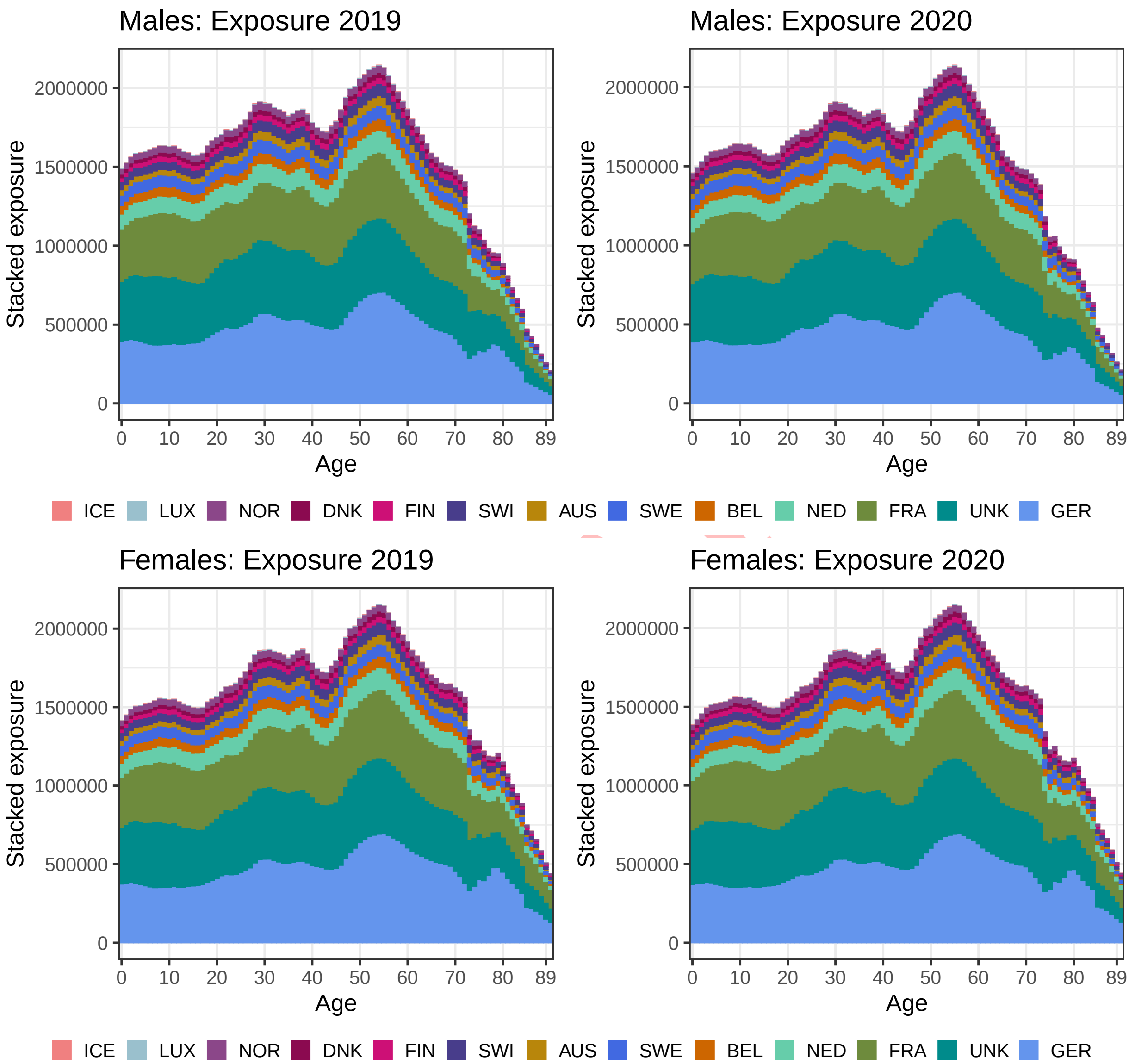


Figure 2: Stacked male (top) and female (bottom) exposure for the thirteen European countries as a function of age.

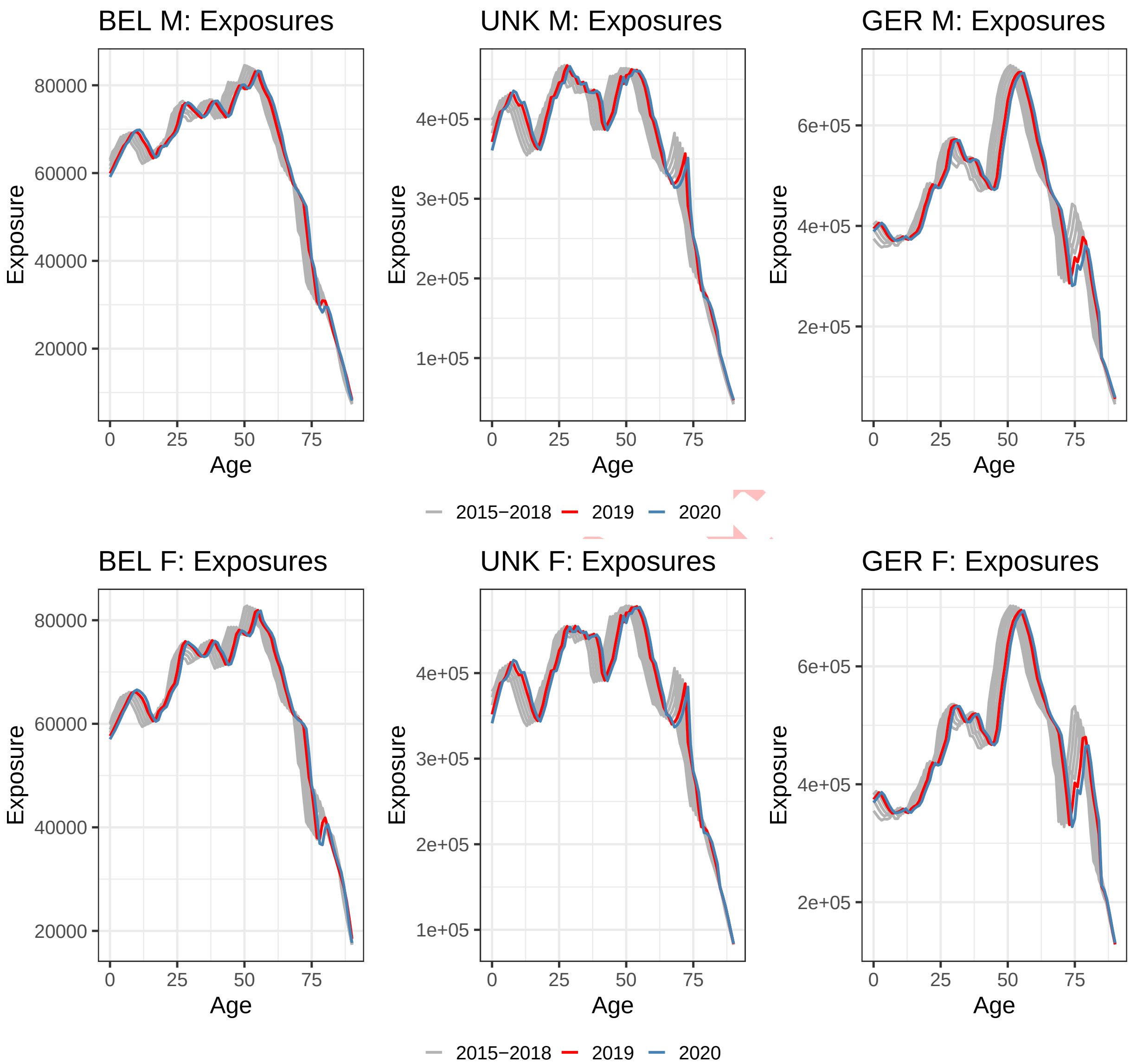


Figure 3: (Virtual) annual exposures $E_{x,t}$ for males (top) and females (bottom) in Belgium (left), the United Kingdom (middle) and Germany (right) as a function of age, years 2015-2020.

