Received: 14 June 2023 | Revised: 26 April 2024 | Accepted: 14 May 2024

DOI: 10.1002/asmb.2876

RESEARCH ARTICLE

WILEY

Gender-inclusive financial and demographic literacy: Monetizing the gender mortality gap

Giovanna Apicella[1] | **Enrico De Giorgi**[2] | **Emilia Di Lorenzo**[3] | **Marilena Sibillo**[4]

[1]Department of Economics and Statistics, University of Udine, Udine, Italy

[2]Department of Economics, School of Economics and Political Science, St. Gallen, Switzerland

[3]Department of Economic and Statistical Sciences, University of Naples Federico II, Naples, Italy

[4]Department of Economics and Statistics, University of Salerno, Fisciano (SA), Italy

Correspondence
Enrico De Giorgi, Department of Economics, School of Economics and Political Science, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland.
Email: enrico.degiorgi@unisg.ch

Abstract

Longevity crucially affects demand for pensions, insurance products and annuities. Consistent empirical evidence shows that women have historically experienced lower mortality rates than men. In this article, we study a measure of the gender gap in mortality rates, we call “Gender Gap Ratio”, across a wide range of ages and for four countries: France, Italy, Sweden, and USA. We show the stylized facts that characterize the trend of the Gender Gap Ratio, both in its historical evolution and future projection. Focusing on an example temporary life annuity contract, we give a monetary consistency to the Gender Gap Ratio. We show evidence that a Gender Gap Ratio that ranges between 1.5 and 2.5, depending on age, translates into a significant reduction of up to 23% in the benefits from a temporary life annuity contract for women with respect to men, against the same amount invested in the life annuity. The empirical evidence discussed in this article documents the crucial importance of working toward a more widespread demographic literacy, for example, a range of tools and strategies to raise longevity consciousness among individuals and policy-makers, in the framework of gender equality policies.

KEYWORDS

gender gap in mortality, financial well-being, demographic literacy

1 | INTRODUCTION

The progressive aging of the population, in all its connotations and implications, engenders complex political, economic and social challenges. The United Nations General Assembly singled out the 2021–2030 decade as “the decade of healthy aging.”[1] National governments, as well as supranational bodies, will thus be tasked with implementing complementary policies under WHO guidance. The overarching aims are ambitious, given that particular initiatives aiming at improving the living conditions of the elderly and tackling inequalities will have to be complemented with a deeper cultural revolution as regards intergenerational equity and active aging. Combined, these will aim at enabling longevity to prevail senescence.

The numbers speak by themselves: worldwide, there will be 1.4 billion (or 16.6% of the world’s population) people over-60s by 2030 and up to 2.1 billion (or 22% of the world’s population) by 2050.[2] In Europe alone[3] the number of over 65s in 2060 should be double that of 2008 (in percentage terms it will increase from 17.3% in 2008 to expected 30.3% in 2060); likewise, in Europe, by the same year, the number of ‘oldest old’, that is, individuals over the age of 80, is expected to increase by approximately 40 millions (in percentage terms it will increase from 4.4% in 2008 to expected 12.5% in 2060).

It has long been clear that these demographic patterns call for significant changes in social structures, public health organization, infrastructure, and care processes, as well as the active and cooperative inclusion of older age groups in a cohesive and inclusive society.

Among the various issues linked with longevity and, in a broader sense, lifecycle, it is particularly interesting to investigate how pension systems currently respond to these critical patterns and what transformations ought necessarily to ensue. Of particular interest in this context is the study of the factors that influence the maintenance of an adequate standard of quality of life in old age so as to have a better perception of financial and insurance products in their actual profiles and to be able to evaluate possible improvement and adaptation to the actual living conditions of individuals (e.g., long term care policies, home pensions).

It is precisely in this context that one encounters a problem within the problem: gender inequality. It is, in fact, no coincidence that the 2030 Agenda for Sustainable Development[1] attributed critical importance to this theme; more specifically, as indicated by the 2021 Aging and Health Report drawn up by the WHO, the Sustainable Development Goals (SDGs) are largely concerned with the protection of health and well-being throughout the life cycle. The fifth SDG, however, places particular emphasis on gender equality, in regarding the latter a driving force underpinning the fullfilment of further SDGs.

Going back to social security aspects, it is important to examine the numbers revealing the magnitude of the gender gap, as these are indicative of the disparities that affect women in specific age groups. Yet, these numbers are symptomatic of much more complex realities.

The OECD Report on Retirement saving outcomes for women[4] demonstrates that women in OECD countries receive an average pension that is 26% lower than that of men. This gap stems primarily from rates of employment, wage differences, women's social and welfare organization, as well as cultural factors that also affect financial literacy.

The evolution of the gender gap in life expectancy over time has been studied from various perspectives: biological, medical-health, geographical, ethnic, and socio-political. Time after time, the results open research horizons that are contextualized in several biological, behavioral, and social frameworks; in the industrialized economies the main themes are the silver economy, the protection of vulnerable population groups and the mitigation of inequalities. Research over the last two decades has shown that male mortality rates are higher than female mortality rates; however, women suffer from lack of physical strength and disability. The phenomenon is well known as male-female health-survival paradox.[5] The male-female health survival paradox is also recognized by recent research, that takes into account activities of daily living (ADL) and instrumental activities of daily living (IADL).[6]

The trend of life duration, together with the state of health and disability, has changed in the light of progressive medical, health and economic advances; the systemic aging of industrialized societies has led researchers to deepen not only state-partial specific life expectancies, but also health expectancies.[7]

Regarding the gender gap, many studies highlight the importance of the women's levels of education, the involvement in the labor market, the marital status.[8] Relevant differences by sex depend on racial/ethnic disparities, as well as educational health disparities[9,10] and the impact of income throughout life.[11]

Just about the latter topic, Schünemann et al.[12] investigate how the gender gap declines with rising income; consequently, differences in survival lead to economic and financial differences, for instance the present discounted value (EDPV) of the payout from a fair annuity, that varies by "gender, health, and level of education."[13]

In this article, we describe the historical dynamics of the gender differences in mortality and study their projections in the future. We use a synthetic indicator (say Gender Gap Ratio, that is, the ratio between the male mortality rate and the female one) that allows firstly to analyze, in an immediate and intuitive way, the past experiences (the last seven decades) up to the present historical context. In addition, we forecast the future trend of the gender gap ratio, after discussing the class of the stochastic process representative of the general survival trend. Interesting stylized facts emerge that characterize the trend of the gender gap, both in its historical evolution and future projection. Finally, focusing on an example temporary life annuity contract, we give a monetary consistency to the gender gap by quantifying its impact in the realization of the financial-demographic balances between an invested sum and the constant periodic amount due in case of life to which this investment gives rise, differentiating the male case from the female case. We show that the difference in mortality rates between males and females, or the evidence we provide about a Gender Gap Ratio that ranges between 1.5 and 2.5 depending on age and country, translates into a significant reduction of up to 25% in the benefits from a temporary life annuity contract. This result highlights the economic relevance of understanding the future evolution of the Gender Gap Ratio. Indeed, a quantitative description of the gender mortality gap's evolution has an informative and educational value, whose nature is both economic and social. More importantly, it enables a better-informed approach to the promotion of gender equality via policies and practices. In this context, in the final Section, we illustrate the most recent literature explaining how the well-known financial literacy and the still not thoroughly addressed demographic literacy can enhance individuals' awareness and decisions, especially in relation to retirement behavior. Our article adds empirical evidence to these research streams and, in particular, fosters the adoption of a more gender-inclusive perspective in the implementation of social development policies targeting individuals' well-being in old age.

2 | GENDER-INCLUSIVE FINANCIAL AND DEMOGRAPHIC LITERACY: LITERATURE AND RELEVANCE

Assessing the evolution of the gender gap in mortality plays an important role for policymakers and financial institutions devising strategies to cope with the economic, social and financial impact of such a phenomenon.

In compliance with the Council Directive 2004/113/EC[14] and the Ruling of the Court of Justice of the European Union (Case C-236/09), the use of gender as a factor in the calculation of insurance premiums and benefits in relation to insurance contracts is illegal, since 21 December 2012, in all the Member States. As explained in European Commission,[15] this rule represents an important step toward clarifying the fundamental right of gender equality under EU law. In OECD,[4] in relation to the policy interventions advocated to narrow the gender gap in pensions, it is stressed that the design of the pay-out phase may impact on the level of retirement income received by men and women. In this respect, women's higher longevity risk is pivotal. As we also show in the remainder of this article, accounting for gender as a pricing factor implies that, against a certain level of assets accrued at retirement, a woman will receive lower periodic annuity payments than a man, although for longer, thus balancing the pension wealth between the two genders. By contrast, using unisex mortality tables for pricing, involves, for the same level of assets at retirement, the same level of annuity payments every year for men and for women, although received for a likely longer time horizon by women. The prospective endurance of the gender gap in longevity thus poses a challenge, related to, inter alia, the cost and the welfare effects of unisex pricing. Schmeiser et al.[16] discuss several implications of gender-neutral pricing for the insurance industry and customers, for example, negative outcomes such as adverse selection effects and market distortions, that require interventions and changes on the part of insurance companies to be faced. These interventions are inhrent to the business strategy and the risk management of insurance companies, for example, toward an equitable spread of risks in their portfolio. Wealth redistribution is also a welfare implication of unisex pricing. Finkelstein et al.[17] shows, with respect to the UK market for compulsory retirement annuities, a redistribution of more than 3% of retirement wealth from men to women and argue that voluntary markets may be affected by amplified welfare consequences. Bruszas et al.[18] measure the disadvantages, for males, of unisex pricing with respect to German participating life annuities, by considering a lifetime utility that accounts for stochastic mortality and the total annuity payments over the entire life span. Based on this approach, when the perspective is extended from a single point in time to the life duration, the men's disadvantages resulting from unisex pricing are substantially lower than the disadvantage observed in the empirical market data. Indeed, the gender longevity gap brings with it a long-term risk, with financial consequences. To soundly assess the effectiveness of how such a risk is tackled within the retirement income provision landscape, it is fundamental to take a stochastic approach, allowing to catch the dynamics of this phenomenon and to reliably portray its future evolution, as shown in the remainder of our article. Reliable estimations of the prospective gender gap in mortality are also useful to enhance individuals' awareness about the cost of the insurance protection, in terms of economic value and fairness, and about its long-term benefits. Schmeiser et al.[16] use survey data concerning five European countries to assess the consumer's degree of acceptance of gender-specific price differences in relation to four insurance products, including health insurance, term life insurance and annuities. The results point out that using gender as a pricing factor within the business lines health, annuity and term life insurance is not accepted by consumers as soon as they can compare, side-by-side, the magnitude of the two gender-specific premiums, involving a statistical assessment of gender-related mortality patterns. Actuarial information about longevity risk can thus be used to induce reflection on the ethical and social value entailed by unisex pricing and the underlying mechanism of cross-subsidization between women and men, and to increase women's participation in the annuity market, given their objectively high survival prospects.

As stressed by OECD,[4] also the design of communication and financial education strategies, especially targeting women (e.g., Reference 19), concretely helps people take action to enhance their retirement readiness. Education strategies should entail information making individuals aware of the main long-term risks they are faced with when planning for their retirement, such as financial and longevity risks and their interactions.

There is by now vast evidence showing that individuals' financial literacy is very low even in advanced economies,[20] despite the importance of possessing a solid knowledge of basic financial concepts to make well-informed economic decisions that affect individuals' life in many respects. The positive impact brought by financial literacy have been documented in several studies, for example, in relation to pension planning, individuals who are more financially literate are also more likely to plan for their retirement[21–23] and display a higher propensity to save.[24] The significant effect of financial education on financial knowledge and financial behavior has also been empirically validated.[25,26] An important stylized fact regarding financial literacy is that women do generally possess lower financial knowledge compared to men,[20,27,28] the so-called gender gap in financial literary. Interestingly, Bucher-Koenen et al.[29] show that around one third of women's gap in financial literacy can be attributed to women's lack of confidence in answering the questionnaire designed to measure financial literacy; see also Aristei and Gallo,[30] Tinghög et al.,[31] and Driva et al.[32] on the role of gender stereotypes in relation to household finance matters.

The empirical evidence on individuals' demographic literacy is more limited. Hamermesh,[33] Hurd and McGarry[34] and Perozek[35] show that subjective beliefs on survival differ from actuarial estimates. Specifically, Hamermesh[33] reports evidence that people extrapolate changing life tables when they determine their subjective estimates, but the resulting subjective distribution is flatter and has greater variance than its actuarial counterpart. Similarly, Hurd and McGarry[34] and Perozek[35] document a gap in survival beliefs where younger individuals aged 50 to 70 tend to underestimate their survival probabilities compared to the actuarial counterparts, while older individuals aged 70 and more tend to overestimate them. Perozek[35] illustrates a gender gap in survival beliefs, as in their sample female participants tend to underestimate survival, while male participants tend to overestimate it. Recent research adopts a behavioral perspective to study the drivers of longevity perception and few papers shed light on the gender connotation of such a phenomenon referring to behavioral biases as over- and under-optimism, or over- and under-confidence (see, e.g., References 36–38). It is indeed consistently demonstrated by the literature that survival expectations (even if biased) affect forward-looking economic behavior, for example, saving and investing for retirement.[39,40] As emphasized in Apicella and De Giorgi,[41] besides the gender gap in financial literacy (also partially explained referring to under-confidence and stereotypes), the gender gap in longevity risk perception may contribute to explain why women tend to be less financially prepared for retirement than men.

