Dottorato in Scuola di Scienze Statistiche

Curriculum in Scienze Attuariali- XXX Ciclo

Stochastic Mortality in a Complex World: Methodologies and Applications within the Affine Diffusion Framework

Tutor
Prof.ssa Marilena Sibillo

Candidato
Giovanna Apicella
Matricola 1642949

Sessione di Dottorato

Anno Accademico 2016/2017

SAPIENZA
UNIVERSITÀ DI ROMA

Tesi di Dottorato

Stochastic Mortality in a Complex World: Methodologies and Applications within the Affine Diffusion Framework

Facoltà di Ingegneria dell'Informazione, Informatica e Statistica
Dipartimento di Scienze Statistiche
Dottorato di Ricerca in Scuola di Scienze Statistiche-Curriculum in Scienze Attuariali-XXX Ciclo

Candidato
Giovanna Apicella
1642949

Tutor
Prof.ssa Marilena Sibillo

Coordinatore della Scuola
Prof. Pier Luigi Conti

Referente curriculum in Scienze Attuariali
Prof. Fabio Grasso

A/A 2016/2017

Contents

Aknowledgements.
The PhD programme has represented a significant training ground for learning and researching and also an important human experience. I have met several people, whose feedback has provided valuable insights into the topics addressed in this Thesis. My gratitude goes to all these people. The following few lines are devoted to especially acknowledge some of them.
I wish to thank my supervisor, Prof. Marilena Sibillo, for her persevering, inestimable, guidance and for instilling a huge store of knowledge in me.
I sincerely thank Dr. Michel Dacorogna for his precious help and advices.
I give thanks to the members of the Department of Statistical Sciences, at University of Rome "La Sapienza", for their support.
Last but not least, I express profound gratitude to my parents, Anna and Gerardo, my lighthouse, whose presence and unconditional love are with me in the everyday life and every time I am faced with a new challenge.

THESIS OUTLINE

Context : *a stochastic world with complex risks.*

The globalization gives rise to multiple networks that make risks more complex. Insurance and reinsurance deal with several risks and need to: (i) aggregate the risks, by modelling their dependence, and (ii) assess the influence of such a dependence on the diversification benefit. For life (re)insurance companies it is important to address the prospective relationship between their primary sources of risk: mortality and interest rate risks.

Motivation : *the need for proper quantitative tools for modelling potentially interactive financial and demographic risk factors.*

The introduction of market-consistent accounting (see e.g. IASB (2004)) and risk-based solvency requirements (see e.g. EIOPA (2009)) has called for the integration of mortality risk analysis into stochastic valuation models. Therefore, both financial and mortality risks require a stochastic representation. Furthermore, in recognition of the fact that, in times of extreme events, risks can be more correlated (for example, as in the case of the collapse of the World Trade Center towers on September 11 2001, when insurance products of lines of business, such as property and life, which were thought to be independent until then, were triggered simultaneously), modelling the evolution of mortality and interest rate dynamics in a common mathematical framework can be a proper choice.

Original Contributions : *stochastic representations of mortality in the affine diffusion framework.*

In several works in the literature, the affine processes, thanks to their computational tractability and flexibility, were used as the processes driving the evolution of both demographic and financial risk factors, with and without explicitly considering a possible dependence between them. The core of this Thesis is to make use of affine processes and exploit their properties for modelling stochastic mortality dynamics. Our original contribution consists in using the affine processes with two different purposes of application: (i) for fitting and comparing past mortality trends among different Countries; (ii) for designing a multiplicative affine model for the future evolution of mortality, by combining two components: the forecast provided by any existing mortality

model, representing the deterministic baseline, and an affine driving process that stochastically affects the baseline over the forecasting time horizon. The so structured model, not only is affine, thus fitting well our targets, but, when assessing its forecasting performance, it proves to be parsimonious and to provide a more accurate forecast with respect to the baseline.

Chapter 1

Introduction

> Everything should be made as simple as possible but not simpler
>
> *Albert Einstein*

1.1 Context and Motivation of this Thesis

1.1.1 Risks and their interactions: diversification benefit and dependence

The notion of risk, as it became prevalent in finance and economics during the XX^{th} century, was well summarized in the $XVIII^{th}$ century by the French philosopher Etienne Bonnot de Condillac, who qualified risk as "The chance of incurring a bad outcome, coupled, with the hope, if we escape it, to achieve a good one". The future outcomes are stochastic, but risk is measurable, since we can know the probability associated with each outcome, both good and bad (see Knight (1921)). So insurance companies take risks, hoping that the outcome will turn out to produce a surplus either industrial or financial, but have also available sophisticated risk modelling tools allowing them to assess the probability of rare events and to manage the consequences of the bad realization of any of the risks they are exposed to (Quantitative Risk

Management).

Nowadays, as it happened to our forebears in the past, we experience extreme events, but the multiple networks arising from the globalization, the Internet communication and the global economic development could propagate any event, either due to natural or manmade catastrophes, in a very short time. Such acceleration allows us to see immediately the possible bad impacts of any catastrophe occurring anywhere in the world and study their interactions (that potentially could have existed also in the past). At the same time, these networks make the risks more complex and could increase the severity of the losses (see Dacorogna and Kratz (2015)).

In handling such complex world, following a scientific approach, aiming at simplifying the problem and making it treatable and experimentally provable, is the appropriate way: learning from the available observations, detecting the main contributions to the phenomenon and modelling such contributions. Exploring the dependence between risks is fundamental for understanding their real impacts and consequences, because dependence reduces the possibility to diversify the risks (Bürgi et al. (2008)). In light of the fact that risks are more dependent in extreme situations, 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.

Anyway, any rigorous scientific approach, consisting in mathematically modelling the behaviour extrapolated from the past data to deduce future scenarios, will have to be coupled with: 1. a taxonomy of the existing risks and their complexity; 2. the development of procedures and guidelines for dealing with them. Therefore, the risk management of the future will have to rely on both qualitative risk assessment and quantitative risk modelling. Scientists and managers will have to be able to reach synergy between entrepreneurship and creativity, being essential for the economic development and the prosperity, and the protection from risks. They are unavoidable components of innovation and progress.

The Solvency II Directive, in force since 1^{st} of January 2016, sets economic risk-based solvency requirements across all EU members; it requires the estimation of the Value-at-Risk (VaR) at the 99.5% probability. Analogously, the Swiss Solvency Test (SST) is a modern supervisory tool adopting a risk-based approach, requiring the estimation of the Tail-Value-at-Risk 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 the dependence between the various risks, especially in the extreme situations, whereby we supsect that dependence is much stronger than in normal situations and thus affect significantly the diversification benefit measured at high thresholds (99 or 99.5%).

In order to estimate such risk measures, (re)insurance companies are required to identify their overall loss distribution functions. Therefore, other than determining the loss distributions for the risks per se they are exposed to (marginal distributions), they have also to aggregate those loss distributions. This means that dependencies between insurance risks have to be recognized.

When aggregating insurance risks, it is crucial to preserve the important information concerning the tails of the distributions. By definition, there is little data in the extremes. It is thus fundamental that mathematical models perform better and better in producing reliable artificial data, without neglecting any useful information for detecting the dependence.

1.1.2 The prospective dependence between demographic and financial risk factors

As we have already mentioned, insurance and reinsurance, due to their very nature, deal with several risks on their balance sheet. To determine the total risk of a portfolio, (re)insurance companies need to establish the rules for aggregating the various risks that compose it, by modelling their dependence; they, furthermore, need to assess the influence of such a dependence on the diversification benefit. As stressed in Bürgi et al. (2008), insurance and reinsurance live and die from the diversification benefits or lack of it in their risk portfolio and, as it is well known among traders in financial markets, "diversification works the worst when one needs it the most", since, in times of extreme events, risks turn out to be more correlated. For instance, the collapse of the World Trade Center towers, on September 11, 2001, has shown that insurance products of lines of business, such as property and life, which had thought to be independent until then, can be triggered simultaneously; at the same time, the crisis on the stock market caused the depreciation of

insurance companies assets. For the general insurance industry, it is thus crucial to be able to measure and model the dependence between risks. There are a number of purposes for which taking into account the dependencies is needed: risk management, capital management, profitability analysis, for example. Moreover, for life (re)insurance companies, it is important to address the prospective relationship between their primary sources of risk: mortality and interest rate risks.

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 Pitacco (2007b)); furthermore as interest rates change, the values of a life insurer's assets and liabilities change, potentially exposing the company to risk (see Berends et al. (2013)). Indeed, also the dependence between mortality and market risks is a very important component of the risk of life business, since it would link the liabilities to the investments made by (re)insurance companies to back those liabilities.

Risk management should be able to understand the impacts and consequences of each source of risk (i.e. the effects that, individually, the prospective increase of longevity (see Appendix A.2) and prolonged low interest rates could have on the liabilities of pension funds and life insurance companies) but also their interactions. Longevity and its consequences are among the most crucial concerns; on the other side of the coin, it would be very relevant to assess also the effects that an extreme event, like the breaking out of a severe pandemic, could have on the economy and on financial markets in our highly connected world. According to the World Bank study carried out by Burns et al. (2008), a severe flu pandemic could kill up to 70mn people and cause global GDP to contract by up to 4.8% in just one year. Could such an unforeseen scenario affect the price of the assets that life (re)insurance companies hold to pay mortality claims? Accordingly, it would be relevant to investigate the possible dependence between mortality and interest rate risks and, even more so, to increase our knowledge into the degree of such a prospective relationship in case of stress.

A further motivation to the study of the dependence between mortality and interest rate risks, can be found in Miltersen and Persson (2005), where

the concept of forward force of mortality, without any limitations on the dependence between the stochastic behavior of the forward force of mortality rate and the stochastic behavior of the forward rate, was introduced. The Authors argued that the death of individual policy holders, for all practical matters, can be considered independent of the development of the financial market and the general macroeconomic development of a society. Instead, the force of mortality rate in the future, for instance, the average number of death per unit of time, is to a large extent dependent on the development of the economy in general and thereby also the development of the financial market, and the future interest rates in particular. If there is such a dependence between changes in the force of mortality rate over time and the development of the financial market, the market price of mortality risk could very well be different from zero. As a result, the pricing of all types of life-insurance products, in particular derivative products related to changes in mortality, would have to rely on a good estimate of this market price of mortality risk.

Empirical studies about dependence in the literature

Several studies deal with the relationship between population and economy (e.g. DellaVigna and Pollet (2007), Favero et al. (2011) and Favero et al. (2016)).

Furthermore, several empirical investigations have been carried out in order to assess the relationship between fluctuations in economy and in mortality.

Already in Ogburn and Thomas (1922), the influence of economic changes on social conditions in the United States was investigated. The Authors aimed at measuring the amount of concurrence in the fluctuations of certain social conditions (e.g. births, deaths, marriages) with the business cycles, by computing correlations. Social changes were thus correlated not to the lasting changes in the economic order but in the oscillatory changes of short duration. As to the results concerning the association between mortality and economic changes, the fluctuations in the curve of the death rates from 1870 to 1920 seemed to correspond somewhat with the business cycles. The correlation was found to be positive and fairly high, amounting to +0.57, and with cycles from nine-year moving averages, being equal to +0.63.

Similarly, the study performed in Tapia Granados (2005) was aimed at ascertaining if short-term oscillations of the death rates, with respect to the

long-term falling trend observed throughout the 20th century, were related to the business cycles. For doing so, age-adjusted total mortality and mortality for specific population groups, ages and causes of death were transformed into rate of change or percentage deviation from trend, and were correlated and regressed on indicators of the US economy during the 20th century, transformed in the same way. The results showed that the oscillations of the four different economic indicators under study had significant associations with the oscillations of the death rates and that mortality is what the economists call a procyclical variable. Indeed, the decline of total mortality and mortality for different groups, ages and causes accelerated during recessions and was reduced or even reversed during periods of economic expansion, with the exception of suicides which increase during recessions.

Bruckner et al. (2013) investigated if the evidence from the United States that period life expectancy improved during the Great Depression (cf. Tapia Granados and Diez Roux (2009)) could apply also to European Countries. By exploiting ARIMA models, the Authors assessed if the changes in period life expectancy observed during the Great Depression were unusual, given the past history of the mortality time series. As evidenced by statistically significant coefficients on the Great Depression variable, period life expectancy in 1930, for females in Italy and France and for males in England and Wales, rose above expected levels. In the remaining tests they performed, period life expectancy during the Great Depression remained within intervals consistent with history. Taken together all the Countries under study, the null hypothesis that, during the Great Depression, life expectancy in Europe remained within expected values could not be rejected.

Ribeiro and di Pietro (2009) and Dacorogna and Cadena (2015) addressed empirical investigations on the relationship between mortality and financial risks.

In Ribeiro and di Pietro (2009), the focus was on "macro" longevity risk, that is the longevity risk associated with national populations, large defined benefit pension plans and insurer annuity portfolios, and on the portfolio benefits that investors could have if going short longevity. The Authors found that longevity risk in securities, such as mortality forwards, is largely uncorrelated to equities and bonds, thus turning out to be a good diversifier for investors. The empirical study they performed consisted in analyzing the correlation between asset prices and changes in mortality rates and, also,

between asset prices and the implied longevity returns from annuities (namely the component of life annuity returns that is not attributable to interest rate changes). The correlation, for different age groups, between changes in annual US mortality rates and annual 10-year constant maturity US Treasury returns, computed by using data going back to 1990, was found to be close to zero, generally not statistically different from zero at the 5% significance level. Analogous results concerned the correlation between changes in US mortality rates and returns on the Dow Jones Industrial Average (DJIA), for different holding periods. But, while mortality changes and DJIA returns showed low correlation in normal periods, a larger dependence between them was found in case of extreme events, namely when the change in mortality rate was more than a 0.5 standard deviation move with respect to the previous year and to an average of mortality rates over the previous five years. Indeed, a spike up in US mortality rates turned out to be associated with same-year negative DJIA returns (-4.6% versus an average performance of 4.9% over the full sample) and was also followed by below average annual returns over the next five years. The Authors concluded that being short longevity could be a good equity hedge during extreme events.

In Dacorogna and Cadena (2015), a statistical approach to explore empirically the dependence between mortality and market risks in case of stress was developed. Exploiting data for six developed Countries, over 80 years, the Authors built a yearly index based on the average life expectancy at birth for the whole population, extracted its worst 10 years and looked at the average performance of various economic and financial indicators (GDP, CPI, stock index and 10y government yield) in the extremes and over the whole sample. Such methodology allowed them to look at the effect beyond a simple short-term dynamics. Compared to the study perfomed in Ribeiro and di Pietro (2009), thus, the Authors explored various economic/financial variables and chose a larger number of Countries than only the US. Furthermore, for selecting mortality spikes, they looked at the 10 worst yearly change in mortality over the sample; such a method, which is based on the empirical quantiles (cf. Ferreira and De Haan (2015)), let them not assume the existence of the second moment of the underlying distribution. The lead-lag correlation analysis, according to the methodology developed in Dacorogna et al. (2001), was performed in order to see if any of the economic or financial indicators had an effect on the mortality indices (lag-analysis), or if the mortality indices had lagged effects on the economic or financial indices (lead-analysis).

Their study showed some direct weak effects on the financial markets and on the correlation, during the worst years for the mortality index, and a lagged effect of mortality on the macro-economic data.

1.1.3 The need for stochastic mortality models

As explained in the previous Section, together with financial risk, *mortality risk* is the primary risk factor for both social security system and private life annuities business.

Mortality risk may emerge in three different ways (see Pitacco et al. (2009)):

- as *process risk* or *insurance risk*, resulting in random fluctuations of the observed mortality rates around expected mortality rates.
  Process risk is a "pooling risk", since its severity decreases when the portfolio size increases;

- as *uncertainty risk*, consisting in systematic deviations of the observed mortality rates from those expected. When adult-old ages are concerned, such a risk is called *longevity risk*.
  The term *uncertainty risk* is used to refer to *model* and *parameter risk* jointly, where the former results in deviations due to the shape of the time-pattern implied by the mortality model, in respect of both age and calendar year, and the latter consists in deviations due to the level of the parameters of the mortality model.
  Uncertainty risk cannot be lowered increasing the portfolio size, since it concerns the portfolio as an aggregate risk (cf. Olivieri and Pitacco (2006));

- as *catastrophe risk*, resulting in sudden and short-term rises in the frequency of death. It is worth noting that while catastrophe risk, being due to period effects, involves short-term shocks, longevity risk endures over longer time horizons.

As stressed in Dahl (2004), dealing with the financial and the systematic mortality risks in such a way to choose both the interest rate and the mortality intensity to the safe side may turn out to be an unsuitable "modus operandi" for life insurance companies, especially when contracts running for 30 years or more are concerned. This phenomenon can be seen for the interest rates, in

light, for instance, of the low returns on bonds we are currently experiencing. If the uncertainty in financial markets poses an immediate issue, the level of mortality intensity gives rise to a more long-term and permanent problem, because of the lower speed at which changes in the mortality intensity seem to occur. Both financial risks and mortality risks deserve careful analyses and require the adoption of proper management solutions.
Under the deterministic approach to mortality modeling, the biometric functions, e.g. q_x, l_x, etc., involved in the actuarial models, even if correctly meant as probabilities, expected values, etc., are only used to calculate expected or best-estimate values and no measure of riskiness is allowed for. When assessing the risk profile of a life insurer, such a deterministic approach should be rejected and replaced by a *stochastic approach* allowing for process risk and uncertainty risk (or systematic mortality risk) as well (see Olivieri and Pitacco (2006) and Olivieri and Pitacco (2012)).
The stochastic approach allows to capture two important features of the mortality intensity: time dependency and uncertainty of the future development. It results that, when stochastic mortality modelling is performed, premiums and future liabilities can be assessed more realistically and accurately, since possible trends in the mortality intensity and the market attitude towards systematic mortality risk can be taken into account; furthermore, the risk of the insurance companies associated with the underlying mortality intensity can be quantified (see Dahl (2004)).
The three following reasons, stressed in Biffis et al. (2010), further motivate the need for the stochastic approach to mortality modelling: the extent of mortality decreasing trend has posed severe challenges to the insurance and pension industry, emphasizing the importance of demographic trends forecasting; the introduction of market-consistent accounting (see e.g. IASB (2004)) and risk-based solvency requirements (see e.g. EIOPA (2009)) has called for the integration of mortality risk analysis into stochastic valuation models; mortality-linked securities have attracted the interest of capital market investors, who in turn demand transparent tools to price demographic and financial risks in an integrated fashion.

Deterministic models for mortality consist in assuming suitable parameterized analytical models or adopting adjusted/projected mortality tables.

Within the stochastic mortality model framework, instead (see Pitacco et al. (2009)):

- observed mortality rates are outcomes of random variables representing past mortality;
- forecasted mortality rates are estimates of random variables representing future mortality
- probability distributions for the random numbers of deaths and a statistical structure linking forecasts to observations must be specified.

Cairns et al. (2006a) pointed out the following criteria that any 'plausible' stochastic mortality model should meet: it should keep the force of mortality positive, be consistent with historical data and be characterized by biologically reasonable long-term future dynamics; as to this point, since historical mortality curves are always upward-sloping at the oldest ages, an 'inverted' mortality curve, in which mortality rates for the elderly fall with age, would not be reasonable. Long-term deviations in mortality improvements should not be mean-reverting to a pre-determined target level, even if this target is time dependent and allows for mortality improvements; instead, short-term deviations from the trend due to local environmental fluctuations might be mean-reverting around the stochastic long-term mean. The reason behind this last criterion is that if mortality improvements have been faster than anticipated in the past then the potential for further mortality improvements will be remarkably reduced in the future. The model should also be comprehensive enough to deal appropriately with the current pricing, valuation or hedging problem. A further criterion might be taken into account, provided that it does not override any of the criteria we have just mentioned: it should be possible to value the most common mortality-linked derivatives using analytical or fast numerical methods.

In the next Subsections a review of several studies proposing stochastic models in continuous time for the future evolution of mortality is provided. In Appendix A.3 a brief review of stochastic models in discrete time is also provided.

1.1.4 Stochastic mortality models in continuous time under the assumption of independence between mortality and interest rates dynamics

Short-rate models for the development in continuous time of the force of mortality have been proposed by Milevsky and Promislow (2001), Dahl (2004) and Biffis (2005).

The motivation underpinning the work of Milevsky and Promislow (2001) was the need for coming to a valuation of options on mortality-contingent claims. Indeed, most US-based insurance companies were offering variable annuity contracts with the long-term option to annuitise the policy at a predetermined rate over a pre-specified period of time. The holders of such policies, with guaranteed annuitisation rates, were thus granted an option on two underlying stochastic variables: future interest rates and future mortality rates. The heuristic idea was to treat the underlying life annuity as a defaultable coupon-bearing bond, where the default occurred at the exogenous time of death. In practice, the option to annuitise was an American-style contingent claim on a corporate bond. The methodology developed in Milevsky and Promislow (2001) consisted in stochastically modelling the future *hazard-plus-interest rate process*, defined as the sum of the instantaneous risk-free rate of interest and the hazard rate (or force of mortality), that were assumed to be independent. From an actuarial perspective, the force of mortality was considered as a random variable forward rate, whose expectation was the force of mortality in the classical sense. Assuming the equivalence between the instantaneous force-of-mortality curve (in insurance) and the instantaneous forward-rate curve (in finance), the instantaneous force of mortality was defined directly from the probability of survival, which in turn was driven by the stochastic hazard rate. In their example, Milevsky and Promislow assumed a CIR process for the short rate and the mean-reverting Brownian Gompertz (MRBG) process for the hazard rate. This process, representing the stochastic analogue of the Gompertz survival function, was expected to grow exponentially, had a variance proportional to the value of the hazard rate, never hit zero, and exhibited mean reversion.

In Dahl (2004), the mortality intensity was modelled by a fairly general diffusion model, which included the mean-reverting Brownian Gompertz model proposed by Milevsky and Promislow (2001) as a special case. In particular,

the concept of *affine mortality structure* for a given cohort was discussed. The definition of an affine mortality structure was almost analogous to the definition of an affine term structure, see e.g. Björk (1998); within such a structure, the survival probabilities for a given cohort could be expressed by relatively simple expressions. The model chosen for describing the dynamics of mortality intensities for fixed cohorts was the extended Cox-Ingersoll-Ross model, being mean reverting around the time and cohort dependent level, and ensuring strict positivity of the mortality intensity for the chosen cohort, provided that a specific condition on the coefficients of the Stochastic Differential Equation was fulfilled (see Maghsoodi (1996)). Furthermore, the extended Cox-Ingersoll-Ross model admitted an affine mortality structure for the mortality intensity, so that closed form expressions for the survival probabilities could be found, provided that the PDEs for the drift and the volatility term could be solved.
In Dahl and Møller (2006) a special case of the general affine mortality structures was studied. Given an initial mortality intensity curve, the Authors assumed that the mortality intensity at a given future point in time and at a given age could be obtained by correcting the initial mortality intensity by the outcome of some underlying process, modelled through a time-inhomogeneous Cox-Ingersoll-Ross (CIR) model. The mortality intensity was described by a CIR model as well and the survival probabilities could be obtained by using standard results for affine term structures.

As mentioned in Section 1.1.3, the release of IASB (2001) and IASB (2004) by the International Accounting Standards Board (IASB) prompted a sound and articulated valuation of insurance contracts, stressing the need to deal explicitly with all sources of risk and advocating the accounting of insurance liabilities at market value. In light of such a regulatory framework, in Biffis (2005), affine jump-diffusions were proposed as effective tools for addressing the trade-off between complexity and computational tractability of pricing and estimation. Indeed such stochastic processes had already proven to be useful in developing tractable and flexible dynamic asset pricing models, with applications to the term structure of interest rates (cf. Dai and Singleton (2000)), the stochastic volatility for currency and equity prices, and to the risk of default of corporate bonds and other credit-risky securities (cf. Lando (1998) and Duffie and Singleton (2003)); in Biffis (2005), affine jump-diffusions were used as driving processes for the evolution of both demographic and financial risk factors. Furthermore, following the stream

of research drawing a parallel between life insurance business and credit-sensitive securities, first developed in Artzner and Delbaen (1995) and used in Milevsky and Promislow (2001) and in Dahl (2004), the *intensity-based* approach to credit risk modelling was exploited to derive closed form expressions for the pricing and reserving of several life insurance contracts, either traditional, indexed and unit-linked.

Such an approach was used also in Luciano and Vigna (2005), where the random evolution of mortality was described by using doubly stochastic (or Cox) processes. Luciano and Vigna found that time-homogeneous mean reverting affine processes 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 Pitacco (2004)), and produce unrealistic survival function at very old ages. Furthermore, Luciano and Vigna detected an exponential behaviour of the force of mortality observed and/or extrapolated from the mortality tables; their study was thus focused on non mean reverting affine processes, with deterministic part increasing exponentially, such as the Ornstein Uhlenbeck and the Feller processes, with and without jumps; these models represent natural generalizations of the Gompertz law. An interesting result coming from the calibration of such models was the fact that the stochastic component of the intensity processes seemed to be properly described by negative jumps. In Luciano and Vigna (2008), three processes were selected for describing future evolution of mortality within a single cohort and between different cohorts, in such a way to assess mortality trend. The chosen processes (the Ornstein Uhlenbeck with and without jumps and the Feller processes) proved to satisfy most of the reasonable criteria that a good mortality model should have; furthermore, the simulation showed that the variability of the number of deaths could significantly increase even when the volatility of the mortality intensity was small.
Fung et al. (2014) proposed a one-factor, non mean-reverting and time homogeneous affine process for modeling the mortality intensity process of an individual; compared to the Feller process examined in Luciano and Vigna (2008), under such a specification, thanks to one additional parameter, the mortality intensity had a non zero probability of reaching zero even when the volatility term was positive.