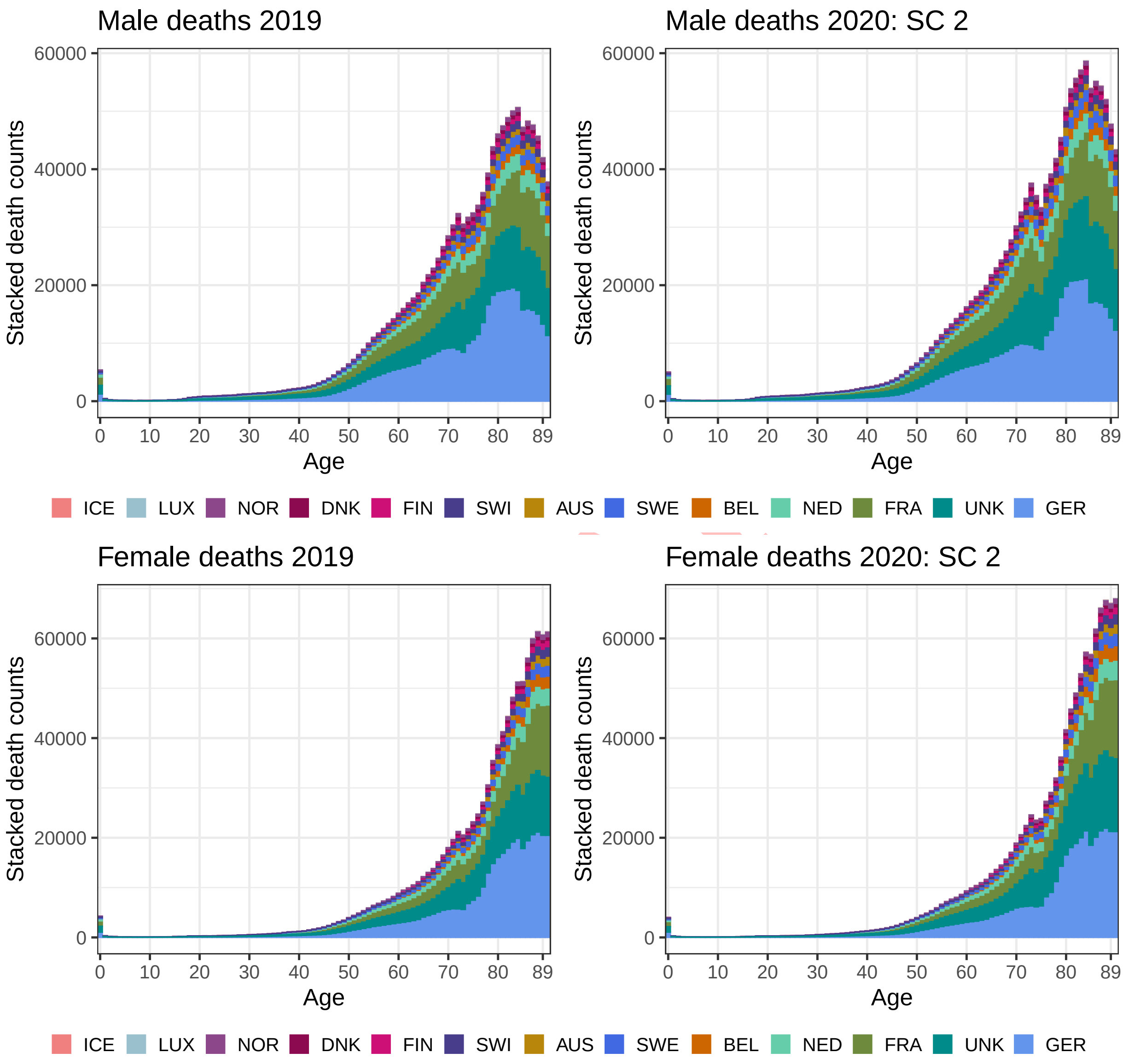


Figure 4: The stacked (virtual) European death counts in 2019 (left) and the second scenario for 2020 (right) for males (top) and females (bottom).

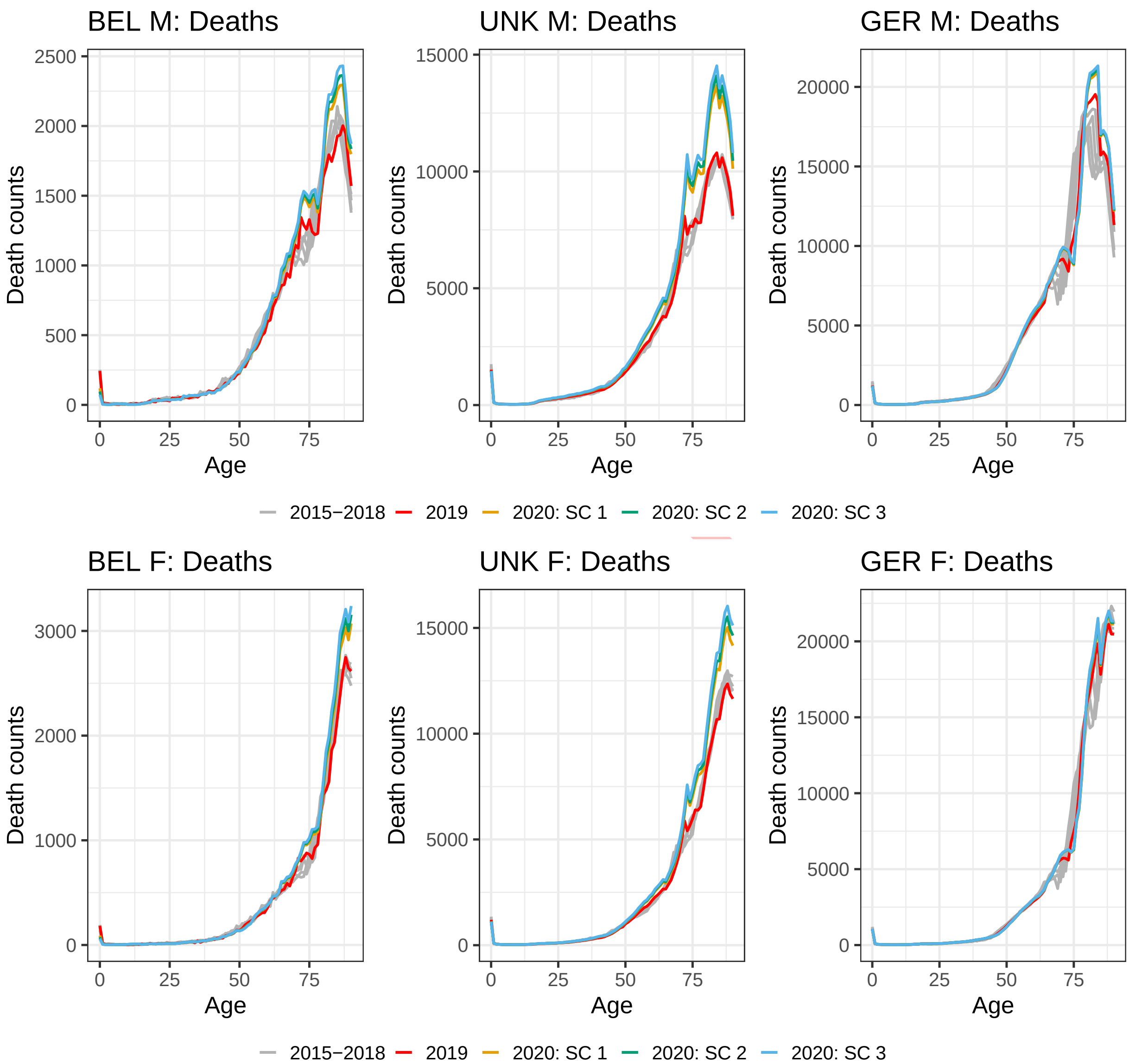


Figure 5: (Virtual) annual death counts $d_{x,t}$ for males (top) and females (bottom) in Belgium (left), the United Kingdom (middle) and Germany (right) as a function of age, years 2015-2019 and the three death scenarios for 2020.

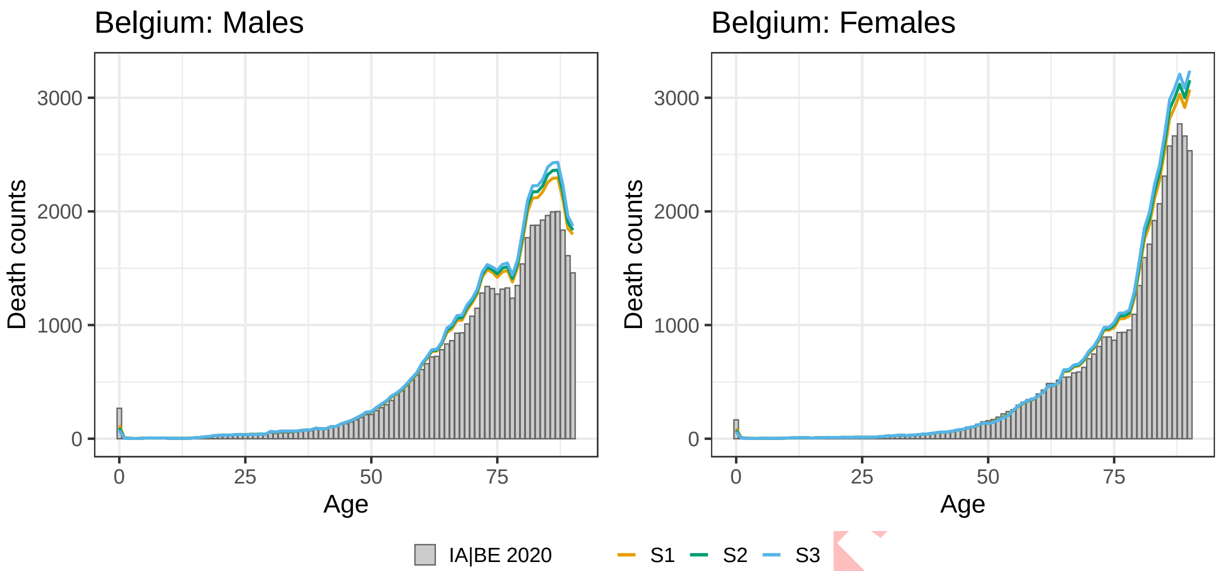


Figure 6: A comparison of the estimated number of Belgian deaths in 2020 under the IA|BE 2020 model with the observed (virtual) deaths in 2020 in each of the tree scenarios.