Overall, the empirical evidence on gender differences concerning financial and demographic literacy further highlights the relevance of understanding the future development of mortality patterns, for example, forecasting the Gender Gap Ratio and its economic implications. The COVID-19 pandemic provides additional arguments to urgently address gender differences in mortality patterns, as women survival rates have been higher in relation to COVID-19[42–44] and to epidemics in general.[45] Hurwitz et al.[46] show that providing longevity risk and life expectancy information impact individuals' financial decisions, while longevity risk information also affects subjective assessments of survival probability. Angelici et al.[47] report on the effectiveness of tutorials targeted to women, explicitly addressing, for example, the gender gap.

The empirical evidence discussed in this article documents the crucial importance of working toward a more widespread demographic literacy, meant as a range of tools and strategies to raise longevity consciousness among individuals and policy-makers, in the framework of gender mainstreaming and gender equality policies. Indeed, we have shown that differing longevity patterns in men and women can lead to different financial outcomes for them. Accordingly, a reliable measure of the expected gender gap in longevity can help governments and financial institutions not only when designing reforms and innovative products but also when introducing such policies to potential customers, who may lack longevity awareness.

3 | DATA

Our analysis benchmarks four countries: France (FR), Italy (I), Sweden (SE), and the USA (USA). Based on the "Global Gender Gap Index" (World Economic Forum[48]), these nations achieve similar results with respect to the gender gap in Health and Survival, but perform quite differently in other socio-economic domains of the gender gap. Accordingly, analyzing these countries allows to detect some "stylized facts" in the evolution of the gender longevity gap, net of the country-specific socio-economic and political environment. For the chosen countries we have available from the Human Mortality Database[49] a wide range of reliable mortality data across time and ages. We exploit data until 2019, being the last year of observation common to all countries under consideration at the time of this study. We benchmark five representative ages (from 45 to 85, being 10-year apart), so that to build knowledge around the gender gap in survival at those steps in the lifecycle when consciousness is needed the most (e.g., in the context of pension schemes, health care policies).

4 | NOTATION AND MATHEMATICAL FOUNDATION

We use notations and methods from life insurance mathematics (see, e.g., Reference 50) and follow an age-period approach to analyze changes in mortality as a function of both age and time. For age $x \geq 0$ and time $t \geq 0$, we denote by T_0 the random lifetime of an individual aged 0 at time $t - x = 0$. The conditional probability

$$ {}_hq_{x,t} := P_t(T_0 \leq x + h | T_0 > x) \tag{1}$$

is the probability that the individual dies before or at age $x + h$, for $h > 0$, conditional on surviving age x. Similarly, the probability

$$ {}_hp_{x,t} := P_t(T_0 > x + h | T_0 > x) = 1 -_h q_{x,t}$$

is the probability that the individual survives age $x + h$ conditional on surviving age x. For any integer $h > 1$, ${}_hp_{x,t}$ corresponds to Reference 50:

$$ {}_hp_{x,t} = p_{x,t} \cdot p_{x+1,t+1} \cdots p_{x+h-1,t+h-1}, \tag{2}$$

where $p_{x+h-i,t+h-i} :=_1 p_{x+h-i,t+h-i}$ is the one-year survival probability for the individual aged $x + h - i$ at time $t + h - i$ for $i = 1, \ldots, h$.

The force of mortality $\mu_{x,t}$ is the instantaneous rate of mortality at a time t, that is:

$$\mu_{x,t} := \lim_{h \to 0} \frac{{}_hq_{x,t}}{h}. \tag{3}$$

If h is sufficiently small, we can write:

$$ {}_hq_{x,t} \simeq \mu_{x,t}\, h. \tag{4}$$

In the actuarial practice, it is frequently assumed that the age-specific forces of mortality are constant within bands of age and time, but can vary from one band to the next one, that is:

$$\mu_{x+\xi_1,t+\xi_2} = \mu_{x,t}, \quad for \quad 0 \leq \xi_1, \xi_2 < 1. \tag{5}$$

Under Assumption (5):

- $\mu_{x,t}$ coincides with the crude death rate $m_{x,t}$, that is, the expected number of deaths divided by the exposure-to-risk, defined as the total number of "person-years" in a population over a calendar year;
- the one-year survival probability corresponds to $p_{x,t} = \exp(-m_{x,t})$.

It follows that the mortality rate $q_{x,t} := 1 - p_{x,t}$ corresponds to

$$q_{x,t} = 1 - \exp(-m_{x,t}).$$

The observed death rate $m_{x,t}$ is commonly computed as the ratio:

$$m_{x,t} = \frac{d_{x,t}}{E^c_{x,t}}, \tag{6}$$

where d_{xt} is the observed number of deaths at age x last birthday during calendar year t, while E^c_{xt} is the so-called central exposure-to-risk (the average size of the population aged x last birthday during year t).

5 | THE MORTALITY MODEL AND THE GENDER GAP RATIO

The main objective of our study is to use well-established quantitative methods to describe the gender mortality gap, with a particular emphasis on some crucial economic implications of such a phenomenon, as the impact of unisex pricing of annuities on the redistribution of retirement wealth. Indeed, annuities play a pivotal role for protection from the financial consequences of longevity risk, being, on average, higher for women than for men. To be able to inform policy-making, it is important to make use of modeling frameworks that are largely deployed by practitioners in the insurance sector, statistical offices and decision-makers, who are tasked with mortality projections, risk management, and economic planning. Within our study, we use the Lee–Carter (LC) model as in Lee and Carter[51] to fit and forecast mortality for the male and female populations of our interest. As stressed by Basellini et al.,[52] national and international statistical offices, as well as private sector practitioners and academics, widely make use of the LC method and its extensions to project mortality. In this respect, Bergeron-Boucher and Kjærgaard[53] mention, as example cases, the national statistical offices of Canada, Denmark, Italy and Sweden, which use the original Lee–Carter model as proposed by Lee and Carter[51] or alternative models to the LC, such as its variants (e.g., References 54–56). Nevertheless, it is shown that although such alternatives may increase accuracy and robustness, no single approach performs best for all countries or at all ages. The state-of-the-art includes several other extensions to the Lee–Carter method. For instance, Pedroza[57] restated the original Lee–Carter method under a Bayesian framework to reflect more accurately the forecasting error associated with the model. Furthermore, in the last two decades, several studies have focused on coherent mortality forecasting ensuring the non-divergence of mortality levels of closely related populations in the long run (e.g., References 56,58,59).

Our study aims to uncover how the patterns of female and male mortality rates related to each other over an extended time horizon, through a data-driven approach. We do not make any prior assumption about the extent by which the mortality rates of male and female populations should move in line in the long run. Furthermore, we do not constrain male and female age-specific death rates to maintain a constant ratio to one another. Our target is, instead, to shed light on the features and the magnitude of such a ratio, as both revealed by the empirical evidence and captured by the smoothed mortality surfaces for each gender generated by mortality models. To this aim, we apply the Lee–Carter stochastic model to the male and female populations individually and use the extrapolated mortality trends of such populations and their projected mortality patterns to infer information about the historical and future gender gap in mortality.

As a model validation procedure, we demonstrate that the Lee–Carter model is suitable for our intended purposes, also when evaluated under rigorous quantitative criteria and based on the data sets of our interest. We compare the performance of the LC (or M1) model against two competitive models within the family of Generalized Age-Period-Cohort (GAPC) mortality models: the Cairns–Blake–Dowd model (CBD, or M5, see Reference 60) and the M8 model.[61] The respective predictor structures are described in Table 1, according to the notation in Cairns et al.[61]

In Table 1, $\bar{x}$ is the mean age in the sample range and x_c a constant parameter to be estimated. Furthermore, the functions $\beta_x^{(i)}, k_t^{(i)}$ (for $i = 1, 2$) and γ_{t-x} are age, period, and cohort parameters, respectively. Under model M1, the change in the general level of mortality over time is described by a univariate time factor $k_t^{(2)}$. M5 model includes two time-varying parameters, $k_t^{(1)}$ and $k_t^{(2)}$, allowing to capture the imperfect correlation in mortality rates at different ages. M8 further incorporates one cohort parameter γ_{t-x}.

The state-of-the-art sets out quantitative frameworks for systematic comparison across mortality projection models, tested on a variety of data sets and over different age ranges (e.g., References 61–64). Methodologies to establish the quantitative goodness of fit include, for instance, the evaluation of information criteria balancing quality of fit and parsimony, and the assessment of the mortality residuals, namely the differences between the realized mortality rates for some given ages and calendar years and their model-generated counterparts. The backtesting of mortality models involves, but is not limited to, the study of the behavior of the ex post forecast errors. The underlying rationale is that a good stochastic mortality model should be relatively simple. Furthermore, a good mortality model, inter alia, should be able

TABLE 1 Predictor structures of models M1, M5, and M8.

Model	Predictor structure
M1	logit $q_{x,t} = \beta_x^{(1)} + \beta_x^{(2)} k_t^{(2)}$
M5	logit $q_{x,t} = k_t^{(1)} + k_t^{(2)}(x - \bar{x})$
M8	logit $q_{x,t} = k_t^{(1)} + k_t^{(2)}(x - \bar{x}) + \gamma_{t-x}(x_c - x)$

to produce parameter estimates and predictions that turn out to be plausible, robust with respect to the period of data and age intervals considered, and consistent with the historical trends and volatility of the mortality data. We thus set out a dynamic framework to ascertain the relative goodness of fit, predictive accuracy and robustness of the LC model, compared to M5 and M8, for different periods of data and ages. Backtesting methods involving dynamic analyses were developed, for instance, in Apicella et al.[65] We implement the backtesting procedure exploiting fixed-length windows of data rolling one-year-ahead through time. In this way, we can verify if the recalibration of the mortality models provide updated parameters being robust over time and able to catch the dynamics of the underlying data. While the look-back windows are used to calibrate the mortality models and assess the quality of fit, the look-forward windows are used to perform out-of-sample tests of the models' predictive accuracy. We arrange the time interval [1948, 2019] in such a way to exploit 40 years of data in the look-back window and 20 years of data in the look-forward one. The first look-back window is made up of the years from 1948 to 1987, and the last one includes years from 1960 to 1999. Overall, we obtain 13 time horizons to be used for the implementation of the backtesting methodology on the data relative to the male and female populations of France, Italy, Sweden and USA. We fit the three stochastic mortality models in the **R** software (https://www.r-project.org/) through the *StMoMo* package.[66,67] The StMoMo function builds an object representing each specified model, based on information on the link function, the predictor structure and the set of parameter constraints. In all cases, we target the one-year mortality rate $q_{x,t}$ and assume that the random numbers of deaths $D_{x,t}$ are independent and follow a Binomial distribution (B) conditionally on $(q_{x,t})$, so that:

$$D_{x,t} \sim B(E^0_{x,t}, q_{x,t}). \tag{7}$$

This implies the use of a logit link function (see, e.g., Reference 68). Parameter estimates are obtained by maximizing the model log-likelihood function, by means of the *fit* function, whose inputs are matrices of deaths and exposures with the integer ages of interest $x = 18, \dots, 90$ on the rows and the 40 calendar years included in the look-back window on the columns. Since M8 is a cohort model, we exclude all cohorts having fewer than five observations, as in Cairns et al.[61]

We use the Bayesian Information Criterion (BIC)[69] to evaluate the in-sample fitting performance and the parsimony of model M1, compared to models M5 and M8. Indeed, as stressed by Cairns et al.,[61] Haberman and Renshaw[64] and following studies, BIC in one the most deployed penalty functions allowing an assessment of the maximum likelihood informed by the number of parameters characterizing the mortality models. We obtain the BIC related to the calibration of each model on the 13 look-back windows under study. Such a computation can easily be implemented within the StMoMo environment through the generic function *BIC*, that provides the following measure[67]:

$$BIC = \nu \; ln(K) - 2 \; \mathcal{L}, \tag{8}$$

where ν represents the effective number of estimated parameters, $\mathcal{L}$ the estimate of the maximum log likelihood and K the number of observations (not counting those cells that have been removed from the analysis). Accordingly, a lower value of BIC is preferable. In total, considering the four countries and the two genders under study, we examine 104 values of BIC for each model. Impressingly, model M1 scores consistently better than the other two models on each look-back window, for all the countries considered and for each gender. To ease the interpretation of findings, we provide a synthetic measure of the BIC for each case. In Table 2, we thus report the average BIC over all look-back windows, in relation to each country (on the rows) and each model (on the columns), distinguished by gender. On average, the BIC characterizing model M1 is always lower than that of M5 and M8 models. In this respect, we also notice the significant advantage of

TABLE 2 Fitting accuracy of models M1, M5, and M8, by gender and country.

Country	Females			Males		
	M1	**M5**	**M8**	**M1**	**M5**	**M8**
FR	**35,437**	311,791	49,911	**43,721**	175,936	62,433
I	**37,731**	199,652	40,205	**50,781**	164,583	79,987
SE	**25,383**	48,220	25,919	**26,841**	48,846	30,762
USA	**58,995**	353,140	79,868	**71,221**	622,225	182,250

Note: Average Bayesian Information Criterion over 13 fixed-length look-back windows (40 years of data).

capturing age-related effects through latent components. Our results are important to validate model M1 as a proper tool to obtain smoothed mortality surfaces by gender, being consistent with the data, throughout the post-World War II time horizon (e.g., to gain insights into the historical trend of the gender mortality gap).

