A General framework for modelling mortality to better estimate its relationship with interest rate risks

Apicella, Giovanna and Dacorogna, Michel M

University of Rome "La Sapienza", DEAR-Consulting

July 2016

December 2016 N°39

SCOR Papers

A General Framework for Modeling Mortality to Better Estimate its Relationship to Interest Rate Risks

By **Michel Dacorogna*** and
Giovanna Apicella°

** DEAR-Consulting,*
former SCOR Scientific Advisor

° University of Rome "La Sapienza"

Abstract

The need for having a good knowledge of the degree of dependence between various risks is fundamental for understanding their real impacts and consequences, since dependence reduces the possibility to diversify the risks.

This paper expands in a more theoretical approach the methodology developed in [6] for exploring the dependence between mortality and market risks in case of stress. In particular, we investigate, using the Feller process, the relationship between mortality and interest rate risks. These are the primary sources of risk for life (re)insurance companies. We apply the Feller process to both mortality and interest rate intensities. Our study cover both the short and the long-term interest rates (3m and 10y) as well as the mortality indices of ten developed countries and extending over the same time horizon. Specifically, this paper deals with the stochastic modelling of mortality. We calibrate two different specifications of the Feller process (a two-parameters Feller process and a three-parameters one) to the survival probabilities of the generation of males born in 1940 in ten developed countries. Looking simultaneously at different countries gives us the possibility to find regularities that go beyond one particular case and are general enough to gain more confidence in the results. The calibration provides in most of the cases a very good fit to the data extrapolated from the mortality tables. On the basis of the principle of parsimony, we choose the two-parameters Feller process, namely the hypothesis with the fewer assumptions. These results provide the basis to study the dynamics of both risks and their dependence.

SCOR
The Art & Science of Risk

1 Introduction

The world we live in is subject to large stochastic shocks: disasters that could occur because of climate changes, failures of complex IT systems, which we massively use in a number of fields, are just few examples. At the same time, risks are unavoidable components of innovation and progress. Societies should accept risks when they invest in new ideas and engage in new adventures; furthermore they should be able to manage the consequences of bad realizations of any of the risks they are exposed to.

As Dacorogna and Kratz stress in [7], risks are becoming more and more complex because of the increasing interconnectedness of the world and the faster time scale whereby actors have little time to adapt. The multiple networks arising from globalization, Internet communication and the economic development could propagate any extreme event, due to natural or manmade catastrophes, in a very short time, bringing about two kinds of effects: on the one hand we can see immediately the possible bad consequences of any catastrophic event occurring anywhere in the world and study their interactions; on the other hand, the severity of the losses could increase. Exploring the dependence between risks is crucial for understanding their real impact and consequences, because dependence reduces the possibility to diversify the risks [19]. The forthcoming research in statistics and probability should make progress in addressing the dependence to better meet the social need for protection against systemic risks.

Within the re(insurance) framework, the importance of having a good grasp of the degree of dependence between various risks has been emphasized by the current solvency regulations. The Solvency II Directive requires the estimation, at a one year time horizon, of the Value-at-Risk(VaR) at the 99.5% probability and the Swiss Solvency Test (SST) the estimation of the Tail-Value-at-Risk (TVaR) at the 99% probability. In line with what we have explained previously, both cannot be reasonably computed without a proper understanding of each of the risks and also a good knowledge of the degree of dependence between the various risks, especially in extreme situations, whereby we suspect that dependence is much stronger than in normal situations and thus affects significatively the diversification benefit measured at high thresholds (99 or 99.5%).

From the life (re)insurance companies perspective, mortality and interest rate risks are the primary sources of risk. The calculations of the expected present values of living benefits are based on three-factors: stock market returns, interest rates and survival probabilities which often extend over a long time horizon. Individually speaking, risks due to the uncertainty in level and in trend of future mortality may heavily affect portfolio results, especially when long-term insurance products, like life annuities, are concerned (see [18]); furthermore as interest rates change, the values of a life insurer's assets and liabilities change, potentially exposing the company to risk (see [3]). Indeed, also the dependence between mortality and market risks is a very important component of the risk of life business, since it links the liabilities to the investments made by (re)insurance companies to back those liabilities and it thus represents a relevant issue for risk management.

Our line of research aims at investigating the long-term dependence between mortality and interest rate risks. As we have explained in [2] providing an overview of the motivation and goals underlying our research, we expand the empirical study done in [6] by a more theoretical approach where we model the dynamics of both mortality and interest rate intensities with the same type of equations and within the same framework. In particular, we exploit the following important similarities between the intensity of mortality and the one of interest rate, stressed in Milevsky and Promislow [16], Dahl [8] and Biffis [4]: both are positive processes, have term structures, are stochastic in nature. For modelling both mortality and interest rate intensities we make use of affine processes.

In this paper, we develop such line of research addressing the stochastic mortality risk modelling. We calibrate two different specifications of the Feller process(a two-parameters Feller process and a three- parameters one) to the mortality data concerning ten developed countries. Having such a wide set of countries and data give us the possibility to find regularities that go beyond one particular case and are general enough to gain more confidence in the results.

The paper is organized as follows. In Section 1, we state the conceptual framework of our research. In section 2, we provide an overview on the state of the art about the use of affine processes in the context of both mortality and financial risks. In Section 3, we describe the features of the two Feller models we exploit within our investigation. In Section 4, we present the data selected for calibrating the models, according to the methodology explained in Section 5. We discuss the results in Section 6 and conclude in Section 7.

2 Literature review

As widely indicated in the recent literature, it is opportune, in complex valuation frameworks, to work with computationally tractable and flexible models. In this context, affine processes are particularly suitable.