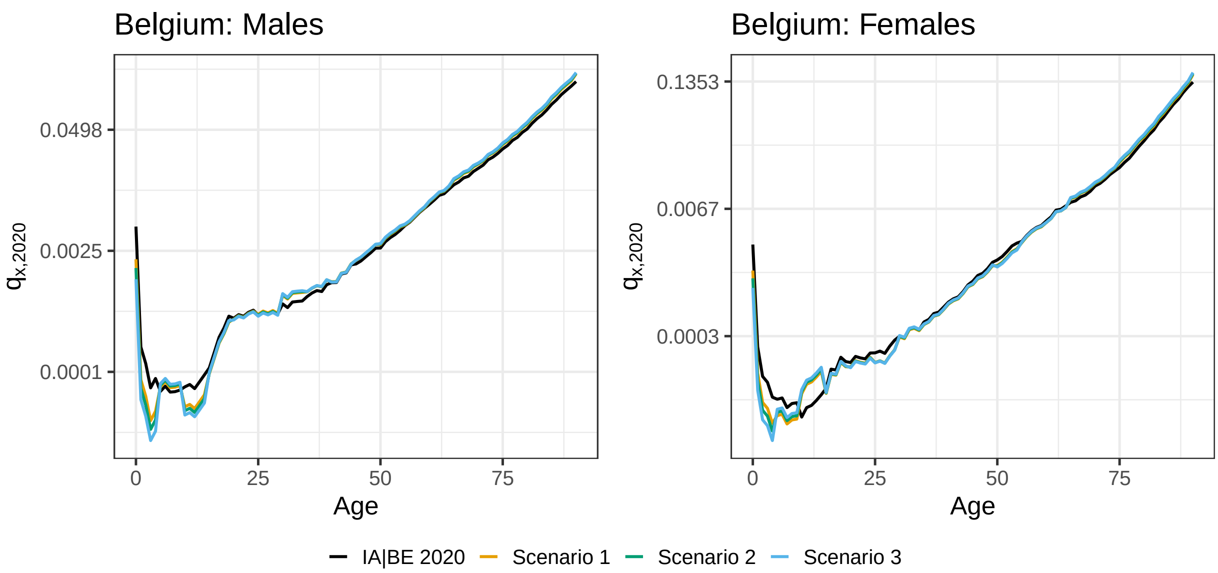


Figure 7: A comparison of the estimated mortality rates $\hat{q}_{x,t}$ in 2020 on log-scale under the IA|BE 2020 model with the observed (virtual) mortality rates in 2020 in each of the tree scenarios.

4 Recalibrating IA|BE 2020 when including the virtual data points: a COVID-19 impact assessment

The IA|BE 2020 model, outlined in Antonio et al. (2020), structures the logarithm of the force of mortality for Belgium, $\mu_{x,t}^{\text{BEL}}$, as follows

$$\begin{aligned} \ln \mu_{x,t}^{\text{BEL}} &= \ln \mu_{x,t}^{\text{EU}} + \ln \tilde{\mu}_{x,t}^{\text{BEL}} \\ \ln \mu_{x,t}^{\text{EU}} &= A_x + B_x K_t \\ \ln \tilde{\mu}_{x,t}^{\text{BEL}} &= \alpha_x + \beta_x \kappa_t. \end{aligned} \tag{1}$$

The dynamics of the common period effect, K_t, are modelled with a Random Walk with Drift ([RWD]). The Belgian period effect, κ_t, follows an AR(k) process with intercept.

Recalibration. When recalibrating the stochastic mortality model we try to stay as close as possible to the original model design and the specific choices made in IA|BE 2020. In line with IA|BE 2020 we use 1988 as the starting year of the calibration period.

Figure 8 shows the estimated European and Belgian period effects, together with a zoom-in on the relevant area. Except for the female Belgian period effect κ_t^F, there is no visible difference in the time dependent parameters for the years 1988-2019. In the year 2020, the European period effect K_t jumps upwards, which results in an overall increase in mortality rates as we will see later on. The largest upward jump appears for the most severe scenario S_3 from Table 1. In addition, a downward shift takes place for the male Belgian period effect κ_t^M. The opposite is true for the female κ_t^F. As the female process κ_t^F moves away from the origin in 2020, there is a wider gap between European and Belgian mortality rates. Figure 9 shows the estimated European and Belgian male and female age effects A_x, B_x, α_x and β_x for the three scenarios. The age effects calibrated along the three scenarios mainly coincide.

Mortality rates. Next, using the estimated Li-Lee parameters in Figures 8 and 9 and using the formulas in Equation (1), we can calculate mortality rates for Belgium:

$$q_{x,t}^{\text{BEL}} = 1 - \exp\left(-\mu_{x,t}^{\text{BEL}}\right).$$

Figures 10 and 11 show the calibrated Belgian mortality rates for ages 25, 45, 65 and 85. We only observe differences in the year 2020 for males, where the most severe death scenario causes the highest increase in mortality rates across all ages. For females, there are some minor differences observed in earlier years. The main reason why we get a counter-intuitive increase of the mortality rates at young ages (e.g. age 25) is due to the multiplication term $B_x \cdot K_t$ in Equation (1). Having a look at Figures 8 and 9, young ages are more punished for the increase in K_t in the year 2020 than older ages.

Projection. Using the calibration period 1988-2020, we have no guarantee to end up with a stable AR(1) process for the male and female Belgian period effect κ_t^M or κ_t^F for each impact scenario investigated in this note. We follow a similar approach as outlined in Appendix C of the IA|BE 2020 report and increase the lag size of either the male or female process until we attain stability. The resulting model choices are listed in Table 2 for each scenario.

Death scenario	Calibration period	TS model for K_t^M and K_t^F	TS model for κ_t^M	TS model for κ_t^F
Scenario 1	1988-2020	RWD	AR(1)	AR(1)
Scenario 2	1988-2020	RWD	AR(1)	AR(4)
Scenario 3	1988-2020	RWD	AR(1)	AR(4)

Table 2: Model specifications of the Li-Lee mortality model for each investigated scenario. RWD refers to the random walk with drift.

Period and cohort life expectancy. We are now ready to assess the effect of COVID-19 on the estimated period and cohort life expectancy in 2020 for males and females. Figure 12 compares the estimated and observed period life expectancies in the years 1988-2020 for the three impact scenarios in this note. In 2020, we under-estimate the period life expectancy of a 0-year old due to, among other things, an increase in estimated mortality rates at young ages (see Figures 10 and 11 and the paragraph about mortality rates). On the other hand, we observe an over-estimation of the period life expectancy in 2020 for a 65-year old. Table 3 then shows the estimated cohort life expectancies in 2020 for a 0 and 65 year old male and female. The period and cohort life expectancies in 2020 decrease due to the excess of mortality in 2020, caused by COVID-19.

Cohort life expectancy in 2020		Males		Females	
		0	65	0	65
IA|BE 2020	Best. Est.	89.91	20.38	91.54	23.14
	$[q_{0.5}; q_{50}; q_{99.5}]$	[88.11; 89.89; 91.46]	[19.57; 20.37; 21.17]	[89.46; 91.53; 93.25]	[22.15; 23.14; 24.07]
Scenario 1	Best. Est.	88.31	19.58	90.31	22.54
	$[q_{0.5}; q_{50}; q_{99.5}]$	[85.51; 88.30; 90.62]	[18.49; 19.58; 20.62]	[87.46; 90.31; 92.58]	[21.40; 22.54; 23.65]
Scenario 2	Best. Est.	87.98	19.43	90.19	22.37
	$[q_{0.5}; q_{50}; q_{99.5}]$	[84.73; 87.95; 90.47]	[18.28; 19.42; 20.62]	[87.11; 90.16; 92.53]	[21.13; 22.37; 23.60]
Scenario 3	Best. Est.	87.71	19.27	89.07	22.00
	$[q_{0.5}; q_{50}; q_{99.5}]$	[84.10; 87.71; 90.35]	[18.04; 19.27; 20.52]	[84.99; 89.08; 92.10]	[20.50; 21.99; 23.50]

Table 3: The cohort life expectancy for a 0 and 65 year old in 2020. The best estimate and the 0.5% quantile, median and 99.5% quantile obtained from 10 000 simulations are shown, for males and females and for each scenario.

