EAA Series

Jean-Paul Laurent
Ragnar Norberg
Frédéric Planchet *Editors*

Modelling in Life Insurance – A Management Perspective

European actuarial academy
developing actuaries

Springer

EAA Series

EAA series is successor of the EAA Lecture Notes and supported by the European Actuarial Academy (EAA GmbH), founded on the 29 August, 2005 in Cologne (Germany) by the Actuarial Associations of Austria, Germany, the Netherlands and Switzerland. EAA offers actuarial education including examination, permanent education for certified actuaries and consulting on actuarial education.

actuarial-academy.com

More information about this series at http://www.springer.com/series/7879

Jean-Paul Laurent · Ragnar Norberg
Frédéric Planchet
Editors

Modelling in Life Insurance – A Management Perspective

Springer

Editors
Jean-Paul Laurent
Université Paris 1 Panthéon-Sorbonne
Paris
France

Ragnar Norberg
Univ Lyon – Université Lyon 1 – ISFA – LSAF
Lyon
France

Frédéric Planchet
Univ Lyon – Université Lyon 1 – ISFA – LSAF
Lyon
France

ISSN 1869-6929 ISSN 1869-6937 (electronic)
EAA Series
ISBN 978-3-319-29774-3 ISBN 978-3-319-29776-7 (eBook)
DOI 10.1007/978-3-319-29776-7

Library of Congress Control Number: 2016933465

Mathematics Subject Classification (2010): 00A06, 00A09

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Foreword

I am on the side of models.

Always have been. By personal inclination and actuarial education I love going deep into the mathematics and intricacies of models. For the intellectual challenge they offer. For knowing what's under the hood. However, I often have not had enough time to do so… I recognize, because it does indeed take time.

In my different capacities as manager and now as a CEO of an insurance company, I regularly come across models, especially when a decision is at stake. I always felt that models could and should help me, that they have something useful to tell to me. But I have to admit that it is not that easy to decipher their words, however hard you try to listen! If you have already experienced using models in the perspective of steering an insurance company, you know well that there are often hot debates on how far the say of models should prevail in decision-making: there are pros and cons; there are experts and ignorants… The good use of a model's advice is quite a tricky issue. More often than not, as the final decision-maker, you feel alone with your trust toward or your defiance against the model's outputs.

To break free from this loneliness, you absolutely need to read this book because academics and experts on models are here convened to express their opinions on the multiple issues of developing, implementing and using models from both theoretical and practical viewpoints… and because they did so in a very accessible way, closer to the b1 language than to the standard vocabulary of a doctorate thesis. Believe me, that feat is the trademark of the best as, to quote Nicolas Boileau in *L'art poétique* (1674): "*whate'er is well conceived is clearly said. And the words to say it flow easily.*"

This book will be instrumental in helping you to make informed decisions based on a thorough and comprehensive understanding of what models bring to the decision table and what they do not. Indeed, you, as managers, should be eager to Know Your Models. The acronym KYM should nowadays be placed on all life insurance companies just like the sentence "*Know thyself*" was carved into the temple of Delphi, house of the Pythia who predicted the future. So enter now the house of models and receive the sought-for knowledge!

For those of you who need a map of the house, go straight to the introduction by Jean-Paul Laurent, Ragnar Norberg and Frederic Planchet. There, chapter by chapter, are unveiled the issues addressed by the different authors of the book, the solutions they envisioned and the road they propose to go on to fix them. At the end of this introduction, I am sure you could not help but be surprised by so many complex unresolved questions touching all areas of modelling. Indeed, before opening this book, you may have first thought that modelling is a known and familiar territory; now you know that there is still a lot of research to do!

Well, quite early, I was convinced that outside academic research, there was no salvation: as the Directive Solvency II turned out to give a central role to models, I sensed that perhaps financial models were taking on too much of the regulation load without being fully adapted to the specificities of the insurance business and that the consequences of massively using them were not sufficiently assessed. I felt both a sense of danger (are we mad to put the solvency of our insurance companies in the hands of market-consistent risk-neutral models?) mixed with a sense of opportunities (if flaws exist, I am sure we could find a way to fix them). That is the reason why BNP Paribas Cardif decided to give its full support to the BNP Paribas Cardif research chair hosted by the ISFA actuarial school.

This book is a five-year research milestone of the chair. It gives the flavour, if not the cream, of the great work accomplished during all these years under the helm of, first, V. Maume-Deschamps, and now C. Robert. I once had the opportunity to attend the Board of the chair. I confess that it was my pleasure and a source of great intellectual inspiration to debate with such an enthusiastic and innovative bunch of researchers. The number of publications they produced during this time span (one hundred) is a clear testimony to the quantity and the quality of their collective work.

I also would like to pay tribute to J.-P. Laurent who, as the head of the editorial committee, did a tremendous job in gathering complementary chapters which, overall, gives a comprehensive impression of where the models stand, where they come from and where they might go to. You will find advice and best practices, warnings and food for further thought. You can read the book from front to back if you want to: in its kind, it is a page-turner. But if you would rather skip around and just read the topics that interest you, that is fine, too.

Whatever your reading approaches, I hope that you will take as much pleasure in reading it as I did. I am also ready to bet that you will find some unexpected treasure hidden in these pages and that you will be taught invaluable lessons.

So now… will you join me on the side of models?

Pierre de Villeneuve
CEO of BNP Paribas Cardif

Contents

Part I Life Insurance Context

1 Paradigms in Life Insurance 3
Ragnar Norberg

2 About Market Consistent Valuation in Insurance 43
Pierre-E. Thérond

Part II Design and Implementation of Life Insurance Models

3 Cash Flow Projection Models 63
Jean-Paul Felix

4 Economic Scenario Generators 81
Thierry Moudiki and Frédéric Planchet

5 From Internal to ORSA Models 105
Frédéric Planchet and Christian-Yann Robert

6 Building a Model: Practical Implementation 125
Patrice Palsky

Part III Model Validation and Steering Processes

7 Ex-ante Model Validation and Back-Testing 151
Stéphane Loisel and Kati Nisipasu

8 The Threat of Model Risk for Insurance Companies 161
Christian-Yann Robert

9 Meta-Models and Consistency Issues 181
Jean-Paul Laurent

Part IV Models and Business Processes

10 Model Feeding and Data Quality 205
Jean-Paul Felix, Nathalie Languillat and Amélie Mourens

11 The Role of Models in Management Decision-Making 223
Bernard Bolle-Reddat and Renaud Dumora

12 Models and Behaviour of Stakeholders 237
David Ingram and Stéphane Loisel

Bibliography 249

Contributors

Bernard Bolle-Reddat BNP Paribas Cardif, Paris, France

Renaud Dumora BNP Paribas Cardif, Paris, France

Jean-Paul Felix BNP Paribas Cardif, Paris, France

David Ingram WILLIS RE, New York, USA

Nathalie Languillat BNP Paribas Cardif, Paris, France

Jean-Paul Laurent Université Paris 1 Panthéon-Sorbonne, Paris, France

Stéphane Loisel Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France

Thierry Moudiki Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France

Amélie Mourens BNP Paribas Cardif, Paris, France

Kati Nisipasu SCOR, Paris, France

Ragnar Norberg Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France

Patrice Palsky Fractales, Paris, France

Frédéric Planchet Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France; PRIM'ACT, Paris, France

Christian-Yann Robert Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France

Pierre-E. Thérond Galea & Associés, Paris, France; Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France

Introduction

The insurance industry of today is science-based and arguably more so than any other high-tech industry, at least when measured by the sheer volume and diversity of the body of theoretical models employed in its operations. There are several explanations for this. Firstly, while manufacturing of physical goods typically involves processes governed by a small set of "exact" natural laws, insurance is all about exchange and trade—hence measurement and management—of various forms of risk in a complex and ever-changing social, technological, regulatory and competitive environment: in any line of insurance there exist many candidate models, none of which is the uncontested true one. Secondly, the advent of modern financial mathematics, new accounting standards, and endeavours to transfer insurance risk to the markets, have widened the scope of actuarial science to include, in addition to the traditional studies of insurance liabilities, studies of asset-liability management, securitization, hedging and investment strategies. Thirdly and finally, new insurance regulations oblige insurers to model risk at all levels from micro to macro and to analyse business objectives and strategies in this theoretical framework.

Against this background the present book discusses how to build models, how to calibrate them, how to feed them with reliable data, and how to use them in decision-making, including warnings against misuses and pitfalls. It goes beyond the commonplace by considering these issues in a management perspective under conditions set by new regulatory regimes. Because models have become of such importance for insurance managers, regulators and investors, how to deal with models goes far beyond actuarial or financial expertise. In this book, senior executives, experts from the insurance industry and academics join forces to bridge the gap between technicalities and operational management. The book aims at a readership that extends beyond the specialists, and places its subject matter in a broad perspective, capturing the views of makers, maintainers, and end users of models at all levels in insurance companies and regulatory bodies. Unlike mainstream texts, it presents the objectives and views of the top management of an insurance enterprise in the contemporary regulatory and competitive environment and also a discussion of the very design of life insurance schemes and policies in a historical perspective.

After the BNP Paribas Cardif research chair was launched in 2010, a number of technical meetings were held, involving researchers from the ISFA actuarial school, University of Lyon and actuarial, ALM and finance executives from the chair's sponsor. They have been a discussion forum where theory and applications were sitting together in the same room. In "Les Pensées", Blaise Pascal made a core and fruitful distinction between "esprit de finesse" (subtle mind, where subtle is often associated with acute or intuitive) and "esprit de géométrie" (geometrical mind). Besides being a gifted mathematician, with breakthrough insights in probability theory and automated processes (the "pascaline" is likely to be the first mechanical arithmetic device), Pascal stated that both forms of intelligence, geometrical and subtle had to work together for greater efficiency. A long time after Pascal, Alan Turing's electromechanical device led the path to make sense of unsorted bunches of big data. Not only paving the way for modern computer science and programming, Turing questioned the differences between artificial and human intelligence. Despite being fed by human expertise, there is quite a lot of geometry in life insurance models, but not that much subtlety. In spite of increased complexity, models remain quite constrained by our current knowledge on meaningful generation of scenarios, specification of risk premiums, the modelling of management actions or of customer behaviour. They are helpful in structuring decision processes and the quantification of outputs, but caution is required at every stage, and prescriptions need to be challenged by the key functions (CEO, CFO, CRO, CAO, CCO). While automated car driving is planned for the (almost) near future, we already rely on a number of automated assistance tools, but the driver's experience and control is still required to face unexpected situations. Managing a life insurance business is by far more complex than getting rid of traffic jams. Thus modelling experts and decision-makers are to work together for a long time. In a famous speech given at the 36th economic policy symposium, at Jackson Hole, Andrew Haldane introduced the following parable entitled "the dog and the Frisbee". Catching a Frisbee involves solving an optimal control problem and a good knowledge of mechanics. However, a dog, with just a bit of practical experience and never having sat an exam in a quantitative topic, performs quite well at this exercise (apparently dogs do better than humans when it comes to Frisbees). Haldane intended to point out the dilemma between the increased complexity of models and the quest for simple, meaningful and robust decision rules required to manage risky businesses.

In Chap. 1 of the book, Ragnar Norberg places life insurance and pensions in a contemporary historical perspective and concludes that today's practices and regulation represent a change at the level of a paradigm shift. In Chap. 2, Pierre Thérond goes further into the new paradigm with its emphasis on market consistency in financial analysis, accounting rules, and reserving. Even though insurance liabilities are seldom traded, there has been a major shift towards a fair value approach to the liability side of insurance companies. This led to a massive use of mark-to-model approaches calibrated to tradable instruments whenever possible. This also expands to prospective evaluation of reserves, required when computing risk-based measures in an internal model or in an ORSA framework. Ragnar Norberg and Pierre Thérond both warn about the theoretical and ideological basis of such approaches. Furthermore, an extra-layer of complexity was subsequently introduced to cope with the long-term nature of the life insurance business and mitigate the perverse effects of full fair value approaches. Ultimately, we might have ended up far away from the requirements of insurance stakeholders: meaningful and comparable measures of risk and profitability.

Following this first part dedicated to the life insurance context, the second part focuses on the design and implementation of life insurance models. Jean-Paul Félix shares his practical experience when it comes to forecasting asset and liability cash-flows under various economic scenarios. He reviews the main challenges and outlines three critical fields of assessment:

- Data and parameters are of an appropriate quality;
- The model is consistent with both the reality of the business and the management rules;
- The runtime is fast enough to meet the different deadlines involved in the computation of reserves and solvency requirements.

Eventually, the right balance between these criteria is the key to success.

The next chapter, written by Frédéric Planchet and Thierry Moudiki, deals with the other key ingredient of best estimate computations, the economic scenario generator (ESG). They point out that the work ahead is not so much to locate an ESG that accurately reflects the behaviour of modelled assets, but rather to provide a simple, coherent explainable framework that sustains asset/liability models. Even though, due to Pillar 1 constraints, scenarios need to be made consistent with market views, the authors claim that building one's own scenarios is key to a genuine mastery of risk management. It compels an appraisal of risk exposure, its assessment and the consequences of model choices.

Due to regulatory constraints on internal model validation, a number of insurance companies are likely to rely on standard formulas for a while. Nevertheless, ORSA involves a significant amount of dynamic modelling regarding the entity's coverage ratio. In the chapter dedicated to ORSA, Christian-Yann Robert and Frédéric Planchet describe a less granular description of risks as well as a proper blend of probabilistic approaches and scenario-based approaches. The goal here is to propose models that are simple, flexible, and enable one to obtain satisfactory results that are also consistent with those of Pillar 1. This is not a question of undermining the standard formula, but of being able to explain the modelling differences needed to be closer to the reality of the business.

Last, but not least, Patrice Palsky deals with practical implementation issues. Well-conducted project development, including efficient team work involving financiers, actuaries and IT professionals, are of prime importance. Life insurance models encompass all aspects of risk, thus robust architecture and advanced engineering must be deployed to prevent such huge software from collapsing under its own weight. Dealing with the operational difficulties through many examples, Patrice Palsky states that the waterfall model development cycle commonly used in IT management is unsuited to financial models. He pushes forward to using adaptive development methods, such as Agile methods, to minimize costs and the risk of project failure.

Unsurprisingly, due to the tricky nature of implementing life insurance models, validation, back-testing and assessment of model risks are of the utmost importance for a safe and proper use of models. This is the purpose of the third part of this book. Stéphane Loisel and Ecaterina Nisipasu deal with ex-ante validation and back-testing procedures. In this dedicated chapter, they emphasize the independent review of internal models or how to report the main findings at the most senior level. Given the one-year horizon and 99.5 % confidence level associated with Solvency II requirements, back-testing based on VaR exceptions cannot be conducted. Deriving loss distributions based on the law of large numbers and central limit theorems might also be misleading due to herding of customer behaviour. Besides providing guidance on model validation issues, the authors discuss breakthrough approaches aimed at sequentially testing for changes in actuarial assumptions or parameters. They add to the renewed interest on early warning indicators of financial distress.

In the following chapter, Christian Robert addresses the issue of model risk in life insurance businesses. Models can lead to wrong decisions and potential adverse consequences when they are incorrect or misused. He reviews the different concepts of model risks, as seen by academics, regulators, financial quants and actuaries. This concerns the model design process, the development and production processes and possible misuses of model results. Then, different paradigms for measuring model risk are discussed, such as the Bayesian model averaging, the robustness and "maxmin" approaches. Finally, he deals with approaches to mitigate model risk from a broad management perspective, with an emphasis on solvency issues. It is also noteworthy that too much standardization and uniformity in the design of models could lead to pernicious systemic model risk.

The chapter written by Jean-Paul Laurent deals with the kind of model risk that comes from using different models with same scope but different risk granularity or distributional assumptions. That kind of model multiplicity and potential inconsistencies within a given life insurance company is shown to be inherent to the development process and diverging operational or compliance constraints. From a management perspective, a strong reliance on the skills of actuaries and quantitative modellers is required since operational risks could arise and are difficult to monitor through standard auditing procedures.

Eventually, following this wide review of business issues related to the conceptual framework that underpins model design, practical issues involved in model development and validation, the last part of this book deals with using models from a strategic perspective.

In a chapter entitled "Model Feeding and Data Quality", Amélie Mourens, Nathalie Languillat and Jean-Paul Felix address information systems. The goal of a

Data Quality program is to support an efficient data management process. The authors detail how such a Data Quality project can be enforced, a necessary condition for educated decisions: For those insurers that dare not enter into further thoughts about data management, the risks are high of wrong strategic decisions (in force business management, acquisitions, launching of new businesses), wrong risk management decisions, supervisory sanctions (capital add-ons), and losing confidence of shareholders, partners or customers.

The two following chapters ending the fourth part of this book provide challenging views regarding how decision-makers deal with the surge of models.

Bernard Bolle-Reddat and Renaud Dumora address these issues from a practitioner's viewpoint and share their experience at a major insurance company. First of all, they remind us that insurance undertaking could not be run today without an extensive use of quantitative models. However, they stress the mass crisis implications due to model standardization or pro-cyclical effects of changes of quantitative regulation of solvency risks. When it comes to running a life insurance company, challenging the models is therefore necessary for the sake of achieving sound decisions. The authors claim that modelling cannot be only a mathematical approach to the future. It has to be a smart mix between both mathematical algorithms to simulate the future central scenario and its sensitivity to small variations of the environment, on the one hand, and a set of stress tests decided by the management of the company, at its highest level, on the other hand. Based on practical examples such as strategic asset allocation or discretionary participation to policyholders, they provide management guidelines: departures from allocation suggested by the model need to be documented and explained. On the other hand, changes in the economic environment lead to changes in the setting of the target annual participation rate. Another major point tackled by the authors is the consistency of decisions based on diverse models. They suggest a global model mapping so that every intended use is adequately covered and to show the areas of relevance for each model so that decision-makers know when to trust it. This is a real challenge to decision-makers who need to be familiar with the ins and outs of the models to fully exert their duty of challenging them. For that purpose, BNP Paribas Cardif sets up off-site sessions of the Executive Committee, with a focus on such technical issues.

Based on their previous academic work and on feedback of the above "bunkering sessions" of the Executive Committee, David Ingram and Stéphane Loisel deal with stakeholder's behaviour regarding model risk and model management. They show how endogenous risks and feedbacks loops, well documented since the financial crisis, can be extended to typical insurance risks, such as surrender. In this chapter, they describe behavioural findings regarding risk attitudes and whether they are suitable to different market environments. They set up simple prescriptive rules, in line with Haldane's quest. The chapter concludes by presenting different attitudes towards models that are not directly linked with attitudes with respect to risk. They also point out the role of cultural backgrounds when it comes to trust or challenge models as a decision tool in life insurance businesses.

We hope that readers, either newcomers to the field, experienced model designers or wise life insurance managers, will enjoy going through these chapters and that the reading will stimulate the subtle skills required for a proper use of models in a business context.

Jean-Paul Laurent
Ragnar Norberg
Frédéric Planchet

Part I
Life Insurance Context

Chapter 1
Paradigms in Life Insurance

Ragnar Norberg

Abstract The present chapter describes the developments that the insurance system has undergone over the past four decades and that have been brought to their conclusion in new regulations representing no less than a spectacular paradigm shift. The new paradigm is compared with its immediate predecessor in a perspective guided by timeless questions: what are the objectives of insurance, and how can they be achieved efficiently? And, more specifically, is it all about investments, market valuation, and solvency of businesses? Or has it also got something to do with social security, broad insurance coverage, transparency, and cost awareness? It is argued here that regimes like Solvency II and International Financial Reporting Standards are just our time's response to problems created by our time's practices (and malpractices), and that a rethink may well lead to a partial reinstatement of the preceding paradigm and add to it innovative ways of realizing the very idea of insurance.

1.1 Introduction

1.1.1 Insurance on the Move

The memory of man is short, especially in financial matters. Insurance is no exception, and readers are therefore alerted that the theories and practises presented in this book are just temporary answers to questions that are as eternal as the very concept of insurance. Indeed, the past some four decades have seen insurance and pensions undergo radical changes driven by multiple and interacting forces—regulatory, corporate, economic, demographic, and technological—the total result of which is a transform at the level of paradigm shift.

R. Norberg (✉)
Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France
e-mail: ragnar.norberg@univ-lyon1.fr

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_1

1.1.2 Insurance in the Recent Past

It is with some nostalgia that I recall the Scandinavian insurance system I was trained to serve when studying insurance mathematics at the University of Oslo around 1970. Paramount to everything was the insurance legislation, introduced early in the 20th century and consisting of three parts. Firstly, the Insurance Treaty Act defining the notion of insurable risk and prescribing general terms and conditions of contracts under which such risk could be covered. Secondly, the Insurance Company Act setting out the purpose of insurance and the obligations placed on insurance companies licensed to serve it, instituting regulation by a central government supervisory office, and requiring that each insurance company be represented by a *responsible actuary* approved by and reporting to the supervisory office. Thirdly, a parliamentary bill assigning the University the task of offering an academic programme leading to the degree of *candidatus actuariae*, which set the scientific requirements to the responsible actuary hence to the actuarial profession.

Under that regime the insurance market was a quiet place. The insurance industry was seen as a pillar in the social security system, directed and strictly monitored by the regulator. Products were standardized and subject to licensing. Premiums were dependent only on basic risk characteristics (age and sex in life insurance) and, based on industry-wide statistics, varied little from one company to another. Underwriting and eligibility criteria were broad, a main objective being that the public in its entirety should be offered insurance coverage at affordable prices. Competition was relaxed and limited to customer service and efficient administration. Inevitably, there was an element of wealth transfer between “good” and “bad” risks. With little to gain by moving policies, customers usually stayed with the company they had chosen in the first place. In pensions they were not even allowed to move. Then as now, life insurance companies and pension funds were major accumulators of capital, but their investments were restricted to low risk asset classes: mortgages to credit-worthy borrowers (the major ones subject to the regulator’s approval) and government bonds (gilt-edged papers, traded mainly in the domestic market, and backed by the central bank). Solvency was not so much of an issue. Not in non-life insurance where contracts are short term so that premiums can be adapted to current risk experience, and large claims (contagion and conflagration) can be covered through reinsurance supplemented with provision of a fluctuation reserve. And not in life insurance and pensions because premiums were placed on the safe side and benefits were adapted to interest rates and mortality rates as they emerged in the course of the contract period (mortality risk and financial risk were eliminated by the very design of the products). With little risk there were no big losses and, correspondingly, no big gains: profits were moderate but persistent in the long run. The system was low cost: some 85 % of premiums collected by private insurers were redistributed to customers as claims and benefits, and in compulsory schemes (e.g. automobile liability insurance and occupational pension funds) the percentage was higher. There was little manoeuvring space for managers: executives were caretakers in houses catering for the public good, and their principal virtues were high actuarial standards, proper book-keeping, cost-awareness, and professional conduct. In such a culture there was not much discussion about what models to use and for what purposes: for one thing, lead management was usually recruited from the higher ranks of actuarial departments, and for another, the system produced an abundance of statistics at an aggregate level and allowed reliable estimation of claims frequencies, loss distributions, and mortality rates that were practically the same in all companies. Premiums and reserves were calculated by actuarial formulas representing broad consensus.

1.1.3 Insurance at Present

Things are different now. A wave of deregulation in the 1980s unleashed a transform of what used to be a part of the welfare state into a competitive industrial sector governed by business objectives. Co-operative endeavours, like shared data bases and exchange of information on customers, have been abolished, with the consequence that statistical estimates have become more uncertain (a Big Data opportunity forfeited). The standardized products have been replaced by a plethora of different products, usually packaged in a manner that disguises the prices of the individual covers. Customer choice is influenced by advertising campaigns, insurance brokers, and a little bit by magazines publicizing expert ratings of goods and services (vacuum cleaners, kitchen rolls, insurance,…). Life insurance products are being priced without sufficient prudence and are equipped with various sorts of guarantees that cannot all be met and have caused failures—almost and outright—of the providers. A combination of advances in technology and competition drives increasingly finer measurement and classification of risks, resulting in increasingly discriminating tariffs. Insured lives are soon to be dissected down to their very atoms using medical records, gene tests, socioeconomic indicators, information on occupation, life style, habits,…(Big Data). In turn all this leads to increased mobility of insurance contracts and intensified efforts to attract and keep customers. Premiums in non-life insurance now depend as much on studies of customer behaviour as they depend on the calculated risk. New insurance regulations oblige life insurers to estimate probabilities of lapses and surrenders of policies, which have become major sources of uncertainty. Another consequence is decreasing total insurance coverage as those flagged as bad risks (typically the not so well off) cannot afford insurance, and increasing total premium income as good risks purchase more insurance. The welfare component is fading out. Insurance companies are mushrooming in good times and disappearing through acquisitions and mergers in bad times. Under these tumultuous conditions insurers have made untold profits and losses and—inviting reflection—the former were not set aside to meet the latter.

While general insurance remained largely viable from a business point of view, life insurance and notably pensions have shown huge deficits caused by "competitive" premiums that proved to be unsustainable in a time with recurrent financial crises and significant changes in mortality. These self-inflicted losses, misperceived as risk, have been responded to in two ways. Firstly, large scale transfer of "mortality risk" to the market, which effectively means that the formerly ring-fenced coffers of the life insurance companies and pension funds have (at long last!) been opened to banks and reinsurance companies. Secondly, the resurrection of the regulators from the ashes and introduction of new regulations focusing on solvency—its *idée fixe*—and based on a combination of assumptions about the efficiency of self-regulation and the perfect market, and about the feasibility of modelling demographic and economic indices in the very long term. Also actuarial science has changed radically and now instructs that assets and liabilities of insurance companies be priced under the *Risk Neutral Measure*. **The what?** I shall explain this term later and for the moment be content to say that it comes from the combination of modern financial mathematics, new accounting standards, and new solvency regulations, all based on the doctrine that the market is perfect and produces fair prices. Of course, all this product development, marketing, risk measurement, and solvency control come with a price tag: under the new paradigm insurance is no longer low cost. The insurance companies themselves are becoming transnational, a case in point being BNP Paribas Cardif with HQ registered in Seoul. Top management is recruited from the managerial class in the global business world, the primary criterion being the success of the individual and the profits made by his/her employer, regardless of what they produce and sell (vacuum cleaners, kitchen rolls, insurance,…). Not surprisingly, the consensus on the role of models has also disappeared, as can be read between the lines (or rather the chapters) of this book. The enthusiasm of the professionals, now back in business after the reintroduction of regulation, can be sensed in many of the chapters, and the frustration of the managers over "intrusive regulation" and an uncontrolled spread of the "model jungle" is expressed in Chap. 11.

1.1.4 Outline of This Chapter

Section 1.2 elaborates on issues raised in this introduction and does so in prose style: the focus is on differences between the old and the new. Section 1.3 recapitulates basic concepts in finance and accounting, especially compound interest, and touches on the notion of arbitrage, which is central in the actuarial theories and insurance regulations under the new paradigm. Section 1.4 presents the idea of insurance, which is diversification of risk in large portfolios, and analyses the main types of life insurance and pension products. Section 1.5 presents the old paradigm's approach to management of non-diversifiable economic and demographic risk. Section 1.6 offers an elementary introduction to the arbitrage pricing theory on which much of modern life insurance mathematics is based. In particular, it aims to demystify the Risk Neutral Measure. Section 1.7 lists some bullet points on the new paradigm and concludes that it needs to be evaluated and rethought with a view to its predecessor.

Inevitably, this chapter will present some pieces of basic actuarial and financial theory, in moderate portions and mostly at a technical level accessible to all readers. Technical sections, deemed helpful but not indispensable for an understanding of the message, are highlighted with a dagger †.

1.2 Preliminaries

1.2.1 Modern Financial Economics 1

In parallel with rapid developments in financial practices there have been great advances in financial mathematics, which now presents coherent theories for the pricing of financial contracts. The basic principle of *arbitrage pricing theory (APT)* goes as follows. The market features a range of *fundamental assets*, equity and bonds, that are liquidly traded in unlimited amounts by agents who all have access to the same information, namely the past an present prices of all assets. If the market is sufficiently rich in assets, then any *derivative security* with pay-off determined by the prices of the fundamental assets can be perfectly hedged (duplicated) by an investment portfolio that is *self-financing*, which means that the portfolio is initiated with a single investment and thereafter dynamically rebalanced with no further infusion or withdrawal of capital. The amount needed to initiate this strategy is the market price of the claim: if, for example, the claim should be priced higher than the initial investment, then one could sell the claim and purchase the portfolio that will settle it, and pocket the difference at no risk. Such a costless and risk-free gain, called *arbitrage*, is not possible in a well-functioning market. The richer the market is in tradeable assets, the greater its hedging capacities. Thus, the introduction of new assets is a way of creating hedging opportunities for risks that otherwise would remain with the businesses that carried them in the first place. Such financial innovation, known as *securitization*, serves to transfer an ever increasing variety of economic risk to the marketplace. A cornerstone in the APT is the fundamental theorem of asset pricing, which states that absence of arbitrage is (essentially) equivalent to the existence of a probability measure that is *equivalent* to the physical probability measure (assigning positive probabilities to the same events) and under which the discounted price processes of the fundamental assets are martingales (future values will hover around the present ones in a purely erratic manner). Accordingly, such a measure is called an *equivalent martingale measure*, and the one used for pricing is called the risk neutral measure. The mathematics will be explained in fuller detail in Sects. 1.3 and 1.6. Let us have a look at the real market that this theory is supposed to describe.

1.2.2 The Casino

The past quarter of a century has seen a dramatic growth and change in the workings of the financial markets, the most eye-catching feature being the emergence of the derivatives market. From its tender beginnings around the mid 1980s, the derivatives market is now the largest capital market in the world, with a notional value about 10 times the total world GDP. Obviously, these derivatives cannot all be tangible (backed in real values). Some of them are pure speculation instruments like the credit default swap (CDS), which allows investors to bet on the default of a bond without having a stake in it. There are enormous profits to be amassed from market manipulation, asymmetry of information, and preferential treatment, a legendary case being Soros' quick 1 billion pounds gain on short sale of the British Pound. Under the subprime crises Lehman Brothers went bankrupt while Goldman Sachs—the former employer of the Secretary of the Treasury, the Chairman of the Federal Reserve, and many other key figures in the central government—received decisive support from the government's troubled asset relief programme (TARP). Let me recapitulate some facts about the crisis itself.

The crises was triggered by the securitization of mortgages in the US. Fuelled by bonuses rewarded for creation of business, mortgage agents knowingly oversold loans to customers with low credit-worthiness and no secured equivalent value in their properties (the housing market was a bubble at the time). This was of little concern to the lenders because they had invented *collateralized debt obligations* (CDO), derivative securities that served to transfer the right to mortgage repayments to unsuspecting investors. This way the mortgagees quickly got rid of their troubled assets and could use the proceeds from the sales to issue more loans generating more commissions, fees, and bonuses. Outstanding accounts totalling some 7.3 trillion dollars were securitized over a relatively short period from the mid 1990s to 2007. At that point an economic down-turn started to bite, with surging unemployment and depreciating house-prices. Subprime customers defaulted on their repayments, and repossessions could not settle their debts. When the investors started to sense that the CDO securities were rather insecurities, worth far less than sales people and rating agencies had made them believe, they rushed to the doors: everybody wanted to sell what nobody wanted to buy and, under the regime of "fair value" accounting, the derivatives became technically valueless. Thus, the crisis was created by a combination of asymmetric information (a finance term for what in this context was fraud), shock volatility (a quant term for what could be called globalization of the herd instinct), and the sheer scale of the CDO market which makes it a major macroeconomic driver.

Another striking feature of the capital markets of today is that floor trading performed by humans is being replaced with electronic trading performed by robots. It is estimated that some 60–70 % of the trades on US exchanges are now made by computers capable of, swiftly and automatically, gathering Big Data, processing it, and acting on it. Had the APT and its hedging offshoots offered models and methods sufficiently realistic to form the basis for practical decisions, then their implementation—a matter of model calibration and numerical computation—would of course have been computerized. However, this is not what electronic trading is about. To a great extent it has to do with high frequency trading, the essence of which is to monitor widely asks and bids and exploit the spreads to make profits in a matter of micro-seconds. This activity, a.k.a. "pinging" or "baiting", is accompanied by market making in the form of placing asks or bids large enough to move the prices and withdrawing them before they can materialize in trades. This is not something that everyone can do: the big players (investment banks and hedge funds) are using superior technology and expertise to create quick and highly probable gains from the slower operations of the small players (pension funds, local governments, and other institutional investors). Absence or non-absence of outright arbitrage is not at all an issue: all that matters is that highly probable small gains will accumulate to certain and limitless profits in the long run. In this perspective it is pathetic to watch the actuarial profession in their orchestrated endeavours to implement new regulations based on APT, the most F.A.Q. being: Where is the risk neutral measure? The answer is: it is in outmoded finance textbooks, and it is in academic papers drafted by actuaries with a working command of financial jargon, but it is nowhere to be found in the complex real world market where there is arbitrage. It is also very telling that the true financial mathematicians have long abandoned the ship: realizing that the potential of APT has been exhausted and being on the lookout for new research topics, they are heading at full sail into the uncharted waters of market microstructure, two-price economy, algorithmic trading, flash trading, front running, slow market arbitrage, rebate arbitrage, and all the rest of it. The only thing that remains unchanged is their bold claim to be catering for the needs of the market hence also the public good.

Trading on the stock exchanges always had a feature of speculation. The problem today is that this has become the dominant feature. The notion of *ownership* has changed completely. Investors are coming and going, some of them in a matter of microseconds, and they have no interest in the wellbeing of the company they hold shares in: the only thing that matters is to buy and sell at the right times. In the good old days an investment came with a responsibility of assessing the viability and/or the creditworthiness of the company in the longer term.

When my new neighbour (NN) moved in next door a decade ago, we introduced ourselves and the following transpired:

- NN: *What do you do for a living?*
- RN: *I am a financial mathematician.*
- NN (eyebrows raised): *Really? Can one earn money from that?*
- RN: *Yes, but not as much as the financiers though. And you?*
- NN: *I work in a casino. I am the security guard, checking that slots and roulettes aren't being rigged. There is no mathematician.*
- RN (after pause): *Actually, I also work in a kind of casino—a big one. There is no security guard.*

1.2.3 Competition and Costs

What economics science has to say about the blessings of competition applies to markets for standardized products like raw materials, energy, and other basic commodities for which delivered quantity is all that counts. It compares monopolies, oligopolies, cartels, and price leadership with an idealized market with infinitely many independent suppliers, and infers that the last mentioned will offer the best price. However, the major part of the market is for goods and services that are not standardized and where competition is as much on quality and design as it is on price. Therefore, product development is essential for a competitive business, and often it is to the unquestionable benefit of the consumer. Sometimes, however, it is not, and that is the case when products that ought to be standardized are being "developed" for the mere purpose of *appearing different* so that prices cannot be directly compared and customers can be guided—and misguided—by advertisements.

It is not at all obvious that customers have gained anything from the intensive product development in insurance under the new paradigm. Things worked well under the previous paradigm, when there was licensing, not only of companies, but of individual products and tariffs by the regulator. A case in point is Swiss health insurance: the terms and the conditions of the standard health insurance contract is laid down by the federal government, and the competitive parameter left to the providers is the price. Totally transparent.

Competition means there are many independent businesses selling the same goods or services. This may be a good idea for small scale production meeting local demand, but it need not be a good idea for production of sophisticated products on a grand scale requiring high-level organization and advanced expertise and technology. In particular, it is not a good idea when there are obvious advantages of large scale operations, which in the case of insurance has two peculiar aspects in addition to the commonplace ones: firstly, the diversification of risk in large portfolios and, secondly, the need for large statistical bases in order to asses risk reliably.

The Evening Standard is one of those newspapers that you pay for when you purchase vacuum cleaners, kitchen rolls, insurance services, etc. Back in 2012 it had a column telling readers that the implementation of Solvency II would cost British insurers the same as it had cost Britain to build the Olympic City in London, and concluding that Solvency II is not what the insurance industry needs. The conclusion can be debated, but the premise given was a misrepresentation. The massive costs of regulation reflects the fragmentation of the industry into a myriad of non-cooperative enterprises employing battalions of experts, managers, salesmen, and marketing people, all doing the same things except offering transparent and socially useful products at reasonable prices.

A recent account of the economic results of European insurers[1] reveals that loss ratios in non-life insurance are around 60 % (they used to be some 85 % under the previous paradigm). This is the price of competition. The report also reveals that the total expenses of the life and pension insurance industry are 130 % of the total premiums collected. It is not clear exactly what the definitions underlying this figure are, but—by any calculation and evaluation, especially those of the corporate world —this must be a complete failure. Had such results been reported for a government run insurance scheme, then it would have caused an uproar.

1.2.4 Modern Financial Economics 2

Like many other branches of theoretical economics, the APT is a means of understanding rather than a device for practical decisions. It aims to explain how consistent prices are formed in an idealized perfect market with frictionless trading (no transaction costs, no taxes, no regulation), an infinite supply of every asset (no limits to short and long positions), no asymmetry of information (all agents know equally much or little, namely the complete history of prices of all primary assets), and all agents are price-takers (no rigging of indices, no market making through massive shorting, no "Government Sachs"). Its "applications" are mainly theoretical exercises on calculation of prices for special products in special models.

Now to the question: what does the risk neutral measure look like? To the arbitrage pricing theorist this is not so much of an issue since he/she is accountable only to the logics and rules of the mathematical game. To most practitioners (investors, traders) it is not an issue at all since they do not need the APT whose main postulate is that there is no free lunch in a perfect market. Their job is to earn money by exploiting the imperfections of the real markets, and the theoretical tools needed for this are taken from econometrics, fast data processing, and fast computation (everything viewed under the physical measure). The APT may be useful in financial engineering, where theoretical calculations can assist issuers of complex derivatives in determining prices that are consistent with those of existing derivatives. In such situations it may even be an issue to use a risk neutral measure calibrated to data.

From what has been said above, it should follow that a traditional insurance company does not need APT hence does not need any risk neutral measure in its normal operations. What the company needs to do in the context of solvency assessment is to value its assets and its liabilities. For those assets that are freely traded, their market prices are available and should be used in the accounts (no theoretical risk neutral valuation). For assets that are not traded, APT is irrelevant and one uses the practical valuation principle prescribed by the relevant excerpt of

---

[1] http://www.insuranceeurope.eu/uploads/Modules/Publications/european-insurance-in-figures-2.pdf.

the accounting standards. For the insurance liabilities, which of course are not traded, one uses the best estimate based on traditional actuarial methods. In this context the notions of “Fair Value”, “Market Consistent Embedded Value”, and “Risk Neutral Valuation” are mere ceremony. They have, however, created much needless confusion.

1.2.5 Galbraith & Co

One can, in my opinion, learn more about economics and finance from small popular books with the debateable broad views of great minds than from brick-stone size elitist books filled with correct mathematical scribbles. As a general background for what is going to be discussed here, I recommend the now classic book by Galbraith (2007). John Kenneth Galbraith (1908–2006) was a professor of economics at Harvard, a public official, a diplomat, and a prolific writer whose unconventional views and outspokenness made him, arguably, the best-known economist of his time. Not only was he a tall man (2.03 m), but also intellectually he rose conspicuously above the technocrat majority of his profession about whom he once said: *Economics is extremely useful as a form of employment for economists*. A proponent of modern American liberalism (social liberalism combined with support for social justice and a mixed economy) and post-Keynesian economics from an institutionalist perspective, he expressed concerns that public debate about the economy and its social ramifications was being commanded by model nerds with a partial view of big issues. He was sceptical of the oracle status granted to natural-science-style econometrics, which implicitly denies that society is an evolutive process governed by intentions and actions of humans. Pragmatically engaged in that process, Galbraith distanced himself from the closed systems of thought held forth by Milton Friedman, Karl Marx, and other philosophical system builders: *Under capitalism man exploits man. Under communism it's the opposite.* For further background on contemporary economics, I recommend the book by Krugman (2008). Readers with sufficient appetite could give Piketty's opus magnum (Piketty 2014) a try.

1.3 Finance with a Single Asset: Banking

1.3.1 Asset Prices and Interest

Suppose financial operations are restricted to one single asset, which can be a bank account, a certain stock, or some composite of stocks. Denote by S_t the price of one share of the asset at time $t = 0, 1, 2, \ldots$, and by $\Delta S_t = S_t - S_{t-1}$ the price change in year t (from time $t-1$ to time t). The annual *interest rate* in year t is the relative change in the value over the year,

$$r_t = \Delta S_t / S_{t-1},$$

One can also write

$$\Delta S_t = S_{t-1}\, r_t \quad \text{or} \quad S_t = S_{t-1}(1 + r_t) \tag{1.1}$$

and $S_t = S_0 \prod_{j=1}^{t}(1 + r_j)$.

1.3.2 Investments

The cost of u_t shares traded at time t is $c_t = u_t S_t$. A positive cost is a purchase and a negative cost is a sale, and the number of shares traded is

$$u_t = c_t / S_t. \tag{1.2}$$

An *investment strategy* over the period from time 0 to time T is given by the costs c_t, $t = 0, \ldots, T$. By (1.2) it can equally well be defined by specifying the shares u_t, $t = 0, \ldots, T$, or the *portfolio*, which is the total number of shares held (those purchased less those sold) at any time $t = 0, \ldots, T$:

$$\theta_t = \sum_{j=0}^{t} u_j = \sum_{j=0}^{t} c_j / S_j\,.$$

The *balance of the account* at time t, denoted by U_t, is the total number of shares held times the current asset price:

$$U_t = \theta_t S_t = S_t \sum_{j=0}^{t} c_j / S_j. \tag{1.3}$$

Upon combining (1.3) and (1.1), one obtains

$$\Delta U_t = U_{t-1}\, r_t + c_t = \theta_{t-1}\, \Delta S_{t-1} + c_t\,, \tag{1.4}$$

which shows how the balance develops with interest and movements on the account.

1.3.3 Arbitrage

An investment strategy is said to be an *arbitrage* if $c_t \leq 0$ for all t, $U_T \geq 0$ for sure, and $U_T > 0$ is possible. This means that an investor can make a profit at no positive cost. Absence of arbitrage is a fundamental requirement in theoretical finance. In the present set-up with only one asset, absence of arbitrage is secured if the prices S_t are strictly positive, which means that the asset cannot become valueless or a debt. Due to (1.1), this is the case if $S_0 > 0$ and

$$r_t > -1$$

(interest greater than −100 %) for all t.

1.3.4 Financial Contracts

A *financial contract* between two parties, let us say a bank (the issuer) and a customer, specifies a sequence of payments b_t, $t = 0, \ldots, T$, to be paid by the former to the latter. The payments can be positive or negative: for a savings contract withdrawals are positive and deposits negative, and for a loan the principal (b_0) is positive and redemptions are negative. For the moment we shall consider only standard saving contracts because they have much in common with the pension and life insurance policies that are a main topic in this book. I interpose that, in fact, it is hard to find other banking products that are equally clear cut. Even consumer loans and mortgages are no longer as plain and innocent as they used to be: a few decades ago they were thrown into the market place, the first consequence being that the primary issuer ceased to be party to the contract, and the ultimate and fatal consequence being the subprime crisis of 2008 and onwards.

Now to the standard saving contract offered by commercial banks. For simplicity we disregard the administration expenses of the bank, which could be included as positive payments. In general, the bank finances the contractual payments with a sequence of costs c_t, $t = 0, \ldots, T$. The balance of the account at time t is

$$U_t = S_t \sum_{j=0}^{t} (c_j - b_j)/S_j.$$

When the contract terminates at time T, all contractual payments must be covered and the account closes with balance

$$U_T = 0. \tag{1.5}$$

This is what is meant by saying that the investment strategy is a *hedge* of the contractual payments.

1.3.5 The Principle of Equivalence

Being a regulated business, the bank can neither engage in usury nor in beneficence, and its total costs should therefore balance at 0 in some sense. In the context of savings contracts this is usually attained by requiring all costs to be null:

$$c_0 = \cdots = c_T = 0\,.$$

Under this constraint the balance of the account is just the accumulated value of deposits less withdrawals,

$$U_t = -S_t \sum_{j=0}^{t} b_j / S_j,$$

which is called the *retrospective reserve* at time t. This reserve should always be non-negative (no overdraft facility). The constraint (1.5) reduces to the so-called *principle of equivalence*,

$$\sum_{j=0}^{T} b_j / S_j = 0\,, \tag{1.6}$$

which means that the deposits match the withdrawals exactly so that neither party makes a profit.

Equivalence implies absence of arbitrage, of course. It can be attained by suitable design of the payments. If the future interest rates are known at time 0, then the b_j can be specified in nominal terms already at the outset. If, as is usually the case, future interest rates are unknown at time 0, then the payments need to be adapted to the interest rates: obviously, the payments cannot be stipulated in nominal terms at time 0, but the number of shares b_j/S_j can, and this is done in *variable interest* loans. I digress here to say that commercial banks issue a wide range of contracts, many of them with payments that do not comply with equivalence and, therefore, must be financed by costs. This is even more the case with investment banks, primarily engaged in risky operations. Business is then viable if positive costs (losses) are covered by negative costs (gains) on the average across all contracts. However, in a bubble-prone economy the banks (too big to fail) are in a win-win position: gains made during booms are pocketed by traders, managers, and owners, while losses made upon bursts are footed by tax-payers. Galbraith again: *The salary of the chief executive of a large corporation is not a market award for achievement: it is frequently in the nature of a warm personal gesture by the individual to himself.*

1.3.6 The Prospective Reserve and Its Proxies

Regulation requires the bank to provide, at any time, a reserve that adequately represents its future contractual liabilities. At time t the bank needs to hold $b_{t+1}/S_{t+1} + \cdots + b_T/S_T$ shares of the asset, and the value of this inventory is

$$V_t = S_t \sum_{j=t+1}^{T} b_j/S_j, \tag{1.7}$$

called the *prospective reserve* of the contract at time t. Upon multiplying through (1.6) by S_t and rearranging a bit, it can be recast as

$$U_t = V_t,$$

which says that, under equivalence, the balance of the account always matches the future liabilities.

Further to the discussion of contract design in the previous paragraph, we can add the following. If the contract specifies the shares b_j/S_j, which means that the nominal amounts b_j are adapted to the prices S_j, then the prospective reserve is known at any time and is precisely the reserve that the bank needs to provide. If, on the contrary, the contract specifies the nominal amounts b_j, then the prospective reserve is known only if the future prices are known. If the future prices are unknown, then the bank needs to provide a *technical reserve* that is a prudent estimate of the prospective reserve based on the available information. This notion needs to be formalized.

1.3.7 Information

The investment, the technical reserve, and other decisions made at time t must be based on the information that is available at that time, denoted by $\mathcal{F}_t$. It is assumed that no information is sacrificed at any time, which means that $\mathcal{F}_{t-1}$ is contained in $\mathcal{F}_t$. It can also safely be assumed that past and present market prices are known at any time, which in the present set-up means that $\{S_0, \ldots, S_t\}$ is contained in $\mathcal{F}_t$. Core finance theory usually makes the much stronger assumption that the past and present market prices are the only information that is available so that $\mathcal{F}_t = \{S_0, \ldots, S_t\}$. In the real world, however, some investors have more information than others and are also more capable of using the information they have.

1.4 Insurance

1.4.1 The Insurance Idea: Diversification in Large Portfolios

Supported by the previous section, we can now explain the workings of insurance schemes, the purpose of which is to mitigate the risk borne by the individual by averaging out its consequences over a large pool of individuals exposed to similar risk. It is essentially a matter of solidarity (or, at least, that is what it used to be). For an accessible and much more complete survey of modern insurance mathematics, the reader is referred to Norberg (2015).

Mr. (50), now 50 years old, plans to invest money to secure himself economically in his old age. His first idea is to deposit the amount S_0 (say) on a savings account today and withdraw the entire amount with earned interest in 20 years. Suppose interest is earned at a fixed rate of 3 % per year. After 1 year the balance of the account is $1.03S_0$, after 2 years it is $1.03^2 S_0$, and so on until 20 years have passed when the balance of the account with compound interest is

$$S_{20} = 1.03^{20} S_0 = 1.8061 S_0. \tag{1.8}$$

Now, quoting Benjamin Franklin, *in this world nothing can be said to be certain except death and taxes*. Mortality statistics show that 75 % of those who are 50 years old will survive to 70, and the prospects under the savings contract are:

- with probability 0.75 Mr. (50) will survive and then possess S_{20},
- with probability 0.25 Mr. (50) will die and lose his savings.

The expected amount at disposal in 20 years is

$$0.75 \cdot S_{20} + 0.25 \cdot 0 = 0.75 S_{20}.$$

A second idea occurs to Mr. (50). He could propose to his peer Mr. (50)*, also 50 years old, that each of them invest S_0 in a joint savings account today, and after 20 years the survivors, if any, share the balance $2S_{20}$ equally. The prospects under this mutual scheme would be:

- with probability $0.75 \cdot 0.75 = 0.5625$ both survive, and Mr. (50) will possess S_{20},
- with probability $0.75 \cdot 0.25 = 0.1875$ Mr. (50) survives Mr. (50)* and will possess $2S_{20}$,
- with probability $0.25 \cdot 0.75 = 0.1875$ Mr. (50)* survives Mr. (50) and will possess $2S_{20}$,
- with probability $0.25 \cdot 0.25 = 0.0625$ both die and lose their savings.

The expected amount at disposal for Mr. (50) in 20 years is

$$0.5625 \cdot S_{20} + 0.1875 \cdot 2 \cdot S_{20} + 0.25 \cdot 0 = 0.9375 S_{20} .$$

Compared with the individual savings account, the expected amount is substantially increased due to the possibility of inheriting the deceased peer, which in life insurance is called *mortality bequest*.

Mr. (50) realizes that the effect of mortality bequest increases with the number of peers. He easily calculates that in a mutual scheme with three peers his expected amount at disposal in 20 years would be

$$0.75 \cdot 0.75 \cdot 0.75 \cdot S_{20} + 2 \cdot 0.75 \cdot 0.75 \cdot 0.25 \cdot \frac{3}{2} \cdot S_{20} + 0.75 \cdot 0.25 \cdot 0.25 \cdot 3 \cdot S_{20}$$
$$= 0.9844 S_{20},$$

which is tantalizingly close to what the individual savings account would earn him if he is lucky and survives. Endowed with this insight, but unable to find more peers, Mr. (50) goes to an insurance company serving a large number of customers and therefore able to make each of them benefit maximally from mortality bequest. The following calculation illustrates how this works. Suppose L_{50} participants, all aged 50, join in a scheme similar to the one described. The balance of the account after 20 years is $L_{50}S_{20}$. With L_{70} survivors, the share per survivor is

$$\frac{L_{50} S_{20}}{L_{70}} .$$

Now, by the so-called *law of large numbers*, the proportion of survivors L_{70}/L_{50} stabilizes as the number of participants L_{50} increases, and in the limit it equals the survival probability 0.75. Thus, in an infinitely large portfolio the share per survivor after 20 years is

$$\frac{1}{0.75} S_{20},$$

and (50) would face the following possibilities:

- with probability 0.75 he survives and gets $\frac{1}{0.75} S_{20}$,
- with probability 0.25 he dies before 70 and gets nothing.

The expected amount at (50)'s disposal after 20 years is

$$0.75 \frac{1}{0.75} S_{20} = S_{20} ,$$

the same as (1.8). Thus, in the insurance scheme the mortality bequest produces the same expected result as the bank account for an immortal. In this sense, a pension scheme eliminates the loss due to untimely death: all the money is kept for the participants (nothing is left to their heirs). In passing, and recalling the Franklin quote, pension schemes also eliminate the inheritance tax that would have kicked in had Mr. (50) and Mr. (50)* gone for their small mutual scheme. In addition to this tax exemption comes that income tax is being postponed from early life to retirement, hence diminished by discounting and by lower marginal rates. Thus, pension funds and life annuities with life offices are very attractive tax avoidance schemes, which may explain why they have survived in spite of their paternalistic nature (pensions to the bread-winner, nothing to the dependants). We proceed to discuss in more detail the nature of the most standard insurance contracts. For the purpose of illustrating the insurance idea, we first look at a well-known contract that is not insurance.

1.4.2 A Bank Savings Contract

Mr. (x), who is now x years old and will retire in m years, opens a savings account based on a certain asset (investment portfolio) with unit price S_j at time $j = 0, 1, 2, \ldots$ A deposit of p_j will be made at each time $j = 0, 1, \ldots, m-1$ and, thereafter, a withdrawal of b_j will be made at each time $j = m, \ldots, T-1$ in order to finance consumption. Equivalence (1.6) is exercised, which means that

$$p_0/S_0 + \cdots + p_{m-1}/S_{m-1} = b_m/S_m + \cdots + b_{T-1}/S_{T-1}. \tag{1.9}$$

The role of the capital gains is clearly displayed. If there were no interest, as if the money were tucked under the mattress, then all S_j would be equal, and (1.9) would reduce to

$$p_0 + \cdots + p_{m-1} = b_m + \cdots + b_{T-1}; \tag{1.10}$$

Mr. (x) can withdraw exactly what he had saved. If the interest is positive, he can withdraw more. The higher the interest rates (the steeper the sequence S_j), the larger the total amount of withdrawals for a given investment. If the asset portfolio consists only of bonds, then the interest rates are non-negative. If it consists of equity, then the interest rates may be positive or negative, the attraction of such an asset class being that high interest (above that of bonds) is deemed more likely than low or negative interest.

In practice the future asset prices are usually unknown, and Mr. (x) faces the problem of designing the deposits and the withdrawals in a manner that will give adequate income over the entire retirement period. Moreover, since T should ideally be his remaining life length, he faces the quite impossible task of predicting this date. Therefore, he should rather do business with a life insurance company which can offer him a pension policy, a special form of savings contract that takes survival prospects explicitly into account. We proceed to describe how this is done, but first we need to understand the actuarial notion of mortality.

1.4.3 Mortality

An insurance company typically serves tens, maybe hundreds of thousands of customers, sufficiently many to ensure that survival rates are stable (the law of large numbers). From life history statistics the actuary constructs a so-called *decrement series*, which takes as its starting point a large number ℓ_0 of new-born and, for each age $x = 1, 2, \ldots, \omega$ specifies the number of survivors, ℓ_x. Here ω is the highest attainable age: $\ell_\omega = 0$. Table 1.1 shows an excerpt of the table used by Danish insurers in the 1980s to describe the mortality of insured males. The second column in the table lists some entries in the decrement series. It shows e.g. that about 65 % of all new-born will celebrate their 70th anniversary. The number of survivors ℓ_x decreases with the age x. The difference $d_x = \ell_x - \ell_{x+1}$ is the number of deaths at age x, shown in the third column of the table. The number of deaths peaks somewhere around age 80, but this does not mean that 80 is the most dangerous age; the actuary measures the mortality at any age x by the *mortality rate* $q_x = d_x/\ell_x$, which is the proportion of deceased in 1 year among those who survive to age x. This rate, shown in the fourth column of the table, increases with the age. Obviously,

$$\ell_{x+1} = \ell_x(1 - q_x). \tag{1.11}$$

1.4.4 Life Annuities and Pensions

The pension contract offered by life offices differs from the savings contract in several respects. The terminology is different in that deposits are called *contributions* or *premiums* and withdrawals are called *benefits*. More importantly, the payments are contingent on survival, and premiums and benefits are calculated by actuarial methods under which the accounting principles for savings accounts are applied, not to the individual contract, but to the entire insurance portfolio as shall now be explained.

Table 1.1 Danish 1982 mortality table for males

x	l_x	d_x	q_x
0	100 000	58	0.000579
25	98 083	119	0.001206
50	91 119	617	0.006774
60	82 339	1 275	0.015484
70	65 024	2 345	0.036069
80	37 167	3 111	0.083711
90	9 783	1 845	0.188617
100	401	158	0.394000

Consider a pension policy issued to an x year old at time 0, specifying premiums p_j due at times $j = 0, \ldots, m-1$, and benefits b_j due at times $j = m, \ldots, T-1$, where $x + T = \omega$ (a whole life policy). The actuarial calculation is based on the idealized assumption that a large number ℓ_x of policyholders aged x purchase identical policies at the same time and thereafter survive/die in exact accordance with the decrement series. Thus, with ℓ_{x+j} survivors at any time j, the total contributions at time $j = 0, \ldots, m-1$ are $\ell_{x+j}\, p_j$ and the total benefits at time $j = m, \ldots, T-1$ are $\ell_{x+j}\, b_j$. Similar to the balance Eq. (1.9), one obtains

$$p_0 \ell_x / S_0 + \cdots + p_{m-1} \ell_{x+m-1} / S_{m-1} = b_m \ell_{x+m} / S_m + \cdots + b_{T-1} \ell_{x+T-1} / S_{T-1}. \tag{1.12}$$

This relationship rests on the principle of equivalence, which was fundamental in life insurance under the previous paradigm. It was laid down in the insurance legislation and enforced through supervision. The rationale is that the gains and the losses on individual contracts should average out to zero in a large portfolio, making insurance a "fair game" between the insured and the insurer. The relationship (1.12) is also essential from a solvency point of view; if the expression on the left is no less than the expression on the right, then the contributions cover the benefits, leaving the company solvent. Equivalence is the benchmark case where the company makes neither loss nor profit. Comparing (1.12) with (1.9), and recalling the discussion following that relationship, it is seen that mortality virtually serves to increase the interest earned on the investments. Indeed, the decrement function ℓ_{x+j} decreases with j so that, ceteris paribus, the withdrawals that can be made under the budget constraint (1.12) outperform those that can be made under the budget constraint (1.9). This is due to the mortality bequest. Obviously, the steeper the decrement function, the stronger the impact of the mortality bequest.

1.4.5 Life Insurance

If Mr. (x) has dependants, then pension may not be his favoured savings vehicle. He would typically be enrolled in a compulsory occupational pension scheme, but—if well informed—he would not purchase a supplementary private life annuity: rather than exchanging inheritance entitlement with anonymous peers, he might prefer saving in a bank or in property in the shared interest of himself and his family. He might even look to secure his dependants further against the possibility of his untimely death. Then only insurance companies can help as they can offer him a life assurance policy designed as follows. A premium b_j is payable at time $j = 0, 1, \ldots, T-1$, and a benefit a_j, called *the sum insured*, is payable to the nominated beneficiaries at time $j = 1, \ldots, T$ if the insured dies in year j. The term T of the policy is usually less than $\omega - x$ as it does not make sense to insure against the certain event of dying at advanced age. Again, the actuarial calculation is based on the assumption that a large number ℓ_x of policyholders aged x purchase identical policies time 0. Equivalence is attained by designing premiums and benefits such that

$$p_0\,\ell_x/S_0 + \cdots + p_{T-1}\,\ell_{x+T-1}/S_{T-1} = a_1 d_x/S_1 + \cdots + a_T\,d_{x+T-1}/S_T.$$

1.4.6 Sensitivity of the Equivalence Premium to Changes in Interest and Mortality

In order to gain some insight into the roles of interest and mortality and the risk associated with them, we shall provide some comparative numerical results based on the equivalence relationship (1.12) for savings products. Assuming ad hoc that future interest rates and mortality rates are known upon inception of the contract, we can design a contract with level premium p and level withdrawal b satisfying equivalence for fixed term T and fixed premium payment period m. The indicator we shall look at is the annual premium needed per annual amount withdrawn,

$$\frac{p}{b} = \frac{\ell_{x+m}/S_m + \cdots + \ell_{x+T-1}/S_{T-1}}{\ell_x/S_0 + \cdots + \ell_{x+m-1}/S_{m-1}}\,,$$

taking $x = 30$, $m = 35$, and $T = 70$. Focusing first on the role of interest, assume that there is no mortality (we are back to banking). If there is no interest (we are back to (1.10)), then $p/b = 1$. If interest is 4 %, then $p/b = 0.254$. If interest is 8 %, then $p/b = 0.068$. These numbers demonstrate the strong impact of the interest rate within a range of plausible values. Next we focus on mortality, assuming that interest is fixed at 4 %. If the mortality is as in Table 1.1 (mean life length is 73 years), then $p/b = 0.115$, which should be compared with 0.254 for the case with no mortality. If the mortality rates are the half of those in Table 1.1 (mean life length is 81), then $p/b = 0.159$. These numbers demonstrate that also mortality has a strong impact, although not as great as interest (the order of magnitude and the volatility of the former is much higher than the latter).

1.5 Management of Economic-Demographic Risk

1.5.1 The Big Issue

Life insurance and especially pension policies are long-term contracts, with time horizons wide enough to see significant changes in financial and demographic indices determining the economic result of the portfolio. Since such indices affect all individuals in the same manner, the risk associated with their uncertain development cannot be mitigated by increasing the size of the portfolio: we are all in the same boat on our voyage through the troubled waters of the economic-demographic history, and the waves cannot be calmed by taking more passengers on board. With a view to the sensitivity study in the previous paragraph, it is clear that management of economic-demographic risk is a big issue in insurance. I shall present actuarial approaches to its resolution.

1.5.2 With-Profit Insurance

This scheme is the traditional approach. I shall describe it in some technical detail because it is widely ignored by actuaries nowadays and probably unknown to most of them. I also find it pertinent to include—in this book on modelling in life insurance—a piece of evidence that traditional actuarial mathematics was decision-oriented, parsimonious in its assumptions, and still sophisticated enough to support deep insights into its subject matter.

Consider a combined policy issued to an x year old at time 0, terminating at time T, and specifying that at time j the insured will pay a premium p_j^* and receive a life annuity payment b_j^* conditional on survival or a sum insured a_j^* conditional on death, all payments in nominal amounts (dollars, euros, pounds,...). These contractual payments are binding to both parties throughout the contract period, hence the insurer cannot counter adverse developments of interest or mortality by raising premiums or reducing benefits. Therefore, the contractual premiums are set on the safe side, sufficiently high to cover the contractual benefits under any likely scenario. The way this is usually done is to enforce equivalence with the unknown indices S_j and ℓ_{x+j} replaced by so-called *technical* indices S_j^* and ℓ_{x+j}^* that are prudently chosen and ideally should represent a worst case scenario. Subsequently, as time passes and the true indices surface, the insurer will see systematic surpluses emerging from the prudent technical assumptions. These surpluses belong to the insured and are to be repaid in the form of *bonuses*, the timing of repayments being at the insurer's discretion.

By (1.6), equivalence is attained under the technical assumptions if

$$\sum_{j=0}^{T}\Big((p_j^* - b_j^*)\ell_{x+j}^* - a_j^* d_{x+j-1}^*\Big)/S_j^* = 0 \tag{1.13}$$

(defining $a_0^* = 0$). This means that there is balance between the technical versions of the discounted retrospective reserve,

$$\widetilde{U}_t^* = \sum_{j=0}^{t}\Big((p_j^* - b_j^*)\ell_{x+j}^* - a_j^* d_{x+j-1}^*\Big)/S_j^*$$

and the discounted prospective reserve,

$$\widetilde{V}_t^* = \sum_{j=t+1}^{T} \Big((b_j^* - p_j^*)\ell_{x+j}^* + a_j^* d_{x+j-1}^* \Big) / S_j^* .$$

Factorize the latter as

$$\widetilde{V}_t^* = \frac{\ell_{x+t}^*}{S_t^*} \upsilon_t^* ,$$

where ℓ_{x+t}^*/S_t^* pertains to the present and

$$\upsilon_t^* = \frac{S_t^*}{\ell_{x+t}^*} \sum_{j=t+1}^{T} \Big((b_j^* - p_j^*)\ell_{x+j}^* + a_j^* d_{x+j-1}^* \Big) / S_j^*$$

pertains to the future and has the obvious interpretation as the technical prospective reserve per survivor at time t. For obvious reasons, the contract is always designed such that the technical reserve is non-negative, hence $\upsilon_t^* \geq 0$.