For dealing with risk analysis and market valuation of life insurance contracts, Biffis (2005) [4] exploited the analytical tractability of affine processes for both financial and mortality risks, stressing how useful the aforementioned stochastic processes proved to be in developing dynamic asset pricing models. Schrager's approach (2006) [20] consisted in applying techniques from the literature on term structure and credit risk to the modelling of mortality rates evolution over time. The model he estimated, using data on Dutch mortality rates, was characterized by a rich analytical structure, clear interpretation of the latent factors and consistency with derivative pricing models.

Following the modelling framework of [4] and the analytical studies of Luciano and Vigna [12, 13] (2005 and 2008) aiming at the selection of the most appropriate affine process to describe the death intensity of individuals, Jalen and Mamon (2009) [11] were among the first to introduce a pricing framework in which the dependence between mortality and interest risks is explicitly modeled. The topic was deepened in Liu et al. (2014) [22] , where affine processes were used for pricing Guaranteed Annuity Options (GAO). In [21], Russo et al. (2011) calibrated affine stochastic mortality models using term assurance premiums.

Within the analysis of the systematic risk in a portfolio of deferred annuities, in [14] Mahayni and Steuten (2013) considered a pure diffusion model and a compound Poisson jump model for the specification of the stochastic mortality rates, while they used a one-factor Hull White interest rate process. All models were calibrated to financial market and demographic data. The Affine Term Structure Model was also used also by Blackburn and Sherris (2013) [5] for modelling longevity risk. The multi-factor dynamic affine mortality model they calibrated to the Swedish mortality data for years 1910-2007 proved to be an appropriate tool for fitting historical mortality rates. Grasselli et al. (2015) [10], in deriving pricing formulas for insurance contracts like GAO, developed more general affine models allowing for the dependence between mortality and interest rates.

3 Mortality modeling

Following [4] and [13], we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a filtration $\{\mathcal{G}_t : t \geq 0\}$ of sub-σ-algebras of $\mathcal{F}$ satisfying the usual conditions. We focus on an individual aged x at time 0 and model his/her random residual lifetime T_x as a doubly stochastic stopping time with intensity λ_x driven by the sub-filtration $\{\mathcal{F}_t : t \geq 0\}$, where $\mathcal{F}_t \subset \mathcal{G}_t$. Specifically, T_x represents the first jump-time of a nonexplosive counting process N having a Poisson distribution with parameter $\int_t^s \lambda_u du$. The process N jumps whenever the individual dies: $N_t = 0$ if $t < T_x$, $N_t = 1$ if $t \geq T_x$.

Such setup is stochastic and dynamic, since the process λ depends on the “state of the world" determining its particular trajectory and on the date $t \geq 0$, representing, for example, the continuous-time counterpart of the calendar year of reference in longitudinal tables. As Biffis, we are adopting a diagonal, or cohort-based, approach (see Pitacco [17]). This means, from an actuarial point of view, that $\lambda_x(t)$ stochastically changes over time, describing the future intensity of mortality for any age $x + t$ of an individual aged x at time 0.

Formally, it is possible to show that the survival probability is given by:

$$S_x(t) = P(T_x > t|\mathcal{G}_0) = E[\mathrm{e}^{-\int_0^t \lambda_x(u)du}|\mathcal{G}_0] \tag{1}$$

Under technical conditions (see Duffie and Singleton [9]), if $\lambda_x(t)$ is modeled as an affine process, the probability of survival in (1) has an affine exponential form and can be computed as follows:

$$S_x(t) = \mathrm{e}^{\alpha(t)+\beta(t)\lambda_x(0)} \tag{2}$$

where the coefficients $\alpha(t)$ and $\beta(t)$ satisfy generalized Riccati Ordinary Differential Equations (ODEs). The latter can be solved at least numerically and in some cases analytically. Therefore, working in an affine framework is very convenient from the computational point of view.

In selecting the features of the affine process for λ we are inspired by the remarks of Cairns et al.[1] and Luciano et al. [12]. In [1] the properties that any plausible mortality model should have were suggested; among such criteria we highlight that a mortality model should: keep the force of mortality positive, be consistent with historical data, have biologically reasonable long-term future dynamics, imply long-run deviations in mortality improvements from those anticipated that are not strongly mean reverting to a pre-determined target level, even if this target is time dependent and incorporates mortality improvements. The latter requirement ensures that, if mortality has improved faster than anticipated in the past, the potential for further mortality improvements will not be reduced in the future.

In [12], time-homogeneous mean reverting affine processes turned out to fail to capture the rectangularization phenomenon, namely the increasing concentration of the deaths around the mode (at old ages) of the curve of deaths (see [17]). Furthermore, an exponential behaviour of the force of mortality observed and/or extrapolated from the mortality tables was detected; non mean reverting processes, with deterministic part increasing exponentially were thus selected; these models represent natural generalizations of the Gompertz law.

Within our application, we choose for the stochastic force of mortality two different specifications of a Feller process: a two-parameters Feller process and a three-parameters Feller process; we calibrate them to the same sample of data and we compare their goodness of fit. In 3.1 and 3.2 we briefly describe the features of such processes, referring the reader to [13] and to Fung et al. (2014) [15], for further details.

3.1 Two-parameters Feller process

Under such setup, the stochastic differential equation (SDE) of the Feller process describing the stochastic force of mortality dynamics is given by:

$$d\lambda(t) = a\lambda(t)dt + \sigma\sqrt{\lambda(t)}dW(t), \tag{3}$$

where $a > 0$, $\sigma \geq 0$ and W is a standard Brownian motion. Compared to the Cox-Ingersoll-Ross(CIR) process describing the dynamics of the spot interest rate, in (3) the parameter representing the long-term mean of the process is missing.