In Schrager (2006), a new model for stochastic mortality based on the literature on affine term structure models was proposed. In contrast to the

approaches of Milevsky and Promislow (2001) and Dahl (2004), considering only a single age at a time, Schrager modelled the mortality intensity of all ages simultaneously. More specifically, the mortality intensity was an affine function of factors driving its uncertainty, whose dynamics were described by a multivariate diffusion process. The instantaneous drift and variance of the factors were affine functions of the factors themselves. When Gaussian factors (e.g following a multivariate Ornstein–Uhlenbeck process) were combined with the functional form for the dependence of mortality intensity on age postulated by Thiele (1872), the Gaussian Stochastic Thiele model was obtained.

1.1.5 Stochastic mortality models in continuous time without limitations on the dependence between mortality and interest rates dynamics

Miltersen and Persson (2005) introduced the concept of forward force of mortality rate without any limitations on the dependence between the stochastic behavior of the forward force of mortality rate and the stochastic behavior of the forward (interest) rate. Inspired by the approach developed in Heath et al. (1992) for deriving the behavior of the term structures of interest rates by no-arbitrage conditions, Miltersen and Persson (2005) derived no-arbitrage conditions on the stochastic behavior of future term structures of forces of mortality rates. By assuming that they could observe the initial term structure of forward forces of mortality rates, their key idea was to use price data from the market to back out the market's best view of the force of mortality rate today and, more importantly, the market's best view of the stochastic behavior of the future force of mortality rate, namely, the brevity (i.e. the risk of premature death) and longevity risk. This and further mortality modelling structures were deeply examined in Cairns et al. (2006a). In such a paper a range of arbitrage-free (or risk-neutral) frameworks for pricing and hedging mortality risk, allowing for both interest and mortality factors to be stochastic, were described and new mortality market models equivalent to the LIBOR and swap market models were introduced.

In Biffis and Millossovich (2006), random fields were used to model the evolution of mortality across different risk-classes or cohorts of insureds, thus providing a random field extension of the stochastic models where the random

intensity of mortality was typically specified with reference to a representative policyholder, such as the stochastic models for dynamic mortality employed by Milevsky and Promislow (2001), Dahl (2004) and Biffis (2005). As stressed by the Authors, treating the intensity of mortality as a random surface allowed to compute explicitly the fair value of some insurance liabilities at aggregate level. Accordingly, the joint effect of financial and demographic risk factors as well as cross-generation (or risk-class) effects were considered in a common framework. In addition, a model for new policyholder arrivals, based on a marked point process, was introduced; mortality, financial markets and policyholder arrivals could be correlated in such a framework.

In order to address the pricing problem of life insurance contracts and provide analytical solutions, Jalen and Mamon (2009) proposed an approach based on the change of the probability measures technique and affine diffusion processes for the description of the evolution of mortality and interest rates. Within the numerical implementation they performed, the prices of basic mortality-linked instruments were computed, under both the assumptions of dependence and independence between mortality and financial factors. Their results suggested that the dependence between mortality and financial factors should not be ignored when pricing and reserving for the instruments with long-term maturities.
In Liu et al. (2014), the dependence between mortality and interest rate risks was further modelled in the valuation of a guaranteed annuity option (GAO). In their setting, the dependence between mortality and interest rates was described by one constant, namely the pairwise linear correlation coefficient.
Starting from the Gaussian approach of Liu et al. (2014) and with respect to a similar purpose, namely deriving pricing formulas for insurance contracts like indexed annuities and GAOs, Deelstra et al. (2016) considered two multifactor specifications that are nested in the general affine framework, namely the multi-CIR and the Wishart model. Among the two models they examined, the Wishart model turned out to be the most flexible and proved to be able to produce the richest structure of dependence.

1.2 The Aim of the Research

In order to be able to benefit from diversification, (re)insurance companies need a good knowledge of the degree of the dependence between the risks they are exposed to, especially in case of crisis, whereas experience has shown that risks can be more correlated.

For a number of purposes (see Section 1.1.2), it is fundamental, for life (re)insurance companies, investigating if mortality and interest rate risks are linked by dependence and, if such dependence turns out to exist, understanding how it affects capital requirements and pricing.

When assessing the risk profile of a life insurer, a stochastic approach to both mortality and interest rates modelling is required (e.g. for assessing premiums and future liabilities realistically and accurately); furthermore, in the literature (cf. Biffis et al. (2010)), the need for developing tools to price demographic and financial risks in an integrated fashion has been stressed.

In this context, a common mathematical framework for modelling the dynamics of stochastic mortality and interest rates, namely for considering the changes in financial and demographic conditions over time, is needed. By common mathematical framework we mean, figuratively, a big box that contains stochastic processes belonging to the same class, e.g. having analogous empirically relevant properties.
For our research, we focus on the class of the affine processes, that have been used in a wide range of applications in financial and actuarial sciences, thanks to their computational tractability and flexibility. For instance, affine processes have been extensively used in modelling the term structure of interest rates. The focus of "affine" term structure models (ATSMs) extends back at least to the studies by Vasicek (1977) and Cox et al. (1985); these and related ATSMs underpin extensive literatures on the pricing of bonds and interest-rate derivatives and are also at the basis of many of the pricing systems used by the financial industry (cf. Dai and Singleton (2000)).

The core of this Thesis is to make use of affine processes and exploit their properties for modelling stochastic mortality dynamics, thus contributing to the "novel" stream of research (with respect to the one relating to the ATSMs) dealing with the applications of the affine processes to the modelling of the evolution of mortality.

1.3 Original Contribution and Numerical Applications

Our original contribution consists in using the affine processes with two different purposes of application: (i) for fitting and comparing past mortality trends among different Countries; (ii) for designing a multiplicative affine model for the future evolution of mortality, by combining two components: the forecast provided by any existing mortality model, representing the deterministic baseline, and an affine driving process that stochastically affects the baseline over the forecasting time horizon. The so structured model not only is affine, thus fitting well our targets, but, when assessing its forecasting performance, it proves to be parsimonious and to provide a more accurate forecast with respect to the baseline. Throughout the Thesis, such contributions are discussed and the related numerical applications are fully explored:

- as a first, in Chapter 2, we calibrate 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. We thus fit coherently the Country-specific past mortality experience and, by comparing the optimal parameters of the calibrated models among the Countries under study, we check the existence of common historical trends. Furthermore, we assess the goodness of fit of each specification of the Feller process for each Country compared to its average performance among all the other Countries under study;
- as a second, in Chapter 3, we illustrate the multiplicative affine model we design for describing the future evolution of mortality. Within such a model, the affine driving factor is tasked with describing the dynamics over time of a measure of the fitting error of the existing mortality model providing the baseline. For our numerical application, explained in Chapter 4, we choose the Cairns-Blake-Dowd (or M5) model as the existing mortality model giving the baseline and the CIR process as the stochastic factor affecting the baseline in a multiplicative way. The resulting model is called mCBD. Using the Italian females mortality data, for fixed ages, and implementing the backtesting procedure, over both a static time horizon and fixed-length windows rolling one-year ahead through time, we empirically test the performance of the CBD

and the mCBD models in forecasting death rates. On the basis of average measures of forecasting errors and information criteria, we demonstrate that the mCBD model is a parsimonious model providing better results in terms of predictive accuracy than the CBD model and showing a stronger potential to gain accuracy in the long-run when a rolling windows analysis (dynamic approach) is performed.

1.4 Notation

As stressed in Pitacco (2004), in a dynamic context, mortality is assumed to be a function of both the age x and the calendar year t. In a rather general setting, a dynamic survival model is a function $\Gamma_{x,t}$, usually with real values, that may represent mortality rates, mortality odds, a force of mortality, a survival function, etc. In the following, we define, in mathematical terms, the realizations of the function $\Gamma_{x,t}$ that are useful within our study.

Let us consider the instantaneous death rate at time t known as the *force of mortality*. Posing ${}_{\Delta x}q_{x,t}$ the probability that an individual aged x at time t dies before age $x + \Delta x$, the force of mortality $\mu_{x,t}$ related to this individual can be expressed as follows:

$$\mu_{x,t} = \lim_{\Delta x \to 0} \frac{{}_{\Delta x}q_{x,t}}{\Delta x} \tag{1.1}$$

If Δx is sufficiently small, we can write:

$${}_{\Delta x}q_{x,t} \simeq \mu_{x,t}\Delta x \tag{1.2}$$

so that the product on the right can be interpreted as a probability. It is commonly assumed that, within each year of age, $\mu_{x,t}$ remains constant, as in (1.3) where k and h are numbers such that, respectively, $x+k$ is a non-integer age and $t + h$ is a non-integer duration:

$$\mu_{x+k,t+h} = \mu_{x,t}, \quad \text{if} \quad k, h \in [0, 1) \tag{1.3}$$

Let $S_t(x)$ be the probability that T_0, the random future lifetime for a newborn, is longer than x attained in the calendar year t:

$$S_t(x) = P[T_0 > x, t], \quad t \geq 0$$

If we assume that $S_t(x)$ is differentiable with respect to x, it is straightforward to show that the force of mortality is given by:

$$\mu_{x,t} = \frac{-\frac{\partial S_t(x)}{\partial x}}{S_t(x)} = -\frac{\partial}{\partial x} \ln S_t(x) \tag{1.4}$$

When $\mu_{x,t}$ is known, equation (1.4) becomes a differential one; integrating between 0 and x, and posing the initial position $S_t(0) = 1$, we have:

$$S_t(x) = \exp\left[-\left(\int_0^x \mu_{u,t} du \right) \right] \tag{1.5}$$

Being:

$$ {}_h q_{x,t} = \frac{S_t(x) - S_t(x+h)}{S_t(x)} \tag{1.6}$$

it results that the probability of death ${}_h q_{x,t}$ can be expressed as in equation (1.7):

$$ {}_h q_{x,t} = 1 - \exp\left[-\left(\int_x^{x+h} \mu_{u,t} du \right) \right] \tag{1.7}$$

Recalling formula (1.3), we have immediately that:

$$p_{x,t} = \exp\left[-\mu_{x,t}\right]; \quad q_{x,t} = 1 - \exp\left[-\mu_{x,t}\right]$$

where $q_{x,t}$ is the probability for a live aged x at time t to die within 1 year, that is before age $x+1$, and $p_{x,t}$ the probability for the same person to be alive at age $x+1$.

We denote by $d_{x,t}$ the number of deaths in one year among the people aged x in t and by $E^c_{x,t}$, the exposed to risk, with x varying from 0 to the ultimate age and t varying with the calendar year (cf. Currie (2014)). The exposed to risk are assumed to be *central exposed to risk*, an average of the living people aged x in t.

The *mortality coefficient* is given by (cf. Olivieri and Pitacco (2010) and Pitacco (2007a)):

$$m_{(x,x+h),t} = \frac{\int_x^{x+h} \mu_{u,t} S_t(u) du}{\int_x^{x+h} S_t(u) du} \tag{1.8}$$

and the *central death rate* is defined as follows, taking into account equation (1.6) and posing $h = 1$ in (1.8):

$$m_{x,t} = m_{(x,x+1),t} = \frac{S_t(x) - S_t(x+1)}{\int_x^{x+1} S_t(u)du} \tag{1.9}$$

It can be demonstrated (cf. Pitacco (2007a)) that the central death rate can be expressed also as in (1.10):

$$m_{x,t} = \frac{d_{x,t}}{E^c_{x,t}} \tag{1.10}$$

In the remainder of the Thesis we refer to central death rates also as crude death rates.

Under the assumption (1.3), the forces of mortality and the central death rates coincide, that is $\mu_{x,t} = m_{x,t}$ (see Pitacco et al. (2009)).

All the age-specific functions we have examined so far, $\Gamma_{x,t}$, if both age x and calendar year t are discrete variables, can be represented by a matrix whose rows correspond to ages and columns to calendar years. Let us assume $\Gamma(x,t) = q_{x,t}$. Then, the annual probabilities of death in the matrix can be read according to three arrangements (cf. Pitacco (2004)):

- a "vertical" arrangement (i.e. by columns),

$$q_{0,t}, q_{1,t}, ..., q_{x,t}, ... \tag{1.11}$$

  corresponding to a sequence of period life tables, each table referring to a given calendar year t;

- a "diagonal" arrangement,

$$q_{0,t}, q_{1,t+1}, ..., q_{x,t+x}, ... \tag{1.12}$$

  corresponding to a sequence of cohort life tables, each table referring to the cohort born in year t;

- a "horizontal" arrangement (i.e. by rows),

$$..., q_{x,t-1}, q_{x,t}, q_{x,t+1}, ... \tag{1.13}$$

  giving rise to the mortality profiles, each profile referring to a given age x.

Within the numerical applications aiming at representing the mortality phenomenon of a given cohort (see Chapter 2), mortality data are collected according to the diagonal arrangement. When interested in describing the mortality profile for a given age (see Chapter 4), the horizontal arrangement is adopted.

Chapter 2

Using Affine Models for Fitting Historical Longevity Trends

2.1 Context and Motivation

This Chapter is grounded in the context described in Chapter 1: a complex world, where the interconnectedness between developed Countries is growing rapidly. Indeed, the World Economy, people's massive mobility, trades and technologies make World's populations become more and more closely linked. Accordingly, survival rates across multiple populations can be affected by common factors in a similar way.

The literature in demographic field has investigated the existence of dependence across multiple populations and, also, of common long-run relationships among different Countries.

In Wilson (2001), a stylized presentation, aiming at explaining the scale and nature of demographic convergence in the second half of the twentieth century, was provided. According to the Author, while huge economic gaps remain between rich and poor Countries, we are moving into a world in which that distinction has a diminishing demographic relevance. Following this initial explanatory investigation, in Wilson (2011) a further study was performed, finding out that it makes sense to view most demographic change over the past half century as falling along a "main sequence" of demographic transition. Using similar graphs as the famous, in astronomy, Hertzsprung–Russell diagram for displaying his results concerning mortality

and fertility trends, the Author concluded that the large majority of the world's population is engaged in a process of demographic convergence.

On the other side of the coin, in Moser et al. (2005), a different outcome was stressed: the distribution of life expectancy at birth worldwide has started to diverge, meaning that global inequality in mortality is increasing. Briefly examining the Authors' quantitative approach to the issue, the global mortality distribution at a point in time was quantified using a dispersion measure of mortality (DMM), whose trends indicate global mortality convergence or divergence. United Nations data for 1950–2000 for all 152 Countries with populations of at least 1 million in 2000 (99.7% of the world's population in 2000) were used for such an analysis. The DMM for life expectancy at birth declined until the late 1980s but has since increased, signalling a shift from global convergence to divergence in life expectancy at birth.

The works we have just mentioned and further studies try to assess to what extent the decrease of mortality over the past 50 years have been accompanied by convergence in the mortality experience of the world's population. This is an important question, since mortality co-movements for multiple populations can have relevant implications for longevity risk management. A number of Authors, in the actuarial field, have thus discussed the advantages of common mortality models for detecting multiple population trends: see, for instance, Li and Lee (2005), Hatzopoulos and Haberman (2013), Chen et al. (2015) and Danesi et al. (2015). In such works, the performance of multi-population stochastic mortality models was investigated.

2.2 Aims and Key Idea

In this Chapter, an approach for exploring mortality dynamics for a pool of Countries is presented, under the mathematical framework of the affine processes (cf. Dacorogna and Apicella (2016) and Apicella and Dacorogna (2016)).

- In recognition to the fact that any sound procedure for projecting mortality must begin with a careful analysis of past trends (cf. Pitacco et al. (2009)), in this Chapter we focus on fitting past longevity trends. In particular, the target we aim at achieving is the extrapolation of the past mortality experience of a given generation, the one of males

born in 1940 over the time horizon [1980,2011], chosen by way of an example.

- Since we know from the literature that mortality patterns and trajectories in closely related populations are likely to be similar in some respects (cf. Li and Lee (2005)), we perform such a study for ten different developed Countries, in order to see if *cross-Country longevity common trends* can be detected. Furthermore, looking 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 key idea is:

1. to individually model the Country-specific mortality experience of the generation 1940 by using affine models for the mortality intensity of that cohort (see, e.g., Dahl (2004)). Within our numerical application, the stochastic mortality intensity of the specific generation under study, in each Country, is described by the two following specifications of the Feller process (whose fitting performance is statistically assessed and compared):
   - a two-parameter Feller process (as in Luciano and Vigna (2008));
   - a three-parameters Feller process (as in Fung et al. (2014)).
2. to explore, *a posteriori*, the prospective similarities between the Country-specific optimal parameters coming from the ten separate calibrations, for each specification of the Feller process.

Borrowing from the literature two specifications of the Feller process that have been already applied to describe the intensity associated with the random residual lifetime of the representative insured belonging to a given risk-class or cohort, we propose a further purpose of application, namely the comparative analysis on the fitted mortality trends of individuals belonging to different, but potentially related, populations.

2.3 Layout of the Chapter

Chapter 2 is organized as follows: in Section 2.4, we describe the mathematical framework. In Section 2.5, we describe the data used in the numerical application and discuss the empirical methodology, namely the techniques applied for calibrating the two Feller models. In Section 2.6 we present the results and illustrate the statistical assessment of the fitting performance of the two competing models. In Section 2.7, conclusions are given.

2.4 Mathematical Framework

As stressed in 1.1.4, according to Milevsky and Promislow (2001), Dahl (2004) and Biffis (2005), there are important similarities between the force of mortality and interest rates: for instance, they have term structures and are stochastic in nature. We exploit such features for building our mortality models: in modelling the stochastic force of mortality, we apply the same mathematical tools used in the credit risk literature to model the time to default. We refer the reader to Brémaud (1981), Jeanblanc and Rutkowski (2000) and Duffie (2001) for details about the mathematical framework.

Following Biffis (2005) and Luciano and Vigna (2008), 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 (2004)). 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 (2.1):

$$S_x(t) = P(T_x > t|\mathcal{G}_0) = E[\mathrm{e}^{-\int_0^t \lambda_x(u)du}|\mathcal{G}_0] \tag{2.1}$$

Under technical conditions (see Duffie and Singleton (2003)), if $\lambda_x(t)$ is modeled as an affine process, the probability of survival in (2.1) has an affine exponential form and can be computed as in (2.2):

$$S_x(t) = P(T_x > t|\mathcal{G}_0) = E[\mathrm{e}^{-\int_0^t \lambda_x(u)du}|\mathcal{G}_0] = \mathrm{e}^{\alpha(t)+\beta(t)\lambda_x(0)} \tag{2.2}$$

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 Luciano and Vigna (2005) and Cairns et al. (2006a). We briefly recall such remarks, that have been more extensively illustrated in Section 1.1.3. Luciano and Vigna (2005) found that time-homogeneous mean reverting affine processes fail to capture the rectangularization phenomenon and detected an exponential behaviour of the force of mortality observed and/or extrapolated from the mortality tables; they thus focused on non mean reverting processes, with deterministic part increasing exponentially; these models represent natural generalizations of the Gompertz law. According to Cairns et al. (2006a) any plausible mortality model should imply long-run deviations in mortality improvements from those anticipated that are not strongly mean reverting to a pre-determined target level, in order to ensure that, if mortality has improved faster than anticipated in the past, the potential for further mortality improvements will not be reduced in the future.

Within the numerical application illustrated in this Chapter, we choose for the stochastic force of mortality two different specifications of a Feller process: a two-parameters Feller process (as in Luciano and Vigna (2008)) and a three-parameters Feller process (as in Fung et al. (2014)); we calibrate them to the same sample of data, for ten different developed Countries, and we compare their goodness of fit. 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.

2.4.1 Two-parameters Feller process

Following Luciano and Vigna (2008), we assume that the stochastic force of mortality λ of a specific generation is described by a Feller process, whose dynamics is given by the stochastic differential equation (SDE) in (2.3):

$$d\lambda(t) = a\lambda(t)dt + \sigma\sqrt{\lambda(t)}dW(t), \tag{2.3}$$

where $a > 0$, $\sigma \geq 0$ and W is a standard Brownian motion.
The solution $\lambda(t)$ of the SDE is given by (2.4):

$$\lambda(t) = \lambda(0)\mathrm{e}^{at} + \sigma \int_0^t \mathrm{e}^{a(t-u)}\sqrt{\lambda(u)}dW(u) \tag{2.4}$$

In the affine framework, the solution of the system of ODEs for α and β in (2.2) is given by (2.5):

$$\begin{cases} \alpha(t) = 0 \\ \beta(t) = \frac{1-\mathrm{e}^{bt}}{c+d\mathrm{e}^{bt}} \end{cases} \tag{2.5}$$

with:

$$\begin{cases} b = -\sqrt{a^2 + 2\sigma^2} \\ c = \frac{b+a}{2} \\ d = \frac{b-a}{2} \end{cases} \tag{2.6}$$

The coefficients b, c, d are negative, so the survival probability is always decreasing in t if and only if condition (2.7) is fulfilled:

$$\mathrm{e}^{bt}(\sigma^2 + 2d^2) > \sigma^2 - 2dc \tag{2.7}$$

If the starting point is nonnegative, this model does not violate the nonnegative 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 (2.5) 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 equality in (2.8) holds:

$$\mu_x = \lambda_0(0)\mathrm{e}^{ax} = \lambda_0(x) \tag{2.8}$$

where μ_x represents the deterministic force of mortality as in equation (1.1).

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

2.4.2 Three-parameters Feller process

In Fung et al. (2014), the Feller process in (2.9) describes the stochastic intensity of mortality of a specific generation:

$$d\lambda(t) = (\bar{a} + \bar{b}\lambda(t))dt + \bar{\sigma}\sqrt{\lambda(t)}dW(t), \tag{2.9}$$

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, in (2.1), that an individual aged x at time 0 survives age $x + t$ is given by (2.10):

$$S_x(t) = A(t)\mathrm{e}^{-B(t)\lambda_x(0)} \tag{2.10}$$

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

$$\begin{cases} A(t) = \left(\frac{2\gamma \mathrm{e}^{\frac{1}{2}(\gamma-\bar{b})t}}{(\gamma-\bar{b})(\mathrm{e}^{\gamma t}-1)+2\gamma}\right)^{\frac{2\bar{a}}{\bar{\sigma}^2}} \\ B(t) = \frac{2(\mathrm{e}^{\gamma t}-1)}{(\gamma-\bar{b})(\mathrm{e}^{\gamma t}-1)+2\gamma} \end{cases} \tag{2.11}$$

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:

$$\lambda_0(x) = \lambda_0(0)\mathrm{e}^{\bar{b}x} \tag{2.12}$$

Looking at the stochastic differential equations in (2.3) and (2.9), we can see that the two specifications of the Feller process for λ, discussed in Sections 2.4.1 and 2.4.2, differ by the presence, in the latter one, of one additional parameter $\bar{a}$, that, according to Fung et al. (2014), 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 condition (2.13) holds:

$$\bar{a} \geq \frac{\bar{\sigma}^2}{2}, \tag{2.13}$$

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

2.5 Empirical Methodology

2.5.1 Collection of 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 HMD (Human Mortality Database (2016). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)) (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 $m_{x,t}$ (where x denotes the "age last birthday" and t 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

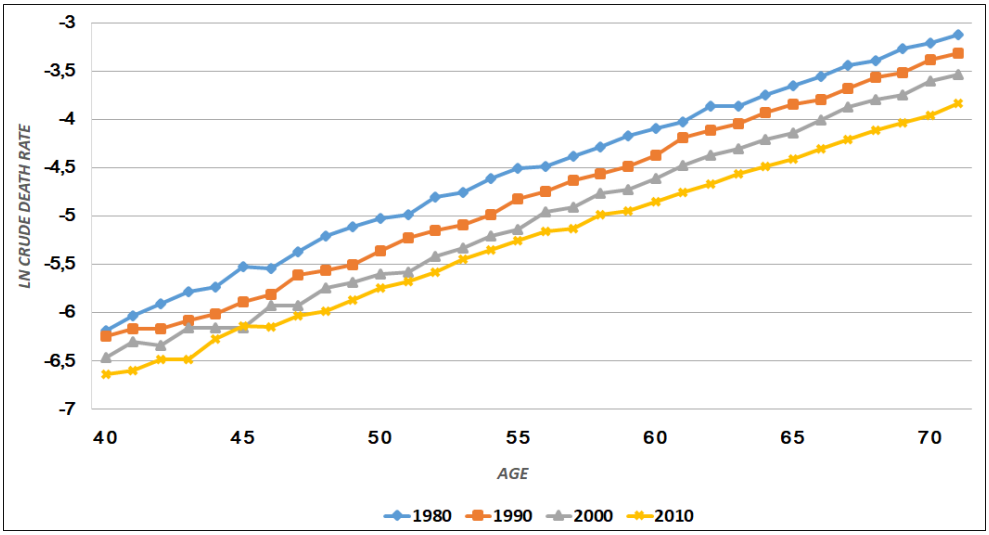


Figure 2.1: Death rates (on the log scale) relating to Canadian males aged from 40 to 71 in four selected years: 1980 (cyan line), 1990 (orange line), 2000 (grey line); 2010 (yellow line)

throughout life, but can be considered a good approximation. We refer to equation (1.10) for the mathematical definition of crude death rate.

In Figure 2.1 we display the logarithm of the observed death rates $m_{x,t}$ relating to Canadian males aged from 40 to 71 in four selected years: 1980, 1990, 2000 and 2010. We can see that a reduction in mortality has occurred over time. By way of an example, we compute the percent change in death rates over the period 1980-2010, as follows:

$$\Delta \ln m_x = \frac{\ln m_{x,2010} - \ln m_{x,1980}}{\ln m_{x,1980}}$$

with $x = 40, ..., 71$.

Since $\ln m_{x,t}$ are negative, a positive percent change indicates a decreasing trend in mortality. The average percent change in death rates is 11.20% for ages from 40 to 49 and 16% for ages from 50 to 59. The greatest decline in mortality occurred for people aged 60-71, with the percent change being 21%.