The discounted *surplus* at time t is defined as

$$\widetilde{W}_t = \widetilde{U}_t - \frac{\ell_{x+t}}{S_t} \upsilon_t^* , \tag{1.14}$$

where

$$\widetilde{U}_t = \sum_{j=0}^{t} \Big((p_j^* - b_j^*)\ell_{x+j} - a_j^* d_{x+j-1} \Big) / S_j$$

is the discounted retrospective reserve generated by the realized rates of interest and mortality, and $\ell_{x+t}\upsilon_t^*/S_t$ is the discounted total reserve that needs to be provided for the realized number of survivors under the prudent first order assumptions. The surplus should be positive, and the way to achieve this is seen from its dynamics to be derived below:

$$\Delta \widetilde{W}_t \approx \frac{\ell_{x+t}}{S_t} \left[(q_{x+t-1}^* - q_{x+t-1})(a_t^* - \upsilon_{t-1}^*) + (r_t - r_t^*)\upsilon_{t-1}^* \right]. \tag{1.15}$$

This transparent expression guides us in placing the elements of the technical basis on the safe side. Firstly, since $\upsilon_t^* \geq 0$, the technical interest should be chosen smaller than the factual interest at all times. Secondly, the technical mortality rates should be chosen bigger than the true mortality rates in years with $a_t^* - \upsilon_{t-1}^* > 0$ and smaller in years with $a_t^* - \upsilon_{t-1}^* < 0$. In particular, for a pure death benefit the technical mortality should be set high, and for a pure pension contract the technical mortality should be set low.

The surpluses that emerge from prudent technical assumptions are to be redistributed to the insured in the form of bonuses (dividends). Their discounted values $\tilde{d}_t$ must meet the requirement of prudence,

$$\sum_{j=1}^{t}(\Delta\widetilde{W}_j - \tilde{d}_j) \geq 0, t = 1, \ldots, T,$$

and the requirement of equivalence (fairness in a proper sense),

$$\sum_{j=1}^{T}(\Delta\widetilde{W}_j - \tilde{d}_j) = 0\,.$$

The bonuses should be designed in a manner that conforms with the purpose of the insurance product. Typically, they will take the form of additional benefits.

I underscore that the only model assumption made is that the law of large numbers applies to a large portfolio under the given realization of the mortality rates. This humble assumption should be no matter of dispute, and without it there would be no rational foundation for insurance. Apart from the specification of the worst case scenario, no specific model assumptions are made about the future development of interest rates and mortality rates. This is very good since forecasting and modelling of interest and mortality rates under the time horizons of insurance and pensions contracts would be pure guesswork and would introduce a superimposed model risk. Galbraith again: *The only function of economic forecasting is to make astrology look respectable*. He would have turned in his grave had he known how actuarial science has been replaced by the crystal ball exercises required by Solvency II. Furthermore, I underscore that the with-profit scheme eliminates the conundrum of solvency risk caused by uncertain and non-diversifiable economic-demographic factors: such risk is placed on the shoulders of the insured, who will have to live—and die—with the benefits that the premiums can sustain under the economic and demographic circumstances. Which is fair enough: the role of insurance is to average out the purely random differences between individuals, not to rescue them from their common fate. The credo of the with-profit scheme is: good risk management starts with writing sensible contracts, and that is where it could end if it were done properly.

The with-profit concept has been blamed for the pension crisis. Unjustly, of course: if you get a ticket for speeding, you do not put the blame on your Mercedes. The with-profit vehicle was good enough, but it was not conducted with sufficient prudence. The truth of the matter is that, in attempts to attract customers, insurers wanted their products to appear cheap. Level-headed actuarial judgement had to give way to salesman manners; premiums were set too low, and bonuses were promised prematurely. This could happen after the regulators went into hibernation in the 1980s.

1.5.3 † Derivation of (1.15)

The dynamics of the discounted surplus in (1.14) is

$$\Delta \widetilde{W}_t = \Delta \widetilde{U}_t - \Delta\left(\frac{\ell_{x+t}}{S_t} \upsilon^*_t\right). \tag{1.16}$$

Since reserves are linear (additive and proportional to the payments), it suffices to analyse the surplus dynamics for two elemental insurance benefits.

First, consider a life endowment that pays 1 at time j contingent on survival. To calculate (1.16) we need the following:

$$\widetilde{U}_t = \begin{cases} 0, & t < j, \\ -\frac{\ell_{x+j}}{S_j}, & t \geq j, \end{cases} \tag{1.17}$$

$$\Delta \widetilde{U}_t = \begin{cases} 0, & 1 \leq t < j, \\ -\frac{\ell_{x+j}}{S_j}, & t = j, \\ 0, & t > j, \end{cases} \tag{1.18}$$

$$\upsilon^*_t = \begin{cases} \frac{S^*_t}{\ell^*_{x+t}} \frac{\ell^*_{x+j}}{S^*_j}, & 0 \leq t < j, \\ 0, & t \geq j. \end{cases} \tag{1.19}$$

For υ^*_t we have the recursive relationship

$$\upsilon^*_t = \frac{S^*_t}{\ell^*_{x+t}} \frac{\ell^*_{x+t-1}}{S^*_{t-1}} \upsilon^*_{t-1} = \frac{1+r^*_t}{1-q^*_{x+t-1}} \upsilon^*_{t-1}, \; 1 \leq t < j, \tag{1.20}$$

the last expression obtained from (1.1) and (1.11). Using this, we next calculate

$$\Delta\left(\frac{\ell_{x+t}}{S_t} \upsilon^*_t\right) = \begin{cases} \frac{\ell_{x+t}}{S_t}\left(\frac{1+r^*_t}{1-q^*_{x+t-1}} - \frac{1+r_t}{1-q_{x+t-1}}\right)\upsilon^*_{t-1}, & 1 \leq t < j, \\ -\frac{\ell_{x+j-1}}{S_{j-1}} \upsilon^*_{j-1}, & t = j, \\ 0, & t > j. \end{cases} \tag{1.21}$$

Entering the expressions in (1.18) and (1.21) for $t < j$ into (1.16), we obtain

$$\Delta \tilde{W}_t = \frac{\ell_{x+t}}{S_t}\left(\frac{1+r_t}{1-q_{x+t-1}} - \frac{1+r^*_t}{1-q^*_{x+t-1}}\right)\upsilon^*_{t-1}. \tag{1.22}$$

Trivially, this expression also holds for $t > j$ as $\upsilon^*_{t-1} = 0$. Finally, for $t = j$,

$$\Delta\tilde{W}_j = -\frac{\ell_{x+j}}{S_j} + \frac{\ell_{x+j-1}}{S_{j-1}}\upsilon^*_{j-1} = \frac{\ell_{x+j}}{S_j}\left(\frac{1+r_j}{1-q_{x+j-1}} - \frac{1+r^*_j}{1-q^*_{x+j-1}}\right)\upsilon^*_{j-1},$$

the same expression as for $t<j$. Thus, (1.22) is a unifying formula.

Second, consider a simple death benefit that pays 1 at time j if death occurs in year j, for which we have

$$\begin{aligned} \widetilde{U}_t &= \begin{cases} 0, & t<j, \\ -\frac{d_{x+j-1}}{S_j}, & t \geq j. \end{cases} \\ \upsilon^*_t &= \begin{cases} \frac{S^*_t}{\ell^*_{x+t}}\frac{d^*_{x+j-1}}{S^*_j}, & 0 \leq t < j, \\ 0, & t \geq j. \end{cases} \end{aligned} \tag{1.23}$$

These expressions are similar to (1.17) and (1.19), with d_{x+j-1} and d^*_{x+j-1} in the roles of ℓ_{x+j} and ℓ^*_{x+j}. Therefore, (1.20) remains the same, (1.18) and (1.21) remain the same for all $t \neq j$, hence (1.22) holds for all $t \neq j$. The case $t = j$, where

$$\Delta\tilde{W}_j = -\frac{d_{x+j-1}}{S_j} + \frac{\ell_{x+j-1}}{S_{j-1}}\upsilon^*_{j-1},$$

requires special treatment. Anticipating that $\Delta\tilde{W}_j$ should have a term like (1.22), let us subtract that from the expression above (and add it again later):

$$\begin{aligned} &-\frac{d_{x+j-1}}{S_j} + \frac{\ell_{x+j-1}}{S_{j-1}}\upsilon^*_{j-1} - \frac{\ell_{x+j}}{S_j}\left(\frac{1+r_j}{1-q_{x+j-1}} - \frac{1+r^*_j}{1-q^*_{x+j-1}}\right)\upsilon^*_{j-1} \\ &\quad = \frac{\ell_{x+j}}{S_j}\left(-\frac{d_{x+j-1}}{\ell_{x+j}} + \frac{1+r_j}{1-q_{x+j-1}}\upsilon^*_{j-1} - \frac{1+r_j}{1-q_{x+j-1}}\upsilon^*_{j-1} + \frac{1+r^*_j}{1-q^*_{x+j-1}}\upsilon^*_{j-1}\right) \end{aligned}$$

(cancel second and third terms in the parenthesis, reshape fourth term using (1.23), and continue)

$$= \frac{\ell_{x+j}}{S_j}\left(-\frac{q_{x+j-1}}{1-q_{x+j-1}} + \frac{q^*_{x+j-1}}{1-q^*_{x+j-1}}\right).$$

We gather

$$\Delta\tilde{W}_t = \frac{\ell_{x+t}}{S_t}\left[\delta_{tj}\left(\frac{q^*_{x+t-1}}{1-q^*_{x+t-1}} - \frac{q_{x+t-1}}{1-q_{x+t-1}}\right) + \left(\frac{1+r_t}{1-q_{x+t-1}} - \frac{1+r^*_t}{1-q^*_{x+t-1}}\right)v^*_{t-1}\right],$$

where δ_{tj} is 1 for $t=j$ and 0 otherwise.

Since reserves are linear, we obtain for the combined policy that

$$\Delta\tilde{W}_t = \frac{\ell_{x+t}}{S_t}\left[\left(\frac{q^*_{x+t-1}}{1-q^*_{x+t-1}} - \frac{q_{x+t-1}}{1-q_{x+t-1}}\right)a_t + \left(\frac{1+r_t}{1-q_{x+t-1}} - \frac{1+r^*_t}{1-q^*_{x+t-1}}\right)v^*_{t-1}\right].$$

The approximate expression (1.15) is obtained upon noting that $1/(1-q) \approx 1+q$ if q is close to 0, and that terms like $(q^*_{x+t-1})^2$ and $q^*_{x+t-1}r^*_t$ are negligible.

1.5.4 Index-Linked Insurance

The bonuses repaid to holders of with-profit contracts serve to adapt the benefits to the development of the non-diversifiable indices. Alternatively, the very terms of the contract could specify payments that depend, not only on individual life history events, but also on the development of interest, mortality, and other economic-demographic factors. An example is given by the so-called *variable annuities*, which relate the benefits to the performance of the insurer's investment portfolio. They emerged in response to the elevated and volatile interest rates seen in many countries from the late 1960s until the early 1990s. (Mortality risk was not perceived as important when interest rates were some hundred times bigger than mortality rates.) As an illustration of the idea, consider the equivalence relation for a pure life endowment of b against level premium p:

$$p\sum_{j=0}^{T-1}\ell_{x+j}/S_j = b\,\ell_{x+T}/S_T. \tag{1.24}$$

The traditional insurance contract specifies p and b at the outset. If the interest rates hence the prices S_j are uncertain, then so is the fulfilment of equivalence. The solution to this problem is to specify only p in the contract and let b be determined as the solution to (1.24). This way the interest risk has been passed from the insurer to the insured. The idea was launched with great fanfare in the insurance world, but it wasn't new: variable interest savings accounts are what banks typically offer and have been offering since banks emerged in antiquity.

I shall pursue and generalize this idea, and I am proud to announce that this book contains at least one piece of independent and novel research. Not of any mathematical greatness, but of potential practical import. A suitable framework is the combined policy considered in the previous two paragraphs. At time 0 the designer of the contract specifies *baseline indices* S_j^* and ℓ_{x+j}^* (or r_j^* and q_{x+j}^*), which typically are best estimates (not on the safe side). Baseline payments p_j^*, b_j^*, and a_j^* are designed so as to satisfy equivalence in this scenario:

$$\sum_{j=0}^{T}\Big((p_j^* - b_j^*)\ell_{x+j}^* - a_j^* d_{x+j-1}^*\Big)/S_j^* = 0. \tag{1.25}$$

(Formally the same as (1.13), but the idea is different.) The contractual payments are specified as

$$p_j = \frac{S_j \ell_{x+j}^*}{S_j^* \ell_{x+j}} p_j^*, \quad b_j = \frac{S_j \ell_{x+j}^*}{S_j^* \ell_{x+j}} b_j^*, \quad a_j = \frac{S_j d_{x+j-1}^*}{S_j^* d_{x+j-1}} a_j^*,$$

having the same profile as the baseline payments, but being tilted in response to the movements of the indices. Now, insert these payments into the equivalence requirement,

$$\sum_{j=0}^{T}\big((p_j - b_j)\,\ell_{x+j} - a_j d_{x+j-1}\big)/S_j = 0,$$

and you will see that—hocus pocus—the problematic indices disappear as everything reduces to (1.25), which was arranged for by the choice of the baseline payments at the outset.

This method is an example of an *automatic balancing mechanism*, a term that covers various ways of adapting payments to indices in order to eliminate solvency risk. There are many conceivable ways of arranging automatic balancing, and optimal design of such mechanisms may even invite academic research. However, the most important thing is, just as for with-profit, that it offers a common sense actuarial solution to the solvency problem and does not require models for something that cannot be modelled realistically.

1.5.5 Guarantees

Variable annuities in their clear cut form are no longer widely sold. In an environment of sharp competition and in the absence of product regulation, insurers developed a range of new products that—supported by marketing apparatus—were perceived as attractive. Index-linked products were equipped with guarantees with the proclaimed purpose of protecting policy holders against the downside risk of low interest. Soon the market was abounding with guarantee-type products with fancy acronyms like GMAB, GMDB, MGSV, GAO, MGIR, the legend being A = annuity, B = benefit, D = death, G = guaranteed, IR = interest rate, M = minimum, O = option, SV = surrender value,… Sales people and actuaries engaging in financial innovation had a great time. Academics engaging in application of option pricing theory to the new products had a great time. Journals publishing papers with all the calculations (well known, but now under the Risk Neutral Measure!) had a great time. And customers enjoying free choice from a range of products—brand new and coming with the coveted guarantees—were told they had a great time. The truth of the matter is that customers didn't have the foggiest idea how to compare all these opaque products or what the guarantees really meant. For an example, consider the variable interest life endowment discussed in the previous paragraph. The endowment b determined by (1.24) increases with increasing interest rates, which is obvious but can also be seen upon multiplying through by S_T and recalling that $S_T/S_j = \prod_{k=j+1}^{T}(1+r_k)$. Suppose the insurer introduces a guaranteed "floor" $\underline{r}$ to the interest rate, which means that the sum insured $\underline{b}$ is determined by

$$p\sum_{j=0}^{T-1}\ell_{x+j}\prod_{k=j+1}^{T}(1+\max(r_k,\underline{r}_k)) = \underline{b}\,\ell_{x+T}. \tag{1.26}$$

Since, for fixed p, $\underline{b}$ is no less than b given by (1.24), the guarantee must be financed by an added premium, e.g. raising the premium rate from the nominal p to the "market consistent" $\underline{p}$ determined such that the expected present value of the premiums equals the expected present value of the benefits in (1.26) (under the risk neutral measure by all means). I skip this elementary exercise and instead turn to the common sense comments this example provokes. Firstly, the guarantee reintroduces financial risk on the part of the insurer (it is only in the theory that it can perfectly hedge itself out of the consequences of its light-hearted promises). Secondly, theoretical pricing of the guarantee requires that interest rates be modelled for the entire term of the contract, which may be exceedingly long. Thirdly, since the guarantee must be paid for, the customer will buy more insurance, only with a different benefit. While the guarantee reduces the variance of the nominal value of the benefit at term, it obviously increases the variance of the present value of the benefit at time 0, which for the clear cut variable annuity is 0. Thus, if interest rates follow inflation, and if the customer's objective is to preserve the purchasing power of the savings, then the guarantee would not be useful. The typical guarantee serves to link the benefits to nominal values, which may be a selling point, but nothing more; I don't know what 1000 € can buy me in 50 years—a kitchen roll?

The guarantee culture has its origins in malpractices that emerged in the decline and fall of the previous paradigm. Obviously, the technical basis in the with-profit scheme is essentially also a guarantee (a GMB if you like), but the point is that it should be tenable, even under a worst case scenario. The problem was that the technical assumptions became less and less prudent. To my recollection comes an episode in 1985 when I was a member of the Norwegian Council of Insurance Supervision. The Council was considering an application for the licensing of a technical basis with annual interest rates starting at 8 % and thereafter descending to the commonly used 4.5 % over 10 years. The chairman, a prominent public figure and former secretary of the Ministry of Social Affairs, exclaimed: *Those people must be completely ignorant of contemporary economical history. Today's interest rates at 10 % and above are unsustainable in the long perspective. Only three decades ago they were 2 %, and that could happen again.* It has happened already, and with the return of depression economics we are even seeing negative interest rates. By the way, and pertaining to marketing, I added: *The applying company claims it is offering cheap insurance. This is misleading advertising because, in properly regulated insurance, customers get what they pay for. If premiums go down due to more optimistic interest assumptions, then the bonuses will also go down.* In fact, the very company might go down as did Equitable Life after they offered GAOs effectively guaranteeing 8 % interest.

1.6 Finance in a Multi-asset Market

1.6.1 APT

As was seen in Sect. 1.3, finance in a single-asset market is little more than book-keeping and calculation of compound interest. The only issue of some theoretical interest is how to model the annual interest rates. Finance in a multi-asset market raises questions of far greater difficulty and hence also theoretical interest: what are possible price processes in a well-functioning market, and what are the principles for writing fair contracts and for pricing of financial derivatives? Such questions are studied in arbitrage pricing theory (APT), a branch of financial economics instituted by the path-breaking work of Black, Scholes and Merton in 1973 and thereafter rapidly developed into a vibrant area of mathematics. Tantalizingly, as the name suggests, the theory builds entirely on the assumption that the market admits no arbitrage. One may ask: how can such a humble assumption power such an impressive body of theory, and is it really sufficient to ensure good performance of the real market? Reasoned answers to these critical questions require a proper understanding of both the APT and the workings of the real financial market. First, I shall present the essentials of APT. For my purpose, which is just to demonstrate the otherworldly nature of this theory, it suffices to work with a toy model that uses only elementary mathematics. For a general and rigorous account based on discrete time models, the reader is referred to Föllmer and Schied (2011).

1.6.2 A Market with Two Assets

The market features two freely tradeable assets. The first asset is a bank account that is risk-free in the sense that the annual interest rate r_j^0 in year j is known at the beginning of that year. Recalling Sect. 1.3, the price process is given by S_0^0 (known at time 0) and

$$S_t^0 = S_{t-1}^0(1 + r_t^0), t = 1, 2, \ldots$$

The second asset is a stock that is risky in the sense that the annual interest rate r_t^1 in year t is known only at the end of that year. The price process of a unit invested at time 0 is given by S_0^1 (known at time 0) and

$$S_t^1 = S_{t-1}^1(1 + r_t^1), t = 1, 2, \ldots$$

All prices are strictly positive, which means that $S_0^i > 0$ and

$$r_t^i > -1, \text{ for all } t, \ i = 0, 1.$$

At the cost of c_t at time t one can purchase u_t^0 shares of the bank account and u_t^1 shares of the stock under the budget constraint

$$c_t = u_t^0 S_t^0 + u_t^1 S_t^1. \tag{1.27}$$

An *investment strategy* over the period from time 0 to time T is given by the *portfolio* (θ_t^0, θ_t^1), $t = 0, 1, \ldots, T$, which specifies the total number of shares held in each asset at any time:

$$\theta_t^i = \sum_{j=0}^{t} u_j^i,$$

$t = 0, \ldots, T$, $i = 0, 1$. The shares u_j^i can be positive (a purchase) or negative (a sale). If a portfolio component θ_t^i is positive, then it is called a *long position* (the shares are owned); if it is negative, then it is called a *short position* (the shares are owed/borrowed). Due to the relationship (1.27), the strategy can equally well be specified by the costs and the total number of shares in the stock, (c_t, θ_t^1), $t = 0, 1, \ldots, T$.

The value of the portfolio at time t is

$$U_t = \theta_t^0 S_t^0 + \theta_t^1 S_t^1,$$

and its dynamics is

$$\Delta U_t = \theta^0_{t-1}\,\Delta S^0_t + \theta^1_{t-1}\Delta S^1_t + c_t, \tag{1.28}$$

starting from $U_0 = c_0$. The relationship (1.28) speaks for itself and can be calculated in the same manner as (1.4).

I interpose that investment in two assets, with short and long positions, opens possibilities of limitless profits or losses. This is the nature of speculation. If short positions were not allowed, as occasionally happens when regulators need to calm panicking markets, then no portfolio could ever have a negative value.

1.6.3 Does This Market Admit Arbitrage?

It doesn't take a deep theorem to see through a shallow notion, so the answer is almost trivial:

Proposition 1.1 *There is arbitrage if and only if one given asset outperforms the other asset for sure in some given year*

Proof Recall the definition of arbitrage given in Sect. 1.3 and that an investment strategy now consists of (c_t, θ^1_t), $t = 0, 1, \ldots, T$. As explained in Sect. 1.3, an arbitrage cannot be based on trading in the bank account only. Therefore, if an arbitrage should exist, it must be a strategy with $c_t \leq 0$ for all t and $\theta^1_t \neq 0$ for some t. Consider such a strategy. By (1.28),

$$\begin{aligned} U_t &= U_{t-1} + \theta^0_{t-1}\Delta S^0_t + \theta^1_{t-1}\,\Delta S^1_t + c_t \\ &\leq U_{t-1} + \theta^0_{t-1} S^0_{t-1}\, r^0_t + \theta^1_{t-1}\, S^1_{t-1}\, r^1_t \\ &= U_{t-1}(1 + r^0_t) + \theta^1_{t-1}\, S^1_{t-1}\,(r^1_t - r^0_t)\,. \end{aligned} \tag{1.29}$$

If both $r^1_t > r^0_t$ and $r^1_t < r^0_t$ are possible for all t, then it is possible that the second term in (1.29) is ≤0 for all t and <0 for some t. By (1.29), this means that

$$U_t \leq U_{t-1}(1 + r^0_t)$$

with strict inequality for some t. As $U_0 = c_0 \leq 0$ and $1 + r^0_t > 0$, it follows that $U_T < 0$ is possible, which means the strategy is not an arbitrage. On the other hand, if for instance $r^1_t > r^0_t$ for sure in a given year t, then there is arbitrage (e.g. take a loan in the bank at time $t - 1$, invest it in the stock, sell the stocks at time t, and repay the bank). □

We shall elaborate on the notion of arbitrage. An investment strategy is said to be *self-financing (SF)* if $c_t = 0$ for all t = 1, …, T, that is, it is fully costed at time 0. For such a portfolio the dynamics (1.28) reduces to

$$\Delta U_t = \theta^0_{t-1}\Delta S^0_t + \theta^1_{t-1}\Delta S^1_t, \, t = 1, \ldots, T, \tag{1.30}$$

starting from $U_0 = c_0$. The relationship (1.30) says that, after the initial cost has been footed, the value of the portfolio depends only on the gains and losses of the investments. Another way of phrasing this is: after time 0 every purchase of one asset is financed by a sale of the other asset. This comes from (1.27), which now becomes

$$u^0_t S^0_t + u^1_t S^1_t = 0, t = 1, 2, \ldots, T. \tag{1.31}$$

It takes just a little bit of thinking to realize the following:

Proposition 1.2 *There is arbitrage if and only if there is an SF arbitrage.*

Self-financing strategies play a key role in the theory of pricing of financial derivatives to be treated next.

1.6.4 Pricing

This aspect of the theory requires that all possible outcomes of the interest rate processes be fully specified. For simplicity, let us assume that the r^0_t are equal to a fixed and known rate r and that the r^1_t can assume a low value d (market is down) or a high value u (market is up) such that (recall Proposition 1.1)

$$d < r < u. \tag{1.32}$$

A financial derivative is a freely tradeable contract specifying that the holder will receive payments that, at any time, depend (only) on the market information at that time. We shall consider only simple derivatives of so-called *European* type: such a derivative pays an amount of the form $H(\mathcal{F}_T)$ at a fixed time T (the term). In the present market the flow of information is given by $\mathcal{F}_t = (r^1_1, \ldots, r^1_t)$.

Proposition 1.3 *To every European derivative there exists an SF strategy that duplicates the pay-off exactly, that is, $U_T = H(\mathcal{F}_T)$ for sure. By the principle of no arbitrage, the cost of this strategy is the unique price of the claim at time 0.*

Remark Plainly, the second statement follows from the first as explained in the first paragraph of Sect. 1.2. It also follows that U_t is the unique price of the pay-off at time t.

Proof The first step is to construct a portfolio $(\theta^0_{T-1}, \theta^1_{T-1})$ at time $T-1$ that duplicates the pay-off at time T. The value of the portfolio at time $T-1$ is

$$U_{T-1} = \theta^0_{T-1} S^0_{T-1} + \theta^1_{T-1} S^1_{T-1}, \tag{1.33}$$

and its value at time T, with no cost added, is

$$\begin{aligned} U_T &= \theta^0_{T-1} S^0_T + \theta^1_{T-1} S^1_T \\ &= \theta^0_{T-1} S^0_T + \theta^1_{T-1} S^1_{T-1}(1 + r^1_T). \end{aligned}$$

The portfolio duplicates the pay-off $H(\mathcal{F}_T) = H(\mathcal{F}_{T-1}, r^1_T)$ if their values coincide in either of the cases "market down" and "market up":

$$\theta^0_{T-1} S^0_T + \theta^1_{T-1} S^1_{T-1}(1 + d) = H(\mathcal{F}_{T-1}, d), \tag{1.34}$$

$$\theta^0_{T-1} S^0_T + \theta^1_{T-1} S^1_{T-1}(1 + u) = H(\mathcal{F}_{T-1}, u). \tag{1.35}$$

From these two equations we solve $(\theta^0_{T-1}, \theta^1_{T-1})$ as functions of $\mathcal{F}_{T-1}$. Upon inserting this solution into (1.33), we obtain the value U_{T-1} as a function of $\mathcal{F}_{T-1}$. Repeating this argument, we next determine the portfolio $(\theta^0_{T-2}, \theta^1_{T-2})$ at time $T-2$ that will duplicate U_{T-1}. Continuing backwards in this manner, we finally determine U_0, which is the cost c_0 in an SF strategy that duplicates the pay-off perfectly. □

1.6.5 † The Risk Neutral Measure

There exists an alternative characterization of arbitrage, which is mathematically more elaborate without adding any deeper insight into arbitrage as such (Proposition 1.1 is the crux of the matter). I shall demonstrate it here because it takes us on a guided tour into the enigmatic APT edifice and the hideouts of its phantom, the Risk Neutral Measure.

The first step is to measure monetary amounts at any time in units of the current price of the bank account to obtain the so-called *discounted* values. Equipping those with a tilde, the discounted asset prices are $\tilde{S}^0_t = 1$ and $\tilde{S}^1_t = S^1_t / S^0_t$. One should expect that the property of being SF does not depend on the monetary unit and, therefore, is preserved under discounting. That this is indeed the case, is seen by dividing by S^0_t in (1.31). Therefore, the discounted value of an SF portfolio has dynamics as in (1.30), only with discounted prices in the place of the nominal ones. Since $\Delta \tilde{S}^0_t = 0$, we arrive at

$$\Delta \tilde{U}_t = \theta^1_{t-1} \Delta \tilde{S}^1_t. \tag{1.36}$$

Thus, the development of the value of an SF portfolio, once initiated, is entirely determined by the increments $\Delta\tilde{S}_t^1$. These increments can be reshaped as

$$\Delta\tilde{S}_t^1 = \frac{S_t^1}{S_t^0} - \frac{S_{t-1}^1}{S_{t-1}^0} = \frac{S_{t-1}^1}{S_{t-1}^0}\left(\frac{1+r_t^1}{1+r} - 1\right) = \frac{\tilde{S}_{t-1}^1}{1+r}\left(r_t^1 - r\right). \tag{1.37}$$

The second step requires that we engage in spiritualism, imagining that the market is governed by a probability measure $\tilde{\mathbb{P}}$ under which the interest rates r_t^1 are independent and identically distributed random variables. Thus, $\tilde{\mathbb{P}}[r_t^1 = u] = \tilde{p}$ and $\tilde{\mathbb{P}}[r_t^1 = d] = 1 - \tilde{p}$ for some $\tilde{p}$ between 0 and 1. The expected value of r_t^1 under $\tilde{\mathbb{P}}$ is $\tilde{\mathbb{E}} r_t^1 = u\tilde{p} + d(1-\tilde{p})$. The magic trick is to make this expected value equal to r, which is possible due to (1.32) and is obtained by choosing

$$\tilde{p} = \frac{r-d}{u-d}.$$

This way the increments in (1.37) will satisfy

$$\tilde{\mathbb{E}}[\Delta\tilde{S}_t^1 \,|\, \mathcal{F}_{t-1}] = 0,$$

which is precisely the defining property of a *martingale*. Being constant, $\tilde{S}^0$ is trivially a martingale. Thus, in this fantasy model the discounted asset prices are martingales: they have zero mean future increments for all outcomes of the past. By (1.36) it follows that the discounted value of any SF portfolio must be a martingale, implying that its expected value at the term T must be the same as its initial value. Obviously, there cannot exist any SF arbitrage and so, by Proposition 1.2, there exists no arbitrage. The results derived here are valid in much more general models and can be stated as follows.

Proposition 1.4 *In a theoretical market model, absence of arbitrage is equivalent to the existence of a probability measure that assigns positive probability (exclusively) to all possible market events and under which the discounted asset prices are martingales.*

This was a somewhat complicated way of saying the same thing as before, but the bonus is that, at last, we can unmask the mysterious Risk Neutral Measure: it is the hypothetical $\tilde{\mathbb{P}}$, which in the present situation boils down to just that little $\tilde{p}$. I reiterate that the risk neutral measure depends only on the possible outcomes of the risky indices and requires that these be fully specified. It is virtually independent of what probabilities, if any, we might assign to the possible outcomes in the real physical world. It is, therefore, of little interest to investors who make systematic profits from trades where gains are more likely than losses.

There is, however, a reason why the Risk Neutral Measure is a useful theoretical concept, namely that it admits a constructive representation of the theoretical price of any financial derivative as I shall now explain. Proposition 1.3 and the remark associated with it say that the price at time t of a derivative with pay-off $H(\mathcal{F}_T)$ at term T is the current value U_t of the SF strategy that duplicates the payoff. The same is true if we work with discounted values. Now, since the discounted value of the SF strategy is a martingale under the Risk Neutral Measure, we have

$$\tilde{U}_t = \tilde{\mathbb{E}}[\tilde{U}_T | \mathcal{F}_t],$$

and since $\tilde{U}_t = (1+r)^{-t} U_t$ and $U_T = H(\mathcal{F}_T)$, we arrive at

$$U_t = \tilde{\mathbb{E}}\left[\frac{1}{(1+r)^{T-t}} H(\mathcal{F}_T) | \mathcal{F}_t\right].$$

I summarize:

Proposition 1.5 *At any time the theoretical price of a derivative is equal to the expected present value of its pay-off under the Risk Neutral Measure.*

1.6.6 Completeness

A market where every financial derivative can be perfectly hedged by a duplicating SF portfolio, is said to be *complete*. Proposition 1.3 states that the toy market considered there is complete. An examination of the proof, especially Eqs. (1.34) and (1.35), reveals the crux of the matter: there are as many freely tradeable assets as there are possible states of the economy at any time (two). The model extends in a straightforward way to a situation where the market can be in n possible states at any time, and where there are m assets: the bank account with fixed interest rate and $m - 1$ risky assets with variable interest rates depending on the state of the economy. Instead of (1.34)–(1.35) one now gets n equations with m portfolio weights. If $m = n$, these equations typically admit a unique solution that determines a duplicating investment strategy, hence the market is complete. In this case one also determines a unique equivalent martingale measure along the same lines as above. If $m < n$, then there is no solution in general, and additional costs are needed underway in order to hedge the pay-off. It turns out that, in this case, there exist infinitely many equivalent martingale measures (candidate risk neutral measures). These sketchy considerations can be summarized as follows.

Meta-Proposition 1.6 *A theoretical market is complete if the number of assets is equal to the number of random sources, and in this case the equivalent martingale measure is unique. If there are fewer assets than random sources, then there are infinitely many equivalent martingale measures.*

The qualifying term "theoretical" is meant to underscore that this and other results presented here are valid in the tiny world of the chosen mathematical model. As I have indicated elsewhere in this chapter, the real world markets are a different ball game, probably with arbitrages hence no risk neutral measure.

1.7 Some Bullet Points on the New Paradigm

1.7.1 The Crystal Ball

The binary market model considered in the previous section is overly simplistic. To borrow a phrase from the managers speaking in Chap. 11, it is a caricature of the real world. Now, financial mathematics boasts a huge collection of advanced mathematical models based on Brownian motion, Lévy processes, jump-diffusion processes, etc. The problem is that all these models, however sophisticated, are caricature representations of the real markets. This goes a fortiori for the models used in modern life insurance mathematics based on financial mathematics since they need to be realistic under exceedingly long time horizons. So why is it that those finance models and methods have become the theoretical basis for the new regulations under the new paradigm? The composite answer can be gathered from the previous sections. Clearly, option pricing and hedging is all about forecasting, and this could not possibly be different. Therefore, theories for option pricing and hedging have to build on models for the future development of market indices, and the same goes for life insurance mathematics built on theories for option pricing and hedging. The reason why management of economic and demographic risk in life insurance and pensions has become a matter of forecasting, is the guarantee culture based on the perfect market delusion that hedging opportunities exist for even the most audacious financial contracts. Actuarial science under the new paradigm is crystal ball gazing, and the serious predicament is that new regulations have set this in stone. I reiterate that the with-profit scheme of the old paradigm and the index-linking idea presented in Sect. 1.5 do not require modelling of future economic and demographic indices. Galbraith: *There are two types of forecasters: those who don't know and those who don't know they don't know.*

1.7.2 Securitization

Proposition 1.6 suggests that an incomplete market can be completed (at least partly) by the introduction of derivatives with prices linked to the risk factors governing non-hedgeable financial claims. The stated purpose of securitization is to offer, on the one hand, diversification opportunities for general investors and, on the other hand, hedging opportunities for parties to risky contracts, e.g. banks and insurers. The idea is that the risk carried by these parties can be transferred to the financial markets, which on a global basis have enormous capacity. For instance, the losses incurred by major catastrophes like hurricane Andrew or the World Trade Center atrocities are only a fraction of the typical daily change in the total capitalization of global financial markets. Now, introducing new tradeable securities also means adding to the pool of financial instruments that investors can speculate in. It is therefore clear that securitization can serve its primary purpose well only if the underlying risk does not admit too much asymmetry of information, market manipulation, and other anomalies for which the existing markets are notorious.

The past two decades have seen the advent of catastrophe derivatives with pay-offs linked to indices for natural catastrophes (floods, earthquakes, hurricanes). Such "CAT" securities exist in many different forms and are largely a success story: they are liquidly traded, serve the combined purpose of diversification and hedging, and meet the essential requirements for asymmetry of information.

1.7.3 Longevity Risk on the Market?

Securitization of systematic mortality and longevity risk is currently gaining momentum in practice and is drawing increasing attention from contributors to the professional and scientific actuarial press. A fairly up to date overview is offered in Sect. 1.6 of Cairns (2008). *CAT mortality bonds*, aiming to offer hedging opportunities for life offices against catastrophic mortality risk in the short term, have fared reasonably well. A *longevity bond*, aiming to offer hedging opportunities for annuity businesses and pension funds against uncertain long term mortality developments, did not sell at sustainable prices and was taken off the market. The mixed experience reflects certain differences between mortality risk and longevity risk to be described next.

The CAT mortality bonds have much in common with their CAT predecessors in general insurance. They are short term and related to limited losses that can comfortably be absorbed by the global capital markets. The underlying risk is caused by purely random phenomena that can be reliably described by a mathematical model calibrated to historical data.

Longevity bonds are different. They need to be exceedingly long term and are related to potential losses at a macroeconomic scale. The underlying index cannot be predicted in the long term from statistical data, but in the shorter term it can to some extent be anticipated from current collateral information to which investors will not have equal access. Thus, there is asymmetry of information that can be exploited by the big players. Major investment banks and reinsurers have positioned themselves with research departments employing experts in an interdisciplinary mix of social policy, demography, medical science, actuarial science and statistics, collating statistical and other information about the current developments in mortality, down to the individual causes of death and within fine segments of the population and the pension industry. All this, not for the betterment of our shared human condition, but for the purpose of making money on superior information, expertise, and information technology: it does not matter in which direction our longevity prospects are moving, all that matters is to know it a bit better and a bit earlier than the other parties in the market. Thus, to the usual pros for securitization of longevity risk there are certain cons to be added that cast doubt on its practical feasibility and on any attempt to analyze its working with traditional financial mathematics models.

Now, securitization is currently not the main vehicle for transfer of longevity risk to the markets. It happens through over-the-counter arrangements between banks and reinsurers on the one hand and pension funds and small direct insurers on the other. Their bargaining is guided by information and expertise (of which the former party has a lot and the latter has little), pieces of information (and rumours) as to what prices some competitors out there might ask or bid, deliberations over business goals and strategy, negotiation skills, and a lot of gut feeling. They might also gaze into their crystal balls, and after the deal is struck they might even gaze into the arbitrage pricing scriptures just to be reassured they were guided to the fair price by the invisible hand of the efficient market (the Risk Neutral Measure).

Actuarial theory for hedging of longevity risk is about forecasting, and it is interesting to see how academics are approaching this hopeless task. Mortality is the resultant of highly complex processes—societal, environmental, scientific, and technological—a mix of trends, sudden irreversible shifts, and purely erratic movements. It lacks completely the stationarity and reproducibility that is needed in statistical forecasting. This notwithstanding, actuarial journals are currently being filled to the brim with papers telling us that mortality is stochastic (which I knew), and that future mortality can be extrapolated in statistical models fitted to death rates in the past (which I didn't know and am surprised to hear). The investigators are observing the mortality phenomenon from the moon, using aggregate population statistics and disregarding all collateral information about individual causes of death available in the Lancet, in Science, and in the newspapers for that sake. The predictive capacities of the models are checked by back-testing, which means fitting the model to data for the period 1950–1980 (say) and comparing its predictions for the period 1980–2010 with actual experience. May I suggest the following: fit the model to mortality data for the twentieth century, extrapolate backwards to the eleventh century, and conclude that the guys who fought the battle of Hastings were some 5 years old on the average.

1.7.4 The Piggy Bank

Longevity risk and interest risk were not big issues in the old paradigm as long as it was implemented properly. They have been created in the business hustle of the new paradigm. Market management of longevity risk has become an issue because life offices and pension funds created non-diversifiable risk through the very design of their products and schemes: they promised too much. The reader should contemplate the following. A pension fund is essentially a savings account, only with the special features that it is a joint account for the members of the scheme and that the survivors share the withdrawals. Arguably, the pension industry is the greatest accumulator of capital on the planet. And now it finds itself in the bizarre situation that it is in deficit (!) and needs to be rescued by the speculators. The backdrop of deregulation, frivolous guarantees, and overreliance in the workings of the market is described elsewhere in this chapter.

I was recently at a longevity conference where a missionary from the City of London triumphantly announced that so far 250 billion pounds worth of pension liabilities have been transferred to banks and reinsurers through longevity swaps. There was no mention of the cost this has incurred and that has to be taken from the pension savings. Also there was no mention of the counterparty risk: obviously Lehman Brothers and all the other big gamblers that went bust in the crisis of 2008 are not in the longevity market. I shall round off with an allegory of what I consider to be the biggest scandal in contemporary insurance history.

Huey, Dewey, and Louie get weekly pocket money from their uncle. Huey, planning for future consumption of ice cream, feeds coins into his piggy bank. Dewey and Louie, impatient to consume ice cream excessively and now, gamble with their money. At some point Huey smashes his piggy bank, counts its contents, and realizes that—alas—he is in deficit! The piggy bank did not earn him as much interest as he had assumed, and he cannot consume as much ice cream as he had promised himself. In fact, the interest rate was zero just like in ordinary customer banks nowadays. Dewey can help out because he just gambled successfully (Louie cannot because he gambled unsuccessfully and is under the uncle's tutelage). Dewey offers Huey a guarantee against the risk of negative interest in his next piggy bank. At a certain price, of course, which can buy Dewey ice cream now.

1.7.5 Insurance in the Future

It is difficult to make predictions, especially about the future. This quote, due to the physicist Niels Bohr, concurs with those of the economist Galbraith and testifies to their universality. I shall, obviously, not lapse into prediction of the future of insurance, which will be shaped by our individual and collective action anyway. What I will do, and have attempted to do in this chapter, is to try and understand the ongoing crises—in finance, in insurance, and in their mathematical theories—in order to see what action can be taken to get us out of the mess we are in. It is not a matter of changing our mathematical models—adding more heavy-tailed distributions, more dependencies, and more of everything that can enable us to measure more accurately the risk we have created. It is also not a matter of changing our philosophy—adding more about the unpredictable, more about the highly improbable (the Black Swan), and more about everything that can prepare us mentally for the next crisis, which will almost surely be created by the risk we have created, the only unpredictable about it being the time of its occurrence, the form it will take, and the damage it will inflict on the ordinary citizen. Getting us out of this is a matter of reinstating proper governance and regulation and of restoring plain common sense. To quote a famous political economist: *The philosophers have only interpreted the world, in various ways; what matters is to change it.*

Acknowledgment The author thanks the BNP Paribas Insurance Chair "Management de la modélisation" for financial support. The views expressed in this chapter are the author's own and do not necessarily reflect those endorsed by BNP Paribas Insurance.

Chapter 2
About Market Consistent Valuation in Insurance

Pierre-E. Thérond

Abstract The latest developments of both prudential (Solvency II) and financial reporting (MCEV, IFRS) frameworks seem to consecrate market consistent valuation as a kind of paragon of insurance liabilities assessment. In this chapter, we initially try to analyze the underlying motivations of this evolution. We show that it results from an objective of harmonization of measurement of quite different insurance contrats. This heterogeneity being the result of heterogeneous national insurance regulations. In the second part, we analyze the limitations of this measurement principle. For that, we mobilize some of the arguments opposed to Fair Value Accounting. Moreover, we insist on the limitations resulting as well from the implementation issues as of their use in a risk management perspective.

2.1 Introduction

Since the late 1990s, insurers have had to deal with significant developments in the valuation standards of insurance liabilities. Whether their purpose is financial communication or to measure solvency, these valuation frameworks all rely on assessment principles of risks borne by insurers that are consistent with the market. Thus the objective sought is twofold. An investor's initial reasoning is: valuations based on market consistent assessments are supposed to provide information that is useful for making investment choices by rendering them comparable to other types of investments on the markets. The second objective is harmonization. Indeed, by requiring a market coherent assessment, the standard setter (whatever it may be) hits two birds

P.-E. Thérond (✉)
Galea & Associés, 12 avenue du Maine, F-75015 Paris, France
e-mail: pierre@therond.fr
URL: http://www.galea-associes.eu
http://isfa.univ-lyon1.fr

P.-E. Thérond
Univ Lyon – Université Lyon 1 – ISFA – LSAF, 50 avenue Tony Garnier, F-69007 Lyon, France

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*, EAA Series, DOI 10.1007/978-3-319-29776-7_2

with one stone: he avoids having to personally determine what is a proper measure of risk, and ensures, in theory, the consistency of this measure, therefore reducing the arbitrariness that would leave risk levels up to freedom of choice. On this second point, insurers intervening on the same market will observe the same market price and therefore, in principle, use the same measure in assessing these risks.

If this development is in line with the dual harmonization process and relies on the same observable market data used in accounting and, more broadly, in financial communication since the end of the 1970s, it leads, in the field of insurance, away from the pre-existing accounting valuation principles.

The problem for which we seek answers here is twofold. Initially we must try to understand the motivations that led the market consistency principle to appear as a kind of paragon of insurance assessment, following in a more comprehensive manner that of the fair value method. For this, we will review one of the specificities of the insurance business: its strict regulations. We will observe the consequences of its structure and the close link between regulations and local accounting practices. Beyond the criticism of market values, and more broadly fair value as a measure of economic activity, we will look at how the transposition of the coherence approach of the markets to the insurance business faces additional pitfalls.

2.2 A Business Structured by Regulations

In this first section, we will try to explain the reasons that led to the advent of market consistent assessments to measure the insurance business. For this, we will review one of the specificities of the insurance sector in general, and specifically life insurance: its heavy regulation by public authorities. This has two major consequences: an overlapping of contract law, accounting and solvency regulations leading to significant diversity between various jurisdictions.

Indeed, given its role in economic development and safeguarding economic actors, the insurance business, regardless of which developed country, is subject to specific rules designed to ensure the strength of its players. Two complementary types of approaches to this regulation can be distinguished. The first consists of supply-side constraints, i.e. the definition of insurance products that may, or may not, be marketed within a given jurisdiction. The second involves regulating the organizations allowed to market these products. This second type of regulation can itself be subdivided into a priori restrictions (whether or not other commercial or financial activities may be offered, maximum exposure to certain market risks, legal form of the companies, and governance methods, etc.) or a posteriori restrictions linked to risk exposure measurements and their level of danger (the solvency margin of the European directives of the 1970s, America's Risk-Based Capital, and the solvency capital of Solvency II, etc.)

Some of these regulations, derived from prudential considerations, have led to a considerable structuring of the supply and functioning of insurance products. When these provisions stem from national legislation, they can lead to significant discrepancies, depending on the jurisdiction, the product, the risks and more generally on the management of insurance companies providing basically similar products.

In this regard, a striking and perhaps the most vivid example is life insurance in the form of savings in euros.

2.2.1 An Illustration of Life Insurance in Euros: The French Case

Often referred to by the press as a Frenchman's preferred investment, life insurance represents more than half of all long-term savings (source: French Federation of Insurance Companies (FFSA) 2014 annual report) and more than 37 % of the financial wealth of French households (source: 2014 annual report of the Observatory of Regulated Savings); the mathematical provisions of life insurance contracts and their capitalisation alone represent over €1.3 trillion.

Having their own specific fiscal framework, these life insurance products in euros are subject to special regulations resulting in particular from prudential considerations. Savings and retirement contracts in euros provide guarantees in the event of life (regulation of savings over time or payment of a life annuity) and in the event of death (payment of savings made to beneficiaries and possibly premium death benefits for annuity contracts). They are taken out by the insured party from a long term investment perspective in order to benefit from a capital distribution or a life annuity (retirement supplement). One of the main expectations of policyholders is to see their savings increase under favourable tax treatment and the ability to redeem the contract at any moment. To provide this service, the amounts collected by insurers are invested on the financial equity and debt markets, as well as in real estate.

In this context, legislation has set conditions on what may be provided to the insured parties. In particular:

- These contracts may not benefit from a minimum guaranteed revaluation rate that exceeds 60 % or 75 % of the semi-annual reference to the average rate of government bonds[1] (C. ass., Art. A132-1) upon subscription;
- The insurer is required to use the statutory mortality tables or those based on the regulations (in the case of experience-based mortality tables) when setting tariffs (C. ass., art. A335-1);
- The compensation due to the insurer in case of early redemption of the savings may not exceed 5 % of the acquired savings and must be zero if redemption occurs over ten years after the effective date of the contract (see C. ass., Art. R-132-5-3).

[1]More precisely, the rate of 60 % (with a maximum of 3.5 %) applies beyond eight years and, for the contracts with periodic premiums or variable capital, whatever their duration.

If the maximum adjustment rate constraints appear, at first glance, rather strong, it would mainly result in taking insurers out of the savings market if it was not supplemented by provisions for profit sharing (from regulations contractually more favourable) and the insurers' methods in this area. Indeed, in consideration of the limitations on the revaluation promises offered at the time of subscription, regulations require allowing the community of life insurance holders in euros to share in the technical and financial results (C. ass., Art. A331-3 et seq.). This obligation does not, however, provide any individual rights to the insured party.[2]

In particular, the insurer has a certain latitude in the way he can redistribute these profits to policyholders, both on the maturity and on the allocation between the various contracts. This results in a relatively widespread practice of providing more benefits to those with low guaranteed minimum rates to improve their overall valuations.

What interests us here is the close link, stemming largely from prudential considerations, between the contractual terms and conditions, the economy of the contract and the accounting. Indeed, there are no profits without an accounting system to determine its extent. In the case in question, these benefits are determined by the accounting benchmark—the French insurance accounting standards[3] (PCA).

We will not detail here the provisions of the PCA, instead we propose to focus on the main provisions of life insurance contracts in euros. They can be schematically summarized as follows:

Technical provisions are made up of mathematical reserves and a provision for profit sharing:

- The mathematical reserve provisions are individually determined and correspond to the acquired savings (sum of premiums paid net of fees, including the various annual revaluations, net of fee income, which were incorporated into the insurance contract—at the minimum these annual revaluations are equal to the minimum contractual revaluation rate) for savings contracts or expected present value of future benefits for pension contracts in service;
- The provision for profit sharing is broadly established and relates to amounts from the technical and financial benefits accruing to the insured parties but was not incorporated into the contracts;

Investments are recorded at their historical cost[4] or at their amortized value for amortizable securities.

To maintain an overall consistency, the regulations set limits on the assessment of the mathematical provisions. In particular, their evaluation must include certain

[2]This was recently recalled by a judgment of the Final court of appeal (Civ. 2nd, March 5th, 2015, n° 14-13.130) confirming a preceding decision of Conseil d'État (CE, May 5th, 2010, n° 307089).

[3]Until 2015, the insurance accounting rules were part of the Code des assurances, like most of the above-mentioned regulatory or legislative provisions. From 2016 and the coming into effect of the new prudential framework Solvency 2, these rules will be transferred to the national accounting standard-setter (Autorité des Normes Comptables).

[4]This measurement is associated with impairment rules for financial assets whose value fell significantly to a point at which the perspective of recovering the acquisition value is unlikely.

assumptions (discount rates and mortality tables) consistent with those of the tariff. Thus, it follows that if the financial performance of the investments[5] is greater than the interest expense from the accretion costs of interest rates, this necessarily creates financial income from which the insured community can benefit for at least up to 85 %. The logic is the same regarding mortality assumptions: an overly conservative assumed tariff compared to the insured portfolio year after year will lead to mathematical reserve bonuses that will increase technical profits, which will return for at least up to 90 % back to the insured community.

The analysis of this regulation demonstrates the overall consistency of the system, especially the link between the insured party's contractual rights and the insurer's commitment to the insured community. The latter is more important than the sum of individual rights. Furthermore, its development is partly a consequence of the insurer's management of its investments (through their allocation, but also their timing) as well as the revaluations made beyond the guaranteed rates (i.e. the individual rights of the policyholders). All this occurs within a competitive environment with the possibility of policyholder redemption and could significantly penalise a low bidding insurer in terms of revaluations.

Another important aspect, which will be discussed later, is that through the accounting of the acquisition cost of investments and the provisions for profit sharing (which allows accruing policyholder profits for a maximum of eight years), this system allows the insurer to spread out redistributions. In practice, this allows the use of previously existing reserves to continue re-evaluating when the markets are down and offset policyholders benefiting from the lowest guaranteed rates, meeting certain expectations of investors as identified by Séjourné (2006, 2007).

2.2.2 Comparative Analysis: The German Situation

The German insurance market offers life insurance policies in euros very similar to those presented in the previous section: a guaranteed rate, a profit sharing clause, an annual ratcheting mechanism, and a similar redemption value to the mathematical reserve, etc. However, an analysis of their regulations enables us to highlight clear differences that lead to problems in Germany different from those encountered in France. The main provisions of these regulations are summarized in Kling et al. (2007), Maurer et al. (2013) and in Berdin and Grundl (2015). In particular, Berdin and Grundl (2015) presents the regulatory reform carried out in 2014.

First of all, let us look at guaranteed rates. For this purpose, Eling et al. (2012) proposes a comparison of regulations on the matter between jurisdictions in the US, Germany, Austria and Switzerland. Regarding Germany, the Federal Ministry of Finance determines the maximum technical interest rate pursuant to

---

[5]In our schematic approach the financial incomes thus consist of the incomes from amortizable assets and the dividends and gains or losses on ceded non-amortizable assets.

Directive 92/96/EEC, on the basis of 60 % of the average of the rates on government bonds. However, there are, as is also the case in France, specific provisions for certain contracts. For example, for single premium savings contracts,[6] the maximum guaranteed rate over eight years is 60 % of the historical average rate of German government bonds (bond maturities between 9 and 10 years). Until the late 2000s, this average was based on a ten year track record. This has been reduced to five years to better reflect the lower interest rates.

Up until the sovereign debt crisis, the practice of German insurers was, as Berdin and Grundl (2015) recall, to guarantee the maximum rate permitted by law.

These regulatory provisions, combined with the practices of the insurance market, put the German life insurers into a difficult situation when interest rates fell. That is the finding made in particular by Kablau and Weiss (2014). They show in particular that at the end of 2012, the market's average technical interest rate was 3.2 % when the average rates of German sovereign debt stood at 1.3 %.

These contracts also contain profit sharing provisions (at 90 %), mortality (at 75 %) and management (at 50 %). The amounts attributable to policyholders can be directly incorporated into contracts to the allocated or non-allocated provisions. Reference should be made to Maurer et al. (2013) for details of the profit sharing mechanism. These provisions are supplemented by the sharing of at least 50 % of unrealized gains on investments (valuation reserve) representing commitments to policyholders leaving a mutual insurance company.

On aspects of profit sharing and unrealized capital gains, the 2014 reform had two significant impacts. The first was to increase from 75 to 90 % the rate of technical profit sharing. The second was to limit the proportion of capital gains on bonds within the base of the valuation reserve. Indeed, only capital gains on bonds, in addition to those required for future compliance with the guaranteed rate commitments, are now integrated in the profit sharing when exiting a contract.

The comparison of French and German regulatory provisions calls for several comments.

The first is the similarity of the existing relationship between accounting provisions (how investments are valued, recorded results) and the economy of the contract. This makes it difficult to define a common accounting framework to act as a national accounting framework, to prepare financial statements and to define profit sharing between insurers and policyholders.

The second comment observes that, although the general logic is the same in both jurisdictions, setting parameters, in particular the maximum interest rate that the insurer can guarantee, leads to rather distant problems in terms of risks. If in 2014 three quarters of the technical life insurance rates are zero and more than 85 % below 1 %,[7] then the average guaranteed rates in 2014 on the German market are greater than 3 %.[8]

[6]These rules also apply to annuity contracts with no surrender options.

[7]Source: *Étude sur les taux de revalorisation des contrats individuels d'assurance - vie au titre de 2014*, Analyse et synthèse n° 47, ACPR, juin 2015.

[8]Source: Moody's.

It follows that if one of the major concerns of French life insurers is maintaining the competitiveness of life insurance while the projected financial returns of their investments decline as a result of the prolonged low interest rates, their German counterparts are more concerned about their ability to immediately honour their guaranteed rate commitments to their policyholders. This results in insurers and supervisors increasing their spending on research and development. Analysis and modelling of policyholder redemption, payments and arbitration behaviour is now an area of particular interest among French life insurers. Expectations made in this area help to determine revaluation policy. A key question is that of the market's elasticity to interest rate variations and, in particular, to offers that reflect the rate at the time proposed by banks (savings accounts, term accounts, etc.) This question is much less central in Germany, where the main problem appears to be that of optimizing financial management (allocation, realization of capital gains) to reduce risk while meeting commitments. A second stage is the changing regulatory framework, as illustrated by the 2014 reform (lowering of guaranteed rates, limiting the rights of redeeming policyholders to valuation reserves) or the transitional measures obtained on the implementation of Solvency II (a sixteen-year transition period obtained by the German market to converge with the best estimate market consistent provisions).

2.3 Fair Value and Its Criticism

At this stage, it appears useful to us to reconsider the concept of fair value which was essential in past decades in the accounting countable frameworks (IFRS and US GAAP in particular) and which provides the conceptual background to the market consistent measurement of insurance portfolios. The use of fair value to measure balance sheet items (assets and some liabilities) has been increasing to culminate at the time of the financial crisis of 2008. The description of this historical movement through accounting frameworks is beyond the scope of this text. The reader will be able to refer in particular to Power (2010), Emerson et al. (2010) and Richard et al. (2014). These authors propose a historical overview of the advent of fair value and analyze the reasons for its rise to power.

In this section, after having pointed out what fair value is and how it is put into work in accounting frameworks, we will give an overview of those criticisms which are opposed to it.

2.3.1 One Fair Value Principle, Several Valuation Techniques

The last definition of fair value appears in IFRS 13, published in May 2011 by IASB at the conclusion of a collaborative work with the American accounting standard-setter (FASB), leading to a single definition and a common practice.

> This IFRS defines fair value as the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date.

This definition is supplemented by provisions on the transaction concerned:

> A fair value measurement assumes that the asset or liability is exchanged in an orderly transaction between market participants to sell the asset or transfer the liability at the measurement date under current market conditions.
>
> A fair value measurement assumes that the transaction to sell the asset or transfer the liability takes place either:
>
> (a) in the principal market for the asset or liability; or
> (b) in the absence of a principal market, in the most advantageous market for the asset or liability.

We have to consider several cases in the effective measurement of an asset at fair value.

The first situation, the simplest, is that of a financial asset which is indeed exchanged, under normal conditions, on an active market. In this case, the fair value corresponds to the value of exit (exit been worth) without adjustment due to costs of transaction (which are in the scope of other IFRS).

When the considered asset is not in this situation, the fair value will have to be estimated by means of techniques of evaluation. It is thus advisable here to recall that the fair value is therefore not synonymous with market value on an active market. In particular, this allows a financial asset which was not systematically exchanged on an active market (at acquisition) to see its value annihilated in the balance sheet of the company, when there are no more exchanges on this security. The standard specifies that the three most widespread techniques of evaluation are: approach by the market, approach by the costs, and approach by the result. Whatever the technique(s) used, the objective of the evaluation remains to estimate the price on which the participants in a normal transaction of sale of the asset would have agreed. This scenario is directly inspired by the neo-classical economic theory as underlined in Richard et al. (2014).

The approach by the market consists in being based on the prices and other observations generated by market transactions on identical or comparable assets. The approach by the costs consists in determining the replacement cost, at the date of evaluation, which would make it possible to replace the value of service of the asset considered. Finally, the approach by the result consists in estimating the value which the market would grant to future amounts (future flows of treasury or products or loads). It is this last approach which will be of interest to us, taking into account its conceptual proximity with the market consistent valuations required to measure insurance portfolios.

Whatever the adopted approach, IFRS 13 requires us to detail in the appendices the fair values obtained, segmented according to a hierarchy (level 1 to 3) corresponding to the informational characteristics of the data used to carry out the evaluations. Thus the data input of level 1 correspond to directly accessible information on active markets, like market prices. The data input of level 2 correspond to information other than the market price, but which remains observable (directly or indirectly). Finally, Level 3 corresponds to the unobservable data. This hierarchy obviously presents the order in preference to considering the data simultaneously, for example, level 2 and 3 data are available at the same time. As an illustration, the corrected estimate of the price of the risk of expected inflation that allows us to estimate the observation of the price of the sovereign obligations indexed on inflation will take precedence over the need to estimate of the fair value which in turn takes precedence over estimates of the chief economist of the company.

In summary, it is advisable to recall that there is not a strict equivalence between fair value and market value. In the situation where there is no value of active market, the fair value is the solution of the problem which consists in determining what this price would be, taking into account available information.

2.3.2 The Criticism

The criticism of the use of fair value in accounting is registered on several levels. The first attempts to express doubts on the relevance of the use of market values (or resulting from observations of market) to reflect the good vision of the financial situation and the performance of a company. The second is that of the complexity of implementing the valuation methods and the lack of objectivity which can result from their use. The third level, most recent, concerns the macro-economic impact and the consequences on global financial stability. The objective of this section is to give a short outline of these criticisms. We will see in the next section which of these criticisms also hold for market-consistency of insurance portfolios.

Regarding the first register of criticisms which one can address to fair value, one can count, initially, the more general objections opposed to the neo-classical theory and its developments (cf. for example Orlean (2011) or Walter and Brian (2008)) on which rests the principle of fair value. In particular, Orlean (2011) and Walter and Brian (2008) detail how the market value can move away from the (non-observable) fundamental value. This divergence can take several forms.

The first of these forms are market noises caused by an informational asymmetry or disturbing agents which do not behave rationally (as assumed in the efficient information theory). In the long term, one would expect these disturbances to disappear.

A second phenomenon, perhaps more problematic, comes from the historical observation of financial bubbles. These would come in particular from the phenomenon of *prophétie auto-réalisatrice*. This concept is easily illustrated by means of the stock exchange quotations. Since the expected income of an investor on a stock primarily comes from its resale, rather than from the future dividends (as in the model of Bachelier (1900)), the maximization of the profit comes more from the anticipation of the evolution of the market value of the stock than from real economic outlooks for the company. Thus an investor who estimates (at the conclusion of a financial analysis, for example) that a title is overpriced may rationally find it beneficial to be a purchaser of the title if it anticipates that the other agents on the market will invest on the title. This mechanism of intersubjectivity is particularly well described in Orlean (2011).

Beyond these considerations relating to the quality of the stock exchange prices, to give a faithful impression of the value which one describes as fair, the question of the relevance of a value of exit arises to measure the financial situation of a company. One of the major arguments in favour of fair value rests on the assumption that the principal recipient of the financial statements is the investor. Thus evaluations based on observations of market must make it possible to lead to values comparable with the other types of investment (of market) which it can choose. For this purpose, it should be noted that if the investors appear indeed, in principles IFRS, like the paramount users, the fair value is used in other countable reference frames (ex: US GAAP) without the things being also clear. Thus, it is interesting to note that, as regards development of accountancies (cf. Richard et al. (2014)), the exit value was initially essential in measurements of solvency: if the company must fulfil its commitments today, does it have the assets enabling it to do it? If this question interests the investor, one might doubt that it constitutes the alpha and the omega of its criteria of appreciation. In particular, Holthausen and Watts (2001) estimate that, compared to what one can expect from an accounting framework, the fair value has a very moderate explanatory power (more particularly when used to measure non-financial assets).

The second register of criticisms is that of the distrust that investors may have regarding values obtained by means of valuation techniques. The issues in this respect are double.