In the affine framework, the solution of the system of ODEs for α and β in (2) is:

$$\begin{cases} \alpha(t) = 0 \\ \beta(t) = \frac{1-\mathrm{e}^{bt}}{c+d\mathrm{e}^{bt}} \end{cases} \tag{4}$$

with:

$$\begin{cases} b = -\sqrt{a^2 + 2\sigma^2} \\ c = \frac{b+a}{2} \\ d = \frac{b-a}{2} \end{cases} \tag{5}$$

The coefficients b, c, d are negative, so the survival probability is always decreasing in t if and only if:

$$e^{bt}(\sigma^2 + 2d^2) > \sigma^2 - 2dc \tag{6}$$

If the starting point is nonnegative, this model does not violate the non-negative constraint of the intensity, but the intensity process can take the value of zero with positive probability; nevertheless, such probability proves to be negligible in the practical applications.

The function β in (4) is increasing in σ, so the survival probability increases with the diffusion part. If $\sigma = 0$, the evolution of $\lambda_0(t)$ is deterministic and coincides with the Gompertz specification. Thus the following equalities hold:

$$\mu_x = \lambda_0(0)e^{ax} = \lambda_0(x) \tag{7}$$

where μ_x represents the deterministic force of mortality:

$$\mu_x = \lim_{h \to 0} \frac{P(x < T_0 \leq x + h | T_0 > x)}{h} \tag{8}$$

When $t \to \infty$, the survival probability tends to $e^{\frac{1}{c}}$.

3.2 Three-parameters Feller process

The following Feller process describes the stochastic intensity of mortality of a specific generation:

$$d\lambda(t) = (\bar{a} + \bar{b}\lambda(t))dt + \bar{\sigma}\sqrt{\lambda(t)}dW(t), \tag{9}$$

where $\bar{a} > 0, \bar{b} > 0, \bar{\sigma} > 0, \lambda(0) = \lambda_0 \in \mathbb{R}^{++}$. Since the process is expected to have no mean reversion, the assumption $\bar{b} > 0$ is made.

Under such setup, the conditional probability that an individual aged x at time 0 survives age $x + t$ in (1) is given by:

$$S_x(t) = A(t)e^{-B(t)\lambda_x(0)} \tag{10}$$

where $A(t)$ and $B(t)$ are the solutions of the following system of Riccati equations:

$$\begin{cases} A(t) = \left(\frac{2\gamma e^{\frac{1}{2}(\gamma - \bar{b})t}}{(\gamma - \bar{b})(e^{\gamma t} - 1) + 2\gamma}\right)^{\frac{2\bar{a}}{\bar{\sigma}^2}} \\ B(t) = \frac{2(e^{\gamma t} - 1)}{(\gamma - \bar{b})(e^{\gamma t} - 1) + 2\gamma} \end{cases} \tag{11}$$

where $\gamma = \sqrt{\bar{b}^2 + 2\bar{\sigma}^2}$.

When $\bar{\sigma} = \bar{a} = 0$, the model reduces to the constant Gompertz specification, as follows:

$$\lambda_0(x) = \lambda_0(0)e^{\bar{b}x} \tag{12}$$

Looking at the stochastic differential equations in (3) and (9), we can see that the two specifications of the Feller process for λ, discussed in sections 3.1 and 3.2, differ by the presence, in the latter one, of one additional parameter $\bar{a}$, that, according to [15], provides the following advantages compared to the two-parameters Feller process: the limit of the survival probability, when t diverges, is zero, and, if the initial point λ_0 is strictly positive and the following condition holds:

$$\bar{a} \geq \frac{\bar{\sigma}^2}{2}, \tag{13}$$

then $\lambda(t)$ is strictly positive for each t, almost surely.

4 Description of the data

In order to get a comprehensive view of the mortality dynamics in developed countries, we investigate the mortality trends of ten countries: Australia, Canada, France, Germany, Italy, Japan, Netherlands, Sweden, UK and USA. To fix the lack of data for Unified Germany, we decide to employ mortality data concerning the population of West Germany.

We select the generation of males born in 1940, who were aged 71 on 30/06/2011 (we are assuming that the individuals were born in the middle of the year). We fix 30/06/1980 as the observation point, at which the age attained by all the individuals was 40, and we fit the observed survival probabilities $S_{40}(t)$ with $t = 1, ..., 32$.

We take mortality data from the Human Mortality Database `http://www.mortality.org` (data downloaded in April 2016). Since the cohort-data we require are not available for all the countries under study, we resort to period life tables. For each country we collect in a unique matrix all the observed crude death rates $m_x(y)$ (where x denotes the “age last birthday" and y the “calendar year") from 1980 to 2011; then we extrapolate the diagonal starting from $m_{40}(1980)$ to $m_{71}(2011)$. This method does not provide exactly the mortality rates experienced by a certain generation observed throughout life, but can be considered as a good approximation.

We compute the probability that an individual aged 40 reaches age $40+t$, with $t = 1, ..., 32$, namely the survival function $S_{40}(t)$, by the product of the annual survival probabilities:

$$S_{40}(t) = \prod_{i=x}^{x+t-1} p_i \tag{14}$$

the p_i's represent the annual survival probabilities in a dynamic context, where they are assumed to be functions of both the age and the calendar year: $p_{40}(1980)$, $p_{41}(1981)$, ... , $p_{71}(2011)$. They are calculated using the collected crude death rates.