2.5.2 Preparation of data

Under the assumption (1.3) and exploiting equation (1.5), we compute the annual survival probabilities $p_{x,t}$, with $x = 40, ..., 71$ and $t = 1980, ..., 2011$ (x and t simultaneously varying), as follows:

$$p_{x,t} = \mathrm{e}^{-m_{x,t}} \tag{2.14}$$

The probability $q_{x,t}$ that an x-aged individual in calendar year t dies before reaching age $x + 1$ is then given by:

$$q_{x,t} = 1 - p_{x,t} = 1 - \mathrm{e}^{-m_{x,t}} \tag{2.15}$$

We can get the annual probability of death q_x also employing the general principles at the base of the HMD life tables (see Wilmoth et al. (2007))(we drop reference to the calendar year y to make the notation simpler), as in (2.16):

$$q_x = m_x/(1 + ((1 - a_x)m_x)), \tag{2.16}$$

In (2.16) a_x denotes the average number of years lived within the age interval $[x, x + 1)$ for people dying at that age and is assumed to be equal to $\frac{1}{2}$ for all single-year ages except for age 0. For the ages of our interest we are able to get the same values for the probabilities of death (equal in the first eight digits) by formulas (2.16) and (2.15).

The probability that an individual aged 40 reaches age $40 + t$, with $t = 1, ..., 32$, namely the survival function $S_{40}(t)$ (see Subsection 1.4), can now be computed by the product of the annual survival probabilities, as in (2.17):

$$S_{40}(t) = \prod_{i=x}^{x+t-1} p_i \tag{2.17}$$

with p_i representing the annual survival probabilities in a dynamic context, where they are assumed to be function of both the age and the calendar year: $p_{40,1980}$, $p_{41,1981}$, $\cdots$, $p_{71,2011}$.

$S_{40}(0)$, namely the probability that an individual aged 40 survives the age he has already attained, is equal to 1.

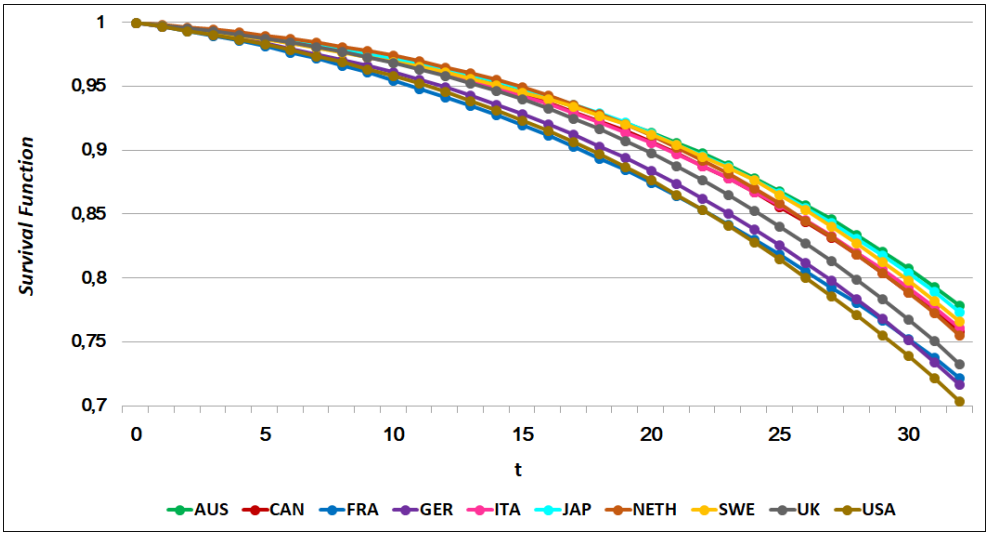


Figure 2.2: The survival function $S_{40}(t)$, with $t = 0, ..., 32$ relating to the males born in 1940 in the ten Countries under study.

In Figure 2.2, we show the survival function $S_{40}(t)$, with $t = 0, ..., 32$, relating to the males born in 1940 in the ten Countries under study. We can see that an individual, that was born in the Netherlands in 1940 and 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 years was around 74%, in Europe, and 73% in North America (Canada, USA).

2.5.3 Calibration of the models

For each Country, we calibrate the two-parameters and the three-parameters Feller process (see Sections 2.4.1 and 2.4.2) to the observed survival function $S_{40}(t)$, computed as in Section 2.5.2, by minimization of the Root Mean Square Error (η). η is thus 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 (2.2) or in (2.10). We solve the optimization problem by using the MATLAB function *fmincon*, that finds the minimum of constrained

nonlinear multivariate functions.

By θ we denote 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 parameters values, the optimization algorithm finds the optimal parameters of the selected model, namely those parameters that, if used as inputs in (2.5) and (2.11), allow us to get a predicted survival function diverging as little as possible from the observed one.

We subject the minimization to the following conditions on the parameters (mentioned in Sections 2.4.1 and 2.4.2):

- 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 (2.7) 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 (2.13) is satisfied (as a nonlinear constraint).

We set the initial value of λ, $\lambda_{40}(0)$, as $-\ln(p_{40})$, according to (2.14).

Table 2.1: 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

In Table 2.1, we report the value of the observed mortality intensity at age 40, $\lambda_{40}(0)$, for the generation under study in all the Countries of our interest. 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 2.2 since all the curves representing $S_{40}(t)$ are almost overlapping when $t = 1$.

2.6 Empirical Results

We report the optimal values of the parameters, η, the AIC and the BIC of the two-parameters Feller process and the three-parameters Feller process relating to each Country, in Table 2.2 and 2.3, respectively.

Table 2.2: Fitting results for the two-parameters Feller process (unit 10^{-2} for a, σ and η_2)

Country	a	σ	η_2	RR_2	AIC_2	BIC_2
AUS	6.24	0.03	0.316	1.14	-253.48	-361.45
CAN	7.81	1.03	0.027	-0.82	-410.63	-518.61
FRA	7.35	2.07	0.065	-0.56	-354.82	-462.80
GER	6.71	0.32	0.148	0.01	-301.93	-409.90
ITA	7.04	0.45	0.094	-0.37	-331.37	-439.34
JAP	7.09	0.31	0.155	0.05	-299.02	-406.99
NETH	9.77	1.66	0.032	-0.78	-400.45	-508.42
SWE	6.24	2.2e-04	0.489	2.31	-225.56	-333.53
UK	8.68	1.46	0.056	-0.62	-364.54	-472.51
USA	6.98	0.84	0.094	-0.37	-331.39	-439.36

NOTE: The red colour is used for highlighting the highest values of: the RMSE η_2 and the measure RR_2.

2.6.1 Estimated parameters of the Feller models

The parameters a and $\bar{b}$ represent the rate of the increase of the force of mortality of the Gompertz model underlying the two-parameters Feller process and the three-parameters one respectively (see equations (2.8) and (2.12)), 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.

Table 2.3: Fitting results for the three-parameters Feller process (unit 10^{-2} for $\bar{a}$, $\bar{b}$, $\bar{\sigma}$ and η_3)

Country	$\bar{a}$	$\bar{b}$	$\bar{\sigma}$	η_3	RR_3	AIC_3	BIC_3
AUS	6.27e-07	6.24	0.01	0.316	1.08	-251.49	-361.47
CAN	1.38e-03	7.30	0.52	0.035	-0.77	-392.83	-502.80
FRA	9.29e-03	5.19	1.36	0.041	-0.73	-381.66	-491.63
GER	4.03e-06	6.68	0.01	0.143	-0.06	-302.30	-412.27
ITA	2.19e-05	6.99	0.03	0.086	-0.43	-334.80	-444.77
JAP	3.06e-06	7.07	0.01	0.150	-0.01	-299.18	-409.16
NETH	4.78e-03	7.90	0.97	0.074	-0.52	-344.82	-454.80
SWE	3.08e-06	6.23	0.01	0.489	2.22	-223.55	-333.52
UK	3.64e-03	7.49	0.84	0.086	-0.43	-334.93	-444.90
USA	8.14e-04	6.73	0.40	0.098	-0.35	-326.47	-436.44

NOTE: The red colour is used for highlighting the highest values of: the RMSE η_3 and the measure RR_3.

We find, for the generation 1940 born in the Netherlands, the smallest value of $\lambda_{40}(0)$ (cf. Table 2.1) and the highest value of a and $\bar{b}$ (cf. Tables 2.2 and 2.3). 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. We can thus infer that, for the generation under study, the ten developed Countries showed similar patterns of mortality dynamics.

2.6.2 Assessment of the fitting performance of the Feller models

Root Mean Square Error

Overall, as we can see in Tables 2.2 and 2.3, our results show very low η's, indicating a good fit to the data extracted from the mortality tables. In particular, we obtain the most accurate fits of both the two and the three-parameters Feller process for Canada, France, Italy, the Netherlands, UK and USA and the least satisfactory fits for Australia and Sweden (in Tables 2.2 and 2.3 we highlight the value of the RMSE relating to these two Countries with the red colour).
In Figure 2.3, we display the observed and fitted probabilities (through the two-parameters Feller process) to survive further 32 years characterizing males born in Canada and Sweden; for these Countries we get, respectively,

the best and the worst fit to the real data. We can see that, when Canada is concerned, the curve of the model-implied survival function (the orange curve) matches the curve of the observed survival function (the blue curve) at each point. A similar behaviour can be found also for the other Countries for which the two specifications of the Feller process show themselves to perform the best. With this respect, at the end of this Chapter, we display the following Figures: Figure 2.4, in which we show the observed and fitted survival function (through the two-parameters Feller process) for all the Countries under study other than Canada and Sweden; Figures 2.5 and 2.6, in which we display the observed and fitted survival functions coming from the three parameters Feller process and concerning Canada and Sweden, in the former, and the remaining Countries under study, in the latter. In all the Figures the observed survival function is represented by the blue curve, while the model-implied survival function is displayed as the orange curve. In Figures 2.5 and 2.6, we can notice that the fitting performance of the three-parameters Feller process is not much different from the one of the two-parameters Feller process; for instance, the blue and the orange curves overlap when Canada is concerned, while the largest divergence between the two curves can be seen, at some points, for Australia and Sweden.

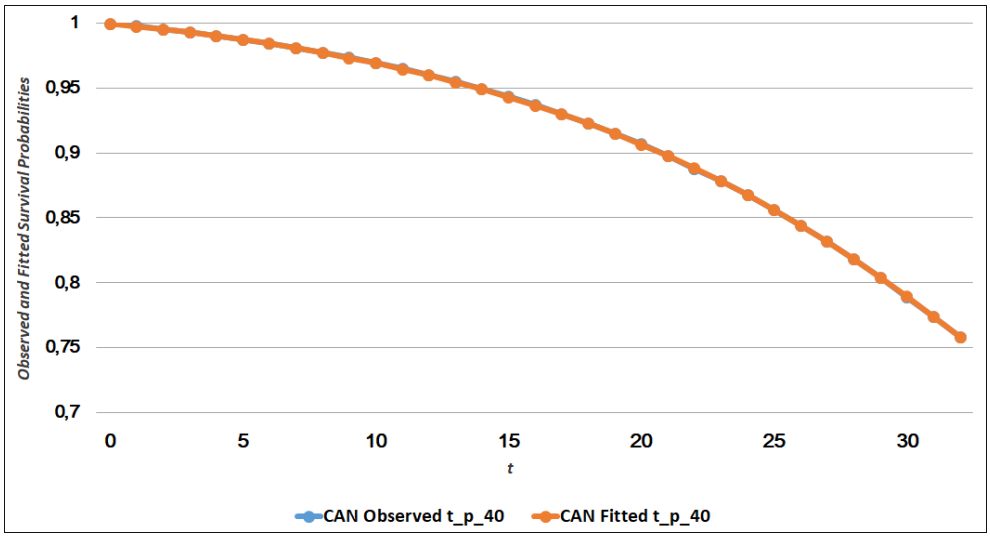
Since we calibrate the two and the three-parameters Feller process to an homogeneous sample of data for all the Countries (we fit the survival function $S_{40}(t)$ of the same generation and over the same time horizon, namely for $t = 1, ..., 32$), we are interested in assessing the goodness of fit of 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 denominate "RR", computed as follows:

$$RR = \frac{(\eta - \bar{\eta})}{\bar{\eta}} \tag{2.18}$$

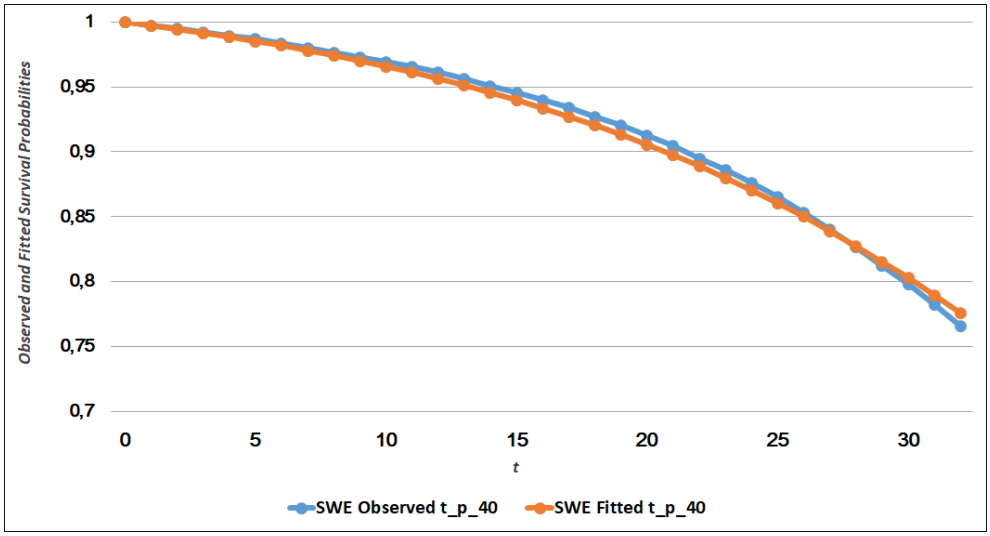
where $\bar{\eta}$ represent the mean of the values of η we find for each Country, as in (2.19):

$$\bar{\eta} = \frac{\sum_{j=1}^{10} \eta_j}{10} \tag{2.19}$$

where $j = 1, 2..., 10$ denotes the Countries arranged in alphabetical order (starting from 1= Australia until 10=USA). A negative value of "RR" indicates that the Root Mean Square Error of the model we are considering is less than $\bar{\eta}$, while a positive value indicates a value of η higher than the average.



(a) Canada



(b) Sweden

Figure 2.3: Canadian and Swedish generation of males born in 1940. Observed and fitted probabilities (through the two-parameters Feller process) to survive further 32 years. Blue line: observed probabilities; orange line: fitted probabilities

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. As we can see in Tables 2.2 and 2.3, 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 the η and $\bar{\eta}$ 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 η and $\bar{\eta}$ 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 possible from the average goodness of fit 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 the Countries, we can consider the average RMSE $\bar{\eta}$ as the benchmark representing a good fit.

The Akaike and the Bayesian Information Criteria

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, the AIC estimates the quality of each model, relative to each of the other models. Thus, the AIC is useful for model selection. Moreover, the 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 and the complexity of the model. The model with the lowest AIC is preferred.

We compute the AIC as in (2.20):

$$AIC = 2z + n \ln(\tau) \tag{2.20}$$

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. z, namely the number of parameters, is equal to 2 for

Table 2.4: Difference between the AIC and the 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%

NOTE: The red colour is used for highlighting negative values of the difference between the AIC (or the BIC) of the two Feller models

the two-parameters specification of the Feller process and 3 for the three-parameters specification; n is equal to number of observations, that is 32 in both cases.

As the 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 the BIC than in the AIC. We compute the BIC in terms of the residual sum of squares (τ):

$$BIC = n \ln(\tau/n) + z \ln(n) \tag{2.21}$$

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

In Tables 2.2 and 2.3 we report the values of the AIC and BIC criteria for the two and the three-parameters Feller models; in Table 2.4 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 the BIC criterion.

The two Feller models provide close values of the AIC and the BIC for all the Countries. Since the AIC and the BIC are negative, a positive difference in Table 2.4 denotes that the AIC (or the BIC) of the three-parameters Feller

process is larger than the AIC (or the BIC) of the two-parameters Feller process. In 60% of cases, the AIC of the three-parameters Feller model is higher than the AIC of the two-parameters Feller model, and in 50% of cases the BIC of the three-parameters model is larger than the BIC of the two-parameters model. We can see that the biggest positive gap between the AIC (BIC) of the two different mortality models concerns the 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.

Interpretation of the Results

On the basis of the evidence provided by the calibration of the two Feller models, differently parameterized, the two-parameters Feller models should be preferred. 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 the AIC and the BIC show. Therefore, among our competing hypotheses we are induced to select the one with the fewer parameters.

2.7 Conclusion

In this Chapter we discuss the first contribution presented in this Thesis, consisting in developing, within the affine diffusion framework, a methodology for performing a comparative analysis of the fitted past longevity trends of different populations.

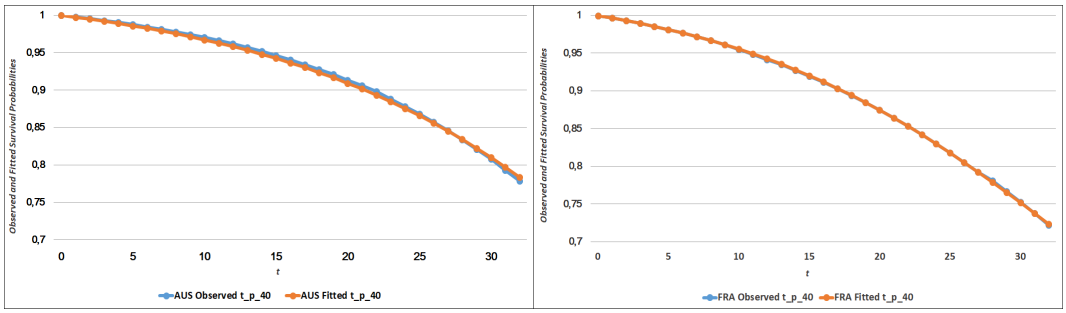
We calibrate two different specifications of the Feller process (with two and three parameters) to the observed survival functions of the generation of males born in 1940 (ages 40-71) in ten developed Countries. Such method allows us to model the individual mortality experience of each Country in a coherent manner with respect to the mortality experiences of the other Countries under study. In order to check the existence of common historical

trends, we compare the optimal parameters of the calibrated models among Countries. The optimal values of the parameters representing the rate of the increase of the force of mortality of the Gompertz model underlying the two specifications of the Feller processes are very close to each other for all the Countries, for both the Feller processes under study.

We assess the fitting performance of the two Feller processes for each Country, by looking at the Root Mean Square Error, whose low values indicate a good fit to the data extracted from the mortality tables. We obtain the most satisfactory fit for Canada and the least satisfactory fit for Australia and Sweden. Furthermore, we determine the relative goodness of fit of each specification of the Feller process for each Country compared to its average performance for all the other Countries under study. The two Feller models provide close values of the AIC and the BIC for all the Countries, meaning that, when using the three-parameters Feller process, the accuracy gained by adding a parameter to the fit is not enough to compensate for the added complexity of the model; therefore, the two-parameters Feller process should be preferred.

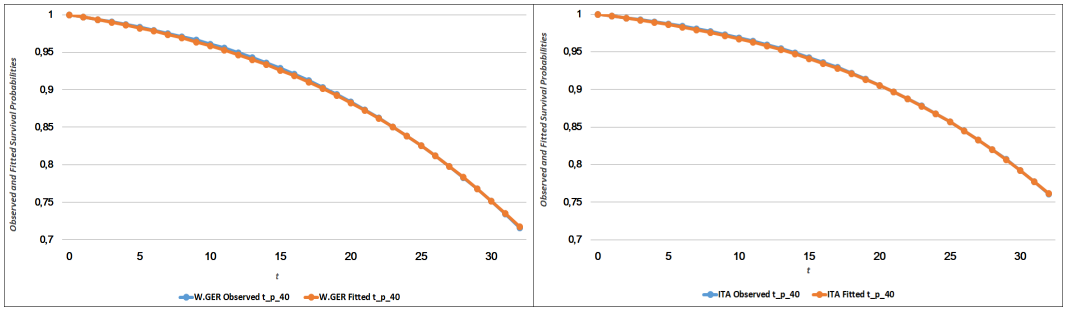
2.7.1 Further research

The study presented in this Chapter leaves scope for further research in several directions. An extension could consist in setting out a backtesting framework in order to evaluate the forecasting performance of the two specifications of the Feller process out-of-sample. In addition, a multi-population model could be used for analyzing the mortality dynamics among the ten Countries under study, in a multivariate context, and the resulting outcomes could be compared to the ones discussed in this Chapter.



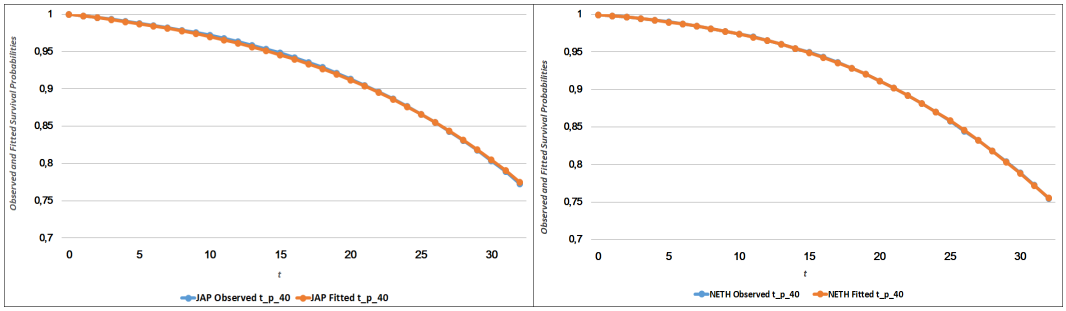
(a) Australia

(b) France



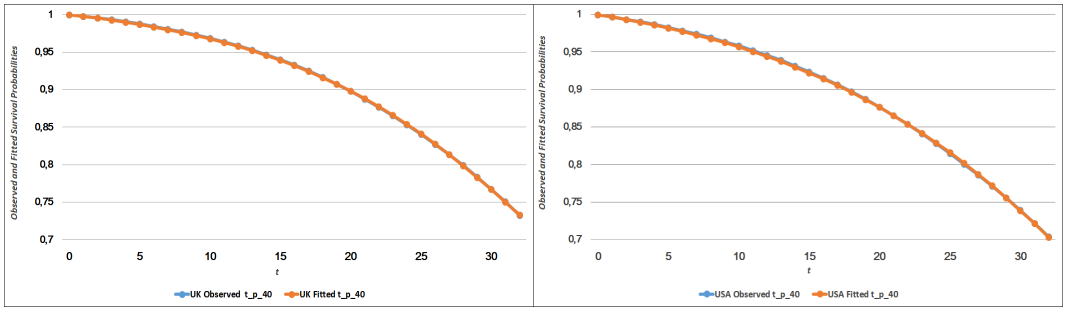
(c) West Germany

(d) Italy



(e) Japan

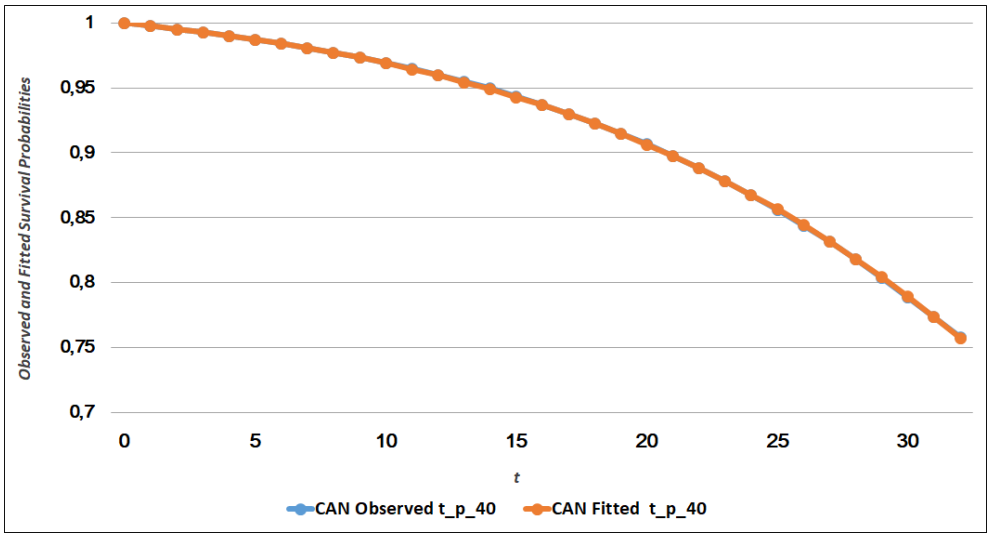
(f) Netherlands



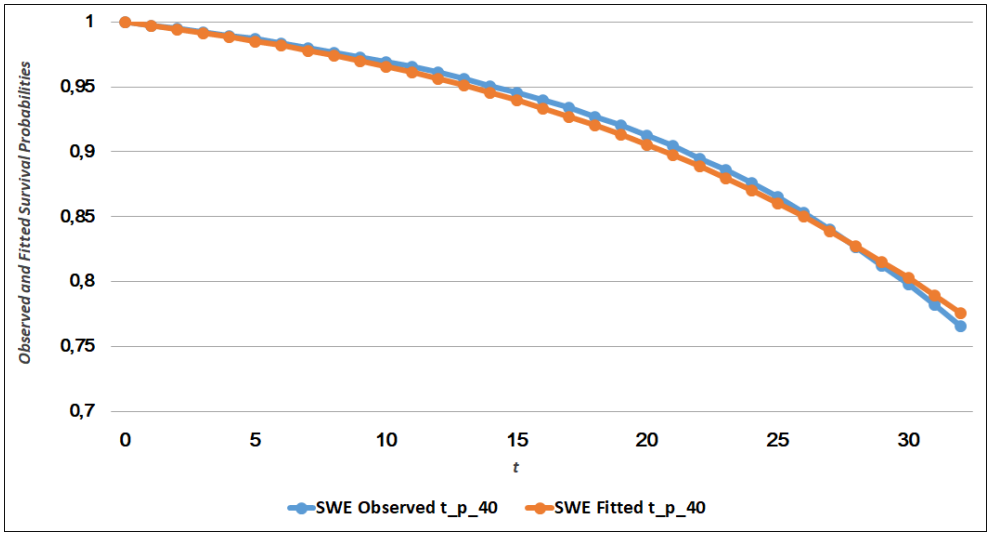
(g) UK

(h) USA

Figure 2.4: Observed and fitted survival probabilities (two-parameters Feller process) for all the Countries under study except for Canada and Sweden. Blue line: observed probabilities; orange line: fitted probabilities