The first kind of issue is that of the comparability of the amounts obtained by means of these techniques of evaluation. One single asset could be evaluated by several techniques (cf. the previous section) which, themselves, can be supplied with different data. Mary Barth, member of IASB during the year 2000, underlines in Barth (2006) the conceptual limits of this possibility and the difficulties which result from it for the users from financial statements.

The second kind of issue results from the phenomenon of moral risk. The share left with the models of evaluation as far as the fair value (when the asset is not directly traded on an active market) may give a manoeuvre margin to handling. It is imagined that this can be the case when, for example, the no-claims bonus or the promotion of the leaders of the company or the people in charge of these evaluations is conditioned by the results of the company (or of one of its segments). A historical example is that of Enron, whose financial statements had been artificially inflated by means of generous estimates of fair value of assets. This case is studied in particular in Haldeman Jr. (2006), Benston (2006) and Gwilliam and Jackson (2008). Beyond the identified behaviours (a posteriori), these authors point out the unreliability of the evaluations of fair value (including those done by external cabinets) and the strong asymmetry of information between the management and the users of the financial statements.

Finally, the last register of criticisms is that of procyclic financial stability and the character lent to the fair value. It is underlined in particular by the political world at the time of the meeting of G20 in London in April 2009, which requires an accounting global framework where the fair value is less prevalent and measurements of depreciation are less violent (cf. Azzaz et al. (2015) for details of the impairment methods of IAS 39 with regard to depreciation of actions accounted at FVTOCI). This in particular led IASB to re-examine the standard on the financial instruments and to publish IFRS 9, whose first application is scheduled for 2018.[9] IFRS 9 leads to a slackening of the strong constraints which meant IAS 39 was able to use the amortized value and refer to the concept of business model of the entity to decide if the titles (which are eligible for amortized value measurement) are measured at amortized cost, fair value through own-funds (FVTOCI) or fair value through income statement (FVTPL). On the other hand, it imposes measurement at fair value through P&L for the other categories of financial instruments (stocks, funds, etc.)

The reproach of procyclicity relies on the following consideration. The use of fair value results in reflecting in the assessments (and partly in the result) the variations noted on the markets. In particular, in the event of crack stock exchange, the image reflected by the use of fair value is affected in that the role of the fair value in the financial crisis of 2008 could be qualified as a catalyst in the transmission of the financial crisis to the real economy. In particular, Plantin (2008a) illustrates the effects of contagion that Fair Value accounting of Financial Institutions can produce. It is, however, interesting to note that they do not grant, on the matter, a better virtue to the historical cost accounting. Overall, it arises from work of Plantin et al. (2008), Laux and Leuz (2009–2010) and Jaggi et al. (2010) that if the use of fair value in times of crisis accentuates the problems of implementation and potential heterogeneity already mentioned, it is more its use at prudential ends for the banking institutions than its accounting purposes which could contribute to the effects of contagion to the real economy.

In summary, one can conclude that the benefits of fair value accounting are tarnished when they result from entity-specific techniques not easily controllable by the users of financial statements. This problem is accentuated in crisis periods (where the active character of certain markets is called into question) and can involve an increased distrust of the investors with respect to the financial institutions. However, as pointed out in Laux and Leuz (2009) and Plantin et al. (2008a), the alternative of the historical cost (coupled with measurements of depreciation) does not seem to be more relevant to the matter.

2.4 The Delicate Transposition to the Field of Insurance

The use of valuations relying on consistency with market observable data is in line with the introduction of fair value, particularly the IFRS. In this regard, it should be recalled that the first major principles of the Embedded Value benchmark (European

[9]This standard is currently under examination by the European authorities and has not yet been adopted by the European Union.

Embedded Value) were published by the CFO Forum in May 2004. The period is significant insofar as the IASB published in July 2004 an exposure draft (ED5 built on the current IFRS 4: Insurance contracts) on insurance confirming the path chosen in 2002 for the 2005 objective: a temporary short-term solution for the measurement and recognition of insurance contracts. The ambition to form a satisfactory accounting framework for insurance liabilities was postponed until 2007. Such, in any case, was the objective of the IASB at the time. At the time of writing, the new IFRS insurance standards (phase 2) have still not been published and will not apply before 2019. The use of Embedded Values for financial reporting remains topical for life insurance groups. Thus harmonizing practices through precise principles was required, giving rise to European Embedded Values to be supplemented in June 2008 by the publication of the principles of Market Consistent Embedded Value, which underwent a revision in October 2009 (financial crisis requires it).

Over the same period work started on the new European prudential framework: Solvency II. At the initiative of the European Commission in 2002, the project launch, in its initial specifications, relied on IAS/IFRS standards that were subject to a regulation on adoption, (see Regulation (EC) no. 1606/2002 of the European Parliament) and of the Council of 19 July 2002 on the application of international accounting standards. This was a delicate exercise for the regulator since IFRS 4 is a standard that derogates from the IFRS Framework allowing the continuation of local practices (with some adjustments) in terms of insurance contract accounting. In this exercise, the various quantitative impact studies (QIS) take up a number of expressions (Current Exit Value, Current Value Fulfilment) or principles (on contract limits) stemming from IASB discussions on phase 2 of the insurance standard. As work on phase 2 of the IFRS for insurance was not going anywhere, the QIS5 specifications published in July 2010 significantly stray from the phase 2 insurance ED published that month, in terms of evaluation principles of the technical provisions. Schematically, we can say that if Solvency II specifies the principles of valuation of investments and technical provisions, the prudential benchmark uses the IFRS assessments for other asset and liability items, with adjustments if IFRS principles have amortization processes to smooth the impact on the income statement. The fact remains that, ever since the QIS2 was published in 2006, Solvency II provides that the evaluation of the best estimate of technical provisions should be market consistent.

The resumption of work of IASB on standard insurance (phase 2) was concretized by the exposure-drafts of 2010 and 2013 and the standard is expected for the end of 2016. There still (cf. the appendix) the required evaluation of insurance portfolios must be coherent with the market data. It is, however, advisable to note that the concept of current fulfilment been worth is not exactly equivalent to that of fair value. In particular, the notion of transaction or transfer is replaced by that of execution of the commitments. Moreover, it is explicitly a question of measuring the current value of future flows of treasury necessary for the achievement of the commitment of the insurer towards the policy-holders. One thus finds here explicitly a technique of evaluation based on flows (cf. the preceding section on the various techniques of evaluation of the fair value), even if the standard authorizes approaches of replicating portfolio type. There is no longer a reference to the market, in the spirit of the equalizer, ensuring comparability and coherence with the financial placements, themselves primarily measured with the fair value.

The principles of evaluation of these three reference frameworks are given in the appendix.

2.4.1 The Implementation

The search for a model allowing an analytical valorization of the contracts of life insurance has been at the core of academic actuarial work in the last few decades. Brennan (1976) opened up the way for the financial techniques of valorization on unit-linked contracts. As regards contracts of savings in euro, one can in particular quote the works of Briys and de Varenne (1997a), Bacinello (2001), Grosen and Jorgensen (2000), Ballotta (2005) and Planchet et al. (2011). All of these nevertheless run up against the complexity of the accounting mechanisms which we have described in the first section, such as the necessary modelling of the behaviour, both of the management of the company of insurance and of the policyholders (cf. in particular, regarding lapse behaviour of policyholders, Sejourne (2006), Planchet and Thérond (2007), Loisel and Milhaud (2011) and Borel-mathurin et al. (2015)).

It results from this that, concretely, most insurers use Monte Carlo techniques by means of stochastic models of projection in which these various behaviours are modelled.

Market consistency of these evaluations is assumed to be ensured by the coherence of the economic scenarios projected with the noted prices of the financial placements held by the insurers to whom the collective right of the policy-holders to the participation in the profits applies. Other risks, i.e. those which are not implicitly integrated in the estimate obtained by these techniques, are taken into account ex-post by the addition to the market of a consistent best estimate of an allowance for risk (risk margin of Solvency II, risk adjustment in IFRS or cost of the non-hedgeable risks in MCEV) of which the concept varies according to the framework. This segmentation is described in more detail in Therond (2008).

These reference frameworks thus operate a segmentation in the manner of treating the risks according to whether or not they can be measured by means of market data. In the second case, a measurement of risk (freely selected by the company in IFRS or constraint in Solvency II) must make it possible to reflect the allowance for risk which the market would require for the transfer or the elimination of these uncertainties.

If one again takes the grid of analysis of the fair value (in terms of techniques of evaluation and hierarchy of fair value), these evaluations use a technique based on cash-flows with data input of the various levels presented to the previous section. A simplified nomenclature is as follows:

- Level 1 data: the price of the financial placements (when they are quoted on an active market, which is generally the case taking into account the existing prudential constraints on the choices of investments of the insurers).
- Level 2 data: the interest rate curve obtained by the observation of the price of the products of interest rate (if liquid maturities are long enough compared with the duration of insurance liabilities), parameters of volatility and dependence of the stochastic models used to project various risks (interest rate, actions, etc.)
- Level 3 data: the modelling of the behaviours of the company (financial management, discretionary part of the revalorization of the savings) and of policy-holders (payment of the bonuses for the contracts with periodic bonuses, repurchase, death, etc.), the modelling of future costs (internal and external) the entity will have to support in order to fulfil its commitments towards the policy-holders, etc.

Consequently, this kind of measurement stands in the field of the criticisms arising from the asymmetry of information between the entity and the user of the evaluations, as presented in the previous section. Nevertheless, it is advisable to recall that, as for accounting financial statements, embedded value calculations are the object of a review of an independent cabinet and the Solvency II evaluations are controlled by the prudential supervisors.

2.4.2 Limits of the Approach

To evoke the limits specific to the insurance of such an approach, it is advisable, initially, to delay on the recipients of such financial statements.

For Solvency II, this is usually the supervisor. Also, the vision of type value of exit is understood relatively. This makes it possible to define the level surplus (equities) on a current and coherent basis with the prices to which the financial placements could be yielded. One could nevertheless regret the absence of a prospective size of this measurement, but this point is assumed to be regulated by the requirement of capital (Solvency Capital Requirement). It remains the case, however, that this measurement, by construction, takes into account the volatility of interest rates and the markets action, itself very volatile. This can result in unseating leaders and supervisors who see the tests of solvency increase or decrease in one period and in a consequent way in the other. The difficulty lies, in particular, in the analysis of this evolution resulting from the evolution of the portfolio, the concrete actions conducted by management and of course of the conditions of market.

Anti-cyclicity features have been implemented since the financial crisis to try to limit the effects of unfavourable conditions of market: volatility adjustments to reflect, in the technical provisions, bonuses of illiquidity of the market or effects of attenuation of the loads of capital for equity risks. In particular, the integration of a spread for illiquidity (whatever its name according to the reference frame considered) thus led to a projection, in so far as the insurer can hold his placements in the duration, of higher incomes than those which would have been obtained without taking it into account. That led in particular to outputs (thus countable) which more easily make it possible to respect the contractual commitments (guaranteed rates) and the satisfaction objectives of the customers (revalorization beyond the guaranteed rates to be used as a target rate of revalorization).

Some of the quantitative reporting statements of Solvency II are nevertheless public. They will thus be added to the publications of embedded been worth and IFRS financial statements in the double sack of the investors. For those dimensions the *black box* image of the evaluation models poses the problems already mentioned in Sect. 2.3. In particular, the level 3 data presented in Sect. 2.3. Indeed, they can be portfolio-specific (e.g. the mortality or the surrending behaviour of the policy-holders) and thus it is difficult to compare them. It can be a question of choice, modelling some financial variable with some given model. They can also reflect the future intentions of management. In this case, management may not wish to communicate them, especially as they could hypothetically lead to the creation of new commitments (within the framework of increasing jurisprudence in favour of the protection of the policy-holders). The investor thus finds himself with an extraordinary mine of information (much more than with national accounting) for which he will have to define the level of credibility that he grants to them.

Another major limit of the approach lies in the following paradox. If these various normative reference frames are based on a financial valorization, if it is not economical, it is necessary to precisely define what is evaluated and what is not. In accounting terms, it is a question of defining what does and what does not cause a liability. The been worth embedded, with its comprehensive view of the current value of the future profits intended for the shareholder under the portfolio of contracts in force, gives a coherent answer to this question, though raising more convention than prediction. Indeed, the scenario is like that of the absence of future production and thus of the assignment only to the policy-holders in force, at the date of evaluation, of part of the future results. In practice, as we saw in Sect. 2.2, things are more complicated because of intergenerational mutualization. One can indeed wonder about the relevance of this scenario which would lead the insurer to project structures of costs quite different from those that he knows and, in particular, to pose the question of a minor revalorization of the contracts.

This paradox is thorough with its paroxysm in IFRS, where the unit of account is the technical provision defined by the current value of the commitments towards the policy-holders under the contracts in force. It seems to us that the vision of IASB based on the legal commitment under the individual contract with the policy-holder plays a crucial role in the difficulties that meets IASB to work out and finalize standard IFRS insurance (phase 2). This reasoning is reconciled with difficulty with the provisions as regards participation in the profits and their properties of smoothing between the generations presented in Sect. 2.2. The level of pregnancy of these provisions under the current market conditions perhaps also explains the geographical heterogeneity of the degree of acceptance by the insurers of the proposals of IASB on the contracts known as participative.

2.5 Annex

This annex aims to provide the reader with a direct reading of the valuation principles consistent with market observations found in the benchmarks of Market Consistent Embedded Value, Solvency II and the IFRS insurance standard (phase 2).

2.5.1 Market Consistent Embedded Value (MCEV)

The principles below are extracts from the European Insurance CFO Forum Market Consistent Embedded Value Principles, published in October 2009.

Value of in-force covered business
Principle 6: The value of in-force covered business (VIF) consists of the following components:

- Present value of future profits (where profits are post taxation shareholder cash flows from the in-force covered business and the assets backing the associated liabilities) (PVFP)
- Time value of financial options and guarantees as defined in Principle 7
- Frictional costs of required capital as defined in Principle 8
- Cost of residual non-hedgeable risks as defined in Principle 9.

Principle 7 specifies the nature of the link between valuation methods that must be retained and observations on the markets.

Financial options and guarantees
Principle 7: Allowance must be made in the MCEV for the potential impact on future shareholder cash flows of all financial options and guarantees within the in-force covered business. The allowance for the time value of financial options and guarantees must be based on stochastic techniques using methods and assumptions consistent with the underlying embedded value. All projected cash flows should be valued using economic assumptions such that they are valued in line with the price of similar cash flows that are traded in the capital markets.

2.5.2 Solvency II

Article 76 of Directive 2009/138/EC states the principle of consistency with the information available on the financial markets.

Directive 2009/138/CE—Article 76 General provisions

1. Member States shall ensure that insurance and reinsurance undertakings establish technical provisions with respect to all of their insurance and reinsurance obligations towards policy holders and beneficiaries of insurance or reinsurance contracts.

2. The value of technical provisions shall correspond to the current amount insurance and reinsurance undertakings would have to pay if they were to transfer their insurance and reinsurance obligations immediately to another insurance or reinsurance undertaking.
3. The calculation of technical provisions shall make use of and be consistent with information provided by the financial markets and generally available data on underwriting risks (market consistency).

It is complemented by Article 22 of the Delegated Regulation (EU) 2015/35 of the Commission.

Delegated Regulation 2015/35—Article 22—General Provisions
3. Insurance and reinsurance companies determine assumptions about future scenarios and parameters of the financial markets that are appropriate and consistent with Article 75 of Directive 2009/138/EC. When insurance or reinsurance companies use a model to forecast future scenarios and parameters of the financial markets, it must meet all the following requirements:

(A) It generates asset prices consistent with those observed in financial markets;
(B) It does not assume any arbitrage opportunities;
(C) The calibration of parameters and scenarios is consistent with the curve of the risk-free rate used to calculate the best estimate referred to in Article 77, paragraph 2 of Directive 2009/138/EC.

2.5.3 IFRS Insurance Contracts (Phase 2)

The excerpts below are taken from the exposure draft of 2013 (updated preliminary decisions of the IASB through June 2015).

Future cash-flows
§22. The estimates of cash flows used to determine the fulfilment cash flows shall include all cash inflows and cash outflows that relate directly to the fulfilment of the portfolio of contracts. Those estimates shall:

(a) be explicit (i.e. the entity shall estimate those cash flows separately from the estimates of discount rates that adjust those future cash flows for the time value of money and the risk adjustment that adjusts those future cash flows for the effects of uncertainty about the amount and timing of those cash flows);
(b) reflect the perspective of the entity, provided that the estimates of any relevant market variables do not contradict the observable market prices for those variables (see paragraphs B43–B53);
(c) incorporate, in an unbiased way, all of the available information about the amount, timing and uncertainty of all of the cash inflows and cash outflows that are expected to arise as the entity fulfils the insurance contracts in the portfolio (see paragraph B54);
(d) be current (i.e. the estimates shall reflect all of the available information at the measurement date) (see paragraphs B55–B61); and
(e) include the cash flows within the boundary of each contract in the portfolio (see paragraphs 23–24 and B62–B67).

Time value of money
§25. An entity shall determine the fulfilment cash flows by adjusting the estimates of future cash flows for the time value of money, using discount rates that reflect the characteristics of those cash flows. Such rates shall:

(a) be consistent with observable current market prices for instruments with cash flows whose characteristics are consistent with those of the insurance contract, in terms of, for example, timing, currency and liquidity; and
(b) exclude the effect of any factors that influence the observable market prices but that are not relevant to the cash flows of the insurance contract.

These principles are supplemented in the annex by the following paragraphs (excerpts).

Uncertainty and the expected present value approach (§22)
§B40. The objective of estimating cash flows to measure the fulfilment cash flows is to determine the expected value, or statistical mean, of the full range of possible outcomes. Thus, the starting point for an estimate of the cash flows is a range of scenarios that reflects the full range of possible outcomes. Each scenario specifies the amount and timing of the cash flows for a particular outcome, and the estimated probability of that outcome. The cash flows from each scenario are discounted and weighted by the estimated probability of that outcome in order to derive an expected present value that is consistent with market variables. (…)

Market variables (paragraph 22(b))
B44. Estimates of market variables shall be consistent with observable market prices at the end of the reporting period. An entity shall not substitute its own estimates for observed market prices except as described in paragraph 79 of IFRS 13. In accordance with IFRS 13, if market variables need to be estimated (for example, because no observable market variables exist), they shall be as consistent as possible with observable market variables.

B45. Market prices blend a range of views about possible future outcomes and also reflect the risk preferences of market participants. Consequently, they are not a single-point forecast of the future outcome. If the actual outcome differs from the previous market price, this does not mean that the market price was 'wrong'.

Part II
Design and Implementation of Life Insurance Models

Chapter 3
Cash Flow Projection Models

Jean-Paul Felix

Abstract The cash flow projection model is simulating the way the insurance undertaking is working and reflects it by generating run-off balance sheets and P&Ls. This simulation implies a very thorough and deep understanding of all dimensions: products, assets, risk factors, markets, environments, regulations, behaviors, accounting... The challenge for the model is then to look at the whole undertaking "through the eyes of the management" and to encapsulate it and how things interact into formulas. This could only be done with a certain level of shortcuts: granularity of projected portfolios, behavioral laws…, in order both to reach acceptable running time and to be sustained by the IT capabilities. However these shortcuts should be carefully tailored so as not to false the results. Difficulties addressed in the article cover: the simulation of French multi-funds savings contracts where cash dynamically flows from General Fund to Unit-Linked funds (and vice versa), the granularity of asset classes, the simulation of structural cash flows from liabilities and notably the estimation of expenses, the stochastic modeling of assets, the determination of behavioral laws: policyholders' profit sharing, surrenders, fund shifting, investment policy, behavior in central scenarios versus behavior in extreme scenarios… Answers to these difficulties should take into account the requirements of consistency (backtesting), reality (how the model reflects the risk profile of the undertaking), rapidity (a model is designed to be used in production under deadline's constraint).

3.1 Introduction

Nowadays models are more and more extensively used to run Life Insurance companies. They are a must for insurance undertakings' managers to answer the increasing assessment issues they are facing in the strategic and day-to-day management of their companies. Indeed, in addition to traditional internal requirements

J.-P. Felix (✉)
BNP Paribas Cardif, Paris, France
e-mail: jean-paul.f.felix@bnpparibas.com

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_3

for a better understanding and monitoring of its activities and its risks (which imply running models), an insurance company must now meet new requirements associated with changing regulatory standards (Solvency II, IFRS4 Phase 2), that impose a new paradigm for assessments: the full fair value. This paradigm is the reason for the forthcoming massive use of models and for their increasing complexity. As in absence of markets, the full fair value of insurance contracts is to be a modelled value (called "marked-to-model") and not a quotation (called "marked-to-market").

The purpose of this chapter is to review the main questions that arise when developing models and that might be subject to internal debates. This list of questions is certainly not comprehensive; however it is a useful starting point both for model designers beginning from scratch and for decision-makers willing to get more insights into these strange animals that are models. When seeking answers to these questions, model designers should keep in mind, disappointing as it may sound, that a model always gives only a limited vision of the Company's activities. Their responsibilities are not to predict the future but to enlarge as much as possible, within the technical and cost constraints, the intrinsic limitations of the model to help decision-making. They should also remind themselves that a decision is rarely made based only on a model's outputs (see Chap. 11).

Yet models' outputs are useful. They are mainly a "marked-to-model" indicator that is the probability-weighted present value of future cash flows generated by what is being calculated (Fig. 3.1).

The structure of this chapter mirrors this definition. The first short section addresses the issue of discounting the present value, which is a key sensitive technical issue but not a modelling issue. The second section is about how probabilities are taken into account in the models. The third section, the longest, lists questions relating to cash flow modelling.

Model features are closely linked to the businesses they are simulating. In this chapter, the focus is on the following three business blocks, each based on different actuarial formulas: General Funds (GF), Unit-linked (UL) and Protection. In this chapter, General Funds and Unit-linked are grouped together to give a global

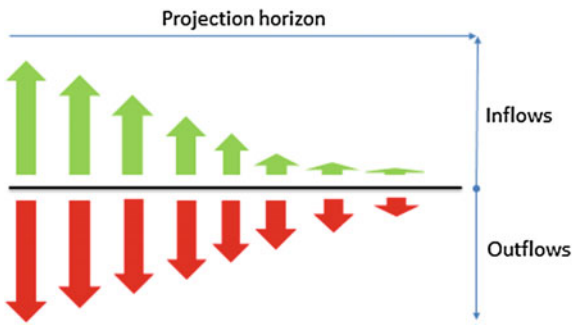

Fig. 3.1 Probability-weighted present value of future cash flows

picture of modelling issues for Savings. Protection is presented separately. Each topic discussed hereafter in this chapter is, when necessary, presented as broken down into these business blocks.

3.2 Discounting Issues

Discounting methods all take into account the two factors determining the time value of cash flows: the preference for immediate owning (do you prefer to own 1 now or €1 plus centimes tomorrow?) and risk aversion (do you prefer €1 with a 100 % probability or €1 plus centimes with a sub-100 % probability?).

The time value reflects the fact that €1 today is worth more than €1 tomorrow. Indeed, €1 today can be invested and produce interest. This potential interest income over time is materialized by the use of the risk free rate curve. The cost of the risk, i.e. the uncertainty, is reflected by the risk premium to be added to the risk free rate. In particular, in a risk-neutral environment, this risk premium is nil as it is considered that economic agents are risk-adverse.

The choice of the discounting rate is a key variable for cash flow discounting. It has a major impact on the outputs of the models. It is mentioned here not because it causes difficulties in the modelling itself, but because there are many hot debates on how to set it. Just look at the discussions on the volatility adjuster in the Solvency II framework… or the illiquidity premium considered by the CFO Forum in the calculation of the MCEV.

3.3 Probability Issues

In the world of life insurance modelling, there are two categories of contracts: with or without asymmetries. General Fund contracts usually embed asymmetries like surrenders, minimum guaranteed rates, policyholders' profit sharing… Most of these asymmetries are linked to financial markets' fluctuations. Guaranteed minimum death or life benefits (called GMxB) are the main option within Unit-Linked contracts. It is correlated to the timing of the insured event (death or survival at the maturity date) within the financial markets' cycle. Both General Fund and Unit-Linked options are theoretically (if not always in practice) hedgeable by financial instruments. Therefore stochastic simulations are required to assess them. In contrast, the main asymmetries of Protection contracts relate not to financial markets but to the allocation of the technical result between the insurer and its commercial network. They are non-hedgeable risks. Stochastic simulations are not necessary here.

The consequence of the existence of hedgeable asymmetries on the structure of the model is that it in fact comprises three sub-models: one for simulating liabilities' cash flows (the liability model), one for simulating assets' cash flows (the asset model) and a last one for simulating interactions between assets and liabilities in order to properly assess the hedgeable asymmetries (the asset and liability management model or ALM model). Both asset and ALM models are to be run in a stochastic mode, which in turn requires having an additional sub-model generating the required scenarios, i.e. the economic scenario generator (ESG). (See Chap. 4).

The use of the ESG produces the distribution curve of the model's outputs. A probability should be given to each point of the curve. Depending on the calculation's environment, risk-neutral or real-world, the given probability is different.

3.4 Cash Flow Issues

3.4.1 Time Horizon

The projection is done over a fixed period, which should match either the farthest expected lapse of the last contract of the portfolio (i.e. run-off mode) or a study-specific term (with or without future new businesses).

Thus, the model should be flexible enough to produce simulations at different horizons, from short term to long term, depending on the needs they are dedicated to answer. This is not really a modelling issue but it is potentially difficult, if not impossible, to set the model parameters and variables in order to reach the same level of precision and robustness in all cases. A systematic behavioural law (such as, for instance, mechanically selling a percentage of the equity portfolio each year) may be consistent in the long term horizon of a contracts' liquidation but may give unrealistic results if we focus only on a particular projection year.

As the model should stop at a certain point in time, there are basically two ways to model the end of the projections: the quick one, i.e. to stop there, and the more realistic one, i.e. to allocate the Company's balance sheet at the stop date between insurers and policyholders. The first method should only be used when outstanding figures are not material. The second method is quite a paradox: it approximates the impact of running-off the portfolio until the end by considering a theoretical situation where the Company is no longer a going concern. The question to answer is then: "To whom does the balance sheet belong?" The answer should be found in the local insurance code (see undertaking liquidation or insurance portfolio transfers sections) and in the contractual clauses.

3.4.2 Modelling Cash Flows

Projected cash flows are balance sheet positions at each observed time step. To properly assess them requires projecting some elements of the profit and loss account as they have an impact on assets and liabilities: for instance, the policyholders' profit sharing account, the level of claims… Obviously projections of either balance sheet items or P&L items are made on a consistent level of granularity.

Fig. 3.2 Assets and liabilities balance sheet

Asset	Liability
Investments	Own funds
	Technical reserves (Gross of reinsurance)
Reinsurers' share of technical reserves	
Receivables and other assets	Debts and other liabilities

The cash flows that we are willing to model are those cash flows that explain the roll-forward of the initial balance sheet, at the beginning of the year, in local norms (i.e. the local Generally Accepted Accounting Principles or local GAAPs) to the final balance sheet, at the end of the year.

The balance sheet is the instantaneous photograph of the Company's wealth, and it is composed of two parts, the assets (what you own) and the liabilities (what you owe), which are balanced by the shareholders' own funds as illustrated in the following table (Fig. 3.2).

The P&L account displays a film of the Company's activities throughout the considered period, here the year, and corresponds to the difference of positive cash inflows and of negative cash outflows; this difference being the profit or the loss of the period.

Schematically, as illustrated in the following table, the P&L account for a life insurer is composed of four modules: 1. Premiums and premiums reserves, 2. Benefits and associated reserves, including policyholders' profit sharing and commercial networks' variable commissions, 3. Net Financial Income, 4. General expenses, including taxes (Fig. 3.3).

Modelling is based upon detailed descriptions of the mechanisms behind each of these balance sheet and P&L items. These descriptions should go deep into the assessment rules by reference to the local insurance code, the contractual clauses,

Profit & Loss account	Examples
Premiums & premiums reserves	Mathematical reserves Unearned premiums reserves
Benefits & associated reserves	Claims reserves Policyholders' profit sharing Commercial networks' variable commissions
Net financial income	-
General expenses, including taxes	-

Fig. 3.3 Profit and loss account

the accounting framework, and the decisions that should be made by top management… This is because modelling all comes down to finding the right formulas to reflect as realistically as possible the running of the Company. Therefore a model designer should talk to (almost) all employees. Indeed, it is necessary to get these descriptions of how cash flows are flowing from different teams within the Company and not only from the actuaries, the primary and traditional source for modelling. Accountants, for example, are most needed in order to explain the complex working of regulatory standards, the rules to abide to for reserves' projection… The Financial Management team is instrumental in describing the impact on the P&L of the structure of the own funds (subordinated debts' interests…). The Management Accounting team is responsible for the cost model. The Top Management should also being questioned upon their decision-making (i.e. discretionary policyholders' profit sharing).

All these teams and some others are to be interviewed in order for them to describe their daily actions and decisions. All the collected information is then to be translated into mathematical formulas in the projection models. These interviews should therefore be well documented, preferably in detailed business requirements or, if not, alternatively in the minutes of targeted technical interviews. In practice, there is rarely only one interview as the right level of depth is generally not immediately perceived by interviewees.

3.4.3 Modelling the Liabilities' Cash Flows

Cash flow modelling depends closely on the features of the marketed contracts. Therefore there is a structural difference between Savings contracts and Protection contracts.

Structure of the Savings Model

The first question about the structure of the model could be expressed as follows:

"Shall we model the multi-funds contracts in a unique environment?, or shall we allow the coexistence of two different environments, the first one for modelling the GF component of the multi-funds contracts, and the second one for modelling the UL component, the two models having no interactions?"

Modelling in a unique environment (i.e. a global model) enables us to take into account statistical switches between all funds, GF and UL alike and also to simulate a more realistic dynamic policyholders' behaviour. Most of the time, the policyholders manage their contracts as a whole. They surrender their contracts globally, making no difference between GF and UL. When they get loans backed by their contracts, the financial hedge is made of each one of the two components, GF and UL. Naturally this more realistic modelling is relevant only if inputs for the model and behaviours' laws are available at the correct level of granularity. Otherwise, the rough use of the global model could ruin its expected benefits and, as a consequence, drastically change the costs/rewards balance expected when deciding to develop it. Modelling switches between all funds of a Savings multi-funds contract also implies modelling switch fees levied on contracts and switch commissions paid to commercial networks.

Experience shows that it is difficult to answer the question of one or two Savings models. Many have a predefined idea of what is best. However, it is not easy to predict what the impacts of each modelling choice are on the key metrics. Designing a proof of concept (POC) might turn out to be impracticable or too costly. Listing all the pluses and minuses of each choice is the best way to get a global picture and at the end of the day, to avoid engaging oneself with a too-theoretical and potentially not workable model.

The choice of the number of modelled funds is another very important question, as it has two major impacts: the first on the data required to run the model and the second on the technical capacity to implement the model.

First, the number of modelled funds defines the granularity of information required to run the model. For example, it forces us to set switch statistical laws by modelled funds. Second, the higher the number of modelled funds, the longer and the costlier is the processing by the model. The model will demand more and more machine power in a proportion that may be exponential to the number of modelled funds. This capability of running the model in a reasonable timeframe and at a reasonable cost is of paramount importance, taking into consideration the average huge number of investment funds marketed in multi-funds contracts (more than one thousand) and notably, in the context of a very short reporting process.

Modelling each of these funds seems to be a sort of utopia. Once this is said, the immediate new question is: "What is the right number of modelled funds, i.e. what is the right aggregation level to allow a proper run and to be realistic enough?" The answer is obviously entity-specific and depends on the more or less comparable nature of the funds.

As we see, right at the outset of the modelling process, the model designer is confronted both by his desire to translate reality in its greater complexity, and his ambition to see his model used extensively to give insights into many management decisions. From a compromise of these two conflicting interests, his desire and his ambition, will come the overall quality of the model.

Structure of the Protection Model

The question about the structure of the Protection model could be expressed as follows:

"Shall we model in a unique environment the multi-risk Protection contracts that cover several risks (death, accidental death, temporary disability, critical illness or involuntary unemployment)? Or shall we model each risk in a separate environment without any interaction between them?"

Modelling in a unique environment allows modelling the interactions between risks more or less precisely. In its rougher version, the multi-risk environment focuses on ensuring that once an insured is dead, he is no longer considered to be in another risk status (disabled or unemployed). In its more sophisticated version, the multi-risk environment allows us to manage the transition, policyholder by policyholder, from one health status to another: for instance, from a sound health state to a degrading health state (temporary or permanent disability).

The point here is to challenge whether multi-risk modelling is really necessary to meet the needs of the model users. If it is eventually considered indispensable, its use, when developed, requires setting the transition probabilities and so, to have the required statistical database reliable and exploited. Of course, a multi-risk model will require top-of-the-notch IT capacity to be used in the best conditions.

Arbitrage between precision and robustness is again at the heart of the decision-making process here.

Structure of the Model and Its Impact on Data

The choice of a structure has consequences on how the models will be fed. The more sophisticated the structure, the higher the volume of the data. In addition to the issue of getting these data is the issue of ensuring that they are at the required quality level. (See Chap. 10).

Modelling Insurance Contracts' Cash Flows

It all starts with inventorying transactions on contracts (collecting premium, surrender payment, claim payment, contract cancellation…) because they are the basic business-related cash flows reflecting decisions made by policyholders. Their assessment key drivers are the actuarial hypotheses and experience, which are known territories. Indeed, actuaries have been producing them for a long time. These assumptions relate to policyholders' behaviour, as encapsulated for instance by surrender laws, top-up premium laws and also periodical premium reduction laws.

Some issues on these basic cash flows deserve to be briefly commented on.

The premium frequency (single, periodical, and top-up) might differ from one contract to another and requires different statistical laws to be modelled. In the case of top-up premiums, the reason for their collection should be identified: are they structural?, in which case, they could be considered as periodical premiums. Or are they conjectural?, in which case, a dynamic collection law should be built-up. This question of the frequency of the premiums is also linked to the question of the contract boundaries, as expressed by the Solvency II texts.

Structural laws (dynamic laws are addressed in the section "Assets and liabilities interactions") are required everywhere in the model, either on Savings or Protection perimeters. Indeed, the model considers mortality laws, structural surrenders laws

Date of calculation	Method 1	Method 2
In-Force claims	Payment rate chronicle	Liquidation triangle
Future claims	Entry rate & recovery curve	Claims ratio

Fig. 3.4 Possible methods for claims projections

(with surrenders penalties or not, with surrenders values or not), annuity option laws, collective portfolio transfer laws, structural switches laws… These laws can be different depending on whether the modelling concerns individual or collective insurance policies.

These structural laws are as good as the quality of both the statistics they are built-upon and the actuarial method used to translate them into series of data. (See Chap. 10).

On Protection perimeter, claims' modelling should distinguish according to the timeline: claims that have been incurred before the projection date—called "in-force" claims, and after the projection date—called "future new" claims. Claims' modelling is not unknown to the Solvency I framework. Indeed, if a large part of the claims reserves is constituted by the reported claims assessed file by file, the balance is composed of unknown incurred claims because they are reported late to the insurer. To evaluate these IBNR (incurred but not reported claims), the insurer uses simulation techniques… which are duplicated in the new context of Solvency II. These techniques may also be used to assess the other components of the outstanding claims reserves when the simulation is not made policy by policy: reported but not settled (RBNS) and in course of payment (ICOP).

Different modelling methods for claims are then possible depending on the date the calculation is done (Fig. 3.4).

In addition to the basic insurance contracts' cash flows, the model designer should also model the setting-up of the technical reserves. This stage is quite significant as the change in value of technical reserves is instrumental to the final value of the P&L.

In particular, for the Savings contracts, the modelling of the change in value of the mathematical reserves (MR) comes down to scheduling the management acts (i.e. the basic cash flows: premium collection, claim payment, switches, etc.…) whose features have been previously defined. The point here is to position them consistently on the timeline based on agreed conventions. For instance, monthly deaths all occur at the beginning of the month, so that no monthly premium is collected. This consistent scheduling is quite time consuming in the development phase. It should explain the roll-forward of the mathematical reserves over the period, as shown in the following table: (Fig. 3.5).

Always on Savings perimeter, a key element in the change in value of MR is their revaluation at the credited rate, i.e. the contract's annual return rate that is made of both the minimum guaranteed rate (MGR), if any, and the policyholders' profit sharing. The modelling of the profit sharing is addressed in the following section relating to assets and liabilities interactions.

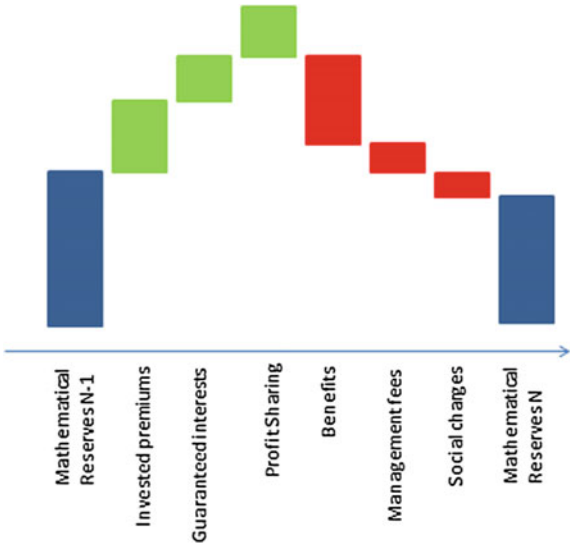

Fig. 3.5 Mathematical reserves changes in value (GF)

The modelling of the MGR revaluation should cover different situations: one-year MGR, multi-year MGR, variable MGR, high MGR, boosted MGR, low MGR, whole life MGR… MGR might also be modelled by policyholders' segments, by distribution channels… The final number of modelled MGRs is decided by the level of aggregation that produces the results at the right level of analysis, as defined by Top Management. Again, the more detailed the granularity, and the model may have technical and IT limits, the more aggregated is the granularity, and the model might support numerous biases with impacts on the analysis and results' validations.

For the Protection perimeter, the premium reserves' modelling is based on well-known formulas taught in all actuarial schools. However, these formulas may have to be duplicated to be applied differently to each component of the unearned premium reserves (UPR): pure premium, commission, fee, capital remuneration. Each component may have a different amortization pattern. Overall, here, again, the combinatory may be quite high.

Overheads' Cash Flow Modelling

Three issues are to be addressed:

Nowadays, overheads are closely monitored and, often, subject to strong action of the Top Management. However, in most forward-looking assessment or in the marked-to-model fair value calculation, one is not authorized to include the benefit of any cost-related programs that are today the reality of all major insurance players: cost reduction programs, increasing operational efficiency programs or stability expenses budget programs, also often named "zero based budgets".

However, for some calculations it appears to be a necessity, as for the Company appraisal value measurement, for instance.

Repartition and costs' allocation are not necessarily done at a consistent granularity with the one chosen in the model. How could this discrepancy be fixed? Do cost models have to converge to a granularity consistent with that of the model or does the model have to change?

Projecting fixed overheads can be conceptually challenging in some regulatory environments where the Company is in run-off but still considered to be a going concern. In this case, how does one determine the right level of fixed overheads to be projected considering that in the real world they should be partly covered by future new businesses' margins?

The scope of overheads to be projected can also be questioned. In the run-off portfolio simulation, the administration costs are those to be closely considered, as the acquisition costs would intervene only marginally. Within administration costs, it is also necessary to isolate the exceptional non-recurring overheads whose projection will give erroneous results. As the definition of exceptional non-recurring overheads might be loose or differ from one company to the other, it is very important to document the reasons supporting their qualification as exceptional and non-recurring.

The projection term (see the previous section) is another decision to be made that has an impact on the overheads' issue. The farther the projection term, the more significant is the impact of the overheads projection on the modelled results. Such an impact can easily be demonstrated by a sensitivity study.

Special care should be given to investigate the analytical processing of overheads (namely, how they are allocated to products) in order to identify any inconsistency or bias compared with other modelling assumptions. Adjustments may require being necessary.

3.4.4 Assets' Cash Flow Modelling?

Assets' cash flow modelling is different from liabilities' cash flow modelling for two main reasons: First, it is a well-known subject which has been taught for decades in the Finance world, especially for the pricing of financial instruments. A modelling conceptual framework has been developed over time and has generated lots of literature. Second, assets evolve in a more unstable environment than liabilities, as they are subject to both financial markets laws and economical cycles.

Assets' cash flows represent cash flows generated by the financial instruments existing in the insurer's investment portfolio at the simulation date. They do not encompass the interactions between those assets and the liabilities that are addressed in the section "Assets and liabilities interactions modelling". Assets' cash flows are assessed using the ESG in order to take into account possible financial markets' fluctuations in the future and the asymmetric feature of their cash flows.

Assets mapping	Assets classes	Assets modeling type
Equities (CAC40, S&P500)	Equities	Equity modeling type
Physical Real Estate	Real Estate	Equity modeling type
Fixed rate government bonds	Fixed rate bonds	Bonds modeling type
Variable rate government bonds	Variable rate bonds	Bonds modeling type
Inflation rate government bonds	Indexed bonds	Bonds modeling type
Barrier caps	Rate options	Rate option modeling type
Alternatives, hybrids	Structured assets	X% * Equity modeling type (1-X%) * Bond modeling type
Equities Unit Linked	Unit Linked	Equity modeling type

Fig. 3.6 Example of asset class grouping

The first step in assets' cash flow modelling is to decide what assets the model will project: either each individual asset of the portfolio, or only asset classes. Modelling each individual financial instrument certainly turns out to be impracticable. Setting asset classes should be made in the most relevant manner as possible, especially regarding the asset allocation rules. To decide the right scope of asset classes, one should begin by looking at the asset mapping as shown by the detailed investment reporting. A homogeneous asset class regroups assets that have the same cash flow pattern and the same risk profile. In other words, they could be modelled with the same formulas. An example of asset class grouping is given in the following chart (Fig. 3.6).

The second step of the modelling is to model the investment strategy, i.e. the interactions between assets.

Indeed, once the classification is made, the choice of the formulas to be used is directly to pick one amongst the well-known closed formulas (Black and Scholes…). The asset modelling should, however, be complemented by additional formulas that reflect the interactions between assets. These interactions are not taken into account by those above-mentioned closed formulas which consider conceptually that assets' values are independent from one another and that assets are held to maturity. These concepts do not reflect the real world that is to be modelled for the insurance companies. Indeed, in the valuation process (Best Estimates Liabilities, Embedded Values…), insurer's assets should not be considered as independent from one another or from the global running of the company.

Most of the time, assets should respect an asset allocation, called an "asset mix", which indicates a target (in proportion to the total asset value) for each asset class. The revaluation of a given class then implies quite automatically whether to sell or to buy some assets of the other classes, otherwise the allocation objectives may not be respected.

Some assets (equities and real estate) could not be held until their maturity… as they do not have a maturity date and the capital gains they generate over their holding period is a significant part of the achieved yield. So they have to be sold at a certain point in time. Therefore asset turnover objectives (i.e. the frequency of asset sales) are to be input into the model to reflect the way assets are managed. These turnover objectives are different depending on the assets, more frequent for liquid equities, less frequent for more illiquid real estate.

Each of these events (asset allocation adjustment, capital gains and losses realization…) reflects a management action that should be encapsulated into a behavioural law, regrouped in the "investment strategy", so that stochastic simulations may be run. For instance, if the strategic asset allocation is to maintain the initial level of bonds, when interest rates are increasing, the market values of the bonds are reduced by the application of the bonds' closed formula. Therefore the bond asset class is weighted less in the Company's asset portfolio and, in some cases, this weight might be out of the authorized range as set by the Company "asset mix". To return to the target mix, some bonds need to be bought which, in turn, means selling other assets in the portfolio in order to get the cash to do so, via equities selling, for example.

Many other possible situations need to be taken into account in the cash flow projection model and require rules and benchmarks to be modelled in the most realistic possible manner, in line with true asset management. For instance, there are two different ways to approximate the asset allocation rule: the "asset fixed" strategy considers that the asset mix is a constant target over the period. The "asset drift" strategy recognizes that in reality the asset mix is evolving in a given corridor.

3.4.5 Assets and Liabilities Interactions Modelling

Assets and Savings Liabilities Interactions

Central to the model, certainly the most technical part of it, assets and liabilities interactions are what make the results of stochastic models. There are many such interactions, the two main ones being as follows.

The first ALM interaction, the investing/disinvesting of liabilities' cash flows into/from appropriate asset classes, should be made according to preset rules.

The second ALM interaction, setting the policyholder's profit sharing based on actual asset yield, implies, in some jurisdictions, meeting some complex regulatory rules (just think about the French insurance code!) and, in all jurisdictions, requires being compliant with the contractual clauses. The level of the policyholder's profit share in turn triggers or fails to trigger dynamic surrenders and switches.

The model designer has to somehow enter into the brain of the policyholders to understand their expectations: as for their GF profit share, they certainly compare it to actual market rates. Depending on the gap, they may surrender their contracts. As for their UL values, they probably compare them to the profit share they receive on their GF component. Depending on the gap, they may decide to switch in or out of the UL component.

The Financial Strategy

These ALM interactions together with the investment strategy form what we call the "financial strategy". This strategy explains movements from the initial balance sheet to the final balance, as displayed hereafter (Fig. 3.7).

- From the initial balance sheet, the first movement is the yearly cash inflows and outflows from recurrent operations, both the asset cash flows (dividends for equities, coupons and maturities for bonds…), and the liabilities cash flows (premiums minus benefits), generating a result.
- The second movement comes from the financial markets' fluctuation over the period and their impacts on the purchases and sales of assets made in order to be in the asset mix range. This movement is made of two components: the first is the change in assets' values due to the application of economical scenarios; the second is the implementation of the investment strategy.
- The third movement is both the calculation and allocation of profit sharing to the mathematical reserves. The calculation depends on all or a proportion of the results generated by the two first movements.

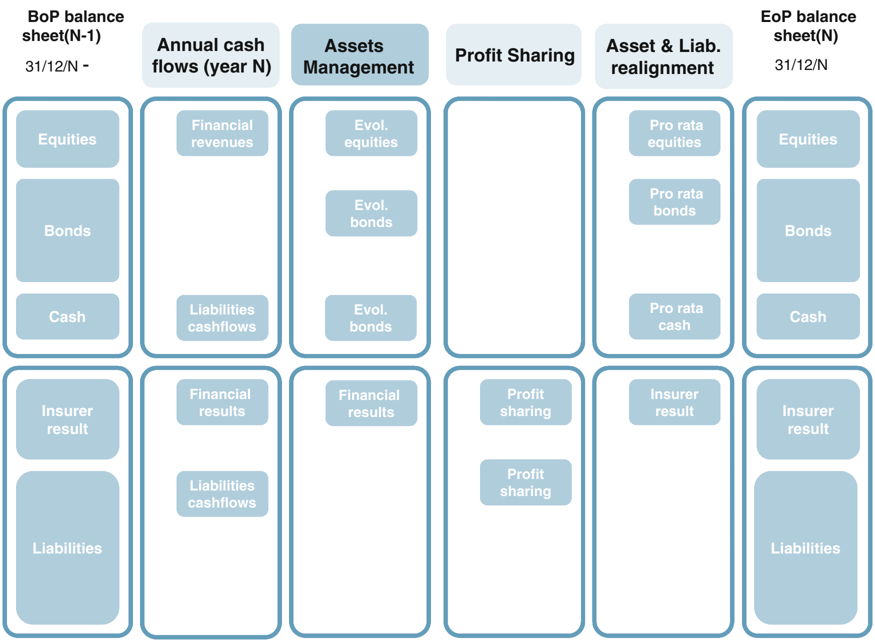

Fig. 3.7 The financial strategy cinematic

- The fourth movement is the assets and liabilities realignment which consists in selling assets (equities, bonds, and cash) proportionally to their respective weight in the total asset amount in order to make assets match with liabilities. Doing so, the assets sold in excess of liabilities crystallize the current period result.
- This financial strategy and its sequence of movements are repeated for each economical scenario, and for each calculation time period up to the term of the projection horizon. This means that for a stochastic run made of one thousand economical scenarios and for a forty year horizon, the financial strategy runs forty thousand times.

Policyholders' Profit Sharing Modelling

The target profit sharing is the value that the Company is willing to give to policyholders at each simulated period in addition to the minimum guaranteed rate (if any) in order to be consistent with its participation policy. This target rate depends on the asset yield. It also reflects Management's anticipation of the competitors' credited rate. It is consistent with:

- The wealth position of the Company at the end of the period, notably the unrealized gains or losses on assets,
- The credited rate of the previous year,
- The financial markets' situation,
- The top management's will to strategically promote some contracts.

The target profit sharing is the target that, if not achieved, might trigger surrenders. Therefore the modelling should reflect the management actions taken to be as close as possible to this target. Indeed, there exists a certain number of levers when, for instance, the recurrent asset income is not considered high enough; such as adjusting the insurer's margin downward, or adjusting the amount of capital gains upward…

In terms of modelling, the credited rate may be deduced from a formula decided at the start of the projection period. This formula is often unchanged during the projection period. This is mainly because changing it would have a significant negative impact on the calculation runtime, despite it having a positive impact on the results. Indeed, decisions in situations around the central scenarios certainly differ from the decisions made in extreme scenarios. Today this variable profit sharing formula is not feasible.

Therefore, the modelling choice for the implemented formula is very important as you will have to stick with it over the tens of projected years. Making an error will have a direct impact on the indicators, such as the liabilities fair value or the value of future profits. For instance, a very high target profit sharing will trigger all levers modelled in order to reach it whatever the financial markets' situation: the insurer's profit will be quite low in these circumstances. In addition, this very high target will over the long projection horizon generate surrenders and switches with a high level of probability. Conversely, a low target profit sharing will increase the value of the calculated indicators (like the future profits), perhaps too much. This might push validation and certification teams together with model users to reject the model.

Assets and Protection Liabilities Interactions

There are many such interactions, the three main ones being as follows.

The first interaction, the investing/disinvesting of liabilities' cash flows into/from appropriate asset classes, should be made according to preset rules.

The second interaction is the variable commissions of partner distributors when there is a contractual remuneration of the technical provisions based on market interest rates like Euribor.

The third interaction is the interaction between creditor payment insurance contracts (CPI) and the underlying mortgage contracts: actual interest rates may push borrowers to reimburse their loans early through renegotiation. This in turn would result in early insurance policy surrenders. Therefore the question may be asked as to whether modelling these dynamic surrenders or not. A "yes" answer implies the availability of banks' statistics on the dynamic surrenders of mortgage loans and the capacity to derive a CPI surrender law from them.

3.5 Conclusion

When seeking to find an answer to the different issues listed in the previous sections, the temptation is to design a dream model with the following features: a robust tool, tested and validated internally and externally, built to mirror the business-model of the Company, consistent with different norms and standards, well documented, simple to feed in, giving rich and exhaustive information on outputs, as fast as possible to run, available everywhere in the whole Company for multi-users needs, embedded in a solid IT environment, with a high evolution potential and finally at a reasonable cost!

This dream model is… a dream. However, it could be approached provided some conditions are fulfilled: Data and parameters are of an appropriate quality; the model is consistent with both the reality of the business and the management rules; the runtime is fast enough to meet the different deadlines.

These achievements are obviously conflicting interests. So, to get the dream model, one should find a balance between them and put the cursor at the right position: if one of the sought-after achievements is pushed too far, the other two might deviate from their starting points and might be degraded as they would impair the model. This balance should be managed by decisional committees like a modelling committee in which priorities of developments, minor or major evolutions, can be investigated and studied in order to define the proper modelling roadmap. Indeed, analysis on the current model is necessary to explain how it works and to allow users to form their own new intuition. R&D is also necessary to prepare future changes to model and to make educated decisions on whether or not to launch them.

Chapter 4
Economic Scenario Generators

Thierry Moudiki and Frédéric Planchet

Abstract The projection of economic and financial risk factors is a key element of prospective analyzes made by life insurers, both for the calculation of reserves under Solvency 2 and for the asset allocation and management of financial risks. This projection is achieved in practice through "economic scenario generators" (ESG), which are inputs for the calculus of the economic value of assets and liabilities and the analysis of the distribution of this value. The calculation of economic values is based on the "no free lunch" assumption and therefore leads to model the risk factors in a riskneutral probability, while the analysis of the distribution of these values requires the projection of these factors under the historical probability. Therefore, the insurer must handle different representations of the risk factors, which requires looking at the characteristics of a risk neutral ESG, those of an "historical" one and the possible need for coherence between these two representations. This is what we propose to do in this chapter.

4.1 Introduction

Analyzing economic and financial scenarios is a decisive element for insurers managing long-term risk because of the importance in this context of the investment proceeds from sums placed as reserves.

This analysis underlies insurers' policies towards asset/liability management (ALM) by allowing them to arbitrage between performance and risk for the various possible asset allocations.

For life insurers (particularly for savings and retirement), the performance of an asset plays an even greater role in risk management since it contributes to determining the level of liabilities through profit-sharing plans underlying profit uprating policies. In this case, the economic scenarios have a direct impact on commitment levels.

T. Moudiki · F. Planchet (✉)
Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France
e-mail: frederic@planchet.net

F. Planchet
PRIM'ACT, Paris, France

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_4

The analysis of economic and financial risk factors potentially affecting an insurer's balance sheet may be initially considered as part of a scenario approach, with each scenario being connected by a clearly identified economic situation. But it is difficult to come up with a significant number of distinct scenarios, each one having to be defined "manually".

Building scenarios "automatically" using a probabilistic description of the risk factors and their interactions provides an alternative approach to producing scenario families in large numbers, with each being assigned a probability of occurrence. We agree, from now on, to refer to "economic scenario generators" (ESG) as a tool for projecting economic and financial risk factors in a probabilistic framework. Here, the term "tool" refers to the actual model itself (the formal mathematical description of risk factor behaviour), the data that feeds it (its programming) and its IT implementation allowing an operational employment of the model.

Typically integrated risk factors are share prices, interest rate term structures, bond counterparty default, liquidity, inflation and property prices. Other economic factors such as unemployment rates and growth rates can also be modelled. However, we will limit ourselves here to explicit financial risks in that they are directly linked to traded (or tradable) asset prices on the market.

In this restrictive framework, the relationship between the risk factors and the price of assets may be more or less direct:

- For equities and real estate, the modelled factor is directly the price of the asset.
- For bonds, in general, these are modelled using a limited number of explanatory factors, typically short-term rates.

Projecting risk factors depends on the use that will be made of the ESG, which leads to defining the concepts of both historical and risk neutral probabilities.

Emphasis here is placed on highlighting some theoretical and practical focal points related to establishing ESGs in this context, focusing especially on calculating reserves. ESGs are developed more comprehensively in Kamega et al. (2009) as well as in Faleh et al. (2012) regarding particular issues of asset/liability management.

4.2 Prices and Price Distributions

As we will discuss below, the calculation of reserves for liquidity needs that mostly involve a financial hazard is homogeneous to a price calculation (this is also presented in detail below). Once reserves have been calculated, a one-year balance sheet is forecast to assess the minimum level of required capital to ensure solvency with a probability of at least 99.5 %.

An asset/liability model used by life insurers must be able to take into account these two aspects, implying in particular:

- Having to calculate prices (assets and reserves).
- Having to calculate the quantiles of the net asset value (NAV) distribution.

Taking into account both of these two elements in the framework of an internal model or of one's own risk and solvency assessment (ORSA)[1] is complex because it leads to manipulating the risk factors for two very different purposes.

As part of a comprehensive modelling whose ambition is to provide distributions of economic value, we must absolutely use a two-level approach:

- building a functional g that provides the price vector as a function of the status variables Y at the time of calculation, $\pi_0 = g(Y_0)$, and
- building a positive dynamic for the risk factors, Y_t.

We can then determine prices for any date via $\pi_0 = g(Y_t)$. For example, within a Vasicek type mono rate factorial model, we have the following short rates:

- Forecast: $dY_t = dr_t = a(b - r_t)dt + \sigma\, dW_t$, W_t being a Brownian motion;
- Price calculations: $dr_t = a(b_\lambda - r_t)dt + \sigma\, dW_t^Q$, with $b_\lambda = b - \frac{\lambda\sigma}{a}$ and $W_t^Q = W_t = \lambda \times t$ being a Brownian motion under the probability of Q, which allows us to construct the functional

$$\begin{aligned} g(r_t) &= P(r_t, T-t) \\ &= \exp\left(\frac{1 - e^{-a(T-t)}}{a}(r_\infty - r_t) - (T-t)r_\infty - \frac{\sigma^2}{4a^3}\left(1 - e^{-a(T-t)}\right)^2\right) \end{aligned}$$

by noting $r_\infty = b_\lambda - \frac{\sigma^2}{2a^2}$.

Building the functional g relies on the conventional assumptions of market finance and in particular the absence of arbitration leading to "risk neutral" probabilities, which make the process of discounted prices a Martingale.

Constructing the dynamics of Y is an econometric problem.

From a mathematical point of view, the models used to calculate prices are often described in a continuous framework that allow having closed formulas available for the prices of certain derivatives (the most famous of them being the Black and Scholes formula) and in a discrete framework for dynamic vision, time-series models providing many tools.

Finally, we can see that, in this theoretical framework, the two representations are linked via the "market price of risk", λ. Thus, a natural approach to building a coherent ESG is to specify a risk factor dynamic as a "historic probability", make an assumption on the form of the market price of risk and then deduce the associated dynamics as a "risk neutral probability". This approach, however, is fraught with many practical difficulties, as we will see below.

[1]*Own Risk and Solvency Assessment*. See Guibert et al. [2014] for a thorough analysis of this method.

4.2.1 Asset/Liability Management

Building an ESG for the purpose of asset/liability analysis is globally reduced to an exercise in econometrics. It is a question of making a probability model that can distort price structures known at the time of the calculation.

Many approaches have been proposed. The Wilkie model (Wilkie 1986) is one of the first examples of this type, followed by the Ahlgrim model (Ahlgrim et al. 2005), developed within the Casualty Actuarial Society (CAS) as part of discussions on Dynamic Financial Analysis (DFA—ALM non-life). The Hibbert et al. (2001) model is a good example of a model that gives special importance to long-term issues. It allows the generation of scenarios for (real and nominal) interest rates, inflation, return on equities and dividend rates. In terms of dynamics, two-factor models are used for interest rates and inflation, a regime change model is used for equities and a Vasicek type one-factor model is used for the dividend rate. Its structure is in fact very close to that of the Ahlgrim model.

We can also cite Campbell et al. (2001) and the Cap model: link (see Ziemba and Mulvey 1998).

Modelling the distribution of a complex asset value affected by multiple risk factors implies, in the context of Solvency II, giving particular attention to the behaviour of this value in the left side of the tail distribution (see Leroy and Planchet 2011). This leads directly to examining the consequences of the dependency structure between the risk factors relating to the solvency capital requirement (SCR). We refer to the high sensitivity of this choice for a value-at-risk at 1 year from the portfolio value, thus on the level of the SCR (see Armel et al. 2011).

For instance, if we consider a portfolio made up of 80 % 5-year EEA government bonds, 10 % French equities and 10 % real estate, managed to maintain this allocation constant, the quantile at one year from loss distribution is equal to 4.7 % of the portfolio's market value with the Gaussian copula, but 12.2 % with the Clayton copula, which can be shown to better represent the dependence structure of the assets in question.

Finally, it should be noted that in the context of asset/liability management, the major difficulty lies not in the construction of the generator itself but in the ability to efficiently find an optimal balance (typically an asset allocation) within such a fairly large space (the number of asset classes). This point is detailed, for example, in Monin (2014) and Faleh et al. (2012).

4.2.2 Calculating Reserves

When Joseph Tchundjang Pouemi wrote "we cannot theorise a reality that has not been previously observed" (Tchundjang Pouemi 1979), he did not yet know that the option markets, relying on the ground-breaking work of Black and Scholes (Black and Scholes 1973), was in the process of proving that, with market finance, theory sometimes precedes observation.

The founding act of modern finance was indeed to demonstrate that when certain technical conditions are met, and when an investor is faced with a commitment to pay a counterparty an uncertain amount at a set maturity, he can put together a portfolio at a certain price without subsequently adding funds, so that the value at the portfolio's term is equal to that of the returns it committed to pay. The portfolio is established and managed to "replicate" these returns. In this context, the "value" on the date these returns are calculated can only be equal to the cost of putting together the replicated portfolio.

This conceptual framework provides a mathematical framework for calculating the "price" or the "value" of a cash flow, but especially rules for constructing a replicate portfolio. It turns out that calculating the price may be reduced to calculating the expected sum of future cash flow discounted at the risk-free rate by choosing a suitable probability, which is usually called the "risk-neutral probability" (see Planchet 2009b).

With financial products, generally, cash flow is expressed in a relatively simple way by using prices of different assets. For example, the cash generated from a European put option having a strike price of K on an underlying S is of the form $F(T) = [K - S_T]^+$.

Adapting this approach to an insurance context, and more particularly to savings and retirement portfolios, runs up against the complexity of describing the cash flows of the contract. These result, in fact, in complex interactions between the market returns of the assets backing the commitments, accounting rules, the applied rate and profit sharing reserves set by the insurer and policyholder decisions in terms of redemption.

Faced with this difficulty, specialists developed a modelling framework summarized by the following diagram:

The ESG, therefore, feeds a cash flow forecasting model and the amount of reserves R is reached by simulation[2]:

$$R = E^{P^A \otimes Q^F}\left(\sum_{j \geq 0} \frac{F_j}{(1+R_j)^j}\right)$$

$$\approx \frac{1}{N}\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{a=1}^{A} \frac{Cashflows_{t,n,a} - Premiums_{t,n,a} + Management\,fees_{t,n,a} - Loading_{t,n,a}}{(1+R_n(0,t))^t}$$

In order to be consistent with market values, the financial risk factors entered into the ESG must be modelled under a "risk neutral" probability.

The ESG must take into account all the financial risks faced by the insurer or at least the risk-free rate, the credit risk, inflation, stock prices and real estate. While financial derivatives are often built around a reduced number of risks (the loss in value of an underlying, or a counterparty default, etc.), this involves considering them together much more comprehensively.

[2]Ifergan (2013) discusses the effectiveness of this scheme from a numerical point of view.

Among these factors, the choice of the rate model is a key element of a "risk neutral" ESG. The model's ability to adequately represent the price of interest rate derivatives is a criterion that respects consistency with market values. But when examining the choice of market finance models, we see that there is no relevant model for all of the fixed income products and that the model is chosen and calibrated according to the product's nature; different models are used for *CAPS*, *Swaptions* and CMS. The model is chosen and calibrated to best represent the price of the instrument for which it is used, without claiming to properly represent the prices of other instruments of different structure.

Applying this approach in the context of making a best estimate calculation of reserves for savings contracts in euros, implies, a priori, having the prices of reset and cyclical redemption options; information that does not exist and the calculation of which is therefore made in a mark-to-model framework without any direct observable data. We must then rely on surrogate data and use observable prices of interest rate products, which can reasonably be expected to behave like priced options.

This choice is somewhat arbitrary and this observation reinforces the normative character of reserve calculations. In other words, if the model set by the regulator is without constraints, the latitudes of choice are significant and deciding whether a model is relevant or not can only be based on criteria of financial theory.