As a next step, we perform the forecasting of models M1, M5, and M8 over the considered look-forward windows. The first look-forward window includes calendar years from 1988 to 2007, while the last one spans from 2000 to 2019. The forecasting function in the StMoMo package allows to obtain matrices of the central projections of the mortality rates $q(x, t)$, from which we compute the corresponding forecasts of central death rates $m(x, t)$. The forecasting procedure involves estimating and forecasting the time-series models describing the dynamics of the period indexes characterizing M1, M5 and M8 and of the cohort index that is additionally included in the predictor structure of M8 model. In all cases, for the period effects we fit a multivariate random walk with drift,[70] while for the cohort effect of M8 model, we estimate an ARIMA(2,0,0) model, as in previous studies (e.g., Reference 71). We assess the forecasting performance of the three models over the considered 13 look-forward windows. Starting from the relative forecasting errors between the realized crude death rates and the model-generated counterparts (that is, projections), we compute the root mean square errors (RMSEs) of each model, with respect to the seven integers ages, being ten-years apart, in the interval [25, 85]. The RMSE is commonly used as a statistical measure of the goodness of an estimator or predictor, also with respect to the performance of mortality models (e.g., References 65,72,73). Overall, accounting for both genders along with all the considered ages, countries and look-forward windows, we have 728 cases where to compare the magnitude of the RMSEs across models M1, M5 and M8. The evidence based on the RMSE points to the fact that none of the models perform the best for all ages. Furthermore, the age being fixed, in some cases, no model consistently dominates the others across the look-forward windows. Nevertheless, if we combine all evidence in relation to the diverse ages and prediction time spans under consideration, we can draw conclusions about the model that is more likely to return the lowest RMSE for each country and gender, by counting the number of the occurrences in favor of such a model out of the relative total. We report such evidence in Table 3, in relation to each country (on the rows) and each model (on the columns), distinguished by gender. Except for French and Italian males, it is more frequently observed that the lowest RMSE is associated to model M1. Altogether, as shown by the last row of Table 3, in 46.23% of the 728 total cases, model M1 provides the most accurate forecast in the set of the models under study (M5 model: 19.23%, M8 model: 34.34%). The domain of the lowest RMSEs range from 1.6 to 35 (unit 10^{-2}), over the total number of examined cases. By way of an example, in Figure A1 of Appendix A, we show the magnitude of the RMSE of models M1, M5 and M8, for the forecasts related to the Italian female population over the lookforward window [1998–2017], across the various ages under study.

The outcomes of the validation procedure confirm that the Lee–Carter model is suitable for the fitting and forecasting purposes of this study, also when evaluated under rigorous quantitative criteria and based on the data sets of our interest.

The two main objectives of this study are to detect general, systematic patterns of the evolution of the gender gap in mortality and to monetize the implications of such a phenomenon in the annuity framework. The core variable of our analysis is thus a measure of the gender gap in mortality. By virtue of its immediate and simple interpretability, we select the "Gender Gap Ratio" (GGR), that is the ratio between male and female death rates. Formally, the observed GGR, for any given age x and calendar year t, is as follows:

$$GGR_{x,t} = m^M_{x,t} / m^F_{x,t}, \tag{9}$$

TABLE 3 Forecasting accuracy of models M1, M5, and M8, by gender and country.

	Females			Males			
Country	**M1**	**M5**	**M8**	**M1**	**M5**	**M8**	
FR	**47**	0	44	23	**37**	31	
I	**63**	1	27	35	19	**37**	
SE	**44**	11	36	**36**	33	22	
USA	**39**	15	37	**51**	24	16	
Total	193	27	144	145	113	106	728

Note: Number of cases in which the RMSE measure is lower for each model compared to the others. Seven ages and 13 look-forward windows considered, for a total of 728 cases.

where $m^M_{x,t}$ is the observed death rate for a male aged x in year t and $m^F_{x,t}$ is the observed death rate for a peer female in the same year. We denote by $\widehat{GGR}_{x,t}$ the GGR based on the estimates of male and female death rates provided by model M1, $\hat{m}^M_{x,t}$ and $\hat{m}^F_{x,t}$. We denote by $\widetilde{GGR}_{x,t}$ the GGR obtained from male and female death rates projections through model M1.

In Sections 6–8, we discuss stylized facts that characterize the trend of the GGR, both in its historical evolution and future projection. In Section 9, we monetize the consequences of such the GGR in the annuity framework.

6 | STYLIZED FACTS: EVIDENCE FROM THE PAST

In this section, we perform a cross-country analysis of the Gender Gap Ratio trends through the time horizon [1948–2019], for France, Italy, Sweden, and USA. We describe the GGR evolution over time by means of general and systematic patterns and track the country-specific dynamics of this gap. In this respect, model M1 generates a smoothed mortality surface for each gender, since it provides an estimate $\hat{m}^M_{x,t}$ and $\hat{m}^F_{x,t}$ for every age x and calendar year t under study. Relating such mortality surfaces leads to smoothed curves of GGR, $\widehat{GGR}_{x,t}$, that describe how GGR has evolved over time. We focus on the five integer ages, being ten years apart, in the range [45, 85] and target robust estimates of GGR that can reflect the change in the underlying data over a time horizon spanning 72 calendar years. To this aim, for each population and gender, we calibrate model M1 over 40-years fitting samples, rolling one-year ahead through time, according to the same reasoning underpinning the dynamic validation procedure described in Section 5. In this case, we can exploit a higher number of time horizons for calibration, since we do not need to reserve data for the model backtesting. Indeed, we have available 33 fitting samples: the first one covers years from 1948 to 1987 and the last one spans years from 1980 to 2019. Each fitting sample, compared to the previous one, is made up of the same years except for the first and the last one. It results that all the years within the span (1948, 2019), apart from the starting and the ending one, are used more than once by the fitting procedure. Accordingly, for each year we obtain at least one and at most 33 corresponding estimates of the mortality rate. The "best estimate" of the mortality rate associated to each calendar year between 1948 and 2019 is thus obtained by averaging all the fitted mortality rates for that year. Such best estimate reflects better mortality dynamics, since these are captured by the dynamic reoptimization of model M1. This procedure is implemented with respect to the female and male populations of each country under study. The best estimate of GGR is obtained accordingly, by the ratio of the male and female best estimates associated to the same age and calendar year.

To uncover more distinctly features of the GGR evolution by age, we compute 5-years moving averages of $\widehat{GGR}_{x,t}$, based on its best estimates. This approach allows to detect more sharply the historical periods when the smoothed Gender Gap Ratio has attained its highest values, in relation to each age. We refer to these averaged values of $\widehat{GGR}_{x,t}$ as $\overline{GGR}_{x,(t-4:t)}$. In Figure 1, we show four graphs, one per country, displaying $\overline{GGR}_{x,(t-4:t)}$ as a function of calendar year t, for ages $x = 45, 55, 65, 75, 85$. In Table 4, we report the maximum value for $\overline{GGR}_{x,(t-4:t)}$ and the respective calendar year t at which such a value is attained, under a country-age perspective.

Despite some cross-country differences, stylized features of GGR can be identified, especially in relation to its age profile. A first phenomenon, we could call the "Gender Gap Ratio expansion", is quite consistently verified for the four countries under study. It consists in the shifting to the right (toward more recent calendar years) of the maximum value attained by $\overline{GGR}_{x,(t-4:t)}$ as age x advances. France is generally characterized by higher values of GGR compared to the other countries under study. As shown in Table 4, in some cases, the four countries have experienced a different timing in attaining the highest values of their GGR. For instance, our estimates suggest that the maximum GGR for age 85 in the USA is temporally placed around three decades in advance compared to France. Except for age 45, the behavior of GGR in the USA has anticipated what occurred progressively later in the other countries, starting with Sweden. Smoothed curves of GGR disclose a general decline in its magnitude in the recent decades with somehow evident exceptions for ages 55 and 65 in the USA and mild evidence in this sense for ages 45 and 55 in Sweden. It is worth noting that analyzing averaged estimates of GGR in place of raw data, while revealing the overall historical trends of GGR free from noise, prevents from catching sharply the very recent behavior of the underlying data. For instance, raw data show a trend inversion in the GGR behavior (from decreasing to increasing) quite contextually in all countries (roughly after 2010) (see Figure B2 of Appendix B). This finding is consistent with the empirical evidence provided by Zarulli et al.[74] and may arise from the positive effects of a steady decline in breast cancer mortality rates that women aged more than 40 have undergone since 1989 in both Europe and the USA.[75,76] We also highlight that raw data reveal a higher relative change over time of female and male death rates for Sweden, compared to the other countries, that may arise from its much smaller volumes of exposures and deaths. Accordingly, GGR unsmoothed data reveal much more marked year-by-year variations in its

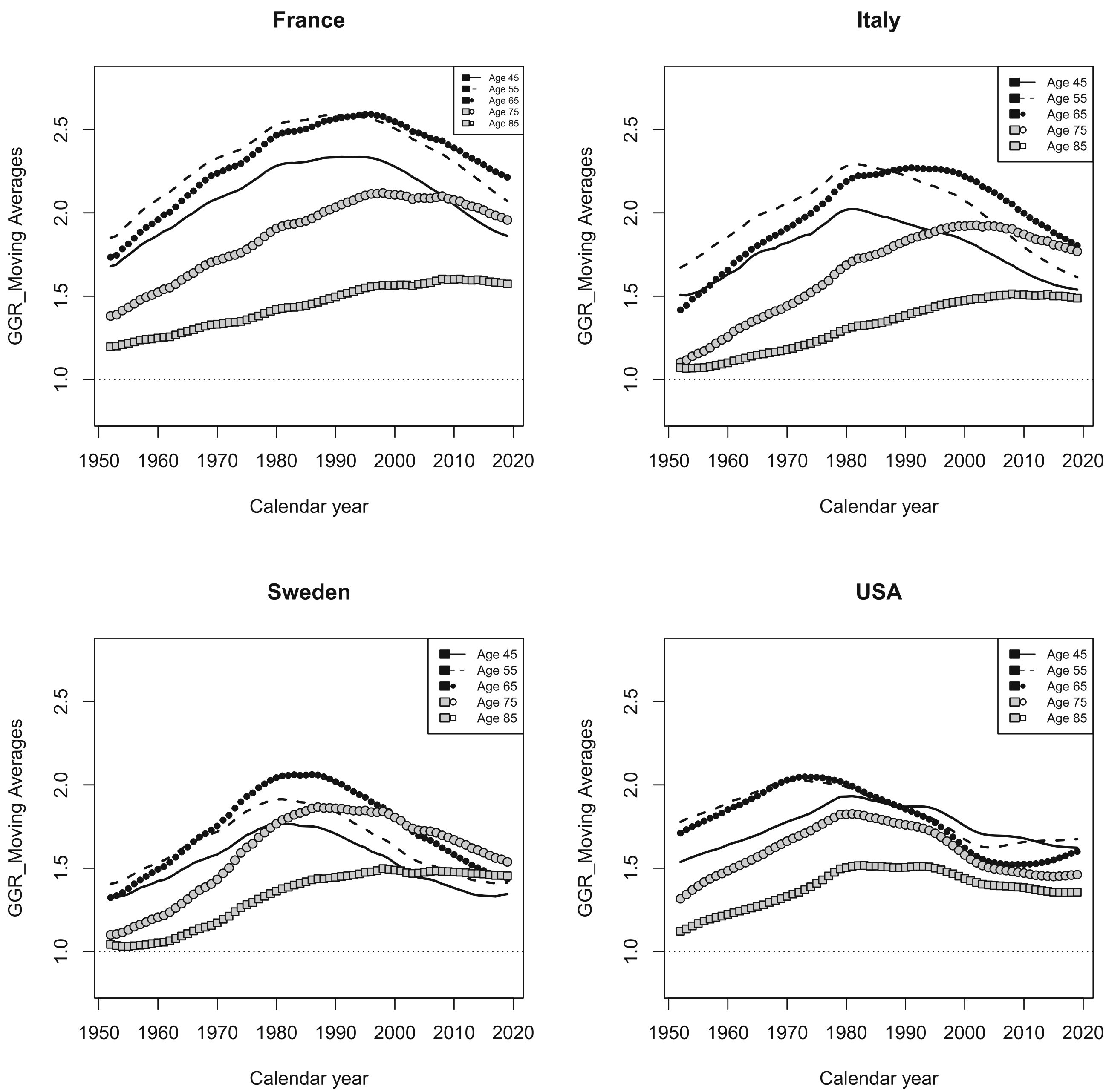

FIGURE 1 Moving averages (5-years) for the Gender Gap Ratio. Such averages are computed on the best estimates of GGR from dynamic reoptimization of model M1. (Calendar year in [1948, 2019]. Ages 45 (black solid line), 55 (black dashed line), 65 (black dotted line), 75 (gray dotted line), 85 (gray squared line). Countries: France, Italy, Sweden, and USA.)

TABLE 4 The table reports evidence on the calendar years where the gender gap ratio reached its maximum values for ages 45, 55, 65, 75, and 85.

	Country (ISO abbreviation)							
	FR		I		SE		USA	
Age	Max $\overline{GGR}$	*t*	Max $\overline{GGR}$	t	Max $\overline{GGR}$	*t*	Max $\overline{GGR}$	*t*
45	2.34	1991	2.01	1981	1.77	1981	1.93	1981
55	2.59	1989	2.29	1982	1.91	1981	2.03	1972
65	2.59	1996	2.27	1991	2.06	1986	2.05	1973
75	2.12	1998	1.93	2002	1.87	1987	1.83	1981
85	1.60	2011	1.51	2008	1.50	1998	1.52	1983

Note: These maxima relate to the best estimates of GGR from model M1 and the 5-year moving averages based on such estimates.

magnitude for Sweden than for the other countries under study, this implying larger values of its relative change over time. We provide the related evidence in Figure B1 of Appendix B.