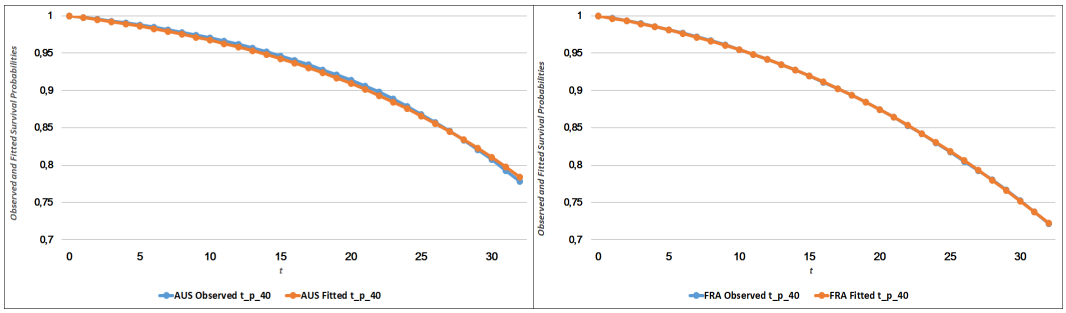


(a) Canada



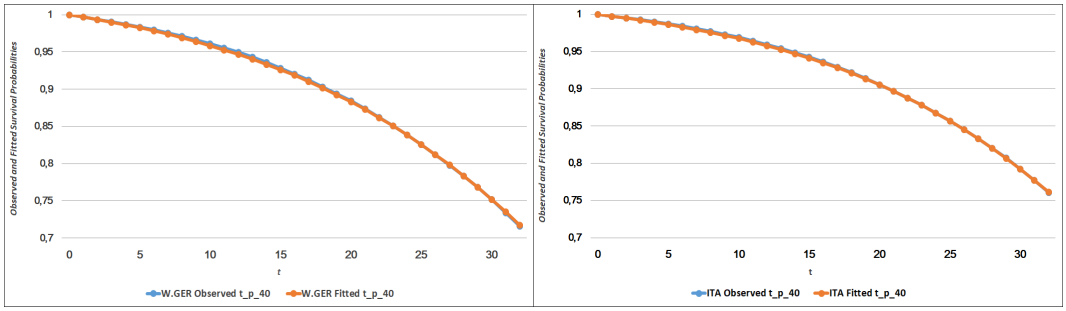
(b) Sweden

Figure 2.5: Canadian and Swedish generation of males born in 1940. Observed and fitted probabilities (through the three-parameters Feller process) to survive further 32 years. Blue line: observed probabilities; orange line: fitted probabilities



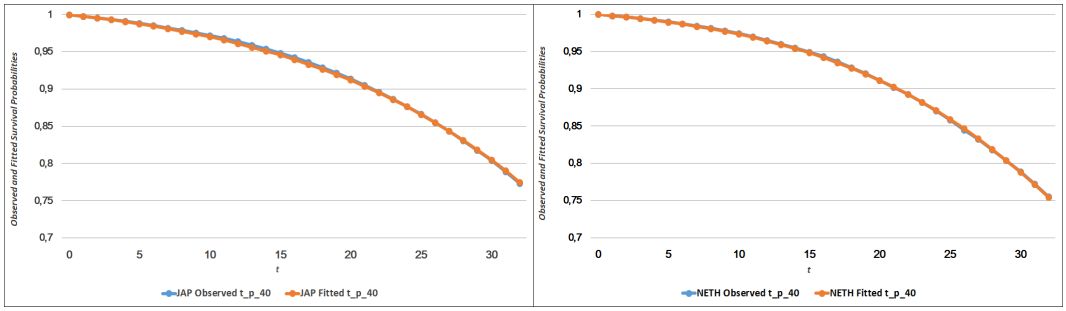
(a) Australia

(b) France



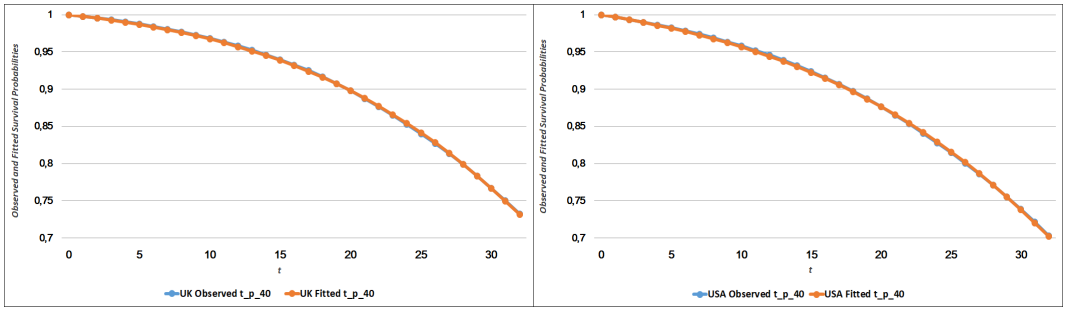
(c) West Germany

(d) Italy



(e) Japan

(f) Netherlands



(g) UK

(h) USA

Figure 2.6: Observed and fitted survival probabilities (three-parameters Feller process) for all the Countries under study except for Canada and Sweden. Blue line: observed probabilities; orange line: fitted probabilities

Chapter 3

Using Affine Models for Improving the Forecast of Longevity

3.1 Context and Motivation of this Chapter

The accuracy of mathematical models in predicting mortality is an important challenge because of the strong impact of life expectancy forecasts for policy making in many significant social sectors. One of the aspects which makes central the modelling of such a stochastic phenomenon is the systematic incidence of longevity on life insurance and pension funds, particularly in the industrialized Countries. The stability and consistency of social welfare systems are put in danger worldwide due to the combined phenomenon of the progressive increase in life expectancy, along with the reduction of birth-rates in industrialized Countries (see for instance Dacorogna and Kratz (2015)).

The consequences in several fields, particularly in financial micro and macro - economic environments, have been matter of discussion for over two decades, especially in the current recessionary phase of the World Economy, drastically reducing resources to fund the needs of the advanced age that is slowly becoming a wide segment of the population. Among several issues, the risk connected with the future of today's new generations, already facing an enduring lack of jobs, constitutes a good example for understanding how rel-

evant a good survival forecasting is for financial planning. De facto, those generations potentially are confronted with the risk of a lack of resources for their primary needs at old ages. This undoubtedly affects social security and public spending. Currently, in the middle of economic crisis and critical issues of public pension systems, we are witnessing an increased demand by workers to protect their own economic status at the end of the working age.

The competitive challenges arising from the survival evolution in industrialized Countries have been debated since 2002, during the Second World Assembly on Ageing (see United Nations (2002)). As pointed out throughout the guidelines of the Green Paper (European Commission (2012)), the effort of the relevant authorities in the European Union (EU), confronted with the forecast of a considerable increase in the number of oldest old (individuals who are older than 80) by 2060, aims at using the evolutionary trend of survival as a new economic driving force. Within this context, new developments in financial and insurance fields are considered. For example, in the Consultative Document drawn up on August 2013 (cf. Basel Committee on Banking Supervision (2013)), it is stated that the ageing phenomenon poses serious social policy and regulatory/supervisory challenges. One way to deal with this question is to foster longevity risk transfer markets to better spread the risk. In this sense, there is no doubt that forecasting models involving demographic changes are important quantitative tools, given any scope of application.

Due to application complexity, constructing, implementing and verifying descriptive systems of the evolution of mortality/survival rates requires a rigorous and, at the same time, easy and effective, methodological approach. Several studies show that classical survival models are currently unfit to accurately represent the life evolution, especially when it comes to advanced ages. A detailed discussion of this subject can be found in Booth and Tickle (2008). Many contributions in the recent demographic and actuarial literature deal with the comparison among different survival models. An interesting contribution in this direction is provided by Dowd et al. (2013). The study is extended also to forecasting models of aggregate type, namely models built on categories of population selected on the basis of death causes and cause-specific mortality (cf. Alho and Spencer (1990), Park et al. (2006)).

3.2 Aims and Key Idea

In this Chapter we design the affine model we propose for describing the evolution of the mortality phenomenon. This model is multiplicative and combines two components: the deterministic output (baseline) of any existing mortality projection model and an affine stochastic factor affecting the baseline over the forecasting time horizon. Such an affine stochastic factor is described by a Cox-Ingersoll-Ross process suitably fitted to a measure of the error of the existing mortality model producing the baseline. A fundamental part of the work is the study of the properties of the so structured affine model and of its capability to reduce the spread between the real mortality phenomenon and the corresponding forecasted one. To this aim, we structure various out-of-sample validation methods. Besides the usual static method, we develop a dynamic one that allows us to catch the change in behavior of the underlying data. The construction method of the affine model we present in this Chapter is quite general, since it can exploit various mortality models as long as their errors behave in a way similar to the CIR process.

3.3 Layout of the Chapter

The remainder of Chapter 3 is organized as follows: in Section 3.4 and 3.5, respectively, we explain the mathematical framework and provide the key features of the empirical methodology we propose. In Section 3.6 a summary of the main statistical contributions and of the related literature is given.

3.4 Mathematical Framework

In this Section we set the mathematical framework and the actuarial assumptions useful for the development of Chapters 3, 4 and 5.

3.4.1 The CIR process and its meaning in the context of Mortality Modelling

The mortality phenomenon is strongly affected by the impact of dynamic variables, which we could roughly classify as economic and biological factors. The mortality stochastic models, although taking into account their influence,

evolve in time implicitly assuming that this influence is "steady", in the sense that the interactions of the biological and economic risk drivers with the mortality phenomenon do not suffer changes in their structure when time passes. The historical data analysis, on the contrary, shows that such interactions dynamically evolve in time. This aspect is well represented by the deviations of the real data from those ones provided by the mortality model. The key idea of this Chapter is to resort to those deviations, considering them useful extra information, for dynamically projecting the model (cf. Apicella et al. (2017)). We pinpoint the ratio:

$$Y_{x,t} = \frac{B_{x,t}}{\mu_{x,t}} \tag{3.1}$$

where $B_{x,t}$ denotes the observed central death rates and $\mu_{x,t}$ the baseline provided by the chosen mortality model for describing the same rates (cf. Di Lorenzo et al. (2006)). As to the meaning of the baseline, we remind that under condition (1.3) (cf. Section 1.4):

- the forces of mortality and the central death rates coincide, that is $\mu_{x,t} = m_{x,t}$;
- the maximum likelihood estimate of the force of mortality $\mu_{x,t}$ is given by the crude death rate for age x in calendar year t (see Pitacco et al. (2009));
- the mortality rate can be computed as follows: $q_{x,t} = [1-\exp(-\mu_{x,t})] = [1 - \exp(-m_{x,t})]$.

The idea is to model the dynamics over time of $Y_{x,t}$ as a Cox-Ingersoll-Ross (CIR) process as in the following stochastic differential equation:

$$dY_{x,t} = \alpha(\beta - Y_{x,t})dt + \sigma\sqrt{Y_{x,t}}dW_t \tag{3.2}$$

which, for $\alpha, \beta > 0$, corresponds to a continuous time first-order autoregressive process where the rv $Y_{x,t}$ is elastically pulled toward a long-term mean, β, at a speed α.

If the starting point of $Y_{x,t}$ is non-negative and the coefficients satisfy condition (3.3):

$$2\alpha\beta \geq \sigma^2 \tag{3.3}$$

$Y_{x,t}$ can never become negative.

Given the probability density of Y_x at time t, conditional on its value at the current time s, for some $s < t$, as in (B.1), straightforward calculations give the expected value as follows (cf. Cox et al. (1985)):

$$E(Y_{x,t}|Y_{x,s}) = Y_{x,s}\exp(-\alpha(t-s)) + \beta(1-\exp(-\alpha(t-s))) \tag{3.4}$$

The reason behind the choice of the CIR process for modelling $Y_{x,t}$ lies in the empirically relevant properties of such a process and in its computational tractability. Indeed, within this structure, provided that condition (3.3) is fulfilled:

- negative values are precluded and if the process reaches zero, it can subsequently become only positive (these are desirable features for $Y_{x,t}$, since mortality rates are, by definition, non-negative);
- the absolute variance of the process increases when the process itself increases;
- there exists a steady state distribution for $Y_{x,t}$.

Two different meanings can be given to the process $Y_{x,t}$, in the specific context of this study:

- as a correction factor:
  our approach leads to the multiplicative model in equation (3.5):

$$\tilde{B}_{x,t} = \tilde{Y}_{x,t}\bar{\mu}_{x,t} \tag{3.5}$$

  for constructing projected values, $\tilde{B}_{x,t}$, starting from the chosen mortality model: its outputs, $\bar{\mu}_{x,t}$ can be corrected simply multiplying them by the CIR process in (3.2) with properly estimated parameters. $\bar{\mu}_{x,t}$ is the deterministic baseline, since it is the output of the mortality projection model, while $\tilde{Y}_{x,t}$ is the affine stochastic driving factor (a Cox-Ingersoll-Ross process); it results that $\tilde{B}_{x,t}$ is stochastic, whose dynamics over time being described by a CIR process.
  The CIR process shows itself to be suitable for this purpose.
  Firstly, it provides a return to the mean dynamics of $Y_{x,t}$. The optimal long-term mean β and speed α, resulting from the calibration

of the CIR process, give us useful information about the dynamics of $Y_{x,t}$ over the observation time horizon; furhermore, they provide crucial directions for adjusting the baseline over the forecasting time horizon, through the correction factor $\tilde{Y}_{x,t}$.
In addition, as already stressed, there are closed form formulas for the probability density of the state variable at time t, conditional on its value at the current time s, for some $s < t$, and also for its expected value and variance. This allows us to obtain the "best estimate" of the correction factor $\tilde{Y}_{x,t}$, for any age x and future calendar year t, simply by computing the conditional expected value of the CIR process $Y_{x,t}$, as in (3.4)

- as a mortality model benchmark:
  Briefly speaking, when the calibration of the CIR process in (3.2) provides an optimal value of the long-term mean β close to 1, we can immediately infer, with the confidence coming from the empirical calibration, that the chosen mortality model is able to fit well the real data themselves.

It is worth noting that when we investigate the mortality profile for a given age and we thus read the matrix of the observed death rates $B_{x,t}$ by rows, the CIR process in (3.2) describes the dynamics over time (as t varies) of $Y_{\ddot{x},t}$, where the subscript $\ddot{x}$ denotes that the age is fixed. When we are interested in representing the evolution of mortality for a given cohort (diagonal approach), the CIR process in (3.2) describes the dynamics over time of $Y_{x,t}$, where the age x is not fixed, but it implicitly moves forward time by time.

3.5 Empirical Methodology

3.5.1 Backtesting procedure

Once designed the so structured affine model, a crucial part of the work is to study its properties, in order to test its ability to enhance the predictive accuracy of the existing mortality projection model providing the baseline, making its behavior closer to the real trend of the mortality phenomenon. We thus set up backtesting methods, for empirically testing the forecasting performance of the mortality model for itself and of our affine model.

In order to be able to compare the forecast of mortality against the realized one, we use the following backtesting procedure: we split the sample of available reliable mortality data into two time intervals and use the corresponding data sets as follows:

- the in-sample data (within the "look-back" window (cf. Dowd et al. (2013)) for fitting $\mu_{x,t}$ from the mortality model and calibrating the CIR process $Y_{x,t}$
- the out-of-sample data (within the "look-forward" window) for estimating how accurately the forecasts produced by the mortality model perform out-of-sample, with and without our correction. It is an "out-of-sample" test because the data used here have not been used to calibrate the model.

We use both static and fixed-length windows rolling one-year-ahead through time (dynamic approach).

Using this methodology allows us to test the models with data that were never seen before as it would be the case in the real use of such models to predict future development of mortality.

We choose the historical central death rates (as defined in (1.9)) as the target variables to fit and test in the predictions.

3.6 Statistical Contributions and Related Literature

In this Chapter, we have explained the key features of the multiplicative affine model we develop for describing the evolution of the mortality phenomenon. The key ideas are: (i) to track, over an observation time horizon, a measure of the fitting error of any existing mortality projection model, such as the ratio between the observed central death rates and the corresponding fitted values obtained as outputs of the chosen existing mortality model, and model this ratio, dynamically, by means of the Cox-Ingersoll-Ross (CIR) stochastic process, that provides a return to the mean dynamics for this quantity; (ii) to design our multiplicative affine model for forecasting mortality, that is based on the deterministic forecast (baseline) provided by the existing mortality

projection model and the CIR corrective factor with properly estimated parameters. (iii) Once designed the so structured affine model, a crucial part of the work is to study its properties, in order to test its capability to enhance the predictive accuracy of the existing mortality projection model providing the baseline, making its behavior closer to the real trend of the mortality phenomenon. For doing so, we design backtesting methods that combine both static and dynamic analyses. The overall applicability and validity of our numerical procedure can be tested on any mortality model, whose errors follow similar dynamics as the CIR process.

Similar perspectives and approaches as those underpinning the methodology presented in this Chapter can be found in the following literature, dealing with (i) dynamic adjustment models (e.g. Engle and Granger (1987), where the error correction model is used in a linear way), (ii) methods of adjustment and prediction of mortality models based on geostatistical techniques (e.g. Debón et al. (2010), where the mortality data are decomposed in a additive way into a deterministic large-scale variation (trend) plus a stochastic small scale variation (error)), (iii) the backtesting of stochastic mortality models (e.g. Dowd et al. (2013)).

Chapter 4

Predicting the Mortality Profile for Different Ages through the CBD and the mCBD Models

4.1 Aims of this Chapter

In Chapter 3 we have illustrated the affine model we have designed for describing the future evolution of mortality and the backtesting methods, combining both static and dynamic analyses, we have set up for testing its predictive accuracy.

In this Chapter, we are going to provide a numerical implementation of the affine model, with the aim of assessing, through the designed backtesting procedure, its forecasting performance with respect to the one of the existing mortality projection model producing the baseline. For doing so:

- we use the Italian females mortality data;
- we select the Cairns-Blake-Dowd (CBD or M5) model as the mortality model providing the baseline; such a model is one of the most popular models in the actuarial field and is suited to pension-related applications;
- exploiting the baseline produced by the CBD model and the CIR factor

suitably fitted to the fitting error of such a model, we build our affine model (mCBD model). We test the performance of the mCBD model in predicting the mortality profile of different ages against the CBD model

4.2 Layout of the Chapter

Chapter 4 is organized as follows: in Section 4.3, we discuss the empirical setup of the application, namely we present the CBD model and describe the statistical measures and the data we use. Sections 4.4 and 4.5 are devoted to the illustration of the techniques applied for calibrating and backtesting the CBD and the mCBD models. In Section 4.6, we construct "adjusted projections" of mortality, through the mCBD model. In Section 4.7 we discuss the graphical and statistical assessment of the predictive performance of the two competing models, the CBD model and the mCBD one. In Section 6 conclusions are given. In Appendix B a deep discussion of the theory underlying calibration and simulation of the CIR process is provided and the main steps of their implementation are explained.

4.3 Empirical Setup

In this Section we describe our empirical methodology. The first step of the method consists in the choice of the mortality model to correct by means of the CIR process. The model we choose for testing our methodology is the CBD model, namely the two-factor stochastic model introduced in Cairns et al. (2006b) for the development of the post-age-60 mortality curve through time and one of the most popular stochastic mortality models within the actuarial literature. Nevertheless, the procedure we propose is suited to any choice of stochastic mortality model. The target of our methodological approach is not to provide a new mortality projection model, but to correct an already existing one, that we consider worth to be chosen for representing the mortality phenomenon.

4.3.1 The CBD (or M5) model as the starting model for mortality

In Cairns et al. (2006b), the CBD model had the predictor structure for the mortality rates $q_{x,t}$ expressed in (4.1) (see also Cairns et al. (2009)):

$$\log \frac{q_{x,t}}{(1-q_{x,t})} = \beta_x^{(1)} k_t^{(1)} + \beta_x^{(2)} k_t^{(2)} \tag{4.1}$$

where the left term of the equation is the *logit* function of $q_{x,t}$ and the terms $\beta_x^{(i)}, i = 1, 2$, modulating the effect of the time indices $k_t^{(i)}, i = 1, 2$, across ages, are assumed to be parametric, as follows:

$$\beta_x^{(1)} = 1 \quad and \quad \beta_x^{(2)} = (x - \bar{x})$$

where $\bar{x}$ is the mean age in the sample range. Basing on equation (4.1), the predictor structure of the CBD model can be formulated as:

$$logit(q_{x,t}) = k_t^{(1)} + k_t^{(2)}(x - \bar{x}) \tag{4.2}$$

Such a model does not require parameter constraints to ensure unique parameter estimates, namely the model has no identification problem, as explained in Cairns et al. (2009).

In formula (4.2), the first factor (the intercept or level term, $k_t^{(1)}$) affects mortality-rate dynamics at all ages in the same way, whereas the second factor (the slope, $k_t^{(2)}$) affects mortality-rate dynamics at higher ages much more than at lower ages.

Remarks

- Building on the observation that log mortality rates are approximately linear at higher ages, the CBD model has been designed for modelling mortality at such ages; it is thus suitable for modelling longevity risk in pensions and annuities. For this reason, within our numerical application, the main focus is on age 65, for which all the graphical and statistical results are shown. Anyway, for illustrative purposes, we demonstrate that, when younger ages are concerned, the methodology we propose works as well.

- The CBD model is an "extrapolative" projection model, which is reliable if past trends continue (cf. Cairns et al. (2008a)), in the sense that the CBD model takes into account the biological and economic components affecting the mortality behaviour, but allows for stable interactions between them. On the contrary, real mortality data suggest that such interactions change with time (e.g. Preston (1975)). Our methodological approach, leading to the mCBD model, adds flexibility to the model itself and aims at adapting it, dynamically, to the changes in societal conditions over time, as they appear from the real data. Indeed, as we will see from the outcomes of our empirical study, the Cox-Ingersoll-Ross process, we use for enhancing the predictive accuracy of the CBD model, thanks to its drift, provides a dynamic adjustment, so that it enables to better capture the dynamics of the system as the forecasting time horizon goes forward. In this sense, the mCBD model is not meant to produce the overall best forecast, but to provide a more accurate forecast with respect to the prediction coming from a selected mortality model, such as the CBD model.

4.3.2 Forecasting errors and information criteria: a general overview

Let us suppose n pairs of estimates or predictions $(P_i, i = 1, ...n)$ provided by a given model and the corresponding observations $(O_i, i = 1, ..., n)$. The average model-estimation (or prediction) error can be written as (cf. Willmott and Matsuura (2005)):

$$\bar{e}_\gamma = \left[\sum_{i=1}^{n} w_i |e_i|^\gamma / \sum_{i=1}^{n} w_i \right]^{\frac{1}{\gamma}} \tag{4.3}$$

where $|e_i| = |(O_i - P_i)|$ are the absolute values (or magnitudes) of individual relative errors, $\gamma \geq 1$ and w_i is a scaling assigned to each $|e_i|^\gamma$ for reflecting its hypothesized influence on the total error.

Let us define the two following estimators of the average error expressed in (4.3): the Root Mean Square Error (RMSE), which is the most commonly used statistical measure of the quality of an estimator or predictor, and the Mean Absolute Error (MAE). For illustration purposes, here we recall

some notations, which we are going to specify within our application in the following Subsection 4.7.1.

For the RMSE we use equation (4.3), putting $w_i = 1$ and $\gamma = 2$; for the MAE, we use equation (4.3), setting $w_i = 1$ and $\gamma = 1$.

Compared to the MAE, in the computation of the RMSE each individual error e_i affects the total error (namely the sum of all the individual errors) in proportion to its square, rather than just its magnitude, so that large errors have a stronger influence on the total square error than small ones. Furthermore, the RMSE increases with the variance of the frequency distribution of the error magnitudes; in other words, the total square error grows as the total error is concentrated within a decreasing number of increasingly large individual errors (for a more complete discussion, we refer the reader to Willmott and Matsuura (2005)).