In Figure 1, we show the survival function $S_{40}(t)$, with $t = 0, ..., 32$, for the males born in 1940 in the ten countries under study. We can see that an individual, that was born in Netherlands in 1940 and

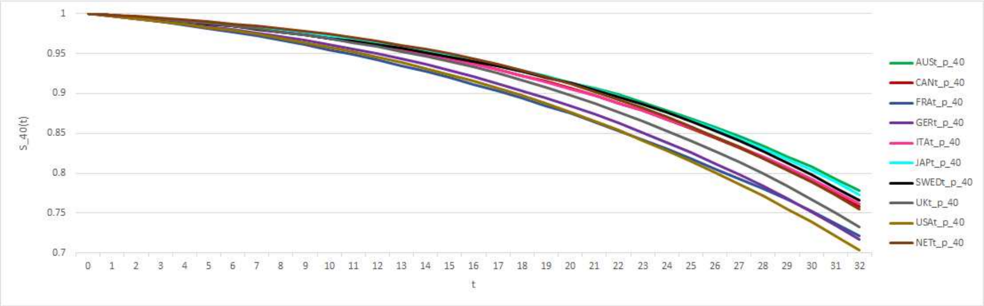


Figure 1: **Survival function $S_{40}(t)$, with $t = 0, ..., 32$ for males born in 1940 in the ten studied countries.**

was still alive in 1980, was characterized by higher probabilities of surviving ages from 41 to 47 than his peers living in the other countries. Australian people belonging to the generation 1940 were the most likely to reach the ages from 60 to 72. The countries for which the survival function took the smallest values are France and USA. The average probability that a 40-aged person in year 1980 survived further 32 year was around 74%, in Europe, and 73% in North America(USA,Canada).

5 Calibration methodology

For each country, we calibrate the two-parameters and the three-parameters Feller process (see Sections 3.1 and 3.2) to the observed survival function $S_{40}(t)$, computed as in Section 4, by minimization of the Root Mean Squared Error(ψ). In (15) we point out such optimization problem we solve using the MATLAB function *fmincon*, that finds the minimum of constrained nonlinear multivariate functions.

$$\min_{\theta} \psi \tag{15}$$

By ψ we mean the square root of the mean of the squared differences between the observed survival probabilities and the theoretical ones, computed, according to the specification of the Feller model we are handling, as in (2) or in (10).

In (15), θ denotes the vector of parameters to be estimated, namely $[a, \sigma]$ in the two-parameters Feller framework and $[\bar{a}, \bar{b}, \bar{\sigma}]$ in the three-parameters Feller setup.

Thus, starting from a vector of initial parameter values, the optimization algorithm finds the optimal parameters of the selected model, namely those parameters that, if used as inputs in (4) and (11), allow us to get a predicted survival function diverging the least possible from the observed one.

We subject the minimization to the following conditions on the parameters (mentioned in Sections 3.1 and 3.2):

- for the two-parameters Feller process: in setting up the optimization algorithm we define the lower and upper bounds for the parameters a and σ such that $a > 0$ and $\sigma > 0$; we check a posteriori that condition (6) holds;

- for the three-parameters Feller process: in the optimization function we require that $\bar{a} > 0$, $\bar{b} > 0$, $\bar{\sigma} > 0$ (as bound constraints) and that condition (13) is satisfied (as a nonlinear constraint).

We set the initial value of λ, $\lambda_{40}(0)$, as $-\ln(p_{40})$.

In Table 1, we report the value of the observed mortality intensity at age 40, $\lambda_{40}(0)$, for the generation under study in all the countries ofinterest. The individuals belonging to the generation 1940, aged 40 in the observation year (time zero), but living in different nations, are characterized by very close mortality intensities. This feature is evident in Figure 1 since all the curves representing $S_{40}(t)$ are almost overlapping when $t = 1$.

Table 1: Initial value of λ, $\lambda_{40}(0)$

Country	$\lambda_{40}(0)$
Australia	0.00238
Canada	0.00207
France	0.00312
West Germany	0.00296
Italy	0.00227
Japan	0.00209
Netherlands	0.00156
Sweden	0.00249
UK	0.00207
USA	0.00306

6 Empirical results

We report the optimal values of the parameters, ψ, the AIC and the BIC of the two-parameters Feller process and the three-parameters one relating to each country, in Table 2 and 3, respectively.

The parameters a and $\bar{b}$ represent the rate of increase of the force of mortality of the Gompertz model underlying the two-parameters Feller process and the three parameters one respectively (see equations (7) and (12)), namely they represent the constant rate at which the probability that an individual dies grows, as age increases, in the deterministic framework described by the Gompertz law.

We find for the generation 1940 born in Netherlands the smallest value of $\lambda_{40}(0)$ and highest value of a and $\bar{b}$. The calibration gives rise to optimal values of a and $\bar{b}$ that are close to each other for all the countries, for both the specifications of the Feller process.

Table 2: Fitting results for the two-parameters Feller process

Country	a	σ	ψ_2	RR_2	AIC_2	BIC_2
AUS	0.0624	0.0003	0.00316	1.1428	-253.48	-361.45
CAN	0.0781	0.0103	0.00027	-0.8161	-410.63	-518.61
FRA	0.0735	0.0207	0.00065	-0.5602	-354.82	-462.80
GER	0.0671	0.0032	0.00148	0.0051	-301.93	-409.90
ITA	0.0704	0.0045	0.00094	-0.3655	-331.37	-439.34
JAP	0.0709	0.0031	0.00155	0.0519	-299.02	-406.99
NETH	0.0977	0.0166	0.00032	-0.7844	-400.45	-508.42
SWED	0.0624	0.0000022	0.00489	2.3145	-225.56	-333.53
UK	0.0868	0.0146	0.00056	-0.6221	-364.54	-472.51
USA	0.0698	0.0084	0.00094	-0.3658	-331.39	-439.36