5 Outlook

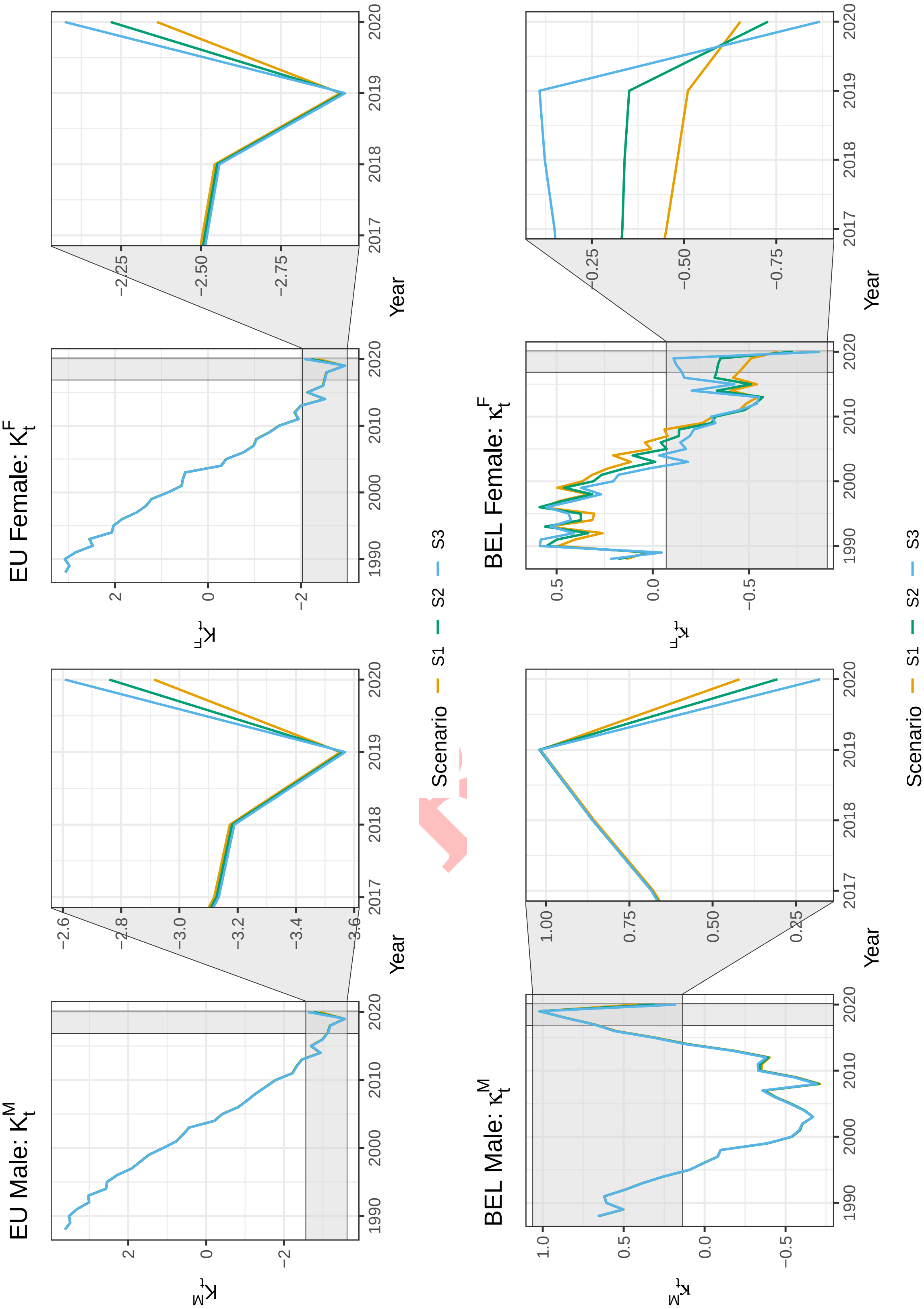


Figure 8: The calibrated European (top) and Belgian (bottom) period effects K_t^M, K_t^F, κ_t^M and κ_t^F for the three scenarios. Calibration period $1988 - 2020$.

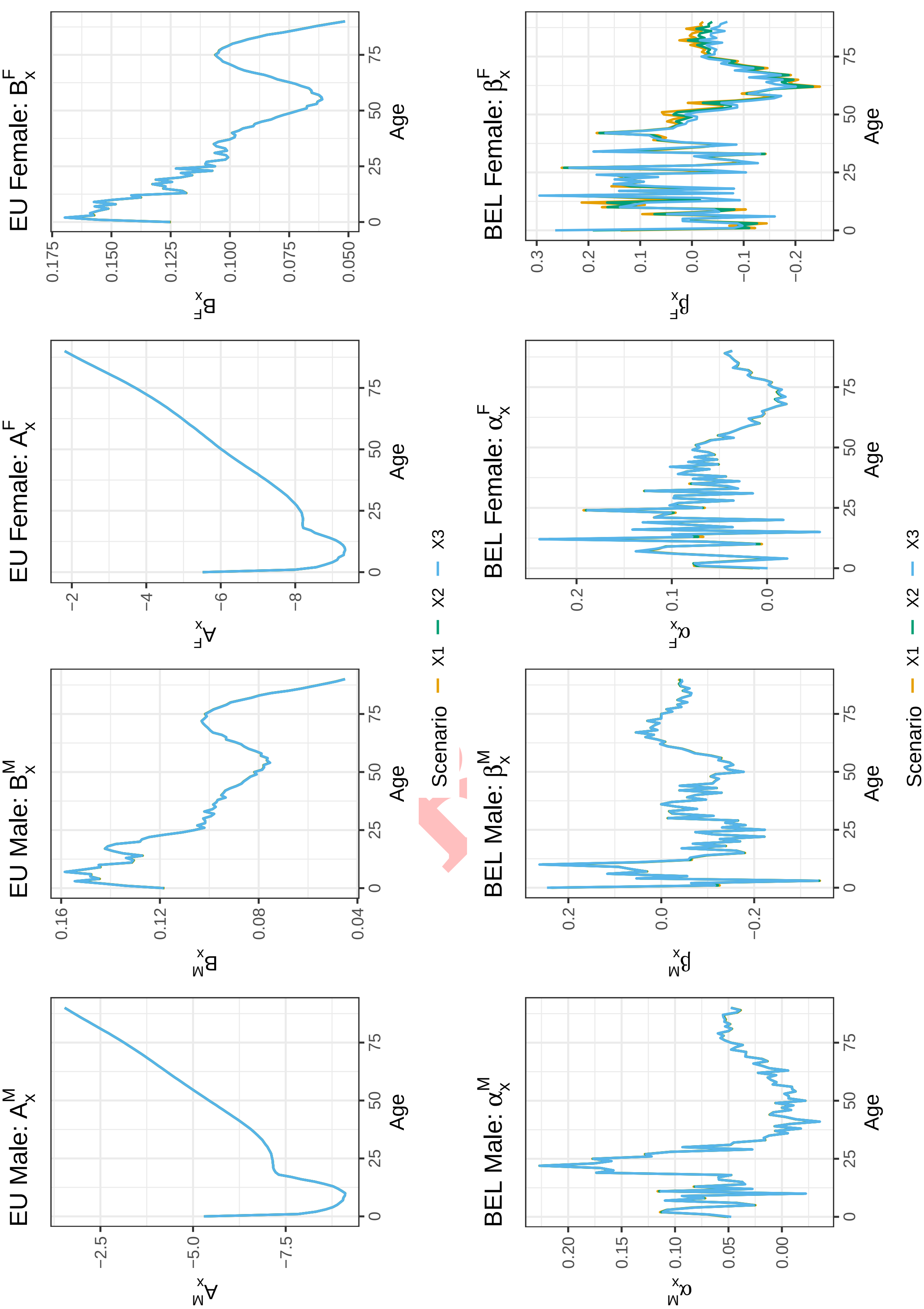


Figure 9: The calibrated European (top) and Belgian (bottom) male and female age effects A_x, B_x, α_x and β_x for the three scenarios. Calibration period $1988 - 2020$.

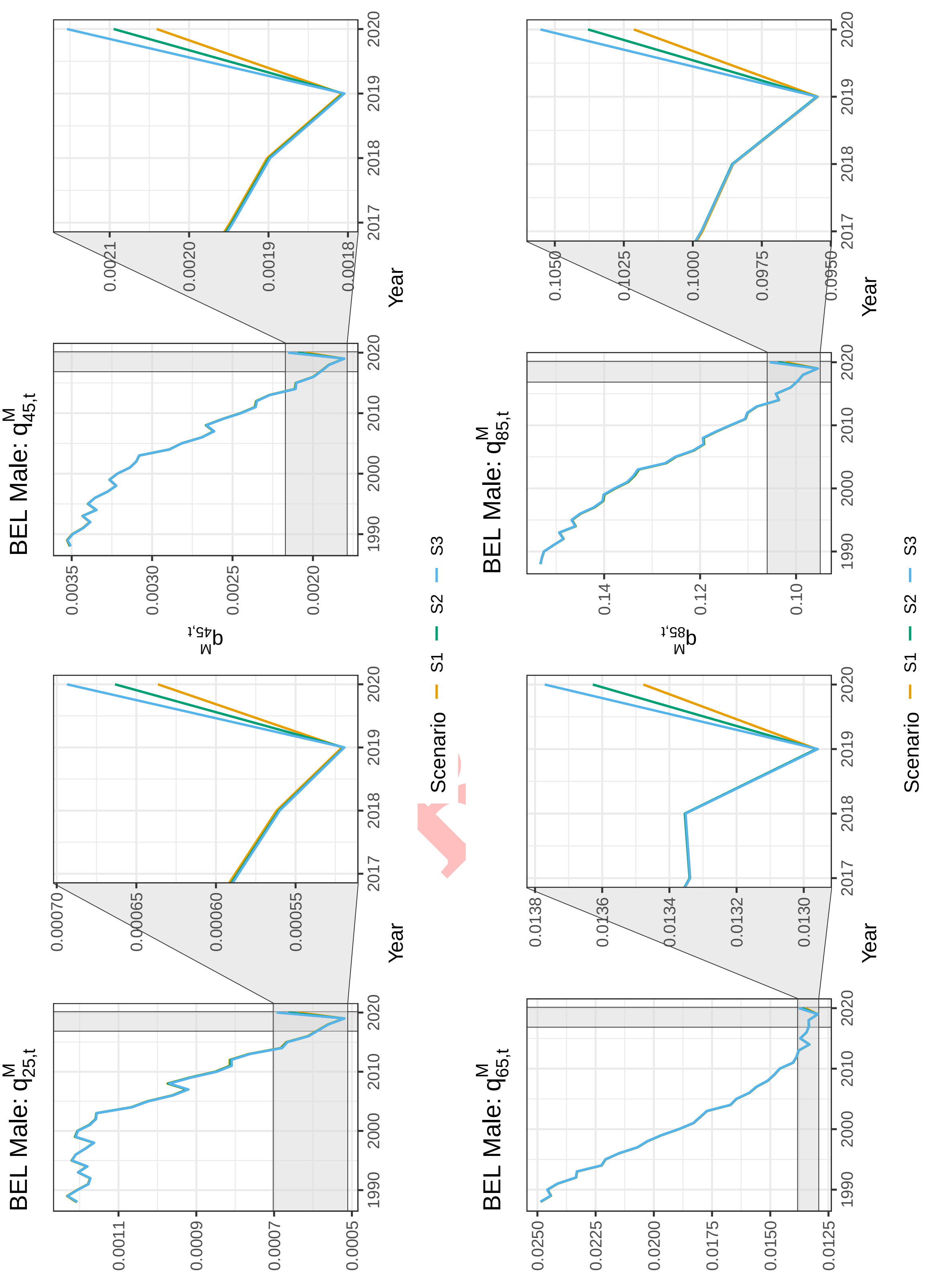


Figure 10: Estimated mortality rates $\hat{q}_{x,t}$ for Belgium, male data, ages 25,45 (top) and 65, 85 (bottom), calibration period 1988-2020.