The mCBD model implies the increase in the number of parameters used to predict the outcome. Concerning this topic, in what follows we assess if the improvements made by adjusting the forecast by means of the CIR process justifies the increase in the number of parameters. As stressed in Burnham and Anderson (2004), on the basis of the famous Box's statement "All models are wrong but some are useful" (see Box (1976)), models, by definition, are only approximations to unknown reality or truth. Neither the CBD nor the mCBD model can perfectly reflect the real dynamics of the survival phenomenon, but, within the model-based inference framework, we can assess which of the two methods performs better on the basis of three principles: *simplicity and parsimony*, *multiple working hypotheses* and *strength of evidence*. Between our two competing hypotheses, the CBD model is the one characterized by the lowest number of parameters (two against the overall five parameters characterizing the mCBD model). In order to deepen the impact of the mCBD larger number of parameters, we look at information criteria. Here, we assess if, within our application, the simplest model for the number of parameters, namely the CBD model, should be preferred, computing the Bayesian Information Criterion (BIC) or Schwarz Criterion:

$$BIC = n \ln(\tau/n) + z \ln(n) \tag{4.4}$$

where: n, the number of points in the look-forward window, is equal to 35 under both the settings; z, the number of parameters, is 2 or 5, according to the concerned model (the CBD model and the mCBD model, respectively); τ

represents the residual sum of squares. We see that the BIC penalizes models with a large number of parameters. Thus, it is well suited to compare between the CBD and the mCBD models. z could also be considered equal to $2n$ and to $2n + 3$ for the CBD model and the mCBD model, respectively, since the two parameters of the CBD model, $k_t^{(i)}, i = 1, 2$, vary with time, but this would not change the spread between the BIC of the CBD model and the one of the mCBD model.

4.3.3 Overall data set

We focus on the *period life tables* relating to the Italian females mortality. As in Chapter 2, we use the mortality data provided by HMD (data downloaded on October 20, 2016). However, we limit ourselves to the time interval [1906,2012], since it is recommended to use with extra caution the data prior to 1906 due to problems of data quality (see Glei (2015a)). We thus collect the following data from HMD: the number of deaths $d_{x,t}$ and the central exposures to the risk of death $E^c_{x,t}$, with $x = 0, ..., 90$ and $t = 1906, ..., 2012$. Within the implementation of the methodology following a static approach, as described in Subsection 3.5.1, we use the data from 1906 up to 1977 as the look-back window (in-sample) and the rest [1978,2012] as the backtesting window (out-of-sample). Furthermore, in order to test the stability of our methodological approach, we implement the backtesting procedure also in a dynamic setting, making use of the rolling windows of data described in the next Subsection.

4.3.4 Rolling windows of data

We perform the backtesting procedure explained in Subsection 3.5.1 over both the whole data set, as described in Subsection 4.3.3, and by using a fixed-length rolling window one-year-ahead through time, in order to verify if the re-optimization on the rolling windows provides updated parameters able to catch better the dynamics of the system than the static optimization.

We arrange our whole data set relating to the time interval $[1906, 2012]$ into two different ways obtaining in both cases 36 time horizons to be used for implementing our overall methodology. The look-back and the look-forward windows are sized according to the following two arrangements:

1. 67 years in the look-back window and 5 years in the look-forward one;
2. 62 years in look-back window and 10 years in the look-forward one.

In order to provide a few examples of how the rolling windows move forward through time, in what follows we briefly describe the first two time horizons within the first and the second arrangement of the overall data set. Concerning the first arrangement, the first time horizon is set such that the corresponding look-back and look-forward windows are the following time intervals: [1906,1972] and [1973,1977] respectively. Rolling one-year ahead with respect to the previous time horizon, the second one is characterized by the look-back and look-forward windows: [1907,1973] and [1974,1978], and so on. The same mechanism rules the second arrangement of the data set, whose time horizons, as previously mentioned, are made up of 62 points look-back windows and 10 points look-forward windows. Thus, the look-back and look-forward windows corresponding to the first time horizon are, respectively, the time intervals [1906,1967] and [1968,1977], while those ones relating to the second time horizon are, respectively, [1907,1968] and [1969,1978].

4.4 Calibration of the Models

In this Section, we describe the calibration procedures used for the two models: the CBD and the CIR model. As previously stressed, when we use a fixed time horizon, the look-back window includes years from 1906 to 1977.

4.4.1 Calibration of the CBD Model Over a Fixed Look-back Window

We perform the fitting of the CBD model using the package "StMoMo", provided in **R**. It is an implementation of the generalized Age-Period-Cohort family of stochastic mortality models (see Villegas et al. (2016a), (2016b)). We concentrate on the one-year mortality rate $q_{x,t}$. For the parameters estimation, we maximize the log-likelihood function based on the assumption that the random numbers of deaths $D_{x,t}$ are independent and follow a Binomial distribution (B) conditionally on $(q_{x,t})$ (Haberman and Renshaw (2011)):

$$D_{x,t} \sim \mathcal{B}(E^0_{x,t}, q_{x,t}). \tag{4.5}$$

In order to obtain the estimation function with the required initial exposures, say $E^0_{x,t}$, we compute them by adding half the matching reported number of deaths to the central exposures (cf. Pitacco (2007a)). Following what we stated above, we start from the inputs: $d_{x,t}$ and $E^0_{x,t}$, with $x = 18, ..., 90$ and $t = 1906, ..., 1977$. Basing on the following choices about the features of the model: a logit link function; a predictor structure as in (4.2); an empty set of parameter constraints (see Section 4.3.1), we get as outputs the fitted mortality rates $\hat{q}_{x,t}, x = 18, ..., 90, t = 1906, ..., 1977$, from which we compute the corresponding death rates $\hat{m}_{x,t}$.

4.4.2 Calibration of the CIR process over the fixed look-back window

As illustrated in Section 3.4.1, when we read the matrix of the observed death rates by rows, the CIR process in (3.2) describes the dynamics of $Y_{x,t}$ over time and for fixed ages. Our application is focused on age 65, chosen within the age range of major interest when pricing long-term insurance products, like life-annuities and longevity bonds, but we show also statistical results for other ages, also much younger than 65: 60, 45, 40, 35 and 18, chosen by way of an example. After fixing the age $\ddot{x}$, we obtain the $Y_{\ddot{x},t}$ time series as in (3.1), by calculating the ratios between the observed crude death rates $B_{\ddot{x},t}$ and the corresponding fitted ones $\hat{m}_{\ddot{x},t}$, for $t = 1906, ..., 1977$. For instance, the ratio relating to Italian females aged 65 in the starting year of the look-back window, that is 1906, is computed as follows:

$$Y_{65,1906} = \frac{B_{65,1906}}{\hat{m}_{65,1906}} \tag{4.6}$$

and, considering the same individuals, the overall time series to which we calibrate the CIR process in (3.2) is: $(Y_{65,1906}, Y_{65,1907}, ..., Y_{65,1977})$.

As explained in detail in the Appendix B.3, we estimate the vector of parameters of the CIR process $\theta = (\alpha, \beta, \sigma)$ by maximization of the log-likelihood function $\ln L(\theta)$ based on the known probability density for the process. We borrow and adapt to $Y_{\ddot{x},t}$ the model for the term structure of interest rates described in Cox et al. (1985)(cf. Section 3.4.1). Within such structure, the probability density of the state variable at time t, conditional on its value at the current time s, for some $s < t$, has indeed a closed form formula (see Ap-

pendix B.1). We perform the optimization by using the **MATLAB** function *fminsearch*, looking for the minimum of $-\ln L(\theta)$.

Relevant Remark : we highlight that we do not set any constraint on the parameters among the inputs of the minimizing function.

In Table 4.1, we show the optimal parameters provided by the calibration of the CIR process $Y_{\ddot{x},t}$ over the look-back window and the corresponding value of the log-likelihood function. We see that all the optimal parameters $\hat{\alpha}, \hat{\beta}$ and $\hat{\sigma}$, are positive and that the condition (3.3) is always fulfilled, verifying *a posteriori* that we did not need to put any constraint on them.

Table 4.1: Italian females mortality data. Estimated parameters and log-likelihood of the CIR process $Y_{\ddot{x},t}$

Age	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$	$\ln L(\hat{\theta})$
65	**0.1687**	**0.8172**	**0.0272**	**2.3366**
60	0.0794	0.7872	0.0410	1.9275
45	0.0174	1.2788	0.0498	1.6627
40	0.0892	1.0869	0.0708	1.2524
35	0.1286	1.2782	0.0638	1.2756
18	0.1538	3.3172	0.2016	-0.2978

NOTE: $\hat{\alpha}$: optimal speed of adjustment; $\hat{\beta}$: optimal long-term mean; $\hat{\sigma}$: optimal volatility; $\ln L(\hat{\theta})$: estimated log-likelihood function

The CIR process describing the dynamics over time of $Y_{\ddot{x},t}$, when $\ddot{x}$ is 18, is characterized by the highest long term mean and the biggest volatility. As stressed in Section 3.4.1, $Y_{\ddot{x},t}$ can be considered as a mortality model benchmark; therefore, we can derive some evidences about the behavior of the CBD model over the look-back period from the value taken at each calendar year by the ratio $Y_{\ddot{x},t}$ as defined in (3.1), or, more generally, from the optimal long term mean of the CIR process (3.2). If $Y_{\ddot{x},t}$ is smaller than 1, the fitted crude death rate coming from the fitting of the CBD model for the calendar year t is larger than the corresponding realized crude death rate; this means that the CBD model overestimates the real data in a given year t. The farther $Y_{\ddot{x},t}$ is from 1, the larger is the distance between the observed central death rate and the fitted one. According to these observations, the

estimated long term mean provides a good indicator of the trend of the CBD model with respect to the real data: within our application, the estimated long term mean is bigger than 1 for all the ages but 60 and 65 (see Table 4.1), denoting the propensity of the CBD model to underestimate the real central death rates observed over the look-back window for all ages except for 65 and 60.

4.4.3 Calibration of the CBD and CIR models over rolling windows

Making use of the techniques explained in Subsections 4.4.1 and 4.4.2, we calibrate the CBD model and the CIR model to the observed death rates and to the factors $Y_{x,t}$, respectively, over the fixed-length windows rolling one-year-ahead through time, as described in Subsection 4.3.4; we deepen, at a later step (see Sections 4.5.3 and 4.7.2), if the optimal parameters provided over time by the calibration adapt and catch better the dynamics of the system over the corresponding forecast periods than in the static case.

4.5 Look-forward Window: Forecasting

In this Section we move on the forecasting procedure. When a fixed time horizon is used, the look-forward window includes years from 1978 to 2012.

4.5.1 Forecasting of the CBD model over a fixed look-forward window

We perform the forecasting of the CBD model 35 years-ahead from 1977, using the package "StMoMo", which, exploiting the outcomes of the fitting procedure described in Section 4.4.1, provides, among other information, a matrix with the central projection of the mortality rates, $\dot{q}_{x,t}$, with $t = 1978, ..., 2012$, from which we compute the corresponding forecasts of the death rates $\dot{m}_{x,t}$ with $t = 1978, ..., 2012$.

4.5.2 The CIR process over the fixed look-forward window

As illustrated in Section 3.4.1, the most convenient way of obtaining the "best estimate" of the correction factor $\tilde{Y}_{\ddot{x},t}$, for any age $\ddot{x}$ and future calendar year t, consists in computing the conditional expected value of the CIR process $Y_{\ddot{x},t}$, as in (3.4). A further method for obtaining analogous results is represented by the simulation. In Cox et al. (1985), it is pointed out that the distribution of the state variable at time t given its current value at time s, for some $s < t$, up to a scale factor, is a non central chi-square distribution. As discussed in Glasserman (2003), such a property can be used to simulate the process (3.2) exactly on a discrete time grid (see the Appendix B.4 for a theoretical discussion about generation of sample paths from a CIR process).

In light of these remarks, by using (3.4), we exploit the optimal parameters provided by the calibration for getting the following "expected" projection: $(\tilde{Y}_{\ddot{x},1978}, \tilde{Y}_{\ddot{x},1979}, ..., \tilde{Y}_{\ddot{x},2012})$, namely our "best estimate" of the realization of the rv $Y_{\ddot{x},t}$ in each calendar year within the look-forward window. We validate these results simulating paths of the square-root diffusion (3.2) over the time interval [1978,...,2012], by sampling from its transition density.

4.5.3 Forecasting of the CBD and CIR models over rolling windows

Exploiting the techniques described in Subsections 4.5.1 and 4.5.2, we perform, respectively, the forecasting of the CBD model and of the CIR process 5 and 10 years-ahead, according to the chosen arrangement of our data set (see Subsection 4.3.4). Such a rolling-window analysis lets us backtest for checking the predictive performance of the two models on 36 different time horizons for each arrangement of the data set.

4.6 The "Adjusted" Projection: the "New" Model mCBD

Based on (3.1), we use the following formula to correct the CBD forecast:

$$\breve{m}_{\ddot{x},t} = \tilde{Y}_{\ddot{x},t} \dot{m}_{\ddot{x},t} \tag{4.7}$$

and, consequently, to construct "adjusted" projections of the death rates $m_{\ddot{x},t}, t = 1978, ..., 2012$. We denote such projections by $\breve{m}_{\ddot{x},t}, t = 1978, ..., 2012$. $\tilde{Y}_{\ddot{x},t}$ is the "best estimate" of the realization of the rv $Y_{\ddot{x},t}$, for a given age $\ddot{x}$ and in a given year t within the look-forward window, obtained as in Section 4.5.2, while $\dot{m}_{\ddot{x},t}$ represents the forecasted death rate for the same age and the same year coming from the CBD model (see Section 4.5.1).

In line with what we have explained in the previous Sections, in the follow-up we will handle, over the look-forward window, the forecasted death rates provided by:

1. the stand-alone CBD (hereinafter CBD) model, from which we get $\dot{m}_{\ddot{x},t}$;
2. the forecasts $\dot{m}_{\ddot{x},t}$ of the CBD model, matched with the conditional expected values (or the Monte Carlo sample paths) generated from the CIR process (3.2), globally providing us with $\tilde{m}_{\ddot{x},t}$ as in (4.7). We denote by mCBD the model exploiting the synergy between the two just mentioned models.

4.7 Assessing the Forecasting Performance of the CBD and the mCBD Models

In this Section, we study the forecasting performance of the CBD and the mCBD models through the procedures described in Section 4.3.

4.7.1 Over the fixed look-forward window [1978-2012]

This Subsection is devoted to the assessment of the quality of the forecasts over the fixed look-forward window [1978-2012], both graphically and statistically. The graphical analysis provides us with an insight of the dynamics of the forecasts of death rates and of the errors associated with them, while the statistical measures give us an estimate of their magnitude.

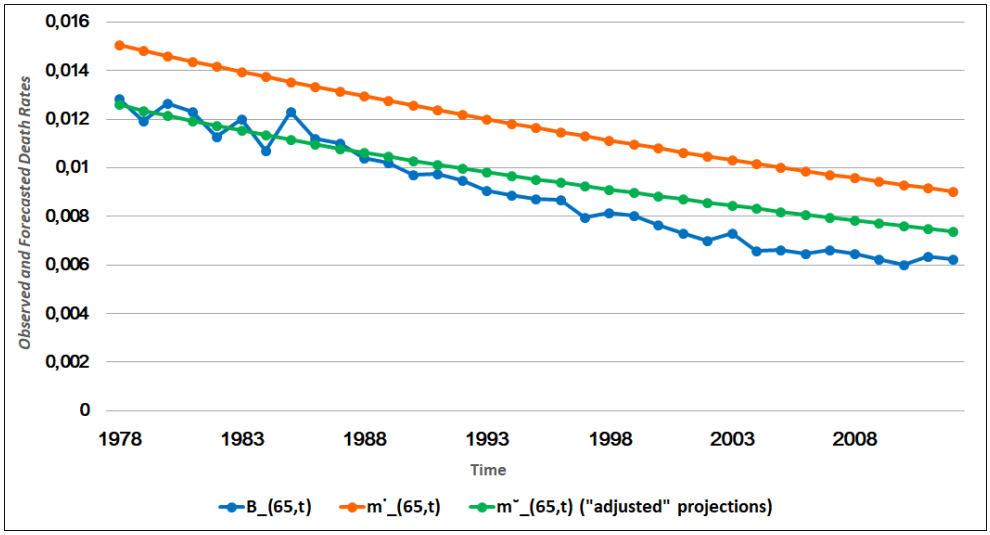


Figure 4.1: Observed and forecasted death rates for Italian females aged 65 over the fixed look-back window [1978,2012]]. Observed death rates: blue line, CBD forecasted death rates: orange line, mCBD forecasted death rates: green line

Graphical results

For each of the six fixed ages under study and for $t = 1978, \ldots, 2012$, we collect the graphical evidences of:

- the realized death rates, $B_{\ddot{x},t}$ (the blue line);
- the projections of death rates, $\dot{m}_{\ddot{x},t}$ (the orange line);
- the "adjusted" projections of death rates, $\breve{m}_{\ddot{x},t}$ (the green line).

In Figure 4.1, we show the results for age $\ddot{x} = 65$, due to the relevance of such an age for long-term insurance products and pension plans and because it is within the ages for which the CBD model was developed. The green line is closer to the blue one than the orange line, meaning that, within this empirical application, just at a first sight it is evident that, thanks to the correction factor $\tilde{Y}_{\ddot{x},t}$, the mCBD model is able to improve the forecasting performance of the CBD model in each calendar year. Indeed, on the basis of the graphs displaying the observed and forecasted death rates for all the ages under study over the same forecasting horizon, we can infer that the mCBD model predicts better than the CBD model in 100% of the ages and throughout the forecasting span.

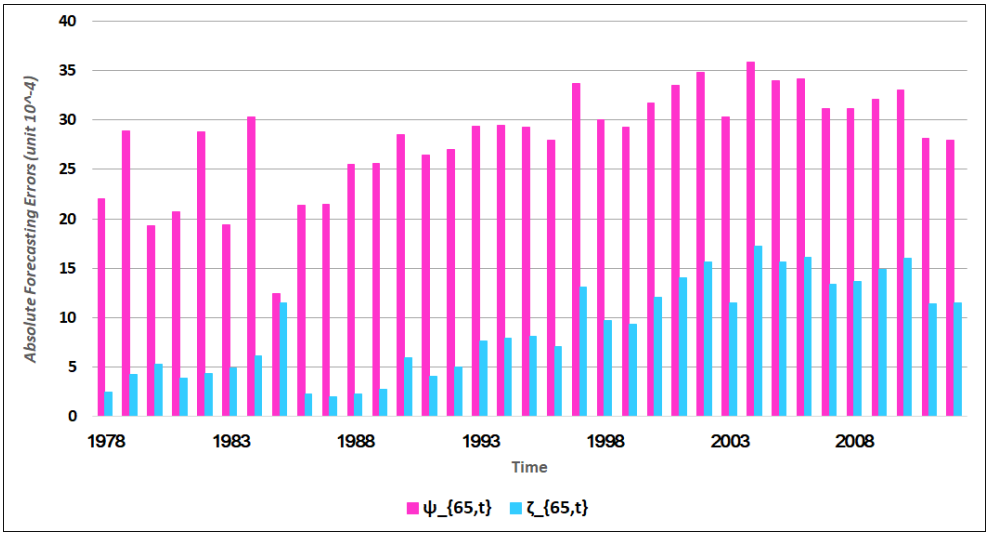


Figure 4.2: Italian females mortality data. Age 65. Absolute forecasting errors of the CBD model and the mCBD model (unit 10^{-4}) over the fixed look-back window (1978,2012). Magenta bars: absolute forecasting errors of the CBD model; cyan bars: absolute forecasting errors of the mCBD model

Statistical Assessment of the forecasting performance: measuring forecasting errors

We define the magnitudes of the forecasting errors of the CBD and the mCBD models, respectively, as $\psi_{\ddot{x},t}$ and $\zeta_{\ddot{x},t}$, where the subscript $\ddot{x}$ denotes that the age is fixed; ψ and ζ are thus only functions of time $t \in [1978, 2012]$. Both $\psi_{\ddot{x},t}$ and $\zeta_{\ddot{x},t}$ are designed to measure the absolute forecasting error between the real crude death rate, observed in a given year of the look-forward window, and the corresponding forecast; in particular:

- $\psi_{\ddot{x},t}$ quantifies the absolute forecasting error between the realized crude death rates, $B_{\ddot{x},t}$, and their projections, $\dot{m}_{\ddot{x},t}$, obtained as output of the CBD model;
- $\zeta_{\ddot{x},t}$ quantifies the absolute forecasting error between the realized crude death rates, $B_{\ddot{x},t}$, and their "adjusted" projections, $\breve{m}_{\ddot{x},t}$, computed through the mCBD model as in (4.7).

In Figure 4.2, we show, again for age 65, the absolute forecasting errors of the mCBD model $\zeta_{\ddot{x},t}$ (the cyan bars) and of the CBD model $\psi_{\ddot{x},t}$ (the magenta

Table 4.2: Italian females mortality data. Measures of the average forecasting error (unit 10^{-4})

Age	RMSE		*Impr.Rate*	MAE		*Impr.Rate*
	ψ^*	ζ^*	IR^*	$\bar{\psi}$	$\bar{\zeta}$	$\bar{IR}$
65	**28.60**	**10.13**	***-65%***	**28.13**	**8.94**	***-68%***
60	10.01	3.08	*-69%*	9.72	2.13	*-78%*
45	3.54	1.69	*-52%*	3.48	1.48	*-58%*
40	3.00	2.18	*-27%*	2.95	2.01	*-32%*
35	2.68	1.79	*-33%*	2.56	1.56	*-39%*
18	2.29	1.26	*-45%*	2.24	1.14	*-49%*

NOTE: ψ^* and $\bar{\psi}$ are, respectively the RMSE and the MAE of the CBD model; ζ^* and $\bar{\zeta}$ are, respectively, the RMSE and the MAE of the mCBD model; IR^*: Improvement Rate when the RMSE is concerned; $\bar{IR}$: Improvement Rate when the MAE is concerned

bars), over the look-forward window $[1978, ..., 2012]$ (unit 10^{-4}). Confirming the evidences provided in Figure (4.1), $\zeta_{\ddot{x},t}$ turns out to be smaller than $\psi_{\ddot{x},t}$ in 100% of both ages and calendar years.

Starting from the absolute forecasting errors, $\psi_{\ddot{x},t}$ and $\zeta_{\ddot{x},t}$, for each age under study, we compute the RMSE and the MAE of the CBD model (ψ^* and $\bar{\psi}$, respectively) and the RMSE and the MAE of the mCBD model (ζ^* and $\bar{\zeta}$, respectively). In Table 4.2, we report such measures of the average forecasting error (unit 10^{-4}) for age 65 and also for younger ages, in order to give an overview on the predictive performance of the mCBD model, that is not limited to age 65. In Table 4.2, we provide also a synthetic measure of the ability of the factor $\tilde{Y}_{\ddot{x},t}$ in correcting, on average and throughout the look-forward window, the outcomes of the CBD model, that is what we call "Improvement Rate" (IR). In particular, we consider the two following measures: IR^* $((\zeta^* - \psi^*)/\psi^*)$, quantifying how much, in percent, the mCBD model is able to reduce the RMSE of the CBD model; and $\bar{IR}$ $((\bar{\zeta} - \bar{\psi})/\bar{\psi})$, measuring the percent gain provided by the mCBD model in the average forecasting performance, when MAE is concerned.

If, for each age under study, we look at ψ^* compared to $\bar{\psi}$ and at ζ^* compared to $\bar{\zeta}$, we notice that for all ages and for both the CBD model and the mCBD model, the RMSEs, ψ^* and ζ^*, are slightly bigger than the corresponding MAEs, $\bar{\psi}$ and $\bar{\zeta}$ respectively. For both the CBD model and the mCBD model, the largest difference between the RMSE and the MAE corresponds

to age 65, for which, in both cases, the distribution of the error magnitudes is characterized by the highest standard deviation. It is important to highlight, in Table 4.2, that the mCBD model is characterized by smaller values of both the RMSE and the MAE than the CBD model, this meaning that it turns out to be more accurate in predicting future death rates for each age under study.

Moreover, it is worth noting that in Table 4.2 both the RMSE and the MAE of the CBD model are bigger for age 65 than for age 18, although this model performs better for older ages. This does not happen if we base our computations on the relative forecasting errors (namely $e_i = (O_i - P_i)/O_i$, cf. Section 4.3.2).

Since, as reported in Table 4.2, both ψ^* and $\bar{\psi}$ are always bigger than, respectively, ζ^* and $\bar{\zeta}$ for all ages, IR^* and $\bar{IR}$ are always negative; the higher such rates are, in absolute terms, the larger the magnitude of the average correction provided by $\tilde{Y}_{\ddot{x},t}$ is. The improvement on the MAE is higher than on RMSE, indicating that the larger movements are worse forecasted than the average movements. In Table 4.2, IR^* and $\bar{IR}$ take the highest absolute values for the following ages: 60 (being characterized by the biggest values of both the measures) and 65; therefore, among the six ages under study, the mCBD model provides the largest improvement of the predictive accuracy of the CBD model for ages 60 and 65, both when the RMSE and the MAE are concerned. We consider this observation important for its implications in the pension demographic assessment.

Relevant remark: By looking at Tables 4.1 and 4.2, we can infer an intuitive and general empirical rule about the features of $\tilde{Y}_{\ddot{x},t}$. As stated in Section 3.4.1, when the optimal long-term mean β is very close to 1, the chosen mortality model fits well to the data within the look-back window used for calibrating $Y_{\ddot{x},t}$. Pulled toward 1, $\tilde{Y}_{\ddot{x},t}$ has a weaker and weaker ability to drive the mortality outcomes of the model toward the real data. Within our application, we find evidence of that with age 40, for which $\hat{\beta}$ is equal to 1.08 and both IR^* and $\bar{IR}$ take the smallest (absolute) value.