It can also be pointed out that where markets are incomplete, the loss of the uniqueness of the "risk-neutral" probability forces a choice between the infinity of reference possibilities or arbitrarily picking out[3] a model for calculating prices from an infinite number of possibilities.

This situation makes it all the more easier to design models tailored to specific situations, both in terms of the type of guarantee (savings in euros, retirement, etc.) as in the type of calculation (pillar 1, pillar 2). Thus, in practice it is necessary, as needed, to collectively manipulate the differentiated models having the same risk factors. The risk assessment of the model associated with this approach is a subject in its own right, discussed in Chap. 8 of this book).

4.2.3 Is There Any Coherence Between the Different Descriptions?

It must be immediately pointed out, as we have seen, that the issue of coherence not only exists between "risk neutral" and "historic" representations but also between different representations of the same risk factors in different parts of the model. Projection models are indeed rather complex, which may involve various sub-models for the same risk factor depending on the calculation being made.

[3]The choice is arbitrary in an insurance context, in the sense that the tight relationship in market finance between price calculation and coverage cost is much more tenuous. Choosing a "risk-neutral" probability in an incomplete market is equivalent, in fact, to choosing risks to be covered and those that will not.

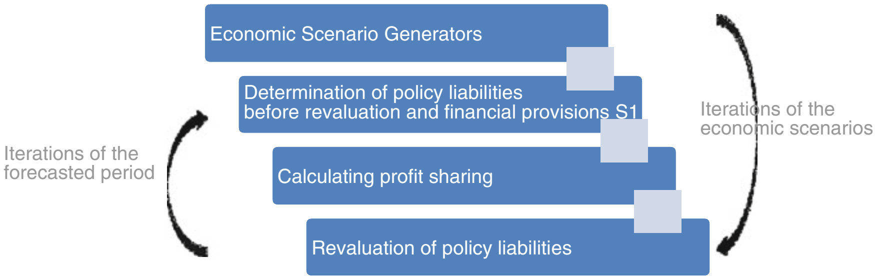

Fig. 4.1 The role of the ESG in an asset/liability model for calculating reserves

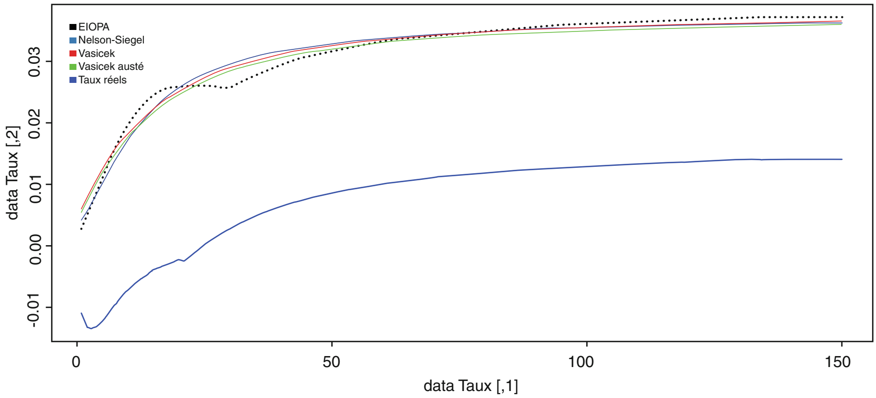

Fig. 4.2 Different representations of the yield *curve*

As an example, consider the model described in Bonnin et al. (2014) as part of ORSA. The term structure of interest rates accounts for different calculations:

- First of all, an arbitrary curve is provided by the regulator in a non-parametric form;
- Its "risk neutral" dynamic as of a particular date is used to calculate reserves due to the presence of asset/liability interactions: the deflator calculation is made within the framework of a single-factor model that requires a short rate "risk neutral" projection model coherent with the yield curve. The Vasicek model is chosen.
- It must also fit in with historical probability in order to measure the interest rate risk and associated capital requirements: the Nelson-Siegel model is used for this (Fig. 4.1).

This leads to handling the various representations of the yield curve based on the model's components (Fig. 4.2).

The interactions between these representations are the following:

- The initial curve provided by EIOPA is used to calibrate the adjustment of the Nelson-Siegel parametric form.

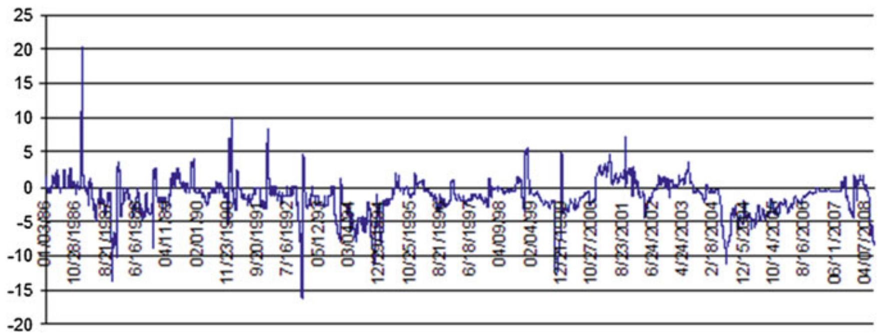

Fig. 4.3 Changes in the price of market risk (see Caja and Planchet [2010])

- This parametric form is used with each forecast to calibrate the Vasicek model from which short-term rates are forecast that are used to calculate discount factors.

In this context it is important to ensure coherence between these various representations. Calibration is thus determined to minimize the differences between these various curves. To ensure consistency when calculating present values between the initial curve from the Vasicek model, this must be corrected to ensure, *ex-post*, that the present value of the unadjusted performance is identical with the zero-coupon of the EIOPA curve and the same coefficients from the Vasicek model. To achieve this, a spread adjustment is added to the zero-coupon rate from the Vasicek model. This spread is set at 0 and assumed constant during the forecast.

The degree of precision of each forecast is adapted to the nature of the calculation made. A forecast based on three factors is thus used for the deviation of the curve based on historical probabilities, while the deflator calculation is based on a mono-factor model.

The quality of a model relies heavily on the relevance of the choices made at this level, with the understanding that using these various representations is incorrect, from a strictly theoretical point of view.

The general question of coherence between the "risk neutral" and "historical" approaches is based on a theoretical view of the "market price of risk" concept (see Caja and Planchet 2010). Note that most of the usual models[4] are based on very strong assumptions that require this factor to be highly consistent; however, experience shows it to be very erratic, as seen in the following Fig. 4.3.[5]

Taking this erratic nature into consideration, however, leads to models unsuited for insurance purposes, due, in particular, to the lack of closed formulas for zero-coupon prices, which complicates calibration and digitalization.

[4]As with the Vasicek model described in the introduction to this section, which gives a constant price to market risk.

[5]Built on the basis of historical Libor rates and used by Caja and Planchet (2010).

Moreover, theory requires that a change in probability that allows switching from an historical measure to a "risk neutral" measure only changes the leeway of the underlying processes that model the risk factors and leaves the volatility process constant.[6] However, in practice we consistently see different historical and implied volatilities.

The table below, taken from Gauthier (2011), illustrates the inclusion of the above comments by showing the choices made by practitioners for the different risk factors within an ESG context:

Finally, the coherence between the risks taken into account by the model and those actually incurred by the insurer must be examined.[7]

Consider, as an example, a specific amount payable at maturity *T*. The purchase of a zero coupon bond with no risk of default provides a perfect hedge for interest rate risk. The price of a bond on the asset varies in the opposite direction of the discount rates measuring the market value of liabilities and assets, and is always equal to the best estimate of reserves.

Now let us assume that the organization invests its assets in a bond having a spread of the reference discount rate. This situation is typical of French insurers invested in French government bonds who must discount liabilities based on a curve reflecting the overall level of German interest rates.

The existence of such a spread is justified by the credit risk, default risk and liquidity risk of the bond issuer. But in our example, counterparty default is the only real risk to the insurer. Credit risk (along with the possibility of a reduced credit rating) does not come into play immediately because it does not influence payments at maturity. Similarly, the liquidity of the security is irrelevant to the insurer in this example because it will be held to maturity. However, if the spread increases, this will cause the value of the asset to fall while the liability remains the same, leading to a shortfall, even though the insurer is perfectly hedged.

This simplified example highlights a difficulty in the prudential framework and it can be seen that this is directly related to the requirement of recognizing assets at their market value. However, market value is of little interest to the insurer, in this case, since he has no intention of selling the security. The commitment coverage rate is thus subject to fluctuations, the causes of which are only partially related to the actual risk, since counterparty default is the only real risk. We can also observe here that what really interests an insurer is the objective probability of default. The valuation process, however, is based on the probability of a "risk neutral" default which adds to the objective default probability an increase related to investor risk aversion, which can vary over time even if the probability of default remains constant. Thus, even if the elements of credit and liquidity are removed from our analysis, the inadequacy of the model remains.

[6]More generally, changing the measure leaves the dependence structure constant.

[7]The rest of this section is based on Planchet and Leroy (2013).

However, even if under Pillar 1 of Solvency II the leeway to remedy this situation is slim, it is possible to consider taking strict account of it under the second pillar in the ORSA evaluations (see Guibert et al. (2014) as well as the next section).

4.3 Calibration and Implementation

Once the model is formally specified, it still remains to be implemented into a software tool and then the settings must be decided (Fig. 4.4).

4.3.1 From Theory to Practice: Implementation of the Model

To illustrate the scheme's implementation as seen in Fig. 4.1, the simplest example that we can suggest is to simulate the price calculation of a zero-coupon bond within a mono-factor model. To be even more specific, we set it within the framework of the Vasicek model based on the short rate dynamic $dr_t = a(b_\lambda - r_t)dt + \sigma\, dW_t^Q$, W_t^Q, which is a Brownian motion under probability Q. The estimation error[8] associated with simulation-based calculation of zero coupon prices, in this context, would look like this:

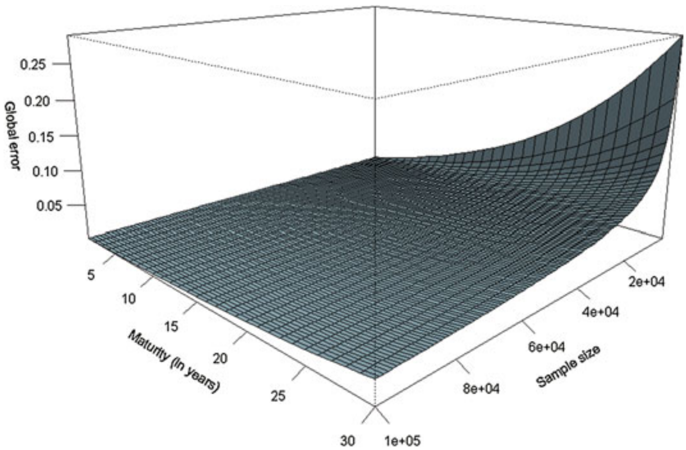

Fig. 4.4 Simulation of a zero-coupon price

[8]Estimation error, here, is understood as the half-width relative to the asymptotic confidence interval at 95 % for the interest size estimator.

Errors logically increase with maturity and decrease with the number of simulation runs.

It can be said that approximating the price of a single zero-coupon through simulation is not easy. After 2000 runs, which constitute the number of simulations for a best estimate calculation, the price of a 30 year zero-coupon is estimated with a relative error of approximately 30 %, which decreases to 10 % after 10,000 simulations. However, the practical consequences of this example should be qualified because if the relative error is significant, the absolute value of a long-term zero-coupon is low and the impact on a best estimate, which combines cash flows from different maturities, is rather small.

Now consider a unit-linked contract with an underlying modelled by a log-normal process and rates modelled by a Vasicek model:

$$S(t) = S_0 \exp\left(\int_0^t \left(r(u) - \frac{\sigma^2}{2}\right) du + \sigma B(t)\right), \; dr(t) = k(\theta - r(t))\, dt + \sigma_r dB_r(t),$$

$B(t)$ and $B_r(t)$ being two Brownian motions under a risk-neutral probability.

The contract is for 10 years, fully redeemed at maturity. In the meantime, the structural redemption rate is 2 % and cyclical redemptions, of up to 5 %, are added on when the value of the unit is lower than the initial value (Fig. 4.5).

Here, we can see that the convergence of an empirical best estimate towards its theoretical value is slow and, after 1000 runs, a gap of about 1.5 % still remains. Consistent with the orders of magnitude observed in a life insurer's balance sheet,

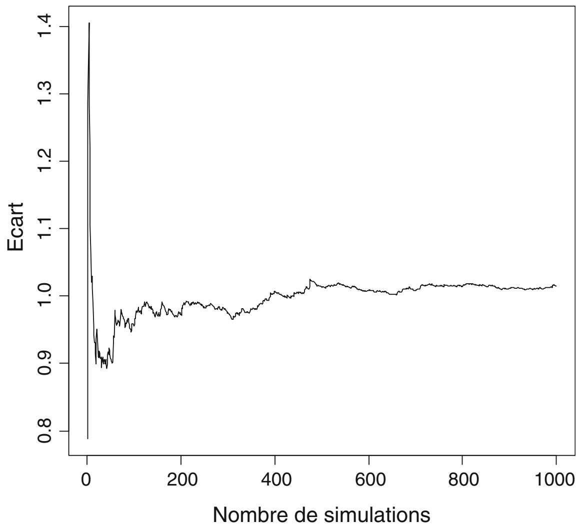

Fig. 4.5 Convergence gap depending on the number of simulation runs

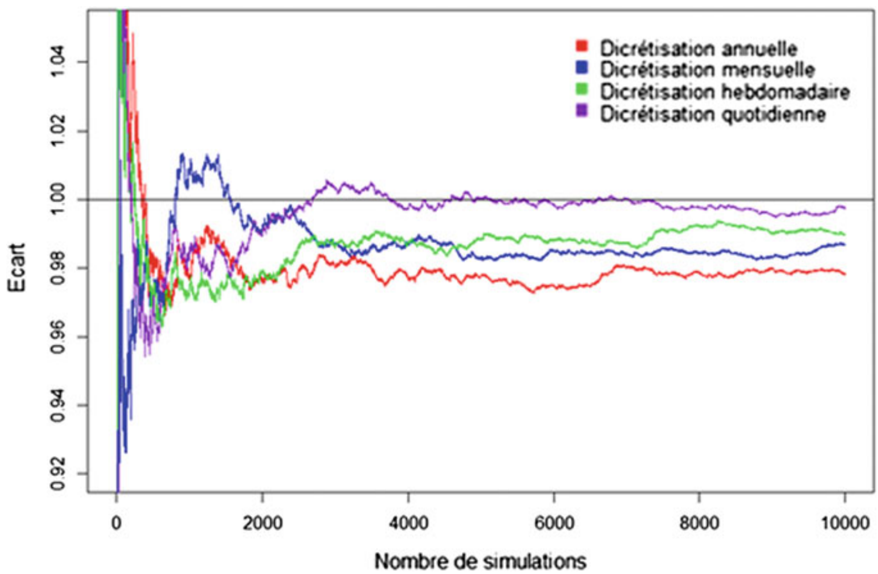

Fig. 4.6 Convergence gap in terms of the line of discretization

and under the assumption that the ratio between equity and technical reserves is from 1 to 10, a gap of 1.5 % on reserves led to a difference of about 15 % on equity.

Furthermore, it should be noted that this is no longer just a sampling error, but a systematic bias that slowly diminishes with the number of runs. Thus in the example above, we see that the actual value of a best estimate is systematically underestimated.

To cut this difference down by a factor of 10, multiply the number of runs by 100. Therefore, it may be useful (essential even) to optimize this scheme (see e.g. Nteukam and Planchet 2012).

The bias, thus demonstrated, should be carefully considered. More specifically, we find that the convergence is also impacted by the discretization choice of the underlying process. It is necessary to distinguish the line of cash flow forecasts (usually annual) and the line of discretization used to approximate the discount factors (see Ifergan (2013), their illustration will be taken up below). The latter is not constrained by considerations linked to performance and can be chosen arbitrarily at the end.

The following figure shows the impact of the choice made in this regard for a savings contract in euros (Fig. 4.6).

In practice, in the forecasting models used to calculate a best estimate, the sampling error is considered to be approximately 0.50 %, to which is added a possible discretization bias.

The care given to the model's operational implementation is, therefore, a crucial element for a coherent approach ensuring the implemented algorithms are compliant with the underlying theoretical model.

4.3.2 Some Remarks About the Calibration

Once the model is built, before using it, it must be fed a set of robust parameters consistent with its use.[9] The parameters fed into an asset model may be selected in three ways, without these different approaches being exclusive from one another.

The most natural approach, at first glance, involves selecting the historical series of values representative of the various modelled asset classes and, on that basis, estimate by statistical methods, usually the maximum likelihood.

For "equity" type assets, for example, this approach is intuitive, because the modelled value (price) can be directly observed on the stock market. For interest rate products using factor models, it is more difficult to compare the variables that determine the yield curve (short and/or long rates) with the amount of interest (the price of risk-free bonds). We are then led to search for the parameter values of the underlying model that best represent the interest variable–bond prices, in our example.

In practice, on a given date, we must then constitute a range of prices (for different maturities, for example) and seek the parameter values that minimize the standard deviation between the model's prices and the prices observed in the market. This approach, implicitly based on prices, is also used for equities by deriving the level of volatility from option prices and a pricing model for such options. It has the undeniable advantage of accurately reflecting prices on the initial date in market consistent approaches.

A direct approach is preferred when estimating the parameters of an "historic" probability generator and is largely a statistical problem, while the implicit price approach is adapted to the calibration of the "risk neutral" generator.

But the direct approach, as well as the implicit approach, relies on the premise of a certain stability of past behaviour that should recur in the future. In times of regime change, as we are seeing at this time, this assumption seems questionable. It is then conceivable to calibrate the model's parameters "according to an expert" or, more precisely, on the basis of assumptions about the development of the economic situation for the forecasted horizon–long-term assumptions that do not overly disrupt the model with values reflecting temporary crisis periods.

Inflation behaviour helps to illustrate this point. By adjusting data available since 1950 for inflation, a long-term inflation of 4.84 % is expected.

---

[9] It may be noted here that the parameters may be different in MCEV and IFRS/Solvency II, even though the underlying model is the same.

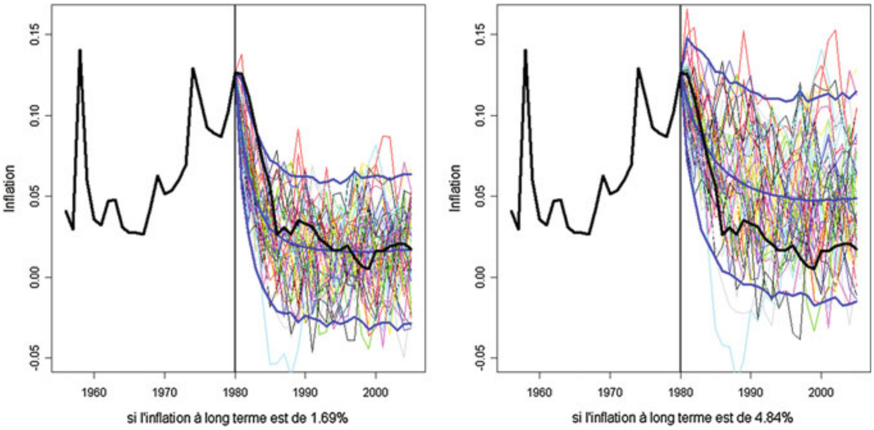

Fig. 4.7 Inflation calibration: impact of the choice of the estimation period (Kamega et al. [2009])

This does not correspond to the long-term growth objectives set by the European Central Bank (ECB) since the 2000s (Fig. 4.7).

The gap is significant and, medium term, it would be better to use a value consistent with the past few years given the ECB's monetary policy objective of maintaining inflation close to and below the 2 % level, without calling into question the ECB's ability to maintain its inflation target.

The prospective use of ESG in a long-term perspective imposes a confrontation between "objective" information provided by economic and financial data and economic forecasts provided by economic analysis.

Calibrating an ESG must also be addressed comprehensively. In fact, the various components of the model exhibit strong interactions: calibrating a model for default and liquidity risks, for example, directly depends on prior calibration for the interest rate term structure—as shown in the figure below, taken from Laïdi (2013), which presents a very detailed analysis of this issue (Fig. 4.8).

Estimation errors can accumulate with ESG, which in practice often leads to unstable parameter estimates and requires the use of certain techniques to stabilize them.

Finally, it should be noted that the implementation of valuations consistent with market values, in practice, leads to significant volatility in the evaluations and, consequently, the balance sheet.

This volatility is a matter of concern for regulators and the industry. By way of example, the latest specifications of the standard formula at the time of writing provide a volatility adjustment (see EIOPA 2014), whose objective is to stabilize calculation results (that is to say, reserves and the NAV). However, in times of crisis, this adjustment serves to neutralize the "surplus" volatility induced by the liquidity risk to bonds, only providing, therefore, a partial solution to this problem.

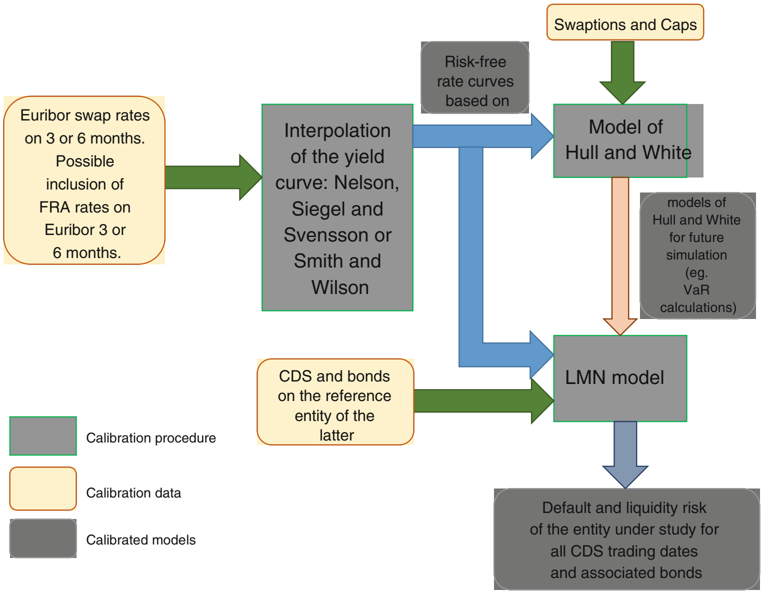

Fig. 4.8 Example of a calibration procedure for a rate model (Laïdi [2013])

Here, we propose an approach that seeks the same goal but allows extended control of the volatility parameters beyond just liquidity risk.

In the case where a stochastic process, used to forecast risk factors, has constant coefficients, which is a relatively common event, the choice of parameters should be logically compatible with this hypothesis over a short period.

For this, it is therefore necessary to choose parameters consistent with observed prices and assume the model's parameters remain stable. Calibration should be relatively stable during its updating on dates that are close. A method using only the last known price is obviously not relevant in this case because it induces parameter volatility that is not representative of the actual risks, as we saw in the discussion at the end of Sect. 4 2 above, and is not consistent with the chosen forecasting model. It would seem better to build a calibration procedure based on a fixed period.

The main idea is that the parameters must be determined to ensure that the prices of the model represent the best possible prices observed, not at a particular time, but over a set period. This is likely to limit model errors when the parameters are assumed to be constant over time.

The practical implementation of this observation requires the use of genetic algorithms (see Laïdi and Planchet (2014) for a detailed presentation in the case of a credit model) and is not immediate, but allows a significant gain in parameter stability and thus a gain in its capacity to exploit trends detected from the model's results.

4.3.3 An Example: ESG Packages

In this section we propose to illustrate, as simply as possible, the issues described above using a generator that ensures consistency with regulatory requirements under Solvency II. For this, we rely on the ESG package,[10] which was developed for this purpose.[11]

The implemented models are:

Nominal short rates: A one-factor model of Heath-Jarrow-Morton (Heath et al. 1992):

$$df(t,T) = \alpha_f(t,T)dt + \sigma_f(t,T)dW_t^{(f)}$$
$$f(0,T) = f^M(0,T)$$

where $f(t,T)$ is the instantaneous *forward* rate observed on date t with a maturity of T; $\alpha_f(t,T)$ $\alpha_f(t,T)$ and $\sigma_f(t,T)$ are both adapted processes, $(W_t^{(f)})_{t\geq 0}$ is a standard Brownian motion under risk-neutral probability, and $T \mapsto f^M(0,T)$ represents the instantaneous *forward* rate term structure observed in the markets for $t = 0$. By setting in this model

$$\sigma_f(t,T) = \sigma e^{-a(T-t)}$$

and by taking

$$\alpha_f(t,T) = \sigma_f(t,T)\int_t^T \sigma_f(t,s)ds$$

the model is equivalent to a Hull-White Extended Vasicek model (Hull et al. [1990]):

$$dr_t = (\theta(t) - ar_t)dt + \sigma^{(r)}dW_t^{(r)}$$

where r is the instantaneous short rate, a is the short rate's speed of mean reversion, σ its volatility, and $(W_t^{(r)})_{t\geq 0}$ is a standard Brownian motion under risk-neutral probability. This model is designated as "HW". The function $t \mapsto \theta(t)$ allows an exact adjustment of the model to the initial rate curve:

[10]http://cran.r-project.org/web/packages/ESG/index.html.

[11]The description of the model as well as the illustration are available at http://www.ressources-actuarielles.net/C1256F13006585B2/0/A5E99E9ABF5D3674C125772F00600F6C.

$$\theta(t) := \frac{\partial f^M(0,T)}{\partial T} + af^M(0,T) + \frac{\sigma^2}{2a}(1 - e^{-2at})$$

The HW model's main advantage is to allow closed formula derivation for bond prices and interest rate derivatives such as caps, floors and swaptions, the cash flows of which are similar to insurance liabilities.

However, it is likely to generate negative instantaneous short rates and, like all one-factor rate models, it generates perfectly correlated future spot rates.

Equities (with reinvested dividends): we use a simple geometric Brownian motion with stochastic interest rates from the Hull-White Extended Vasicek model (Merton 1973)

$$dS_t = r_t S_t dt + \sigma^{(E)} S_t dW_t^{(E)}, \quad dW_t^{(E)} dW_t^{(r)} = \rho dt$$

$(W_t^{(E)})_{t \geq 0}$ is a standard Brownian motion having a risk-neutral probability; equities and short-term rates have an instantaneous correlatiosn equal to ρ. This is the *Black-Scholes Hull-White* (BSHW) model.

According to Brigo et al. (2007) the BSHW model is useful for assessing options with long maturities, as is often the case in insurance, for which the stochastic nature of interest rates can have a significant impact on prices.

It is possible, using a closed formula, to determine the prices of calls and puts whose underlying assets are modelled by BSHW (Merton (1973) and Brigo et al. (2007)). Having closed formulas available is very important for calibrating the models to market data, as will be illustrated below.

Real Estate: we use a simple geometric Brownian motion with stochastic interest rates from the Hull-White Extended Vasicek model:

$$dS_t = r_t S_t dt + \sigma S_t dW_t^{(P)}.$$

The difference with the BSHW model is that there is no instantaneous correlation between the interest variable and the short rates. According to Koivu et al. (2005) and Djehiche et al. (2005), property price movements are similar to those of stock prices. Returns consist of large price fluctuations associated with relatively stable cash income. However, unlike equities, cash makes up a large part of the total return on this type of asset. Djehiche et al. (2005) also discusses the calibration of this type of model.

Returns on bonds: the NML model of Longstaff-Mithal-Neis (see Longstaff et al. 2005) is used, which is based on:

- a default intensity

$$d\lambda_t = (\alpha - \beta\lambda_t)dt + \sigma\sqrt{\lambda_t} dW_t^{(\lambda)}$$

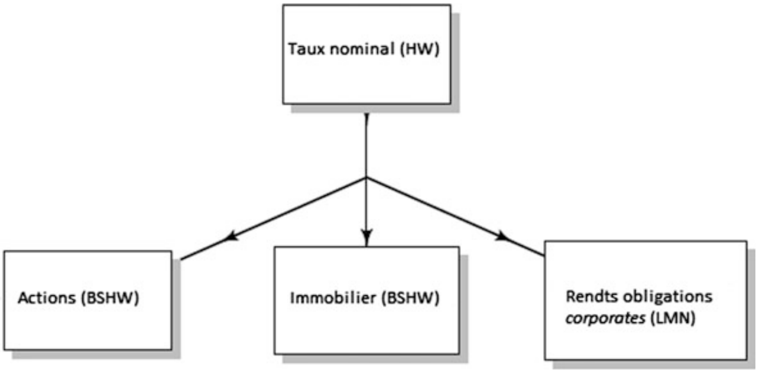

Fig. 4.9 ESG package structure

- and an additional liquidity spread

$$d\eta_t = \eta_t dW_t^{(\eta)}.$$

The instantaneous rate defining the return on corporate bonds is equal to $r_t + \lambda_t + \eta_t$ where r_t is the instantaneous short rate (nominal) given by the HW model, λ_t is the intensity of the Poisson process governing instances of default and η_t is the incremental return required by corporate bondholders to compensate for the liquidity risk they take investing in such bonds (Fig. 4.9).

Ultimately, the structure of the package is shown below:

Illustration of the Calibration: BSHW Model

The BSHW model provides a closed formula for the price of calls and puts (see Merton (1973) and Brigo et al. (2007)). In particular, for calls on shares not issuing dividends (e.g., an index):

$$C^{BSHW}(0,T,K) = S_0\mathcal{N}(d_1(0,T)) - KP(0,T)\mathcal{N}(d_2(0,T))$$

In this formula, S_0 is the value of the underlying asset for the date $t = 0$, T is the call's expiration, K is the exercise price, and $P(0,T)$ is the price at $t = 0$ for a zero coupon bond with maturity T given by the HW model. The latter equals the market price using the HW model. We have:

$$d_1(0,T) = \frac{\frac{log(S_0)}{KP(0,T)} + \frac{1}{2}v^2(0,T)}{v(0,T)}$$

and

$$d_2(0,T) = d_1(0,T) - v(0,T)$$

where

$$v^2(0,T) = V(0,T) + \eta^2 T + 2\rho\frac{\sigma\eta}{a}\left[T - \frac{1}{a}(1 - e^{-aT})\right]$$

and

$$V(0,T) = \frac{\sigma^2}{a^2}[T + \frac{2}{a}e^{-aT} - \frac{1}{2a}e^{-2aT} - \frac{3}{2a}]$$

Here, the BSHW model is calibrated on the basis of implied volatilities of the call options on the DAX, observed on July 5, 2002, and provided in Sep (2003). The value of the underlying is equal to $S_0 = 4468, 17$.

Expiry	Jul-02	Aug-02	Sep-02	Dec-02	Mar-03	Jun-03	Dec-03	Jun-04
Time	2 weeks	1 month	3 m	6 m	9 m	12 m	18 m	24 m
Date	19-07-02	16-08-02	20-09-02	20-12-02	21-03-03	20-06-03	19-12-03	18-06-04
T	0.0389	0.1139	0.2083	0.4583	0.7111	0.9583	1.4556	1.9528
r	0.0357	0.0349	0.0341	0.0355	0.0359	0.0368	0.0386	0.0401
Strike								
3400	0.6625	0.4875	0.4204	0.3667	0.3431	0.3267	0.3121	
3600	0.6007	0.4543	0.3967	0.3511	0.3279	0.3154	0.2984	0.2921
3800	0.5084	0.4221	0.3718	0.3327	0.3155	0.3027	0.2919	0.2880
4000	0.4541	0.3869	0.3492	0.3149	0.2963	0.2926	0.2819	0.2800
4200	0.4060	0.3607	0.3330	0.2999	0.2887	0.2811	0.2751	0.2775
4400	0.3726	0.3396	0.3108	0.2871	0.2788	0.2722	0.2661	0.2686
4500	0.3550	0.3277	0.3012	0.2781	0.2718	0.2661		
4600	0.3428	0.3209	0.2958	0.2740	0.2688	0.2627	0.2580	0.2620
4800	0.3302	0.3062	0.2799	0.2631	0.2573	0.2533	0.2504	0.2544
5000	0.3343	0.2959	0.2705	0.2540	0.2504	0.2464	0.2448	0.2462
5200	0.3460	0.2845	0.2624	0.2463	0.2425	0.2385	0.2373	0.2422
5400	0.3857	0.2860	0.2578	0.2399	0.2357	0.2327	0.2312	0.2351
5600	0.3976	0.2860	0.2607	0.2356	0.2297	0.2268	0.2241	0.2320

The missing values were determined by linear interpolation. Calibration is made on the price, not on the implied volatilities. This minimizes the objective function:

$$\frac{1}{2N}\sum_{i=1}^{N}\left(\frac{C^{BSHW}(0,T,K,\Theta) - C^{M}(0,T,K)}{C^{M}(0,T,K)}\right)^2$$

where

$C^M(0, T, K)$ is the call's market price, for expiration T and the exercise price K, as determined from the implied volatility of the market and the Black-Scholes formula.

$C^{BSHW}(0, T, K, \Theta) = C^{BSHW}(0, T, K)$, where $\Theta = (a, \sigma^{(r)}, \sigma^{(E)}, \rho)$ is the model's set of parameters.

a is the speed of the short rate's mean reversion, $\sigma^{(r)}$ is the short rate's volatility, $\sigma^{(E)}$ is the volatility of the shares, and ρ is the coefficient of the instantaneous linear correlation between the share price and the short rate.

This function is difficult to optimize because it has many local optima. Here a differential algorithm is used in the `DEoptim` package, which does not use objective function derivatives in the search for the solution, but rather a genetic algorithm.[12] To find the appropriate parameters, the number of iterations and algorithm generations must be increased, and the convergence of the parameters and the objective is observed using the `plot.DEoptim` function.

The code used to make the estimate is available from the authors upon request.

Validation of the Simulated Prices

We will now see how prices from the calibrated model are validated. The principle is to build a confidence interval of 95 % around:

$$\mathbb{E}\left[e^{-}\int_0^T r_s ds \frac{S_T}{S_0} - 1\right]$$

and

$$\mathbb{E}\left[e^{-}\int_0^T r_s ds \frac{max(S_T - K, 0)}{C^M(0, T, K)} - 1\right]$$

being the mean relative deviations between the observed prices and the prices simulated by the package ESG, for T and K fixed. We therefore obtain:

[12] http://cran.r-project.org/web/packages/DEoptim/index.html.

```
set.seed(123)

# Simulation of short rates and equity prices
sim <- rStock(horizon = 2, nScenarios = 10000, ZC = ZC,
      vol = res$optim$bestmem[2], k = res$optim$bestmem[1],
      volStock = res$optim$bestmem[3], stock0 = S, rho = res$optim$bestmem[4])

# Deterministic sport rates
r1 <- r[(tau == 0.9583)]
r2 <- r[(tau == 1.9528)]

#  Deterministic discount factors (observed zero-coupon prices)
df1 <- exp(-r[(tau == 0.9583)]*0.9583)
df2 <- exp(-r[(tau == 1.9528)]*1.9528)

# Stochastic discount factors
df.stoch <- exp(-t(apply(sim$shortRatePaths, 1, cumsum)))
df1.stoch <- df.stoch[, 1]
df2.stoch <- df.stoch[, 2]

# Percentage difference between simulated and observed
# zero-coupon prices
colMeans(cbind(df1.stoch/df1-1, df2.stoch/df2-1)*100)

## [1] -0.6127 -0.1419
```

Confidence interval of 95 % around:

$$\mathbb{E}\left[e^{-\int_0^T r_s ds}\frac{S_T}{S_0}-1\right]$$

for $T = 1$ and $T = 2$ years:

```
apply(cbind(sim$stockPaths[ , 2]*df1.stoch/S - 1, sim$stockPaths[ ,
3]*df2.stoch/S - 1)*100, 2, function(x) t.test(x))
```

```
## [[1]]
##
##  One Sample t-test
##
## data:  x
## t = 4.024, df = 9999, p-value = 5.767e-05
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.5498 1.5943
## sample estimates:
## mean of x
##     1.072
##
##
## [[2]]
##
##  One Sample t-test
##
## data:  x
## t = 5.101, df = 9999, p-value = 3.439e-07
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.222 2.747
## sample estimates:
## mean of x
##     1.984
```

Confidence interval of 95 % around:

$$\mathbb{E}\left[e^{-\int_0^T r_s ds}\frac{max(S_T - K, 0)}{C^M(0, T, K)} - 1\right]$$

for $T = 1$ and $T = 2$ years, and an exercise price equal to 3400:

```
(strike0 <- strikes[1])

## [1] 3400

mat.pricesBS <-  matrix(priceBS, nrow = nb.strikes, byrow = TRUE)
mat.pricesBSHW <- matrix(priceBSHW, nrow = nb.strikes, byrow = TRUE)
p1 <- mat.pricesBS[strikes == strike0, (tau == 0.9583)]
p2 <- mat.pricesBS[strikes == strike0, (tau == 1.9528)]
Payoff <- pmax(sim$stockPaths-strike0, 0)[ , -1]
apply(cbind(Payoff[, 1]*df1.stoch/p1 - 1, Payoff[, 2]*df2.stoch/p2 - 1)*100, 2,
    function(x) t.test(x))

## [[1]]
##
##  One Sample t-test
##
## data:  x
## t = 0.7819, df = 9999, p-value = 0.4343
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -1.023  2.380
## sample estimates:
## mean of x
##    0.6785
##
##
## [[2]]
##
##  One Sample t-test
##
## data:  x
## t = 3.638, df = 9999, p-value = 0.0002765
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.873 6.251
## sample estimates:
## mean of x
##     4.062
```

The differences between the prices obtained by simulation and the market prices are relatively low for the selected parameters.

Ideally, the confidence interval should contain a 0. Even if it were possible to obtain such a result, it is important to note that due to the annual forecast adopted in version `v0.1` of the `ESG` package, and the interpolation of the yield curve, the estimation of stochastic discount factors (including evaluating the integrals) remains rather rough.

4.4 Conclusion

Overall, in the end it seems to us that the purpose of these reflections and the work ahead is not so much to locate an ESG that accurately reflects the behaviour of modelled assets,[13] but rather to provide a simple, coherent explainable framework that allows fuelling asset/liability models with credible scenarios shared widely by the market, at least for Pillar 1–the common base.

A goal could be set to establish a number of minimum conditions that an ESG should respect, i.e. that however imperfect it may be, it should be robust enough to ensure its users consistent and comparable economic scenarios, which report financial risks with reasonable effectiveness.

Such a framework requires coherent management with a "snapshot" view based on observed prices from the financial markets ("risk neutral" ESG) with a dynamic vision that aims to provide information on price distribution at various future dates ("historical" ESG). This particular need is at the core of ORSA, as we shall see in the next chapter.

As we reach the end of this chapter, it appears that putting together one's own ESG is within the scope of every organization. Building one's own scenarios is key to a genuine mastery of risk management by imposing an a priori reflection on risk exposure, their modelling and the consequences of choices made at this level.

[13]Taken from Planchet and Leroy (2011).

Chapter 5
From Internal to ORSA Models

Frédéric Planchet and Christian-Yann Robert

Abstract Pillar 1 of Solvency II framework set out quantitative requirements for calculation of the Best Estimate (BE) of liabilities and the Solvency Capital Requirement (SCR) (a 99.5 % value-at-risk measure for net assets over one year) using either a standard formula given by the regulators or an internal model developed by the insurance company. Its implementation for a life insurance company leads to complex cash flows projection models that aim to allow an accurate calculation of the BE (the SCR is most often determined using the standard formula). Meanwhile, the Own Risk and Solvency Assessment of Pillar 2 left the insurance company to define an optimal entity-specific solvency constraint on a multi-year time horizon. Its implementation requires in particular to model the distortion of the distribution of the entity's coverage ratio over the period of the strategic plan. For this, it is needed to use much more aggregated models for the insurance company, not only to take into account computation and time constraints but also to ensure a minimum robustness to projections. Therefore, models adapted to these constraints are developed within a framework of probabilistic approaches and scenario analysis. The insurer is faced with the need to coordinate these aggregated models with the analytical view of Pillar 1. We propose in this chapter to discuss and challenge the way to bring back together models needed for the quantitative requirements of Pillar 1 with quality requirements of global and appropriate risk management system of Pillar 2.

5.1 Introduction

Until today, Pillar 1 of the Solvency II Directive focused the main efforts of the insurance companies towards modelling. This Pillar introduced a prescriptive approach based on flow projections to calculate a best estimate, and based on

F. Planchet (✉) · C.-Y. Robert
Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France
e-mail: frederic@planchet.net

F. Planchet
PRIM'ACT, Paris, France

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_5

considering a margin of additional risk when calculating the provisions and defining capital requirements to ensure the company's solvency up to one year with a 99.5 % probability. This level is generally calculated using the standard formula, along with some specific possible adjustments to the risk module within the context of a partial internal model.

This first Pillar imposed an "economic" evaluation of the balance sheet and, with life insurance, it enabled real progress in terms of identifying and measuring risks by defining a common methodology for calculating provisions leading, at least in principle, to a certain transparency and comparability of values. It also helped to make insurance companies more aware that for each quantifiable risk a capital requirement was needed and could be calculated quite simply using the standard formula.

However, in terms of managing the business of the company, decision-making and risk management, it remains insufficient. Indeed, even if the flow projections are carried out to the end of the commitments, it still remains a one year short-term solvency calculation for which medium-term control and management rules are only partially taken into account.[1]

The purpose of Pillar 2 of Solvency II is to lay the foundations for a formalized and accepted framework of risk control for insurance companies. It requires managing risk by establishing an accountability framework: a risk management function, an actuarial function, an internal control function, and a comprehensive and effective risk management system. This system must ensure sound and prudent management with permanent risk control and it should help ensure that the business is well capitalized according to the objectives set out in its strategic plan.

The Own Risk and Solvency Assessment (ORSA) is central to Pillar 2, allowing the construction of a robust prospective model that can forecast scenarios showing the coverage ratio becoming distorted over time.

Unlike Pillar 1, which imposes modelling methodologies and capital requirement calculations,[2] ORSA allows greater freedom in the choice of risk metrics, and calculation and forecasting methods—provided they adapt to the characteristics and particular situation of the company. In particular, it aims to help insurance companies:

- To better understand their own risks and the actions to be taken to manage or respond to an unfavourable situation;
- To determine, beyond what is used in the standard formula, a specific amount of capital to their risks adapted to the strategic plan;
- To strengthen a culture of risk management at all levels and develop a genuine Enterprise Risk Management[3] approach (ERM).

[1]These rules are viewed within the context of projection models for best estimate provision calculations (see Guibert and Planchet 2014).

[2]However, it should be noted that it does not impose any requirements for the calculation of provisions.

[3]More information about ERM can be found at: http://fr.wikipedia.org/wiki/COSO.

Articles applying to the Solvency II Directive that make reference to ORSA are Articles 45, 246 para. 4 and 120. In practice, three separate evaluations are applied in the ORSA context (see French Prudential Control Authority (ACPR) (2014), the headings of which are listed below):

- Evaluation of the overall solvency needs (OSN);
- Evaluation of ongoing compliance with regulatory obligations regarding Solvency Capital Requirements (SCR), Minimum Capital Requirements (MCR) and requirements for the calculation of technical provisions;
- Evaluation of the extent to which the organization's risk profile deviates from the assumptions underlying the SCR calculation.

The ORSA results must be sent to the Prudential Control Authority (ACPR). The ACPR recommends that the three reports be concise and different from ORSA'S internally documented written policy. The links between the forwarded Pillar 1 and 2 documents may be summarized as follows (see Fig. 5.1).

Based on a comprehensive risk analysis of mainly pricing, provisioning, commercial and financial risks, the challenge of ORSA will be to provide models consistent with Pillar 1, which provides reference values, and is robust enough to allow projections over several years—a potentially complex task.

This is not only about calculating the net asset value (NAV) provisions and the SCR, but also about providing information on forecasting these values into the future. This has to do with setting up appropriate models according to the context such as non-life insurance *versus* life insurance, financial risks versus technical risks, the number of commitments already made *versus* new production, etc.

The aim of this section is to describe how these models reconcile those developed under Pillar 1 of the Directive with the establishment approach of an internal model.

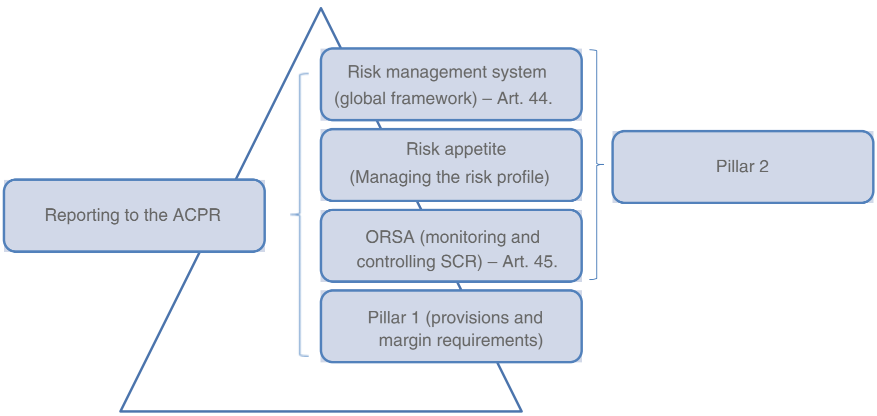

Fig. 5.1 Coordination of documents submitted to the supervisory authority for Pillars 1 and 2

This section is organized as follows: first we will review the ORSA structure from a global perspective, emphasizing in particular the coordination of the qualitative and quantitative aspects of this process. Thereafter, and in greater detail, we will outline the quantitative angle of this process, presenting the thinking behind the models the insurers developed to meet the production needs of the various indicators. Lastly, we discuss the impact of these choices on model positioning and their coordination with the first Pillar's inventory calculations, using either the standard formula or eventually an internal model.

5.2 What Is ORSA?

ORSA's main objective is to set up a governance system that uses a forward-looking risk self-assessment process in order to:

- Define and adapt both the strategic plan and risk policies;
- Effectively manage the business of the company;
- Create and conceptualize possible new products; and
- Meet regulatory requirements.

The ORSA process can be schematically represented as follows (see Fig. 5.2).

It thus appears that implementing ORSA is not merely a question of producing calculations and reports, but involves establishing an iterative process, itself composed of processes and sub-processes involving multiple company functions—functions that are both actors and users of ORSA.

Four requirements underpin this process:

- The need for an exhaustive list of individual risk profiles because an oversight or misidentification of a risk profile can be as serious as a poor quantitative assessment;

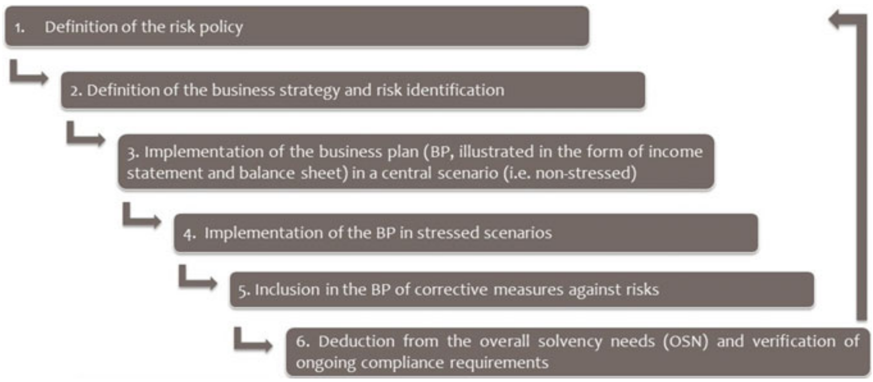

Fig. 5.2 Structure of the ORSA process

- The need to evaluate and manage global solvency requirements using a proactive approach towards strategic planning so that strategic goals become achievable without putting the company at risk of insolvency;
- The need for permanent monitoring and an ongoing approach to solvency management between the time the annual accounts are closed and the time the financial statements are fully filed with the supervisory authorities;
- The need to analyze and measure changes in the company's risk profile according to the assumptions underlying the SCR calculations.

Overseeing the ORSA process requires defining:

- The input and output elements of the process;
- The reliance on other processes—the budget process, for example;
- The sequence of the operations, the checkpoints and the validation steps;
- The responsibilities assigned to the process such as steering and contributions, etc.; and
- The associated planning.

The Administrative Management and Supervisor Body (AMSB) is directly involved in the process. It is vital that the AMSB be aware of all significant risks and be conscious of the fact that, by managing the process and questioning its results, it plays an active part in ORSA.

5.2.1 The Logic Behind the Three Evaluations

Meeting the demands described above means having to produce quantitative indicators at various levels.

More precisely, the effective implementation of ORSA requires determining the three quantitative evaluations and drafting their respective reports.

1st Evaluation: Overall Solvency Needs (OSN)

The OSN takes into account the company's risk appetite and risk profile. Its coverage is made with capital (stress tests are set up) and a range of mitigation measures, which might typically include:

- Revising the underwriting rules to avoid anti-selection;
- Strengthening provision monitoring;
- Setting up an adjusted reinsurance program;
- A cost control plan;
- Adjusting prices; and
- Revising asset allocations, etc.

The OSN evaluation thus leads to a demonstration that, according to the business plan, the entity's risk appetite is being met and that it is also able to meet its strategic objectives even under adverse situations. This demonstration is based on the risk profile and is reflected both by a suitable equity capital level (quantitative coverage) and by means of suitable risk control (qualitative coverage).

At this point, two interpretations can be retained to define the OSN:

- It can be seen as a capital complement that the insurance company must maintain beyond the SCR to make sure that even under stressed scenarios it satisfies its risk appetite requirements. Example: 70 % of the SCR, knowing that the risk appetite is 150 % of the SCR and that even under stressed situations solvency does not fall below 150 %;
- It can also be understood as a new capital measure that reflects risk appetite in the same way as the SCR.

The second choice is often retained and the OSN is considered to be a measure of equity capital in the company, used to gain a better understanding of its risks—risk estimates unquantified by the standard formula, different matrix correlations for example, arising from evaluation 3 confronting the standard formula, etc. According to this logic, the OSN is greater than or equal to the SCR.

The company compares its equity capital across all scenarios to the estimated OSN to make sure it has sufficient capital. If not, it lists the measures it intends to take to cover.

The company is free to choose the methods it decides to use to conduct the OSN evaluation as long as these methods are proportionate to the nature, extent and complexity of the risk being considered and that it is able to demonstrate the relevance of these methods.

2nd Evaluation: Continual Respect for Statutory Solvency

The second evaluation explicitly introduces a dynamic aspect and a related concern of prospective compliance with the requirements of Pillar 1, taking into account the business plan. It is no longer a matter of simply verifying the initial creditworthiness of the entity, but of demonstrating that compliance with the various regulatory and corporate requirements is guaranteed over the duration of the business plan.

For example, this assessment may be done through the multi-year stress test results by measuring the shock impact on the central scenario—increase of the S/P, costs, and lower financial yields, etc.

Its implementation is potentially complex, since it implies calculating solvency indicators at various dates in the future, which are random variables when seen from the original date. Models may be more or less complex depending on the approach to this randomness in the calculation, and on whether one wants to provide information on the distribution of the indicator or simply calculate a few key values for specific situations (and, more particularly, opposing ones). This point will be discussed below.

3rd Evaluation: Analysis of the Variances Between the Risk Profile and the SCR Calculation Assumptions

This evaluation's objective is to check the quality and suitability of the risk profile against the regulatory measures. This leads to verifying the suitability of each SCR sub-module calculation and discussing the reasons for which the standard formula is adequate or not. This verification primarily concerns major risks such as financial risks, underwriting risks and operational risks. Complementary, strategic or environmental risks, for example the establishment of an Inter-professional National Agreement (ANI),[4] key-man, regulatory or reputation risks, should also be part of the discussion.

This evaluation is not simply academic and should not constitute a general criticism of the standard formula or the calculation principles of the SCR, but should raise awareness of the standard formula's partial suitability to the particular risk of the entity concerned in measuring the financial impacts of shocks looking 1 year into the future at a 99.5 % level.

The Prudential Control Authority (ACPR) particularly stressed the fact that a high solvency ratio was not a sufficient reason for not making this evaluation and that it was necessary to ask the question whether the standard formula has a potential modelling error (see ACPR 2014).

5.2.2 Risk Appetite

The strategic plan's objectives must be formulated within the context of risk appetite. This, however, is not explicitly defined in the Directive and the definitions used by the various market players are not always homogeneous. We will adopt the definition proposed by the Institute of Actuaries (see Institute of Actuaries 2014):

> An insurance company's risk appetite corresponds to the level of maximum risk that the entity agrees to take to achieve its strategic objectives. It is determined by the company's corporate governance, and expressed in the measures it takes towards risks and limits.

The formulations usually encountered rely on various complementary measures and additional criteria such as: a coverage rate floor on the solvency margin, eventually accompanied by a probability of occurrence over the projection period (solvency dimension), a deviation of the result compared to the estimated value (result dimension) or even a loss of shareholder value (equity capital dimension).

Risk appetite formulas also set a horizon or time limit to the objectives. This can be the duration of the strategic plan for the solvency dimension, an annual horizon for the results dimension or enterprise value.

---

[4]Inter-professional National Agreement (http://direccte.gouv.fr/IMG/pdf/ANI_securisation_de_l_emploi-2.pdf) of 11/01/2013.

Criteria levels are set, either on an a priori basis, for example by maintaining the minimum coverage ratio of the solvency requirement—with values set relative to the solvency margin S1 from 200 to 300 %, or to the SCR from 125 to 200 %—or are deduced from the business plan and the stress tests that were previously carried out.

5.2.3 Business Plan and Risk Profile

The company's objectives are expressed in a strategic plan and/or a business plan. These are the reference documents for the prospective component of the evaluations to be carried out. Nevertheless, using these references is not straightforward; most of the time it requires adaptation:

- Existing business plans are, a priori, not well aligned to the typology of risks;
- Several "models" or "versions of models" coexist (strategic plan *vs* updated budget, for example).

Similarly, with regard to risk mapping (risk profile), what already exists is not always usable 'as is', mostly because they are often operational risk mappings that do not take into account technical and financial risks. These mappings also tend to be very detailed using a bottom-up approach compiled from the business process, or very macro using a top down approach. The use of mapping on quantitative terms is not always directly possible due to a low base of incidents and the existence of expert evaluations without a common metric. Like business plans, these descriptions are not always related to the Pillar 1 risk matrix.

It should also be stressed that in the context of ORSA, it is not only necessary to identify and describe the risks, but also to specify the risk reduction measures in place.

One of the results of setting up ORSA should be to bring together the structure of these documents to make them part of the overall risk management process.

5.3 ORSA'S Quantitative Approaches

Here the focus has been on forward-looking calculations (evaluation no. 2); any capital additions used to determine the OSN do not in principle fundamentally change the Pillar 1 calculations, even if different techniques can be used. For some risks, i.e. sovereign risk, the insurer can simply reuse the standard formula risk modules such as the spread module or the counterparty risk module.

Implementing the forward-looking quantitative aspects of ORSA involves:

- Calculating an economic report, which requires a potentially complex model for determining provisions;

- Calculating the SCR, which also requires a model—possibly the standard formula;
- Projecting forward the above calculations in a deterministic (scenarios) or probabilistic logic to build coverage ratios over a period of several years.

Various levels of analysis can be considered and constructing a relevant model requires prior consideration of the major risk factors and their interactions.

More precisely, depending on the context and nature of the risk to be modelled, specific modelling choices can be made to describe the asset, the calculation of provisions and thus the net asset value (NAV) and the SCR calculation, such as the proportionality rule, the standard formula or an "internal model".

At least two forecasting architectures are possible:

- A scenario approach having a central scenario and adverse scenarios; and
- Setting up stochastic models to forecast balance sheets.

The choice strongly depends on the portfolio's complexity:

- For property and casualty (P&C) or for a straightforward personal insurance portfolio with activity concentrated in one type of risk, the stochastic approach allows a focus on the major risks and produces a wealth of information on the distribution of future budgets (see Guibert et al. (2012) for an illustration of possible modelling types).
- In more complex situations, the scenario approach allows a broader use of Pillar 1 forecasting tools.

5.3.1 Scenario Approach

Reasoning about explicit scenarios and not on various scenarios derived from risk factor assumptions facilitates the analysis of the impact of these factors. This makes explaining the process to third parties easier, particularly appropriation of the ASMB approach. That is why this approach is widely used by organizations. Below, we provide some clarity as to how the central scenario and the adverse scenarios should be structured.

5.4 Structuring the Central Scenario

The central scenario should be consistent with the basic assumptions of the entity's business plan. In particular, it must reflect the strategic assumptions set by the AMSB and the relevant economic and market assumptions before any consideration of impacting events. In a sense, it is the implementation of the business plan's central assumptions.

It involves forecasting the “financial statements” (Solvency I) and Solvency II and the income statement. This would require having the flow of liabilities broken down by subscription—the initial amounts and then the years of future contributions.

In the absence of strong asset/liability interactions, forecasting technical provisions is possible by simply assuming that the risk margin is proportional to the duration or to the best estimate; calculating the best estimate is done by updating the ongoing flow.

In a savings context, the evaluation made with the Pillar 1 tools can be relied on to determine the projected balance sheets and the SCR, which involves making Pillar 1 calculations on each forecasted date—typically 3–5 in addition to the initial calculation. Simplified approaches might be considered by building on a projection of adjusted policy liabilities as a second step to take into account the value of options and guarantees (e.g. following Bonnin et al. 2014a).

The SCR calculation can be made at first by pegging the reported SCR ratios to the best estimate (liability risks) or to the market value (market risks) observed in the initial balance sheet.

More refined approaches can be imagined by initially using a standard formula calculation at each forecasted date. At this point it can be seen that the most sensitive factor in the progression of the coverage ratio is the NAV and that the SCR is more stable overall. Therefore, the challenge lies more in a better appreciation of the change in NAV at each date than the level of the SCR itself. The key point is thus the calculation of provisions projected “all along” the path of the scenario.

5.5 Building Stress Scenarios

It is necessary to test the resistance of the balance sheet and coverage ratios over several possible development scenarios against various risk factors that impact them to address the permanent forecasted solvency requirements (evaluation no. 2).

This is a question of describing adverse, but not extreme, situations and measuring the impact on key values of the balance sheet (assets, liabilities, and equity) and on coverage (SCR and MCR).

The scenarios considered should cover all of the main risks that might have an impact on solvency. In financial matters, we consider rate levels with their slope, the value of equities and real estate with their volatility, and the level of their spreads, etc. As far as possible, we also consider cross effects. As regards insurance risks, consideration is made for the loss ratio, market share levels, the amount of management fees, and the levels of redemptions, etc. However, to remain operational and provide clear signals, the number of scenarios should remain limited.

Defining severity levels can be related to the probability that such a scenario will play out, but this is in no way compulsory. It is possible to rely on the shocks of the standard formula to recalibrate at a different level.[5]

In cases of calibrated scenarios with reference to a probability of occurrence, calibrating the shock amidst several collectively degraded risk factors is not easy. Such an operation involves defining the dependence structure according to the risk factors.

The severity level can also be determined by expert opinion, which is the case for economic scenarios combining several factors including inflation and long rate levels, and for "catastrophe" type scenarios like pandemics and terrorism, and scenarios associated with legal and tax risks, as well as scenarios for operational risk as in risks for which only a small amount of data is available or for which the environment is very variable.

5.5.1 *Probabilistic Approaches*

Probabilistic approaches intend to provide coverage rates over a period of several years providing a more comprehensive view of the current and future solvency of the company. Inherently, these approaches are more complex and costly in terms of implementation and the time it takes to calculate. They run up against the diversity of contracts, commitments, the nature of the risks, and their management methods, etc. In addition, with life insurance and its numerous asset/liability interactions, this creates additional difficulties. They require a configuration that is able to provide an average coherent development with the central scenario of the deterministic approach, but should also provide variability to all risk factors on the company's creditworthiness.

Therefore, it is necessary to simplify the calculations related to the number and the choice of scenarios generated, on the analytical approximations of conditional expectations, and on the number of risk factors considered, etc. We describe below a few proposed methods for evaluating the SCR or the provisions (Pillar 1) that can be included in a dynamic framework.

Choice of the Extreme Scenarios

When calculating the SCR and under the simulations in simulations method, Devineau and Loisel (2009) propose selecting primary scenarios that generate extreme results likely to be used when calculating equity capital. They begin by observing that knowing the entire distribution of equity for a year is unnecessary, since only the 99.5 % quantile needs to be determined. The method involves setting a standard to identify first year scenarios that are dangerous risk factors affecting

[5] See http://actudactuaires.typepad.com/laboratoire/2011/03/recalibrer-les-chocs-du-mod%C3%A8le-standard-.html.

equity the most. Although they provided very good results in the example studied in Devinau and Loisel (2009), the method seems difficult to develop within an insurance company framework having a diversified business.

The Weighted Monte Carlo Method

When calculating the SCR and the simulations in simulations method, one way to reduce calculation time is to reduce the number of secondary simulations. Avellaneda et al. (2001) proposes to reduce the number of scenarios using a weighting system to reflect constraints such as a lack of asset arbitration. This method only allows a moderate reduction in the number of simulations to reach a sufficient level of accuracy.

The Least-Squares Monte Carlo Method (LSMC)

When calculating conditional expectations as seen, for example, in the calculation of the NAV or the best estimate, the idea is to provide an approximation of these conditional expectations by linear combinations of known functions of major risks on the basis of a least squares regression. This is a financial technique used to assess exotic options (see Longstaff and Schwartz 2001) and applied for the first time to insurance by Bauer et al. (2010) in a simplified framework and then by Nteukam et al. (2014) on a real portfolio. It has emerged as the method of reference for ORSA and for calculations made within an internal model.

An analysis of the strength of the parametric form is paramount as it defines the ability of the form to be used outside the scenarios used to calibrate the model. The model should be regularly confronted with reality and tested within the realm of the risk factors.

Replicating Portfolios

The idea behind this replicating portfolio method is to create a portfolio of shadow financial assets on the market that has the same market value as the portfolio of liabilities or the revalued net assets (Revelen 2009). These fictitious financial assets must be able to be valued using closed formulas to avoid calculating under the risk-neutral probability of the value of commitments and thus reduce the calculation time. The choice of assets is essential and must be made according to the nature and convexity of the options included in the liabilities.

The main limit of this approach lies in its inability to price risks other than those related to listed financial instruments such as biometric risks (risk of mortality), and the risk of increasing fees, etc. It should be pointed out that the LSMC approach described above may be seen as a generalization of the replicating portfolios using “assets” (risk factors) better suited to describing liabilities.

Closed Formula Methods

The principle behind these methods is to use closed formulas for conditional expectation calculations in a wider context than that used for financial products. The idea is to propose comprehensive modelling approaches aggregating the interest variables of the insurer's balance sheet using a limited number of parameters to describe in a synthetic way the risks to the company and to obtain closed formulas for the main items of the balance sheet. The non-financial risks are modelled by means of simple models with dynamics that can be based on those used for financial instruments like shares or rates (see Bonnin et al. 2014a, b).