7 | WHAT TO EXPECT IN THE FUTURE

In this Section, we discuss our findings about the projected trends of GGR over the span 2020-2039. We calibrate model M1 to the mortality data relative to the male and female populations of France, Italy, Sweden and USA, from 1980 to 2019, thus exploiting 40 years of data. We forecast model M1 for the following 20 years, until 2039. For each age, we compute the ratio of male to female forecasted death rates, this giving us the corresponding projection of GGR, $\widetilde{GGR}_{x,t}$ associated to a given future year t. In Figure 2, we show for each country, with respect to the ages under study, both the estimated path of

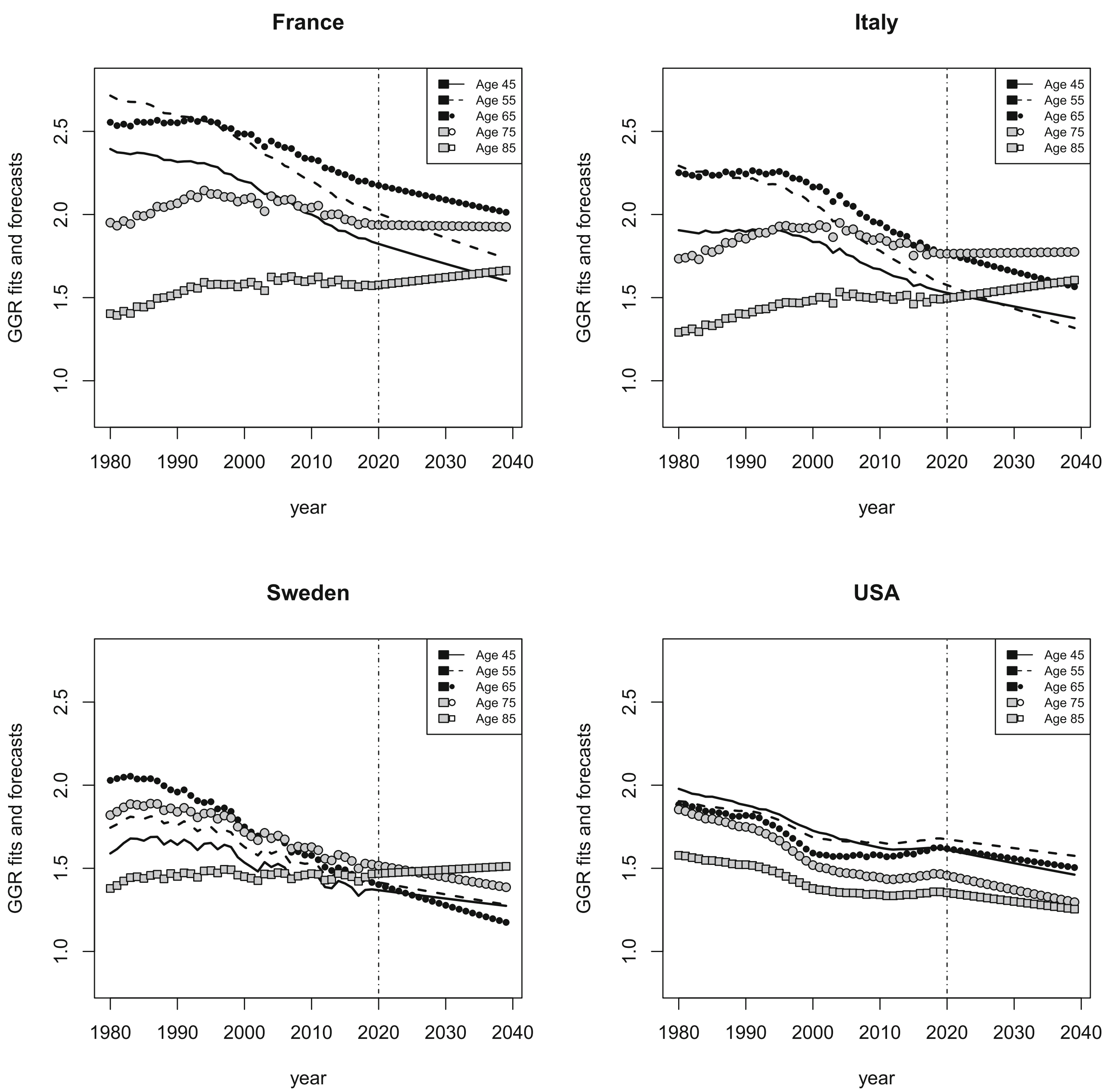

FIGURE 2 Fitted and forecasted path of GGR. (Calendar years in [1980,2039]. 2020 marks the start of the forecasting period. Ages 45 (black solid line), 55 (black dashed line), 65 (black dotted line), 75 (gray dotted line), 85 (gray squared line). Countries: France, Italy, Sweden, and USA.)

GGR, $\widehat{GGR}_{x,t}$, for $t = 1980, \ldots, 2019$, and its forecast, $\widetilde{GGR}_{x,t}$, over the next 20 years. A vertical line marks the beginning of the forecasting period. We use the black solid line for age 45, the black dashed line for age 55, the black dotted line for age 65, while we associate the gray dotted line to age 75 and the gray squared line to age 85.

Our findings point to some commonalities between countries, with respect to the age pattern of the projected GGR. This is the case for France and Italy, which show a predicted declining behavior in GGR for ages 45, 55, and 65, at odds with the mildly increasing trend predicted for age 75 and the steeper raise expected for age 85. For the USA, based on our forecasts, within 2039 it is expected the convergence in GGR at ages 75 and 85. A general decreasing trend of GGR is found for Sweden, with the only exception of age 85. Previous literature shows that mortality differences at older ages increasingly drive the gender longevity gap (e.g., Reference 77).

8 | MAKING SENSE OF THE COVID-19 IMPACT ON THE GENDER GAP RATIO: SOME ATTEMPTS

In the previous section, we have obtained forecasts of female and male mortality rates over the time period [2020, 2039], based on the probabilistic structure of model M1 and on the past realizations of death rates until 2019. After this year, the entire world has been confronted with the COVID-19 pandemic disease, that has emphasized sex disparities in mortality rates, as shown by empirical evidence (see, e.g., References 42 and 78). In this section, we try to make sense of if, and eventually how much, accounting for COVID-19-informed past mortality data changes the magnitude of the Gender Gap Ratio expected in the future. We restrict our analysis to Sweden, being the only country for which we have available mortality data up to 2021, at the time of this study. For each gender, we thus calibrate model M1 on the 40 most recent data points with respect to 2021 and forecast it from 2022 to 2039. In Figure 3, with respect to the ages under study, we show both the estimated path of GGR, $\widehat{GGR}_{x,t}$, for $t = 1982, \ldots, 2021$, and its forecast, $\widetilde{GGR}_{x,t}$, starting from 2022. A vertical line marks this year as the beginning point of the forecasting period. We use the black solid line for age 45, the black dashed line for age 55, the black dotted line for age 65, while we associate the gray dotted line to age 75 and the gray squared line to age 85. Compared to the findings discussed in the previous Section, mortality information updating leads to changes in the expected magnitude of GGR. For each age, we compute how much the predicted GGR under the COVID-19 scenario has varied with respect to its ante-COVID19 predicted value for each year in [2022, 2039]. For the ease of interpretation, we compute the mean of these relative differences throughout [2022, 2039] and provide this synthetic measure for each

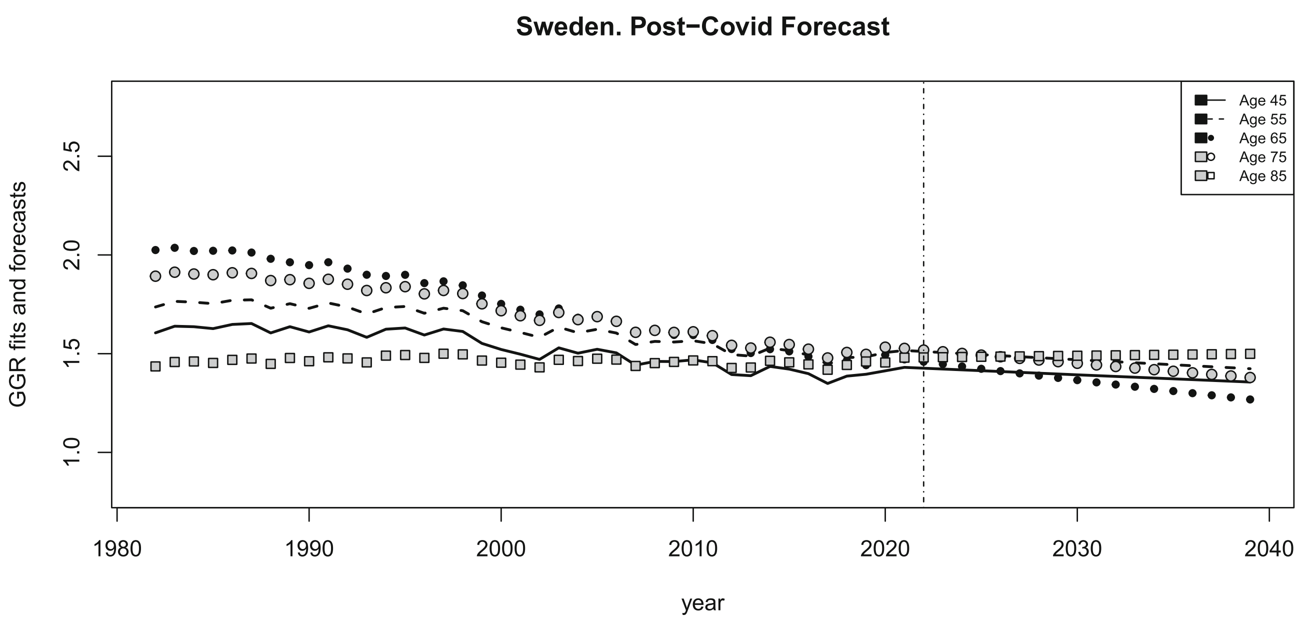

FIGURE 3 Fitted and forecasted path of GGR in Sweden, accounting for COVID-19 mortality. (Calendar years in [1982, 2039]. 2022 marks the start of the forecasting period. Ages 45 (black solid line), 55 (black dashed line), 65 (black dotted line), 75 (gray dotted line), 85 (gray squared line).)

age under consideration. We assess that the expected increase in the magnitude of GGR is around 6% of its ante-COVID-19 predicted value for age 45, 9.5% for age 55 and 7.0% for age 65. Very slight changes affect the predicted GGR for ages 75 and 85. Indeed, accounting for 2020 and 2021 in the calibration phase leads to a predicted GGR, being on average 0.3% higher and 0.2% lower than shown in the previous section, for ages 75 and 85 respectively. These results may suggest that for the younger ages under study, the estimated future values of GGR displayed in Figure 2 approximate from below the gender gap in mortality that could realize in the future. This may result from the consequences of the COVID-19 pandemic on mortality, according to age patterns that echo the one of all-cause mortality,[79] and also from further causes that may play an increasing role in shaping mortality in coming decades, for instance global climate change. In Appendix C, we provide a short discussion about the prediction uncertainty associated with M1 forecasts in the case of Sweden.

9 | MONETIZING THE GENDER GAP RATIO: WHAT ADJUSTMENTS IN LIFE ANNUITIES?

In this section we intend to frame the Gender Gap Ratio in an artificial insurance context, in such a way that a monetary value is associated with this measure. This approach will give us a better understanding of the practical impact of gender inequality in a real world where actuarial valuations are gender neutral. The focus of this section will be to analyze the following problem: the same amount is invested to obtain a regular and constant flow linked to an individual's lifetime. What is the difference that the gender gap existing at the issue time produces on the amount of this flow, depending on whether the contract refers to a man or a woman?

We consider a temporary immediate life annuity providing regular amounts to the annuitant for 10 years upon his/her survival, in exchange for the initial sum of 1000 USD. Our evaluation framework aims at stressing the sensitivity of the annual amount to the Gender Gap Ratio: as a consequence, the interest rate is assumed to be constant over time and symbolically equal to 2% in each case study. Leaving the interest rate fixed in all actuarial valuations will allow us, more clearly, to assess the impact of only the gender differences component in survival, represented by *GGR*. Each calendar year t between 1987 and 2019 represents an issue time, at which the life annuity is evaluated. Within our analysis, under gender-specific demographic assumptions, we define the mathematical relation between a single amount payable at the issue time and a constant amount available at the end of each year, if the life annuity holder is alive, for at most 10 years. Such relation reflects the so-called financial-demographic equivalence principle, that defines an explicit relation between the two mentioned amounts: indeed, 1000 USD is the expected present value of the temporary life annuity, whose constant annual cashflow is $A_{x,t}^{M,M1}$. Annuitants being aged differently at the same issue time or underwriting the life annuity contract in different years are entitled to receive a different annual cashflow. For this reason, although being constant throughout the policy duration, $A_{x,t}^{M,M1}$ has the subscripts t to reflect the year of the life annuity issue and x to denote the age of the annuitant in this year. Furthermore, in the superscript, M stands for males and $M1$ stands for the mortality model employed (LC model). With respect to males, we define the following relation:

$$1,000 = A_{x,t}^{M,M1} a_{\overline{x,t:10}|M1:2\%} \tag{10}$$

where $a_{\overline{x,t:10}|M1:2\%}$ denotes the expected present value (or actuarial value) of a temporary life annuity issued on a man aged x in year t and paying one dollar at the end of each year if he is alive, for at most 10 years. The right-hand subscripts (M1: 2%) define the demographic and financial settings under which such an actuarial value is obtained, namely model M1 survival probabilities and a constant discount rate equal to 2%, respectively. Indeed, with respect to each year t of the life annuity issue, we use the data related to the preceding 39 years and to the observation year for calibrating mortality model M1 and for forecasting it for 10-years ahead. Over this future time horizon, we obtain the forecasted mortality profile of male annuitants, that is, if we set the calendar year t at which the life life annuity is issued as time 0 with respect to the policy duration: $\tilde{m}_{x+1,0+1}^{M,M1}, \dots, \tilde{m}_{x+9,0+9}^{M,M1}$, for every considered age x at time 0. This is consistent with the methodological settings described in the previous Sections. Survival probabilities throughout the policy duration are built from very well-known relations in actuarial mathematics (Equation (2)).