Table 4.3: Italian females mortality data. Bayesian Information Criterion

Age	$BICsa$	$BICm$	ΔBIC	*Change Rate*
65	**-402.89**	**-464.83**	**61.94**	***-15.4%***
60	-476.39	-548.26	71.87	*-15.1%*
45	-549.05	-590.04	40.99	*-7.5%*
40	-560.67	-572.24	11.57	*-2.1%*
35	-568.59	-586.11	17.52	*-3.1%*
18	-579.68	-610.94	31.26	*-5.4%*

NOTE: "$BICsa$" is the Bayesian Information Criterion of the CBD model; "$BICm$" is the Bayesian Information Criterion of the mCBD model; ΔBIC is the first difference between "$BICsa$" and "$BICm$"; *Change rate* represents the percent difference between "$BICsa$" and "$BICm$"

Bayesian Information Criterion

As the next step, we look at the BIC as defined in (4.4) to assess the quality of the improvement in our forecast. We denote by "BICsa" ("sa" for stand-alone) the Bayesian Information Criterion related to the CBD model and by "BICm" the one related to the mCBD one. Table 4.3 reports "BICsa", "BICm", ΔBIC, that is the first difference between "BICsa" and "BICm", the percent *Change rate* (computed as $(\Delta BIC/BICsa)$). We can see that "BICm" is smaller than "BICsa" for all the ages under study; the *Change rate* is indeed always negative. The highest percent change between "BICm" and "BICsa", 15.4%, is found for age 65. When $\Delta BIC > 10$, it is usually accepted that there is strong evidence against the model with the highest BIC. ΔBIC suggests that, for all the ages under study, we have a decisive evidence against the *null hypothesis*, which is the one with the highest BIC value, namely the CBD model.

As illustrated in Kass and Raftery (1995), the Bayes factor is a summary of the evidence provided by the data in favor of one scientific theory, represented by a statistical model, as opposed to another. The Bayesian Information Criterion provides a rough approximation to the logarithm of the Bayes factor, which is easy to use and does not require evaluation of prior distributions (we refer the reader to Kass and Raftery (1995), Weakliem (1999) and Raftery (1999) for a detailed discussion about Bayes Factor and BIC).

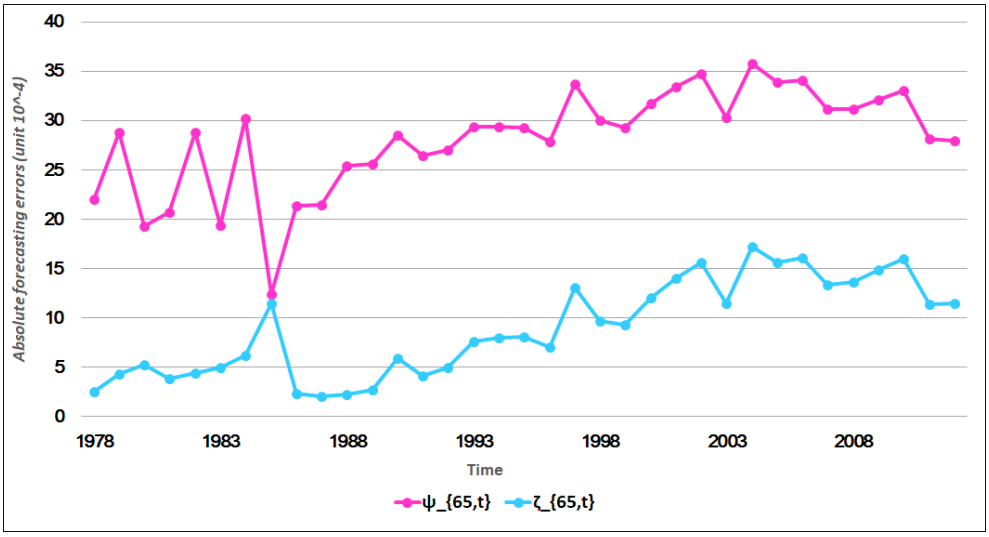


Figure 4.3: Absolute forecasting errors trend over the look-forward window [1978,2012] for the Italian females aged 65. Magenta line: absolute forecasting errors of the CBD model; cyan line: absolute forecasting errors of the mCBD model

4.7.2 Assessing the forecasting performance over rolling windows

In this Section we describe the results of the forecasting procedure performed over rolling windows, as presented in Subsection 4.5.3.

Before illustrating such results, let us briefly focus on Figure 4.3, displaying how the absolute forecasting errors of the CBD model (the magenta line) and those of the mCBD model (the cyan line) behave over the time interval $[1978, 2012]$ for age 65, chosen as the fundamental age for our study, due to its actuarial relevance. $\zeta_{\ddot{x},t}$, plotted against time, presents a trend that can be observed for all the ages under study. As previously mentioned, performing our methodology over the fixed-length windows rolling one-year-ahead through time described in Subsection 4.3.4, allows us to verify if the optimal parameters provided over time by the calibration can adapt and catch better the dynamics of the system over the corresponding forecast periods.

Let us now examine in detail the resulting predictive performance of the CBD and the mCBD model: $\tilde{Y}_{\ddot{x},t}$ turns out to improve the predictive performance of the CBD model in all cases. As an example, we can look at: the age 65 and the 33rd time horizon, made by the 62-years look-back window [1938,1999]

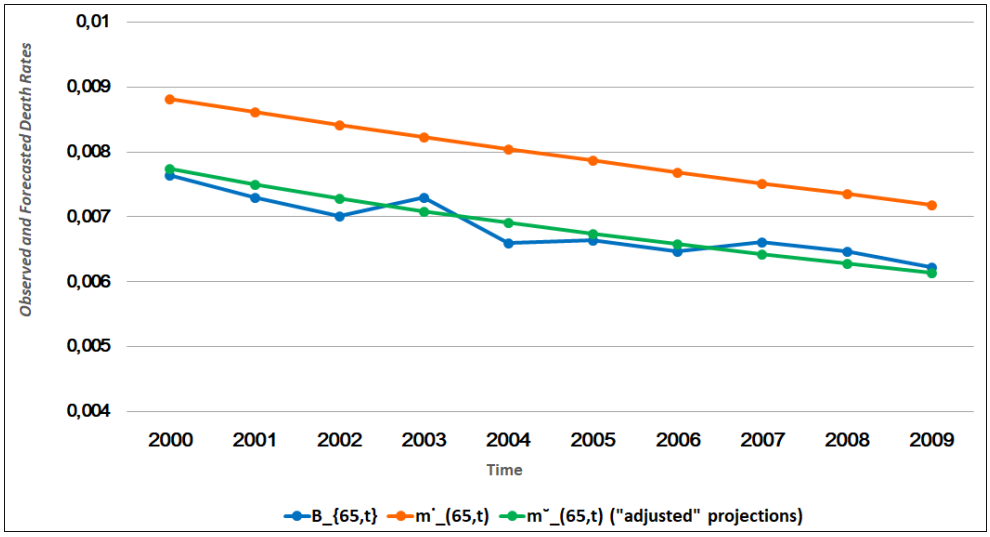


Figure 4.4: Italian females mortality data, age 65 and 33rd 10-years look-forward window: realized and forecasted death rates. Blue line: realized crude death rates; orange line: projections of crude death rates obtained as output of the CBD model; green line: "adjusted" projections of crude death rates coming from the mCBD model.

and the 10-years look-forward window [2000,2009]. As to this settings, Figure 4.4 presents over the look-forward window:

- the realized death rates, $B_{65,t}$ (the blue line);
- the projections of death rates, $\dot{m}_{65,t}$ (the orange line), by the CBD model;
- the "adjusted" projections of death rates, $\breve{m}_{65,t}$ (the green line), by the mCBD model.

In Figure 4.4, we see that $\breve{m}_{65,t}$ is, for each $t \in [2000, 2009]$, much closer to the corresponding realized crude death rate than $\dot{m}_{65,t}$. In Figure 4.5 we also show, for age 65 and setting as unit 10^{-4}, $\psi_{65,t}$ (the absolute forecasting error between the realized crude death rates, $B_{65,t}$, and their projections, $\dot{m}_{65,t}$, obtained as output of the CBD model) and $\zeta_{65,t}$ (which quantifies the absolute forecasting error between the realized crude death rates, $B_{65,t}$, and their "adjusted" projections, $\breve{m}_{65,t}$, computed through the mCBD model), $t = 2000, ..., 2009$. In Figure 4.5, we see how $\psi_{65,t}$ (magenta line) and $\zeta_{65,t}$ (cyan line) behave over the 33rd 10-years look-forward window, namely the period [2000,2009], for age 65. We note that the absolute forecasting errors

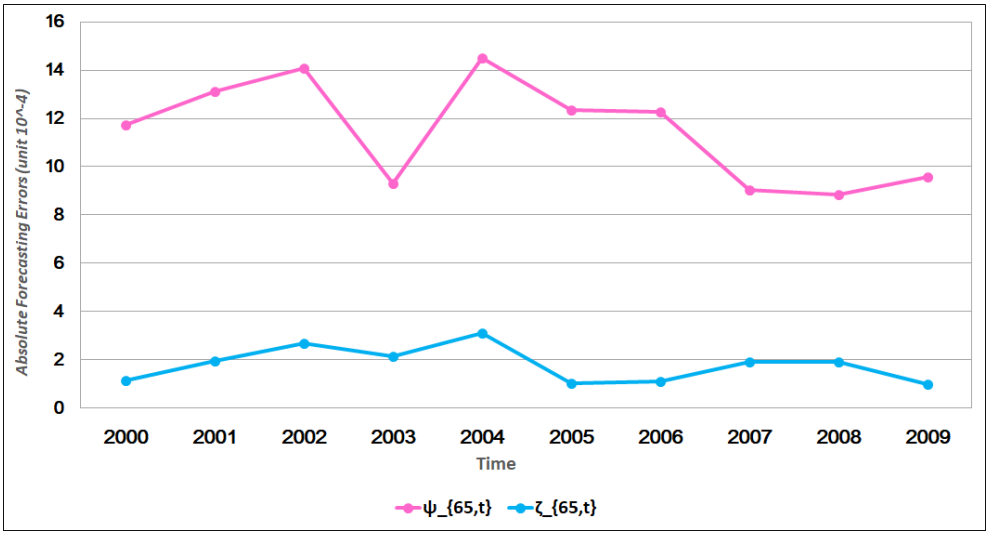


Figure 4.5: Italian females mortality data, age 65 and 33rd 10-years look-forward window (2000,2009): absolute forecasting errors trend. Magenta line: absolute forecasting errors of the CBD model; cyan line: absolute forecasting errors of the mCBD model.

of the mCBD model have lost their trend.

For all the other time horizons (whether the look-forward windows are made up by 5 years or 10 years) we find that, after implementing our methodology over rolling windows, the absolute forecasting errors of the mCBD model have no trend any more. We can thus infer that, when making use of rolling windows, $\zeta_{\ddot{x},t}$ looses its trend and the dynamics of the system is better captured by the re-optimization on the rolling windows (dynamic case) than it is in the static case.

4.7.3 Predictive accuracy under the static and the dynamic approach

As we have explained in Section 4.3, we implement the backtesting procedure over both an unchanging calibration sample (static approach) and fixed-length windows for the calibration sample that is rolling one year ahead through time (dynamic approach). In Section 4.7 we have assessed the forecasting performance of the CBD and the mCBD models, finding out that, under both the static and the dynamic approach, the mCBD model provides

better results, in terms of predictive accuracy, than the CBD model.

The last step of our study consists in comparing the forecasting performance of the CBD and the mCBD models between the static and the dynamic case, in order to understand if the predictive ability of our models deteriorates going further and further in time and, in such a case, how much accuracy we are able to gain using the dynamic approach. For this purpose, we compute and compare the RMSE of our models in the static and in the dynamic case, by exploiting two different methods.

With the first method, the computation of the RMSEs of the two models is performed over the forecasting time horizon [1978-2012], that is the look-forward window we use for backtesting the CBD and the mCBD models in Sections 4.5.1 and 4.5.2, under the static approach. As illustrated in Section 4.7.1, using a fixed look-forward window gives rise, for the CBD and the mCBD models, to the RMSEs reported in Table 4.2. Such outcomes are compared to the RMSEs we get under the dynamic approach described in Section 4.5.3, when 36 10-years look-forward windows are concerned. Under the dynamic approach, the look-back windows roll one-year ahead through time, so each of them, compared to the previous one, includes the same years except for the first and the last ones. It results that all the years within the time interval (1978, 2012), apart from those at the beginning and at the end, are involved in the forecasting procedure more than once; we thus have available, for them, at least one and at most 36 corresponding realizations of the absolute forecasting error, except that the forecasting horizon changes: the second year in the first case becomes the first year when we move the calibration sample from one year. In order to associate only one forecasting error to the corresponding year, for each year we compute the RMSE across the 36 look-forward windows of the dynamic approach for both the CBD and the mCBD models. In Table 4.4 we report, for age 65, the RMSEs of the CBD model and the mCBD one, in both the static and the dynamic case, and the rate we call "Accuracy Gain". For each model, the "Accuracy Gain" is the percentage change of its RMSE when moving from the static to the dynamic approach. Using the dynamic approach allows us to enhance the predictive accuracy of both the CBD and mCBD models; in particular, the forecasting performance of the mCBD model has been remarkably improved under the dynamic setting. The problem with this procedure is that we compute an average error over different forecasting horizons. In order to compare the

Table 4.4: Italian females mortality data. Age 65. First method: RMSEs of the CBD and the mCBD models under the static and the dynamic approach (unit 10^{-4})

Model	Static Approach	Dynamic Approach	*Accuracy Gain*
CBD	28.60	16.82	*-41%*
mCBD	10.13	4.08	*-60%*

Table 4.5: Italian females mortality data. Age 65. Second method: RMSEs of the CBD and the mCBD models under the static and the dynamic approach (unit 10^{-4})

Span	CBD			mCBD		
	Static	Dynamic	*Acc.Gain*	Static	Dynamic	*Acc.Gain*
1968/77	26.58	26.58	*0.0%*	4.99	4.99	*0.0%*
1978/87	30.49	21.46	*-29.62%*	5.70	5.83	*2.28%*
1988/97	35.76	15.13	*-57.69%*	12.02	1.68	*-86.02%*
1998/07	39.66	12.92	*-67.42%*	18.89	2.32	*-87.72%*

errors over similar forecasting horizons, we design a second method.

With the second method, we implement the backtesting procedure under the following settings: under the static approach, we use the time interval [1906-1967] as a fixed lookback window (the calibration sample) and the time interval [1968-2012] as a fixed lookforward window; under the dynamic approach, as above explained, we arrange the overall data set described in Section 4.3.3 in such a way to exploit 36 different time horizons: each of them is characterized by 62-years lookback windows and 10-years lookforward windows. The lookback and the lookforward windows of the first time horizon coincide with the unchanging lookback and lookforward windows we employ under the static approach. The last lookback window is made up by the years from 1941 to 2002 and the last lookforward window includes years from 2003 to 2012.

We compare the static and the dynamic case, by looking at the RMSEs of the CBD and the mCBD models over four consecutive 10-years forecasting spans, starting from 1968. Since the mortality data concerning the Italian population (as available when dowloaded) end in 2012, we cannot set up a

further 10-years forecasting span after 2007. In Table 4.5 we report, for the four forecasting spans under study, the RMSEs of the CBD model and the mCBD one under the static and the dynamic approach and the "Accuracy Gain". Concerning the first span, [1968-1977], the CBD and the mCBD model equally perform in the static and the dynamic case since, under both the approaches, the calibration procedures involve mortality data relating to the same lookback window, so that the same resulting parameters are exploited for forecasting 10-years ahead. The different reasonings underlying the static and the dynamic approaches and the resulting different outcomes become evident from the second forecasting span. Indeed, the forecasts corresponding to the second decade are obtained: in the static case, still using the parameters of the CBD and the mCBD models provided by their calibration to the look-back window [1906-1967], while, in the dynamic case, exploiting the parameters of the CBD and mCBD models resulting from their calibration to the lookback window [1916-1977]. Similar remarks can be done for the subsequent forecasting spans: while in the static case the forecasting procedure resorts always to the same historical information, regardless of the prediction horizon we are interested to, in the dynamic case it is based on the latest 62 data points preceding the first year of the prediction span. Under the dynamic approach, we are thus able to update the historical data sample to be used for the calibration, filtering out the old information and incorporating the most recent one, time by time. In Table 4.5 we report the RMSEs of the CBD model and the mCBD one, in both the static and the dynamic case, for age 65. As the "Accuracy Gain" in Table 4.5 suggests, the farther in time we forecast, the more the precision of the CBD model is enhanced if we select the dynamic approach as our "modus operandi". This behavior is to be expected, since, in the static case, the forecasting horizon becomes longer and longer. Except for the second prediction span, the dynamic approach always enables to enhance the forecasting performance of the mCBD model; for such a model, the improvement of its predictive accuracy takes a very remarkable magnitude, meaning that the mCBD model not only outperforms the CBD model in forecasting central death rates, but also shows a stronger potential to gain accuracy in the long-run under the dynamic approach.

4.8 Conclusion

In this Chapter we have designed and implemented the affine model called mCBD, we propose for describing the evolution of the mortality phenomenon. This model is multiplicative and combines two components: the deterministic output (baseline) of the CBD model and an affine stochastic factor affecting the baseline over the forecasting time horizon. Such an affine stochastic factor is described by a Cox-Ingersoll-Ross process suitably fitted to a measure of the error of the CBD model. Thus the mCBD affine model exploits as driving process the CIR one, that is tasked to describe the dynamics of the ratio between the observed central death rates and the corresponding fitted values obtained as outputs of the CBD model on the look-back window; on the look-forward window the baseline produced by the CBD model is matched with the conditional expected value of the CIR process (or the Monte Carlo sample paths generated from it). We have exploited the Italian females mortality data and made use of backtesting procedures, combining both static and dynamic analyses, for empirically testing the performance of the CBD and the mCBD models in forecasting the mortality profile of different ages (with the main focus on age 65). By means of average measures of the forecasting errors and information criteria, we demonstrate that using the mCBD model provides better results in terms of predictive accuracy than the CBD model. Following Box (1976), we believe that all models are wrong and thus that scientists cannot obtain a "correct" model by excessive elaboration. Therefore, as William of Occam suggests, we seek a parsimonious description of the natural phenomenon, that is, in our case, the mortality phenomenon. Combining the CBD model with a CIR process lets us obtain good results exploiting only three additional parameters compared to the CBD model. Even when correcting the results for the additional parameters, the BIC shows that the gained accuracy compensates the additional complexity of the model. Moreover, using a dynamic setting for the optimization remarkably improves the quality of the mCBD model, making it suitable for pricing of longevity bonds and planning for retirement pension funds.

To conclude, we point out that our procedure is not restricted to the choice of the CBD model. The overall applicability and validity of our numerical procedure can be tested on any mortality projection model (cf. Apicella et al. (2018)), provided the errors can be modelled as a CIR process, and any mortality data set.

Chapter 5

Further Research

In this Chapter, we are going to explain, providing also some examples of application, the further research to which this Thesis is believed to leave scope.

5.1 Testing the mCBD Model on Different Mortality Data Sets

The mCBD model could be tested against the CBD model by using mortality data sets concerning different Countries, in order to strengthen the validity of the methodology proposed in Chapter 3 and, also, investigate geographic patterns of mortality.

By way of an example, making use of the same calibration and forecasting techniques illustrated in Chapter 4, we empirically test the predictive performance of the CBD and the mCBD models, exploiting the Swedish females mortality data provided by HMD (data downloaded on October 20, 2016). Due to problems of data quality, we limit ourselves to the time interval [1861,2014] (cf. Glei (2015b)). We show the results concerning age 65, which is the age of our major interest, obtained by implementating our methodology under the static approach: under this setting, we use data from 1861 up to 1962 as the look-back window (in-sample) and the rest [1963,2014] as the look-forward window (out-of-sample).

Table 5.1: Swedish mortality data. Optimal parameters and log-likelihood of the CIR process $Y^{swe}_{65,t}$ fitted on the look-back window (1861,2014)

Age	$\hat{\alpha}_{swe}$	$\hat{\beta}_{swe}$	$\hat{\sigma}_{swe}$	$\ln L(\hat{\theta}_{swe})$
65	0.4472	0.7943	0.0679	1.5804

NOTE: $\hat{\alpha}_{swe}$: optimal speed of adjustment; $\hat{\beta}_{swe}$: optimal long-term mean; $\hat{\sigma}_{swe}$: optimal volatility; $\ln L(\hat{\theta}_{swe})$: estimated log-likelihood function

In Table 5.1, we report the optimal parameters and the corresponding value of the log-likelihood function provided by the calibration, over the look-back window [1861,1962], of the CIR process (3.2) (that we denote by $Y^{swe}_{65,t}$ when Swedish data are concerned).
$\hat{\alpha}_{swe}, \hat{\beta}_{swe}, \hat{\sigma}_{swe}$ are positive and condition (3.3) is fulfilled. As already seen with the Italian data and, in particular, with ages 60 and 65 (see Section 4.4.2), the estimated long-term mean, $\hat{\beta}_{swe}$, is smaller than 1, denoting the propensity of the CBD model to overestimate the real death rates observed over the look-back window. Specifically, $Y^{swe}_{65,t}$ is characterized by a value of $\hat{\beta}_{swe}$ that is very close to the estimated values of the long-term mean we obtained for ages 60 and 65 in the Italian case, but by a much higher speed of adjustment $\hat{\alpha}_{swe}$.

In Figure 5.1, we show, over the look-forward window [1962,2014], the graphical evidences of:

- the realized death rates, $B^{swe}_{65,t}$ (the blue line);
- the projections of death rates provided by the CBD model, $\dot{m}^{swe}_{65,t}$ (the orange line);
- the "adjusted" projections of death rates provided by the mCBD model, $\breve{m}^{swe}_{65,t}$ (the green line)

The green line is closer to the blue one than the magenta line; accordingly, within this empirical application, the mCBD model enhances the predictive accuracy of the model itself, throughout the look-forward window.

This is confirmed by the graphical evidences of Figure 5.2, in which we display the absolute forecasting errors of the CBD model $\psi^{swe}_{65,t}$ (magenta bars) and those of the mCBD model $\zeta^{swe}_{65,t}$ (cyan bars). $\zeta^{swe}_{65,t}$ turns out to be smaller

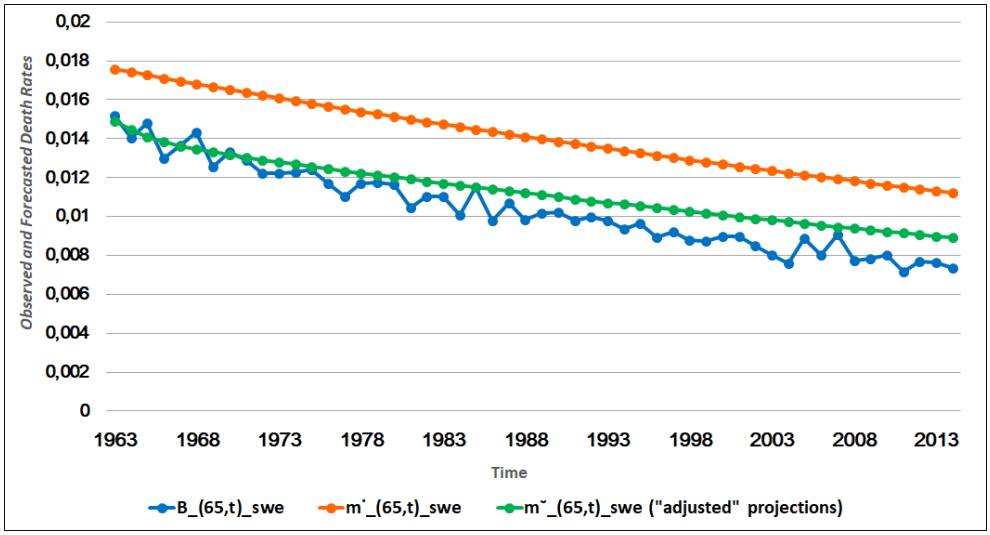


Figure 5.1: Observed and forecasted death rates for Swedish females aged 65 over the look-forward window [1962,2014]). Observed death rates: blue line; CBD forecasted death rates: orange line; mCBD forecasted death rates: green line.

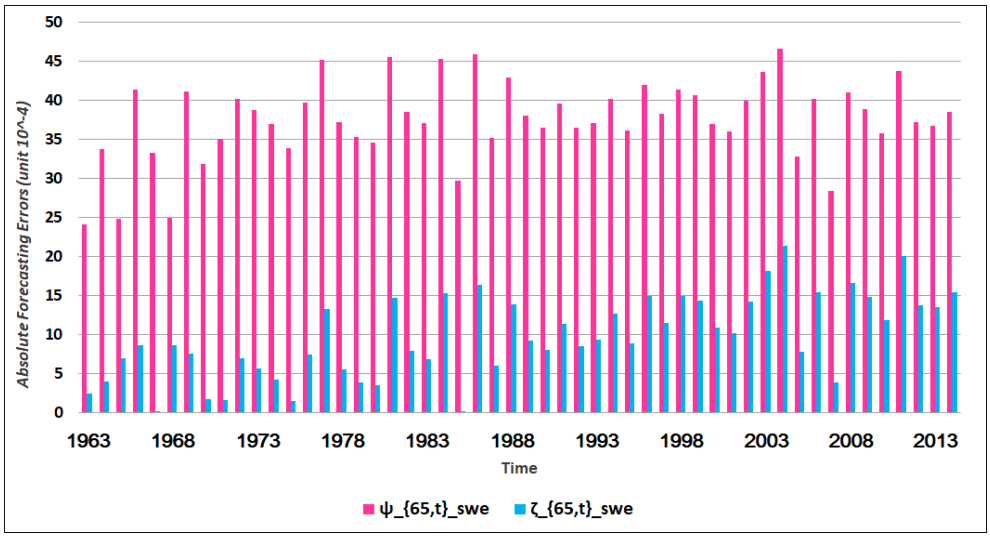


Figure 5.2: Sweden. Age 65. Absolute forecasting errors of the CBD model and the mCBD model (unit 10^{-4}) over the fixed look-forward window (1962,2014). Magenta bars: absolute forecasting errors of the CBD model; cyan bars: absolute forecasting errors of the mCBD model

than $\psi^{swe}_{65,t}$ in 100% of calendar years.