Table 3: Fitting results for the three-parameters Feller process

Country	$\bar{a}$	$\bar{b}$	$\bar{\sigma}$	ψ_3	RR_3	AIC_3	BIC_3
AUS	6.27E-09	0.062	5.35E-05	0.00316	1.083	-251.49	-361.47
CAN	1.38E-05	0.073	0.00522	0.00035	-0.771	-392.83	-502.80
FRA	9.29E-05	0.052	0.01356	0.00041	-0.728	-381.66	-491.63
GER	4.03E-08	0.067	0.00013	0.00143	-0.058	-302.30	-412.27
ITA	2.19E-07	0.070	0.00033	0.00086	-0.433	-334.80	-444.77
JAP	3.06E-08	0.071	0.00012	0.00150	-0.011	-299.18	-409.16
NETH	4.78E-05	0.079	0.00971	0.00074	-0.516	-344.82	-454.80
SWED	3.08E-08	0.062	0.00012	0.00489	2.223	-223.55	-333.52
UK	3.64E-05	0.075	0.00844	0.00086	-0.434	-334.93	-444.90
USA	8.14E-06	0.067	0.00399	0.00098	-0.355	-326.47	-436.44

Overall, our results show very low ψ's, indicating a good fit to the data extracted from mortality tables. In particular, we obtain the best fit of both the two and the three-parameters Feller process for Canada, France, Italy, Netherlands, UK and USA: for these countries, as Figures 2 and 3 display, the curve of the model-implied survival function matches the curve of the observed survival probability at each point. We get the least satisfactory fit for Australia and Sweden.

Since we calibrate the two and the three-parameters Feller process to an homogeneous sample of data for all the countries (we fit the survival function $S_{40}(t)$ of the same generation and over the same time horizon, namely $t = 1, ..., 32$), we are interested in assessing the goodness of fit of the Feller process for each country compared to the results we obtain for all the other countries under study; we thus resort to the measure we call "RR", computed as follows:

$$RR = \frac{(\psi - \bar{\psi})}{\bar{\psi}} \tag{16}$$

where $\bar{\psi}$ represent the mean of the values of ψ we find for each country:

$$\bar{\psi} = \frac{\sum_{j=1}^{10} \psi_j}{10} \tag{17}$$

with $1, 2..., 10$ denoting the countries arranged in alphabetical order (starting from 1= Australia until 10=USA). A negative value of “RR" indicates that the Root Mean Squared Error of the model we are considering is less than $\bar{\psi}$, while a positive value indicates a value of ψ higher than the average. The higher the RR is, in absolute terms, the farther the quality of fit for a given country is from the average performance of the Feller process over all the countries. For both the specifications of the Feller process, RR for Australia and Sweden is high and bigger than 1, confirming that the models for such countries are characterized by the worst quality of fit. Generally speaking, the smallest distance between ψ and $\bar{\psi}$ does not necessarily denote that the goodness of fit of the model is the most satisfactory one; for example, in the two-parameters Feller process setup, Germany is the country with the minimum gap between ψ and $\bar{\psi}$ but is not the country for which we get the best fit (which is, instead, Canada): this means that Germany is the nation for which the goodness of fit of the Feller process diverges the least from the average goodness of fitness of all the calibrated models, including the best and the worst “performances". Anyway, since within our application we are able to obtain very good results for most of countries, we can consider $\bar{\psi}$ as the benchmark representing a good fit.

The Akaike information criterion (AIC) is a measure of the relative quality of statistical models for a given sample of data. Given a set of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection. Moreover, AIC is founded on information theory: it offers a relative estimate of the information lost when a given model is used to represent the process that generates the data. In doing so, it deals with the trade-off between the goodness of fit of the model and the complexity of the model. The model with the lowest AIC is preferred.

We compute the AIC as follows:

$$AIC = 2k + n \ln(\tau) \tag{18}$$

where τ is the residual sum of squares.. When each candidate model assumes that the residual are distributed as independent identical normal distribution (with mean zero) and we deal with the least squares model fitting, the maximum likelihood estimate for the variance of the model's residual distributions is τ/n. k, that is the number of parameters, is equal to 2 for the two-parameters specification of the Feller process and 3 for the three-parameters specification. n is equal to number of observations, 32, in both cases.

As AIC, the Bayesian information criterion (BIC) or Schwarz criterion is a criterion for model selection among a finite set of models and introduces a penalty term for the number of parameters in the model; such penalty term is larger in BIC than in AIC.We compute BIC in terms of the residual sum of squares (τ):

$$BIC = n \ln(\tau/n) + k \ln(n) \tag{19}$$

where $k = 2,3$, according to the selected model, and $n = 32$.

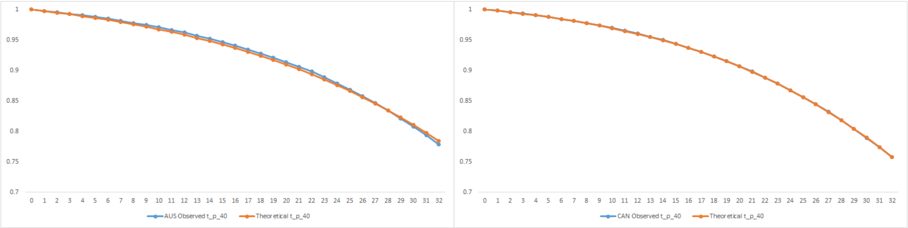
In Tables 2 and 3 we report also AIC and BIC criteria and in Table 4 we stress the value of the difference between the AIC computed for the two-parameters Feller models and that one we obtain for the three-parameters Feller models. We calculate such difference also for BIC criterion.