5.6 ORSA, Pillar 1 and Internal Model: Need for Coordination?

Pillars 1 and 2 will lead to important choices involving model positioning. The first Pillar's need for inventory calculations and the need for a prospective evaluation of the main items in the economic balance sheet have to be coordinated to ensure coherence, either within the framework of an intensive use of the standard formula or an eventual internal model.

Note that such coordination can be understood on two levels:

- Some models are unique to ORSA, for example economic scenario generators in historical probability or introducing uncertainty to a biometric risk (on this last point see, for example, Planchet and Tomas 2014);
- Other models are versions of the existing Pillar 1 models suited to the ORSA framework, for example an approximation on the basis of closed formulas for calculating best estimate provisions for a savings account contract valued with a classic Pillar 1 ALM model.

The focus points are not identical in both situations and the second is a priori more binding. It imposes checking that the risks assumed in both models are identical and that the outcomes of the approximate model, that of Pillar 2, provide a satisfactory approximation of the complete model.

5.6.1 Risks and Capital Requirements

ORSA introduces the concept of quantifiable and non-quantifiable risks. The SCR, in principle, should cover significant and quantifiable risks. A quantifiable risk is, therefore, defined as one that can be properly "covered" by capital—a loss over a one year horizon with a 99.5 % confidence level in Pillar 1, along with other possible definitions, both in terms of the extent of the risk and in the level of the risk in Pillar 2.

A non-quantifiable risk is one that cannot be properly "covered" by a quantitative solvency requirement. In this family of risk we find reputation risk, risks associated with poor strategic decision-making, and certain operational risks. The Directive stipulates that for these risks, companies must demonstrate an ability to implement appropriate management initiatives under the overall solvency needs OSN.

The concept of quantifiable risk must therefore be distinguishable from the concept of calculable risk. For example, operational or catastrophic risks are difficult to assess, or even to calculate, but are explicitly regarded as quantifiable by the Solvency II Directive.

OSN thus covers a global need for solvency having a wider scope than Pillar 1, since it covers all significant risks, whether they are quantifiable or not. Its assessment may be variously expressed as a potential loss that can be covered by capital or by an appropriate management initiative. It should be remembered that a significant risk for Pillar 1 is defined as a risk that could influence the judgement and decision of the AMSB.

It should also be noted that the risk of sovereign default, not included in the standard formula, must be taken into account under the overall solvency needs OSN, just as "commercial" risk is considered when acquiring a new business.

5.6.2 Projections

Evaluations under ORSA differ from those when calculating the SCR by the time horizon of the calculation (that of the strategic plan) and by taking into account the ongoing activity of the company. They therefore need to model the new business and be able to forecast the risk factors related to the company's activity over several years.

Sophisticated and thorough methods developed to assess SCR, including the standard form, and technical provision calculation models, generally cannot be reused to make projections in future years because they require important IT resources and because they are not prone to being forecasted due to the high number of parameters and information that needs to be collected. They generally have an important model risk.

On top of this, models adapted to calculating equity capital under Pillar 1 are not the most effective when forecasting equity as seen in the work of Bakshi et al. (1997).

Finally, the necessary level of detail makes it difficult for even those in charge to interpret the results of the models and hence there is a risk of not being understood by all stakeholders in charge of decision-making.

Moreover, the capital requirement results aside, ORSA must also define action protocols to address stress scenarios. For example, consideration may be made within the policy of risk to set alert thresholds that should lead to action taken by management. When the alert thresholds are approached, the model should allow an analysis of the different levers and initiatives that management could then implement. Only sufficiently flexible and interpretable models will help meet this objective.

5.6.3 ORSA and Internal Models

As previously mentioned, internal models and ORSA address different issues and in the end have little in common. The two concepts come together, however, on the following two topics:

- The identification of major risks associated with an assessment of capital needs; and
- ORSA evaluation no. 3, which would lead to identifying a significant deviation between measurements of risk actually borne and applying the standard formula.

However many points distinguish the ORSA internal model:

- The internal model allows a static calculation of a balance sheet and equity requirements at the inventory date, while ORSA is forward-looking in the sense that it is interested in the distortion of balance sheets;
- The internal model is strictly quantitative, whereas the ORSA process involves quantitative elements;
- ORSA is (save for some exceptions) not intended to quantify high quantiles of a loss distribution but to remain at the core of the distribution.

It can also be noted that, whether the entity uses an internal model or the standard formula to calculate the SCR, ORSA has no reason to change. What distinguishes the two calculations is simply the required calculation, the calculation that ORSA provides.

5.6.4 An Example of Economic Scenario Generators

Putting together economic scenarios (see Chap. 4) provides an illustration of the type of coordination between "Pillar 1" models and their extensions to meet the needs of Pillar 2.

Whichever approach is adopted (simulations in simulations, expert opinions on trajectories, or trajectory models) it is necessary to forecast financial instrument prices over the duration of the strategic plan. The comprehensive economic scenario generator (ESG) is thus summed up by the following diagram (see Fig. 5.3).

Based on prices for a given date (ESG historical data), the risk neutral ESG feeds a function for calculating market prices of financial instruments. In passing, we can see an analogy of this scheme with the classic pattern of representing balance sheet forecasts (see Guibert et al. 2014, p. 5) (see Fig. 5.4).

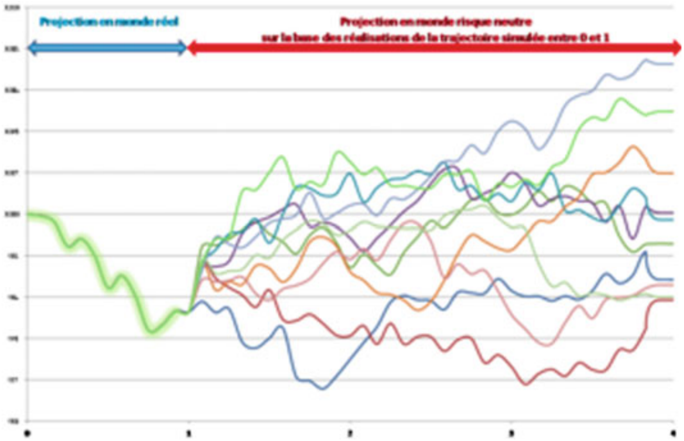

Fig. 5.3 ESG structure

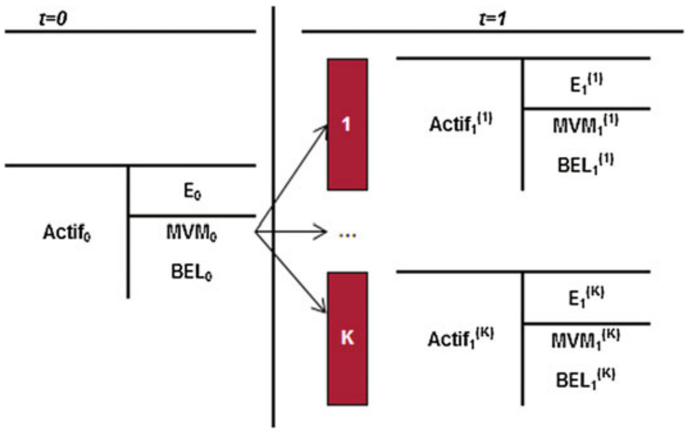

Fig. 5.4 Balance sheet forecasting within the ORSA structure

Therefore, setting up an ESG for ORSA gives rise to two problems:

- Which level of complexity should be used for the processes?
- Which forecasting dynamics for a risk neutral ESG?

Complexity of the Process in Play

The market has a habit of disconnectedly building two separate generators, which is one way of solving the difficulties associated with the price complexity of risky markets. Beyond this principle, which process to choose remains unresolved. However, a trend seems to be appearing that generally involves:

- Choosing simple processes for their functions, the goal being to enhance the best estimate of insurance liabilities, thus a hope within a context of outlining a significant normative dimension;
- Select processes potentially more complex for their variables, the objective here being to interpret changes to risk and thus calculate the quantiles of the distribution of returns.

Which Forecasting Dynamics for a Risk Neutral ESG?

As shown in the previous chart, the risk neutral ESG has a dynamic conditioned by achievements of the historical ESG: some of the parameters of the market price function depend on the values of the process.

This unpredictable nature depends on the processes selected at the first stage. Thus, the parameters of the following price model will remain constant as long as the historical ESG does not have a stochastic volatility process:

Indeed, within the underlying theoretical framework, going from historical probability to risk-neutral probability retains the volatility and, therefore, if this varies in historical probability, it is logical to also vary the risk-neutral probability to ensure a coherent approach.

We can, therefore, maintain that the ESG structure in historical probability forces configuring the risk neutral ESG used to calculate commitments: introducing volatility as a risk factor for forecasting historical probability implies a parameter recalibration of the "risk-neutral" generator, whereas in the absence of this factor, only the initial values are affected by the level of projected factors (see the section on ESG and Guibert et al. 2014).

5.6.5 Use of Proxies: An Illustration

In most cases, ORSA's forward-looking vision requires building a simplified version of a Pillar 1 model, especially to calculate provisions. We can illustrate this by using the deformation process of the risk neutral probability rate curve to calculate the best estimate commitments.

As explained in the section on ESG above, the projection model of flows used to calculate best estimate provisions for life insurance is powered by financial risk factor trajectories from "risk-neutral" ESG. The architecture of the model allows greater flexibility in the choice of the rate model and relatively complex approaches can be taken on this point in order to more precisely represent price structures of the relevant derivatives.

For ORSA, the need to recalculate provisions based on the strategic plan (within evaluation no. 2) to forecast balance sheets and coverage ratios leads to making a rough estimate. This is what is proposed by Bonnin et al. (2014b) for pension contracts by bringing the calculation of the reset option value closer to a closed formula. The willingness to obtain a closed formula highly limits in this example the dynamics of the "risk-neutral" deformation of the yield curve and even the yield curve itself, since the initial curve, provided by the European Insurance and Occupational Pensions Authority (EIOPA), is replaced by an adjustment from a Vasicek model.

It is proposed to discuss an example that is based on Sect. 2.2.2 of Chap. 3 of Guibert et al. (2014). The structure in nominal terms for year t $(R(t,\tau), \tau \geq 0)$ is represented in the model proposed in different ways depending on modelling requirements (the notations are not developed here):

- The baseline assumption at the initial date $(R(0,\tau),\ \tau \geq 0)$ is a given structure according to the Smith-Wilson model, without a simple parametric structure;
- The deformation of this structure over the forecasted years is based on Nelson-Siegel's parametric representation

$$R(t,\tau) = r_0(t)\varphi\left(\frac{\tau}{\tau_1}\right) + l(t)\left(1 - \varphi\left(\frac{\tau}{\tau_1}\right)\right) + c(t)\psi\left(\frac{\tau}{\tau_1}\right),$$

- Deflator calculations for services rely on a representation of Vasicek based on the short-term rate $dr_t = a(r_\infty - r_t)dt + \sigma dW_t$ that leads to the parametric form of the rate (see Planchet et al. 2011).

$$R(t,\tau) = r_\infty - \frac{\sigma^2}{2a^2} + \frac{\tau\sigma^2}{4a}\varphi(\tau\, a)^2 - \varphi(\tau\, \mathrm{a})\left(r_\infty - \frac{\sigma^2}{2a^2} - r_t\right).$$

The interactions between these representations are the following:

- The initial curve provided by EIOPA is used to calibrate the adjustment of the Nelson-Siegel parametric form.
- This parametric form is used with each forecast to calibrate the Vasicek model from which short-term rates are forecast (in risk-neutral probability) that are used to calculate discount factors.

In this context it is important to ensure coherence between these various representations. Calibration is thus determined to minimize the differences between these various curves. To ensure consistency when calculating present values between the initial curve and the curve from the Vasicek model, this must be corrected to ensure, *ex-post*, that the best estimate value $P_n(0,j)$ is identical with the $\sum_{j\geq 1} P_n(0,j)E^{P^a}\left(P_j\right)$ of the EIOPA curve and these same coefficients from the

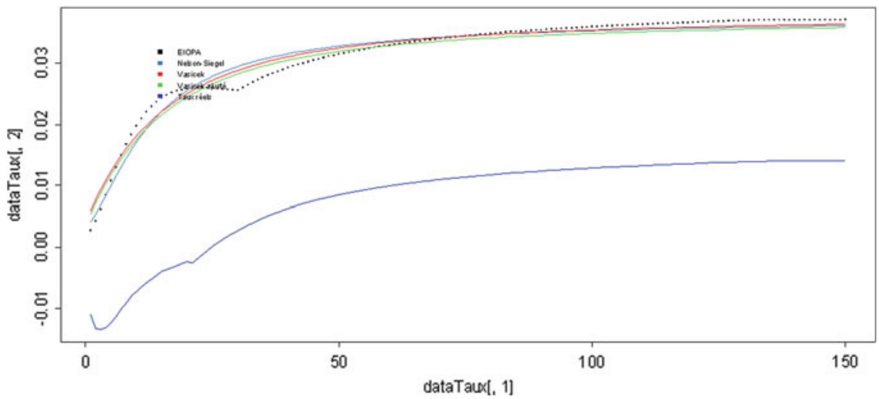

Fig. 5.5 Representations of the risk-free rate

Vasicek model. To achieve this, a spread adjustment is added to the zero-coupon rate from the Vasicek model. This spread is set at 0 and assumed constant during the forecast.[6]

As an illustration, using the EIOPA curve on 31/12/2012 as a reference, the following adjustments are obtained (see Fig. 5.5).

As this example shows, the same object is represented in different ways according to how the need and consistency is controlled via the level resulting from policy liabilities. As part of a more complex model involving multiple risk factors, this validation can be significantly more difficult to justify.

5.7 Conclusion

The objective of Pillar 2 of Solvency II is to allow the insurance companies to take a step back from Pillar 1 and its standard formula. ORSA should allow these companies to maintain their own vision of the overall risks to which they are subject, and their ability to cope with adverse events.

The standard formula provided solutions to assessing significant risks shared by the market by introducing a normative approach, but it required a rather cumbersome operational implementation. ORSA does not offer methodologies or models to evaluate their comprehensive and permanent solvency requirements, but it sets a framework.

[6]From a theoretical point of view one could use the generalized Vasicek model to use the EIOPA curve as a specific, but in practice, as this model is based on the instantaneous forward yield curve and since this is not known it should be approximated from the annual rate, the loss of accuracy would be modest in terms of the loss associated with the increased complexity of the model.

The Pillar 1 tools cannot be used directly and must be modified to provide consistent results compared to the reality of the market and the specifics of the company. These adjustments are motivated not only by the computational time constraints (it is not uncommon to be confronted with machine times of up to a week to produce results using the standard formula for a savings portfolio) but also by the poorly suited nature of a forward-looking Pillar 1 model.

There is, therefore, a strong need for models that are simple, flexible, and enable one to obtain satisfactory results that are also consistent with those of Pillar 1. This is not a question of undermining the standard formula, but of being able to explain the modelling differences needed to be closer to the reality of the business.

Chapter 6
Building a Model: Practical Implementation

Patrice Palsky

Abstract Financial modelling is an art as well as a science. In the Life Insurance world it is pushed to the utmost complexity, requiring exacting work and long term collaboration of many skilled professionals. Alas, it is only half the story: software development is the other face of modelling. It should be the easier part, but that's far from beeing guaranteed.

6.1 Introduction

In finance and economics, the terms “model” or “modelling” can be easily misunderstood. This is may be due to the considerable difference between theoretical models and operational models used in business.

Of course, a model remains a simplified partial representation of an object or system artificially isolated from its context. However, a theoretical model is used to establish broad principles. Its strength lies in its universality and in the intuitions it provides without having to transcribe it into computer language.

The Black-Scholes model is a familiar example of a purely theoretical model based on a radically simplified representation of financial markets. For evaluating put and call options, the Black-Scholes equation can easily be summed up. In terms of software development, the subject of this article, the corresponding calculation could be encoded in less than 15 min[1] with any such procedural language.

Conversely, a model used to create realistic simulations of complex systems is less concise and its computer application much more onerous.

A multitude of elements have to be taken into consideration, the scientific interest of which is marginal, e.g. regulations, taxation, and trading days, etc., although the combined impact of these elements is crucial for the results obtained in the “real world”.

[1]Although, in fact, this formula is the culmination of several decades of academic research!.

P. Palsky (✉)
Fractales, Paris, France
e-mail: p.palsky@fractales.fr

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_6

For a climatologist, a construction engineer or an actuary, the difficulty of creating a model does not lie so much in the sophistication of the mathematical tools used, as in the number and variety of items, functions, events and relationships that together need to be represented.

From this point of view, life insurers' asset-liability models are by far the most complex models of the financial world. Indeed, the nature of life insurance contracts combines with ambitious European regulations to create incredibly exacting requirements:

- A compliant asset-liability model must take into account the totality and diversity of the insurer's assets and liabilities.
- It must make detailed forecasts reaching decades into the future.
- It must calculate the forecasted cash flows per contract, taking into account the expected behaviour of the insured parties and predictable management actions, particularly in terms of investments and interest rates.
- It also has to calculate the technical and financial results across a number of accounting standards and operate both in run-off based on existing contracts and on a going concern basis taking into account future premiums.
- Finally, it must adjust for deterministic economic scenarios as well as stochastics, historical probability and risk neutral events.

The complexity of asset-liability models is not due to perfectionist actuaries or the vast imaginations of legislators. It is caused by the complex financial options found in life insurance contracts as well as the pooling of financial results.

By way of comparison, the models banks use to calculate portfolio NAVs or measure financial risks do not face the same constraints:

- According to regulations, credit risks, trading risks, and operational risks are evaluated separately;
- The various product categories can be analyzed separately using different models—a simplified approach to combining risks can even be used;
- Jointly simulating resources and commitments is not usually necessary in the absence of significant interactions[2];
- It is not necessary to model future operations, as they have no impact on the outcome of the present calculations.

Designing an asset-liability model on the scale of a large life insurer requires bringing together specialized skills. In IT terms this complexity will inevitably lead to high development and maintenance costs.

However, what is even more worrisome: a comprehensive model[3] could involve writing several hundred thousand lines of code. A robust architecture and advanced

[2]Aside from exceptions. In France, this exception is the Home Ownership Savings Plan, the balance of which depends on the behaviour of customers in terms of savings and mortgages.

[3]Namely, covering the Pillar 1 calculations, the quantitative portion of the Own Risk and Solvency Assessment (ORSA) and at least partially using financial decision support tools.

engineering must be deployed to prevent such huge software from collapsing under its own weight.

In practice modelling projects call on intensive resources, last many years and, in the end, they have a significant risk of failure.

Of course, the same thing can be said about many other types of software! Nevertheless, computer scientists and organizers have developed extensive project management expertise, and the success rate of software development is steadily improving.

It remains to be seen whether proven IT methods can be effectively implemented in the field of modelling. In fact, in most cases IT departments only recently stepped in, if at all, to develop asset-liability models.

6.2 Caveat Emptor

Before embarking on designing and developing a model, it is a good idea to impartially assess the company's requirements then set objectives in terms of applications not features.[4]

Only as a second step does it become possible to take note of the required features and then compare software solutions from various sources to finally put together a plan of action.

If this principle seems rather over obvious, it is not consistently applied. The idea of "starting from what exists" is economically attractive, even if it is more often a recipe for disaster…

Well in advance of Solvency II and before the first ALM models, three types of forecasting tools were being used by insurance companies:

- Actuarial tools for inventories that operate by discounting future claims and benefits;
- Financial models used for optimization (the efficient frontier), forecasting yields and market values, or risk calculations such as Value at Risk or Expected Shortfall;
- Models used to calculate embedded value, both "traditional" or market consistent.

6.2.1 Inventory Tools

Under the insurance accounting standards, technical provisions are calculated according to regulatory standards: mortality tables, no redemptions before the contractual deadline, no extensions, rates offered equal to the guaranteed minimum rates, and discount rates at a set technical rate, etc.

[4] A developer's ultimate goal is the business application, the features are only the means to achieve this.

Yet, these so called “prudential” evaluation standards are not necessarily all that prudential. They are sometimes far removed from economic realities and the operation of contracts—things required to adequately assess risks.

By definition, an inventory is made policy by policy, unheeding of future payments, much less concerned with probable payments on contracts not yet concluded. Still, the partial pooling of financial products between contracts in euros and even between successive contract generations implies that only a comprehensive model and going concern[5] balance sheet is capable of measuring financial risks.

Inventory tools can calculate provisions to the penny. However, this accuracy may make sense for *earned value of savings*, but for *provisions*, as the term is commonly understood,[6] this is a total illusion; what statisticians would call *exactly false*.

Inventory software provides a poor basis when putting together an asset-liability model, not only because a lot of work is needed to complete the functions, but also due to its rigid and prescriptive philosophy far removed from that of an economically realistic model.

6.2.2 Financial Models

A wide variety of financial software exists to manage collective investment funds, optimize risk-return ratios, and evaluate credit,[7] pricing and derivative trading risks.

Reusing some of the functions of these tools in an asset-liability insurance model can be expected, at least for simulating investments, which is normally half the battle. Experience has revealed the impracticality of this approach because these software programs have the following features in common:

- A pure market value approach is insufficient in life insurance in that compiling accounting values and preparing an income statement are required to forecast profit sharing;
- A static base[8] that does not allow globally forecasting financial risk due to the partial pooling of financial income previously referred to regarding inventory tools;
- Too short of a horizon as far as insurance liabilities are concerned, sometimes even resorting to a “snapshot” view such as a one day or one week trading risk horizon.

[5]Meaning, for example, that modelling future premiums allows measuring the dilution risk in bond yields in a situation of declining rates.

[6]This refers to the amount needed to cover the likely cost of liquidating commitments.

[7]Particularly in the context of Basel II or III.

[8]Cash flow and future investments and divestments are not taken into account.

Even scenario generators used in certain financial software to evaluate long options or measure credit risk were found inadequate due to their market by market specialization and/or their limited horizon.

6.2.3 Embedded Value Calculation Tools

A classical method of evaluating the future "embedded" results of a portfolio of contracts, the embedded value calls for a realistic projection of benefits and results.

Moreover, in their latest incarnations, the EV market and EV European are close to the Fair Value evaluation method recommended by Solvency II.

This type of tool may become a starting point for developing Solvency II Pillar 1 applications—calculating Best Estimates and capital requirements according to the standard formula. However, do keep in mind the following points:

- The Best Estimate is usually calculated within the context of an overall balance sheet model, including some reserves and provisions that contribute to the pooling of financial income between entries.[9] Conversely, EV calculations are made by isolating equity on one side and contracts on the other, often the latter are then evaluated by sub-portfolios.
- Applications for EV tools are limited to assessments, most often on a static basis, not including new business. They do not cover forecasts, risk studies or the quantitative aspect of ORSA.
- Embedded values are not often calculated, and therefore do not impose performance constraints. In addition, it frequently involves multiple manual operations and does not provide the expected traceability guarantees of a regulatory application.

6.2.4 Asset-Liability Models

Initial asset-liability models were first created in the early 1990s to help with financial decision-making, along with two complementary objectives:

- Measure the risks linked to financial markets and customer behaviour.
- Optimize asset allocation according to the nature of the commitments.

The former models ran on deterministic or historical probability stochastic scenarios. Their purpose was to provide an overall long-term cash flow forecast within a going concern approach.

[9]The capitalization reserve, profit sharing reserve, and the liquidity risk provision, etc.

This purely decisional application scope does not require high precision cash flow forecasting, or a comprehensive depiction of the balance sheet.

These types of models can also serve to inspire Pillar 2 (ORSA) quantitative applications, but generally they do not have the necessary qualities to seriously address Best Estimate or capital requirement calculations.

The special features of the Solvency II compatible models are such that they are rarely interesting to develop from existing business applications, even if they work just fine in related fields.

The problem lies as much with the radical newness of the regulatory applications as with the resulting operating requirements for life insurance models.

From a purely calculation performance point of view, it seems better to start a new project rather than have to deal with inadequate technology.

6.3 Structural Choices

6.3.1 Define an Application Scope

The initial application scope choices present multiple consequences, as much in the model development phase as in its operational phase. Yet, the models' regulatory and decisional applications have diversified with Solvency II and this trend may not be over…

Although there are intermediate solutions, the application scope choice can be summarized by two diametrically opposed approaches:

Specialized Scope

One can deliberately confine the model to the regulatory applications of Solvency II's Pillar 1. The advantage is being able to reduce the size and complexity of the project; alas this solution is not a panacea.

Indeed, business intelligence applications, including ORSA, should be assigned to other tools that are, in principle, simplified but whose compatibility with the original model must be continuously monitored.

All the same, users should increase inputs, pit the different models against each other and "juggle" the numerous files. The system is still just as complex, but it has shifted from the software to the users.

Extensive Scope

Conversely, the idea of entrusting all regulatory and BI applications to a single model[10] seems a bit too ambitious and risky, but as a corollary it has a stage by stage development.

[10]In practice, rather than the same "model", it uses the same asset-liability simulation "engine" for all applications. Indeed, the internal model or Pillar 2 calculations are not necessarily made using the Monte Carlo approach that is used for Best Estimates.

The idea is to establish an initial base consisting of the deterministic forecast of the balance sheet and income statement, and then successively implant more advanced applications such as historical probability stochastic calculations, then risk-neutral probability calculations and ORSA, etc.

The difference with the previous approach is the need to think ahead to the constraints imposed by the latter stages, particularly in terms of time steps and performance.

Below is a non-exhaustive list of applications that can be assigned to one or more models.

Regulatory Applications

Pillar 1 Applications

As necessity (or sometimes a directive) is the mother of invention, an asset-liability model has now to evaluate commitments according to European Insurance and Occupational Pensions Authority (EIOPA) technical specifications.

Such a model should enable a Best Estimate life insurance evaluation before and after the shocks of the Solvency II Standard Formula. It will, therefore, also be used to calculate the Solvency Capital Requirement (SCR) under the standard formula.

Pillar 2 Applications

This is the purely quantitative part of ORSA with three main aspects:

- At all times measure the SCR coverage, however, in practice quarterly or even monthly.
- Forecast the solvency ratio over a period of 1–5 years along with a budget scenario and contrasting scenarios such as a panel of stress tests,[11] or even historical probability stochastic scenarios.
- Calculate risk limits and monitor their observance, knowing they will probably have to be recalculated several times a year due to the instability of the SCR under Solvency II.

Pillar 3 Applications

It should be noted straight off that it is pointless to expect an asset-liability model to fully handle the regulatory reporting responsibility.

Indeed, this reporting is made up of over 90 % straight accounting information, and preparing the regulatory statements is handled by specific software, the constraints and features of which are very different from that of an asset-liability model.

On the other hand, it is necessary to upload the results of the model into the regulatory reporting system at the level of analysis and within the time limits set by the EIOPA specifications.

[11]This application should also cover the stress test requirements of the European Central Bank and the Prudential Control Authority (APCR).

IFRS 4 Applications—Phase II
Although the base notions of international accounting standards are similar to Fair Value, they most certainly will include significant differences to the Solvency II assessment methods, especially for savings contracts.

This IASB draft standard is very controversial, hence it has an uncertain schedule…

Decision Support

Besides the Pillar 2 applications proposed above, the BI applications found here have already been presented to businesses, however they need to be adapted to the new regulatory framework. We can mention

- Traditional or market consistent Embedded Value calculation[12];
- Cost-benefit analysis, whether measuring technical and financial margins on existing contracts or forecasting these elements to define new products.
- Strategic asset allocation.
- Decision support for discretionary benefits and investments management according to insurance accounting standards.

6.3.2 Listing Features

The diversity of the previously considered applications enables measuring the scale of the developments necessary to arrive at a versatile asset-liability model, compliant with Solvency II technical specifications.

It is impossible to fully detail the functions of such a model within the context of this article. From the list of application domains contained in Sect. 6.3.1.2, however, some general conclusions can be drawn about the properties required of the model.

Treating the Regulatory Applications

Pillar 1
To make a Fair Value assessment of insurance liabilities, the model must represent the evolution of contracts over several decades, if not all the way to a contract's close.

In principle, only cash flow can be evaluated in risk-neutral probability[13]: future premiums, benefits paid, and contract management fees. However, accounting

[12] This remains very useful for indexing the incentive awards…

[13] Market consistent evaluation methods cannot be applied to accounting results, only to actual financial flows.

results according to insurance accounting standards will also be forecast, since they are the basis for the rights of policyholders to profit sharing.

Contract descriptions must at least reflect the concept of homogeneous groupings found in EIOPA technical specifications.

Moreover, commitment assessments must take account of the main causes of variability in cash flow, such as:

- Fluctuations in financial markets, unless there is no profit sharing clause.
- Insurer's investment policy, not only in terms of asset allocation, but also as regards available accounting options.[14]
- Discretionary benefits policy, which depends on the economy, achievable capital gains and use of the profit sharing reserve.
- Behaviour of customers in terms of exercising contractual options such as payments, early redemption, reductions, and extension, etc.

The above mentioned options can only be coherently assessed through a Monte Carlo approach in risk-neutral probability. Indeed, analytical methods are virtually powerless when dealing with combined real options that are not being exercised rationally.

To apply the Monte Carlo method, stochastic economic scenarios for several decades are required, calibrated in risk-neutral measure. These scenarios describe the development of key financial variables, not to mention inflation, because over time this significantly impacts contract management costs.

Conceptually, the scenario generator does not have to be integrated into the model. However, scenario samples must be recalibrated to account for rate shocks. Indeed, technical provisions must be subject to as many "Monte Carlo" assessments as there are shocks in the standard formula, either a dozen or more depending on the composition of the assets portfolio.[15]

Finally, ensuring the SCR calculation within the time limits prescribed by regulations is not enough. Throughout the year, the model should forecast the year-end SCR with adequate precision.

Indeed, it is well known that the new solvency ratio has a higher volatility than that of Solvency I.[16] Accordingly, General Management will not like waiting until January to find out the company's past Solvency Capital Requirements (SCR).

Pillar 2—Using a specialized model

Even with regulatory reporting, it is acknowledged that the quarterly Best Estimate and SCR assessments are simplified calculations. A second model can be created, simpler than the previous one, that is dedicated to estimating quarterly results.

[14]Capital gains or losses, allocations to the capitalization reserve, managing the liquidity risk provision, etc. These options are valuable to the insurer.

[15]Any simplification or alternative method to the Best Estimate after shock assessment added to the calculation of changes in net position of variances is difficult to master.

[16]At least if one disregards the transitional measures.

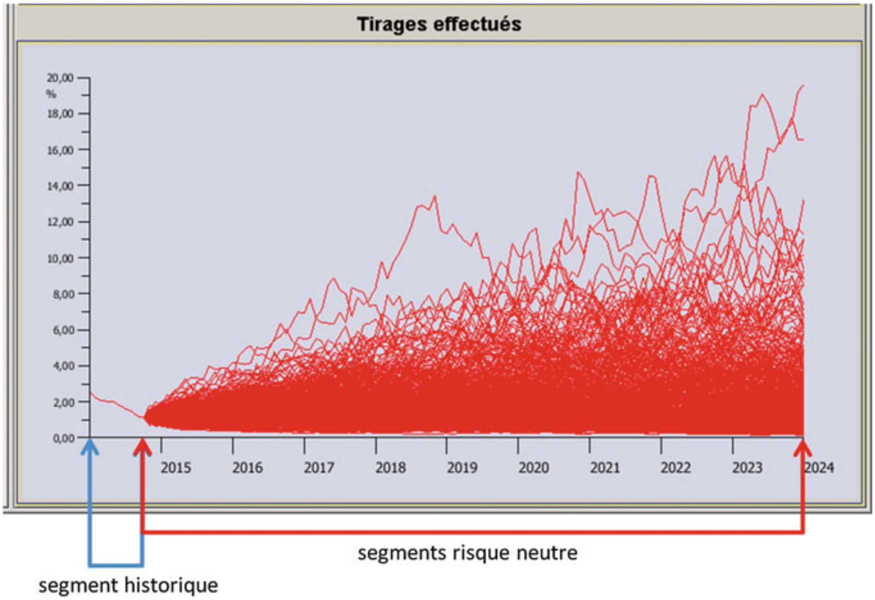

Fig. 6.1 Sample 10Y yield rate scenarios for a quarterly Best Estimate (BE) calculation

This second Pillar 2 model can also be used to forecast future SCRs by using the same simplifications to assess the impact of Standard Formula shocks on a simulated Balance Sheet.

Pillar 2—Using a single model
However, even if the Pillar 1 model does not operate quarterly or monthly, it can estimate the SCR sub-annually. The technique involves reproducing the past quarters using deterministic simulation, then switching to risk-neutral probability to calculate the Best Estimate and SCR.

Each of the scenarios the model simulates will be composed of two parts, a first "historic" going concern segment simulation, followed by a random segment in run-off corresponding to a "risk neutral" simulation performed by the scenario generator (Fig. 6.1).

By way of example, the above sample (the chart only shows 10-year rates) uses an historical segment from January 2014 to September 2014, then runs risk neutral probability stochastics calibrated to September 30, 2014.

Regarding forecasting future SCRs over several years, it can be achieved deterministically (budget scenario or stress test) with a Pillar 1 model. The technique is not too different from that described above: it is just a question of inserting a new "hypothetical" scenario segment into the simulations.

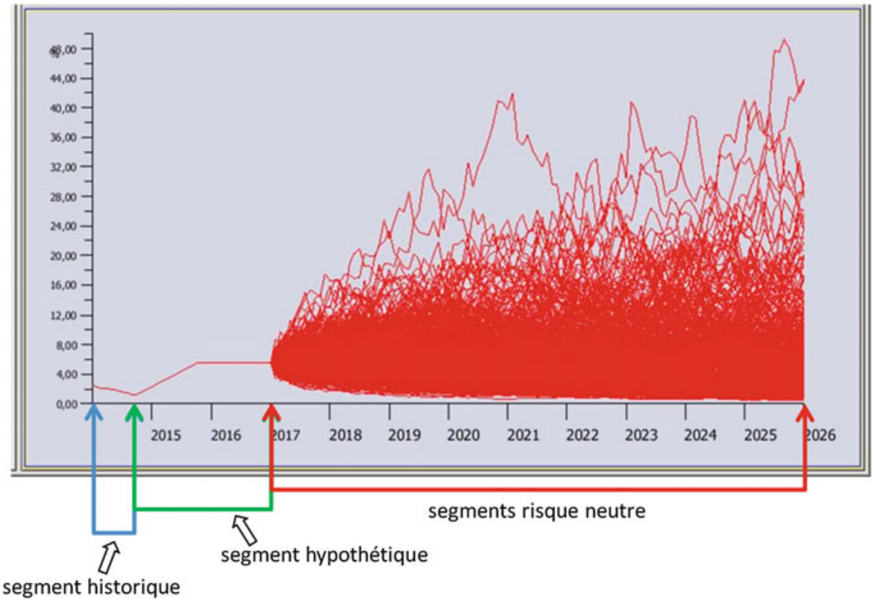

Fig. 6.2 Sample 10Y yield rate scenarios BE computation

This time the scenarios will be broken into three parts: an historical segment reproducing the previous quarters, then a hypothetical segment (budget or stress test), and finally a random segment drawn by the risk neutral probability scenario generator (Fig. 6.2).

The sample above uses an historical segment from January 2014 to September 2014, then a hypothetical segment (rising rate stress test), and finally runs a risk neutral probability stochastic calibrated to December 31, 2016.

Using a single model for Pillar 1 and Pillar 2 makes the calculations reliable and simplifies the user's task. In return, this choice imposes heavy performance constraints, due to the number of forecasted SCRs for ORSA, even limited to a few stress test scenarios.[17]

Moreover, in practice these performance constraints preclude a direct approach to the probability distribution of the projected solvency ratio, which would require replacing the hypothetical segments with historical probability stochastic scenarios (Nested Monte Carlo).

---

[17]Forecasting must be made over 1–5 years, not only with a central scenario, but also with a set of contrasting scenarios (a stress test panel).

However, the solvency ratio probability distribution can be approximated using different methods, e.g. Curve fitting or Least Square Monte Carlo, always retaining the "engine" of the Pillar 1 model simulation.

Pillar 3
As already mentioned, preparing regulatory reports should be done with specialized software. The model will be used to quarterly supply the input for the reporting application with the technical provisions and the capital requirements.

Including the model results in a regulatory processing chain requires a certain amount of traceability, which can take several forms:

- Secure archiving of the model versions, the parameters, and the assumptions used for the regulatory calculations;
- Secure archiving of the model's detailed results, which take up a substantial amount of memory in the case of stochastic simulations;
- The ability to reproduce the calculations upon demand—reducing storage requirements but increasing performance requirements.

Lastly, among the annual regulatory statements, a report dedicated to analyzing Best Estimate variations.[18] Compiling this report is particularly complex in life insurance.

To supply the input, the model will have to calculate several Best Estimates by varying certain simulation parameters, and should be able to forecast the BE out 1 year.

Treating the BI Applications

On a different register, BI applications are no less demanding in terms of features. Indeed, the Solvency II regime profoundly alters the analytical framework of the existing tools.

Embedded Value
Regarding embedded values, the problem involves measuring the cost of capital, especially through a market consistent approach.

The difficulty is similar to calculating Risk Margin: SCR forecasting techniques are required that are similar to those used in the ORSA.

Financial Decision Support
In terms of investment, insurers' tactical or strategic choices depend on a detailed analysis of profitability and risk[19] by asset class or even line by line, with a sub-annual step for genuinely operational applications.

---

[18]This is the VA-C2C statement entitled "Analysis of changes due to technical provisions".

[19]The risk in question is not limited to the volatility of market prices. It also includes issues of asset-liability matching.

Typically, French institutional investors analyze investments not only in terms of expected financial gain, but also based on forecasted book yields under the insurance accounting standards.

These factors are not new, yet Solvency II introduces a third criterion: financial returns net the SCR Market cost associated with the investment.

The calculation of an investment's contribution to the Solvency Capital Requirement (SCR) will in any event be set rather high in order to avoid any risk of exceeding the limits set by the company.

Within the context of controlling financial risks, it would be useful to assess the probability of exceeding an investment's allocated risk budget.

6.4 The Project Team

6.4.1 Professional Expertise

Designing an operational model requires bringing a variety of skills together, since all of the company's activities must be represented.

In this respect, it is not enough to just have a good "general understanding" of insurance. One needs in-depth experience of the mechanisms involved in the following areas:

- Asset-liability management (ALM)
- Actuarial (pricing, calculating provisions)
- Investment Management
- Reinsurance
- Local insurance regulations
- Solvency II technical specifications (all Pillars…)
- Financial mathematics and stochastic processes
- Taxation rules for Institutional investors
- IFRS and local accounting standards
- Financial statements Consolidation

These special skills are rarely exclusive to a single individual. A model's design must be the result of the work of a multidisciplinary team.

Entrusting the model's design to a specific operational department (technical, financial, or ALM …) runs the risk of focusing the model on some points while oversimplifying it on others.

For a large company, it is best not to allow a specific operational department to manage this project.[20]

[20]Nevertheless, the needs of the latter should not be neglected.

6.4.2 Separate Testing and Acceptation Process

It is not enough to design a model; it still needs to be turned into operational software that is coded with a programming language. Most actuaries and financial engineers are adept in one or more high-level computer languages[21] suited to mathematical applications, which they use every day.

These are spreadsheets and related macros, as well as languages more or less specialized such as Visual Basic, Matlab, R, and Gauss...

Alternatively, there is actuarial software available in the form of dedicated development environments and customizable modules.

These two kinds of tools can be used by non-IT professionals. Consequently, programming can be entrusted to business line specialists.

It should be noted forthwith that this organization must not infringe on basic security. This principle consists in assigning the testing and acceptation functions to people other than those in charge of software development.

The point here is not to discuss a division of formalized roles such as Project Manager versus Project Contractor, which is not necessarily suited to modelling projects, but simply to maintain independent controls.

Separating software programming from the role of control occurs naturally when code writing is assigned to outside consultants, yet this separation is also necessary when development is done in-house, even if it is not assigned to IT professionals.

6.4.3 Project Management

Managing a modelling project is a delicate matter, particularly due to the diversity of the business lines concerned (see Sect. 6.3):

- It is not so easy to get specialists, possibly speaking various languages, from different highly technical areas to sit down together on the same project.
- In addition, it is often necessary to bring in outside consultants to handle less common techniques such as implementing stochastic scenarios, for example.
- Finally, the development itself (writing the code) may involve IT professionals whose concerns are different from those of other team members.

In all cases, the responsibility for coordinating the work will be that of a qualified manager who is experienced, patient and persistent.

For certain sized companies, it is not a good idea to temporarily assign this project to one of the business line specialists. This is in fact a full-time job, at least until the model is up and running in its statutory dimension.

[21]This is not a value judgment. In IT language “high level” means “intuitive, concise and comparable to natural language” versus “low level” which refers to languages close to the binary codes used by the machines.

Moreover, this manager should be invested with hierarchical authority. Indeed, many points should be subject to trade-offs and decisions as soon as possible, otherwise the project could get bogged down.

Among the important choices to be made is defining a coherent IT approach aimed at the capacity and performance of the model to develop.

6.5 IT Development

6.5.1 Software Quality

As seen in Sect. 6.4, the code can be written by the business line specialists themselves, with or without the help of skilled IT professionals. However, code created with personal programming tools is not very reliable because this is not the purpose of that type of tool.

This flaw is made worse by the complexity of asset-liability models, which is incommensurate to projects usually handled by an actuary or a financial analyst on a personal workstation.

So it is normal to consider what an IT professional can bring to model construction, particularly in terms of reliability and performance. But even before tackling this issue, we must admit that experienced IT professionals can contribute something beyond technological expertise: they have an overall vision of the information system.

For an IT officer, software cannot be taken alone, but must be seen as a cog in a bigger wheel, a living element of the company's information system, itself in perpetual development.

Business line specialists may only be concerned with the expected features when assessing software quality. But the chief information officer (CIO) will emphasize the wider notion of "usability", which in practice covers a multitude of operational qualities:

- The **suitability** or relevance of the software is the degree of its functional scope, as well as a level of performance consistent with the **criticality** of the applications and volume to be processed.[22]
- **Ergonomics** measures the quality of man-machine interfaces. It ensures ease of use, greater productivity and makes learning easier for users.
- **Adaptability** is the software's ability to develop to support functional changes in scope (upgradability), growth of the operations (scalability), and technical environment changes (portability). It goes without saying that adaptability determines the software's sustainability.

[22]An application is critical if it lies on a decisive path of a processing chain that is vital to the company.

- **Usability** itself of course represents the ease of implementing the software into its technical environment. On a daily basis, this translates into many requirements such as reliability, availability, low crash rates, ease of resuming processing, the ability to communicate with other applications according to different protocols, maintainability, and the facility and rapidity to correct anomalies.

Finally, the many aspects of **security** must be taken into account, including the absence of viruses and malware for code security, as well as access, data, and processing security. By the same token, there are also the problems of **traceability** and the quality of backup and archiving procedures.

For many years, asset-liability models were used for decisional type studies that did not require the industrialization of processing.

Impact questionnaires and tests carried out in the context of Solvency II led to increasing the number of features, without introducing new constraints on reliability, performance and integration with the Information System.

This will change starting in 2016, because the Solvency II regulatory reporting schedule imposes an intense operating rhythm. Consequently, what is the use of having the most advanced model if it does not have the minimum qualifications necessary for regular operation?

Secondarily, the presence (or absence) of these qualities helps to determine the overall operating cost of the software, which tends to far exceed the initial development costs.

6.5.2 Computer Engineering

Languages and Development Environments

Spreadsheets and high-level computer languages used in a personal capacity have their place in a company, particularly as tools for research and prototyping complex applications.

Yet these languages are not designed to develop software embedded in an information system, still less in a processing chain whose purpose is to produce accounting and regulatory results.

Even from a purely business point of view, professional computer languages, such as objective C, C++, C#, Java, and Delphi to name a few, are preferable for the following reasons:

- Professional languages provide processing and performance capacities far greater than personal tools.
- They also offer very precise management of material resources (memory and processors) essential to optimizing performance.

- They are better suited to large projects when handled by an integrated development environment professional (IDE).[23] A modern IDE allows several developers to work collaboratively and merge their work on a continual basis.
- IDEs provide powerful tools for controlling, testing and optimizing that significantly improve the quality of the code.

Procedural Programming Versus Object Paradigm

Classical programming is referred to as "procedural". This involves creating a collection of procedures or independent functions, acting in a certain order on the initial data that are themselves reported separately one from another. Each procedure uses the results of the previous procedures, transforming them into new intermediate outcomes, to be reused by the next procedure through to the final result.

In an asset-liability model, the procedures are numerous and quite varied. Furthermore, the settings and scheduling of these procedures can be dynamically changed by the interim results, which make this type of program extremely complex.

But procedural programming acts like a chain. Any malfunction to the ordering or the results of a single link causes the entire process to fail. This failure is particularly egregious when it comes to correcting or updating a procedure, because any manipulation can reverberate unexpectedly down the chain of procedures.

Object oriented programming for its part no longer works on data, but on "objects". An object is a record that brings together and encapsulates in a single block both the data and the set of procedures or methods for the treatment of these data.

An object thereby protects the integrity of the data it contains, because they are only accessible by the methods listed and defined in terms of an object type or class.

When a programmer processes an object such as a mixed life insurance policy, he uses a method encapsulated in the object itself (e.g. "Get_Buyback_Value") for which he does not need to know the details.

The same method will work differently, calling on various data if the object is a unit savings account. But the programmer does not have to consider this difference as long as he does not have to change the method itself.

What should be remembered from this summary presentation is that object-oriented programming enables a very modular approach to program development, greatly facilitating software maintenance and development due to the relative independence of objects.

NB: certain languages—SmallTalk, Objective C, impose object-oriented programming, while others—C, C++, enable this method as well as the procedural method.[24]

---

[23]IDE: integrated development environment. This is a suite of software used for computer development. The main IDEs currently being used are Eclipse, MS Visual Studio, and NetBeans.

[24]And the mixture of both, which is generally not recommended.

6.5.3 Development Method

Waterfall Method

Although up to now we have said only good things about IT professionals, we regrettably must say that their usual approach to "industrialized" development is susceptible to derailing even the best designed modelling projects.

Corporate IT managers often program using the waterfall method, as they often do not know any other. Inspired by the planning techniques used by major construction sites (!), the waterfall method is a sequential approach to development that can be summed up as follows:

Review of Requirements → Functional Specifications → Detailed Analysis → Programming → Tests → Reception.

The analysis phase in this approach seeks very fine detail in order to provide developers with precise instructions that are so specific they only have to mechanically carry out the programming, theoretically without taking any initiative. Indeed, the aim is also to reduce costs by outsourcing programming, which can then be relocated.

The weaknesses of waterfall development are well known:

- The inventory of needs is rarely comprehensive. Nevertheless, once completed, this inventory becomes a contractual framework from which there is no question of working outside of.
- Detailed analysis is tedious and interminable on large projects. The resulting document is illegible to users who are nevertheless responsible for validating it.
- Developers are deprived of all sense of responsibility and unable to detect errors or inconsistencies in the detailed analysis.
- The development cycle is a rigid sequential procedure, where any error inevitably spreads throughout the whole system.
- Non-blocking design errors are only belatedly detected by users during the test phase. This "tunnel vision effect" could be disastrous if the analysis is called into question.

Of course, IT project managers have learned to minimize these defects by closely monitoring the various stages, by multiplying intermediate tests and by involving users in the development process.

Unfortunately, what is wrong with the waterfall development method is compounded by the characteristics of modelling projects.

First of all, the inventory of needs is particularly difficult to put together due to the lack of experience of the users themselves. The actual needs will not be known until after a few years into the Solvency II regime. In fact, level 3 measures (implementation measures and standard techniques) are by no means stabilized, while in some areas regulation remains in a soft focus allowing several interpretations.[25]

In terms of market consistent evaluations the state of the art has by no means been attained. Translating needs in terms of features is open for debate: for example, some simplifications are commonly used to lighten calculations such as flexing methods, using proxies or replicating portfolios for liabilities, and sensitivity-based modelling. What is the application scope and to what extent are they statutorily acceptable?

Performance requirements are poorly evaluated; they depend on, among other things, the operating rhythm of decision support tools, and the acceptability of the simplification methods discussed above.

Drafting a detailed analysis document under these conditions is challenging. Especially since the area of financial modelling is unfamiliar to business consultants and corporate IT professionals, and has a particularly steep learning curve for asset-liability insurance models.

The programming, testing and correction phases are long. It could only suffer from an organizational structure in which the developers are disconnected from the line of business and considered to be interchangeable, irrespective of their skills in computer engineering.

Finally, the project's duration is often a major obstacle for the waterfall development method because between the needs inventory phase and reception, the regulatory environment and the company's business may change to the point where the software has been rendered obsolete before even being commissioned.

Agile Methods

Formalized since the early 90s,[26] Agile development methods involve primarily trying to meet evolving customer needs through an organic development process that is not systematically planned.

The development cycle is permanent and parallel, several functions beeing covered simultaneously at different stages of development (design ↔ development ↔ testing ↔ user feedback).

The idea is to create software that can regularly add new functions, or enhance existing ones, while remaining operational.

[25]For example, replacing the usual "risk tolerance" vocabulary with "risk appetite" is not enough to accurately define a new concept and does not at all explain how to operationally translate it into risk limits.

[26]There are variants of the Agile methods; the main ones are known by the following names: RAD, DSDM, SCRUM, XP, ASD, and Crystal Clear.

In order to always have an operational application, it is common to implement a new feature in its simplest form, and the users must accept this. This implies advancing the functions step by step[27] until users are satisfied with the level of adequacy.

In turn, developers must accept that functionality is never definitively "delivered"; it remains subject to change or even to being completely reformed at any time depending on feedback tests or environmental changes. If it exists, detailed analysis only plays a secondary role, whereas satisfying the needs of users remains the acid test.

It is quite clear that with this approach the concept of "project" itself could disappear, as it becomes divided into multiple sub-projects that gradually combine with the software's scalable maintenance.

In practice, the programming team is in constant contact with the users or with the business line specialist in the case of modelling projects. The biggest problem with this type of development is maintaining the best possible communication with the developers. Everyone must be a team player with a good understanding of the business issues and the technical constraints.

Thus the Agile method turns its back on industrialization development techniques. Developers are important as they maintain constant contact with the designers and intervene at the function discussion stage as well as the programming stage to propose solutions or explain technological limits.

A team of developers dedicated to the project must be considered, who are accustomed to working with business line specialists and thoroughly understand the basic concepts and vocabulary of insurance.

Information system managers are often uncomfortable with this organizational model, which permanently uses part of its resources. "Assigning" computer technicians to a project whose contours and timeframe are poorly defined also poses very real personnel management and career orientation problems.

At any rate, there is no other way to ensure proper communication between the model developers and the IT professionals.

6.5.4 Model Documentation

Here we are only talking about the model's technical documentation and its applications, outside the framework of the highly restrictive internal models.

The model's documentation is part of a larger whole that describes the overall risk control system—management organization, principles, methodologies, systems and risk management procedures, sources and data quality, and automated and manual processes.

[27] This is referred to as "iterative and incremental" development.

General Documentation of the Model

The model should be the subject of a regularly updated summary presentation providing the following:

- A schematic representation of the model's overall organization, input data, information channels and computer architecture;
- A list of applications and a description of its key features; and
- A chronology of the successive versions and changes to its methodology.

The correct balance must be reached between an exhaustive, superfluous, and illegible amount of documentation, and a summary presentation containing a few general statements. In practice, documentation may be made up of several types of documents corresponding to various levels of detail and targeted uses.

The detailed functional analysis mentioned in Sect. 6.5.3.1. cannot be considered documentation, although its volume is quite impressive. Indeed, it is not structured in a way to foster a better understanding of the modelling principles, and updates are not integrated—at most they are annexed in chronological order.

It would be a good idea to accompany the model's documentation with information for the executive management, in accordance with the governance principles set forth in Article 44 of the Solvency II Directive, *even if the company does not intend to seek approval for the internal model*.

Documentation of Procedures

Procedures contributing directly to the model's operation should be described in relatively high detail to serve as an operational guide for users or internal control:

- Input and control procedures of the initial data;
- Back-testing procedures and monitoring results;
- Backup and archiving procedures;
- Test procedures and receiving new developments.

Application and Method Documentation

This involves documenting recurrent applications, both regulatory and decisional. For example: calculating the Overall Management Provision, calculating the Best Estimate and Risk Margin, quarterly stress testing, budget forecasts, embedded values and so forth.

It is advisable to describe the principles, settings and calculation methods specific to each application—scenario characteristics, discount methods, specific assumptions, and risk indicators, etc.

Here again, the purpose of this document is to serve as a user reference or even for training support. A real effort must be dedicated to presenting these documents.

Model Logbook

The model's development over time can be recorded in a log having the following sections.

- Developments: description of new releases and features, method changes, feedback and an evaluation of the impact of the developments;
- Bugs and operational issues: Chronological list and description;
- Back-testing: summary of the results and feedback;
- Annual report: test results, summary of audit findings and recap of areas for improvement.

6.6 Conclusions

Insurance asset-liability models are the most complex corporate financial models due to the nature and diversity of life insurance contracts, the partial pooling of financial products, but also because the risk approach in the Solvency II regime imposes a comprehensive simulation and an exhaustive representation of the balance sheet.

It would be easy to underestimate the resources required to create a Solvency II compliant model. Before launching a development project, sufficient resources in terms of time and qualified personnel must be considered.

Furthermore, the budget requirements are far from negligible. Development costs are not the only expense to forecast, the ongoing maintenance of this type of software is an important part of overall operational costs.

Insofar as the model must reproduce virtually all of the company's operations, its design must be entrusted to a multidisciplinary team of business line specialists in fields as varied as actuarial, financial management and accounting consolidation.

It is possible to build a model without resorting to IT professionals, by working with certain high-level computer languages used by financiers and actuaries, or by having recourse to specialized development platforms.

However, if we are to exceed the performance limitations of the tools listed above, and ensure an adequate level of reliability for the regulatory and accounting applications, the use of professional programming languages and working with IT engineers becomes indispensable.

But the waterfall model development cycle commonly used in IT management is unsuited to financial models, which paradoxically increases the risk of project failure where IT and organizational consultants operate according to this classical scheme.

It would be better to turn towards adaptive development methods using techniques generally known as Agile methods. However, this solution means employing a team of dedicated IT professionals having a certain experience with the business lines.

Part III
Model Validation and Steering Processes

Chapter 7
Ex-ante Model Validation and Back-Testing

Stéphane Loisel and Kati Nisipasu

Abstract As the famous statistician George Box said, "All models are wrong, but some of them are useful". The goal of the validation process is not to say whether a model is right or wrong. Besides, model elaboration and validation must be regarded as an ongoing process, whose goal is to continuously improve and update the model. In this chapter, we shall focus on several aspects: the formal a priori validation process, back-testing issues and model risk assessment.

7.1 A Priori Formal Validation

There are many different definitions of formal internal model validation. We focus on the case of the validation of the internal model of a large insurance or reinsurance group in the context of Solvency II. In this framework, the validation of an internal model aims at checking the reasonableness and accuracy of the internal modelling and the results thereof, on an independent and proportionate basis as performed by a competent professional. It does not aim to provide parallel modelling or results but it should support the constant development and enhancement of the model and should be organized with enough anticipation in order not to threaten the timely production of results. It is performed by the company and is a pre-requisite for the Solvency II application to internal model use approval (see Article 124 of the Solvency II directive). Note that we do not refer here to the process by which the regulator decides or not to allow the company use her internal model.

The first step of the internal model validation process is to define the scope of the validation. In the case of a group, it is necessary to validate not only the group internal model but also the entity level internal models for each entity in the

S. Loisel (✉)
Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France
e-mail: stephane.loisel@univ-lyon1.fr

K. Nisipasu
SCOR, Paris, France

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_7

Solvency II application zone. Depending on the size of the company, the set of countries where it operates and the coverage of the internal model (which may be partial or total), some entities or some risks may be descoped because they are treated using the Solvency II standard formula or because they are outside of the scope of Solvency II: if, for example, an entity is based in America, then the risks associated with this entity must be included in the group internal model and the group approach to these risks must be validated, but the local entity (solo) internal model at the American entity level does not require validation in the Solvency II process.

Even if some motivations come from classical Enterprise Risk Management ideas, most deadlines of the validation process are imposed by the Solvency II calendar. The validation process must be adapted to this calendar. In particular, one must carefully set the validation objectives at the right level. Validation issues that cannot reasonably be addressed before a deadline must be documented and studied during the next evaluation wave.

In terms of process, the validation must be carried out by a team that is independent from the modellers. This does not mean that modellers cannot take part in the validation process. For example, a person who has only been involved in one particular risk module could perform low level verifications for other, independent modules. However, it is important to have a dedicated team, led by someone from the company who knows the business model of the company very well. This person must have project management skills (because one must respect many deadlines and the validation requires the coordination of several experts from various departments), as well as the ability to interact with modellers, board or executive committee members and regulators. The validation team may be composed of individuals working in the entity and of external experts or consultants, to ensure that one has enough in-house proficiency in the subject and benefits as well from the inputs of external experts.

Let us come back to the notion of independence of the validation team. It seems obvious at first sight to require complete independence between the validation team and the operational teams that have built and run the internal model. The validation team may be composed of external consultants and of members of internal validation teams. The company may choose to allocate several persons to the sole internal model validation function. Often, some part-time validators are also selected from other teams of the company that have not participated in the corresponding part of the internal model.

One might think that it is easy to respect this independence principle to the letter. However, in practice, some issues immediately arise. For example, consider an independent validator who recommends some model changes during year N. Can this validator be regarded as independent again in year N + 1 for the next validation? This of course depends on his/her degree of implication in the model change and on the materiality of the corresponding model change. Consider a validation expert who just joined the validation team from another entity which took part in the modelling in his previous position. Can this member be considered as independent immediately? Should this person mainly validate areas of the internal model close or far from his/her field of expertise (to favour, respectively, expertise or independence)? A very strict notion of independence could be counter-productive both in terms of expertise and in terms of cost. Of course, being too loose on the independence control would completely threaten the objectivity of the validation.

Therefore, the independence of the validation must strike a balance between any potential conflict of interest that might arise in the course of the validation of the internal model on the one hand and a disproportionate level of segregation of duties on the other hand. Rules must be stated in advance in a policy, written carefully with a prospective view (in particular, to avoid that almost nobody can be considered as independent after the first validation!).

It would take too long to check every small detail of the model. Any internal model validation process has to rely on some proportionality rule, to focus on the most material issues. It is in practice very difficult to choose the materiality thresholds and to justify that they are the most appropriate ones. One possible compromise consists in combining qualitative and quantitative arguments. Quantitative arguments generally consist in materiality tests: if a change in the model causes a change of more than x% in the final output (for example, the group Solvency Capital Requirement, for some other relevant computations or metrics), then the materiality is assessed to be high. It is possible to define different levels of materiality and to use these levels in a progressive way. When the quantitative approach is impossible or brings limited information, qualitative arguments may be used as a complement. It is very important to document these qualitative assessments with the same level of detail as the quantitative ones in the validation report.

In practice, two main difficulties may arise regarding proportionality principles. In the case of a group, one often faces the issue that some risks and some changes are material for some entities and not for the group (due to diversification), and vice versa. Besides, some tests considered as non-material initially or in the last review could become material in the new environment due to nonlinearities in a complex, high-dimensional internal model. If time is limited, it could happen that one chooses to focus on tests for variables that are known to be material, and ignore others that are inappropriately considered as non-material.

Most extreme risks faced by insurance and reinsurance companies cannot be studied (at the 99.5 % risk level) with statistical accuracy, just because of the lack of data. We do not have 10,000 data points for nuclear terror attacks, pandemics or for some natural catastrophe risk, and this is fortunate! Therefore, one must rely on expert judgment. Contributions from terrorist risk experts for example, extreme value theory specialists and black box software cannot be evaluated by statistical tests. First, they must be challenged by alternative approaches. Secondly, evaluating the governance of expert judgment process is the most important part of the validation of their conclusions. One must check that the experts understood the question, that their answer has been correctly interpreted, and that their arguments make sense and were not influenced by the way they were contacted or chosen. This is crucial because otherwise many people could say that the model validation is useless because these expert judgment decisions contribute to the main part of the results.

Besides, the validation team must always be concerned with the use that operational experts and decision-makers can make of the model. A good risk measure ceases to be a good risk measure as soon as it becomes a target. Consequently, one must think about the potential consequences of people gaming the model.

A model cannot represent the real world. Consequently, coming back to George Box's quote, a model is always wrong as we necessarily make mistakes while simplifying the problem. One must always ensure that costs of mistakes are smaller than benefits from simplifications. In a validation process, one must in particular check that the approach is consistent with the (arbitrary) framework that has been chosen at the beginning (for example, regarding financial or economic valuation principles). The limitations of the model and the impact of the choice of the initial framework must be made clear to all stakeholders. This can be done in the easiest cases by showing the impact of the change of some assumptions regarding correlation models or risk distributions, for example. At the higher level, one must think about initial choices that may have a significant impact on all components of the model, for example the economic scenario generator.

The validation team must also test the technical exactitude and the accuracy of the results, under the assumptions that have been made. This can be achieved by a battery of tests at the local level, by independent computations, and by different tests of the order of magnitude of the results at the overall level.

Most validation processes have several steps: the components of the model are reviewed independently from one another, but the full structure and outputs of the model are also reviewed. It is important to evaluate the impact of some parameters or assumptions on the final outputs of the model. This is related to the concept of robustness, which is related to the abovementioned materiality levels.

To be more precise, a wide range of tools, commensurate to the specific question to be analyzed, may be used. Back-testing is particularly appropriate when one has enough data to perform it. It is described in the next section. Sensitivity testing corresponds to varying some parameters, some distributions or some correlation approaches in the internal model and checking the impact on the final outputs. Stress and scenario testing may be carried out: what would happen if volatility increased by 20 %, if stock markets went down by 10 %, and so on… The difficulty of this approach is that it should be based on a kind of conditional distribution of the other risk drivers given what is considered in the scenario. Very often, the distribution of the other risk drivers in the initial risk environment is conserved and is not replaced with the conditional distribution (which may also be irrelevant, as it is better to condition by events of the type "this variable level is worse than x" than by events of the type "this variable is exactly at level y").

It is also possible to use a comparison, for example with benchmarks and theoretical portfolios, to perform some qualitative tests and qualitative reviews, or to carry out stability testing. Eventually, some reverse stress tests may indicate which events or which type of model changes could lead to some highly material change in the output.

The model should in practice be validated at least every year for the annual Solvency Capital Requirement calculation, or whenever there is a model change with impact on the Solvency Capital Requirement or a significant change in the risk profile of the company. Tests and results from previous validations can be reused whenever applicable. After the first full review, it is possible to organize reviews based on the changes that have been made or need to be made, due to the evolution of the economic, legal and regulatory environment, and following the recommendations that have been made during the initial full validation.

After 3–5 incremental reviews, it is necessary to organize a new full, in-depth review that looks very much like the initial review, apart from the fact that it is made easier by the expertise acquired by the modelling and validation teams. Note that it is of course necessary to provide assurance that the incremental reviews are sufficient and that no significant matter is missed with this approach and requires a new full review. The regulator will logically require to justify that the quality of the incremental reviews is satisfactory and meets the level of the Solvency II internal model validation standards.