Being the initial sum given, we solve Equation (10) by $A_{x,t}^{M,M1}$ to determine the amount available at the end of each year, for the man in case of life, for up to 10 years, under the demographic assumption that male mortality patterns can be described by model M1. We obtain $A_{x,t}^{M,M1}$ for each calendar year t between 1987 and 2019 set at the issue time and for all integer ages $x \in [45, 80]$ attained in these years.

According to the same reasoning and the same kind of life annuity, we define the mathematical relation between the initial amount of 1,000 USD and the constant amount available at the end of each year in case of life of a female, as follows:

$$1,000 = A^{F,GGA}_{x,t} a_{\overline{x,t:10}|GGA:2\%}. \tag{11}$$

where $a_{\overline{x,t:10}|GGA:2\%}$ denotes the expected present value (or actuarial value) of a temporary life annuity issued on a woman aged x in year t and paying one dollar once a year if she is alive, for at most 10 years. We solve Equation (11) by $A^{F,GGA}_{x,t}$, so that to determine the amount available at the end of each year in case of life under a specific demographic assumption about future female mortality trajectories. Indeed, as denoted by the right-hand superscript GGA, $A^{F,GGA}_{x,t}$ is computed under what we call "gender-gap-adjusted survival probabilities." Our assumption is that the Gender Gap Ratio observed at the issue time will govern the relationship between male and female death rates also in the future. Accordingly, female death rates are obtained from rescaling projected male death rates by a constant factor throughout the policy duration, that is the Gender Gap Ratio realized at the issue time ($GGR_{x,0}$). The "gender-gap-adjusted death rates" are denoted by $\tilde{m}^{F,GGA}_{x+k,0+k}$. Formally, given that x is the age of the policyholder at the issue time 0 and for any integer $0 \le k < 10$, the following relation links $\tilde{m}^{F,GGA}_{x+k,0+k}$ to the male death rate $\tilde{m}^{M,M1}_{x+k,0+k}$:

$$\tilde{m}^{F,GGA}_{x+k,0+k} = \tilde{m}^{M,M1}_{x+k,0+k} / GGR_{x,0}. \tag{12}$$

Survival probabilities are derived accordingly, based on Equation (2). As for males, we compute $A^{F,GGA}_{x,t}$ for $x \in [45, 85]$ and each calendar year t between 1987 and 2019.

In Table 5, we report the obtained values for $A^{M,M1}_{x,t}$ and $A^{F,GGA}_{x,t}$, solving Equations (10) and (11), for $x = 45, 55, 65, 75, 80$ and $t = 1989, 2009, 2019$. We show the outcomes obtained in relation to Italian mortality data, by way of an example. For the sake of the analysis, we report also the magnitude of the Gender Gap Ratio ($GGR_{x,t}$) realized in calendar year t for age x. We see that, the calendar year t being fixed, the larger age x the higher the spread between $A^{M,M1}_{x,t}$ and $A^{F,GGA}_{x,t}$, as an effect of the gender differences in mortality captured by $GGR_{x,t}$. We remark that comparing $A^{M,M1}_{x,t}$ to $A^{F,M1}_{x,t}$, namely to annual amounts obtained by projecting female mortality rates through model M1, gives rise to very similar findings.

As a measure of the discrepancy between the annual cashflows for male and female annuitants, we obtain the ratio between $A^{M,M1}_{x,t}$ and $A^{F,GGA}_{x,t}$ and show its age-dependent patterns in Figure 4 as the calendar year t of the life annuity issue moves forward. For all the examined ages, $A^{M,M1}_{x,t}/A^{F,GGA}_{x,t}$ is larger than one, this denoting, in all cases, a higher annual cashflow for men than for women, against the same amount invested in the life annuity. Nevertheless, the gender gap in mortality reverberates its highest monetary effects at the oldest ages. With respect to the evolution of $A^{M,M1}_{x,t}$ and $A^{F,GGA}_{x,t}$ over time, we notice a more marked stability for middle (that is 45 and 55) than for older ages (that is 65 and above). We remark that the year by year change in $A^{M,M1}_{x,t}/A^{F,GGA}_{x,t}$ reflects two demographic components: the change in male mortality rates' level, as the historical male death rates underlying the forecasting procedures are updated as the issue year t goes forward, and the variation in the observed relationship between male and female death rates as described by $GGR_{x,t}$.

Therefore, in relation to the same kind of life annuity, we further investigate the relationship between the Gender Gap Ratio and $A^{F,GGA}_{x,t}$, by fixing time. Indeed, we set the calendar year in which the life annuity is issued in 2019. Under this static perspective, we design a stress test, to determine the sign and the amount of change that increments in GGR produce in the amount of $A^{F,GGA}_{x,2019}$.

On the one hand, the theoretical amount $A^{M,M1}_{x,2019}$ derives from the previous analysis in relation to calendar year 2019 (see, e.g., Column 8 of Table 5 for given values of x and t); on the other hand, we obtain the amount $A^{F,GGA}_{x,2019}$ under different scenarios for the realization of $GGR_{x,2019}$, namely values between 1 and 2.8, being 0.2 steps apart.

TABLE 5 Annual amounts for men ($A^{M,M1}_{x,t}$) and women ($A^{F,GGA}_{x,t}$) at different issue times t, set in years 1989, 2009, and 2019.

	Annuity issue year (t)								
	1989			2009			2019		
Age (x)	$GGR_{x,t}$	$A^{M,M1}_{x,t}$	$A^{F,GGA}_{x,t}$	$GGR_{x,t}$	$A^{M,M1}_{x,t}$	$A^{F,GGA}_{x,t}$	$GGR_{x,t}$	$A^{M,M1}_{x,t}$	$A^{F,GGA}_{x,t}$
45	2.06	113.68	112.47	1.65	112.56	112.07	1.87	112.37	111.88
55	2.27	118.66	114.53	1.79	114.84	113.28	1.73	114.07	112.91
65	2.27	130.24	119.44	2.06	121.24	116.09	1.78	118.93	115.57
75	1.82	161.69	137.84	1.89	141.91	127.03	1.74	135.43	124.90
80	1.62	196.51	161.08	1.69	168.89	143.94	1.58	159.35	140.82

Note: Age at issue $x = 45, 55, 65, 75$ or 80. $GGR_{x,t}$ denotes the realized Gender Gap Ratio for age x and year t.

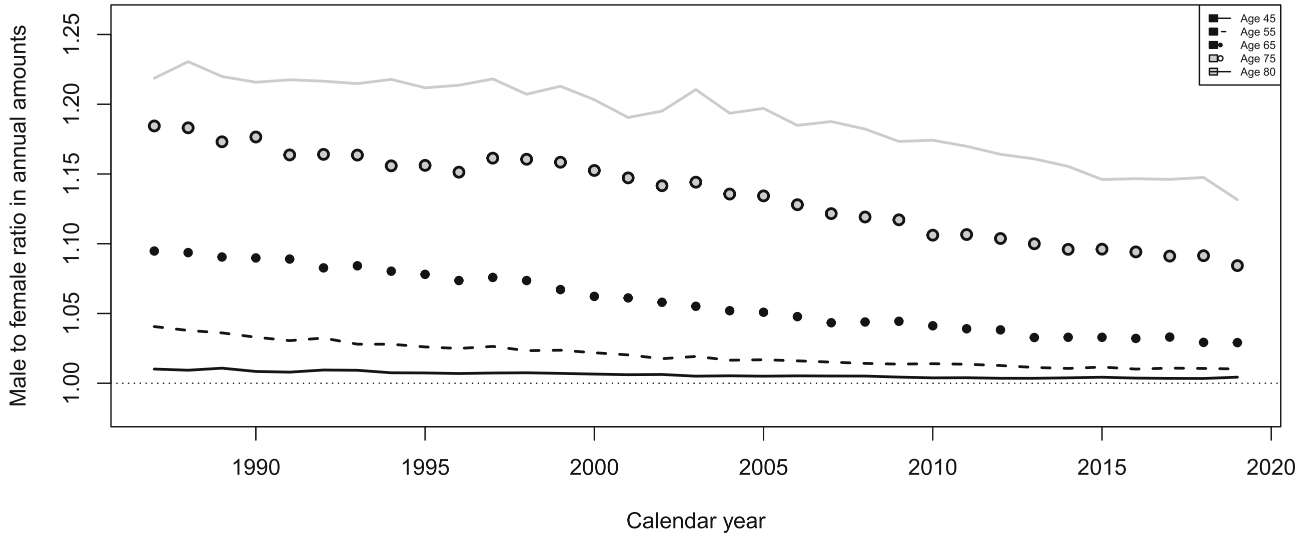

FIGURE 4 Annuity Gap. Historical male to female ratio $A^{M,M1}_{x,t}/A^{F,GGA}_{x,t}$ based on model M1. (The Figure shows the historical male to female ratio $A^{M,M1}_{x,t}/A^{F,GGA}_{x,t}$ for the life annuity payout based on the realized Gender Gap Ratio $GGR_{x,t}$ realized in year t for age x. The male mortality rate evolves from t to $t+10$ according to model M1. We use Italian mortality data. Age at issue are 45 (black solid line), 55 (black dashed line), 65 (black dotted line), 75 (gray dotted line), and 80 (gray solid line).)

Indeed, in this application, female death rates are obtained from rescaling projected male death rates by a constant factor throughout the policy duration being one of ten possible realizations of the Gender Gap Ratio in 2019. We denote by $\tilde{m}^{F,GGA}_{x+k,2019+k}$ the female gender-gap-adjusted death rates. Formally, given that x is the age of the policyholder at the issue year 2019 and for any integer $0 \le k < 10$, the following relation links $\tilde{m}^{F,GGA}_{x+k,2019+k}$ to the male death rate $\tilde{m}^{M,M1}_{x+k,2019+k}$:

$$\tilde{m}^{F,GGA}_{x+k,2019+k} = \tilde{m}^{M,M1}_{x+k,2019+k}/GGR_{x,2019}, \tag{13}$$

where $GGR_{x,2019} = 1, 1.2, \dots, 2.8$. Based on the same equivalence principle previously defined (cf. Equation (11)), we obtain:

$$1,000 = A^{F,GGA}_{x,2019}\, a_{\overline{x,2019:10}|GGA:2\%} \tag{14}$$

where $a_{\overline{x,2019:10}|GGA:2\%}$ denotes the expected present value (or actuarial value) of a temporary life annuity issued on a woman aged x and paying one dollar once a year, for at most 10 years, under gender-gap-adjusted survival probabilities. We remark that increments in $GGR_{x,0}$ express a reduction in female death rates, and thus an increase in women's survival.

In Figure 5, we display what we call "Annuity Gap", namely $\Delta A = A^{M,M1}_{x,2019} - A^{F,GGA}_{x,2019}$, on the z-axis, as a function of the age at the issue time $x \in [45, 85]$ (x-axis) and $GGR_{x,2019} = 1, 1.2, \dots, 2.8$ (y-axis). When $GGR_{x,2019} = 1, A^{M,M1}_{x,2019} - A^{F,GGA}_{x,2019} = 0$. The more advanced the age and the higher $GGR_{x,2019}$, the larger ΔA. The Figure shows an age-specific dependence of ΔA on $GGR_{x,2019}$. Indeed, at each 0.2 increment in $GGR_{x,2019}$ on the $y-axis$, corresponds a more dramatic change ΔA for older ages than for younger ones.

The two life annuity contracts in the example, referring to men and women, differ only by the underlying survival probabilities. Our numerical applications show that, all the other conditions being equal and, in particular against the same amount invested in the life annuity, the gender gap in mortality results in a lower annual cashflow for women than for men; these deductions imply that, all the conditions being equal, a woman would live longer than men, but with a lower income. It is also relevant to highlight that, from a quantitative point of view, such an impact is found to be significant already for annuity durations of 10 years. Accordingly, it is reasonable to expect that the effects of the Gender Gap Ratio would be even amplified when longer annuity durations, for example, pension annuities, are concerned.