Table 4: Difference between AIC and BIC of the two specifications of the Feller process

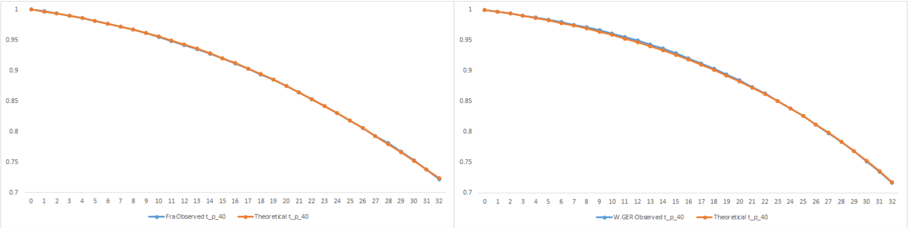
Country	Difference AIC	in Percent	Difference BIC	in Percent
Australia	1.98	-0.8%	-0.02	0.0%
Canada	17.81	-4.5%	15.81	-3.0%
France	-26.84	7.6%	-28.84	6.2%
West Germany	-0.37	0.1%	-2.37	0.6%
Italy	-3.43	1.0%	-5.43	1.2%
Japan	-0.16	0.1%	-2.16	0.5%
Netherlands	55.63	-13.9%	53.63	-10.5%
Sweden	2.01	-0.9%	0.01	-0.0%
UK	29.61	-8.1%	27.61	-5.8%
USA	4.92	-1.5%	2.92	-0.7%

The two Feller models provide close values of AIC and BIC for all the countries. Since AIC and BIC are negative, a positive difference in Table 4 denotes that AIC (or BIC) of the three-parameters Feller process is larger than AIC (or BIC) of the two-parameters Feller process. In 60% of cases, AIC of the three-parameters Feller model is higher than AIC of the two-parameters Feller model, and in 50% of cases BIC of the three-parameters model is larger than BIC of the two-parameters model. We can see that the biggest positive gap between AIC (BIC) of the two different mortality models concern Netherlands, which means that the two-parameters Feller model gives better AIC and BIC results. We see that the differences are small and some times of different signs indicating that the accuracy gained by adding a parameter to the fit is not enough to compensate for the added complexity of the model.

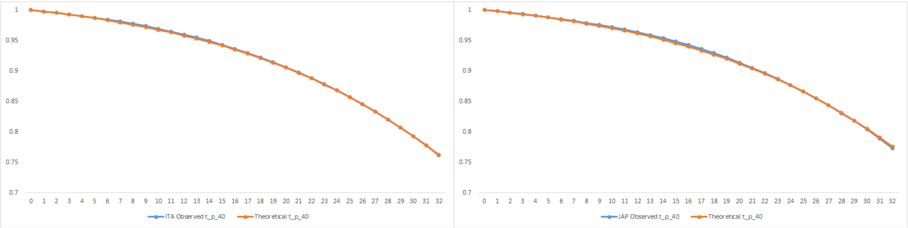
On the basis of the evidence provided by the calibration of the two Feller models, differently parameterized, we choose the two-parameters Feller models for our further applications. The reason behind this choice is mainly the statistical principle of parsimony. Occam's razor suggests "Shave away all but what is necessary". Actually, although the three-parameters Feller process proves to be a more satisfactory stochastic model for the mortality intensity process from a theoretical point of view, within our application it does not turn out to improve enough the quality of the fit, as AIC and BIC show. Therefore, among our competing hypotheses we are induced to select the one with the fewer parameters.