The validation of the Internal Model during each operational run serves to gain confidence over the processes within the Internal Model and the results it provides. On an on-going basis, when no change has occurred in the underlying mathematical assumptions and when no significant event has occurred, the aspects of the internal model that shall be reviewed cover at least the data quality, calibrations, expert judgments and adequacy to risk profile.

An ad hoc check of the model validity may be conducted. When a significant event has occurred since the previous run of the model (e.g. an exceptional loss, a significant change in environment, an M&A transaction, a significant portfolio change).

When a new feature or enhancement of the model occurs, the validation of the overall new model version must then be made and must cover all the validation steps.

A the end of the validation exercise, a follow-up process needs to be implemented to monitor the actions implemented following the recommendations of the validation and provide a status of these recommendations during the subsequent validations.

In terms of organization, roles and responsibilities, internal model validation looks like risk management, which features several lines of defence. Likewise, one first finds the teams in charge of the production and the operational runs of the model, followed by the validation team. Depending on the organization of the risk management process in the company, the conclusions of the validation team may be reported to a kind of risk committee or to the executive committee, before being reviewed by the board or by some risk committee of the board. Of course, there is a lot of diversity in terms of organization at the top management level, due to the large variety of insurance entities and groups subject to Solvency II.

Internal model documentation represents a very large number of pages, and so does model validation documentation. Consequently, one must find a concrete way to help the executive committee and the board to understand, evaluate and react to the main conclusions of the internal model validation. One solution developed for example at SCOR consists in the creation of a 4-person committee, composed of the head of the validation process, the former Chief Risk Officer and 2 academics.

This external expert committee evaluates the validation process and provides guidance regarding the main actions to undertake, to improve the quality of the validation, reported to the executive management in a restricted number of pages annexed to the validation report. Validating a model requires knowledge of the business model of the company: some model might be adapted for one company but not for another one which features some risk accumulation in one risk segment, for example. Size and market share do matter, and this is too often not sufficiently taken into account by classical risk measurement approaches.

It is therefore very interesting to include in this kind of external committee someone who knows both the history of the company and also the risk culture in the company, which may decrease the risk of underestimating or of bypassing some newly important risk due to overconfidence. In the same spirit, it is also beneficial to choose the academic members in different countries, with different backgrounds and specialties, to avoid too frequent, inappropriate consensus due to too similar training and culture.

7.2 Ex-post Validation: Back-Testing

To start with, let us immediately state an obvious but very important point: it is possible to back-test some computations based on mathematical expectations for frequent data, but it is impossible to back-test 1 year, 99.5 % VaR computations.

In finance, the Basel regulation incorporated back-testing ideas in the late 90 s: banks who had too many exceptions received penalties (red and yellow) that led to the increase of their capital requirement multiplier. For the daily 99 % VaR, one observes on average 2 exceptions per year (based on 250 trading days per year), and the back-testing already suffers in that case from the low power of the test. If one miscomputes the 99 % VaR and uses the 97 % VaR instead, then the probability of exiting the confidence interval of the number of exceptions (for the 99 % VaR) only increases from around 10 % (for the 99 % VaR) to 12 % (for the 97 % VaR). Consequently, according to Bayes' rule, observing many exceptions does not mean much regarding the accuracy of the 99 % VaR computation.

The following picture (source: http://www.value-at-risk.net/backtesting-value-at-risk/) shows the daily profit and loss of a financial portfolio and the corresponding 95 %-VaR computed after JP Morgan's RiskMetrics methodology (Fig. 7.1).

We observe 10 exceedances out of approximately 125 trading days (6 months), which is more than the expected number of occurrences in principle ($125 \times 5\ \% \approx 6$). However, it would be hard to distinguish between Type I and Type II errors in this practical example.

In insurance, the 99.5 % VaR at the yearly scale is of course impossible to back-test during the current human lifetime because it corresponds to failure every

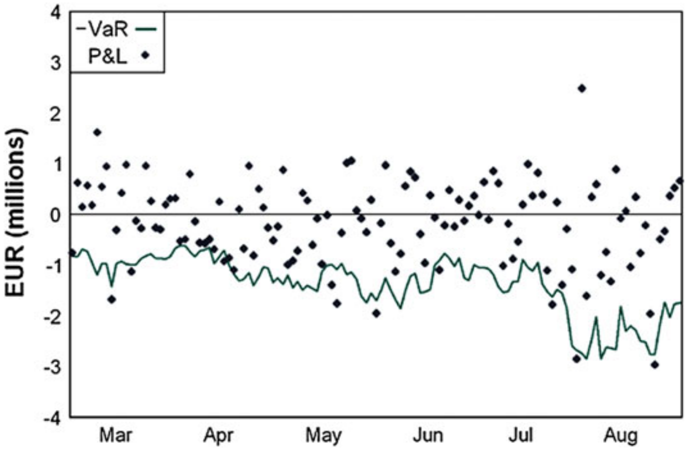

Fig. 7.1 Profit and Loss and daily 95 %-VaR of a financial portfolio over a six-month period

200 years on average. Even if we lived for 20,000 years, back-testing would still have a limited efficiency because of changes of regime and non-stationarities. One could argue that we observe hundreds or thousands of companies every year, each of them having a probability of ruin that should be less than 0.5 %. However, it is important to remember that the outcomes of those companies are far from being independent as they are strongly influenced by common factors (economic, legal, natural risk drivers).

In the absence of enough frequent data to proceed with rigorous back-testing, top managers may choose to perform some informal, non-rigorous back-testing according to their own expectations. If their risk culture is mainly influenced by Gaussian-like risk environments, managers may face the following pitfall: the abovementioned lack of independence may completely change the properties of the best estimate. Usually, the average is thought to be in the heart of the distribution: in the Gaussian setting, after aggregating many small events or the decisions of thousands of independent agents, the central limit theorem guarantees that the result is very likely to be in a confidence interval around the average. In the example illustrated in Fig. 7.2, the surrender rate in a stressed scenario is assumed to be on average 40 %. When policyholders make their decisions independently from one another, the bell-shaped black density curve tells us that with a high level of confidence the surrender rate should be between 35 and 45 %. However, when decisions of policyholders are influenced by a common guru, which breeds correlation between policyholders' decisions, one switches from the black, unimodal density to the red, bi-modal density curve. In this bi-modal world, 40 % remains the average, but there is a very high probability that the final outcome is either between 32 and 36 %, or between 47 and 52 %. It is in fact very unlikely that the surrender rate will stay between 36 and 44 %… In such a situation, the average remains a weighted sum of the potential outcomes, but loses its property of being close to the

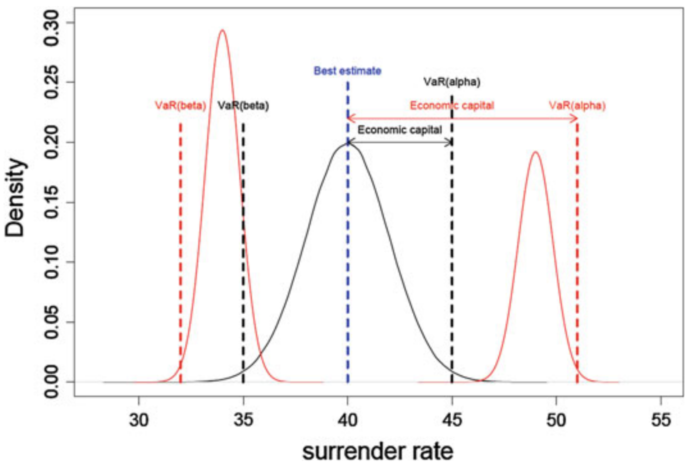

Fig. 7.2 Unimodal density (*black*) and bi-modal density (*red*) of surrender rate distribution in a stressed scenario

mode of the density of the distribution. In a more spread situation, this phenomenon could lead top managers to over-react if they do not take into account that they were in a 2-scenario configuration.

It is, however, possible to back-test the best estimate, and to sequentially test for changes in the actuarial assumptions or parameters, if we have enough events per year. We now describe one particular back-testing strategy that may be used to sequentially test changes in some actuarial assumptions and in some outputs of the model, like the best estimate. Sequential testing is a whole research field, dominated by Russian, Greek, American and Chinese probability and statistics schools. For a detailed survey of the classical results, we refer to the following three books:

- Tartakovsky, Alexander, Igor Nikiforov, and Michèle Basseville. *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, 2014.
- Wald, Abraham. *Sequential analysis*. Courier Corporation, 1973.
- Poor, H. Vincent, and Olympia Hadjiliadis. *Quickest detection*. Vol. 40. Cambridge: Cambridge University Press, 2009.

One may distinguish two different settings: Bayesian quickest detection and non-Bayesian (also called robust) detection. In the first case, one has some prior information or strong belief about the distribution of the time at which the change may occur. Very often at the beginning, in both cases, one assumes that the pre-change parameter and the post-change parameters are known. One of the main focuses of this field now concerns multi-sensor detection problems: how to combine the online information by different sensors in California to build an optimal earthquake detection system? These techniques could be adapted in the near future to Enterprise Risk Management in insurance and finance, where one faces the similar problem of combining different correlated Key Risk Indicators and weak

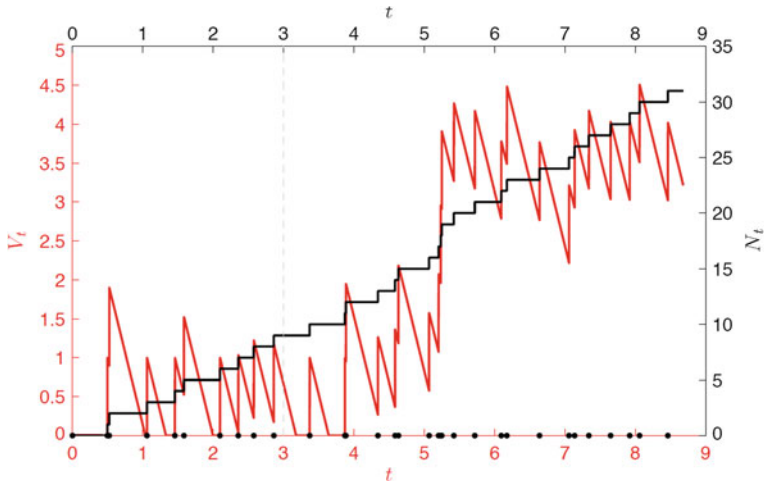

Fig. 7.3 Simulated cusum process *V* (*red*) with a change-point at time 3 and constant intensity 3 (before the change). The post-change intensity is set equal to 4.5. The *left axis* of coordinates is associated with the cusum process and the *right axis* is related to non-decreasing process *N* (*black*). The dots in the abscissa indicate the jump epochs

signals in order to react fast enough to changes of regime without too many false alarms.

To make it more concrete, we illustrate these back-testing tools with one particular example of use of the so-called cumulated sums (Cusum) process, with two figures drawn from El Karoui, Loisel and Salhi (2015).

In Fig. 7.3, we plot the evolution of the number of events up to time t (in black). We want to detect as quickly as possible a change in the Poisson intensity from 3 to 4.5. Without any information on the change time distribution, if one uses Lorden (1971)-type criterion, El Karoui et al. (2015) show that the quickest detection strategy is given by sounding an alarm when the (red) associated cusum process V upcrosses a certain horizontal level m. This level m should be chosen to respect a false alarm constraint: one does not want to sound a false alarm more often than once every p years. The more stringent the constraint is on p (the larger the value of p), the larger the corresponding value of m must be.

In non-mathematical terms, the Lorden criterion corresponds to minimizing the worst potential average delay between the change time and the detection time. This explains why the cusum process is reflected in zero. If this was not done, then the non-reflected version of the cusum process, which corresponds to a kind of log-likelihood ratio between the two probability measures (pre- and post-change), would on average drift towards—infinity: if the change occurred after a long amount of time, the non-reflected process would be very far below zero, and it would take a very long time before it upcrosses level m. The monitoring strategy is in fact quite simple to implement for an insurance company: the cusum process decreases at a slope that depends on the different parameters and is reflected at zero.

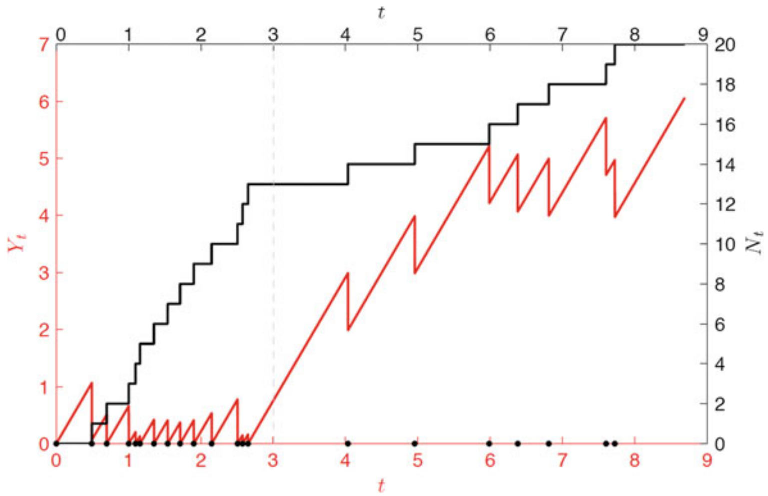

Fig. 7.4 Simulated cusum process *Y* (*red*) with a change-point at time 3 and constant intensity 3 (before the change). The post-change intensity is set equal to 1.5. The *left axis* of coordinates is associated with the cusum process and the *right axis* is related to non-decreasing process *N* (*black*). The *dots* in the abscissa indicate the jump epochs

Each time there is a claim (or a monitored event), the cusum process jumps upwards with a unit jump size.

The cusum process is also optimal in the case where one wants to detect a reduction in the claim intensity. Figure 7.4 shows the claim number process *N* (black) and the associated cusum process *Y* (red). This time, the cusum process has an upward drift and downward jumps. It is also reflected in zero. In contrast to the previous case, the cusum process *Y* upcrosses level *m* continuously, which makes the associated optimality proof easier to obtain than the one for process *V*, which upcrosses level *m* with a jump.

Ideally, this back-testing methodology should be combined with different sets of information coming from inside and outside the company. In practice, the compilation of the different Key Risk Indicators, weak signals and quantitative detection mechanisms has to be done by the management team and reviewed by the board, with the influence of the risk attitudes and of the culture of the different members of those groups (see the chapter on Risk attitudes and the following paper: **The Social Construction of Strategic Surprises: Learning from the CIA, 1947–2001,** Philippe Silberzahn, *EM LYON Business School,* Milo Jones, *IE Business School,* Paper presented at the Academy of Management, Boston, MA, USA, August 3–7, 2012).

Chapter 8
The Threat of Model Risk for Insurance Companies

Christian-Yann Robert

Abstract Insurance companies have increasingly used quantitative decision-making tools for a number of years. They have routinely taken advantage of models for a large number of business activities (underwriting policies, transferring risks, determining reserve adequacy, managing assets and liabilities, valuing risk exposures and financial instruments,..), but they now appeal to models for more ambitious ends like development of new products or strategic planning. Moreover, the new quantitative regulatory requirements of Solvency II, as well as various stakeholders' expectations (including rating agencies, analysts, financial markets,...) push companies to develop more and more sophisticated models, not only for more complex products, but also for improved enterprise risk management. The expanding use of models reflects the range to which models can improve business decisions. But they also can lead to wrong decisions and potential adverse consequences when they are incorrect or misused, which is known as model risk. An active management that addresses these consequences has to be organised by insurance companies. In this chapter, we first try to understand what model risk is. In a second part, we discuss several approaches to measure model risk. In a third part, we raise the problem of model risk management and present several procedures to mitigate it. In a last part, we discuss the issue of model risk for the new regulatory framework Solvency II.

8.1 Introduction

Banks heavily rely on models in most aspects of decision-making for a lot of activities, including underwriting credits, measuring risks, valuing exposures and positions, calculating overall capital requirements, and allocating capital for business decisions,…. In addition, the valuation and pricing of complex financial products are only carried out with sophisticated mathematical models because

C.-Y. Robert (✉)
Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France
e-mail: christian.robert@univ-lyon1.fr

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_8

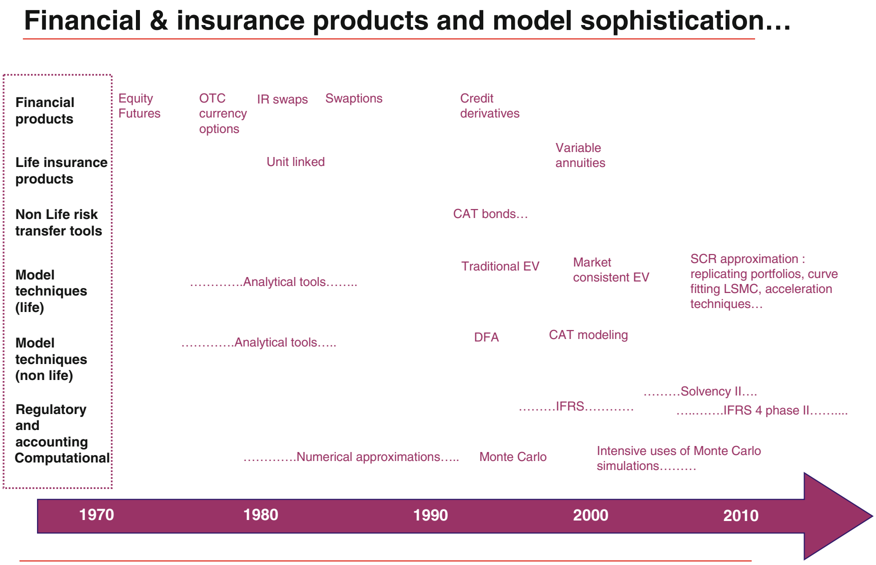

Fig. 8.1 Financial and insurance products and model sophistication (Grigorov 2012)

market prices for these products are unavailable. In recent years, it has often been heard that errors in models and ineffective model risk management were the main reasons for the recent financial crisis. In particular, during the credit crunch, attacks concerning mathematical and quantitative models for credit risks arose, and models have begun to be accused of being too simple and to give wrong prices for complex credit derivatives, such as mortgage-related derivatives.

Insurance companies have also increased the use of quantitative decision-making tools for a number of years. They routinely use models for a lot of business activities, including underwriting policies, transferring risks, determining reserve adequacy, managing assets and liabilities, valuing risk exposures and financial instruments,… (see Fig. 8.1). But they now appeal to models for more ambitious scopes like organizing new product developments or building strategic plans. Moreover, the new quantitative regulatory requirements of Solvency II Directive (Pillar I and SCR, Pillar II and the Own Risk and Solvency Assessment,…), as well as various stakeholders' expectations (including rating agencies, analysts, financial markets,…) push companies to develop more and more sophisticated models, not only for more complex products, but also aiming for better enterprise risk management.

The expanding use of models reflects the extent to which models can improve business decisions. But they can also lead to wrong decisions and potential adverse consequences when they are incorrect or misused. Model risk has become a growing topic for financial institutions, who are becoming more and more concerned. Model risk mismanagement can have strong consequences for financial losses, poor business and strategic decision-making, or damage to the financial institution's reputation. It is for this reason that these institutions place increasing importance on an effective management of model risk.

In this chapter, we first try to understand what model risk is. There are several distinct possible meanings of this expression. We explain what we can learn from finance and what regulators say. Then we discuss several approaches to take into account and measure model risk. We also briefly present what economists say about the impact of model risk on decisions. In the third part, we raise the problem of model risk management and propose several ways to mitigate it. In the last part, we discuss the issue of model risk for the new regulatory framework Solvency II.

8.2 Understanding Model Risk

Before discussing what model risk is, it is important to (re)define what a model is (see also the Introduction in this book). From a mathematical point of view, a model is a theoretical construct representing complex processes by a set of variables and a set of quantitative relationships between them. But from a practical point of view, a model appears as a simplified framework designed to illustrate complex relationships, and mathematical techniques are not the unique ingredient to make it operational.

For this chapter, we follow the definition given in (Parkinson 2011) by the Office of the Comptroller of the Currency of the U.S. Department of the Treasury:

"The term model refers to a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates."

A model that wants to give essence to business decisions has to include three components:

- an information input component, which delivers assumptions and data to the model;
- a processing component, which transforms inputs into estimates;
- a reporting component, which translates the estimates into useful business information.

Models are never perfect and model quality can be assessed in many ways: precision, accuracy, discriminatory power, robustness, stability, reliability, etc. Precision and accuracy apply to models that forecast future values, while discriminatory power applies to models that rank order risks, and robustness and stability apply to models that may use data outliers as input.

One should keep in mind that it is important to understand the capabilities and limitations of a model, but only given its simplifications and assumptions.

8.2.1 What Can We Learn from Quantitative Financial Analysts About Model Risk for Pricing Models?

In a recent book on model risk management in finance (Morini 2011), Mossimo Morini analysed what quants (quantitative analysts) said about models (in particular when they are right or wrong) and distinguished two approaches to defining model risk.

The first approach is called the value approach and is based on papers and books on model risk and model validation written by the famous quant Emanuel Derman (see Derman (2001) and his most recent book Deman (2011)). Derman was one of the pioneers of quantitative modelling between the end of the eighties and the nineties. When he was head of risk management at Goldman Sachs, he was in charge of several trading desks that often have substantial positions in long-term or exotic derivative securities that are marked to market only by means of mathematical models. One of his missions was to verify the fair value of these securities. In a precursor paper Derman (1996), he delineated some of the procedures involved in constructing a financial valuation model and gave recommendations:

1. The financial context: 'Understand the securities, the markets and the way market participants think about valuation and risk factors.'
2. The variables: 'Isolate the most important variables that participants use to analyze value and risk; Decide which of these variables are susceptible to mathematical modelling; Separate the dependent variables from the independent variables; Also decide which are directly measurable and which are more in the nature of human expectations, and therefore only indirectly measurable; For some variables, the uncertainty in their future value has little effect on security values, and they can be treated as known to a good approximation. For other variables, uncertainty is critical. Specify the variables that can be treated as deterministic and those that must be regarded as stochastic.'
3. The mathematical model: 'Develop a qualitative picture that represents how the independent variables affect the dependent ones; Think about how to get the market values of independent observable variables, and how to deduce the implied values of indirectly measurable ones; Formulate the picture mathematically; Decide what stochastic process best describes the evolution of the independent stochastic variables. Consider the difficulties of solving the model, and then perhaps simplify it to make the solution as easy as possible. But only reluctantly give up content for the sake of an easy or elegant analytical solution.'
4. The software: 'Develop a scheme for analytic or numerical solution; Program the model; Test it; Embed it in the software and human environment.'

Derman explained that, once the payoff is accurately described, once the software is reliable and the model is well calibrated to the liquid underlying products, there exists a residual risk that appears to be the core of model risk. Morini then deduced that model risk is the risk that the model is not a realistic/plausible representation of the factors affecting the derivative's value. His opinion is that this approach to model risk is probably the one shared by most practitioners of finance.

The second approach is called the price approach and is based on a paper on model risk written by Riccardo Rebonato (see Rebonato 2002). Rebonato was one of the fathers of the development of interest rate models and he brought major contributions to this field. For him, 'Model risk is the risk of occurrence of a significant difference between the mark-to-model value of a complex and/or illiquid instrument, and the price at which the same instrument is revealed to have traded in the market'. It appears that, as long as the market agrees with the model valuation, there could not be model losses arising from mark-to-market, even if the market model is unreasonable or counterintuitive. The problem arises with very complex or illiquid products, for which market prices are not observed frequently. The model price of a derivative and its market price can suddenly be very different, and a large loss is then needed to be written in the balance-sheet: this is the essence of model risk for him.

Morini deduced that model risk is the risk of using a model different from the market consensus, and when it is forced to be compared with the market (because of a transaction or because the market consensus has become visible), this difference may turn into a loss. The recommendations given by Rebonato to mitigate model risk are then:

- to gather as much information as possible on the approach currently used by the majority of the market players (by getting information about the deals which are struck in the market, by monitoring broker quotes, by attending conferences and other technical events where practitioners present their methodologies for evaluating derivatives,…),
- to evaluate how current accepted pricing methodology might change in the future, to anticipate the patterns in changes in model consensus that may lead to large model losses.

8.2.2 What the Office of the Comptroller of the Currency of the U.S. Department of the Treasury Says

In a guidance on Model Risk Management (Parkinson 2011), the Federal Reserve and the Office of the Comptroller of the Currency (OCC) define model risk as "the potential for adverse consequences from decisions based on incorrect or misused model outputs and reports". It has been identified by the Office that model risk arises for two reasons:

(1) a model is incorrect (all the factors that affect valuation have not been taken into account, some of stochastic variables have been wrongly approximated as deterministic, incorrect dynamics for a factor has been assumed, some assumptions on the market conditions may have become invalid,…); it therefore produces inaccurate outputs with respect to its objective and business uses;

(2) a model is correct but it is implemented and/or used incorrectly or inappropriately (errors appear in the numerical solution to a correctly formulated problem, errors are hidden in thousands of lines programs,….) or there may be a misunderstanding about its limitations and assumptions.

As a consequence, model risk increases with greater model complexity, higher uncertainty about inputs and assumptions, broader extent of use, and larger potential impact.

Banks must be aware of the possible adverse consequences of decisions based on models that are incorrect or misused. The supervisory guidance is intended for them, but also for banking organization and supervisors since they assess organizations' management of model risk.

To highlight the degree of importance the OCC places on model risk management, the guidance explicitly states that "Model risk should be managed like any other risk". The OCC gives in the guidance the key aspects of an effective management. It should include:

- robust model development, implementation, and use: banks have to set up model development and implementation processes that are consistent with the situation and goals of the model users. A sound development process includes: a clear statement of purpose to ensure that the model is developed in line with its use; a suitable economic, financial, mathematical theory underlying the model; robust model methodologies and processing components; a rigorous assessment of data quality; an appropriate documentation.
- effective validation: model validation is defined as the set of processes and activities intended to verify that models are performing as expected, in line with their design objectives and business uses. All model components (inputs, processing, outputs, and reports) should be subject to validation (this applies equally to models developed by the banking organizations themselves and to those developed by vendors or consultants. It is underlined that validation involves a degree of independence from model development and use. Validation must be done by staff who are not responsible for model development or use and do not have a stake in whether a model is determined to be valid (see Chaps. 3 and 6 in this book).
- sound governance, policies, and controls: developing and maintaining strong governance is fundamentally important. Strong governance must include:

  (a) explicit support and structure to risk management functions through policies defining relevant risk management activities, procedures that implement those policies, allocation of resources, and mechanisms for testing that policies and procedures are being carried out as specified;
  (b) documentation of model development and validation that is sufficiently detailed to allow parties unfamiliar with a model to understand how the model operates, as well as its limitations and key assumptions.

Finally, it is mentioned that a model risk management should be applied as appropriate, taking into account the organization's size, nature, and complexity, as well as the extent and sophistication of its use of models.

8.2.3 What the Casualty Actuarial Society, the Canadian Institute of Actuaries, and the Society of Actuaries Say

Very recently, Markus Stricker, Shaun Wang and Stephen Strommen wrote a paper titled "Model validation for insurance enterprise risk and capital models". The research project was sponsored by the Casualty Actuarial Society, the Canadian Institute of Actuaries, and the Society of Actuaries. Contrary to the OCC paper that says that model risk occurs primarily for two reasons, the authors outline five sources of model risk:

1. Conceptual risk: the risk that the modelling concepts/methods are not suitable for the purpose of the application. This risk includes decisions regarding which risk factors have to be taken into account, which mathematical models should be kept, which risks need to be simulated,…
2. Implementation risk: this may be split into two parts: (i) the risk that wrong algorithms were chosen to implement the specified modelling concepts; (ii) the risk that appropriate algorithms were chosen, but they contain coding errors and bugs.
3. Input risk: the risk that the input parameters (data, statistic parameters,…) are inappropriate, incomplete, or inaccurate.
4. Output risk: the risk that the key figures and statistics that can be produced by the model do not support the business purpose or are too sensitive with respect to the provided input parameters.
5. Reporting risk: the risk that the representation of the output for the business users is incomplete or misleading.

The authors explained that the separation between output risk and reporting risk may be difficult to distinguish, because both deal with the outputs of the model. The main difference is that they are addressed to two different groups of people with different skills and responsibilities, and that different levels of detail are provided to these two groups. The detailed output of the model must depict the risk situation very precisely for a technical expert. It must be condensed into a few statistical key figures to be reported to management. But, if the underlying assumptions are not explained, the statistical key figures may lead management easily to misinterpretation.

From the authors' point of view, reporting risk is closely related to the use test under Solvency II. In this framework, an insurance company that uses an internal model must demonstrate that its model is widely used in and plays an important role in its system of governance, in its risk management system, and in its decision-making processes. But the internal model also has to satisfy several tests concerning statistical quality standards, calibration standards, profit & loss attribution, validation standards, documentation standards…

For the authors, although the situations of insurers and of regulators are close to those in the financial services industry, insurers should not blindly adopt the risk models used by banks. Some of the investment risks faced by banks and by insurers are common, e.g. for life insurance and pensions. But even where investment risks overlap, there are major differences in liability maturities (the insurance models usually deal with much longer maturities) and in liquidity considerations that make banking models less applicable to insurers. Moreover, banks are not exposed to insurance claims risks such as disaster-related property damage, health care costs, or longevity. The authors conclude that only a model validation process can help determine when and whether a model used elsewhere can be used appropriately for insurance risk management.

The authors presented three historical examples of model risk failure of governance. They first briefly discussed the role of models in AIG's difficulty with credit default swaps and the role of models at JP Morgan Chase in the case of the London Whale trades. They finally provided the case study of Universal Life with secondary guarantees, and examined the many ways models and model risk played a role in the developing story of this new life insurance product. These examples are important because they show where the chains of risk model management failed.

8.2.4 To Sum up

Insurance companies rely heavily on models for a wide range of applications. The level of sophistication of models used varies in practice widely from relatively simple spreadsheet tools to complex projection models. Regardless of the level of sophistication, it is clear that adverse consequences of an intensive use of models are mainly the possibility of financial losses, incorrect business decisions or damage to the company's reputation. Common ingredients of model risk are:

- errors in the model design process such as errors in the mathematical theory, statistical analysis,..;
- errors in the model development process such as errors in the computer code underlying a model,…;
- errors in the model production process such as errors in data inputs and assumptions, or errors in model execution;
- errors in the model applications such as misuses of model results, or use of models whose performance does not meet company standards.

8.3 Is it Possible to Measure Model Risk in the Model Design Process?

8.3.1 Measures of Model Risk

Although model risk has become more and more important for practitioners as well as for researchers, there is no clear and well-accepted notion of how model risk should be measured. The measure of model risk is an important question that a priori only concerns model risk due to potential errors in the model design process.

In practice, the measure of model risk is addressed by comparing the results of different models. More often, model risk is investigated by varying model parameters (what is called parameter sensitivity) in the context of risk measurement and management. More recently, the effect of changes in the probability law that defines the underlying model has been investigated.

Three different paradigms have been proposed for evaluating uncertain outcomes in the presence of model risk. The first one, which consists of averaging over possible models, is referred to as Bayesian Model Averaging. The second one is derived from the statistical literature on robustness. The last one is based on the worst case or "maxmin" approach.

Bayesian Model Averaging (BMA)

Bayesian Model Averaging is a technique that is used in the presence of uncertainty for a model selection process. Standard statistical practice ignores model uncertainty and selects a model from some class of models to proceed as if the selected model had generated the data. This approach leads to over-confident inferences and decisions that are more risky than one thinks they are. BMA averages over many different competing models incorporate model uncertainty. It improves out-of-sample predictive performance.

However, several issues arise when trying to apply this approach in practice as in any Bayesian method. First, this method not only requires specifying a prior distribution on parameters of each model, but also a prior probability distribution on possible models, which could be more delicate. The second issue is computational: the posterior distributions involved in the formulas are not explicit and sampling from them requires the use of Markov Chain Monte Carlo algorithms, which are in general computationally intensive.

The problem of parameter and model uncertainty in insurance with a Bayesian approach has been addressed e.g. by Cairns (2000).

The Robustness Approach

The Robustness Approach is proposed in Glasserman and Xu (2014) and has the following goals:

- "bound the effect of model error on specific measures of risk, given a baseline nominal model for measuring risk"
- "identify the sources of model error to which a measure of risk is most vulnerable and identify which changes in the underlying model have the greatest impact on this risk measure".

For the first goal, an upper and lower bound are calculated on the set of risk values over a range of model errors within a certain 'distance' of the baseline model. These bounds are somewhat analogous to a confidence interval; but whereas a confidence interval quantifies the effect of sampling variability, the robustness bounds quantify the effect of model error. In general, the distance that is used to quantify changes in probability laws is the relative entropy. In Bayesian statistics, the relative entropy between posterior and prior distributions measures the information gained through additional data. The relative entropy is interpreted as a measure of the additional information required to make a perturbed model preferable to a baseline model. Thus, relative entropy appears as a measure of the plausibility of an alternative model.

For the second goal, the changes to a nominal underlying model that attain the bounds on the measure of risk have to be identified. Simply quantifying the potential magnitude of model risk is of limited value if it is not possible to identify the sources of model vulnerability that lead to the largest errors in measuring risk.

The Worst-Case Approach

In this approach one considers a range of models and minimizes the loss encountered in the worst-case scenario (see e.g. Cairns (2000) for an example in finance). With respect to model averaging procedures, worst-case approaches are more conservative, more robust, and require less sophisticated inputs on the part of the user. Thus, they are more interesting for the design of a robust, systematic approach for measuring model uncertainty.

8.3.2 Model Risk and Ambiguity

In decision theory and economics, ambiguity aversion (also known as uncertainty aversion) describes a preference for risks whose laws are known over risks whose laws are unknown. The distinction between ambiguity aversion and risk aversion is important. Risk aversion comes from a situation where a probability can be assigned to each possible outcome: it is the reluctance of a person to accept a bargain with an uncertain payoff rather than another bargain with a more certain, but possibly lower, expected payoff. Ambiguity aversion applies to a situation when the probabilities of outcomes are unknown and it is defined through the preference between risky and ambiguous alternatives, after controlling for preferences over risk. Ellsberg (1961) has shown that aversion to ambiguity clearly plays a role in decision-making. Aversion to ambiguity can strongly affect decision-makers' behaviour and resolve some paradoxes of classical decision theory.

In Robert and Therond (2014), they considered ambiguity concerning the probability distribution that is used to evaluate risk measures. Almost all models used in the theory of risk measures assume that the distributions of risks are perfectly known. They rather assumed that there exists an uncertainty about a parameter of the risk distribution and that this uncertainty may modify decisions of the individuals if they are averse to ambiguity. This paper examines the effect of ambiguity-aversion on the valuation of risk reduction. In particular, we give an example where ambiguity-aversion and risk-aversion have different consequences on effort to reduce risk.

8.4 Model Risk Management or How to Mitigate Model Risk

We present two views on model risk management. The first one comes from a paper titled "The evolution of model risk management" for banks (PwC's Financial Services Regulatory practice (2013)). The starting point of this paper is that the supervisory guidance issued by the OCC in 2011 is a consequence of systemic weaknesses in bank model risk management programs observed, despite the seminal Risk Bulletin on Model Validation issued by the OCC in 2000. The paper discusses what the new model risk management guidance means holistically for banks and provides a roadmap for evolving the model risk management function.

The second one is extracted from the paper by Stricker, Wang and Strommen "Model validation for insurance enterprise risk and capital models" (Stricker et al. 2014) and gives an exhaustive process guide to insurance companies for setting up a suitable model risk management.

8.4.1 Model Risk Management for Banks

From PwC's Financial Services Regulatory practice (2013), an efficient model risk management for banks should include three levels: the first level is the responsibility of model developers and users, the second level is the responsibility of the Model risk management Unit ("MRMU"), the third level is the responsibility of the Internal Audit.

First level: model developers and users

Model risk should primarily be mitigated or controlled at its source, that is:

- by the personnel responsible for designing and building models for the bank (or, for vendor models, personnel responsible for selecting vendor models),
- by the personnel responsible for using models in such activities as business decision-making, financial and regulatory reporting, risk management, or valuations.

The authors recommend injecting more formality into the model development and implementation control framework through the development of specific model risk control processes. Banks should write documents to specify the exact set of control activities expected for model developers and users in the following areas to control/mitigate model risks at their source:

- Model design and methodology.
- Internal and external data inputs and assumptions.
- Model estimation and testing, including statistical validity, outcomes analysis, sensitivity analysis, and stress testing.

- Model deployment requirements, such as choice of model production platform, production data input requirements, choice of model software, model coding requirements, etc.
- Selection and deployment of vendor models.
- Model pre-implementation testing.
- Model documentation requirements, both development and implementation.
- Model output reporting.
- On-going model risk monitoring processes, including monitoring of: model performance, model operational process risks, model weaknesses and limitations, and compliance with any model risk mitigants required by the second line of defence.
- Identification and communication of key model risks, weaknesses, and limitations to model users, including risks associated with inherent model imprecision, key assumption sensitivities, and any risks identified through other first and second line control activities.
- Model approval process.
- Model change management process.

Second level: Model risk management unit

A model risk management unit should be created to:

- "Monitor the effectiveness of the first line control framework through the performance of independent model risk identification and measurement activities (the primary role of the Model Risk Management Unit is to perform periodic model validation testing on a one-, two-, or three-year cycle– or when material changes occur to a model)."
- Validation can be considered from both a quantitative and qualitative point of view. Back-testing and benchmarking are key quantitative validation tools. In back-testing, the predicted risk measurements will be contrasted with observed measurements using a workbench of available test statistics to evaluate the calibration, discrimination and stability of the model. The aim of benchmarking is to compare internal risk measurements with external risk measurements in order to better evaluate the quality of the model rating system. These methods aim at assessing the consistency of the estimated parameters and models with those obtained using other estimation techniques or other risk measures (see Chap. 4 in this book).
- "Ensure that appropriate short-term model risk mitigants and long-term remediation plans are deployed by model developers and users to reduce residual model risks to acceptable levels."
- "Engage regularly with senior management and board governance/oversight groups with respect to the bank's model risk levels, proposed model-use mitigants designed to reduce model risk levels, and overall model risk management activities."

- The third level, the Internal Audit, audits the contents and compliance with model risk management policies, procedures and standards within the first and second levels.

8.4.2 Model Risk Management for Insurers

In Stricker et al. (2014), the authors observed that there is no commonly accepted standards for a validation process in the insurance industry. They explained that most often they are just referred to as validation principles that essentially express that the range and rigor need to be in line with the potential risk. They therefore proposed an exhaustive process guide to mitigate model risk.

The validation process should follow the major subcategories of model risk mentioned above and it has to respect the dependencies between these subcategories:

```
                 Implementation risk
               ↗                     ↘
Conceptual risk                        Reporting risk → Output risk
               ↘                     ↗
                      Input risk
```

A consequence of these dependences is that if the validation for a risk subcategory fails, then there is no need to validate the dependent subcategories.

(i) *Conceptual Risk*

Rather than describing the application setting in general terms, the documentation of the purpose of the model should do the following:

- "Include reference to the relevant sections from underwriting, investment, or risk management guidelines or other appropriate documents that make reference to the use of the model output."
- "Include a description of the users: at this stage of the validation, it is necessary to verify only that the reports are addressed to a well-defined audience."

The focus of the concept validation should do the following:

- "Check whether the concept documentation makes reference to external sources."
- "Check that the concept documentation describes how the modelling pieces are connected and why they can be used together."
- "Check whether there is enough emphasis on describing the limitations of the concepts."
- "Check vendor model concepts."

(ii) *Implementation Risk*

The authors observed that most actuaries and risk managers implement their models in their departments, which are usually not staffed with IT professionals. It follows that a realistic assumption is that programing codes may contain errors. Implementation risk seems to be the most underestimated model risk for them.

The focus of the implementation validation should be to perform the following:

- "Check that the risk modelling experts have been involved in the selection of the algorithms that implement the modelling concepts."
- "Check whether everything in the model development is versioned."
- "Check whether the accountability for code changes, bug fixes, or improvements is clear"
- "Check whether there is an automated test procedure in place that is run in regular time intervals or, even better, after every new version is checked into the versioning system."
- "Check the test coverage"
- ….

(iii) *Input Risk*

The principles for the validation of the inputs can be expressed very simply: internal and external data must be demonstrably appropriate, accurate, and complete. Although it is easy to understand, the authors themselves confessed that it is not easy to derive explicit guidance for the validation of this risk.

It is important to be able to detect quickly any substantial deviation of input parameters from their values used in the models (e.g. for pricing and reserving). Moreover, parameters that stay the same with respect to the initial assumptions on the risks do not necessarily have to be correct because it could be that changes in the underlying portfolio require changes in parameters. Such changes are sometimes difficult to detect and assess in a validation process and it is recommended that the validation team ask the others teams (the reserving team, for example) for a summary of major changes in the ex-post source data used for the recalibration of models. The input risk model is less important in situations where the observations come quickly (calculation of reserves) than for situations where the observations come less frequently (like for the economic capital evaluations).

Back-testing processes that compare historical results to those produced by the current model in order to validate both the reasonableness and the implementation of the assumptions are very useful.

(iv) *Output Risk*

Validation of output risk should be an additional and subsequent process that has been performed to assess the implementation and input risks. The assessment of output risk needs to check whether skilled people can interpret the model outputs in the context of their intended application.

The validation should include the following checks:

- "Check that the correct input data set and model version are referenced by the outputs."
- "Check whether the outputs can be reproduced."
- "Check whether breaches of input parameter limits are indicated in the output."

- "Check whether documentation exists concerning the selection of input parameters against which the sensitivities of the outputs are measured."
- "Check whether the sensitivities are documented."
- "Check materiality of input parameters based on the sensitivities."
- "Check whether the ranges of the output key figures are made available."

(v) *Reporting Risk*

This is the final step in a validation process and for the authors the most crucial. They recalled that this step is not outside the scope of model validation even if it is closely related to the use test. If this step is not included in the validation process, then the model's use in the intended real-world application cannot be assessed. The main issue is that the people who build models and review them have a background that is almost exclusively quantitative (actuaries, financial engineers,…), and those who receive and use reports on the models often have a more limited quantitative background.

The validation should include the following checks:

- "Check that the reports clearly state which model and data version were used."
- "Check that business users are made aware of situations in which some of the parameters are outside a comfort range or even outside the agreed limits."
- "Check whether the frequency and timing of the reports is in line with the decisions they support. It must be stressed that this has nothing to do with an automated decision process."
- "Check whether the results are communicated using institutionally accepted metrics that are readily understood by all end users. Metrics that may capture or describe the risks well, but are not commonly known or used in the company, introduce reporting risk."
- "Check whether the report uses any means to convey how robust the key figures are. Simply providing point estimates of the key figures does not give enough information to decision-makers. They have to be made aware of the fact that estimation errors for the parameters and different modelling assumptions yield ranges of outcomes."

8.5 Model Risk and Solvency II

8.5.1 Model Risk and the Standard Formula

The Solvency II standard formula consists of a number of risk modules whose outcomes are aggregated step by step to reach a single capital requirement: the Solvency Capital Requirement (SCR). There are modules that are determined by calculating how a prescribed scenario would affect the insurer's balance sheet (e.g. in the case of equity risk, the scenario is a sharp fall in the stock market). The outcome of such a module is the decline in the insurer's own funds as a result of the scenario. There are also modules whose outcome is the product of an explicit calculation instruction. The outcomes of all risk modules are then aggregated step by step.

The standard formula for SCR aims to capture the material quantifiable risks that most undertakings are exposed to. The standard formula might, however, not cover all material risks a specific undertaking is exposed to. A standard formula is, by its very nature and design, a standardized calculation method, and is therefore not tailored to the individual risk profile of a specific undertaking. For this reason, in some cases, the standard formula might not reflect the risk profile of a specific undertaking and consequently the level of own funds it needs.

The insurer should assess deviations between its risk profile and the assumptions underlying the SCR calculation on a qualitative and quantitative basis. If this assessment indicates that the insurer's risk profile deviates materially from the assumptions underlying the SCR calculation, the insurer should also quantify the significance of the deviation. The insurer has to understand the assumptions underlying its SCR calculation and consider whether those assumptions are appropriate. To do this, the insurer will have to compare those assumptions with its own understanding of its risk profile. This process needs to prevent an insurer from simply relying upon regulatory capital requirements as being adequate for its business.

The areas in which differences between the insurer's risk profile and the assumptions underlying the SCR calculation may arise to which the insurer needs to give due consideration are:

- from risks that are not considered in the standard formula; and
- from risks that are under/overestimated by the standard formula compared to the risk profile.

The assessment process should include (see EIOPA-Bos-13/25 2014):

(a) an analysis of the risk profile and an assessment of the reasons why the standard formula is appropriate, including a ranking of risks;
(b) an analysis of the sensitivity of the standard formula to changes in the risk profile, including the influence of reinsurance arrangements, diversification effects and the effects of other risk mitigation techniques;
(c) an assessment of the sensitivities of the SCR to the main parameters, including undertaking-specific parameters;
(d) an elaboration on the appropriateness of the parameters of the standard formula or of undertaking-specific parameters;
(e) an explanation why the nature, scale and complexity of the risks justify any simplifications used;
(f) an analysis of how the results of the standard formula are used in the decision-making process.

If the conclusion of these qualitative and quantitative assessments is that there are significant deviations between the risk profile of the insurer and the SCR calculation, the insurer needs to consider how this could be addressed. It could decide to align its risk profile with the standard formula, to use insurer specific parameters, where this is allowed, or to develop a (partial) internal model. Alternatively, the insurer could decide to de-risk.

In EIOPA-14-322 (2014), Eiopa presents the underlying assumptions of the standard formula that are used for the SCR calculations. This document should help insurers in their analysis of the adequacy of the SCR calculation with their risk profile, and should be read in conjunction with the Guidelines on ORSA (see the section "Model Risk and the ORSA" below).

A. The correlation structure

Correlations that are intended to measure the connection between different risk modules are used in the aggregation steps. Basically, these correlations provide an estimate of the probability of different risks occurring simultaneously. In cases where risks always occur 'simultaneously' the outcomes of the modules must be aggregated to determine the capital requirement. If risks do not occur simultaneously, they are said to be diversified and the capital requirement is lower than the sum of the modules.

The aggregation formula in the standard formula is based on the assumption that the dependence between the distributions can be fully captured by linear correlations. But it is well known that linear correlations are insufficient to fully reflect the dependence between distributions and there exist a number of examples where the use of linear correlations could lead to incorrect aggregated results, i.e. producing either an under-estimation or an over-estimation of the capital requirements at the aggregated level. Therefore, linear correlations are not an appropriate choice for the aggregation of risks in many cases.

Several questions for the aggregation issue should be raised by insurers:

- Is dependence between the distributions non-linear? Are there tail dependencies? Risk aggregation with dependence uncertainty refers to the study of the sum of individual risks with known marginal distributions and unspecified dependence structure. This field has, for example, been investigated by Bernard et al. (2014) and Embrechts et al. (2014).
- Is the shape of the marginal distributions significantly different from the normal distribution? What happens if marginal distributions are skewed?

B. Risks not explicitly formulated in the standard formula calculation

We have already mentioned that not all quantifiable risks have been explicitly formulated in the standard formula. As a consequence, some risks which are not explicitly included in the standard formula may be relevant for a particular undertaking. Examples are: inflation risk, reputation risk, liquidity risk, contagion risk, legal environment risk.

For some of these risks, it has been assumed in the standard formula that the exposure is not always material enough to justify a separate and more granular SCR quantification within the context of the standard formula. For other risks, they are not explicitly formulated in the standard formula calculation because it is very difficult to assess them in general for all undertakings.

The ORSA explains that it would be inappropriate to cover these risks through pillar 1 capital requirements but these should be covered instead through pillar 2 requirements, in particular through risk management requirements for an appropriate monitoring and disclosing of the risk profile of the undertaking.

We give here some examples of other sets of assumptions that have been made without discussing them (see EIOPA-14-322 (2014) for other risk modules and more details).

C. Market risk

It is assumed that the sensitivity of assets and liabilities to changes in the volatility of the market parameters is not material.

D. Interest rate risk

The main underlying assumptions for the Interest rate risk module can be summarized as follows:

- "Only interest rate risk that arises from changes in the level of the basic risk free interest rates is captured."
- "Volatility and changes in the shape of the yield curve are not covered explicitly in the interest risk sub-module."
- "The undertaking is not exposed to material inflation or deflation risk."

E. Life underwriting risk

The underlying assumptions for the Life underwriting risk module can be summarized as follows:

- "The calibration of the Life underwriting risk parameters captures changes in the level and trend of the parameter. It is assumed that the volatility risk component is implicitly covered by the level, trend and catastrophe risk components. This is considered to be acceptable, since volatility risk is thought to be considerably lower than the trend risk."
- "The dependence of benefit payments on inflation is not material."
- "The insurance portfolios are well-diversified with respect to: age, gender, smoker status, socio-economic class, level of life insurance cover, type of insurance cover, degree of underwriting applied at inception of the cover and geographical location."

F. Longevity

The underlying assumptions for the longevity risk sub-module can be summarized as follows:

- "The annual mortality improvements follow a normal distribution."
- "For the simplified calculation of the capital requirement for longevity risk it is assumed that the average age of policyholders within the portfolio is 60 years or more."
- "It is furthermore assumed that the average mortality rate of the respective insured persons does not increase by more than 10 % each year."
- …

It is important to underline that it is not on the scope of the risk margin to take into account model risk potentially existing in the SCR. The risk margin is defined as the hypothetical cost of regulatory capital necessary to run-off all liabilities (following a financial distress of the company, it should be determined in a way that enables the insurance obligations to be transferred to a third party or to be put in run-off).

8.5.2 Model Risk and the ORSA

At the heart of the prudential Solvency II directive, the Own Risk and Solvency Assessment (ORSA) is defined as a set of processes for decision-making, strategic analysis, for managing overall solvency needs and regulatory capital requirements.

The ORSA includes assessments of the insurer's adequacy of current and future economic capital resources given its risk appetite and tolerance levels. Current and future capital levels should take into account longer term business strategies for the calculation of the SCR, risk tolerances, current and new business plans and risk mitigation processes and techniques. An insurer should understand how its risk management and capital management approaches and processes interact to optimize its capital base, to transfer or retain risks.

When performing the ORSA, the insurer takes into account the business strategy and any strategic decisions influencing the risk situation and regulatory capital requirement, as well as overall solvency needs. The Board and senior management need to be aware of the implications of the strategic decisions on the risk profile and regulatory capital requirements and overall solvency needs of the company. Any strategic or other major decisions that may affect the risk and/or own funds' position of the insurer should be evaluated. Depending on the magnitude or significance of the proposed decision, this may not always require a full ORSA process and management may perform a simplified analysis.

At the heart of Solvency II prudential, risk and solvency assessment (ORSA) is defined as a set of decision-making, strategic analysis, management of overall solvency needs and requirements regulatory capital.

When the insurer uses the standard formula for the calculation of the SCR, the insurer has to determine whether the risk profile deviates significantly from the assumptions underlying the SCR. There are a number of reasons that could account for the differences that have nothing to do with deviations of the risk profile, such as (see EIOPA-BoS-13/25 2013):

(a) "The insurer may operate at a different confidence level or risk measure for business purposes compared to the assumptions on which the SCR calculation is based. For instance, it may choose to hold own funds for rating purposes, which represents a higher confidence level than that used to calibrate the SCR.
(b) The insurer may use a time horizon for its business planning purposes that differs from the time horizon underlying the SCR.
(c) In the ORSA, the insurer may consider any agreed management actions that could influence the risk profile."

When the insurer uses an internal model for the calculation of the SCR, the insurer has to assess the assumptions underlying its calculation of the SCR according to its internal model in order to ensure they remain adequate and that the internal model continues to reflect its risk profile appropriately.

The insurer also needs to demonstrate that the internal model plays an important role in the ORSA. An internal model is in itself a tool for the ORSA and the ORSA is a tool for the internal model in the sense that the performance of the ORSA gives input to the on-going exercise of ensuring compliance with the tests and standards. According to the requirements, internal model users have to comply, at the approval date and in an on-going concern, with the use test, statistical quality standards, calibration standards, profit and loss attribution test, validation standards and documentation standards. Each feature of the ORSA could play an important role in this exercise.

In the case of the use of an internal model, the ORSA includes the assessment of (see EIOPA-BoS-13/25 2013):

(a) "the impact of the excluded material risks or major lines of business on the solvency position;
(b) the interrelationship between risks which are in and outside the scope of the model;
(c) the identification of risks other than those covered by the internal model, which may trigger a change to the internal model."

8.6 Conclusion

We have seen that model risk is an inherent part of day-to-day life for insurance companies. Sources of this risk include bad data, incorrect assumptions and methodology, poor implementation and usage… An effective model risk management must include at least:

- Independence of various functions, in particular, the model development, risk control, audit and model risk control functions (whether it is internal or external to the institution).
- Emphasis on documentation at each stage in the model life-cycle (model validation and measurement of model risk should be conducted regularly).
- Dissemination of model risk culture and user education along with model results (in particular recognition of models as a "work-in-progress" that need to be continually re-examined and improved, rather than as a one-time effort).

Finally, note that regulators want companies to use their risk models widely to ensure that they take care to build suitable models and thus hope to minimize the model risk. But if they also prescribe a specific model (e.g., the standard model for Solvency II), and if a large part of the companies use the same prescribed model, regulation could lead to a pernicious systemic risk model.

Chapter 9
Meta-Models and Consistency Issues

Jean-Paul Laurent

Abstract Rather than considering a "model" as a one-piece object, we can translate and adapt the concept of meta-models, commonly used in computer science, to the field of insurance management. We actually deal with a number of interconnected models. These models involve common concepts such as risk and value, assets and liabilities, reserves, management actions, etc. To avoid cacophonies (i.e. operational inefficiencies), every piece has to be placed in the right order. Depending on objectives and context, different levels of modelling will be required. Coherence in the modelling process does not mean uniformity. It is vital to understand correctly how models can effectively enhance business performance, yet not be blurred by undue complexity.

9.1 Life Insurance Models in a Nutshell

Before going into the topic of this chapter, let us recall the two main modelling issues in the life insurance world:

- Computation of reserves since the scope of mark-to-model and best estimate approaches encompasses most of insurance liabilities.
- Computation of risk measures and the corresponding solvency requirements: pillar 1, such as standard SCR formula or internal model approaches, pillar 2/ORSA approaches in Solvency II. To a certain extent, EIOPA dynamic scenario-based stress tests involve the same prerequisites as ORSA.

These two items may be combined to assess solvency over a prescribed time horizon, in most cases a single year in Basel II and III or Solvency II contexts, leading to the well-known issues of nested simulations.[1] To be more specific, under

[1]See Bauer et al. (2012), Broadie et al. (2011).

J.-P. Laurent (✉)
Université Paris 1 Panthéon-Sorbonne, Paris, France
e-mail: laurent.jeanpaul@free.fr

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_9

standard modelling approaches, risk measure computation requires a great deal of reserve calculations, which may lead to process jamming. Clearly, when it comes to multi-period assessment of risk, as in the ORSA framework, computational and modelling issues blow up.

As discussed in the previous chapters of this book, a wide range of specialized modelling approaches exist, focusing on items such as economic scenario generation, management actions, loss quantile-based or standardized risk measures. To ease the exposition, we will pay extra attention to static or dynamic representations on one hand and on the dual view regarding risk models, i.e. internal models and standard formulas on the other. These two themes are well documented and allow for comparisons between insurance and finance fields. Even though the computation of risk-based capital charges is more challenging on theoretical grounds, proper accounting of best estimate insurance liabilities is by far the largest management issue for life insurance companies.

The term meta-model has different meanings, the most common refers to a model of a model, which can be understood as a simplified version, easier to grasp than the original one. For instance, a Vasicek model could be seen as a simplified version of a more general multifactor Gaussian HJM model,[2] leading to much simpler closed forms for interest rate option prices. This is related to the concept of nested models. As for Markov chains, the concept of embeddability discusses whether or not a discrete time Markov chain could be seen as a restriction to a discrete time scale of a continuous time Markov chain.[3] A similar issue is whether we can relate discrete and continuous time versions of a stochastic process or not. For instance, an Ornstein-Uhlenbeck standard mean reverting process or square root type is an AR(1) process when considered over a discrete time scale.

While the above is meaningful in our context, we will consider meta-modelling as the science or art that deals with model design and use. Our concern is not to determine the winner among a set of non-nested competing models, but to investigate the implications of using a number of models simultaneously. We are clearly faced with consistency issues and even though model choice is mostly driven by ease of implementation, model multiplicity might end up with undue overall complexity.

Eventually, meta-models refer to a highly specialized field of research. A model is viewed here as an engine providing some outputs. A meta-model or a surrogate model provides much easier to compute approximations to true model outputs. Response surface methodology, artificial neural networks, multivariate adaptive regression splines and radial basis function approximations are among the numerous techniques involved. As for pricing and risk measurement purposes in life insurance or finance businesses, kriging, a technique derived from geostatistics

---

[2]See Andersen and Piterbarg (2010) for a comprehensive review of interest rate models in theory and in practice.

[3]This is related to the existence of an infinitesimal generator. See Kijima (1997) for mathematical details.

and the modelling of spatial data, is an appealing route. Unlike usual spline interpolation methods, it provides some confidence intervals on meta-model outputs. This is of high importance: when it comes to the assessment of reserves or the risks faced by a company, not being able to assess the magnitude of the approximation leaves a strong sense of discomfort. However, kriging approaches need to comply with some specific constraints when it comes to finance and insurance models. For instance, output approximate prices should not present arbitrage opportunities and the benefits of diversification should be acknowledged when it comes to approximating risk measures. These issues are currently on the academic research agenda, but still in limbo and will not be discussed thereafter.[4]

This chapter is organized as follows. In Sect. 9.2, we will discuss the interplay between models and markets. This, of course, concerns the computation of best estimates, mark-to models calibrated up to some extent to market observables. As for risk models and especially standard formulas, we will investigate the extent to which they are properly connected to statistical features, i.e. actual changes in market risk factors. This could be broadly seen as consistency between models and data. In Sect. 9.3, we will consider practical constraints that drive model building: pragmatism is associated with model multiplicity. Regulatory constraints on internal risk models and their departure from standard formulas (benchmarking, regulatory floors) also shape model construction. Section 9.4 investigates the practical implication of using models based on different premises, models need to be recalibrated to one another, making model governance trickier and leading to operational complexities. Section 9.5 considers the increased use of ad hoc parametric approaches to approximate the output of internal pricing or risk models. Section 9.6 discusses the pros and cons of modelling shortcuts, i.e. replacing the output of a questionable model by a parameter under management control. Section 9.7 reviews the standard approximation techniques, such as replicating portfolios, LSMC or the use of simplified closed-form expressions for prospective risk assessments such as ORSA. We end up with a brief reminder of practical challenges regarding model multiplicity in life insurance businesses.

9.2 Models and Markets

The computation of reserves is likely to be the key challenge to life insurance companies. It is, therefore, worth recalling the theoretical context that underpins valuation procedures and how models come into play. The main risk drivers, or risk factors, are the changes in interest rates and the aggregate fluctuations in mortality rates or related biometric quantities. To ease the exposition, we will leave aside other possible risk factors that might be related to changes in currencies, stock or real estate prices.

[4]Applications to finance and insurance are not well developed yet. When dealing with quantile based risk measures, we mention Chen et al. (2012), Chen et al. (2013), Liu (2010), Liu et al. (2010).

Nowadays it is quite common for actuaries to think in terms of stochastic discount factors or "state-price deflators" or equivalently "pricing kernels" to follow the academic jargon. This allows to price contingent liabilities and in appropriate cases, to price them in a way that is consistent with observable and reliable market prices. In the mere case of fixed payables, the general approach collapses to standard discounting. Focusing on the simplest case is of primary importance, since it is quite likely that the main driver of reserve uncertainty lies in the choice of the risk-free discounting curve:

- This starts from the choice of market data, either a set of collateralized or uncollateralized swap rates, index on Libor or overnight rates, yet it could also relate to rates inferred from currency-based bond prices. The way interbank default risk is taken into account, the crucial choice of an ultimate forward rate and how rates are interpolated and extrapolated up to that ultimate rate are obviously key points and are not technical involved. In Europe, EIOPA provides such a curve for the relevant currencies.[5] Clearly, the way market rates are translated to regulatory discount rates is of primary importance for the computation of reserves and, here, more importantly when designing hedges, which requires the calculation of risk sensitivities. When the market rate over a given time horizon moves by one basis point, the corresponding regulatory discount rate over the same time horizon may not change accordingly. Thus, both levels of discount rates and their changes are concerned at that stage.
- The second issue of importance is that, on top of this model-based regulatory discount curve, insurers are led to add various items, the so-called volatility adjustment and, under a number of restrictions, a matching adjustment so that the discount rate of liabilities becomes tied to the rate of return on the assets held to match liabilities. These items have led to numerous interactions between the insurance industry and the European regulators over recent years.

One potential issue regarding long-term guarantee products[6] is the reliance on the long part of the EIOPA curve, as this is largely model-based, due to the use of a normative ultimate forward rate (UFR) and the Smith-Wilson interpolation/extrapolation scheme and leads to low-volatility of long maturity risk-free discount rates. Thus, even without accounting for shifts in credit spreads, it creates some discrepancies between assets and liabilities of similar maturities and cash-flow schedules (see Lageras 2014).

[5]Colloquially, anyone investigating competing models has the impression of barging into a pet shop. The classical review paper by Hagan and West (2006) accounts for about 20 yield curve interpolation methods, leaving aside numerous parametric models and more statistical approaches (including the stochastic kriging briefly discussed below). They were not aware of the contribution of Smith and Wilson (2001) which underpins the official EIOPA interpolation methodology. It is conspicuous that actuarial yield curve modelling streams were apart from those followed by fixed income quants routinely involved in the pricing of trillions of swaps or billions of corporate or sovereign bonds.

[6]See https://eiopa.europa.eu/Publications/Reports/EIOPA_LTGA_Report_14_June_2013_01.pdf for some context.

As such, it comes as no surprise that this model-driven and somehow ad hoc shortening of liability duration needs to be compensated by a well-calibrated volatility adjustment. We should incidentally note that the volatility adjustment is not only required to balance credit spread volatility, but as an extra modelling layer to mitigate the unexpected consequences of a questionable yield curve stripping methodology. This is typical of constraint-based modelling (see below) whereas repair patches alleviate the perverse effects of previously set convenient shortcuts. Software and high-tech industries provide us with a number of textbook cases where such processes swerve off the road. To keep on track, an overall view of the issues at hand, a clear and sharp understanding of the technicalities and a perfectly driven implementation process are required.

Therefore, market driven input interest rates are filtered out by regulatory technical rules. This is to account for the specificities of the life insurance industry (long-term commitments, illiquidity of life insurance contracts) and the limited reliance on the longer part of the yield curve or on market implied credit spreads. Even though a number of theoretical arbitrage-free dynamic models of the yield curve predict constant ultimate forward rates, the current level of such UFR and the speed of convergence to this UFR in the Eurozone could be revisited. Hopefully, such discounting rules will be stabilized in the EU. Nevertheless, what drives risk is the possibility of changes of rules under changing economic environments and the degree of national discretion. It may be a difficult modelling task to lead an Own Risk and Solvency Assessment Exercise (ORSA). It is almost impossible to guess how key parameters would be recalibrated in stressed environments. Thus, forward looking risk exercises have to be understood as being set up under a fixed regulatory framework.