Table 5.2: Swedish mortality data. Age 65. Measures of the average forecasting error (unit 10^{-4})

Age	RMSE		*Impr.Rate*	MAE		*Impr.Rate*
	ψ^*_{swe}	ζ^*_{swe}	IR^*_{swe}	$\bar{\psi}_{swe}$	$\bar{\zeta}_{swe}$	$\bar{IR}_{swe}$
65	37.92	11.05	*-71%*	37.58	9.72	*-74%*

NOTE: ψ^*_{swe} and $\bar{\psi}_{swe}$: respectively, RMSE and MAE of the CBD model; ζ^*_{swe} and $\bar{\zeta}_{swe}$: respectively, RMSE and MAE of the mCBD model; IR^*_{swe}: Improvement Rate when the RMSE is concerned; $\bar{IR}_{swe}$: Improvement Rate when the MAE is concerned

The last step of this short numerical application consists in examining the measures of the average forecasting errors (RMSE and MAE) and at the Bayesian Information Criterion (BIC), in order to assess and compare the forecasting performance of our two competing models. We report such statistical measures in Tables 5.2 and 5.3, respectively.
The mCBD model is characterized by smaller values of both the RMSE and the MAE than the CBD model, this meaning that it shows itself to be more accurate in predicting future death rates for age 65. In particular, compared to the CBD model, the mCBD one allows to reduce the RMSE by 71% and the MAE by 74% (see IR^*_{swe} and $\bar{IR}_{swe}$, respectively, in Table 5.2).

Table 5.3: Swedish mortality data. Age 65. Bayesian Information Criterion (BIC) of the CBD and the mCBD models

Age	$BICsa^{swe}$	$BICm^{swe}$	ΔBIC^{swe}	$ChangeRate^{swe}$
65	-571.89	-688.30	116	*-20%*

NOTE: $BICsa^{swe}$ is the Bayesian Information Criterion of the CBD model; $BICm^{swe}$ is the Bayesian Informatin Criterion of the mCBD model; ΔBIC^{swe} is the first difference between $BICsa^{swe}$ and $BICm^{swe}$; $Changerate^{swe}$ represents the percent difference between $BICsa^{swe}$ and $BICm^{swe}$

As reported in Table 5.3, the BIC of the mCBD model, $BICm^{swe}$, is smaller than the one related to the CBD model, $BICsa^{swe}$; the percent spread between $BICm^{swe}$ and $BICsa^{swe}$, in absolute value, amounts to 20%. More-

over, ΔBIC^{swe} is much bigger than 10, so we have strong evidence in favor of the mCBD model.

5.2 Exploiting the mCBD Model for Predicting the Mortality Dynamics for a given Cohort

The corrective ability of the mCBD model could be evaluated when describing the mortality phenomenon of a given cohort. In this sense, the mCBD model would not be supposed to compete with the extensions of the CBD model allowing for a cohort effect, such as the M6, M7 and M8 models described in Cairns et al. (2009). Instead, the implementation of the mCBD model would be useful for exploring the potentiality of our corrective methodology also when cohort life tables are used. The results we present are just preliminary outcomes, that may be improved, choosing a different mortality model to provide the baseline, for instance.

By way of an example, making use of the same calibration and forecasting techniques illustrated in Chapter 4, we make use of the methodology presented in Chapter 3 for predicting the mortality dynamics of the generation of Italian males born in 1940. We collect the following data from the HMD (data downloaded on October 20, 2016): the number of deaths $d_{x,t}$ and the central exposures to the risk of death $E^c_{x,t}$, with $x = 0, ..., 72$ and $t = 1940, ..., 2012$, and arrange them into two matrices. We get the observed central death rates $(B_{0,1940}, B_{1,1941}, ..., B_{72,2012})$, by equation (1.9), after extrapolating the diagonal of each matrix. Also for this numerical application, we show the results obtained by implementating our methodology under the static approach: under this setting, we have used data from 1940 up to 1989 as the look-back window (in-sample) and the rest [1990,2012] as the look-forward window (out-of-sample).

When we are interested in representing the evolution of mortality for a given cohort (diagonal approach), the CIR process in (3.2) describes the dynamics over time of $Y_{x,t}$, where the age x is not fixed, but it implicitly moves forward time by time. Within this application, the time series to which we calibrate the CIR process (3.2) (we call $Y^{coh}_{x,t}$) is: $(Y^{coh}_{0,1940}, Y^{coh}_{1,1941}, ..., Y^{coh}_{49,1989})$. In Table

Table 5.4: Italian males mortality data. Generation 1940. Optimal parameters and log-likelihood of the CIR process $Y_{x,t}^{coh}$ fitted on the look-back window (1940,1989)

Cohort	$\hat{\alpha}_{coh}$	$\hat{\beta}_{coh}$	$\hat{\sigma}_{coh}$	$\ln L(\hat{\theta}_{coh})$
1940	1.4892	0.4208	0.2866	0.7886

NOTE: $\hat{\alpha}_{coh}$: optimal speed of adjustment; $\hat{\beta}_{coh}$: optimal long-term mean; $\hat{\sigma}_{coh}$: optimal volatility; $\ln L(\hat{\theta}_{coh})$: estimated log-likelihood function

5.4, we report the optimal parameters and the resulting value of the log-likelihood function coming from such calibration. The optimal parameters, $\hat{\alpha}_{coh}, \hat{\beta}_{coh}$ and $\hat{\sigma}_{coh}$, are positive and fulfill condition (3.3). Compared to the other case studies discussed in this Thesis, we find the highest value of the speed of adjustment $\hat{\alpha}_{coh}$, meaning that $Y_{x,t}^{coh}$ is pulled faster toward its long-term mean; this feature affects the path of the correction factor $\tilde{Y}_{x,t}^{coh}$ over the look-forward window.

As previously mentioned, the look-forward window includes years from 1990 to 2012. Over this time horizon and for x moving forward, year by year, from 50 to 72, we get and display in Figure 5.3:

- the realized death rates, $B_{x,t}^{coh}$ (the blue line);
- the projections of death rates, $\dot{m}_{x,t}^{coh}$ (the orange line);
- the "adjusted" projections of death rates, $\breve{m}_{x,t}^{coh}$ (the green line).

In most of the calendar years, the green line is closer to the blue one than the orange line. This means that, within this empirical application, thanks to the correction factor $\tilde{Y}_{x,t}^{coh}$, the mCBD model is able to improve the forecasting performance of the CBD model in most cases (for all the years within the look-forward window except for those from 1991 to 1999). We can thus infer that the mCBD model predicts better than the CBD model in 61% of the calendar years we examine.

We define the magnitudes of the forecasting errors of the CBD and the mCBD models, respectively, as $\psi_{x,t}^{coh}$ and $\varsigma_{x,t}^{coh}$; compared to the other case studies, within this application the age is not fixed, but it implicitly moves forward year by year.

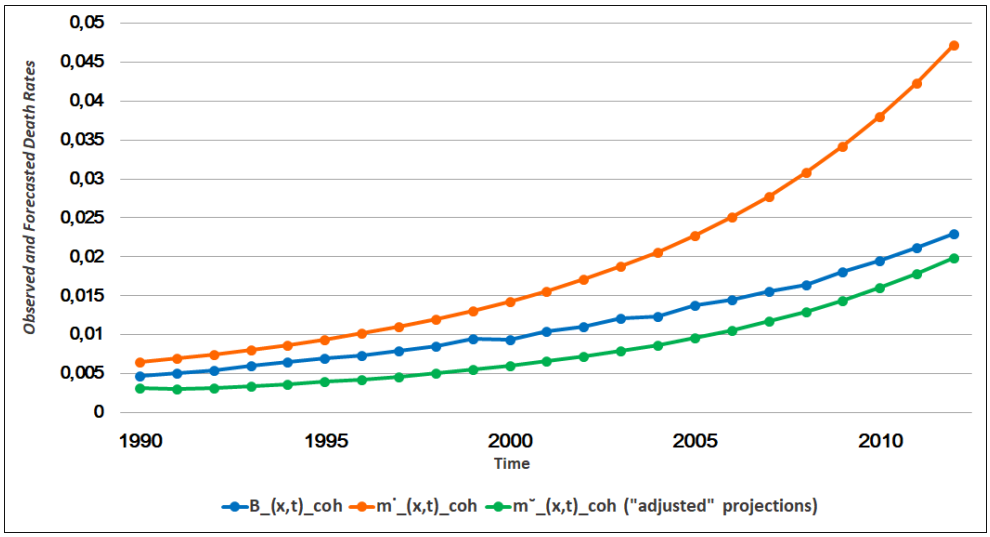


Figure 5.3: Observed and forecasted death rates for the generation of Italian males born in 1940 over the look-forward window (1990,2012). Observed death rates: blue line; CBD forecasted death rates: orange line; mCBD forecasted death rates: green line.

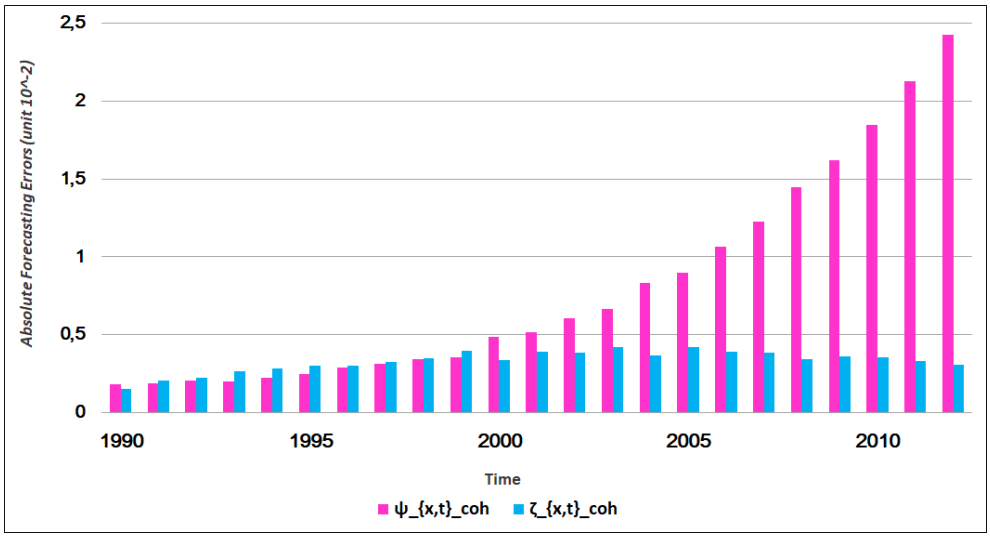


Figure 5.4: Italian mortality data. Generation of males born in 1940. Absolute forecasting errors of the CBD model and the mCBD model (unit 10^{-2}) over the fixed look-back window (1990,2012). Magenta bars: absolute forecasting errors of the CBD model; cyan bars: absolute forecasting errors of the mCBD model

In Figure 5.4, we display the absolute forecasting errors of the mCBD model $\zeta^{coh}_{x,t}$ (the cyan bars) and those of the CBD model $\psi^{coh}_{x,t}$ (the magenta bars), over the look-forward window $[1990, 2012]$ (unit 10^{-2}). Confirming the evidences provided in Figure 5.3, $\zeta^{coh}_{x,t}$ turns out to be smaller than $\psi_{x,t}$ in 61% of calendar years. While $\zeta^{coh}_{x,t}$ is more stable over the look-forward window, $\psi^{coh}_{x,t}$ begins to remarkably increase starting from 2000. Therefore, the further in time we move, within the look-forward window, the larger the gain in the predictive accuracy provided by the mCBD model is, with respect to the CBD model.

In Table 5.5, we report the measures of the average forecasting error and the Improvement Rate.

Table 5.5: Italian mortality data. Generation of males born in 1940. Measures of the average forecasting error (unit 10^{-2}) and Improvement Rate

Cohort	RMSE		*Impr.Rate*	MAE		*Impr.Rate*
	ψ^*_{coh}	ζ^*_{coh}	IR^*_{coh}	$\bar{\psi}_{coh}$	$\bar{\zeta}_{coh}$	$\bar{IR}_{coh}$
1940	1.04	0.34	*-67.59%*	0.80	0.33	*-58.64%*

NOTE: ψ^*_{coh} and $\bar{\psi}_{coh}$ are, respectively the RMSE and the MAE of the CBD model; ζ^*_{coh} and $\bar{\zeta}_{coh}$ are, respectively, the RMSE and the MAE of the mCBD model; IR^*_{coh}: Improvement Rate when the RMSE is concerned; $\bar{IR}_{coh}$: Improvement Rate when the MAE is concerned

Before explaining our results, we remind that, compared to the MAE, in the RMSE's computation each individual forecasting error affects the total error (namely the sum of all the individual errors) in proportion to its square, rather than just its magnitude, so that large errors have a stronger influence on the total square error than small ones. Furthermore, the RMSE increases with the variance of the frequency distribution of the error magnitudes.

The mCBD model is characterized by very close values of the RMSE, ζ^*_{coh}, and the MAE, $\bar{\zeta}_{coh}$. In line with what stated above, this occurs since the distribution of the absolute forecasting errors of the mCBD model, $\zeta^{coh}_{x,t}$, has a low standard deviation (6.9e-04), turning out to be 90% smaller than the standard deviation (68e-04) of the corresponding errors of the CBD model, $\psi^{coh}_{x,t}$. Graphical evidences of the different variability of the two distributions can be found in Figure 5.4.

Both ζ^*_{coh} and $\bar{\zeta}_{coh}$ are smaller than ψ^*_{coh} and $\bar{\psi}_{coh}$, respectively, this meaning

Table 5.6: Italian mortality data. Generation of males born in 1940. Bayesian Information Criterion (BIC) of the CBD and the mCBD model

Cohort	$BICsa_{coh}$	$BICm_{coh}$	ΔBIC_{coh}	$ChangeRate_{coh}$
1940	-203.96	-246.38	42.42	*-20.80%*

NOTE: $BICsa_{coh}$ is the Bayesian Information Criterion of the CBD model; $BICm_{coh}$ is the Bayesian Information Criterion of the mCBD model; ΔBIC_{coh} is the first difference between $BICsa_{coh}$ and $BICm_{coh}$; $Changerate_{coh}$ represents the percent difference between $BICsa_{coh}$ and $BICm_{coh}$

that the mCBD model is more accurate in predicting future death rates. Indeed, the mCBD model reduces by 67.59% and by 58.64%, respectively, the RMSE and the MAE of the CBD model (see IR^*_{coh} and $\bar{IR}_{coh}$ in Table 5.5). Therefore, the mCBD model provides the largest improvement in the average forecasting performance of the CBD model when the RMSE is concerned.

In Table 5.6, we report the Bayesian Information Criterion related to the CBD model ($BICsa_{coh}$), the one related to the mCBD ($BICm_{coh}$); in addition, we provide the first and the percent difference between them (ΔBIC_{coh} and $Changerate_{coh}$, respectively). $BICm_{coh}$ is smaller than $BICsa_{coh}$ and the $Changerate_{coh}$ is indeed negative; furthermore, $\Delta BIC_{coh} > 10$, denoting a decisive evidence in favor of the mCBD model. We can thus infer that, when resorting to the mCBD model, the gained accuracy by adding the three parameters of the CIR process compensates for the added complexity of the methodology.

5.3 Further Developments for the Research

- The methodological approach presented in Chapter 3 could be applied to several mortality projection models, in order to assess to which extent and how it enhances the predictive accuracy of such models, thus gaining more insight behind its corrective capacity.
- In Chapter 4, we have modeled the dynamics of the rv $Y_{\ddot{x},t}$ over time, for different ages, independently. A further advance could consist in modelling the entire surface of ratios $Y_{\ddot{x},t}$, in order to capture possible

age correlations between the "errors" of the existing mortality model we aim to correct.

- In addition, with respect to the same methodology, we could work for increasing the frequency of mortality data, by using, for instance, techniques such as the interpolation or smoothing.
- A further study could concern the dependence between mortality and interest rate risks in the extremes. We could exploit the methodological approach we have designed for improving the predictive accuracy of the existing mortality models for obtaining artificial data. A similar approach, within the same mathematical framework, that one of the affine processes, could be used for getting artificial data for interest rates. We could then explore *a posteriori* the possibility that a relationship between the mortality and interest rates projected patterns can exist.

Chapter 6

Conclusion

In this Thesis, we have addressed the modelling of stochastic mortality, a key issue for life insurance, pension funds, public policy and fiscal planning. Indeed, the prospective increase of longevity can be an advantage for individuals, but represents also a relevant social achievement. The stability and consistency of social welfare systems are put in danger worldwide due to the combined phenomenon of the progressive increase in life expectancy, along with the reduction of birth-rates in industrialized Countries. This phenomenon needs to be interpreted in the context of the connected world in which we live, where the multiple networks arising from the globalization, the Internet communication and the global economic development propagate any event in a very short time, making risks more complex. Due to their very nature, insurance and reinsurance deal with several risks on their balance sheet and, in order to determine the total risk of a portfolio, they need to establish the rules for aggregating the various risks that compose it. As stressed in Biffis et al. (2010), the introduction of market-consistent accounting (see e.g. IASB (2004)) and risk-based solvency requirements (see e.g. EIOPA (2009)) has called for the integration of mortality risk analysis into stochastic valuation models; moreover mortality-linked securities have attracted the interest of capital market investors, who in turn demand transparent tools to price demographic and financial risks in an integrated fashion. Accordingly, a common mathematical framework for modelling the dynamics of stochastic mortality and interest rates, namely for considering the changes in financial and demographic conditions over time, is needed.

Our research is grounded in this context and focused on the class of the affine processes, that have been used in a wide range of applications in financial and actuarial sciences, thanks to their computational tractability and flexibility. Within the class of the affine processes, we have proposed two contributions to the stochastic representation of mortality.

1. **Using affine models for fitting historical longevity trends**.

   In Chapter 2, by using two different specifications of the Feller process, with two and three parameters, we have modelled the stochastic mortality intensity of a given generation, the one of males born in 1940 over the time horizon [1980,2011], chosen by way of an example. We have performed such a study for ten different developed Countries, in order to see if cross-Country longevity common trends could be detected.

   Main results:

   - overall, our results showed very low RMSEs, indicating a good fit to the data extracted from the mortality tables for all the Countries under study;
   - the calibration gave rise to optimal parameters that were close to each other for all the Countries, for both the specifications of the Feller process; therefore, for the generation under study, the ten developed Countries showed similar patterns of mortality dynamics over the examined time horizon;
   - the two Feller models provided close values of the AIC and the BIC for all the Countries; on the basis of the statistical principle of parsimony, we have thus concluded that among the two competing models, the one with the fewer parameters should be preferred.

2. **Using affine models for improving the forecast of longevity**

   In Chapter 3, we have designed a multiplicative affine model for describing the evolution of the mortality phenomenon.

   Technology:

   - tracking, over an observation time horizon, a measure of the fitting error of any existing mortality projection model, such as the ratio between the observed central death rates and the corresponding

fitted values obtained as outputs of the chosen existing mortality model, and model this ratio, dynamically, by means of the Cox-Ingersoll-Ross (CIR) stochastic process, that provides a return to the mean dynamics for this quantity;

- building the multiplicative affine model for forecasting mortality, that is based on the deterministic forecast (baseline) provided by the existing mortality projection model and the CIR corrective factor with properly estimated parameters. The construction method of the affine model is quite general, since it can exploit various mortality models as long as their errors behave in a way similar to the CIR process;
- testing the capability of the so structured affine model to enhance the predictive accuracy of the existing mortality projection model providing the baseline, making its behavior closer to the real trend of the mortality phenomenon, by using out-of-sample validation methods. Besides the usual static method, we have developed a dynamic one that allows us to catch the change in behavior of the underlying data.

In Chapter 4, we have provided a numerical implementation of our affine model. For such a numerical application we have chosen the Cairns-Blake-Dowd (or M5) model as the existing mortality model giving the deterministic baseline, that, combined with the CIR stochastic factor, gives rise to a new model, called mCBD. Using the Italian females mortality data, for fixed ages, and implementing the backtesting procedure, over both a static time horizon and fixed-length windows rolling one-year ahead through time, we have empirically tested the performance of the CBD model and the mCBD models.

Main results:

- on the basis of average measures of forecasting errors and information criteria, the mCBD model shows itself to provide better results, in terms of predictive accuracy, than the CBD one;
- combining the CBD model with a CIR process lets us obtain good results exploiting only three additional parameters compared to the CBD model. Even when correcting the results for the addi-

tional parameters, the BIC shows that the gained accuracy compensates the additional complexity of the model;

- the mCBD model shows a stronger potential to gain accuracy in the long-run than the CBD model when a rolling windows analysis (dynamic approach) is performed; the mCBD affine model thus proves to be suitable for the pricing of longevity bonds and the planning for retirement pension funds;
- as explained in Chapter 5, the construction technology and the backtesting method of our affine model pave the way for a wide range of further applications in the actuarial field.

To conclude, in this Thesis we have explored and tested the properties and capabilities of some affine models in fitting and forecasting mortality data both by themselves and as dynamic driving processes multiplying a deterministic baseline. Combining models and mixing techniques have proved to give satisfactory results and showed a concrete potential to bring the research forward. Our future research is thus oriented to use approaches that combine Monte Carlo simulations and benefit from the synergy between different techniques.

Appendix A

An insight into mortality dynamics and effects

A.1 What We Know About Mortality So Far

Mortality has proven to be a dynamic process. Although all estimates of mortality or longevity dating back to the pre-industrial period (roughly, before 1750) should be vievew with caution because of their inaccuracy, in order to have just a rough perception of how far we have progressed historically in terms of longevity, we can consider that life expectancy at birth among early humans was probably in the 20s (cf. Wilmoth (2000)). According to the World Health Organization (2016), the average life expectancy at birth of the global population in 2015 was 71.4 years, 7.5% bigger than it was in 2000 (66.4 years). Furthermore, if, from 2000 to 2015, the European population has been affected by an increase in the life expectancy at birth in the amount of 6.2% (from 72.3 up to 76.8), in the same time interval the increase in the life span experienced by the African population has been even bigger, in the amount of roughly 19% (from 50.6 up to 60). Based on these demographic evidences, we can infer that, over the course of history, the human life span has remarkably extended.

Whether human longevity will keep on increasing in the future, as it has done in the past, is the subject matter for an ongoing discussion among demographers. On one hand, demographers as Tuljapurkar et al. (2000)

and Oeppen and Vaupel (2002) believe that the length of human life has no natural upper limit; on the other hand, supporters of the opposite stream of research, as Olshansky et al. (2005), argue that the future life expectancy might level off or fall. Lutz et al. (2008) show that the world's population will continue ageing throughout the 21st century; furthermore, the speed of ageing is likely to increase over the next decades and to decelerate in most regions by mid-century.
The different viewpoints emerging from this debate suggest that there is clearly a great deal of uncertainty about future trends in longevity.

Longevity is random and, as proof of that, we summarise some remarks of Cairns et al. (2008b). They stressed some *stylised facts* suitable for describing the mortality dynamics in most developed Countries throughout the twentieth century, when the population of the industrialized world underwent the most striking mortality transition:

- mortality rates have decreased dramatically at all ages;
- for specific ages, the variations of the mortality improvement rate over time have been more significant in some decades than in other ones; furthermore, mortality improvement rates related to different ages have been significantly different. Such evidences are deeply discussed in Currie et al. (2004).
  To gain insight behind the definition and the modelling of improvement rates in the mortality context, we also refer the reader to Richards et al. (2005), Haberman and Renshaw (2012), Mitchell et al. (2013);
- aggregate mortality rates show considerable volatility from one year to the next.

Currie et al. (2004) also constructed confidence bounds for the future development of mortality rates, getting wider when forecasting further and further in time.
On the basis of this and the further evidences provided in this Thesis, we are thus motivated to believe that future mortality improvements cannot be forecasted with any degree of precision (cf. Cairns et al. (2006a)) and that rigorous approaches to mortality forecasts should allow for the *stochastic nature of mortality* (cf. Olivieri and Pitacco (2006)).

A.2 Who is Affected by Increasing Longevity

The need for a sound and accurate *stochastic representation of mortality* has acquired a special significance for policy making in several fields.

The prospective increase of longevity can be an advantage for individuals, but represents also a relevant social achievement. Such a phaenomenon, combined with the reduction of birth rates in the industrialized Countries puts in danger the stability and consistency of social welfare systems worldwide (see Dacorogna and Kratz (2015)).

As stressed in the report of Swiss Re (Singleton et al. (2010)), if life expectancy will continue to increase there will be a shift in 'dependency ratios', so that the older generation's demand for retirement income and certain public services – especially health-related services – will increase substantially. This could deplete government finances, resulting in significant increases in tax and social security contributions on a diminishing working population.
The increase of life span is a growing concern also for employers sponsoring their own employees' retirement income through a defined-benefit plan. Defined-benefit plans promise to pay employees an income on retirement, typically at a level based on their length of service and salary, for the rest of their lives. Each additional year of life expectancy adds about 3–4% to the present value of the liabilities of a typical defined benefit pension fund (International Monetary Fund (2012)). Longer life expectancy means plans may not have made enough financial provision and the costs may have to be covered by employers or through younger generations paying higher contributions.
In some cases the employee has more responsibility in respect with the financial risk associated with providing his/her retirement income. It is indeed increasingly common for employers to offer only defined-contribution schemes, meaning that employees' pensions will depend on their accumulated funds; individuals may bear the risk to outlive their accumulated assets.
Employers and individuals can purchase annuities, thereby transferring the risk for providing retirement income to an insurer. A life annuity is an insurance product providing a regular payment throughout the whole lifetime of the insured in return for a lump-sum payment at the beginning of the contract. The increasing demand for such products is encouraging innovation in the insurance industry, in order to find out new ways of financing longevity

risk.
In its report, Swiss Re emphasises also that no single party has the solution to longevity funding issues and that it will take broader initiatives – through public and private bodies working together – to ensure that living longer remains a benefit to society, rather than a financial burden.