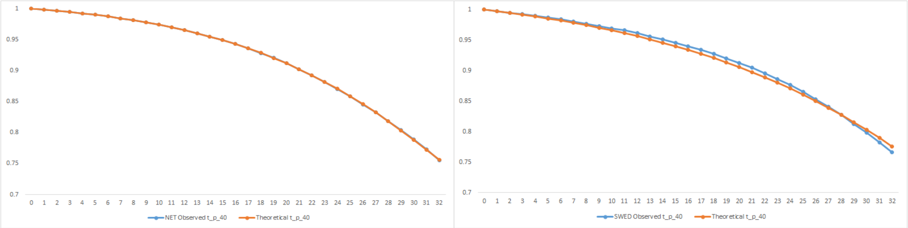
(a) Australia **(b)** Canada



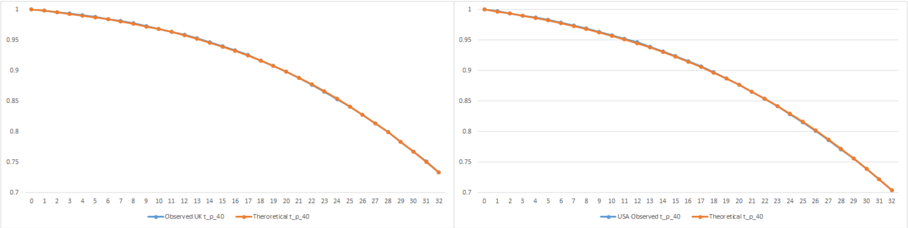
(c) France **(d)** West Germany



(e) Italy **(f)** Japan

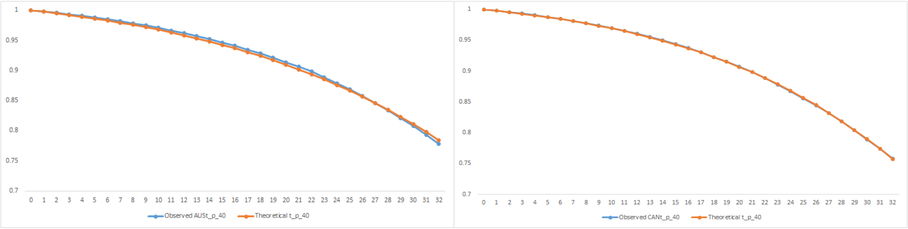


(g) Netherlands **(h)** Sweden

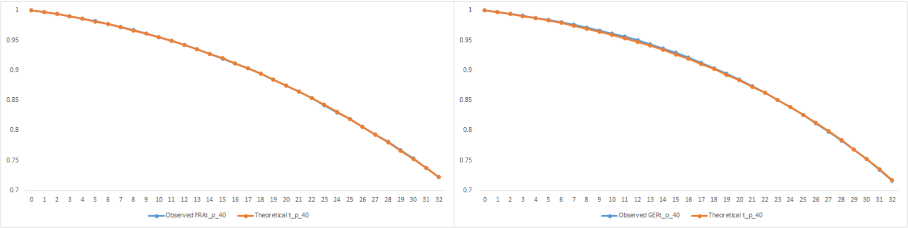


(i) UK **(j)** USA

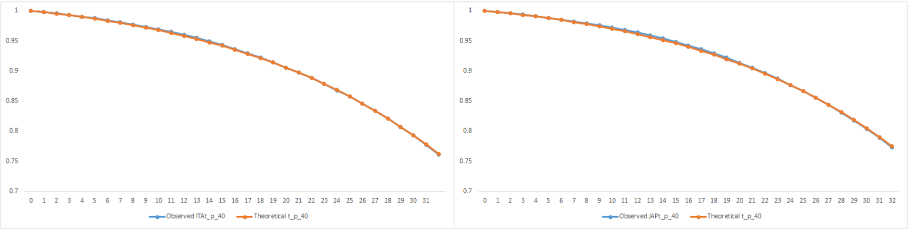
Figure 2: Observed and fitted survival probabilities (two-parameters Feller process).



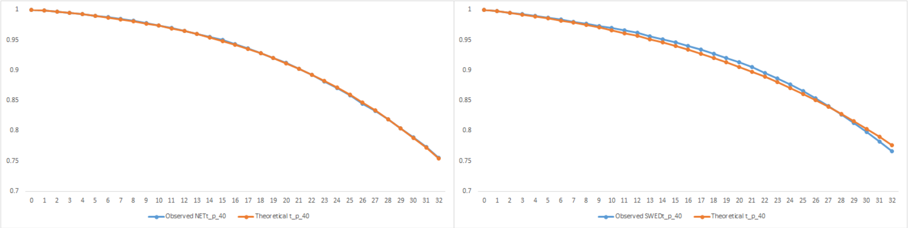
(a) Australia **(b)** Canada



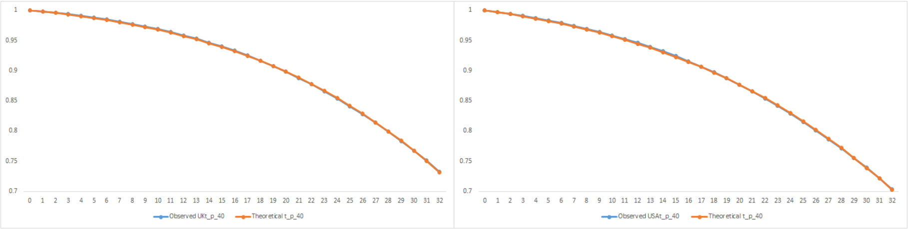
(c) France **(d)** West Germany



(e) Italy **(f)** Japan



(g) Netherlands **(h)** Sweden



(i) UK **(j)** USA

Figure 3: Observed and fitted survival probabilities (three-parameters Feller process).

7 Conclusion

In this paper, we propose a unified way of dealing with the stochastic mortality risk modelling and present results for ten developed countries. It represents a first important contribution to the development of a common modelling framework for the dynamics of mortality and interest rates. Specifically, we are interested in addressing simultaneously the modelling of the dynamics of mortality and interest rates in an affine diffusion setup and in measuring the degree of dependence between them implied by the chosen theoretical models. It is a way to explore effects that have not yet been seen in the data.

As a first, in this study, we calibrate by minimization of the Root Mean Square Error two different specifications of the Feller process (with two and three parameters)to the observed survival functions of the generation of males born in 1940 (ages 40-71) in ten developed countries. The choice of this process is dictated by the fact that similar equations can also be used to model the dynamics of interest rates (Cox Ingersoll Ross model). Comparing the optimal parameters of the calibrated models among countries, we check the existence of common historical trends. Furthermore, we assess the relative goodness of fit of the Feller process for each country compared to its average performance for all the other countries under study.

For most of the countries we find a very good fit to the data extracted from the mortality tables. On the basis of the principle of parsimony, we choose the two-parameters Feller process for the applications we will perform within our further research. This framework is paving the way to study the relation between interest rates and mortality as we will also model interest rates with similar processes. Having all these variables (10 mortality intensity, 10 short-term (3m) interest rates, 10 long-term (10y) interest rates) modelled within the same framework, we can then explore with Monte Carlo simulation the dependence between all the variables and learn more about its short and long termdynamics.

References

[1] D. Blake A. J. G. Cairns and K. Dowd. Pricing death: Framework for the valuation and securitization of mortality risk. *Astin Bulletin*, 36(1):79–120, 2006.

[2] G. Apicella and M. Dacorogna. A Comprehensive Study of Mortality Dynamics in Ten Developed Countries Using the Feller Process. *in Ibit 2016 Proceedings, Paestum*, 2016.