The second issue of importance is the range in which interest rates are likely to vary over the coming years. This range could be partly inferred from macroeconomic analyses; forward guidance by central banks is quite useful from this perspective. On the other hand, one-factor models, quite useful due to ease of implementation, could lead to implausible trajectories. AR(1) models will do a poor job when it comes to economics, mean-reversion parameters are prone to statistical noise and more involved state-space models are difficult to handle.

Going back to market observables such as interest rate options should then be considered when dealing with the pricing of life insurance liabilities. It also provides a direct route that bypasses the use of probabilistic models and stochastic discount factors. It will be later advocated in this chapter that typical life insurers are both short of out-of-the-money caps and out-of-the-money floors and thus would be better off if interest rates remain inside a reasonable corridor. Stated differently, extra rate volatility and pronounced volatility smiles are not the best context for looking for long-term solvency.

To investigate the above issue, let us, for instance, consider a commitment to pay one euro at a 5 year horizon, we assume the payment date is certain, provided that the 3 months (Euro) Libor rate is above 4 %. Provided that the interest rate derivatives, here the cap and floors market, is liquid enough to provide reliable prices, we will readily get the price of the above contingent (to the level of rates) 5 year zero coupon, as today's price of a 5 year so-called digital cap with a 4 %

strike. This is mere algebra. Please note that at this stage, we do not need to be bothered with a number of confusing concepts such as risk-neutral densities, historical (or statistical) probability measures and how these objects are related through risk premiums. We do not even need to call upon probability theory! As can be seen from this example, the higher the price of the digital cap, the higher will be the risk-neutral probability of rates being above that 4 % strike. This is tautological. Then, this risk-neutral probability is the product of the actual probability of the reference Libor rate to be above the prescribed 4 % threshold times a risk premium (to be investigated below).

As can be seen from the digital cap example, the higher the price of out-of-the-money caps and floors the higher the priced dispersion of future Libor rates. A properly calibrated pricing kernel (or risk-neutral ESG) will account for such features. The intuition behind risk-neutral densities and pricing kernels may not be fully understood. It is then tempting to criticize the approach as too abstract. Let us put things differently. If one trades interest rate floors with zero strike at a positive premium, then the market puts a positive weight on negative rates. Rather than risk-neutral pricing, we should talk about market consistent pricing of liabilities. Unless we distrust market prices, say due to irrational exuberance or market illiquidity, an ESG should comply with such material information to provide best estimates (understood as market consistent) of insurance liabilities. It is sometimes advocated that models are performative, for instance interest rate option prices would be driven by disputable pricing models. Let us go the opposite way: If insurance companies were led to hedge against extreme changes in rates, it would raise the price of these hedges. Consequently, market-implied out of the money volatilities would develop, leading to a self-fulfilling fear of strongly diverging rates.

When we need to switch to the generation of meaningful interest rate scenarios, for the purpose of long-term solvency assessment (ORSA, Stress Tests) in the real world, we need to solve a puzzle. Up to now and under some restrictions, we advocated that market prices could be meaningful inputs to assess insurance liabilities. The problem is that, unless some more or less arbitrary restrictions are put on risk premiums, one cannot infer actual probabilities from the risk-neutral ones. Stated slightly differently, such inference is highly subject to model risk. We will use a dual perspective to the construction of real world ESG together with risk-neutral ones.

- When we remain stuck to a standard Brownian framework, say a Vasicek type model to illustrate our purpose, we are faced with a huge restriction. As a consequence of the Girsanov theorem, volatility needs to be the same under risk-neutral and real worlds. This may lead to the kind of implausible explosive scenarios that are very much disliked by life insurers' executive committees.
- On the other hand, when we turn back to discretized versions of the original model, say random samples of rate trajectories, we have much more, too much, flexibility in computing risk premiums. We could reweight the probabilities of the sample paths quite arbitrarily, only being subject to weights that sum up to one. For instance, if monetary analysis states that rates are to remain for a while in a narrow corridor, we would put almost all probability mass on such scenarios

and almost exclude explosive or negative rate scenarios in an ORSA exercise. This goes into the lines set by Avellaneda et al. (2001) and weighted Monte Carlo approaches. The drawback of flexibility is that management risk appetite and discretion regarding risk premiums really makes it an own (subjective) risk assessment.

Another matter of concern regarding the coherence of the best estimate approach relates to the interplay between financial risks and biometric risks (mortality, longevity, dependence). As for the latter, it is often assumed that the law of large numbers applies at contract level. Let us try to be more specific regarding the previous assertion. Best estimates of insurance liabilities involve a market-consistent derivation as for financial risks and statistical approaches, say mortality tables for biometric risks. When it comes to the computation of life-contingent liabilities, an expectation is involved, thus everything looks fine with the best estimate approach. On the other hand, when looking at the more granular level of a single life insurance contract, there is some kind of discreteness in the payments, death is a zero-one event and an individual would not die gradually according to the certified relevant mortality table. Thus, when it comes to hedges, so that the best estimate does not remain a concept but could be translated in today's cash, things become blurred. One might think of using life reinsurance, balance guaranteed swaps or other forms of securitization and risk transfers, but actually such markets are still in limbo. Besides, this is the reason why insurance companies are asked for an extra risk-margin since the only credible transfer is some other insurance company taking over the commitments to the insured. We may also notice that such issues are at the core of the validity of matching adjustments, i.e. whether liability cash-flows can be well predicted.

Clearly, a proper assessment of the reliability of life tables is a legitimate concern when it comes to the computation of best estimates of insurance liabilities. Technical issues at hand are well-documented and will not be further discussed in this chapter.

One of the biggest concerns regarding reserves is the plurality of reference frames: prudential such as in Solvency II, financial when it comes to MCEV and accounting as with IFRS 4 Phase 2.[7] For instance, even though under IFRS, the cost of capital approach is a valid option for risk-margin computations, which should, in principle, be consistent with the Solvency II framework, principles to account for diversification or reinsurance benefits may differ. Consequently, models need to be rerun under different assumptions, obscuring the outputs. Regarding the key issue of discount rates, the scope of volatility adjustments might differ under the two metrics. As can been seen from this example, model multiplicity often arises from compliance constraints.

---

[7]Academic literature on evolving insurance accounting standards is scarce. We can mention Dal Moro et al. (2014), though they focus on reinsurance rather than the life insurance business. However, easy-to-access professional documentations and discussion papers by consultancies are widely published on the web.

While the internal model approaches to solvency go along the previous lines, standard formulas are often remotely related to pricing and risk theory. We recall that, on the academic side, there is now a well-established theory of risk measures. As for the now classical arbitrage-free pricing theory, this theory relies heavily on probability tools. When it comes to applications, internal models can be implemented thanks to Monte Carlo (in most cases for insurance) or through filtered historical simulation (in most cases for market risks over short-time horizons). In the standard formula of Solvency II, coupling risk measures for risk classes as if insurance risks could be modelled by multivariate Gaussian distributions with regulatory prescribed correlation parameters is a clear sign of that departure from standard statistical approaches that underlie internal models. A similar trend is at work in the banking world, either regarding the Basel III capital charge on securitizations[8] or market and default risk treatments in the standard formulas of the FRTB (Fundamental Review of the Trading Book).[9] In a number of cases, as securitization in the trading book, standard approaches will become compulsory and in other cases standard formulas will be involved in floors to the outputs of internal models. Thus, the lack of risk sensitivity could become an issue. It may lead to improper capital allocations and misestimating diversification benefits.

Given the chosen horizon (1Y for Solvency II, default risk in Basel II and III) and the stated confidence level (99.5 % in the case of Solvency II) standard formulas cannot be formally back-tested (see the Chapter "Ex-ante model validation and back-testing" by Loisel and Nisipasu). On the other hand, standard formulas can be calibrated thanks to QIS so that they would still provide sensible results for actual lines of business. Standard formulas are meta-models of risk measures.[10] As such, they need regular patches and updates when market environment changes: EIOPA issued a document entitled "The underlying assumptions in the standard formula for the Solvency Capital Requirement calculation" in July 2014, and some notes on the latest standard formula calibration in September 2014.[11] These documents provide some important insights into the interest rate risk capital charge. It was calibrated on data prior to 2009, thus not accounting for the Eurozone 2011–2012 crisis or for the subsequent credit and quantitative easing. It was based on relative changes of market rates, including 30 years of maturity interest rate swaps. The current version of the standard formula postulates a floor on the absolute positive changes in rates of 1 %. Besides the trickiness of the approach, this calls for two remarks:

---

[8] http://www.bis.org/bcbs/publ/d303.pdf.

[9] http://www.bis.org/publ/bcbs265.pdf.

[10] We refer to Devineau and Loisel (2009) for a discussion of the interplay between Solvency II standard formulas and internal models. Using the standard approach would clearly simplify computation of risk margins as compared with using internal risk models.

[11] https://eiopa.europa.eu/Publications/Standards/EIOPA-14-322_Underlying_Assumptions.pdf and http://fr.slideshare.net/andrewcoffey1/notes-on-the-latest-standard-formula-calibration.

- There is a clear consistency issue between the assumption of a constant, non-market based ultimate forward rate and the assumption that the 90-year rate might change in a ±80 bps range at a 1-year horizon.
- Since regulators needed to accommodate to evolving market environments and close to zero or negative rates, some adjustments were deemed necessary. Consequently, it is wise to question the calibration of the standard formula in an ORSA framework. This is the typical recalibration issue when using meta-models. They may behave understandably at some point in time, once properly calibrated thanks to QIS exercises, but may fail to behave properly in a dynamic risk assessment framework. The major risk and modelling issue in the long term might actually be how the alpha term (driving the convergence to the ultimate forward rate) and the level of that ultimate forward rate are to be monitored…

In October 2013, the Basel Committee issued a discussion paper entitled "*The regulatory framework: balancing risk sensitivity, simplicity and comparability*".[12] The motivation was largely due to increasing defiance on internal models, especially the divergence in RWA (Risk Weighted Assets) in banking and in the trading book. This could be equally of interest to insurers.

Standard formulas aim at being simple, though keeping a reasonable degree of risk sensitivity. Robustness, especially regarding correlation parameters or the frequency of extreme risks, is often a key issue. Since estimates of correlation parameters or default frequencies could be highly unstable, there is a global trend towards the use of regulatory prescribed parameters. This is obviously the case with standard approaches.

As for internal models, it is likely that modelling choices will be more and more constrained: An internal model that involves a large number of risk factors and deals with risk at a low level of granularity is likely to be more risk-sensitive but less robust, i.e. more prone to a number of arbitrary technical choices, and more difficult to run and manage (model governance). For that purpose, the Basel Committee compels a number of technical choices regarding internal models such as the maximum number of factors (two in the case of default risk within the trading book). Also, granularity and materiality of risk factors are closely monitored. Thus, on one hand, the estimation of correlation matrices (involved in diversification benefits) is constrained but, on the other, bank regulators are concerned about material risks under the radar, typically sovereign default and credit risks when it comes to European entities.

Standard formulas (should) give little room to interpretation and thus allow for greater comparability across regulated entities. In the banking world, disclosure of standard capital charges will become mandatory even for institutions using internal models. Moreover, too large departures between the two approaches will be more closely monitored and could lead to floors. Consequently, standard capital charges are part of the benchmarking of risk models (including RCAP, computation of

[12] http://www.bis.org/publ/bcbs258.pdf.

RWA on hypothetical portfolios[13] and benchmarking exercises conducted by EBA[14]).

Due to the above, and since standard capital charges for market risks are to be considered as a credible alternative to internal models, the Basel Committee has recently come to much more granular market risk models (SBA, Sensitivity Based Approach[15]). As in Solvency II, shocks on risk factors and correlation parameters are given and aggregation across risk classes is carried out using a quadratic formula. This scenario-based approach is in the same vein as the CME margin model, SPAN,[16] and departs from the filtered historical simulation approaches (and to a much lesser extent to Monte Carlo) that prevail within financial institutions. The challenge is not to fall under undue complexity, for instance an excessive number of arbitrary inputs, when specifying standard formulas.

Overall, the juxtaposition of risk models built under different premises makes model governance trickier. In the insurance business, we are still at the dawn of this evolutionary process. Compared to the banking sector, regulatory solvency requirements are currently less binding than in the banking world where financial institutions need to build up huge amount of capital to meet the new regulatory constraints. On the other hand, the insurance industry is much more concerned by long-term solvency issues. Comparability and simplicity of risk models are thus clear issues: Currently EIOPA ORSA guidelines stay at a qualitative level and focus on organizational principles.[17] It might well be that modelling choices will eventually be restricted in the curse of on-going supervisory processes.[18]

9.3 Constraint-Driven Modelling

We advocate that model multiplicity (and potential inconsistencies) arises from pragmatism on one hand and from diverging compliance constraints on the other.

Pragmatism implies looking inside the modeller's toolbox for the most suitable way to solve specific issues. Modellers could actually switch from one framework

---

[13]See http://www.bis.org/publ/bcbs267.pdf and http://www.bis.org/publ/bcbs256.pdf.

[14]https://www.eba.europa.eu/-/eba-consults-on-technical-standards-on-supervisory-benchmarking-of-internal-approaches-for-calculating-capital-requirements.

[15]See http://www.bis.org/publ/bcbs265.pdf, http://www.bis.org/bcbs/publ/d305.pdf and http://www.bis.org/bcbs/qis/biiiimplmoninstr_feb15.pdf.

[16]http://www.cmegroup.com/clearing/span-methodology.html.

[17]https://eiopa.europa.eu/Publications/Consultations/EIOPA-BoS-14-259_Final%20report_ORSA.pdf.

[18]To quote Eling et al. (2007), "We anticipate that the models with greatest predictive power will be highly complex, likely including some aspects of dynamic cash-flow. Complexity itself, however, does not guarantee a good model. Also, even if the model is reasonably successful at identifying financially weak companies, such ability does not necessarily justify its costs. Complexity tends to require more data and results in higher costs to develop and maintain the resulting system, for both the insurers and the regulators".

to another, introducing surrogate models for computational ease. To take an illustration, one might think it would be convenient to simulate rates through some AR(1) mean reverting Gaussian process up to the typical 1 year horizon and then compute swaptions or caps intended to hedge convexity risks through a BGM or Black type approach relying on log-normal forward Libor or swap rates. Thus, in many instances, models rely upon different premises. When it comes to probability models, quite often a non-nested, non-meaningful encompassing model is provided which obviously hinders comparisons.

Keeping a minimal degree of overall consistency involves mapping procedures and frequent recalibration of models (one on another, onto market prices). It is more an art than a science. Strong and reliable modelling skills are highly recommended to deal with a large number of interconnected parameters and to keep the business on track. Moreover, mapping procedures may not be well suited when it comes to prospective assessment of risk and dynamical approaches of balance sheet items.

Consequently, there is a strong propensity to rely on market practices and well-recognized professional standards within a given business activity, the academic doxa being only in the background.[19] Within insurance companies, key functions in the system of governance (second pillar of Solvency II) will be involved in setting up standards regarding the management of models. In Europe, EIOPA together with national supervisors is also to be strongly involved. This will parallel the expanding encompassing benchmarking exercises set up in the banking sector, as was the case with the AQR[20] for the Eurozone banking sector: The AQR was a unique opportunity for a comprehensive review of all pricing models. Major Eurozone banks were required to accurately document their modelling assumptions in a standard format. Nowadays, any supervised entity would need to justify departures from commonly accepted modelling choices. To enforce financial stability, EBA has called for prudent valuation adjustments of fair valued positions. These reserves are deducted from the numerator of capital ratios. Among these adjustments, one is entitled "model risk AVA" (Additional Valuation Adjustment).[21] It is noteworthy that as long as there is no divergence among market

[19]In the US, a number of cases regarding the inconsistency of active management practices, thus not conforming to the efficient market hypothesis, went to court. The evidence is mixed. Depending on the context, US courts can presume market efficiency and reject a presumption of prudence. To prevail against the efficient-market defense, participants would have to show active management can triumph over market averages. See http://www.pionline.com/article/20140707/PRINT/307079997/court-backs-efficient-market. It is unclear whether we could face similar issues regarding the computation of reserves in a European context.

[20]https://www.eba.europa.eu/regulation-and-policy/market-risk/draft-regulatory-technical-standards-on-prudent-valuation https://www.ecb.europa.eu/pub/pdf/other/assetqualityreviewphase2manual201403en.pdf?e8cc41ce0e4ee40222cbe148574e4af7.

[21]https://www.eba.europa.eu/documents/10180/642449/EBA-RTS-2014-06+RTS+on+Prudent+Valuation.pdf.

participants regarding valuation models, there is no need to make an additional valuation adjustment for model risk. This is a strong incentive for converging modelling approaches and the actuarial community is to play a key role for that purpose.

9.4 The Recalibration Puzzle

In many cases, constraint-driven modelling approaches will need to be closely monitored for changing market environments and models might need to be recalibrated to one another, to remain consistent at every point in time.

To stay in line with the above illustrative example, this would mean calibrating swaptions or cap volatilities depending on the simulated level of rates at a 1-year horizon. Such a mapping procedure of log-normal volatilities onto, say, the level and slope of interest rates can be achieved by different means, thus we introduce some hidden modelling complexity, with a number of more or less arbitrary and uncontrolled assumptions, while the starting point was ease of computation. It is also worth noting that the impact on interest rate sensitivities and thus ALM policies are likely to be one order of magnitude beyond the best estimate prices.

One of the main issues in the corporate CDO business has been the surge of bespoke CDOs where the underlying credit portfolio was not a standard CDS index, such as iTraxx Europe or CDX NA IG, for which one could easily access liquid market quotes and then rather easily infer a structure of implied base correlations. As for illiquid bespoke CDOs, a derivation of the required bespoke correlations has been made thanks to a number of mapping procedures from the above implied correlations. Thus the marking of bespoke CDO tranches involved a blend of market data and modelling choices, difficult to control, due to the illiquidity of the bespoke tranche. As for most insurance liabilities, even though a number of inputs are calibrated to reliable market quotes, a high degree of model uncertainty remains (see the chapter on model risk).

On the other hand, model governance process will make it quite difficult to update models when required. Let us go back to our favourite story of normal/log-normal rates. In early 2015 the negative rates region expanded. Not only overnight but also 1-month interbank rates became negative, not to mention core sovereign bond yields. Of course, if one would input a negative-forward Libor rate in a Black formula, this would result in a disruption of the computation… This might be sorted out by, say, using a shifted log-normal model at the price of a recalibration of volatilities. Changes need to be documented, reported and approved. When it comes to supervisory reporting and approval, this might lead to some kind of suspicion regarding modelling choices in the first place. Overall, this leads to what could be called sticky modelling choices.

Regarding risk-models under standard formulas, we already advocated that they might need to be recalibrated under changing market environments.[22] The updating process is at regulatory discretion. Clear operational constraints imply that this will remain in the dark, even though this could undermine the outcome of an ORSA exercise.

Ultimately, determining the extent to which the use of different models in different places does not break the coherence of the global pricing or risk management framework is a matter of human expertise and thus is liable to operational risks. One question to be asked is whether every person involved is fully aware and able to assess the consequences of making different models live together, recalibration issues, and an inconsistent range of possible values. There is a legitimate concern that the experts involved in different departments, actuaries, ALM, financial engineers, communicate appropriately and report subsequently to key functions, including AMSB. Everyone should keep in mind the Columbia Shuttle disaster and the miscommunication of key points from the bottom up to decision level.[23] Conversely, top management financial expertise needs to be set at the right level to prevent any malfunctioning and so that black boxes do not run the business (see the chapter on the role of models in management decision-making by Renaud Dumora and Bernard Bolle-Reddat within this book).

9.5 Fitting Models to Models

At this stage, having discussed a number of approaches and issues regarding pricing and risk models, it is worth distinguishing between two approaches to models in insurance or banking businesses:

- Models could be closely related to probability and statistical concepts.
- Models could rather be seen as a processing engine fed by some inputs and intended to produce meaningful outputs. This second view takes it roots from computer science, engineering or systems theory.

[22]The low yield exercise conducted within 2014 EIOPA stress tests provides some interesting insights with respect to the recalibration of key parameters under extreme but plausible scenarios: On page 11 of https://eiopa.europa.eu/Publications/Surveys/LIR%20Stock%20taking%20exercise%202014.pdf it is stated that "several NSAs reported changes in the valuation approach for technical provisions in the last 2–3 years, of which two introduced the ultimate forward rate (UFR) method while another introduced an optional and temporary floor for the discount rate to calculate the technical provisions. These measures are primarily introduced to give some relief to insurance undertakings in a low yield environment".

[23]"As information gets passed up an organization hierarchy, from people who do analysis to mid-level managers to high-level leadership, key explanations and supporting information are filtered out. In this context, it is easy to understand how a senior manager might read this PowerPoint slide and not realize that it addresses a life-threatening situation." Report of Columbia Accident Investigation Board, Volume 1, page 191, http://www.nasa.gov/columbia/home/CAIB_Vol1.html.

When it comes to approximation and the construction of simplified models, a statistician will think in terms of distance between probability cdfs under a suitable metric. The approximate model may be nested within the original or could be related to another probability model under which computations are made easier,[24] for instance leading to closed-form expressions for insurance liabilities. Also, it is likely that a well-educated modeller will think of prices in terms of expectations and of VaR as a quantile of a loss distribution to be determined. In the same vein, approximation of cash flows will involve, say, deviation from the mean and the variance of the error (difference between true and approximate cash-flows).

Under the second approach, the approximate model does not need to be related to what sounds like a properly specified standard academic pricing or risk model. In the theoretical mathematical finance framework, discount bond prices of different maturities are computed as the risk-neutral expectation of a stochastic discount factor. Many approaches developed for yield curve interpolation, such as splines, were not constructed with a dynamic arbitrage-free interest rate model in mind:

- As for the Nelson-Siegel framework, a static parametric interpolation scheme might eventually be made consistent with the above dynamic setting; but this was not the prime modelling purpose.
- A Vasicek formula could be used for yield curve reconstruction. It involves volatility and mean-reversion parameters. When fitting the observed rates, say, thanks to a least square best fit, we may end up with values grossly inconsistent with swaption prices. Here, we are faced with a model with different parameter inputs depending on the use, computation of bond prices or computation of swaptions. The Vasicek model is then viewed as a proxy of a better but difficult to handle more general interest rate model. Actually, under the objective-based inference approach (see Gouriéroux and Laurent 1996), different parameters may be required when using a misspecified model for different purposes.

In banking regulation, there is now a long story of fitting simple and convenient formulas to more complex models. One can think of the celebrated "maturity adjustment" in the computation of Basel II default risk weights. This is to account for the fact that the maturity of loans is greater than the 1 year prescribed horizon. Thus, non-defaulted loans could suffer from downgrades or credit-spread risk,

---

[24] A stochastic volatility model, thus implying several risk factors, possibly with path dependence, can be approximated by a local volatility model, thus with only one risk factor (underlying asset) and a Markovian dynamics. This is related to the so-called Markovian projection technique (see Piterbarg 2006). Under the approximate model, all call and put options are priced accordingly to the original model. The two models will only be used when it comes to pricing path-dependent options or when considering risk-management issues. This is a typical example where the two models being considered are built under the standard mathematical finance framework. In the interest rate risk context, we refer to Andersen and Piterbarg (2010) for a review of cross-calibration of options prices, with the purpose of dealing with log-normal rates and Black-Scholes type formulas.

which is not explicitly taken into account in the banking book.[25] Different regulatory treatments regarding default and credit spread risk provide huge incentives regarding location of risk and are therefore a clear concern to life insurers.

Another important related issue is the computation of capital charges for default risk of securitization tranches.[26] The so-called SSFA (Simplified Supervisory Formula Approach) is at the top of the hierarchy, i.e. the one to be used by major financial intermediaries. Given that securitized tranches were deeply involved in the 2008 subprime crisis, such highly technical issues have huge implications regarding the ability to maintain on-balance sheet positions within banks or whether they should be held by end-investors including life insurance investors. Not entering into undue technical details, the capital charge depends on the seniority of the tranche and involves an exponential decay with a prescribed regulatory parameter driving the decay. This simple parametric model is first fitted to another more sophisticated model such as Gordy and Jones (2003) and then the final calibration involves QIS to assess the amount of extra capital required.[27] Here, we are typically in a framework where a simpler model is calibrated to a more complex one. It is rather the ease of use (implementation, monitoring of key parameters by supervisors) that drives the specification of the simpler model, rather than an economic assessment of risks.[28] Beyond providing comparisons between insurance and banking standard approaches to default risks, it illustrates a regulatory trend.

Whether what we currently see in the banking industry and the building of regulatory risk models is applicable to the life insurance industry is an interesting but difficult-to-answer question: A number of Solvency II issues are still on the agenda, not to mention the conduct of stress-test, ORSA exercises or systemic risk

---

[25] See pages 10, 11 of "An Explanatory Note on the Basel II IRB Risk Weight Functions" (2005), http://www.bis.org/bcbs/irbriskweight.pdf: "Maturity adjustments are the ratios of each of these VaR figures to the VaR of a "standard" maturity, which was set at 2.5 years, for each maturity and each rating grade. (…). In order to derive the Basel maturity adjustment function, the grid of relative VaR figures (in relation to 2.5 years maturity) was smoothed by a statistical regression model. (…) The regression formula for the maturity adjustments in the Third Consultative Paper is different from the one in the Revised Framework of June 2004." Thus, that adjustment is related to the ratio of 1 year VaR for a 1-year maturity loan to the 1-year VaR of a 2.5 years maturity loan. This ratio is computed under a structural credit risk model and then regressed onto the logarithms of the default probabilities. One interesting issue regarding this important regulatory feature is the (in)ability to monitor the required recalibrations over a long period of time: It may be that the experts involved first are no longer in place years after implementation of the proxy model, and that details regarding the purpose of the approach and implementation practicalities are lost.

[26] http://www.bis.org/bcbs/publ/d303.pdf.

[27] See http://www.bis.org/publ/bcbs_wp22.pdf for details regarding the calibration.

[28] The standard approach to counterparty risk of derivative exposures (SA-CCR, http://www.bis.org/publ/bcbs279.pdf) provides some further examples of calibration of models onto models. This is documented in a companion Basel Committee paper: Foundations of the standardized approach for measuring counterparty credit risk exposures http://www.bis.org/publ/bcbs_wp26.pdf. Besides being the standardized capital charge for counterparty credit risk, the model is likely to be used in the now-celebrated leverage ratio. It will thus drive the ability of investment banks to hedge financial risks within life insurance companies via off balance sheet instruments.

assessments. EIOPA stress tests, although based on a few meaningful unpleasant but plausible scenarios, could provide some alternative or complement ORSA with stochastic scenario generation. Insurance stress tests are thus quite useful to benchmark more sophisticated ORSA models: Easy-to-grasp simplified models can be viewed as sanity checks.

9.6 Model Parameters Set at Managerial Discretion

As discussed previously, multiple models are being used, sometimes with the intent of speeding up computations, sometimes models may have different purposes, say valuation or risk management. Besides consistency issues, it is clear that modelling a number of important quantities such rate of appreciation or the surplus participation rate on outstanding contracts is daring. Thus, it seems wise to directly manage such quantities. Instead of being a model output, it becomes an input, set at managerial discretion. At least, it makes it unequivocal that we will not rely on a black box and we can assert that the managed parameter has a clear and understandable meaning. Moreover, it could lead to further simplifications in the assessment of life insurance liabilities. If, say, the surplus participation rate depends on the previous rate of returns on assets, we might be faced with a complex path-dependent payoff. Introducing a constant parameter could dramatically ease computations. Let us briefly investigate the potential drawbacks of such a tempting shortcut:

- While the above approach would certainly be meaningful when the number of such key parameters is small, we could be faced with managing a large number of such parameters. Then, to stay in line with the above example, we would need to introduce some connections among parameters. Say, participation rate would be higher in countries where competition among insurers is tougher. Then complexity is back, we now need a model to relate the parameters.
- Regarding compliance and supervisory approval, it puts more pressure on the key functions. Are the chosen parameters set appropriately? Can we rely on expert opinion? What is the decision process regarding updating?

9.7 Approximation Issues for Pricing Models in the Finance and Insurance Contexts

In this section, we will first recall well-known concepts and approaches of approximation in a general financial context and we will then deal with insurance specific topics. Even though it is dedicated to life insurance modelling issues and we will try to keep technicalities to a minimum, it is a bit more academic by nature.

It will also provide a flavour of the prevailing dualism regarding modelling approaches: in short, KISS as illustrated in the previous section versus rocket science.

Fortunately enough, most life insurance liabilities with embedded optionality do not involve complex payoffs, such as a large number of risk factors and a blend of long/short risks that exacerbate correlation modelling issues and for which approximation methods usually collapse due to the curse of dimensionality. In most cases, insurance liability payoffs involve a smooth degree of path dependence and risks can be adequately captured by the level of current rates.[29] As a consequence, it may well be that the simplified version of the model grasps the essential features of the original.

There are a number of techniques developed by finance quants and subsequently adapted and expanded to cope with life insurance specificities, among which are approximation of cash-flows, the idea behind the replicating portfolios, or approximation of pricing formulas, such as LSMC (Least Square Monte Carlo). We also refer to Planchet and Robert's chapter in this book (From internal to ORSA models) regarding the use of closed-form formulas computed under simplified model assumptions.

The replicating portfolio technique consists in approximating a given payoff by projecting it (i.e. minimizing some suitable distance) on a set of base portfolios/payoffs/risk profiles for which prices are either directly inferred from market observables or easily derived from the pricing model. A linear combination of the base payoffs provides the approximation of the more complex payoff. Then, the approximated price is simply the price of the approximating portfolio. Let us recall that best estimate computation of reserves is associated with a linear pricing rule. Thus, the approximated price is a linear combination of the prices of the base payoffs with the same coefficients as those involved in the approximation of the payoff to be priced.

When the portfolio to be priced can be perfectly replicated, the price of the replicating portfolio is the true price, i.e. there is no pricing error, i.e. no discrepancy between surrogate and true pricing. This idea dates back to the pioneering work of Breeden and Liztenberger (1978). They computed the pricing density for a fixed-time horizon from call option prices with the same maturity. This was further extended to a multi-period framework by Dupire (1994). In the same vein as Breeden and Liztenberger (1978) or Derman et al. (1995), Carr and Madan (2001) provided a strikingly simple result directly relating the payoff $f(S_T)$ of a complex payoff/risk

[29]Day-one shocks on interest rates as considered in a number of stress-tests clearly do not involve complex dynamics. For a number of life insurance companies, a long-lasting period of extremely low rates ("Japanese style"), followed by a taper tantrum, as experienced in the 2013 US market (delayed "inverse scenario" following EIOPA terminology) would be damaging, as it would be likely to be associated with a sharp increase in bond spreads in the periphery (both on sovereigns and corporates). This shows that path-dependency cannot be formally ruled out. We refer to EIOPA 2014 stress tests for further discussion:

https://eiopa.europa.eu/Publications/Surveys/Stress%20Test%20Report%202014.pdf.

profile to standard call option payoffs $(S_T - K)^+$, where S_T stands for the price of the underlying at option maturity and K the call optionstrike:

$$f(S_T) = f(0) + f'(0) \times S_T + \int_0^\infty f''(K)(S_T - K)^+ dK.$$

Consequently, the price of $f(S_T)$ at time zero (we assume for simplicity that the underlying asset does not pay any dividend) is given by:

$$\text{Price}(f) = f(0)e^{-rT} + f'(0) \times S_0 + \int_0^\infty f''(K)C(K,T)dK,$$

where r stands for the default-free short term rate and $C(K, T)$ for the price of a call option with strike K and maturity T. This idea of using the prices of simple options to compute trickier ones, such as those involved in insurance liabilities (capital guarantees, triggered returns and redeemable features), dates back to Ross (1978). Pelsser (2003) provides an application in an insurance context.

An appealing feature of this approach is the direct connection between prices of complex liabilities and market prices of more standard traded options. The prices of the latter would be obtained directly for the market, leading to a model-free valuation. This bypasses using the notion of rather abstract pricing densities and risk-neutral probability, as mentioned earlier for digital caps.

However, as can be seen from the previous formula, the approach requires knowing call option prices for all strikes. The set of observable liquid option prices is quite small and some form of interpolation and, more importantly, arbitrary extrapolation technique of option prices is required to get a complete set of options prices. It is well known that rather standard interest rate derivatives such as CMS (constant maturity swaps) are quite sensitive to the chosen extrapolation scheme.[30] CMS are involved when coupon rates are indexed on long-term rates, making it possible to deal with exposures to the change of the slope of the yield curve.

Also, while the Carr and Madan (2001) decomposition formula is straightforward when dealing with a single risk factor, it becomes trickier in higher dimensions. Bakshi and Madan (2000), Nachman (1988), Ross (1976), Zhang (1998) argue that correlation derivatives are required to decompose hybrid payoffs. However, since such complex options are not routinely traded, we are faced with a severe limitation of the approach. Going back to simplicity needs to project insurance liability payoffs onto, say, the current (at the same time the payoff is being paid) level of the short-term interest rates.

[30] We refer to Andersen and Piterbarg (2010) for the static replication of CMS rates thanks to swaptions. The approach dates back to Amblard and Lebuchoux (2000).

This is more or less the idea behind the replicating portfolio approach: It involves an approximation of the payoff through a linear combination of base payoffs that could be easily priced or calibrated to market quotes. As mentioned earlier, the practical scope seems rather limited but well-suited to most life insurance liabilities and embedded interest rate optionality. Among actuarial studies dedicated to this approach, we mention Boekel et al. (2009), Botvinnik et al. (2014), Natolski and Werner (2014), Oechslin et al. (2007) and Schrager (2008).

Among the issues at hand, when dealing with replicating portfolios, we could consider, say, Hermite polynomials instead of call option payoffs. This can make sense if the payoff to be approximated is a smooth function. Whenever this payoff depends upon a Gaussian variable, such as the short rate in the Vasicek model, Hermite polynomials appear to be a reasonable choice: It is known that these polynomials are orthogonal under the Gaussian measure and form a basis of the space of all potential payoffs; among which of course, the insurance liabilities to be evaluated should rely. The theory that underpins the approach is not part of the core syllabus of actuarial studies but, nevertheless, linear operators are quite standard and well known to mathematicians. Payoffs can be computed through a series expansion that is truncated to provide an easy-to-price approximation. While everything is neat from a mathematical perspective, practical implementation needs to address the following items:

- What is the most suitable choice of polynomials? There are a number of competitors to the Hermite family mentioned above.
- What is the correct level of truncation, i.e. the finite number of terms in the expansion?
- Can we control the approximation error, without of course having to compute the true price? In other words, could we provide ex-ante bounds to the difference between the approximation and the true prices.

As mentioned earlier, in today's applied finance and insurance modelling, pragmatism heads theoretical concerns. Models are implemented first and it is only in a second stage that their empirical performance is assessed. We refer to Li (2014) and Beutner et al. (2015) for an investigation of the above issues.

A dual approach, initiated in the finance world by Longstaff and Schwartz (2001), known as LSMC or Least Squares Monte Carlo, involves an approximation of the prices rather than of the payoffs.[31] We recall that computation of capital requirements, based on internal models, involves evaluating liabilities for every simulation node, which is computationally intensive. Thus, some form of interpolation from a number of well-sampled forward values of liabilities is required. This can be achieved through functional regression and kernel based estimation. While the approach performs quite well under one-factor interest rate models and when payoffs are not path-dependent, computational performance collapses in

[31]See also Stentoft (2004), Glasserman and Yu (2004).

higher dimensions: Thus, the need to project a one-factor interest rate model[32] and the payoff onto the factor, whenever the payoff exhibits some form of path-dependence.[33] We refer to Bauer et al. (2010) for applications to the life insurance context.

A general approach to the approximation issue relies on the analysis of the pricing operator.[34] In that framework, Darolles and Laurent (2000) look for optimal approximations of payoffs and pricing formulas and show the duality in the two approaches. In the case of interest rate models, they deal with the case of Vasicek type models (mean-reversion of the short rate), where the Hermite polynomials are actually optimal and the case of a Brownian motion with reflecting barriers (corridor type dynamics). While this kind of stationary dynamics is suitable for approximation, it is also shown that the standard Black-Scholes dynamics does not comply with the technical requirements for the approximation methodologies to behave properly. We also refer to Pelsser and Schweizer (2015) who compare LSMC and replicating portfolio techniques in an insurance context.

Rather than trying to approximate payoffs, as in the replicating portfolio approach, or pricing formulas as in LSMC, a simpler model could be considered from scratch. This is standard in the financial world. We already mentioned the log-normal approximations of forward swap rates in a BGM setting (see Andersen and Piterbarg 2010). The computation of CVA (Credit Valuation Adjustments) for portfolios of complex interest rate derivatives within trading books of investment banks follows the same lines. CVAs are required to account for counterparty credit risk, especially in the case of uncollateralized trades. Under the standard set-up, the valuation adjustment involves the expectation of the positive exposure (EPE). Positive exposure is the maximum of zero and the present value of the considered book. Stated slightly differently, we need to compute a call option price (with zero strike) on a book of interest rate derivatives, which can include tricky options. This computation is usually done under a much simpler model, say, of Vasicek-type, than the official internal model used to price exotic interest rate derivatives. Bonnin et al. (2014) have been quite successfully using this approach: they could dramatically speed up ORSA computations thanks to the use of closed formulas for a best estimate of savings contracts under a proxy model.

[32]The Markovian projection technique is theoretically appealing but leads to non-linear interest rate dynamics. On the other hand, the use of easy to deal with one-factor interest rate models (say of Vasicek type) implies dealing with recalibration at node points.

[33]In a banking context, Adam et al. (2009) consider the pricing and hedging of non-maturing deposits. Depending on the return on money market funds, deposits can be redeemed. Dynamic hedging of interest rate risks leads to some path-dependence in the optimal hedging portfolio. Approximation of the corresponding payoff by means of computing conditional expectation of the payoff on current Libor rates is investigated. This leads to a replicating portfolio based on caps and floors and eliminates the path-dependency issue at the cost of a loss of accuracy.

[34]See Aït-Sahalia et al. (2008) and the references therein for a review of methods and issues when using linear pricing operators in a Markovian setting. Hansen and Jagannathan (1997) also provide some background regarding approximation methods and pricing kernels.

Up-to-now, we have more or less discussed numerical issues. Still focusing on embedded optionality in life insurance contracts, we need to assess the relevance of the ESG and whether management actions or customer behaviour is adequately described. The ability to efficiently monitor participation benefits' features and surrender options are of primary importance to mitigate the short-term negative effects of rate increases on fixed income investments. These are partly model-based and involve a mix of customer behaviour and of management actions. Commitment from management to follow the actions as set up in the model is not formal. Depending on difficult to quantify parameters, such as the regulatory environment at the time decisions need to be made and the intensity of competition among insurers, it may be optimal to deviate from the originally scheduled actions. Ultimately, such models need to be articulated with the dynamics of default-free rates to provide trustworthy best estimates of liabilities and associated duration measures. Clearly, there is much more model risk as we move away from the field of actively traded fixed income products.[35]

We have a poor view of the magnitude of the errors introduced at the various stages of the computation. This section illustrates the technical complexities involved and the bottom-up communication issues mentioned previously concerning the Columbia Shuttle disaster. Regarding business processes and operational risks, we need to rely on experts to assess the reliability of the approaches and the robustness of the outputs. This is likely to displease ASMB and regulators, but this is the way things are.

9.8 Conclusion

Regarding the development and the use of models within the life insurance business, the learning curve involves a trial and error process and, almost as usual in finance and insurance modelling, empiricism heads theory. Model multiplicity involves more or less ad hoc recalibration procedures. It is quite difficult to make ex-ante assessment of potential inconsistencies. From a management perspective, a strong reliance on the skills of actuaries and quantitative modellers is required. Operational risks could arise and are difficult to monitor through standard auditing procedures (see Christian Robert's chapter on the threat of model risk for insurance companies). Special emphasis should be put on validation standards, sanity checks and bottom-up information so that ASMB is aware of the key modelling and business challenges (see the chapter by Bolle-Reddat and Dumora within this book).

[35]Would it mean that we might drop our guard with respect to, say default-free rates? Unfortunately, understating volatility of long-maturity rates would just add to the modelling maze: Errors add-up.

Part IV
Models and Business Processes

Chapter 10
Model Feeding and Data Quality

Jean-Paul Felix, Nathalie Languillat and Amélie Mourens

Abstract As seen in previous chapters of this book, models assess, over a long term projection period, future cash flows in different scenarios with various types of data. Their results are used by top managers to make important decisions at the Company level. Obviously feeding those models is a big issue to be addressed. Moreover, as the saying goes “garbage in, garbage out”, meaning that the quality of results is directly related to the quality of data. So data quality is a vital subject for many top managers of insurance undertakings (life and non life as well as mutual), because they need to trust in risk models results and use them in their decision making process. The article covers all the major questions related to data quality that you may have. In particular, it explains why data quality should be considered as a process and not a commando-like operation, because there is no absolute level of quality and because after the targeted level of data quality is achieved, it has to be maintained at this level in a changing environment. Firstly the article focuses on data definition (contract and asset information, endogenous or exogenous parameters for example), existing standards, best practices that can be put in place by the Company, and on the data life cycle that need to be well understood and mastered by top management. Secondly, the article elaborates also the advantages of the launch of data quality projects, the building of clear data quality governance and of an optimal documentation. The article focuses on the importance of beginning on a well-defined and representative perimeter in order to experience the method, to achieve in a constrained period of time and after that, to increase the perimeter of data on which the process of quality has to be put in place. The article also considers existing solutions proposed by the market such as packaged data management solutions.

J.-P. Felix (✉) · N. Languillat · A. Mourens
BNP Paribas Cardif, Paris, France
e-mail: jean-paul.f.felix@bnpparibas.com

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_10

10.1 Introduction

As written in previous chapters, models are used both in strategic and risk decision-making process. In the short term, the frequency of models' intervention as well as the importance of their contribution will grow. The decision is generally made based on various indicators and stands at the far end of a quite long process beginning with data collection and models' feeding with data. As the saying goes "Garbage in, Garbage out", meaning that the relevance of calculated indicators the Management is provided with is conditioned to the quality of both models and input data. Besides the obvious own interest of insurers regarding Data Quality, the Solvency II regulation puts it as a first priority requirement, whereas previous regulations are quite silent about it. Accordingly, it generates a new source of pressure on insurers, on a topic that should not be this new for them. This pressure is derived from the high level of requirement of the Solvency II literature and is expressed by anticipated high workload, high cost, and short delays. Insurers are certainly asking themselves how they would tackle this challenging issue, all the more challenging that, most often than not, it should also be accompanied by a deep change in mind and in organizations.

This chapter dedicated to model feeding has the objective of giving a practitioner's overview of the main issues faced by insurers when implementing a Data Quality management process and of solutions they might implement.

10.2 Launching the Data Quality Project

10.2.1 The Reasons for Launching a Data Quality Project

The goal of a Data Quality program is to support an efficient data management process that leads to educated decisions. It is of high importance for insurers to really know their data: what it is, where it is stored, in which processes it is used, with which aim, at which granularity, its confidence level, etc. Based on this knowledge, the data management process could be designed to be as swift, smooth, safe (with low operational risks), accurate and stable in time as possible. Data is an asset for the company (Fig. 10.1).

Where does the Data Quality process begin and where does it end? Boundaries of Data Quality have to be defined.

For those insurers that dare not enter into further thoughts about data management, the risks are high of wrong strategic decisions (in force business management, acquisitions, launching of new businesses), wrong risk management decisions, supervisory sanctions (capital add-ons), and losing confidence of shareholders, partners and customers.

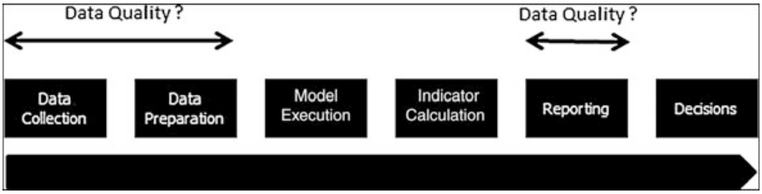

Fig. 10.1 Risks decisions process

10.2.2 The Sine Qua Non Involvements in the Data Quality Project

The recipe of a good Data Quality project requires four ingredients:

The commitment of all necessary teams to contribute: at minimum, the Users (Actuarial and ALM Depts.), the Models Dept. (if existing), and the IT Dept. should be involved. Most of the time, these resources should be completed with external expertise. The commitment of the Users is instrumental in the success of the project because through the years they have become experts on each data perimeter and as such they will offer a significant contribution, as we will see in the section "3-Data Quality: a loop approach". However, a critical issue is to manage the availability of these Users, which is normally under constraint, the Business as Usual deadlines often being in conflict with the Project's ones.

A budget potentially high: insurers try to optimize the management of their Data in term of costs and rewards. However the balance often tips in favour of costs as they may be quite huge, either the cost of changing the data systems or the cost of running them;

A multi-year perspective: "Rome was not built in a day", nor is your data management system;

A mind shift: The Company should recognize that data quality is of paramount importance which is not always the case: current data quality is often considered to be a fatality that is managed in silos without a global perspective. It has to be viewed as a business opportunity all the more with the emergence of Big Data. Therefore the first impulse of the project should come from the Top Management and a sponsor should be named and chosen among them.

10.2.3 Definition of the Ambition of the Project

As for any project, a clear definition of the ambition of the data project is crucial. Data is everywhere around us in the Company. Therefore the potential perimeter of the project is multi-dimensional—and it is hard to decide where its boundaries stand, what the level of the ambition is within, and what is the roadmap to achieve it (Fig. 10.2).

Modelled Lines of Business

The scope of the data to be reviewed is linked to the scope of the modelled lines of business, which still remains quite a large scope. A selection should be made in order to start improving Data Quality only on a more limited perimeter and/or to define a priority order. Its criteria are the level of complexity of the distribution channels (as it impacts the IT systems), the materiality of the business lines, their profitability and their underlying risks. These criteria are the same as those used when deciding whether or not to model the business lines.

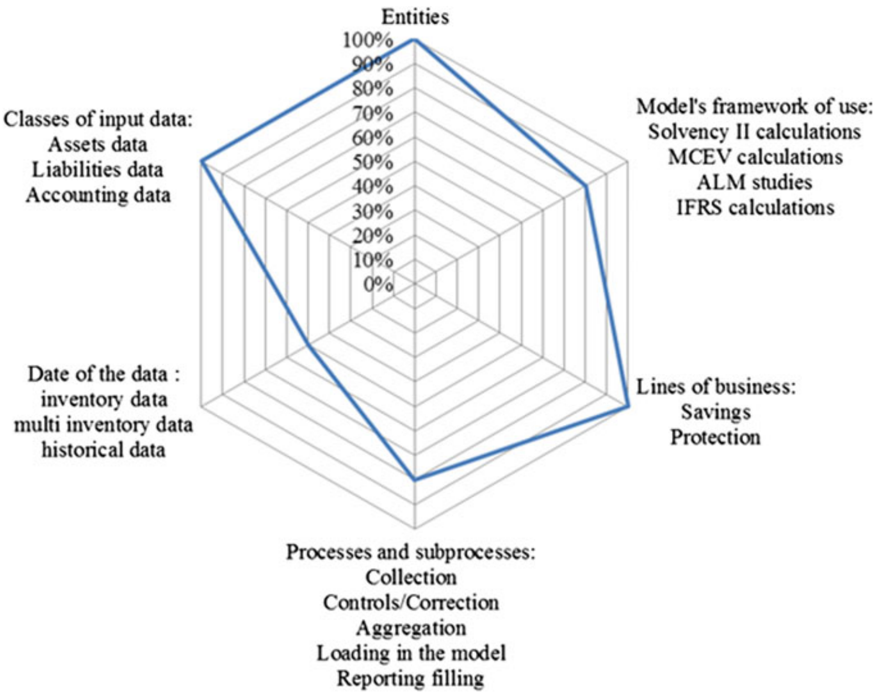

Fig. 10.2 Axes of the Data Quality project perimeter

Processes and Sub-processes

Data processing covers both data inputs and data outputs. However, as the processing principles are the same for inputs and outputs, this section focuses on inputs only. Indeed, the process and sub-processes' chart below illustrates the input data processing.

The mapping of the data processing and its different areas is necessary to decide the roadmap of the project. The part of the process (data collection, data enrichment and controls, data aggregation, data loading) and of the sub-perimeter(s) (accounting, assets, saving liabilities,…) in the scope of the project should be selected as there are many possible entry points (Fig. 10.3).

One could decide to review the Data Quality of a given sub-perimeter, for instance the accounting data processing, at each of its stages (collection, enrichment and control…). One could also decide to review the data management of a subset transversal stage, for instance the control phase for every sub-perimeter. In practice, the roadmap is made of different lots each of which covers one or several stages on one or several sub-perimeters, for instance, Data Collection, Enrichment and Control for the Savings contracts. The choice and the priority of the lots should be based on the level of automation, the level of documentation, the level of maturity of the stage, and the stability of the stage in the forthcoming years (in terms of IT solutions, in terms of functional process, etc.).

The purpose of the above-referred data processing's mapping is also to ensure that there are no uncovered areas and to challenge the current data process. In particular, one might decide that some stages of the data processing could usefully be made in the model's environment. The boundaries of the model are indeed a key issue of IT architecture that should ensure that there is neither duplication of tools nor any hole in the process.

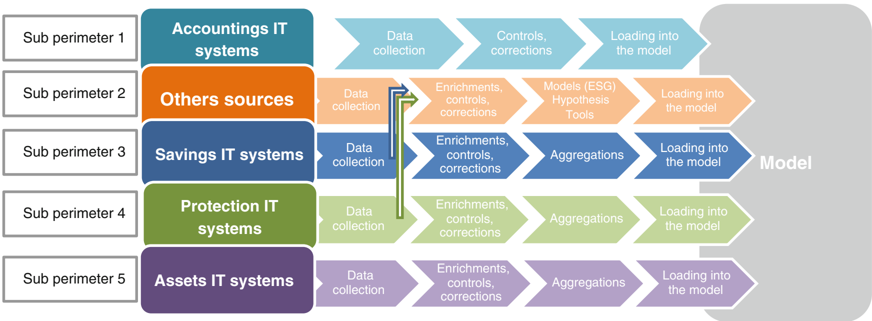

Fig. 10.3 Feeding of the model process

Categories of Input Data

The input data candidates for the data quality approach can be split into three categories:

Referential data, which is enterprise data usable in many different processes and not specific to the projections. For example: the database containing the names and codes of the distributing partners (Fig. 10.4).

Actual data gives a view of the insurer's situation (in particular its balance sheet) at a precise date. For example: the data describing its liabilities (number of contracts, amounts of reserves, policyholders' age, etc.) and its assets (nature and characteristics of the assets held line by line, amounts in balance sheet value/market value, etc.). A great part of this actual data is accounting data.

Projection data is necessary for the model to perform simulations such as financial (see Chapt. 4 "Economic Scenario Generators") and non-financial hypotheses (mortality/lapse laws etc.) and to model parameters (excluding referential data that is a category by itself).

In general, insurers apply by priority the quality approach to the data necessary to build model points which are instrumental to run the model. This data is mainly derived from accounting systems and actuarial IT systems and is mostly actual data.

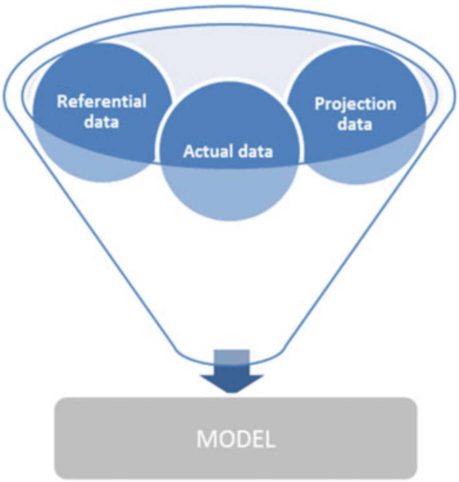

Fig. 10.4 Categories of input data

Date of the Data

Users manipulate all these categories of data under different forms as inventory data and historical data (in general, monthly or quarterly data gathered on a multi-year period).

The availability of historical data is critical for the definition of the projection hypothesis. For example, the building of lapses law or the calculation of Undertaking Specific Parameters within Solvency II Framework requires a deep data history.

Historical data is also used to compare the actual input data or results with those of the last exercise (for example, compare the portfolio of contracts at year end N with the portfolio at year end N-1).

In general, insurers give the priority to inventory data and adapt the approach for historical data (taking into account that many of them have been checked in the past). Once more, the decision is also driven by the cost of treating the associated volumes of data. As a consequence, there are clear benefits to begin as soon as possible to collect historical data in conditions that will ensure its quality.

More recently, internal constraints of financial disclosures and Solvency II alike impose very short reporting delays. This often led insurers to accelerate processes within which data is anticipated and extracted at some dates before inventory date (e.g. end of September). This method increases the number of historical data required and questions the scope of the quality review (at every date, at the end of September, at the end of December?).

Framework of Use

The framework of use of the model (Solvency II, MCEV, ALM, IFRS, etc.) should be considered before designing the quality approach because it also impacts the data preparation process and the type/format/granularity of data needed.

However, the data required in a given framework is not completely different from the data required in another one. There is some core data (for example: data on contracts or on assets) that is used by all frameworks and whose quality should be closely reviewed.

The retreatments of core data (aggregation, spread risk-neutral adjustment …) together with other additional data are then specific to a given framework. It is therefore necessary to measure the proportion of core versus specific data required by a framework before deciding whether or not to take it into account in the data quality review. In particular, for reasons of computation time or data volume limitations or regulatory requirements, the granularity of data may have to be adapted. For example, in the Solvency II Framework, considering the high number of runs to perform, the level of data aggregation may turn out to be higher than for MCEV needs (requiring fewer simulations).

Deployment in Entities

Do international insurance groups have to deal with the question of systematically deploying the approach in all of their subsidiaries and branches?

On one hand, insurers may await some immediate benefits of having a transversal and consistent approach regarding Data Quality in terms of cost optimization and regulator assessment.

But, on the other hand, the more international entities are involved in the Data Quality project, the more difficulties are to be fixed. First, from a governance point of view, lots of top managers have to be convinced to invest on Data Quality. This may turn out to be a hard task. Second, technical issues are multiplied: local specificities have to be dealt with (e.g. local accounting referential, local supervisors' level of requirements, etc.). Lastly, the starting current Data Quality level in all of these entities may be heterogeneous, which requires adapted approaches.

Therefore, in general, the solution adopted is to select a sample of pilot entities, enthusiastic about this initiative with a good or medium data and IT systems quality before extending the perimeter of entities covered in the medium or long term. The aim here is to obtain rapidly quick wins to benefit of the momentum hence created to embark the less enthusiastic.

Setting Priorities

To make a decision, the project Steering Committee leans greatly on the Users: the deliverables have to fit their needs and make explicit their priorities.

The project increases its chance of success by obtaining formalized expressions of requirements that form the first brick of a precious documentation: the stage of documentation generally imposes to take a few steps back and to express clearly and precisely the needs. This formalization is all the more important as the structuring choices have to be made by the Steering Committee based quite solely on this material covering different teams with specific needs and points of view.

As the approach is deemed to operate on the long term, it is also important to keep a record of choices and trade-offs.

The needs of the Users being strongly driven by data required by the models (nature of the data, formats and granularity), model designers have a great added value in this phase of expressions of requirements.

10.3 Data Quality: A Loop Approach

A Data Quality approach is typically gradual and iterative and can be represented by a cycle. Due to the absence of an absolute and stable level of Data Quality, the approach aims at increasing the level of quality and also to maintain it in a fluctuating context (Fig. 10.5).

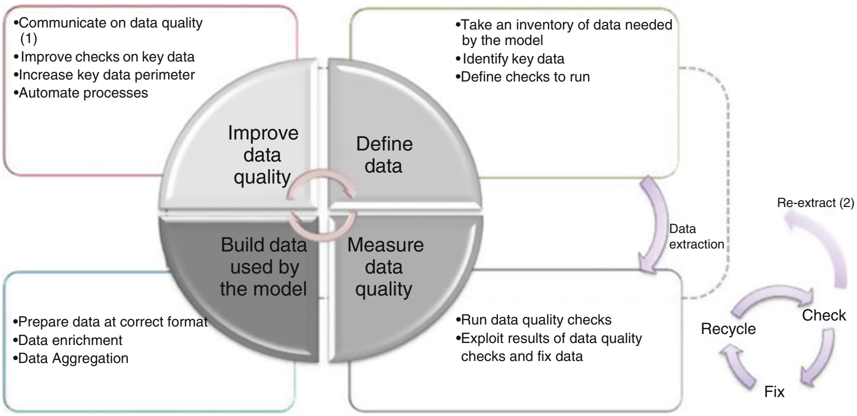

Fig. 10.5 Data Quality loop process. (*1*) Get to the root (PMS) in case of recurrent bugs in order to fix them in the source system (*2*) Re-extract in case of corrupted extractions (e.g. too few lines)

To maximize the chances that this approach is shared among all the Users, it may be helpful to bring them together in order to build this coordinated approach and agree on its components: stages of the process, governance, Data Policy, reporting on Data Quality.

The process is not always meant to change a lot. A first important step is to obtain a description of the existing process. During this phase of description, some improvements can appear.

10.3.1 First Step: Define Data

Data Dictionary and Key Data

As a consequence of the number of teams mobilized by the Data Quality project, they quite quickly need a common vocabulary for their exchanges regarding data. This is one of the needs fulfilled by a data dictionary. Another of its advantages is that it centralizes all the key information on data.

The data dictionary indicates data by data: the data definition, its source (that may be an IT system, an internal or external file), its range of possible values, the list of retreatments it needs all along the process, the controls that have to be performed, its owner, etc. This dictionary is not specific to the models in which data is entered.

Most often these dictionaries are not immediately available on data manipulated in the processes we consider; however, the Users, being the owners of data, should be in charge of their production or completion under the coordination of the project.

Regarding the mass of data in the scope, the best approach is an iterative process on a growing perimeter, at each step beginning with the most sensitive data, with a progressive validation.

The identification of the key data is quite difficult to perform at the launch of the project because many criteria are available and could be relevant:

- Materiality of some lines of business compared to the others: Data related to the most important activities of the insurer should be prioritized;
- Availability of data, stability/concentration of the sources/processes (see the next paragraph);
- Input data/output data: treating input data first may be a logical choice;
- The sensitivity of data.

The sensitivity of data is quite an intuitive criterion but is also quite difficult to assess in practice. It can be expressed in terms of impact on indicators like Best Estimate of liabilities, Solvency Capital Requirement, present value of future profits or other indicators. As noted before, it is of high importance to define clearly in which framework and processes the data quality approach is implemented, what can guide the choice of indicators.

The measure of impacts can be achieved through sensitivity calculations and other studies. This can be long and costly to obtain, especially if the indicators are based on the results of calculations performed with a projection model. A bias can also be introduced by the model if it itself is not already entirely controlled by the teams (the quality of the model has to be taken into account).

In the first place, studies can be replaced by experts' opinions: the project should capitalize on the experience of the Users through workshops and could build a first list of key data after having aggregated the different lists provided by the Users.

Among the advantages of the loop approach, the list of key data may be adapted at each repetition of the process

IT Systems and Data Sources

Once the list of data in the scope of the project is stabilized, the project has to define the best sources for them.

A first issue can be the multiplicity of sources and the differences in amounts they show. The challenge is to choose the best one. The possible and satisfying comparison to accounting data is often a decisive criterion of choice.

Another issue is the short/medium term, known or potential, evolution of IT systems (mergers, migrations, etc.). The progression of the project can be really facilitated by the existence of an updated IT systems cartography and roadmap.

In general, assets data is quite concentrated in IT systems but regarding liabilities the situation can be more contrasted depending on the business models, for example, if an activity is run through different partners. In that case, it is less the multiplicity of sources than the varieties of the formats, the precision (numbers of columns) and the granularity of data received (numbers of lines) that complicates the progress of the project.

Insurers are also tempted to build Data Warehouses that concentrate and store a large quantity of data in the same place. The construction of such a Data Warehouse implies to be quite pragmatic. As it may be seen as quite costly, a multi-year project is necessary that follows an iterative approach: capitalizing on the existing databases, extending them based on specific needs, and creating a common overarching Data Warehouse.

The best approach is to:

- Prioritize the automation of data flows coming from stable and concentrated IT systems;
- Progressively stabilize the formats of the files sent by the distributing partners and make them as homogeneous as possible;
- Include data necessary for comparisons (as we will see below).

Key Controls

The list of controls to be performed has to be decided through workshops with the Users. It may start with a blank page and be based on best practices or anticipated supervisors' requests. More practically, it may be a selection of the most relevant controls already executed by the Users:

- Reconciliation tests: between actuarial bases and accounting bases
- Comparison of amounts between two periods: amounts in year N versus N-1
- Counting of errors and omissions
- Likelihood tests: comparison between the data and contractual documents, respect of a range of values
- Format tests
- Conservation of amounts from IT systems to model entry
- Completeness tests: percentage of the perimeters covered with available data

These controls are quite mechanical and quantitative. Some more qualitative criteria may be usefully added:

- Data is up to date
- Its source is unique
- Exclusions of data are justified
- Data is accessible
- Data can be interpreted (correctly understood by the model)

At each control should be associated synthetic quantitative and/or qualitative indicators (for example: a number of missing values) and a threshold value that shall not be exceeded.

This combination of tests, criteria and thresholds forms the internal Norm of Data Quality. This Norm can be adapted at each run of the process:

- Introduction of a new test
- Adjustment of the indicator and/or the threshold. The threshold being estimated by expert judgment in the first loop may generate too many alerts or at the opposite, may be too loose to capture material errors.

The evaluation of the criteria can be objective or subjective and, moreover, it can be performed at different granularity levels. However, the more fine-grained the controls, the more difficult the synthetic vision will be to obtain. In this case some weights should have to be given to some controls or lines of business or entities, etc.

Regarding anticipated data or data used for periodical reporting, the Norm may be adapted to remain consistent with the methodologies used. In particular, if this reporting is built by means of proxies or requires different data, some constraints could be eased up (accounting reconciliation may not be possible when considering projected data) or some tests should be added.

10.3.2 Second Step: Measure Data Quality

As indeed is the whole process, the phase of control is also progressive and iterative. The first concrete implementation of the controls can be burdensome because of its manual nature. Therefore, the project should aim at implementing a Data Management Tool that would allow for an increasing scope of the controls, an accelerated process of controls and an improved internal control. Throughout the development of this Data Management Tool some functionalities, groups of tests or business perimeters will be made available to the Users. It's important at this stage to maintain the membership of the Users by giving them perspectives on the future milestones regarding the development of the Data Management Tool.

The implementation timescale of a Data Management Tool shall not be underestimated as:

- It takes some time to understand the controls required by the Users
- Data flows to feed the tool automatically are numerous
- Local data specificities entail substantial work if the Company is international
- Data Management tools are intended to be used as much as possible by the Users and they have to familiarize themselves with it as soon as possible.