Understanding the drivers of longevity and handling reliable quantitative tools for projecting its future trend is one way to deal with this important issue and for working towards solutions, that are essential to the assessment of the magnitude of the associated risk. Just for having a shallow perception of its relevance, let us consider that, according to Singleton et al. (2010), underestimating life expectancy by just one-year can increase liabilities of insurers and pension funds by up to 5%. For a pension plan with USD 1 billion of assets, an extra USD 50 million would need to be funded.

A.3 Stochastic Mortality Models in Discrete Time

The majority of work in the field of stochastic mortality modelling has focused on *short-rate models*, targeting the *spot mortality rates* or the *spot force of mortality*.

The Lee-Carter (LC) model (Lee and Carter (1992)) represents one of the most influential stochastic methods for forecasting mortality in discrete time. The LC method describes the logarithm of a time series of age-specific death rates as the sum of an age-specific parameter, independent of time, and another component that is product of a time-varying mortality index reflecting the general level of mortality and an age-specific parameter modulating the effect of the mortality index across ages (in other terms, such a parameter represents the age-specific patterns of mortality change). The estimation of the systematic non-random structure underpinning the LC model is based on the first set of singular vectors generated by the single value decomposition (SVD) of an appropriate data matrix. The resulting estimate of the time-varying mortality index is modelled and projected as a stochastic time series by means of standard Box-Jenkins or ARIMA methods implying that changes in the mortality curve at all ages are perfectly correlated. Future

death rates are obtained from the forecast of the general level of mortality and the estimated age-specific parameters. The Lee-Carter model is tipically intended for long-term projections of aggregate mortality indicators like life expectancies, but it is not meant to produce reliable forecasts of series of death rates for a particular age. Furthermore, the LC method implicitly assumes that the random errors are homoskedastic, but this assumption turns out to be unrealistic, since the logarithm of the observed mortality rate is much more variable at older ages than at younger ages, because of the smaller number of deaths that can be observed at old and very old ages.

Brouhns et al. (2002a) and Brouhns et al. (2002b) proposed a method for improving the Lee-Carter approach, using a Poisson random variation for the number of deaths, in place of an additive error term. In such a framework, the force of mortality is assumed to be log-bilinear and the random number of deaths has a Poisson distribution whose parameter is given by the product of the central number of exposed to risk and the force of mortality.

For fixing the potential for inflexibility with respect to age noted by Lee (2000) and Booth et al. (2002), Renshaw and Haberman (2003) investigated whether age differential effects could be incorporated into the LC forecasting methodology by transferring the second additive set of singular vectors from the error structure to the additive systematic non-random structure of the model.

In Cairns et al. (2009), eight stochastic mortality models that decompose mortality improvements into one or more age-, period-, and cohort-related effects were quantitatively compared, in order to explain mortality improvements for males aged 60–89 in England and Wales (EW) and in the United States. The Authors stressed that no single model standed out as being best under all the selection criteria considered. However, different models had different strengths. It is therefore needed to balance up the strengths and weaknesses of each model to form a conclusion, and to some extent it is up to potential users of the models to decide the weights to place on the different criteria.

Appendix B

mCBD Model: Theoretical Remarks Useful for the Calibration and Simulation of the corrective CIR process

B.1 Transition Law of $Y_{\ddot{x},t}$

We apply the theory of the term structure of interest rates explained in Cox et al. (1985) to the variable of our interest, $Y_{\ddot{x},t}$.
The probability density of $Y_{\ddot{x}}$ at time t, conditional on its value at the current time s, is thus given by:

$$f(Y_{\ddot{x},t}, t; Y_{\ddot{x},s}, s) = ce^{-u-v}(\frac{v}{u})^{q/2} I_q(2(uv)^{1/2}) \tag{B.1}$$

where:

$$\begin{aligned} c &\equiv \frac{2\alpha}{\sigma^2(1-e^{-\alpha(t-s)})} \\ u &\equiv cY_{\ddot{x},s}e^{-\alpha(t-s)} \\ v &\equiv cY_{\ddot{x},t} \\ q &\equiv \frac{2\alpha\beta}{\sigma^2} - 1 \end{aligned} \tag{B.2}$$

and $I_q(\bullet)$ is the modified Bessel function of the first kind of order q.

Following (Glasserman (2003), p.122), the transition law of $Y_{\ddot{x},t}$ in (3.2) is:

$$Y_{\ddot{x},t} = \frac{\sigma^2(1-e^{-\alpha(t-s)})}{4\alpha}\chi'^2_d\Big(\frac{4\alpha e^{-\alpha(t-s)}}{\sigma^2(1-e^{-\alpha(t-s)})}Y_{\ddot{x},s}\Big) \tag{B.3}$$

where χ'^2_d denotes a noncentral chi-square random variable with d degrees of freedom (namely $2q+2$ in Cox et al. (1985)):

$$d = \frac{4\beta\alpha}{\sigma^2} \tag{B.4}$$

and non-centrality parameter λ (cf. $2u$ in Cox et al. (1985)):

$$\lambda = \frac{4\alpha e^{-\alpha(t-s)}}{\sigma^2(1-e^{-\alpha(t-s)})}Y_{\ddot{x},s} \tag{B.5}$$

Therefore, denoting $\sigma^2(1-e^{-\alpha(t-s)})/(4\alpha)$ by w, given $Y_{\ddot{x},s}$, $Y_{\ddot{x},t}$ is distributed as w times a noncentral chi-square random variable χ'^2_d.

B.2 Log-likelihood Function of the CIR Process

The likelihood function for $Y_{\ddot{x},t}$ time series with N equally spaced observations $[Y_{\ddot{x},t_i}; i = 1, ..., N]$, with time step Δt, is:

$$L(\theta) = \prod_{i=1}^{N-1} f(Y_{\ddot{x},t_{i+1}}|Y_{\ddot{x},t_i};\theta,\Delta t) \tag{B.6}$$

and the corresponding log-likelihood function is:

$$\ln L(\theta) = \sum_{i=1}^{N-1} \ln f(Y_{\ddot{x},t_{i+1}}|Y_{\ddot{x},t_i};\theta,\Delta t) \tag{B.7}$$

Exploiting (B.1), we obtain the log-likelihood function of the CIR process (3.2):

$$\ln L(\theta) = (N-1)\ln c + \sum_{i=1}^{N-1}[-u_{t_i} - v_{t_{i+1}} + 0.5q\ln\frac{v_{t_{i+1}}}{u_{t_i}} + \ln I_q(2\sqrt{u_{t_i}v_{t_{i+1}}})] \tag{B.8}$$

where $u_{t_i} = cY_{\ddot{x},t_i}e^{-\alpha\Delta t}$ and $v_{t+1} = cY_{\ddot{x},t+1}$.

B.3 Calibration of the CIR Process

After building the log-likelihood function (B.8) in **MATLAB** as in Kladivko (2012), we estimate the parameter vector θ by maximization of such implemented objective function over its parameter space, as follows:

$$\hat{\theta} \equiv (\hat{\alpha}, \hat{\beta}, \hat{\sigma}) = arg \max_{\theta} \ln L(\theta) \tag{B.9}$$

being equivalent to minimization of the negative of the log-likelihood function:

$$\hat{\theta} \equiv (\hat{\alpha}, \hat{\beta}, \hat{\sigma}) = arg \min_{\theta} - \ln L(\theta) \tag{B.10}$$

As already anticipated in Section 4.4.2, we perform the optimization by using the **MATLAB** function *fminsearch*, searching for the minimum of the unconstrained objective function, since we verify a posteriori that the parameters provided by the calibration are all positive and that condition (3.3) is always fulfilled. We find the initial estimate of θ, to be used as the starting point of the optimization, in the way suggested in Kladivko (2012), namely using Ordinary Least Squares on the discretized version of (3.2). For each age under study, the inputs of our optimization routine are:

- the $Y_{\ddot{x},t}$ time series over the look-back window, to which the CIR process has to be calibrated;
- the time step, that is set equal to 1, since we compute the $Y_{\ddot{x},t}$ time series starting from yearly central death rates

B.4 Simulation of the CIR process

As anticipated in Section 4.5.2, in order to simulate the process (3.2), we can resort to equation (B.3) and sample from a noncentral chi-square distribution with d degrees of freedom and non-centrality parameter λ. In particular, as explained in Glasserman (2003), when $d > 1$, sampling of a noncentral chi-square can be performed by sampling of an ordinary chi-square χ^2_{d-1} and of an independent $N(0,1)$ random variable Z; equation (B.11) thus holds:

$$\chi_d'^2 = (Z + \sqrt{\lambda})^2 + \chi_{d-1}^2 \tag{B.11}$$

Within our empirical application, d always turns out to be bigger than 1; therefore, in all cases, we perform simulations of square-root diffusion (3.2) by means of (B.3) and (B.11), adapting to our needs the sequence of steps outlined in (Glasserman (2003), p.124).

List of Figures

List of Tables

Bibliography

Alho, J. M. and Spencer, B. D. (1990). Error model for Official Mortality Forecasts. *Journal of the American Statistical Association*, 85, Issue 411:609–616. Available at `http://www.jstor.org/stable/2289992`.

Apicella, G. and Dacorogna, M. (2016). A Comprehensive Study of Mortality Dynamics in Ten Developed Countries Using the Feller Process. *Proceedings of the XVI Iberian Italian Conference on Financial and Actuarial Mathematics*, pages 7–11. ISBN: 9788861970601. Available at `http://www.unisa.it/uploads/13833/ibit_proceedings.pdf`.

Apicella, G., Dacorogna, M., Di Lorenzo, E., and Sibillo, M. (2017). Using Interest Rate Models to Improve Mortality Forecast. *Available at SSRN: https://ssrn.com/abstract=3070891*.

Apicella, G., Dacorogna, M., Di Lorenzo, E., and Sibillo, M. (2018). Increasing the predictive accuracy of the Lee-Carter model: methodology and preliminary results. *Accepted for presentation at Mathematical and Statistical Methods for Actuarial Sciences and Finance (MAF 2018)*.

Artzner, P. and Delbaen, F. (1995). Default risk insurance and incomplete markets. *Mathematical Finance*, 5(3):187–195.

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

Biffis, E. (2005). Affine Processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37:443–468.

Biffis, E., Denuit, M., and Devolder, P. (2010). Stochastic mortality under measure changes. *Scandinavian Actuarial Journal*, 4:283–311.

Biffis, E. and Millossovich, P. (2006). A bidimensional approach to mortality risk. *Decisions in Economics and Finance*, 29, Issue 2:71–94.

Björk, T. (1998). *Arbitrage Theory in Continuous Time*. Oxford University Press, New York.

Booth, H., Maindonald, J., and Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, 56:325–336.

Booth, H. and Tickle, L. (2008). Mortality Modelling and Forecasting: a Review of Methods. *Annals of Actuarial Science*, 3, Issue 1–2:3–43. Available at `https://doi.org/10.1017/S1748499500000440`.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, 71:791–799. Available at `http://www.jstor.org/stable/2286841`.

Brémaud, P. (1981). *Point Processes and Queues-Martingale Dynamics*. Springer-Verlag, New York.

Brouhns, N., Denuit, M., and Vermunt, J. K. (2002a). Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 2:105–130.

Brouhns, N., Denuit, M., and Vermunt, J. K. (2002b). A Poisson log-bilinear approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, 31(3):373–393.

Bruckner, T. A., Noymer, A., and Catalano, R. A. (2013). Life Expectancy during the Great Depression in Eleven European Countries. *Population and Development Review*, 39, Issue 1:57–74.

Bürgi, R., Dacorogna, M. M., and Iles, R. (2008). Risk Aggregation, Dependence Structure and Diversification Benefit. Chap. 12, pages 265-306, in "Stress testing for financial institutions", edited by Daniel Rösch and Harald Scheule, Riskbooks, Incisive Media, London.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods*

and Research, 33(2):261–304. available at http://dx.doi.org/10.1177/0049124104268644.

Burns, A., Mensbrugghe, D. v. d., and Timmer, H. (2008). Evaluating the economic consequences of avian influenza. *Preprint of the World Bank. Washington, DC*. Available at http://siteresources.worldbank.org/EXTAVIANFLU/Resources/EvaluatingAHIeconomics_2008.pdf.

Cairns, A., Blake, D., and Dowd, K. (2008a). The Cairns-Blake-Dowd Model. Available at https://www.actuaries.org.uk/documents/cairns-blake-dowd-model-handouts.

Cairns, A. J. C., Blake, D., and Dowd, K. (2008b). Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, 2008, 2–3:79–113. DOI: 10.1080/03461230802173608.

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

Cairns, A. J. G., Blake, D., and Dowd, K. (2006b). A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *Journal of Risk and Insurance*, 73(4):355–367.

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., and Balevich, I. (2009). A Quantitative Comparison of Stochastic Mortality Models using Data from England and Wales and the United States. *North American Actuarial Journal*, 13, Issue 1. DOI:10.2139/ssrn.1340389.

Chen, H., MacMinn, R., and Sun, T. (2015). Multi-population mortality models: A factor copula approach. *Insurance: Mathematics and Economics*, 63:135–146.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385–407. Available at http://www.jstor.org/stable/1911242.

Currie, I. D. (2014). On fitting generalized linear and non-linear models of mortality. *Scandinavian Actuarial Journal*, pages 356–383. 10.1080/03461238.2014.928230.

Currie, I. D., Durban, M., and Eilers, P. H. C. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling*, 4(4):279–298.

Dacorogna, M. and Apicella, G. (2016). A General Framework for Modeling Mortality to Better Estimate its Relationship to Interest Rate Risks. *SCOR Papers no.39*. Available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2888794`.

Dacorogna, M. and Cadena, M. (2015). Exploring the dependence between mortality and market risks. *SCOR Paper no 33*. Available at SSRN: `https://ssrn.com/abstract=2730520`.

Dacorogna, M., Gençay, R., Müller, A., Olsen, R., and Pictet, O. (2001). *An introduction to high frequency finance*. sec 7.4.2, Academic Press, San Diego CA.

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

Dahl, M. (2004). Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics*, 35:113–136.

Dahl, M. and Møller, T. (2006). Valuation and hedging of life insurance liabilities with systematic mortality risk. *Insurance: Mathematics and Economics*, 39:193–217.

Dai, Q. and Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55(5):1943–1978.

Danesi, I. L., Haberman, S., and Millossovich, P. (2015). Forecasting mortality in subpopulations using Lee - Carter type models: A comparison. *Insurance: Mathematics and Economics*, 62:151–161.

Debón, A., Martínez-Ruiz, F., and Montes, F. (2010). A geostatistical approach for dynamic life tables: The effect of mortality on remaining lifetime and annuities. *Insurance: Mathematics and Economics*, 47:327–336.

Deelstra, G., Grasselli, M., and Van Weverberg, C. (2016). The role of the dependence between mortality and interest rates when pricing Guaranteed Annuity Options. *Insurance: Mathematics and Economics*, 71:205–219.

DellaVigna, S. and Pollet, J. M. (2007). Demographics and industry returns. *American Economic Review*, 97(5):1167–1702.

Di Lorenzo, E., Sibillo, M., and Tessitore, G. (2006). A Stochastic Proportional Hazard Model for the Force of Mortality. *Journal of Forecasting*, 25:529–536.

Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M. (2013). Backtesting Stochastic Mortality Models: an ex post evaluation of multiperiod-ahead density forecasts. *North American Actuarial Journal*, 14(3):281–298. DOI:10.1080/10920277.2010.10597592.

Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, third edn.

Duffie, D. and Singleton, K. J. (2003). *Credit Risk*. Princeton University Press.

EIOPA (2009). Solvency II Directive (Directive 2009/138/EC). *European Insurance and Occupational Pensions Authority*. Available at `https://eiopa.europa.eu/regulation-supervision/insurance/solvency-ii`.

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55(2):251–276.

Favero, C. A., Gozluklu, A. E., and Tamoni, A. (2011). Demographic Trends, the Dividend-Price Ratio, and the Predictability of Long-Run Stock Market Returns. *Journal of Financial and Quantitative Analysis*, 46(5):1493–1520.

Favero, C. A., Gozluklu, A. E., and Yang, H. (2016). *IMF Economic Review*, 64, Issue 4:732–776.

Ferreira, A. and De Haan, L. (2015). On the block maxima method in extreme value theory: PWM estimators. *The Annals of Statistics*, 43(1):276–298.

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

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer, ISBN: 978-0-387-21617-1.

Glei, D. A. (updated by Gabriel Borges and Magali Barbieri on October 12, 2015b). About mortality data for Sweden. *Available at* *http://www.mortality.org/hmd/SWE/InputDB/SWEcom.pdf*.

Glei, D. A. (updated by Gabriel Borges on September 27, 2015a). About mortality data for Italy. *Available at* *http://www.mortality.org/hmd/ITA/InputDB/ITAcom.pdf*.

Haberman, S. and Renshaw, A. (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, 48:35–55.

Haberman, S. and Renshaw, A. (2012). Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics*, 50(3):309–333. DOI:10.1016/j.insmatheco.2011.11.005.

Hatzopoulos, P. and Haberman, S. (2013). Common mortality modeling and coherent forecasts. An empirical analysis of worldwide mortality data. *Insurance: Mathematics and Economics*, 52:320–337.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60(1):77–105.

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

Jeanblanc, M. and Rutkowski, M. (2000). Modeling of default risk: an overview. *Mathematical Finance: Theory and Practice*, pages 171–269. In J. Yong and R. Cont (eds.).

Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430):773–795. Available at `http://www.jstor.org/stable/2291091`.

Kladivko, K. (2012). Maximum likelihood estimation of the Cox-Ingersoll-Ross process: the MATLAB implementation. *Available at* *https://it.mathworks.com/matlabcentral/fileexchange/37297-maximum-likelihood-estimation-of-the-cox-ingersoll-ross*.

Knight, F. H. (1921). Risk, Uncertainty and Profit. *art Schaffner and Marx, Boston*.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research*, 2(2–3):99–120.

Lee, R. D. (2000). The Lee-Method for forecasting mortality, with various extentions and applications. *North American Actuarial Journal*, 4(1):80–93.

Lee, R. D. and Carter, L. R. (1992). Modelling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(14):659–675.

Li, N. and Lee, R. (2005). Coherent mortality forecasts for a group of populations: an extension of the Lee – Carter method. *Demography*, 42(3):575–594.

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

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

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

Lutz, W., Sanderson, W., and Scherbov, S. (2008). The coming acceleration of global population ageing. *Nature*, 451:716–719.

Maghsoodi, Y. (1996). Solution of the extended CIR term structure and bond option valuation. *Mathematical Finance*, 6:108–109.

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

Miltersen, K. R. and Persson, S. A. (2005). Is Mortality Dead? Stochastic Forward Force of Mortality Determined by No Arbitrage. *Working Paper, University of Bergen*.

Mitchell, D., Brockett, P., Mendoza-Arriaga, R., and Muthuraman, K.

(2013). Modeling and Forecasting Mortality Rates. *Insurance: Mathematics and Economics*, 52, Issue 2:275–285.

Moser, K., Shkolnikov, V., and Leon, D. A. (2005). World mortality 1950-2000: divergence replaces convergence from the late 1980s. *Bulletin of the World Health Organization*, 83, Issue 3:202–209.

Oeppen, J. and Vaupel, J. W. (2002). Broken limits to life expectancy. *Science*, 296:1029–1031.

Ogburn, W. F. and Thomas, D. S. (1922). The influence of the business cycle on certain social conditions. *Journal of the American Statistical Association*, 18(139):324–340.

Olivieri, A. and Pitacco, E. (2006). Stochastic mortality in life insurance. *Statistica Applicata*, 18(4).

Olivieri, A. and Pitacco, E. (2010). *Introduction to insurance mathematics. Technical and financial features of risk transfers*. Springer, ISBN 978-3-642-16028-8.

Olivieri, A. and Pitacco, E. (2012). Life tables in actuarial models: from the deterministic setting to a Bayesian approach. *AStA Advances in Statistical Analysis*, 96:127–153. DOI 10.1007/s10182-011-0177-y.

Olshansky, S. J., Passaro, D., Hershaw, R., Layden, J., Carnes, B. A., Brody, J., Hayflick, L., Butler, R. N., Allison, D. B., and Ludwig, D. S. (2005). A potential decline in life expectancy in the United States in the 21st century. *New England Journal of Medicine*, 352:1103–1110.

Park, Y., Choi, J. W., and Kim, H. (2006). Forecasting Cause-Age Specific Mortality Using Two Random Processes. *Journal of the American Statistical Association*, 101(474):472–483. Available at `http://www.jstor.org/stable/27590710`.

Pitacco, E. (2004). Survival models in a dynamic context: a survey. *Insurance: Mathematics and Economics*, 35:279–298.

Pitacco, E. (2007a). *Matematica e tecnica attuariale delle assicurazioni sulla durata di vita*. Lint, ISBN 978-88-8190-182-1.

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

Pitacco, E., Denuit, M., Haberman, S., and Olivieri, A. (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press, ISBN:9780199547272.

Preston, S. H. (1975). The Changing Relation between Mortality and Level of Economic Development. *Population Studies*, 29(2):231–248.

Raftery, A. E. (1999). Bayes Factors and BIC - Comment on "A critique of the Bayesian Information Criterion for Model Selection". *Sociological Methods and Research*, 27(3):411–427.

Renshaw, A. E. and Haberman, S. (2003). Lee-Carter mortality forecasting with age specific enhancement. *Insurance: Mathematics and Economics*, 33(2):255–272.

Ribeiro, R. and di Pietro, V. (2009). Longevity risk and portfolio allocation. *JP Morgan, Investment Strategies no.57*. Available at `https://www.jpmorgan.com/`.

Richards, S. J., Kirkby, J. G., and Currie, I. (2005). The Importance of Year of Birth in Two-Dimensional Mortality Data. *British Actuarial Journal*, 12, Issue 01:5–38. DOI: 10.1017/S1357321700004682.

Schrager, D. F. (2006). Affine stochastic mortality. *Insurance: Mathematics and Economics*, 38:81–97.

Singleton, M., Thomsen, T., and Yiasoumi, C. (2010). A short guide to longer lives: Longevity funding issues and potential solutions. *Swiss Re. Available at `http://media.swissre.com/documents/Longer_lives.pdf`*.

Tapia Granados, J. and Diez Roux, A. V. (2009). Life and death during the Great Depression. *Proceedings of the National Academy of Sciences USA*, 106(41):17290–17295.

Tapia Granados, J. A. (2005). Increasing mortality during the expansions of the US economy, 1900-1996. *International Journal of Epidemiology*, 34, Issue 6:1194–1202.

Thiele, T. N. (1872). On a mathematical formula to express the rate of mortality throughout life. *Journal of the Institute of Actuaries*, 16:313–329.

Tuljapurkar, S., Li, N., and Boe, C. (2000). A universal pattern of mortality decline in G7 countries. *Nature*, 405:789–792.

Basel Committee on Banking Supervision (August 2013). Longevity risk transfer markets: market structure, growt drivers,and impediments, and potential risks. *Available at* *http://www.bis.org/publ/joint34.pdf*.

European Commission (2012). Green Paper: "Confronting demographic change: a new solidarity between the generations". *Available at* *http://eur-lex.europa.eu/legal-content/IT/TXT/?uri=URISERV%3Ac10128*.

Human Mortality Database (2016). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). *Available at* *http://www.mortality.org* *or* *http://www.humanmortality.de*. (data downloaded on October 20, 2016).

IASB (2001). Draft statement of principles. *International Accounting Standards Board*. Available at `www.iasb.org.uk`.

IASB (2004). International Financial Reporting Standard n. 4. *International Accounting Standards Board*. Available at `https://www.iasplus.com/en/standards/ifrs/ifrs4`.

International Monetary Fund (2012). Global Financial Stability Report. *World Economic and Financial Surveys, April*.

United Nations (2002). Report of the Second World Assembly on Ageing, Madrid, 8-12 April 2002. *Available at* *https://documents-dds-ny.un.org/doc/UNDOC/GEN/N02/397/51/PDF/N0239751.pdf?OpenElement*.

World Health Organization (2016). Global Health Observatory (GHO) data. *Available at* *http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends/en/*.

Vasicek, O. A. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188.

Villegas, A., Millossovich, P., and Kaishev, V. (2016a). Package 'StMoMo'.

Available at *https://cran.r-project.org/web/packages/StMoMo/StMoMo.pdf*.

Villegas, A., Millossovich, P., and Kaishev, V. (2016b). StMoMo: An R Package for Stochastic Mortality Modelling. *Available at* *https://cran.r-project.org/web/packages/StMoMo/vignettes/StMoMoVignette.pdf*.

Weakliem, D. L. (1999). A Critique of the Bayesian Information Criterion for Model Selection. *Sociological Methods and Research*, 27:359–397.

Willmott, C. J. and Matsuura, K. (2005). Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in assessing average model performance. *Climate Research*, 30(1):79–82. DOI:10.3354/cr030079.

Wilmoth, J. R. (2000). Demography of longevity: past, present, and future trends. *Experimental Gerontology*, 35:1111–1129.

Wilmoth, J. R., Andreev, K., Jdanov, D., and Glei, D. (2007). Methods Protocol for the Human Mortality Database. With the assistance of Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., Vachon, P. Available at `http://www.mortality.org/Public/Docs/MethodsProtocol.pdf`.

Wilson, C. (2001). On the scale of global demographic convergence 1950–2000. *Population and Development Review*, 27(1):155–172.

Wilson, C. (2011). Understanding Global Demographic Convergence since 1950. *Population and Development Review*, 37, iss.2:375–388.