[3] K. R. Berends and Plestis T. Rosen R.J. McMenamin, R. The Sensitivity of Life Insurance Firms to Interest Rate Changes. *Economic Perspectives, Available at SSRN:* `http://ssrn.com/abstract=2386163`, 37(2), 2013.

[4] E. Biffis. Affine processes for dynamic mortality and actuarial valuations. *Insurance:Mathematics and Economics*, 37:443–468, 2005.

[5] C. Blackburn and M. Sherris. Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics*, 53:64–73, 2013.

[6] M. Dacorogna and M. Cadena. Exploring the dependence between mortality and market risks. *SCOR Papers n33*, 2015.

[7] M. Dacorogna and M. Kratz. Living in a stochastic world and managing complex risks. *Available at SSRN:* `http://ssrn.com/abstract=2668468`, 2015.

[8] M. Dahl. Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics*, 35:113–136, 2004.

[9] D. Duffie and K.J. Singleton. *Credit risk*. Princeton Press, 2003.

[10] M. Grasselli G. Deelstra and C. Van Waverberg. The role of the dependence between mortality and interest rates when pricing guaranteed annuity options. *SSRN Archive*, 2015.

[11] L. Jalen and R. Mamon. Valuation of contingent claims with mortality and interest rate risks. *Mathematical and Computer Modelling*, 49:1893–1904, 2009.

[12] E. Luciano and E. Vigna. Non mean reverting affine processes for stochastic mortality. *Carlo Alberto Notebook 30/06 and ICER WP 4/05*, 2005.

[13] E. Luciano and E. Vigna. Mortality risk via affine stochastic intensities: calibration and empirical relevance. *Belgian Actuarial Bulletin*, 8(1):5–16, 2008.

[14] A. Mahayni and D. Steuten. Deferred life annuities: on the combined effects of stochastic mortality and interest rates. *Review of Managerial Science, Vol. 7, Issue 1, pp 1-28*, 2013.

[15] K. Ignatieva M.C. Fung and M. Sherris. Systematic mortality risk: An analysis of guaranteed lifetime withdrawal benefits in variable annuities. *Insurance:Mathematics and Economics*, (58):103–115, 2014.

[16] M. A. Milevsky and S. D. Promislow. Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics*, 29:299–318, 2001.

[17] E. Pitacco. Survival models in a dynamic context: a survey. *Insurance: Mathematics and Economics*, 29:270–298, 2004.

[18] E. Pitacco. Mortality and Longevity: a risk management perspective. *In Proceedings of the 1st IAA Life Colloquium, Stockolm*, 2007.

[19] M. M. Dacorogna R. Bürgi and R. Iles. Risk Aggregation, dependence structure and diversification benefit. *in "Stress Testing for financial institutions", edited by Daniel Rösch and Harald Scheule, Riskbooks, Incisive Media, London*, chap. 12:265–306, 2008.

[20] D. F. Schrager. Affine stochastic mortality. *Insurance: Mathematics and Economics*, 38:81–97, 2006.

[21] S. Ortobelli S. Rachev F. J. Fabozzi V. Russo, R. Giacometti. Calibrating affine stochastic mortality models using term assurances premiums. *Insurance: Mathematics and Economics*, 49:53–60, 2011.

[22] R. Mamon X. Liu and H. Gao. A generalized pricing framework addressing correlated mortality and interest risks: a change of probability measure approach. *Stochastics An International of Probability and Stochastic Processes*, 86:4:594–608, 2014.

SCOR

The Art & Science of Risk

SCOR Papers

SCOR Paper N°19 - May 2012
Microscopic longevity modeling and its practical applications

SCOR Paper N°20 - June 2012
How Long Will We Live? A Demographer's Reflexions on Longevity

SCOR Paper N°21 - February 2013
Microscopic longevity modeling and its practical applications

SCOR Paper N°22 - April 2013
Non-Life Insurance Market Growth in China: can we predict it?

SCOR Paper N°23 - May 2013
A new Dividend Strategy in the Brownian Risk Model

SCOR Paper N°24 - May 2013
Does risk diversification always work? The answer through simple modelling

SCOR Paper N°25 - July 2013
Financial Valuation in a Volatile Scenario of the Guaranteed Minimum Withdrawal Benefit (Gmwb) Policy

SCOR Paper N°26 - September 2013
The risk-free rate: an inescapable concept?

SCOR Paper N°27 - October 2013
Are great earthquakes clustered?

SCOR Paper N°28 - February 2014
Solar storms and their impacts on power grids - Recommendations for (re)insurers

SCOR Paper N°29 - September 2014
A game-theoretic approach to non-life insurance market cycles

SCOR Paper N°30 - September 2014
Explicit Föllmer-Schweizer decomposition of life insurance liabilities through Malliavin calculus

SCOR Paper N°31 - January 2015
The use of economic scenarios generators in unstable economic periods

SCOR Paper N°32 - March 2015
An Integrated Notional Defined Contribution (NDC) Pension Scheme with Retirement and Permanent Disability

SCOR Paper N°33 - April 2015
Exploring the Dependence between Mortality and Market Risks

SCOR Paper N°34 - November 2015
A Change of Paradigm for the Insurance Industry

SCOR Paper N°35 - January 2016
The Globalization of Infectious Diseases

SCOR Paper N°36 - February 2016
Spatial Risk Measures and Applications to Max-Stable Processes

SCOR Paper N°37 - March 2016
ALM Model: The contribution of Fuzzy Logic to behaviour modelling

SCOR Paper N°38 - April 2016
Explicit diversification benefit for dependent risks