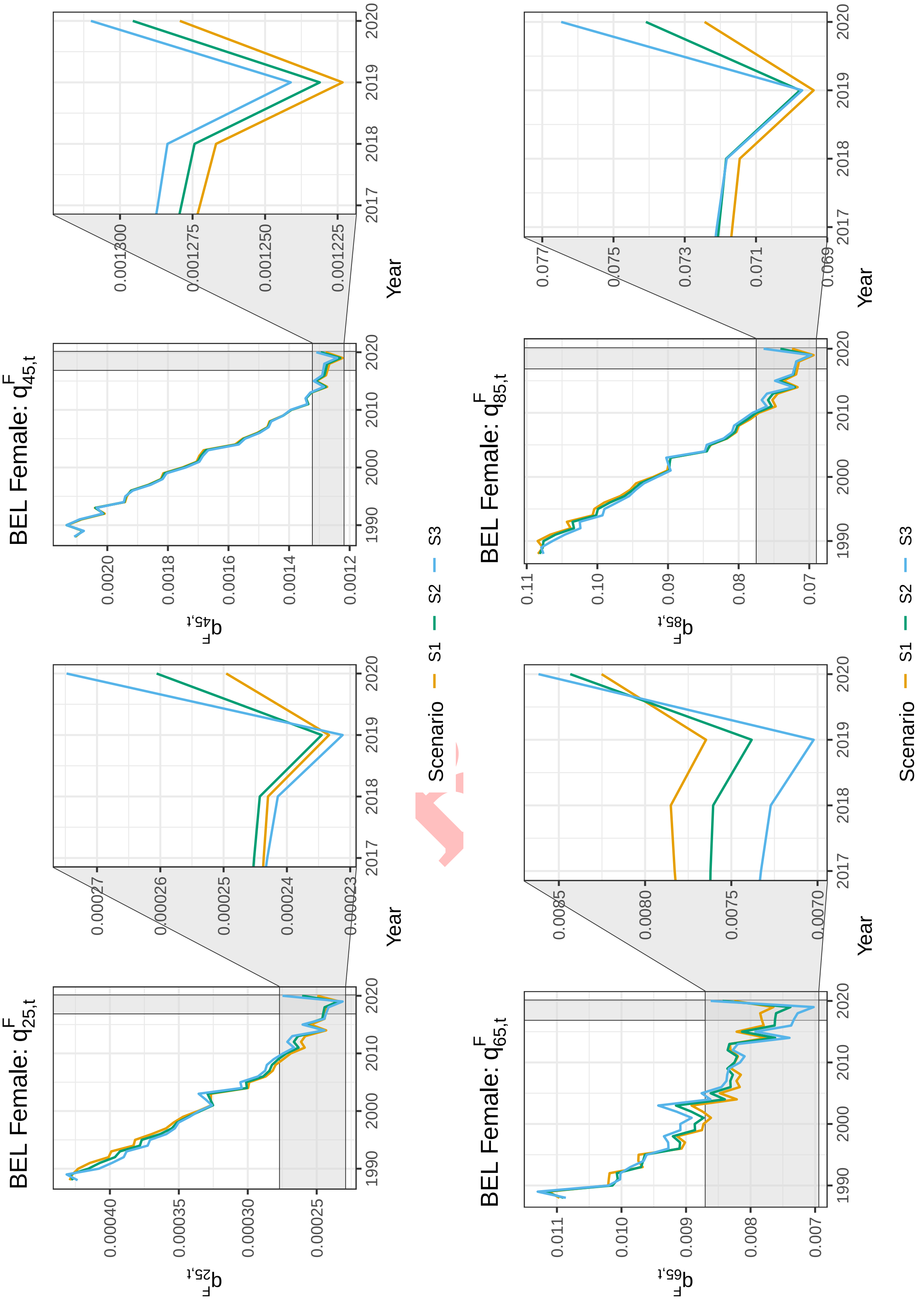


Figure 11: Estimated mortality rates $\hat{q}_{x,t}$ for Belgium, female data, ages 25,45 (top) and 65, 85 (bottom), calibration period 1988-2020.

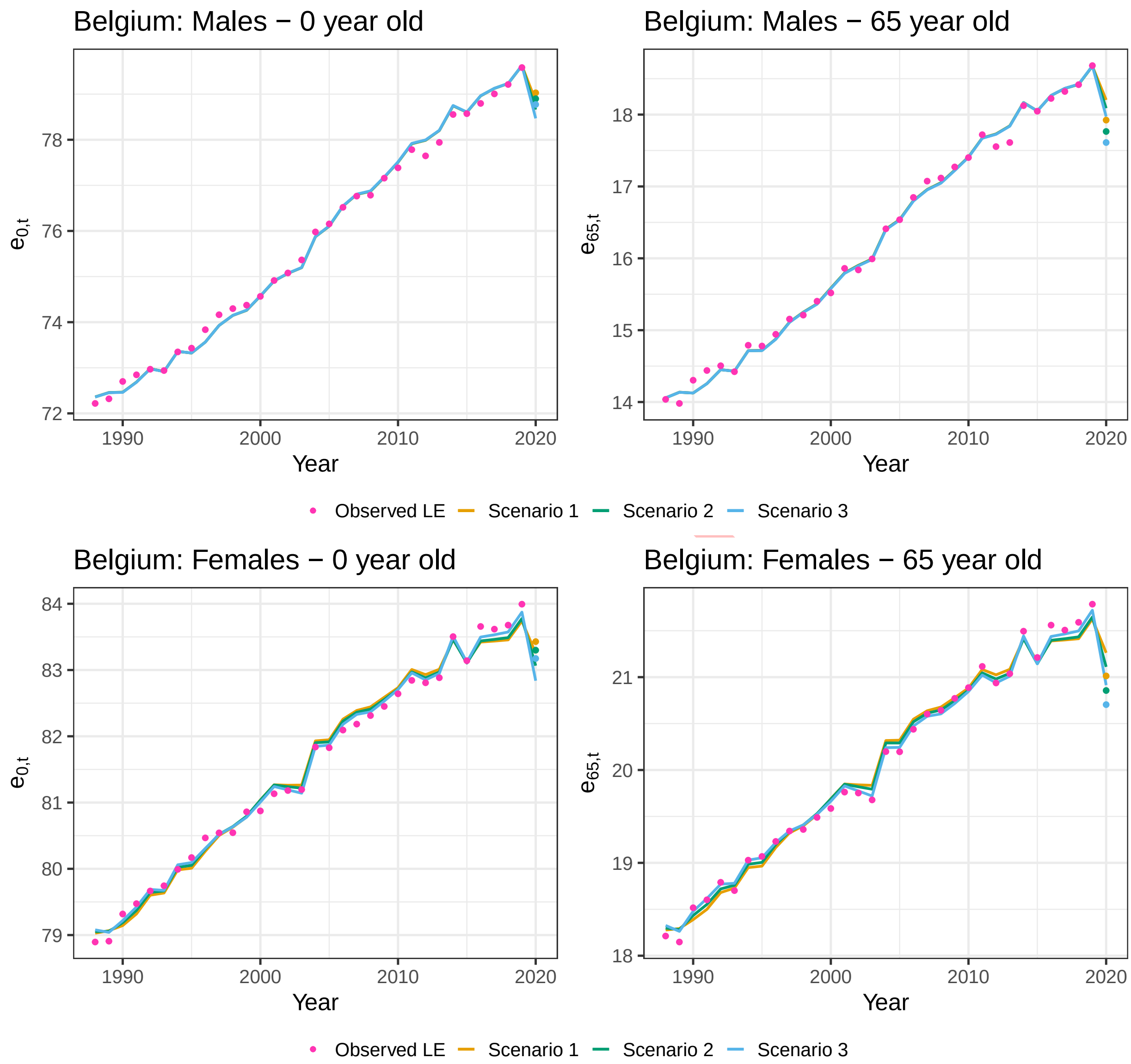


Figure 12: Period life expectancies for the three scenarios in the years 1988-2020. The dots visualize the observed life expectancies. The orange, green and blue lines illustrate the best-estimates of the period life expectancies.

References

Antonio, K., Devolder, L., and Devriendt, S. (2020). The IA|BE 2020 mortality projection for the Belgian population.