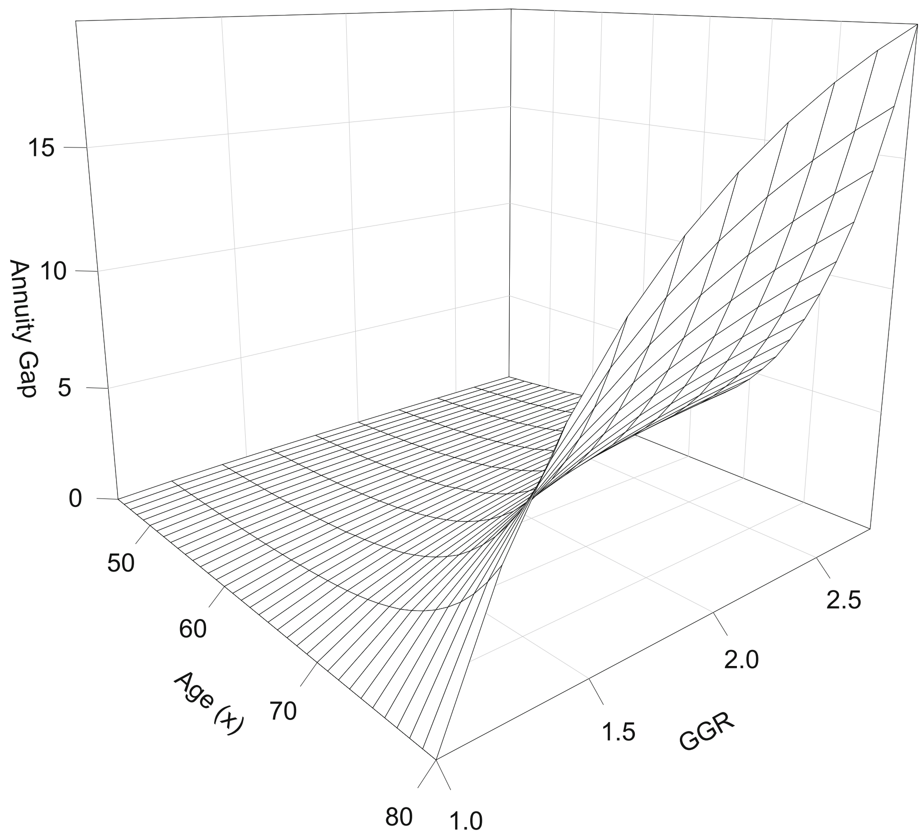

FIGURE 5 Annuity gap for 2019 based on hypothetical values for GGR. (The Figure shows the life annuity gap $A^{M,M1}_{x,2019} - A^{F,GGA}_{x,2019}$ in 2019 based on different scenarios for the Gender Gap Ratio $GGR_{x,t}$ at time $t = 2019$, that is, $GGR_{x,2019} = 1, 1.2, \dots, 2.8$. We use Italian mortality data.)

10 | CONCLUSIONS

In this article, we study the ratio between male and female death rates (Gender Gap Ratio or GGR) for four countries: France, Italy, Sweden, and the USA. We benchmark five representative ages (from 45 to 85, being 10-year apart). We detect general, systematic patterns in the GGR's evolution over the post-World War II time horizon [1948, 2019], for example, the expansion phenomenon, namely the random shifting to the more recent calendar years of the maximum value attained by the Gender Gap Ratio between 1948 and 2019 as the age advances. Such a maximum is followed by a mildly or remarkably decreasing trend of the GGR, except for age 45, being affected by a trend inversion occurring, quite contextually, that is, for some years after 2010, for all countries. Overall, in current times, the magnitude of the Gender Gap Ratio is not remarkably different with respect to its observed magnitude in 1948. France marks the country with the largest gender gap in mortality and Sweden the one with the narrowest gap, for almost all ages in the range $[45, 85]$. After verifying the Lee–Carter model provides a good stochastic description of GGR, we obtain its projection for 20 years, up to 2039, revealing that cross-country differences in the magnitude of the Gender Gap Ratio will continue to exist, despite some commonalities with respect to the age pattern of the projected GGR. We also perform the forecasting procedure including the mortality data related to 2020 and 2021, that accounts for the effects of the COVID-19 pandemic disease and were available for Sweden at the time of this study. Such mortality information updating leads to an upward shifting of the expected magnitude of GGR, for ages below 75.

Finally, focusing on an example temporary life annuity contract, we give monetary consistency to the gender gap by quantifying its impact in the realization of the financial-demographic balances between an invested sum and the constant periodic amount due in case of life to which this investment gives rise, differentiating the male case from the female case. We show evidence that a Gender Gap Ratio that ranges between 1.5 and 2.5, depending on age, translates into a significant reduction of up to 23% in the benefits from a temporary life annuity contract. This result emphasizes the economic importance of assessing the future evolution of the Gender Gap Ratio. Indeed, a quantitative description of the gender mortality gap evolution has an informative and educational value, whose nature is both economic and social. More importantly, it enables a better-informed approach to the promotion of gender equality via policies and practices. Our article adds empirical evidence to the research streams explaining how the well-known financial literacy and the still not thoroughly addressed demographic literacy can enhance individuals' awareness and decisions and, in particular, fosters the adoption of a more gender-inclusive perspective in the implementation of social development policies targeting individuals' well-being at silver ages.

CONFLICT OF INTEREST STATEMENT

The authors report there are no competing interests to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Human Mortality Database at https://mortality.org/.

ACKNOWLEDGMENT

Open access funding provided by University of St Gallen.

ACKNOWLEDGMENTS

We extend our gratitude to the editor and the two anonymous reviewers for their insightful comments. Emilia Di Lorenzo gratefully acknowledges financial support from the Grant Agreement "Progetto Partenariati Estesi - PNRR Project 'Age-It - Ageing well in an ageing society'" (Codice MUR: PE0000015, CUP E63C22002050006). Enrico De Giorgi gratefully acknowledges financial support from the Swiss National Science Foundation (grant no. 189093).

ORCID

Giovanna Apicella https://orcid.org/0000-0002-0524-4552
Enrico De Giorgi https://orcid.org/0000-0001-8017-9618

REFERENCES

1. World Health Organization. Ageing and Health. 2021 Available from https://www.who.int/news-room/fact-sheets/detail/ageing-and-health
2. United Nations, Department of Economic and Social Affairs, Population Division. *World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/NO.3*. Tech. rep. Online ed.; 2022. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf
3. European Commission. Human development and ageing. 2022 Available from https://ec.europa.eu/info/research-and-innovation/research-area/health-research-and-innovation/human-development-and-ageing_en.
4. OECD. *Towards Improved Retirement Savings Outcomes for Women*. OECD Publishing Paris; 2021. doi:10.1787/f7b48808-en
5. Oksuzyan A, Brønnum-Hansen H, Jeune B. Gender gap in health expectancy. *Eur J Ageing*. 2010;7:213-218. doi:10.1007/s10433-010-0170-4
6. Scheel-Hincke LL, Möller S, Lindahl-Jacobsen R, Jeune B, Ahrenfeldt LJ. Cross-national comparison of sex differences in ADL and IADL in Europe: findings from SHARE. *Eur J Ageing*. 2020;17(1):69-79. doi:10.1007/s10433-019-00524-y
7. Payne CF. Expansion, compression, neither, both? Divergent patterns in healthy, disability-free, and morbidity-free life expectancy across U.S. Birth Cohorts, 1998–2016. *Demography*. 2022;59(3):949-973. doi:10.1215/00703370-9938662
8. Munnell AH, Sanzenbacher G, Zulkarnain A. What factors explain the decline in widowed women's poverty? *Demography*. 2020;57(5):1881-1902. doi:10.1007/s13524-020-00915-2
9. Benyamini Y, Burns E. Views on aging: older adults' self-perceptions of age and of health. *Eur J Ageing*. 2020;17:477-487. doi:10.1007/s10433-019-00528-8
10. Raymo JM, Wang J. Loneliness at older ages in the United States: lonely life expectancy and the role of loneliness in health disparities. *Demography*. 2022;59(3):921-947. doi:10.1215/00703370-9937606
11. Walczak D, Wantoch-Rekowski J, Marczak R. Impact of income on life expectancy: a challenge for the pension policy. *Risks*. 2021;9(4):65. doi:10.3390/risks9040065
12. Schünemann J, Strulik H, Trimborn T. The gender gap in mortality: how much is explained by behavior? *J Health Econ*. 2017;54:79-90. doi:10.1016/j.jhealeco.2017.04.002
13. Heiss F, Venti S, Wise D. *The Persistence and Heterogeneity of Health among Older Americans*. NBER Working Papers 20306. National Bureau of Economic Research; 2014.
14. Council of the European Union. Council directive 2004/113/EC. *Off J Eur Union*. 2004; L. 373/37.

15. European Commission. Sex Discrimination in Insurance Contracts: Statement by European Commission Vice-President Viviane Reding, the EU's Justice Commissioner, on the European Court of Justice's ruling in the Test-Achats case. 2011 https://ec.europa.eu/commission/presscorner/detail/it/MEMO_11_123
16. Schmeiser H, Störmer T, Wagner J. Unisex insurance pricing: consumers' perception and market implications. *The Geneva Papers on Risk and Insurance- Issues and Practice*. 2014;39:322-350. doi:10.1057/gpp.2013.24
17. Finkelstein A, Poterba J, Rothschild C. Redistribution by insurance market regulation: analyzing a ban on gender-based retirement annuities. *J Financ Econ*. 2009;91(1):38-58. doi:10.1016/j.jfineco.2007.12.006
18. Bruszas S, Kaschützke B, Maurer R, Siegelin I. Unisex pricing of German participating life annuities - Boon or bane for customer and insurance company? *Insurance Math Econom*. 2018;78:230-245. doi:10.1016/j.insmatheco.2017.09.012
19. OECD/INFE. OECD/INFE Policy Guidance on Addressing Women's and Girls' Needs for Financial Awareness and Education. 2013 https://www.oecd.org/daf/fin/financial-education/G20-Women-Girls-Fin-Ed-Policy-Guidance-2013.pdf
20. Lusardi A, Mitchell OS. The economic importance of financial literacy: theory and evidence. *J Econ Lit*. 2014;52(1):5-44. doi:10.1257/jel.52.1.5
21. Lusardi A, Mitchell OS. Financial literacy and retirement preparedness: evidence and implications for financial education. *Bus Econ*. 2007;42:35-44. doi:10.2145/20070104
22. Lusardi A, Mitchell OS. Planning and financial literacy: how do women fare? *Am Econ Rev*. 2008;98(2):413-417. doi:10.1257/aer.98.2.413
23. Lusardi A, Mitchell OS. Financial literacy around the world: an overview. *J Pension Econom Finance*. 2011;10(4):497-508. doi:10.1017/S1474747211000448
24. Bernheim BD, Garrett DM. The effects of financial education in the workplace: evidence from a survey of households. *J Public Econ*. 2003;87(7):1487-1519. doi:10.1016/S0047-2727(01)00184-0.
25. Ghafoori E, Ip E, Kabátek J. The impacts of a large-scale financial education intervention on retirement saving behaviors and portfolio allocation: evidence from pension fund data. *J Bank Financ*. 2021;130:106195. doi:10.1016/j.jbankfin.2021.106195
26. Kaiser T, Lusardi A, Menkhoff L, Urban C. Financial education affects financial knowledge and downstream behaviors. *J Financ Econ*. 2022;145(2, Part A):255-272. doi:10.1016/j.jfineco.2021.09.022
27. Bucher-Koenen T, Lusardi A, Alessie R, van Rooij M. How financially literate are women? An overview and new insights. *J Consum Aff*. 2017;51(2):255-283. doi:10.1111/joca.12121
28. Preston AC, Wright RE. Understanding the gender gap in financial literacy: evidence from Australia. *Econ Rec*. 2019;95(S1):1-29. doi:10.1111/1475-4932.12472
29. Bucher-Koenen T, Alessie RJ, Lusardi A, van Rooij M. Fearless women: financial literacy and stock market participation. *NBER Working Paper Series*. 28723; 2021. doi:10.3386/w28723
30. Aristei D, Gallo M. Assessing gender gaps in financial knowledge and self-confidence: evidence from international data. *Financ Res Lett*. 2022;46:102200. doi:10.1016/j.frl.2021.102200
31. Driva A, Lührmann M, Winter J. Gender differences and stereotypes in financial literacy: off to an early start. *Econ Lett*. 2016;146:143-146. doi:10.1016/j.econlet.2016.07.029
32. Tinghög G, Ahmed A, Barrafrem K, Lind T, Skagerlund K, Västfjäll D. Gender differences in financial literacy: the role of stereotype threat. *J Econ Behav Organ*. 2021;192:405-416. doi:10.1016/j.jebo.2021.10.015
33. Hamermesh DS. Expectations, life expectancy, and economic behavior. *Q J Econ*. 1985;100(2):389-408. doi:10.2307/1885388
34. Hurd MD, McGarry K. The predictive validity of subjective probabilities of survival. *Econ J*. 2002;112(482):966-985.
35. Perozek M. Using subjective expectations to forecast longevity: do survey respondents know somethingwe don't know? *Demography*. 2008;45(1):95-113. doi:10.1353/dem.2008.0010
36. Apicella G, De Giorgi EG. A behavioral gap in survival beliefs. *J Risk Insur*. 2024;91(1):213-247. doi:10.1111/jori.12459
37. Grevenbrock N, Groneck M, Ludwig A, Zimper A. Cognition, optimism and the formation of age-dependent survival beliefs. *Int Econ Rev*. 2021;62(2):887-918. doi:10.1111/iere.12497
38. Heimer RZ, Myrseth KOR, Schoenle RS. YOLO: mortality beliefs and household finance puzzles. *J Financ*. 2019;74(6):2957-2996. doi:10.1111/jofi.12828
39. Post T, Hanewald K. Longevity risk, subjective survival expectations, and individual saving behavior. *J Econ Behav Organ*. 2013;86:200-220. doi:10.1016/j.jebo.2012.11.012
40. Spaenjers C, Spira SM. Subjective life horizon and portfolio choice. *J Econ Behav Organ*. 2015;116:94-106. doi:10.1016/j.jebo.2015.04.006
41. Apicella G, De Giorgi E. Gender attitudes toward longevity and retirement planning: theory and evidence. In: Corazza M, Perna C, Pizzi C, Sibillo M, eds. *Mathematical and Statistical Methods for Actuarial Sciences and Finance*. Springer International Publishing; 2022:19-24.
42. Danielsen AC, Lee KMN, Boulicault M, et al. Sex disparities in COVID-19 outcomes in the United States: quantifying and contextualizing variation. *Soc Sci Med*. 2022;294:114716. doi:10.1016/j.socscimed.2022.114716
43. Jin J-M, Bai P, He W, et al. Gender differences in patients with COVID-19: focus on severity and mortality. Front Public Healthc; 2020;8:152.
44. Peckham H, de Gruijter NM, Raine C, et al. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nat Commun*. 2020;11(1):6317. doi:10.1038/s41467-020-19741-6
45. Zarulli V, Barthold Jones JA, Oksuzyan A, Lindahl-Jacobsen R, Christensen K, Vaupel JW. Women live longer than men even during severe famines and epidemics. *Proc Natl Acad Sci*. 2018;115(4):E832-E840. doi:10.1073/pnas.1701535115