Duyck, J., Paul, J.-M., and Vandresse, M. (2020). Demografische vooruitzichten 2019-2070: Actualisering in het kader van de COVID-19-epidemie. https://www.plan.be/uploaded/documents/202006020559070.REP_POP1970Covid19_12154_N.pdf.

Koninklijk Actuarieel Genootschap (2018). Prognosetafel AG2018. https://www.ag-ai.nl.

Koninklijk Actuarieel Genootschap (2020). Prognosetafel AG2020. https://www.ag-ai.nl.

van Delft, L. and Huijzer, S. (2020). Impact of COVID-19 on Dutch mortality tables. https://be.milliman.com/-/media/milliman/pdfs/articles/impact-of-covid-19-on-dutch-mortality-tables.ashx.

A Data sources

	Exposures				**Deaths**			
Country	2017	2018	2019	2020	2017	2018	2019	2020
AUT	HMD	EURO	STMF	STMF	HMD	EURO	EURO.W	EURO.W
BEL	HMD	HMD	STAT	STMF	HMD	HMD	STAT	EURO.W
DNK	HMD	HMD	HMD	STMF	HMD	HMD	HMD	EURO.W
FIN	HMD	HMD	HMD	STMF	HMD	HMD	HMD	EURO.W
FRA	HMD	HMD	STMF	STMF	HMD	HMD	STMF	STMF
DEU	HMD	EURO	STMF	STMF	HMD	EURO	STMF	STMF
ISL	HMD	HMD	STMF	STMF	HMD	HMD	EURO.W	EURO.W
LUX	HMD	HMD	HMD	STMF	HMD	HMD	HMD	EURO.W
NLD	HMD	HMD	STMF	STMF	HMD	HMD	EURO.W	EURO.W
NOR	HMD	HMD	STMF	STMF	HMD	HMD	EURO.W	EURO.W
SWE	HMD	HMD	HMD	STMF	HMD	HMD	HMD	EURO.W
CHE	HMD	HMD	STMF	STMF	HMD	HMD	EURO.W	EURO.W
GBR	HMD	HMD	STMF	STMF	HMD	HMD	STMF	STMF

Table 4: Overview of the data sources for each country in the multi-population model. The data sources 'HMD', 'EURO' and 'STAT' are discussed in detail in Antonio et al. (2020). 'HMD' is the primary data source. We use the other two data sources to supplement the deaths and exposures with annual data for more recent years. 'STMF' refers to the Short-Term Mortality Fluctuations and 'EURO.W' to the weekly death statistics from Eurostat. Hence, the latter two data sources provide death statistics per week and for age buckets and are used to extend the data sets until the year 2020. We use 'EURO.W' as data source for the weekly death counts in 2019-2020 whenever they (more or less) correspond with those reported by STMF.

B Constructing virtual exposure points

We create virtual annual exposures $E_{x,t}$ for individual ages 0-90, years 2019-2020 and for each country in our set of European countries (Section 2).[11] We explain our strategy for creating virtual exposure points for Belgium, but we follow a similar approach for any other country in the multi-population model (see Table 4 in Appendix A).

Figure 13 shows the exposures in Belgium as a function of age over some recent years. The exposures given in this figure are observed data from HMD and Statbel (see Table 4), so no assumptions (e.g. migration) are behind these plots. The exposure function has a similar pattern shifted to the right with one age in each subsequent year t. This is in line with our intuitions since people aged x in year t become part of the group at risk aged $x+1$ in year $t+1$, in case of survival.

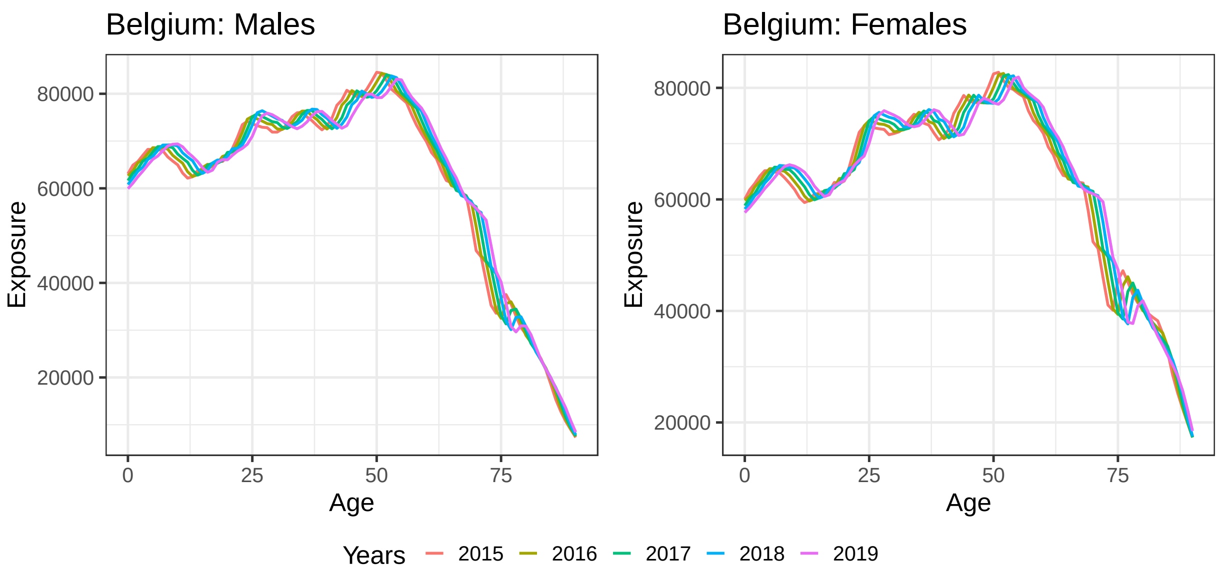


Figure 13: The exposures $E_{x,t}$ of Belgium for ages 0-90 and years 2015-2019. Data from HMD until the year 2018 and from Statbel for the year 2019.

We use this observation to create virtual annual exposures for 2020. Figure 14 graphically explains the strategy for males. First, we start from the given annual exposure $E_{x,t}$ in the most recent observation year available (in casu: 2019 for Belgium, using Statbel) as a function of age (red dashed line). Second, we shift this exposure curve of 2019 to the right with one age to end up in the orange dashed line. This newly created curve is undefined at the age zero because of the right shifting. Therefore, in a third step, we linearly extrapolate the orange dashed line to zero. This choice is justified by the linear pattern of the exposure function at young ages. This leads to the brown point in Figure 14. Lastly, we match the new exposure function of 2020 with the exposure numbers in age buckets we retrieved from the STMF, as provided in Table 5.

[11] For some countries, we do have the annual exposures $E_{x,t}$ at individual ages in 2019 from either HMD, Eurostat or Statbel (see Table 4). For these countries (Belgium, Denmark, Finland, Luxembourg and Sweden), there is no need to create a virtual exposure point in 2019.

Hereto, we consider an age bucket $[x_i, x_j]$ and define the virtual annual exposure $E_{x,t}$ as:

$$\begin{aligned} E_{x,t} &= \hat{E}^f_{x,t} \cdot b_{i,j} \\ b_{i,j} &= \frac{E_{[x_i,x_j],t}}{\sum\limits_{a=x_i}^{x_j} \hat{E}^f_{a,t}}, \end{aligned} \tag{2}$$

for $t = 2020$ (for example), $x \in [x_i, x_j]$ and where $\hat{E}^f_{x,t}$ denotes the fitted exposure at age x and time t as obtained from the orange dashed line, extended with the brown point at age 0. Intuitively, we vertically shift a section of the orange dashed line, corresponding to a certain age bucket, such that the summed exposure within this age bucket corresponds to the total exposure in the same age bucket of Table 5. The right panel of Figure 14 shows this strategy for the age bucket $[0, 14]$, where the purple line shows the final virtual exposure $E_{x,t}$ at individual ages for Belgium in the year 2020.

Age bucket	Male Exp.	Female Exp.
$[0, 14]$	988 713.02	944 379.40
$[15, 64]$	3 699 434.72	3 638 808.41
$[65, 74]$	568 101.96	618 244.99
$[75, 84]$	305 175.72	399 015.96
85+	112 577.56	223 565.55

Table 5: The male and female Belgian exposures in 2020 in age buckets, obtained from STMF.

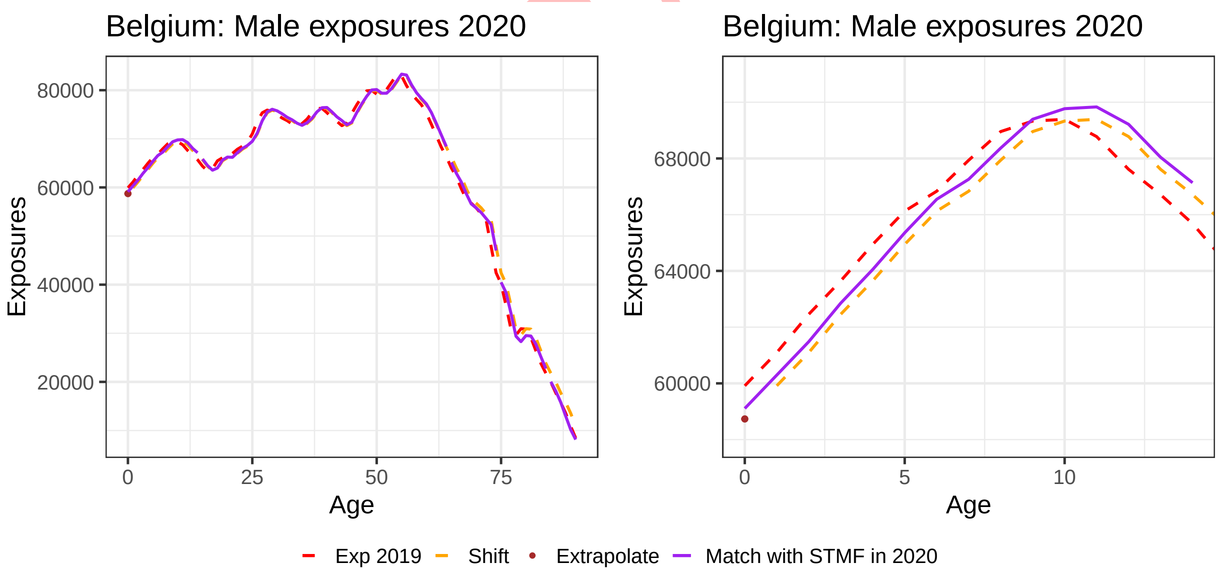


Figure 14: The virtual exposure $E_{x,t}$ for Belgium in the year 2020, for males. At the right, we show a snapshot for the age range 0-14.

We apply a slightly different strategy for the exposure $E_{85+,2020}$ reported for ages in the open age bucket 85+ on STMF. The underlying idea is that we want to distribute the extra exposure

$E_{85+,2020}$ (from STMF) $-$ $E_{85+,2018}$ (from HMD) evenly across the ages 85+, i.e. ages $85, \ldots, 110$ (the assumed maximum age).[12] Hereto, we calculate the shift c_{85+}, as follows

$$c_{85+} = \frac{1}{110 - 85 + 1} \left(E_{85+,2020} - E_{85+,2018} \right).$$

We then apply this shift c_{85+} to go from $E_{x,2018}$ to $E_{x,2020}$:

$$E_{x,2020} = E_{x,2018} + c_{85+},$$

for ages $x \in \{85, 86, \ldots, 90\}$.

We repeat this procedure for all 13 European countries. When there is no exposure data available for 2019 on the HMD (see Table 4 in Appendix A), we start from the exposure curve for 2018 reported on HMD or Eurostat and repeat the procedure two times to generate $E_{x,t}$ for years 2019-2020.

C Constructing virtual death counts

We construct virtual death counts $d_{x,t}$ at individual ages 0-90, years 2019-2020 for each of the 13 European countries.[13] We again focus in the explanation of our strategy on the construction of virtual deaths for Belgium in 2020, but a similar approach can be followed for any other European country.

Figure 15 shows the observed deaths in Belgium across ages 0-90 and over the years 2015-2019 for males (left) and females (right). In line with our discussion about the pattern of the exposure curve in Figure 13, we observe a time-effect in these death counts, e.g. the bumps in the deaths pattern move to the right each consecutive year. We must keep this in mind when constructing virtual deaths for 2020.

We propose the following strategy. Using the IA|BE 2020 model, we project the fitted force of mortality $\hat{\mu}^{\text{BEL}}_{x,t}$ into the future. Under the assumption of a piecewise constant force of mortality, we obtain

$$\hat{\mu}^{\text{BEL}}_{x,t} = m^{\text{BEL}}_{x,t} = \frac{d^{\text{BEL}}_{x,t}}{E^{\text{BEL}}_{x,t}}.$$

Since we created virtual annual exposure points for Belgium in the year 2020 in Appendix B, we can easily go from the force of mortality $\hat{\mu}^{\text{BEL}}_{x,2020}$ to virtual death counts $\hat{d}^{\text{BEL}}_{x,2020}$ according to the IA|BE 2020. In a next step we match these expected deaths $\hat{d}^{\text{BEL}}_{x,2020}$ with the information we retrieve from the weekly deaths data on Eurostat or STMF. For Belgium, we work with the weekly death counts in age buckets of length 5 from Eurostat, see Table 6.

Since 2020 is still ongoing, we do not have full information about the number of deaths in 2020. Hence, we first need to define some scenarios for the total number of deaths in 2020. We start from the number of deaths in the first 26 weeks of 2020 (see Table 6) and then extrapolate these deaths along some impact assessments of the second COVID-19 wave. We start from

[12] Here, we work with exposures in 2018 from HMD for simplicity reasons.

[13] For some countries, we do have the annual death counts $d_{x,t}$ at individual ages in 2019 from either HMD or Statbel (see Table 4). For these countries (Belgium, Denmark, Finland, Luxembourg and Sweden), there is no need to create virtual death counts in 2019.

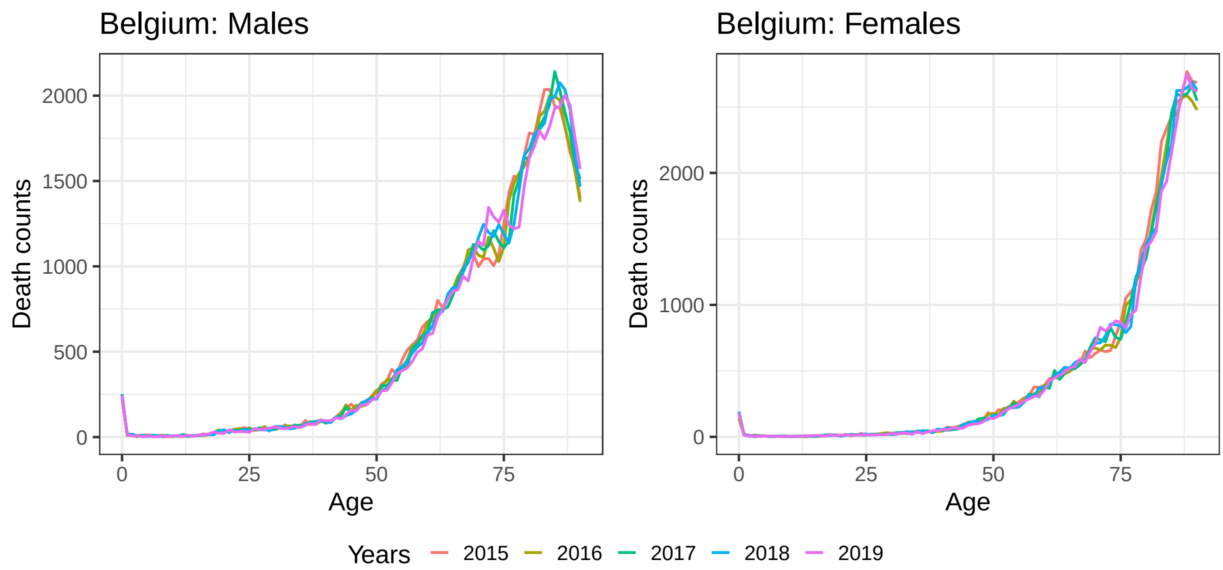


Figure 15: The deaths $d_{x,t}$ of Belgium for ages 0-90 and years 2015-2019. Data from HMD until the year 2018 and from Statbel for the year 2019.

Age bucket	Male deaths	Female deaths
$[0, 4]$	41	36
$[5, 9]$	16	9
$[10, 14]$	8	21
$[15, 19]$	43	22
...	...	...
$[75, 79]$	3 787	2 830
$[80, 84]$	5 326	5 028
$[85, 89]$	5 721	7 518
90+	4 797	10 668

Table 6: Male Belgian deaths in first half of 2020. Eurostat data.

two extreme scenarios, i.e. an unrealistic, optimistic scenario S_0, mainly defined for notation purposes, and a worst case scenario S_∞:

$$\hat{d}^{S_0}_{[x_i,x_j],2020} = \sum_{w=1}^{26} d_{[x_i,x_j],2020,w} + \left(\sum_{w=27}^{52} d_{[x_i,x_j],2019,w} \right) \cdot \frac{E_{[x_i,x_j],2020,w}}{E_{[x_i,x_j],2019,w}}$$

$$\hat{d}^{S_\infty}_{[x_i,x_j],2020} = 2 \cdot \sum_{w=1}^{26} d_{[x_i,x_j],2020,w}.$$

In scenario S_0, we act as if COVID-19 would have disappeared in the second half of 2020. The number of deaths in the last 26 weeks of 2020 is set equal to the number of deaths in the last 26 weeks of 2019, adapted to the current exposure. In the worst case scenario S_∞, we assume an equal number of deaths in the first and second half of 2020, causing a second COVID-19 of approximately the same size in the second half of 2020. From these two extreme scenarios, we construct three more plausible scenarios:

$$\begin{aligned}
\hat{d}^{S_1}_{[x_i,x_j],2020} &= \hat{d}^{S_0}_{[x_i,x_j],2020} + 50\% \cdot \left(\hat{d}^{S_\infty}_{[x_i,x_j],2020} - \hat{d}^{S_0}_{[x_i,x_j],2020}\right) \\
&= \frac{\hat{d}^{S_0}_{[x_i,x_j],2020} + \hat{d}^{S_\infty}_{[x_i,x_j],2020}}{2} \\
\hat{d}^{S_2}_{[x_i,x_j],2020} &= \hat{d}^{S_0}_{[x_i,x_j],2020} + 75\% \cdot \left(\hat{d}^{S_\infty}_{[x_i,x_j],2020} - \hat{d}^{S_0}_{[x_i,x_j],2020}\right) \\
&= \frac{\hat{d}^{S_0}_{[x_i,x_j],2020} + 3\hat{d}^{S_\infty}_{[x_i,x_j],2020}}{4} \\
\hat{d}^{S_3}_{[x_i,x_j],2020} &= \hat{d}^{S_0}_{[x_i,x_j],2020} + 100\% \cdot \left(\hat{d}^{S_\infty}_{[x_i,x_j],2020} - \hat{d}^{S_0}_{[x_i,x_j],2020}\right) \\
&= \hat{d}^{S_\infty}_{[x_i,x_j],2020}.
\end{aligned}$$

Starting from the scenario S_0, we hence add 50%, 75% or 100% of the excess of deaths in scenario S_∞ compared to scenario S_0. In these three scenarios, the number of deaths in the second COVID-19 wave thus roughly[14] corresponds to 50%, 75% or 100% of the number of deaths in the first COVID-19 wave respectively. Figure 16 shows the results for these three scenarios, where the death counts are still grouped in age buckets. In line with our expectations, we obtain the highest number of deaths in the worst case scenario S_∞. In the youngest age buckets, the opposite is true, possibly due to measures causing e.g. less deaths due to accidents.