46. Hurwitz A, Mitchell OS, Sade O. Testing methods to enhance longevity awareness. *J Econ Behav Organ*. 2022;204:466-475. doi:10.1016/j.jebo.2022.10.014
47. Angelici M, Del Boca D, Oggero N, Profeta P, Rossi MC, Villosio C. Pension information and women's awareness. *J Econom Ageing*. 2022;23:100396. doi:10.1016/j.jeoa.2022.100396
48. World Economic Forum. Global gender gap report 2021. *Insight Report March 2021*. World Economic Forum; 2021.
49. Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). 2022 Available from http://www.mortality.org and http://www.humanmortality.de
50. Pitacco E, Denuit M, Haberman S, Olivieri A. *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press; 2009.
51. Lee RD, Carter LR. Modeling and Forecasting U.S. Mortality. *J Am Stat Assoc*. 1992;87(419):659-671. doi:10.1080/01621459.1992.10475265
52. Basellini U, Camarda CG, Booth H. Thirty years on: a review of the Lee-Carter method for forecasting mortality. *Int J Forecast*. 2022;39(3):1033-1049. doi:10.1016/j.ijforecast.2022.11.002
53. Bergeron-Boucher M-P, Kjærgaard S. Mortality forecasting at age 65 and above: an age-specific evaluation of the Lee-Carter model. *Scand Actuar J*. 2022;2022(1):64-79. doi:10.1080/03461238.2021.1928542
54. Lee R, Miller T. Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*. 2001;38:537-549.
55. Li N, Lee R, Gerland P. Extending the Lee-Carter method to model the rotation of age patterns of mortality decline for long-term projections. *Demography*. 2013;50:2037-2051.
56. Li N, Lee R. Coherent mortality forecasts for a group of populations: an extension of the Lee-Carter method. *Demography*. 2005;42(3):575-594.
57. Pedroza C. A Bayesian forecasting model: predicting U.S. male mortality. *Biostatistics*. 2006;7(4):530-550. doi:10.1093/biostatistics/kxj024
58. Carter LR, Lee RD. Modelling and forecasting US sex differentials in mortality. *Int J Forecast*. 1992;8(3):393-411.
59. Dowd K, Cairns AJ, Blake D, Coughlan GD, Khalaf-Allah M. A gravity model of mortality rated for two related populations. *N Am Actuar J*. 2011;15(2):334-356. doi:10.1080/10920277.2011.10597624
60. Cairns AJG, Blake D, Dowd K. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *J Risk Insur*. 2006;73(4):687-718. doi:10.1111/j.1539-6975.2006.00195.x
61. Cairns AJG, Blake D, Dowd K, et al. A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *N Am Actuar J*. 2009;13(1):1-35. doi:10.1080/10920277.2009.10597538
62. Dowd K, Cairns AJC, Blake D, Coughlan GD, Epstein D, Khalaf-Allah M. Evaluating the goodness of fit of stochastic mortality models. *Insurance Math Econom*. 2010a;47(3):255-265. doi:10.1016/j.insmatheco.2010.06.006
63. Dowd K, Cairns AJG, Blake D, Coughlan GD, Epstein D, Khalaf-Allah M. Backtesting stochastic mortality models. *N Am Actuar J*. 2010b;14(3):281-298. doi:10.1080/10920277.2010.10597592
64. Haberman S, Renshaw A. A comparative study of parametric mortality projection models. *Insurance Math Econom*. 2011;48:35-55.
65. Apicella G, Dacorogna M, Di Lorenzo E, Sibillo M. Improving the forecast of longevity by combining models. *N Am Actuar J*. 2019;23(2):298-319. doi:10.1080/10920277.2018.1556701
66. Villegas AM, Millossovich P, Kaishev VK. Package StMoMo. 2016a https://cran.r-project.org/web/packages/StMoMo/StMoMo.pdf
67. Villegas AM, Millossovich P, Kaishev VK. StMoMo: an R package for stochastic mortality modelling. 2016b https://cran.rproject.org/web/packages/StMoMo/vignettes/StMoMoVignette.pdf
68. Currie ID. Smoothing constrained generalized linear models with an application to the Lee-Carter model. *Stat Model*. 2013;13(1):69-93. doi:10.1177/1471082X12471373
69. Hayashi F. *Econometrics*. Princeton University Press; 2000.
70. Cairns A, Blake D, Dowd K, Coughlan G, Epstein D, Khalaf-Allah M. Mortality density forecasts: an analysis of six stochastic mortality models. *Insurance Math Econom*. 2011;48(3):355-367.
71. Scognamiglio S, Marino M. Backtesting stochastic mortality models by prediction interval-based metrics. *Qual Quant*. 2023;57:3825-3847. doi:10.1007/s11135-022-01537-z
72. Levantesi S, Lizzi M, Nigri A. Enhancing diagnostic of stochastic mortality models leveraging constrast trees: an application on Italian data. *Qual Quant*. 2023;58:1565-1581. doi:10.1007/s11135-023-01711-x
73. van Berkum F, Antonio K, Vellekoop M. The impact of multiple structural changes on mortality predictions. *Scand Actuar J*. 2016;2016(7):581-603. doi:10.1080/03461238.2014.987807
74. Zarulli V, Kashnitsky I, Vaupel JW. Death rates at specific life stages mold the sex gap in life expectancy. *PNAS*. 2021;118(20):e2010588118. doi:10.1073/pnas.2010588118
75. Ekwueme DU, Guy GP, Rim SH, et al. Health and economic impact of breast cancer mortality in young women, 1970–2008. *Am J Prev Med*. 2014;46(1):71-79. doi:10.1016/j.amepre.2013.08.016
76. Wojtyla C, Bertuccio P, Wojtyla A, La Vecchia C. European trends in breast cancer mortality, 1980–2017 and predictions to 2025. *Eur J Cancer*. 2021;152:4-17. doi:10.1016/j.ejca.2021.04.026
77. Zarulli V, Lindahi-Jacobsen R, Vaupel JW. Onset of the old-age gender gap in survival. *Demogr Res*. 2020;42(25):727-740. doi:10.4054/DemRes.2020.42.25
78. Gebhard CE, Hamouda N, Gebert P, Regitz-Zagrosek V, Gebhard C. Sex versus gender-related characteristics: which predicts clinical outcomes of acute COVID-19? *Intensive Care Med*. 2022;48(11):1652-1655. doi:10.1007/s00134-022-06836-5

79. Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. *Proc Natl Acad Sci.* 2020;117(36):22035-22041. doi:10.1073/pnas.2006392117

How to cite this article: Apicella G, De Giorgi E, Di Lorenzo E, Sibillo M. Gender-inclusive financial and demographic literacy: Monetizing the gender mortality gap. *Appl Stochastic Models Bus Ind*. 2024;1-23. doi: 10.1002/asmb.2876

APPENDIX A. FURTHER EVIDENCE ON THE FORECASTING PERFORMANCE OF MODELS M1, M5, AND M8

In Figure A1, we show the magnitude of the RMSE of models M1, M5, and M8, for the forecasts related to the Italian female population over the lookforward window [1998–2017], across the various ages under study in the range [25,85].

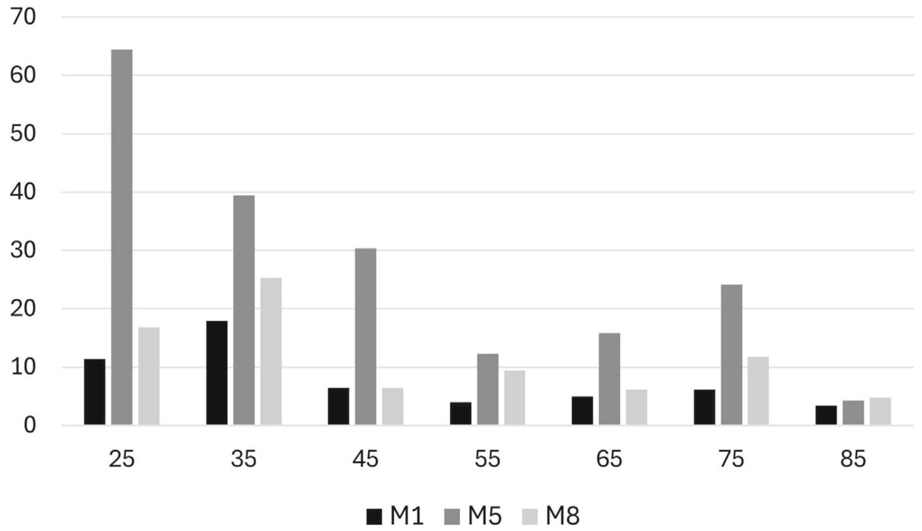

FIGURE A1 RMSE for models M1, M5, and M8 (unit 10^{-2}). Population: Italian females. Lookforward window: [1998–2017]. Integer ages in the interval [25, 85], being ten-years apart (x-axis).

APPENDIX B. SOME EVIDENCE ABOUT GGR UNSMOOTHED DATA

In this section, we show some evidence, in a cross-country perspective, on GGR unsmoothed data. GGR values are obtained according to Equation (9), namely from the ratio of male to female crude death rates.

B.1 Relative change of GGR

We observe how GGR values vary year by year. We measure such a change by computing the relative difference of GGR in one calendar year with respect to its value one year before, for the five ages between 45 and 85 being 10 years apart. We then average such relative differences on 10-year windows (from 1980 to 2019) to gain insights into the persistence of GGR behavior over time. Sweden is characterized by a much smaller volume of exposures and deaths than the other countries under study. Due to the intrinsic features of male and female underlying data, for Sweden GGR experiences much more marked year-by-year variations than for the other countries. This implies larger values of its relative change over time, for all the ages under study and especially for age 45, as shown by Figure B1.

B.2 Trend inversion

In this section, we show 5-years moving averages of the observed values $GGR_{x,t}$. With respect to what shown in section 6, this approach describes how the unsmoothed GGR behave over the time horizon [1948, 2019], in a cross-country perspective.

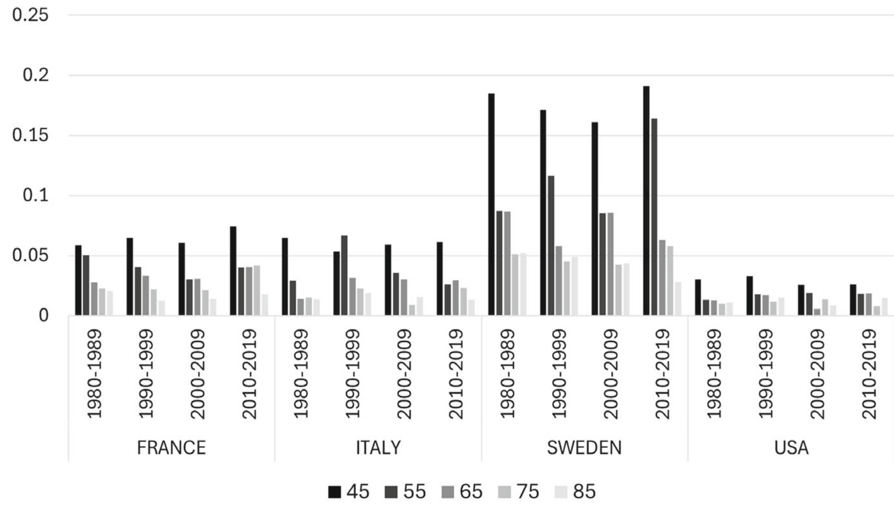

FIGURE B1 Average relative change of unsmoothed GGR over time.