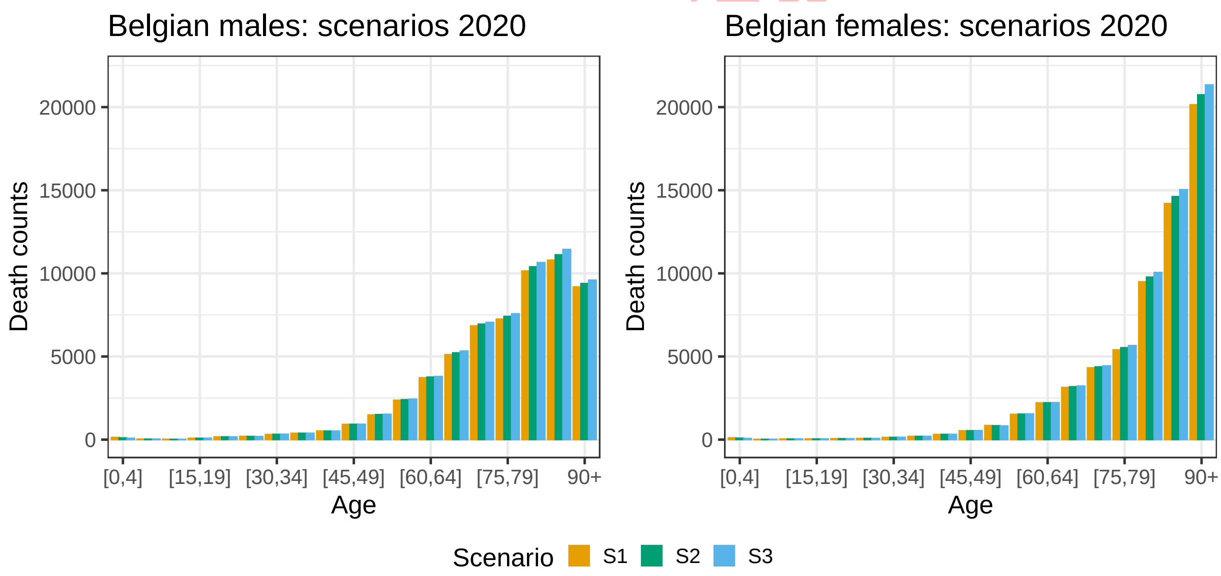


Figure 16: Scenarios for the total male Belgian deaths in 2020, based on the male deaths in the first 26 weeks of Eurostat data.

Having constructed the three death count scenarios, we now return to the construction of the virtual deaths in 2020 at individual ages 0-90. Figure 17 graphically explains this construction for scenario S_1 for males. The red line shows, just for visualisation purposes, the observed number of male deaths in Belgium for the year 2019. Using the IA|BE 2020 forecasting strategy for Belgium, we first project the force of mortality $\hat{\mu}^{\text{BEL}}_{x,t}$ for the year $t = 2020$. This corresponds

[14] Roughly, because history tells us that the number of deaths in the first 26 weeks of a particular year is clearly higher than those in the last 26 weeks of that same year, especially at the highest ages.

to the orange line in Figure 2. Just as with the exposures in Appendix B, we then match the orange death curve of 2020 with the first scenario for the death counts in 2020, as provided in Figure 16.

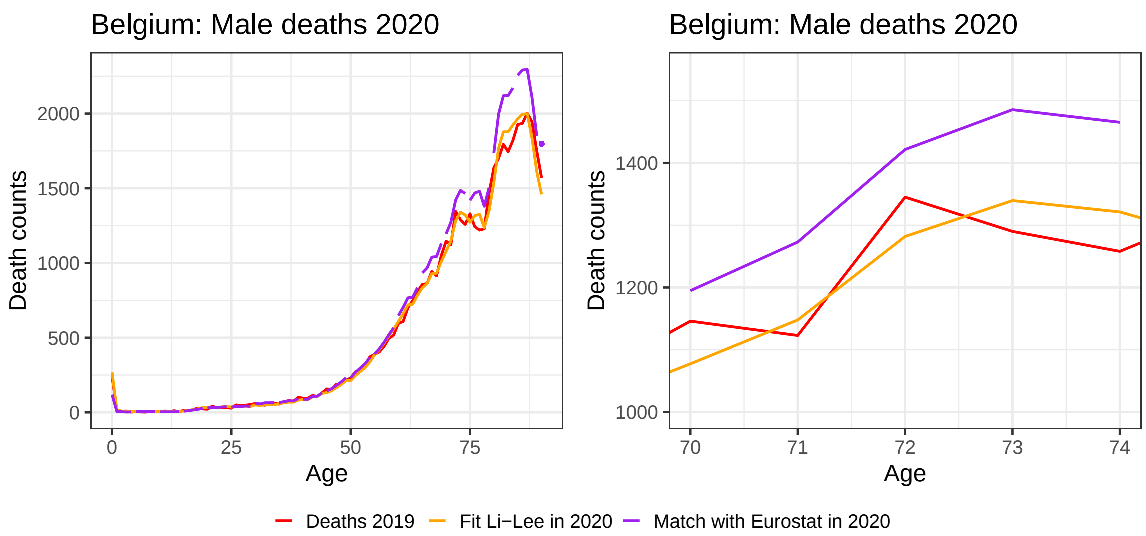


Figure 17: Construction of the virtual deaths in Belgium for 2020, according to the first scenario S_1.

We consider an age bucket $[x_i, x_j]$ and define the virtual annual deaths $d_{x,t}^{S_k}$ for scenario S_k, with $k \in \{1, 2, 3\}$, as:

$$\begin{aligned} d_{x,t}^{S_k} &= \hat{d}_{x,t}^{f} \cdot b_{i,j}^{S_k} \\ b_{i,j}^{S_k} &= \frac{d_{[x_i,x_j],t}^{S_k}}{\sum\limits_{a=x_i}^{x_j} \hat{d}_{a,t}^{f}}, \end{aligned} \tag{3}$$

for $t = 2020$, $x \in [x_i, x_j]$ and where $\hat{d}_{x,t}^{f}$ denotes the fitted deaths at age x and time t as obtained from the orange line in Figure 17. Intuitively, we again vertically shift a section of the orange line, corresponding to a certain age bucket, such that the combined number of deaths within this age bucket corresponds to the total number of deaths in the same age bucket of Figure 16. This matching principle results in the purple line in Figure 17 for the first scenario S_1. The right panel of Figure 17 shows, by means of example, the results for the age bucket $[70, 74]$. We repeat this procedure for each scenario, leading to three scenarios for the deaths $d_{x,t}$, now evaluated at individual ages.

The last age bucket 90+ is again an open age bucket, implying a modification of the strategy outlined in Equation (3) to define the virtual death count at age 90 in 2020 for each death scenario S_k:

$$\begin{aligned} d_{90,2020}^{S_k} &= d_{90,2018} + c_{90+}^{S_k} \\ c_{90+}^{S_k} &= A^g \left(d_{90+,2020}^{S_k} - \sum_{a=90}^{\infty} d_{a,2018} \right), \end{aligned}$$

where A^g is a gender-specific rate, which we assume to be scenario-independent. For example, $A^g = 0.20$ means that 20% of the deaths, at ages 90 or higher, occur at age 90. Based on historical values,[15] we select $A^m = 0.18$ and $A^f = 0.14$ for males and females respectively.

We repeat this procedure for every European country in the study with the following country-specific data adjustments. For some of the countries, we do not have the deaths $d_{x,t}$ at time $t = 2019$ yet. In those cases, we construct a Li-Lee mortality model for the country of interest with a shorter calibration period, ending with the year 2018. In addition, each country has its own start year of the calibration period for stability reasons, e.g. the year 1988 for Belgium. We can then construct death counts for the year 2019 and 2020 (for each scenario) by projecting the force of mortality for the years 2019-2020 and by performing the matching principle at both years.

Moreover, for three of the European countries, namely Germany, France and the United Kingdom, we work with the weekly death counts in age buckets from the STMF, rather than from Eurostat. For these countries, we apply the strategy outlined above but work with larger age buckets.

[15] We look at the ratio of European deaths at age 90 to the European deaths at ages 90 or higher, in the last year of the HMD data. We take a European average to end up with a robust, stable choice.