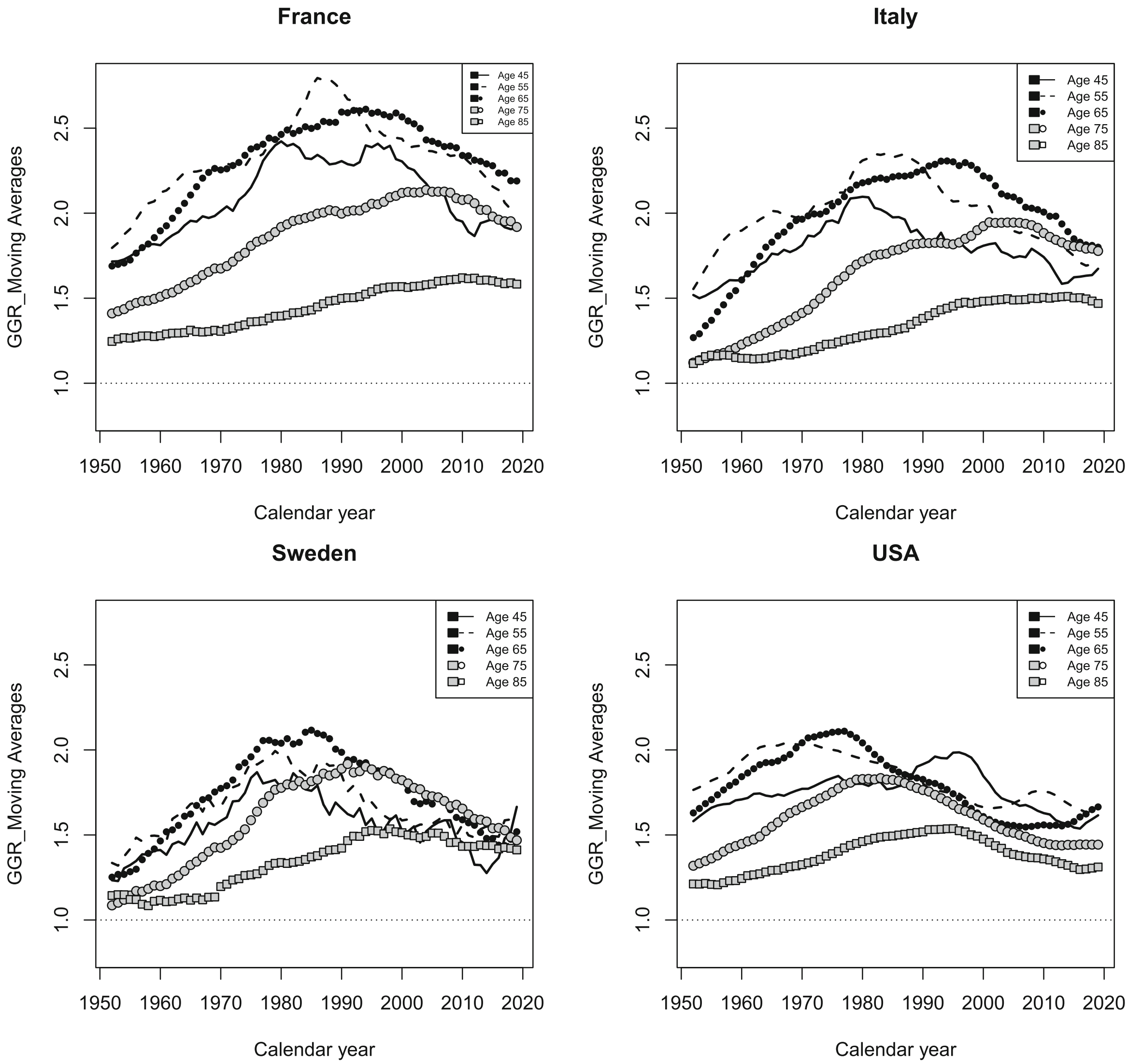

FIGURE B2 Moving averages (5-years) for the Gender Gap Ratio. Such averages are computed on the observed values of GGR. (Calendar years in [1948, 2019]. Ages 45 (black solid line), 55 (black dashed line), 65 (black dotted line), 75 (gray dotted line), 85 (gray squared line). Countries: France, Italy, Sweden, and USA.)

APPENDIX C. PREDICTION UNCERTAINTY

GGR forecasts $\widetilde{GGR}_{x,t}$ are obtained by computing the ratio of male to female forecasted death rates for the corresponding age x and calendar year t. In this section, we obtain predictions intervals at difference confidence level (50%, 80%, and 95%), showing the uncertainty deriving from the error in the forecast of the period index of model M1, in the illustrative case of Sweden. We use two different calibration samples made up by 40 data points: [1980–2019] and [1982–2021], to account for the years highly impacted by the COVID-19 pandemic disease. We implement an analogous procedure as in Reference 66 to simulate 10,000 trajectories from model M1. Figures C1 and C2 show the related fan charts. Each figure consists of two panels (left for females and right for males) to ease comparisons. Ages 45, 55, and 65 fans are displayed within the same graph, and, likewise, ages 75 and 85. The left and right panel are made homogeneous across genders with respect to the domain of values of the death rates reported on the y-axis. For these reasons, prediction confidence

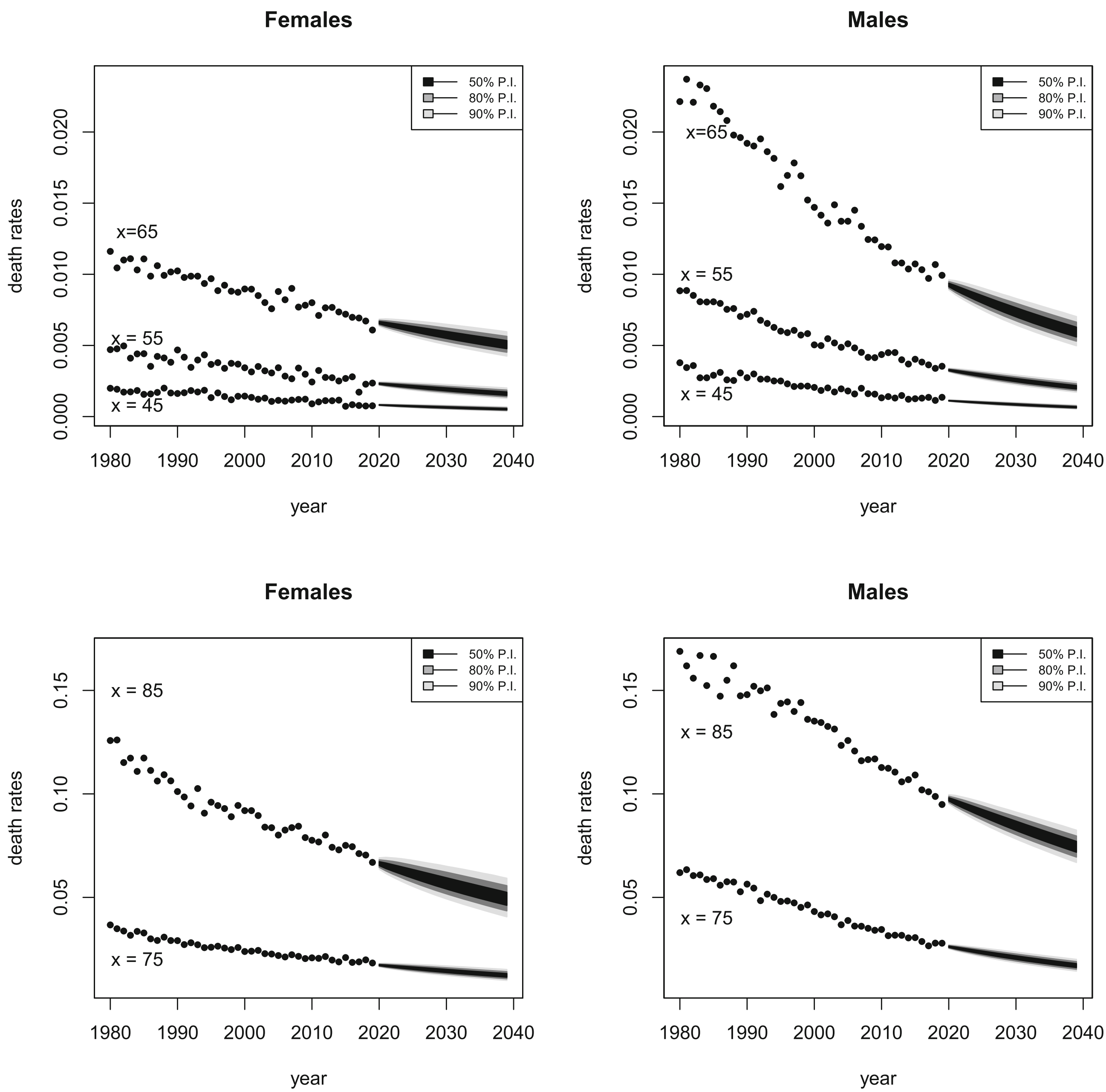

FIGURE C1 Fan charts for death rates at various ages from model M1 fitted to Swedish female and male population for the period 1980–2019. The dots show historical death rates on such time horizon. Prediction Intervals (P.I.) at the 50%, 80%, and 95% level, for the period 2020–2039, are represented with black, dark gray, and light gray colors in the fan.

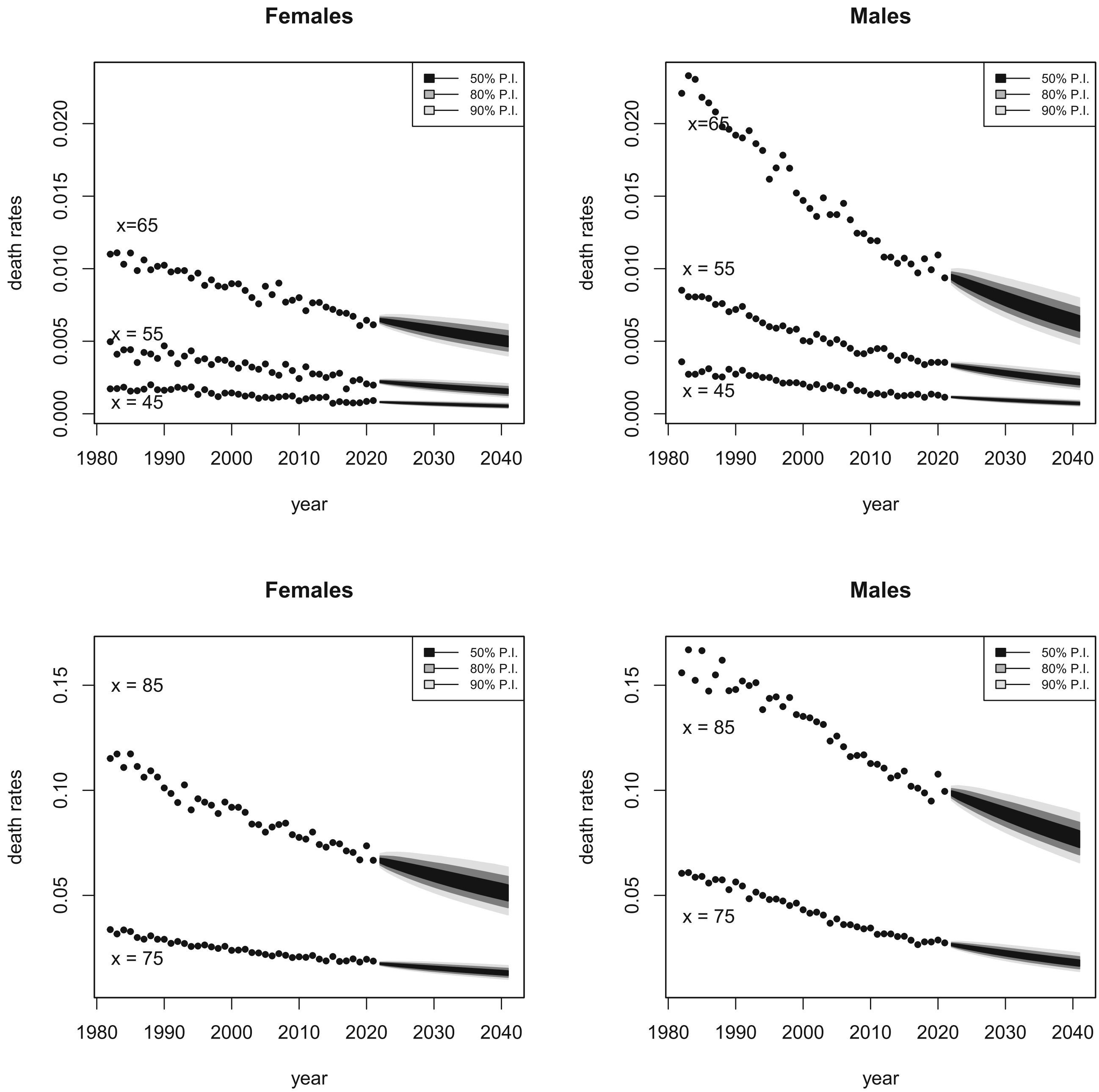

FIGURE C2 Fan charts for death rates at various ages from model M1 fitted to Swedish female and male population for the period 1982–2021. The dots show historical death rates on such time horizon. Prediction Intervals (P.I.) at the 50%, 80%, and 95% level, for the period 2022–2041, are represented with black, dark gray, and light gray colors in the fan.

intervals may graphically appear compressed for some ages, for example, age 45, and for females more than for males. In all figures, the dots show historical date rates on the calibration sample, while prediction intervals at the 50%, 80%, and 95% level, for the corresponding forecasting time horizon, are displayed by the shading in the fan. We can see consistency between the predicted uncertainty and the observed volatility on the calibration sample. This result holds quite generally for the countries under study and the ages we consider. Under the COVID-19 scenario, fitting the model on a time horizon that includes calendar years 2020 and 2021 seems to moderately increase predicted uncertainty, especially at ages 65+.