10.3.3 Third Step: Build Data Used by the Model

The different steps that could potentially be covered by a Data Management Tool are the following:

- Concentration of the data after automatic extractions;

- Analyses/retreatments/standardization/enrichment/controls/reconciliation and correction of the data;
- Aggregation;
- Production of the model points and automatic sending to the model;
- Production of a Data Quality report and follow up;
- Production of an audit trail for all these steps.

In particular, the aggregation step (choice of a level of granularity) is unescapable regarding insurance liabilities projection (computation time, volume of output data) and a source of material error:

- The rules of aggregation of each characteristic have to make sense: some characteristics require an addition; some require a weighted average, etc.
- A too high level of aggregation could lead to an incorrect reflection of the policyholders' behaviour: for example, in the case of savings contracts with minimum guaranteed rates, if the rates are averaged, the guaranteed option will not be correctly valued.

As a consequence, this step is no exception to the rule: it has, first, to be documented and submitted to some controls and as much as possible, automated.

The automation may be progressive and iterative as all the other items of the project are.

Some regulatory texts, notably Solvency II, focus on data security. The evidence of an automatized data process combined with a rigorous audit trail may now be required by supervisors. The audit trail may, for example, help in managing the changes of versions of the data when they need to be reloaded and ensure that the final version of the data is identified, stored and available in case of review.

It would not be enough for the project to stop at the users' doors with the delivery of good quality baseline data. On the contrary, the project, and more specifically the Data Management Tool, should accompany the Users as much as possible in the step of data preparation.

10.3.4 Fourth Step: Improve Data Quality

The reporting on Data quality is instrumental in the efficiency of the loop approach. It answers different needs:

- enabling to capitalize on past executions of the process;
- documenting the controls.

Moreover, it may constitute the proof that the controls were performed and in that case, it is an integral part of the Data Governance (elaborated further). For example, in the case of an international Company, this reporting may be a basis for exchanges between entities and corporate teams or operational teams and internal control teams in order to define together appropriate action plans.

The determination of its format and content is iterative:

- Some controls may need to be manually reported at first;
- After a while, some controls initially selected by expert judgment may seem more relevant than others in the report and should be highlighted, or some may not be easily understood by the recipients of the report, notably at hierarchical level, and should be adapted for better understanding.

The report may be produced by the Data Management Tool.

Discussions may be necessary between Users, project team and teams developing the Data Management Tool, if any, to adapt the report (introduce/remove controls, change the format, etc.).

In the case of an international Company, it may take some time for local teams to familiarize themselves with the content of the report and some training sessions may therefore be necessary.

Data Quality reports are generally produced at the finest level of granularity:

- Type of data (model points, hypothesis);
- Lines of business (savings, protection, etc.);
- Perimeters (geographical, entity, etc.).

Some consolidation is necessary at the end to efficiently support the Data Governance and allow the report to be disclosed to Top Management. Some materiality indicators may be introduced to help this consolidation (set by intervention of expert judgment in the beginning). They can be based on reserves, on premiums, on risks, etc. They allow moderation of bad opinions or underline the necessity of an action plan. A follow-up of the trend of the quality level is also very useful.

For example (Figs. 10.6, 10.7 and 10.8).

Perimeter geographical, entity	Materiality indicator	Asset Data	Savings liabilities	Protection liabilities	Hypothesis	Total
P1	1	Y	G	G	G ↗	G
P2	1	Y	G	G	Y	Y
P3	2	Y	Y	R ↘	R ↘	R ↘
P4	2	G	Y	G	Y	Y
P5	3	G	G ↗	Y	Y	Y
Total		Y	G	G	Y	Y

Fig. 10.6 Example of Data Quality reporting

Code	Interpretation
G	Satifactory quality
Y	Medium quality // some action plans are necessary in medium term
R	Low quality // some action plans are necessary in the short term

Fig. 10.7 Colour code

Fig. 10.8 Trend code

Trend code	Interpretation
↗	Improving
-	
↘	Degrading

10.4 Data Governance

During the time of the project, the Steering Committee is composed of representatives of the Users, Models Team, IT Teams, project Team and the Management and makes a lot of structuring decisions regarding the data project: scope, roadmap and trade-offs.

On the medium term (post project), responsibilities have to be assigned or clarified at Company level between the Users, Models Team, IT Team and the Management. In the case of an international Company, as the organization is more complex, these responsibilities have to be expressed precisely at corporate versus local level and for each data perimeter.

Usually, it takes the form of a Responsibility Assignment Matrix of the process (Fig. 10.9).

The Data Governance should rely on four pillars:

- A sponsor of the Data Governance, ideally a member of the Top Management.
- An operational team, in charge of the governance and the monitoring of its implementation throughout the Company. This team should be dedicated.

Tasks	Responsible	Accountable	Consulted	Informed
T1	-	-	-	-
T2	-	-	-	-
T3	-	-	-	-
T4	-	-	-	-
T5	-	-	-	-
T6	-	-	-	-

Fig. 10.9 Example of RACI

- Data Stewards, responsible for the data on their domain (lines of business, geographical zones, etc.). The Data Steward is in charge of the definition and of the follow-up of quality indicators on his perimeter and coordinates improvement plans when necessary. Data Stewards are generally part of the Users.
- The Data Quality Management Committee (Fig. 10.10)

Some responsibilities on technical aspects not detailed hereafter may be allocated to IT Teams: technical integrity of data, guarantor of the accessibility of data, etc.

Data Management Committee is necessary to steer the Data Quality process. It should meet regularly. Its main missions are to:

- Have a global picture of the Data Quality throughout the Company
- Define objectives and priorities; allocate budgets
- Decide, follow and arbitrate the actions plans.

The Committee should have the necessary influence on teams and with this aim, the chairman could be a Top Manager. In many cases, he could be the Chief Actuary.

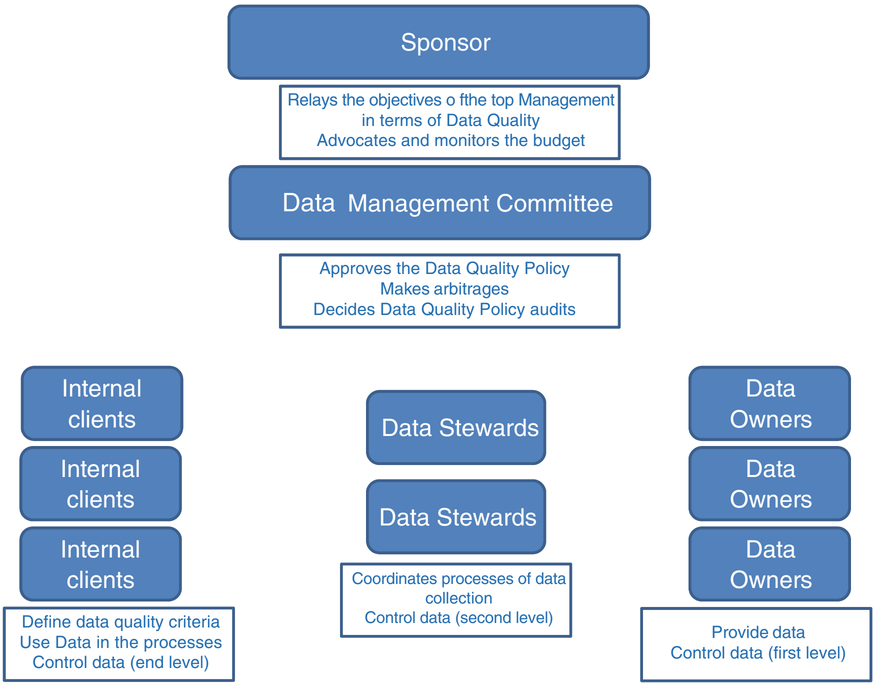

Fig. 10.10 Actors of the Data Governance and main missions

The Committee calendar may be connected to closings (depending on the framework chosen). For example, the validation of the Committee could be a necessary step before using data in the process of closing.

The Committee relies on the report evocated before, consolidated at Company/group level and allowing to identify the areas where action plans are necessary/to prioritize. This report may be sent to the Top Management with adaptations.

Many types of action plans are possible:

- Addition of some controls (manual or automatic)
- Request of systematic corrections in upstream IT systems
- Request of enrichments through other internal/external sources
- Upgrade of the Data Management tool's version
- Conclude or modify an SLA with a data provider
- Merge/Migration of IT systems

Many criteria may be taken into account regarding the decision of implementing of action plans:

- Sustainability of the needs
- Level of priority assessed by the Users
- Delays
- Budget
- Some of them may directly be requested by supervisors.

10.5 Conclusion

In conclusion, the complexity of setting a Data Quality project is quite high. You need ambition, but you have to be pragmatic otherwise you will not be able to move people to work on this apparently endless and tedious task. You need to have a global picture of where the Company stands on Data Quality, but you have to act on local scopes, otherwise you may lose both time and money. You need to get Top Management unfailing support but you also have to commit operational teams. You need to start working with a project (methodologies and specific budgets), but you have to implement a permanent process on a run mode (there is no end in the search for Data Quality).

What will remain is indeed the permanent process:

- The process itself: general design, actors, responsibilities
- The implemented Data Management Tool
- The Data Quality Policy

In terms of documentation, some deliverables appear essential, not only because they are structuring components of the tool but also because National Authorities should check that they exist and are up to date:

- Data dictionary
- IT systems cartography
- The documentation of the process and the procedures related
- The Data Policy including the Data Governance
- The Data Quality Report
- Technical specifications of the Data Management Tool

When facing the Augean stables of Data Quality, many found themselves armless (we are not all Heracles!). Therefore, it is important to start somewhere, to reach valuable achievements that could inspire others. An iterative approach selecting what works and what don’t is at the end of the day the key of the success.

Chapter 11
The Role of Models in Management Decision-Making

Bernard Bolle-Reddat and Renaud Dumora

Abstract Managing is deciding. Today no decision is made in the insurance world without using models. Decision-makers are facing many challenges in this situation: models might evolve from entity-specific to standardized under the pressure of regulators and the lack of diversity of providers, being less relevant and trigerring unintended sheep-like behaviours. Models are good at simulating the future from the past but they fail to simulate the unexpected non-linear phenomenon. Often more than one model is used but their consistency is not guaranteed. Understanding the model is a key issue for decision-makers. Models say what a pure rational based decision should be ignoring strategy and politics, discarding the true beliefs of the person making the decision. Models could not duplicate the decision-making process; however they could be the Fifth Solvency II key function.

Managing is deciding. Some Insurance CEO decisions are easy picks; others are far-reaching decisions committing the Company over the long term, and making assumptions on what the financial, social, regulatory and economic conditions might be at this far horizon. Some decisions, he makes alone; for others he shares the decision-making process with co-decision-makers, often an Executive Committee or the Board of Directors. Some of his decisions are based on few criteria; many others on a large volume of data, assumptions and scenarios. More often than not, the CEO gets insights from a model's runs to simulate the probable or the remote impacts of the decisions.

Indeed, models are new attendees at the table of the decision-makers. Their presence may provoke different kinds of reactions: at one extreme of the spectrum of attitudes, some decision-makers might welcome a *Deus Ex Machina* that will scientifically bring beyond-question answers and therefore they would accept blindly what the models say; at the other extreme, others might blankly reject the unwelcome and caricatured view put on the table by the models as they believe that these machines could never encompass the real complexity of the world; only the

B. Bolle-Reddat (✉) · R. Dumora
BNP Paribas Cardif, Paris, France
e-mail: bernard.bolle-reddat@bnpparibas.com

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_11

human brain could approach it. Anyway, accepting models in the management team leaves no one untouched and changes the decision-making process.

So using models in the decision-making process raises questions. The purpose of this chapter is to address the related issues from a practitioner's viewpoint and to share the experience of BNP Paribas Cardif.

11.1 Could an Insurance Undertaking Be Run Today Without Using Models?

The answer seems quite straightforward as models are everywhere around us in the Company. Almost no decision is made without the insight brought by models. In almost every area of the field of the insurance business one can find some form of modelling, more or less complex.

Big Data and new analytics power are drastically changing relationships between customers and insurance companies. New customers' expectations and increasingly intrusive regulations are pushing companies to rethink their customer journey so that they sell the right product, at the right client, at the right time and through the right sales-channel. Of course, this new customer journey has to be fast and simple. How can it be achieved without models? Indeed, the last stronghold, Marketing, has long opened its gates to models as the opportunity offered by Big Data requires algorithms to exploit it to help design products that will better meet clients' needs.

Of course, the way traditional actuarial missions are performed is now being transformed by the new access to Big Data and by the associated increased capabilities of models. Thanks to the Internet of Things which allows actuaries to measure and monitor the behaviour of customers in real time, thanks to open data which leverages internal information, and in some cases, thanks to electronic personal information (now obtained from online consumption and later probably from social network exchanges), the pricing and reserving processes are now to produce real-time and individual-specific solutions. Property and Casualties actuaries, Life actuaries and Health actuaries cannot quote, forecast or reserve without models: models turn out to be both the best way to get the best idea of what the future might be and the only way to process the mass of data made available by the new technology.

Underwriting processes and claims management are also evolving in the same direction, integrating the benefits of models. High quality service standards require a significant amount of automated decision-making, based on bigger and bigger sets of available data. Models are already helping operations to orient the different customer claims to the best process in term of service quality, efficiency and risk control. This optimized processing based on scoring methodologies has been operational for decades in the banking industry.

Asset Management has chronologically been the laboratory for the use of model testing or even trespassing their limits, as they were in widespread use for pricing financial instruments and in assisting portfolio decision-makers to pick the right assets at the right moment. In a Darwinian move, models were then imported by ALM[1] people to help in setting efficient asset allocations. These allocations optimize both risks and rewards, aiming at constraining traders and ensuring their alignment with the Company's risk appetite.

Finance was also conquered quite early by models as they were needed to prepare and follow-up budgets and business plans. Improving their communication to investors was also a driving force that pushed Finance Departments of listed insurance companies to extend their use of models. They were seeking to develop non-GAAP[2] measures to better explain and sell the insurance business model to potential investors. This quest eventually took the form of the European CFO Forum Embedded Value (EV). With the publication of the EV, models entered into the pages of financial communications and were unveiled to a larger public. This was just the beginning, as models are set to play a major role in book-keeping as the Internal Accounting Standards Board (IASB) is imposing the fair value paradigm for the preparation of the financial statements of European listed companies. The point is that, when there is no deep and liquid market, a balance sheet's items cannot be marked-to-market. They do have to be marked-to-model. This makes models instrumental to the implementation of the incoming new standard, IFRS4 phase 2, as their scope will cover the assessment of all insurance liabilities.

Finally, Risk Management is the latest *Terra Incognita* to be colonized by models: The European Union's Solvency II Directive makes models mandatory to assess the solvency of European insurance undertakings, extending the scope of the limited behavioural simulation achieved in financial models.

So, because available information is increasing exponentially, because clients expect more customized products and services, because regulations are requesting new quantitative angles of analysis, models are now part of the whole insurance value-chain from the marketing targeting process to the accounting and prudential disclosures. Thus, decisions today are hardly made without referring at some point to models.

11.2 Could Models' Inputs in Decision-Making Become a Source of Mass Crisis?

As seen, a review of recent decades shows that models, day after day, progressively proved most of the time to bring valuable information to all stakeholders and sometimes generated unattended consequences. Regulations, surfing on the last financial crisis that underlines failing and (in some cases) absent supervision,

[1]ALM: Assets & Liabilities Matching.

[2]GAAP: Generally Accepted Accounting Principles.

reinforce their role by making them mandatory. In addition, for the sake of allowing a better control, rules were set for modelling.

The danger of these new regulations is that in the extreme, they might make models less entity-specific and therefore less relevant for an entity's decision-making. This standardization of models throughout the Insurance Industry might result from imposed modelling rules by regulations, but also from the relatively limited offer of modelling capabilities. Indeed, in a span of only a few years a middle-size insurance company is going to equip its teams with all kinds of models for marketing, actuarial affairs, finance, and Solvency II purposes. In this very challenging equipment process, it is confronted with the fact that there is no real diversity in the offer, either external or internal, either of products or of people: all companies in a specific market will appoint the same consultancy firms, recruit actuaries and data-scientists who graduated from the same universities or, even buy the same software. One striking example is the situation of the Economic Scenario Generator (ESG), used by European companies for Solvency II calculations, which has been developed by a very small number of providers.

So in addition to making models less relevant to entity-specific decisions, this trend to model standardization might also at worst push insurers to adopt sheep-like behaviour as all the models of the Industry would end up saying the same thing at the same time.

This pro-cyclicity phenomenon has been illustrated in the European Long Term Guarantee Assessment (LTGA) put in place in 2013 to test Solvency II technical guidance. Based on the LTGA results, supervisors eventually agreed to insert some measures to reduce the undue volatility of the Solvency II balance sheet, namely the equity dampener and the interest rate spread volatility adjuster. However, some continue to believe that these patches are not strong enough and that, in the long run, the European Insurance Industry will face difficult times when financial crisis comes. To avoid a crackdown of the new framework right from the start, Solvency II will implement a very sound measure, called 'grandfathering', which will allow the Insurance Industry to progressively adapt their asset management, their ALM and their capital management to the new rules. A more brutal implementation would have generated a dangerous massive move of all investment policies and would certainly have impacted the financial markets.

Of course, models can add to the diversity of a collective decision just like the multicultural and interdisciplinary diversity that all Boards of Directors are now seeking. Models are an asset for the decision-making process. But, if, as a decision-maker, you believe both that all models (including yours) are formatted to reproduce cycles and that the current cycle will soon be reversed, then contradict the model and you might prove to be right!

Therefore, challenging the models is absolutely necessary for the sake of achieving a sound decision.

11.3 Can Models Predict the Future and Can They Help the Company to Prepare for the Future?

A new population is progressively occupying more and more desks in the technical departments of insurance companies. These new employees, naming themselves ‘model-designers’, are in charge of creating the projection models for ALM, business plans, embedded value calculation, reserving… They are mostly engineers, statisticians or actuaries. They are good at simulating the future by using the past, excellent at building forty-year-long scenarios consistent with the past, and the best at building central scenarios based on stochastic simulations. Like modern oracles, they are even able to make sense of the thousands of scenarios deemed to simulate all potential futures of their enterprise. But they are not the ones to imagine the future of the Company.

Indeed mathematical models are generally not designed to simulate the unexpected, the nonlinear phenomenon, and the Black Swan events. They are based on historical series, on large financial or macro-economic databases, and on econometrical algorithms. Very rare events (those that have never happened) are very difficult to account for in the simulation. Many hot debates between experts have already challenged the insufficient weight of the distribution tails without reaching an undisputable conclusion. Still, modelling is generally understood as mathematical only, forgetting the Black Swan crisis.

That is why modelling cannot be only a mathematical approach to the future. It has to be a smart mix between both mathematical algorithms which simulate the future central scenario and its sensitivity to small variations of the environment, on one hand, and a set of stress tests decided by the management of the Company, at its highest level, to materialize the impact of their fears and of the ruptures they see as possible, even if remote, on the other hand. However, foreseeing the unknown Unknowns, as Donald Rumsfeld named them,[3] is a challenge for all. Most reinsurance and insurance companies have now set-up an emerging risk identification process to support this challenge and to help senior management to imagine new stress-based scenarios.

So models are not there to predict the future. There are here to shed light on a large spectrum of possible future states of the Company. But they are to be complemented by forced analysis on the impact of disruptive unknown-by-the-mathematical-model scenarios. The decision should only be made with both inputs.

[3]Donald Rumsfeld made this famous statement while serving as George W. Bush’s secretary of defense: ‘As we known, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns- the ones we don’t know we don’t know.’

11.4 How Consistent Are Different Decisions When They Are Made Based upon Different Models?

Diversity of models may be an issue for decision-makers. How could they be sure that each one of their decisions is consistent with the others if they are backed by different models?

Some models are deterministic, others stochastic; some optimize the impact on next year's result, others are focused on the present value of running off insurance portfolios over more than forty years; some are run in the risk-neutral environment, others in the real world-risk environment; some simulate local GAAP impacts, others look at IFRS impacts; some run given parameterized dynamic laws, others run different calibration of these same laws; some have been sophistically coded, others are rough proxies; some can be developed with a simple Excel sheet, other use huge parallel computing machines; some simulate a closed book, others consider future new business…

Decision-makers need to see some order in this jungle of models. However, this does not mean constructing a Babel Tower of models that encapsulates all models into one single ultimate model that fits all purposes. This, we believe, would be doomed to fail as it would mean fixing too many and too complex conflicted issues on a very large scale.

More modestly, at BNP Paribas Cardif, we believe that different models will always coexist, at least for practical reasons. In consequence, they should be organized in a global Model Mapping so that every intended use is adequately covered, any consistency between models that should be there is ensured and documented, every stress test studied is codified: its intensity is well understood by decision-makers thanks either to a quantitative metric (i.e. a percentile) or to a qualitative shared definition (e.g. 'severe but not extreme').

The goal of the global Model Mapping is also to show the areas of relevance for each model so that decision-makers know when to trust it. At BNP Paribas Cardif we therefore created a team dedicating to modelling our Group models and to oversee local models.

The consistency issue also exists when considering a sole decision.

It is common for a model that gives the outputs for the decision to be made to use itself as inputs for other models. The most obvious example is the scenarios used in a stochastic simulation that are generated by a separate model developed and run, either internally like at BNP Paribas Cardif or externally by a specialized firm. You may also think about dynamic laws that are designed separately.

Ensuring the consistency between the different assumptions of the different models is necessary to achieve the required level of overall reliance. Some links between models are quite straightforward: think about the implicit risk correlations implemented in the ESG and the correlation matrix used to calculate the Basic Solvency Capital Requirement. Could they be very different?

Some links are more complicated: let's take the example of the dynamic surrenders laws. These laws are designed based on historical track records of real-life behaviour. Should they be adjusted to fit the risk-neutral world when they are used in a solvency calculation?

Dedicating resources to back-test and to make explicit the level of consistency between models is a no-brainer and a true way to improve models. Some academic research is certainly needed to help bridge the different models' conceptual frameworks. However hard it may be, decision-makers deserve to understand the reasons for the differences between models so that they are better armed to efficiently challenge their outputs to eventually reach the best decision. In the same way, decision-makers also deserve to fully understand the precision of the outputs of the models, the sensitivity to any assumptions made and the quality of these assumptions.

11.5 Is It Possible for a Decision-Maker to Understand Models?

Those decision-makers that have to use models' inputs to make their decisions are certainly asking themselves, more often than not, the following terrifying question: 'is it really possible to understand models?' This is an all the more important question for non-specialist top managers as the EU Solvency II Directive makes it a legal requirement that the Company's supervisory and management body collectively understands the models. As this collective understanding may rely on a happy few, some boards might consider appointing model experts. The worst risk would then be the illusion of truth or even the illusion of precision.

However, at BNP Paribas Cardif, we believe it to be more interesting in order to foster debates between decision-makers that they all receive appropriate training. We started it a couple of years ago, organizing off-site sessions of the Executive Committee, which we named 'Bunker Days' as the purpose was to force the committee to focus on technical issues.

The first conclusion that we drew from this experience is that it is very difficult in practice to set the bar for what one should learn at a minimum to make an educated decision. It all comes down to knowing what technical choices and rules are significantly limiting the outputs of the models and what parameters are materially influencing their results. That is still a lot to learn!

The second conclusion is that you need to find a five-legged sheep that knows how the model works in all its technicalities and how to make them understandable quite fast (in a span of a few hours) to non-specialists. Usually you get one (a hardcore technician) or the other (a great communicator). Increasing the understanding of the algorithms that stand under the hood of models is really necessary as decision-makers have to be familiar with the ins and outs of the models to fully exert their duty of challenging them.

11.6 Is the Decision What the Model Says?

Models' outputs do not always have the same weight in decision-making.

In strategic decisions, such as whether or not to buy a Company or an insurance portfolio, a model's outputs feed the battery of quantitative indicators under review, which are looked at in conjunction with qualitative considerations: commercial positioning, know-how, legal issues… The context of the decision-making is here often political: it implies a bargaining with a seller (or a co-decision-maker) based on information that is asymmetric, maybe biased or incomplete and also based on agendas sometimes diverging and hidden. Therefore one should not be surprised to see a decision eventually made quite different from what it would have been had the sole input of the models been taken into account.

In other instances, it seems that the decision is to be made on a pure rational basis by a decision-maker who is behaving himself as a true *homo oeconomicus*; in that circumstance, it could be expected that the decision should be what the model says.

This is the case, for example, when setting the strategic asset allocation. The goal of the decision-making is clearly defined: it is to find the optimized asset portfolio based on two criteria: profitability and risk profile. This sounds like a pure mathematical problem to solve.

However, at BNP Paribas Cardif, we experienced lengthy debates on ALM around the model's proposal of asset allocation, as it turned out that there were explicit or implicit diverging views on certain parameters:

The probabilities of the scenarios studied.

- In a stochastic modelling, each scenario has the same probability (1/n, n being the total number of scenarios). The calibration of the ESG is therefore instrumental to the outputs of the model, and there are many different calibrations that would pass the market-consistency test and be considered as fit for running the model. Which one of these calibrations is the good one?
- In addition, lately some[4] have challenged the relevance of the market-consistency requirement in certain uses of the models (such as the solvency calculation), notably the requirement that a market-consistent ESG should generate a certain number of scenarios with negative interest rates and extravagantly high interest rates. When seeking to find the price of a financial instrument, these specific scenarios are not a problem. When having to simulate management actions in a solvency calculation, they are.
- Even without considering these technicalities, some decision-makers, simply from past experience, by expertise or by their own personal attitude towards risks, may have a more or less optimistic or pessimistic view of the probability of occurrence of a given state of the financial world compared to the ESG.

[4]Félix and Planchet (2015).

The assumptions used by the model.

- Assumptions should represent decision-makers' views on the future.
- In most cases, whenever there is a relevant market, they are based on the market consensus. This quite mandatory reference to markets (i.e. the market-consistency rule) comes from the efficient-market hypothesis (EMH), which asserts that market prices reflect all publicly available information and sometimes even hidden or "insider" information. Who knows better what the future might hold?
- However history and academic research[5] show that markets are imperfect as they are also like individuals facing cognitive biases and information bias. Therefore, why would the views of a decision-maker be less legitimate than the markets' predictions?

The dynamic laws.

- Dynamic laws aim to duplicate the decision-making process of the client (will I surrender my contract or not?) and of the top management of the insurance Company (what strategy will I follow for investing and disinvesting cash inflows and outflows? What participation will I give to policyholders?). This duplication is as fair as it could be, but there is always a gap between the modelling and its shortcuts and the complexity of human behaviour. Therefore how could it be avoided that a given decision-maker feels betrayed by a simplified simulated management action?

The decision criteria

- Whatever the quality, the realism and the precision of an ALM Model, its outputs are as relevant as the multi-dimensional criteria it referred to, to optimize the asset allocation. Today IT technologies give us the capability to extend almost ad libidum the scope of an optimization's criteria: extreme risk, short term volatility, profit and loss, value for shareholders, value for our customers. Solving this combinatorial problem with a sole calculation is technically possible but difficult to use in a decision process. Indeed, the different angles of analysis have to be explicit and to give together a global picture to the decision-maker. This may make it necessary to use shortcuts in the sequence of optimization.

Indeed, there are good reasons to challenge the outputs of the model. Collective debates around the model inputs and outputs are absolutely a must, first at the technical level and then at the political decisional level. These discussions often lead to a decision which is different from a simple reading of the direct output of the model. So what to do if the decision made for the strategic asset allocation is not what the model says? There are two different schools: one is to change the model so that it fits the decision. This is the position of use tests fans. The other school asserts

[5]See behavioural economists' works.

that a model could never (or only by chance) give the decision. The model is only one participant amongst others to the decision. The robot's opinion has not to be over-weighted by human decision-makers. Provided that the difference between the decision and the model's proposal stays in a certain range considered to reflect a normal 'challenge zone', no change to the model is necessary. Indeed, only identified material errors and decided model turnarounds are subject to modifications.

At BNP Paribas Cardif, we adhere to the second school. Therefore, in order to make the decision-making process more explicit, we ensure that the reason for the difference is explicit and commented upon in the minutes of the decision.

Regarding the model's outputs as only one simple input to the decision could render the audit trail of the final decision less clear and specifically for the review by external bodies. In particular, EIOPA and national supervisors have to adapt themselves to this new decision-making process, which despite the increased role of models and the required use test, should not, we believe, be automatic. This implicit 'multiple-brain' principle, just as the Solvency II Directive explicitly sets the 'four-eyes' principle for the internal control framework, is the basis for the Solvency II key functions framework. The different experiences and technical skills of the Chief Executive Officer, the Chief Risk Officer, the Chief Finance Officer and the Chief Actuary who all participate in the ALM collective decision, and the different viewpoints they embody, guarantee the quality of the decision, perhaps more than the precision of the model. If challenging models is also supported by the models' certification process, it is first the responsibility of fit and proper key function representatives.

There is one instance in which the outputs of the model are mandatory to the decision: the computation of the Solvency II solvency capital requirement (SCR) and eligible own funds. Applying the standard formula approach or the internal model approach, a model is required to perform the calculations, approved *ex ante* by supervisors in the case of the internal model or controlled *ex post* in the case of the standard formula. In either case, internal model or standard formula, the results of the model is what is reported in the Quantitative Reporting Templates (QRT), without any change.

Strangely enough, even in this clear-cut situation, the Directive is blurring the picture of the reliance of models and how they may be trusted by decision-makers. It adds a supplementary layer above the standard formula assessment: this assessment (what is reported) has to be completed by the own risk and solvency assessment (ORSA): it is mandatorily required that assumptions of the standard formula be challenged. In the end, the ORSA solvency position might even differ significantly and thus trigger decisions quite different from those made based on the standard formula solvency position.

The conclusion to be drawn is that the models do one thing: they are telling a story, a simple story, understandable, with explicit shortcuts. They have to. That is what their role is about, i.e. to enlighten a decision with the many possible outcomes depending on the scenarios and the assumptions. Sensibility studies are in that respect very useful. The decision-maker's responsibility is then to pick the outcome he feels best fit what he is foreseeing and what his objectives are.

11.7 How Far Could a Model Duplicate the Decision-Making Process?

Let's take the example of the modelling of the French discretionary participation to policyholders to illustrate three limits to the realism of simulations by insurance companies.

Some years ago, when setting the parameters of its models, BNP Paribas Cardif decided to set the target for the annual participation of its French savings portfolios as a percentage of the spot rate of the French Government ten year bonds (OAT). This simple rule appeared to be appropriate when compared with the historical records over quite a long period of time.

Then, more recently, having gone through a couple of financial crises, a back-test showed that this parameter now needs some refining. Based on the Executive Committee members' work on their decision-making in that area, it was concluded that the target annual participation should still refer to the spot OAT but also to the previous years' annual participations and to internally forecasted competitors' participations. The equity and real estate markets' situation was also considered to be valid input.

Today, we are pointing at a first limit to the realism of insurance models that we consider to be a flaw of the classical stochastic framework as applied to the solvency calculation. In this framework, all scenarios have the same probability (1/n, n being the total number of simulated scenarios). This means that extreme scenarios have a global impact higher or lower than it should be considering their real life probability of occurrence. We consider that to fix the situation we should change the target's formula depending on scenarios. Indeed, in a real life situation, the decision-makers would certainly not set the target at the same value in scenarios around the central scenario and in extreme scenarios. However, implementing this conditional target would raise new mathematical issues relating the algorithms to implement that is still, we believe, a void research field. Academic work on this issue is very welcome. However, any solution eventually found would have to be run in the fast and challenging timetable imposed to produce the solvency reporting.

Behind the equi-probability issue lies a second limit to the realism of insurance models which in fact represents their main weakness: the simulation of management and customers' intertwined behaviours in extreme scenarios.

Firstly, it is a business issue. Because these scenarios are scenarios that no one has ever experienced in the past, the model designer obviously has some difficulty coding what could happen in these unknown situations. For example, referring again to the modelling of French General Fund profit sharing, who could say today what would be the respective behaviours of insurers and customers in the case of a drastic and durable increase of interest rate of, say, 400 bps after twenty years of smooth slow interest rate decrease?

Secondly, it is a practical issue. Model designers are de facto limited as they are facing a very large combination of all possible management actions and customers' reactions. Their parameterization has to be limited in term of the parameters' number so that models may be eventually run in a reasonable timeline.

The third and last limit to the realism of insurance models is due to the complexity of the stochastic models themselves. Indeed, their realism depends on so many items: the input data quality, the precision of the replication of management actions and customers' behaviours, the quality of the event generator, the perfection of the ESG, the ability to disclose the right outputs in a sufficiently explicit manner so that checking is possible… Stochastic models are usually documented by hundreds if not thousands of coding pages. Therefore biases and mistakes might come from anywhere; some of them are certainly difficult to detect as it is hard for both the experts and the management to achieve a holistic view of the technical internal consistency of the model.

For instance, focusing and investing time and money on replicating as precisely as technically possible management actions or customers' behaviours is worthwhile only if the sophistication of this replication is homogeneous with what is achieved elsewhere in the model for items of the same significance level. It is the responsibility of model designers to alert users and decision-makers to the weak links in the model. It has to be explicit.

Understanding the actual inconsistencies within the model and the target achievable level of sophistication is instrumental to developing a cost-efficient fit-to-decision-making model. You have to know what makes a difference: there is no use in spending lengthy discussions on some detailed assumptions when shortcuts are made elsewhere on far more sensitive issues.

Models try to duplicate human behaviour, this is their *raison d'être*. Their outputs try to point at the decisions that will be eventually made by the human brain. This duplication process is fruitful even if today the technical road seems still very long before models could be changed by robots making decisions in place of Directors. However, this road is still valuable as it brings an increased knowledge of the risks underwritten and of their risk profile. It also forces decision-makers to think about their decision-making process and to better understand their potential biases.

11.8 Conclusion: The Model as the Fifth Solvency II Key Function?

Models are there to help decision-makers. They give information that, like any information, should be put into perspective and carefully weighted before the decision is made. The more sophisticated the model the more insights it brings, provided that the level of sophistication does not hinder a good understanding of how the information has been obtained. Whatever the complexity of the calculations made by the model, high or low, challenging the outputs is the first responsibility of decision-makers so that they get the full information including what the model did not do (i.e. its limitations).

Not following the advice of the model does not mean making a bad decision. Models do not have a holistic view of the situation as they do not (or rarely) get political or strategic inputs, for instance. They process information in a rational decision-making mode, assuming that all economical agents are (and capable of) doing so. We now know that this is hardly the case in real world.

Finally, the course of action in using a model to make decision comes down to understanding its area of reliance, to investigate what could be the impact of the decision when out of this area, to study specific stress tests illustrating specifically feared scenarios by the decision-makers, to identify goals for the decisions that are not been considered by the model… and then to make the decision.

Now that the Model is seated at the decision table, it will remain there for good. It brings valuable inputs for the management of the business. Almost no decision is made today without at some point referring to the output of a model. Its interactions with decision-makers, CEO, CFO, CRO, CAO & CCO, are instrumental to making an educated decision; they therefore should be fostered and documented within the decision-making procedures.

We believe that in these circumstances, the Model has proved that it can become a separate Fifth key function in addition to (under Solvency II) the Chief Actuary, the Chief Risk Officer, the Chief Auditor and the Chief Compliance Officer.

In order to rise to the level of a key function, like the other members of the key functions' team, the Model has to be 'fit and proper'. We may consider that the 'proper' rule is not applicable. However, the 'fit' rule is applicable and its answer may turn out not to be as straightforward as first thought, despite huge progress in modelling science, the always increasing power of computation and the easy access to Big Data. Models must still be improved and controlled and furthermore, its key function companions and users must be educated.

We are at the very beginning of a new era of model management. As science fiction author Isaac Asimov did to robot management with the Three Laws of Robotics, we may consider designing the Three Laws of Models as a safety measure, to keep this new player under control. In addition, each insurance undertaking has to rethink its organization rules and its decision processes to make room for the models.

The Three Laws of Models we propose are:

- A model helps to make a decision: the decision-making responsibilities are always borne by human decision-makers.
- A model should not be used without being properly understood by human decision-makers.
- A model should not mislead a human decision-maker; therefore its outputs should always be challenged and complemented by ad hoc analysis and stress tests.

Chapter 12
Models and Behaviour of Stakeholders

David Ingram and Stéphane Loisel

Abstract Stakeholders' behaviour has several impacts on model risk and model management. The asset side of the balance sheet may be impacted by feedback loops on financial markets. Customer behaviour is one of the top risks faced by a company. Consequently, it must be carefully addressed in the model. In addition to these behavioural risks that have to be taken into account in the models, we also study in this chapter the behaviour of decision-makers of the company through their attitudes with respect to risk and with respect to models.

12.1 Endogenous Risk and Feedback Loops in Financial Markets

Copycat behaviour may create feedback loops and correlation crises, during which correlation between risks suddenly increase, in spite of the fact that those risks were assumed to present low correlation or independence in the classical regime. Large investors may have some price impact on markets, creating or reinforcing correlations in financial markets. To introduce endogenous risks and feedback loops in insurance and finance, let us use the analogy between the Millennium Bridge evacuation and LTCM's failure to develop by Danielsson and Shin (2003). In their excellent paper, Danielsson and Shin (2003) state that endogenous risk is relevant when individuals are similar in terms of beliefs and constraints. In addition, individuals must be similar in terms of potential reactions to events, and their concerted reactions should have a feedback impact on the market or on the system to create endogenous risk. Their Millennium Bridge example is striking: when this bridge linking St. Paul's cathedral and the Tate Modern Gallery opened, the bridge had to be evacuated after a few hours, as oscillations around the axis of the bridge were

D. Ingram
WILLIS RE, New York, NY, USA

S. Loisel (✉)
Univ Lyon – Université Lyon 1 – ISFA – LSAF, Lyon, France
e-mail: stephane.loisel@univ-lyon1.fr

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7_12

endangering the persons walking on the bridge. What is the probability that one thousand people on that bridge walk exactly in step and make the bridge move that much? If one assumes independence, then this probability is close to zero. But given feedback loops, this probability is in fact close to 1. At the beginning, the bridge is almost horizontal, but not 100 % horizontal. Everyone on the bridge will be subject to the same slight slope and without noticing, everyone on the bridge will push a little bit more on the left foot, say, to compensate. As everyone reacts in the same way almost at the same moment, it pushes the bridge a bit more in the other direction, causing everyone to adjust their stance again, which in turn reinforces the oscillation… What is very interesting in this example is that oscillations were created and amplified within the system, in the absence of any exogenous phenomenon like high tide, extreme wind or earthquake. Danielsson and Shin (2003) compare this bridge evacuation with the failure of the LTCM (Long Term Capital Management) fund. This fund used to be regarded as a kind of second central bank of liquidity in the 90s in the US. With several Nobel prize winners on its board, this fund gained power thanks to convergence trading, a kind of dynamic arbitrage strategy exploiting spreads in interest rate markets that were supposed to vanish after some years. The strategy consisted in taking a long position in the cheapest product and a short position in the most expensive one. This enabled the fund to make some profit at acquisition time, and if one waited long enough, there would be zero profit or loss at the end as prices were supposed to converge. The problem was that the fund was beginning very large and that many copycat funds were following similar strategies. As the spreads were narrowing because of the action of arbitragists, LTCM had to venture into uncharted territories and was hit by the Asian and Russian crises at the end of the 90s. Some margin calls forced them to exit some leveraged trades, causing some adverse price movement: the price of the most expensive product went up and the price of the cheapest one went down, causing more distress given their respective long and short positions in those assets. This led to more margin calls, and to more adverse price movements as copycat funds were experiencing similar issues and amplifying the problem. More distress, more margin calls, more trades to be exited, more distress, and so on caused LTCM to fail due to a scenario that had an extremely small probability under the classical risk models that did not take into account endogenous risk.

The LTCM story could lead one to think that feedback market impacts are always dangerous and associated with a violent crisis or event. This is far from being true. Indeed, some other feedback impacts have a stabilizing role. For example, many investment managers have strong investment guidelines. In France, insurance companies tend to invest from 6 to 10 % in stocks, 3 and 10 % in real estate, some percentage sometimes in what they call diversification funds, and the rest in bonds, with some more limits. If stock markets go up, the share of stocks in the investment portfolio might exceed the limit. The manager would then sell some stocks, which participates in a downward potential movement. Conversely, if stock prices go down, such rules will lead managers to purchase more stocks, pushing them up again. This fix-mix strategy, widely adopted by companies around the world, has in fact a stabilizing impact. Cont (2010) nicely shows how concerted actions of investors (for example, when investors follow a fixed-mix investment strategy) can create correlations between asset prices, even when there is no fundamental correlation initially. Of course, some other strategies can amplify crises: if investors follow some investment guidelines with loss limits, then one faces the risk that when prices go down, the selling barrier of more investors is crossed, and their selling orders pushes prices further down, causing more loss limit breaches and more selling orders. We will come back to feedback risk later on in this chapter when we introduce the surprise game and the impact of the change in risk attitudes.

12.2 Endogenous Risk and Customer Behaviour in Insurance and Banking

We have observed mass customer behaviour in the past in the banking industry. One famous example is the bank run experienced by the bank Northern Rock. In Greece, massive cash withdrawals occurred in the first semester of 2015 as the fear of a freeze of bank deposits increased in the population, after the Cyprus precedent. In the insurance industry, the risk of mass surrender is taken seriously by supervisors and insurers, even if such a phenomenon has not been experienced in developed countries recently. Loisel and Milhaud (2011) compare bank run risk and mass surrender risk, investigating the method to switch from deterministic lapse rates to stochastic lapse rates. For a modeller who needs to build a stochastic model from a static model, it is very tempting in general to replace a deterministic number (like the surrender rate given a certain level of interest rates) with a Gaussian-type random variable, implicitly assuming conditional independence between the decisions of policyholders given the economic context. However, in the presence of a crisis or of a particular situation (like a sudden increase of interest rates), policyholders might be influenced by their pairs and also by gurus or specialized newspapers, as happened in the UK mortgage market: customers massively followed the recommendations of specialized newspapers, even when they were disputable. In that case, some correlation crisis may occur: the decisions of policyholders that were regarded as conditionally independent with respect to interest rate levels suddenly become more correlated, at the instant where their marginal probability to surrender is increased. Loisel and Milhaud (2011) show that the switch from a Gaussian-type world to a bi-modal world (where each mode corresponds to one potential recommendation of the guru or specialized analyst) can create huge differences in the appreciation of the 99.5 % Value-at-Risk for the surrender risk. Note that mass surrender is not the only mass behaviour risk; another one is the risk of mass non-surrender, where policyholders surrender very seldom contracts that offer high guaranteed interest rates that had been granted a long time ago.

12.3 Lapse, Sales, Fraud and Moral Hazard

Lapse risk in Property and Casualty insurance and surrender risk in savings portfolios can also be managed with the help of analytics. Dutang (2010) studies the determinants of lapse in the motor insurance industry. One particular point is the effect of cross selling: customers who have more than one contract with an insurance company may lapse two times less than other customers. Lapse is also more frequent after a claim, depending on the degree of satisfaction of the customer and her willingness to accept the potential increase in the price of her contract for the following years.

In addition to risk management, customer behaviour becomes very important for sales, particularly with the development of online banking and direct insurance sales. Very often, in digital insurance conferences, one asks the room the following question: who physically visited his/her bank last year? Usually, less than 5 % of the audience stands up. When one asks the question: who went on his/her bank website during the last month?, everybody stands up. The behaviour of customers who visit websites or ask for a quote is key for the sales strategy. Internet leads are very expensive and it is crucial for the insurance to choose the right set of strategies to attract and retain customers. The change in the sales channels' relative weights is likely to modify customer behaviour. It will be a challenge to master this change in the next decade.

One must not forget other types of customer behaviour: fraud and moral hazard in particular. Some ideal, low risk and long-lasting customers who were offered privileges by an insurance company suddenly exhibited extremely bad loss ratios, because their behaviour changed due to moral hazard. On the contrary, it is sometimes possible to influence customers in a way that is beneficial to both parties, thanks to risk prevention plans or investment advice. However, in some other situations, influencing the policyholders would not be possible as it would correspond to a recommendation that is in contradiction with their interest. Therefore, good risk management starts with good writing of contracts, and the potential consequences of customer behaviour in different contexts must be anticipated from the beginning. Fraud detection is now common practice but remains difficult to implement. In some markets like cell phone insurance, the price is strongly affected by fraud. In internal models, one rarely distinguishes fraudulent claims from classical claims, except when there is a particular risk or previous evidence and present developments.

Another important issue is the impact of beliefs and attitudes with respect to risk on creation and evolution of models. We present the approach of Ingram and Underwood (2012) on risk attitudes. Adapting ideas from Thompson et al. (1990) (in particular the so-called cultural theory, plural rationality, or anthropological approach), Ingram and Underwood (2012) classify individuals into four risk attitude categories: maximizers, conservators, managers and pragmatists.

12.4 Four Different Perspectives on Risk

The four basic risk perspectives were first identified in the 1980s in research about the variety of human groups. Clear patterns emerged in the data and have proved quite resilient over time that included a common perspective of risk. Within businesses, most people tend to identify with one of the following perspectives:

- **Maximizers**. This perspective does not consider risk very important—*profits* are important. Businesses managed according to this perspective will accept large risks, so long as they are well compensated. Managers who hold this perspective believe that risk is mean reverting—gains will always follow losses —and the best companies will have larger gains and smaller losses over time.
- **Conservators**. According to this perspective, increasing profit is not as important as avoiding loss. Holders of this view often feel that the world is filled with many, many dangerous risks that they must be very careful to avoid.
- **Managers**. Careful balancing of risks and rewards is at the heart of this perspective. Firms that hold this view employ experts to help them find risks offering the best rewards, while at the same time managing these risks to keep the firm safe. They believe that they can balance the concerns of the first two groups, plotting a very careful course between them.
- **Pragmatists**. This perspective is not based on a specific theory of risk. Pragmatists do not believe that the future is very predictable—so, to the greatest extent possible, they avoid commitments and keep their options open. They do not think that strategic planning is especially valuable, but rather seek freedom to react to changing conditions.

Ingram and Thompson illustrate the four risk attitudes with the diagram presented in Fig. 12.1. In each case, the company is represented by a ball, evolving in an economic environment. For conservators, for example, the company evolves in a very risky environment. The unstable balance might be threatened by any important risk taking, which leads to enforce very strict risk limits in this risk attitude.

Each of the different perspectives prefers a different strategy for dealing with risk. Firms led by Maximizers seek out risk, believing that no risk is inherently unacceptable—every risk presents an opportunity, and the trick is to negotiate appropriate compensation. Conservator firms shun risk of all sorts. Manager firms carefully manage and calibrate both the amount and type of risk. Pragmatist firms seek diversification but otherwise have no overarching strategy—they operate tactically, reacting to each new development.

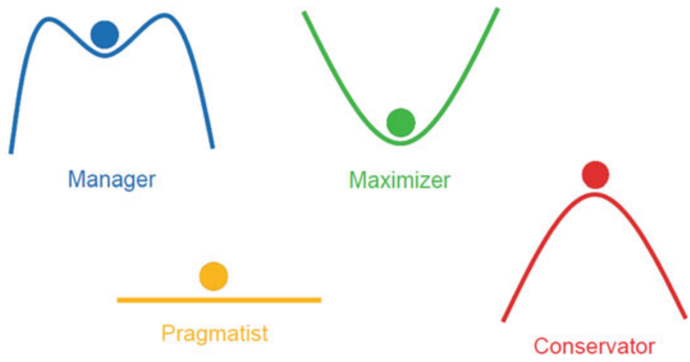

Fig. 12.1 Four risk attitudes

12.5 Four Risk Strategies

Careful examination of risk management practices in a large number of firms reveals that there are four different strategies that fall under the general heading of risk management:

- **Loss Controlling**. This is the most traditional form of risk management; it seeks to identify and mitigate the firm's most significant risks. This includes activities such as safety programs that seek solely to reduce losses. One characteristic of these processes is that they often seek to get "everyone" involved. This type of risk strategy is favoured by Conservator firms. It is particularly appropriate for managing risks that are acute and severe.
- **Risk Accepting**. Many financial firms favour an approach to risk that focuses mainly on getting the price of risk correct. For financial sector firms, this can lead to complicated models of risk and reward. A Risk Accepting strategy is most often applied on a transaction-by-transaction or project-by-project basis. Non-financial companies will choose projects that will be highly profitable, if they succeed. This type of risk strategy is favoured by Maximizer firms. It works well for risks that are relatively benign.
- **Risk Steering**. Under this strategy, the major strategic decisions of the firm go through a rigorous planning process coupled with intense analysis. Risk decisions are made based upon careful cost benefit and risk reward analysis. Perhaps this is why many think that Risk Steering is "real" ERM. Risk Steering ERM is highly favoured by academics and consultants; Manager firms find it appealing, but firms that hold any of the other three risk attitudes do not. This strategy is particularly appropriate for a highly complex portfolio of risks.
- **Diversification**. Spreading risk exposures among a variety of different classes of risks, and avoiding large risk concentrations, is another traditional form of risk management. Formal diversification programs will have targets for the spread of risk with maxima and minima for various classes of risks. The newer ERM

discipline adds the idea of interdependencies across classes, providing better quantification of the benefits of risk spreading. Pragmatists tend to favour diversification because it maximizes their tactical flexibility, but they avoid reliance on any particular risk mitigation process and often mistrust quantitative measurement of risk. Firms whose risks are highly uncertain would often choose this strategy.

But it is important to note that very few firms keep the same risk attitude or the same risk strategy over the long term. Simply put, the world does not stand still.

12.6 Changing Risk Environments

The existence of the four different risk perspectives can be explained. All four are correct, but not all at the same time. That is because over time, the risk environment changes. Many people think that either things are "normal" or they are "broken". But people do not necessarily agree about what is "normal". An observer viewing the world through the lens of the Conservator might say that extreme hazard and danger are the "normal" state of affairs—while a Maximizer, finding this view timid and overly pessimistic, might argue that profitability is "normal" and hazardous conditions prevail only when the market is "broken".

Expanding our view to allow more than the two normal/broken states allows for the possibility that both the Conservation view and the Maximization view can make sense. Consider a model of the risk environment with four risk regimes:

1. **Boom Times**. Risk is low and profits are going up.
2. **Recession**. Risk is high and profits are going down.
3. **Uncertain**. Risk is very unpredictable; profits might go up or down.
4. **Moderate**. Both risk and profit fall within a predictable range.

Such a model seems to be a reasonable description of the phases of business and economic cycles. As the cycle moves through these four different states, external conditions match the worldview of each of the four different risk perspectives. Each perspective has been right part of the time—and will be again, at some point in the future. But none of the risk perspectives is perfectly adapted to external conditions all of the time and surprises will occur (Fig. 12.2).

Purists with the Manager point of view may object that their view takes into account the full range of the cycle. But economic cycles are not simple sine curves; the period and amplitude are irregular, unexpected events do occur, and there are always "unknown unknowns". Model risk can never be eliminated, and restricting risk strategy to a Manager-only view obscures this important fact.

A Risk Steering ERM program works especially well in the Moderate risk environment when risks are fairly predictable. But in a Boom Times environment, firms following such a program will unduly restrict their business—not as much as Conservator firms, but certainly more than Maximizer firms—and more aggressive competitors will be much more successful. In the Recession environment, a Risk

EXPECTED WORLD \ ACTUAL WORLD	UNCERTAIN	RECESSION	BOOM	NORMAL
PRAGMATIST (Fatalism)	ALLIGNED (No surprises)	Expected windfalls don't happen	Unexpected runs of good luck	Unexpected runs of good and bad luck
CONSERVATOR (Egalitarianism)	Caution does not work	ALLIGNED (No surprises)	Others prosper (especially individualistic strategists)	Others prosper (especially hierarchical strategists)
MAXIMIZER (Individualism)	Skill is not rewarded	Total collapse (when none was expected)	ALLIGNED (No surprises)	Partial collapse
MANAGER (Hierarchy)	Unpredictability	Total collapse (when only partial was expected)	Competition	ALLIGNED (No surprises)

Fig. 12.2 Surprises arising from actual world (*column*) for individual/companies who believe in some risk attitude (*row*)

Steering ERM program again advocates a middle path; this may mean the firm sustains too much damage to be positioned to take full advantage of the market when it turns. When times are Uncertain, a firm following a Risk Steering ERM program will be frustrated by frequent surprises and a world that does not quite fit the model. Competitors not tied to a particular view of risk will fare better, making decisions in the moment with maximum flexibility.

Why do corporations adhere to a particular risk perspective? The firm may have been formed during an environment aligned with their perspective. Alternatively, the company may have suffered traumatic damage during a period of dissonance between an old perspective and the risk environment and then made a shift, perhaps under the direction of new leadership. The firm may have been wildly successful at some point in the past, and now cling stubbornly to the strategy that worked for them then. Corporate culture tends to be self-perpetuating: individuals are drawn to employers with a perspective that makes sense to them—and those in a position to make hiring decisions typically prefer to hire staff whose views mesh with their own.

Yet there are always companies who are following strategies that are not well aligned with the environment. Some of these firms muddle along with indifferent results and survive until their preferred environment comes back. Others sustain enough damage that they do not survive; some change their risk perspective and risk strategy to take advantage of the new environment. Meanwhile, new firms enter the market with risk perspectives and risk strategies that are aligned with the current environment.

Since many of the poorly aligned firms shrink, die out, or change perspective—and since new firms tend to be well-aligned with the current risk regime—the market as a whole adjusts to greater alignment with the risk environment via a process of "natural selection". Besides, the feedback effect presented earlier with the Millennium Bridge analogy reinforces the strength of the oscillations between the different regimes. Having the right attitude at the right time is not an easy task and contributes not only to reduce risk but also to increase average profit. With students from Cornell, we have revisited Ingram and Thompson's surprise game with simple rules for the main US stock that govern investors' decisions. Not only do we observe cycles as expected, but what is particularly impressive is that using simple, automatic rules about changing risk attitudes enables us to obtain a significant value of alpha, that is a significant overperformance of the dynamic investment strategy. The automatic rules for transitions from some attitude to another one in the different risk environments are formalized with the following transition matrices (Fig. 12.3).

This must of course be relativized as one does not know what would have happened if most stakeholders had followed this strategy at the same time in the real world. Perhaps the impact of feedback loops would have downgraded the profit of the investor. Nevertheless, this small exercise nicely highlights the advantage of adopting the right risk attitude at the right time.

Fig. 12.3 Transition matrices between different risk attitudes for the four risk environments

Boom: (>20%)

	Max	Man	Prag	Cons
Max	99	0.8	0.15	0.05
Man	40	55	4	1
Prag	20	30	45	5
Cons	5	20	10	65

Moderate: (10% - 20%)

	Max	Man	Prag	Cons
Max	90	8	1.5	0.5
Man	3.5	95	0.75	0.75
Prag	3	15	80	2
Cons	2	10	3	85

Bust: (< -10%)

	Max	Man	Prag	Cons
Max	65	20	10	5
Man	1	55	4	40
Prag	5	30	45	20
Cons	0.05	0.8	0.15	99

Unknown: (-10% - 10%)

	Max	Man	Prag	Cons
Max	80	10	8	2
Man	2.5	85	10	2.5
Prag	2	5	90	3
Cons	0.5	1.5	3	95

12.7 Attitudes with Respect to Models

The development of internal models in the insurance industry is accompanied by the blossoming of diverse attitudes of employees and decision-makers with respect to models. How can we characterize attitudes with respect to models? One postulate is that risk attitudes may explain most of the attitudes of agents with respect to models. Therefore, we propose a taxonomy for attitudes with respect to models that is similar to the one for attitudes with respect to risk.

Maximizers tend to think that models often overestimate some risks and prevent underwriters from running their business appropriately, because of too large safety margins.

Conservators tend to think that the model does not take into account extreme risks and some other top risks well enough and underestimate some risks, justifying some risk taking that should not be allowed in reality.

Managers think that models enable the entity to optimize the risk-return of the company in a certain confidence zone, when combined with opinions from experienced practitioners. Two problems may arise: the reluctance of experienced practitioners to accept conclusions coming from analytics (like in the book *Moneyball: the art of winning an unfair game*[1]) as well as the overconfidence in the results of the model or the inability to explain them due to their complexity.

Pragmatists do not believe it is possible to model the world and they think that financial and insurance variables are not amenable to forecasting analysis. Therefore, from their point of view, models represent both a loss of time, money and energy. They would prefer that the company focuses on preparing to adapt its strategy and business model to unavoidable changes in the insurance and financial markets.

With this taxonomy of attitudes with respect to models, one immediate question is the following: do individuals with conservator attitude with respect to risk also present conservator attitude with respect to models? Do we observe a strong correlation between attitudes with respect to risk and attitudes with respect to models?

12.8 Surveying Attitudes with Respect to Risk and Models

We present the design and the conclusions drawn from a survey that aims at determining the risk attitudes and the attitudes with respect to models of an individual. This online survey builds on an initial questionnaire developed by David

[1]***Moneyball: The Art of Winning an Unfair Game*** (2003) is a book by Michael Lewis, published in 2003, about the Oakland Athletics baseball team and its general manager Billy Beane. Its focus is the team's analytical, evidence-based, sabermetric approach to assembling a competitive baseball team, despite Oakland’s disadvantaged revenue situation. A film based on the book starring Brad Pitt was released in 2011.

Ingram. After questions about the participant's characteristics, including number of years of experience and type of position, the participant has to strongly disagree (score = −2), disagree (score = −1), be neutral (score = 0), agree (score = +1) or strongly agree (score = +2) with forty statements. Each risk attitude is represented by ten statements that are supposed to match the corresponding risk perspective. There are also twenty-four statements, created by Ingram, Loisel and the Cardif Chair team, on attitudes with respect to models, six per considered attitude. To determine the risk attitude of a participant, we use a simple scoring technique and sum the scores for the corresponding questions. If the score is 5 or above, then we conclude that the individual agrees with the corresponding risk attitude. We treat pragmatists differently, who are individuals who do not really believe in any of the three other perspectives. Some individuals (roughly 50 %) present so-called blended risk attitudes, like Maximizer/Manager. We only keep the blended attitudes if the difference between the two scores is not too large. Some other retreatments are necessary and are not detailed here, please see Clot et al. (2015a) for details.

The main conclusion of this survey and of the initial one conducted by David Ingram is that risk attitudes differ in the US and in Europe. The maximizer point of view is more represented in the US than in Europe, and the conservator point of view is more represented in Europe than in the US. This is not surprising and confirms intuition. Another interesting conclusion of this survey is that risk attitudes are very different from one position to another, and also depend on the number of years of experience. Detailed results are available in Clot et al. (2015a). Understanding these different risk attitudes may be useful to ease discussions and collaborations between different stakeholders with different positions or origins.

Attitudes with respect to models do not seem to be as directly linked with attitudes with respect to risk as one could imagine. We observe some common points with attitudes with respect to risk: people from different continents, with different positions, have different attitudes with respect to models. However, we note a great proportion of manager attitudes with respect to models, which may be due to the high proportion of actuaries taking the survey. This could be considered as a signal of overconfidence of actuaries in models, as the manager attitude is the only one that is positive with respect to models. However, statements corresponding to this attitude with respect to models state that models can be useful when they are complemented by experience and when their limits are studied carefully, which seems more reasonable. Some other important conclusions of the study of Clot et al. (2015b) are that individuals who have experienced the same events tend to have correlated attitudes with respect to models. Besides, one potential interpretation of the results of the study is that some individuals who experience outcomes that do not match their risk attitude may first change their model attitude into pragmatist instead of changing their risk attitude: it is easier to blame the model than to blame oneself…

12.9 Some Important Remarks

Risk attitudes may be influenced by salary incentives and by responsibilities in case of failure. We all know some traders and underwriters who behave like maximizers with the money of the company and like managers with their own money, and we all know top managers who act as managers with their own money and as conservators with the money of the company, because their multi-year bonus could be lost in case of a large claim. This may lead them to purchase an aggregate stop-loss for the overall business for example, even if this is far from being optimal for the company from an economic point of view. Similar decisions may come from the fear of a trial in the case of a very unlikely event.

Note that during the debriefing of this experiment, it has been pointed out several times that attitudes with respect to risks and models are somehow time and context dependent. It will be interesting to test the evolution of risk attitudes and model attitudes over time, with repetitions of the survey.

It is a good thing to have individuals with different attitudes in an executive committee to reach reasonable decisions. However, some problems can arise when those attitudes strongly oppose one another. Understanding these different attitudes can be the first step of a conciliation process before efficient decision-making.

To conclude, we mention that this classification is far from being an exact science. Many stakeholders present mixed characteristics (like maximizer-manager, for example). Running this survey about risk attitudes and attitudes with respect to models in a management team must be seen as a nice way to initiate a fruitful discussion about risk, internal risk models and ERM. Note that a similar study would be interesting concerning attitudes with respect to analytics and big data.

Bibliography

Acpr. 2014. Préparation à Solvabilité II : enseignements de l'exercice d'ORSA pilote. *Information Document*. Version 1.0.
Adam A., Houkari M., Laurent J.P. 2009. Hedging interest rate margins on demand deposits. *Available at SSRN 1361660*.
Agenos X. 2010. Appétit pour le risque et gestion stratégique d'une société d'assurance non-vie. *Actuary Memorandum*. CEA.
Ahlgrim KC., D'Arcy SP., Gorvett RW. 2005. Modeling Financial Scenarios: A Framework for the Actuarial Profession. *Proceedings of the Casualty Actuarial Society* 92.
Aït-Sahalia Y., Hansen L.P., Scheinkman J.A. 2008. Operator methods for continuous-time Markov processes. *Handbook of financial econometrics*. 1: 1–66.
Amblard G., Lebuchoux J. 2000. Models for CMS caps. *Risk-London-Risk magazine limited-*. 13 (9; SUPP/1): 6–7.
Andersen L., Piterbarg V. 2010. Interest Rate Modelling. *Atlantic Financial Press.,* volumes 1, 2 & 3.
Armel K., Planchet F., Kamega A. 2011. Quelle structure de dépendance pour un générateur de scénarios économiques en assurance ? *French Actuarial Bulletin*. Vol. 11, No. 22.
Avellaneda M., Buff R., Friedman C., Grandechamp N., Kruk L., Newman J. 2001. Weighted Monte Carlo: A new technique for calibrating asset-pricing models. *International Journal of Theoretical and Applied Finance*, 4 (01): 91–119.
Azzaz J., Loisel S., Thérond P.E. 2015. Some characteristics of an equity security next-year impairment. *Review of Quantitative Finance and Accounting*, 45(1):111–135.
Bachelier L. 1900. Théorie de la spéculation. Annales de l'Ecole Normale Supérieure, 3e série :21–86.
Bacinello A.R. 2001. Fair pricing of life insurance participating policies with a minimum interest rate guaranteed. ASTIN Bulletin, 31(2):275–87.
Bakshi G., Cao C., Chen Z. 1997. Empirical Performance of Alternative Option Pricing Models. *Journal of Finance*. Vol. 52, Issue 5.
Bakshi G., Madan D. 2000. Spanning and derivative-security valuation. *Journal of Financial Economics*. *55* (2): 205–238.
Ballotta L. 2005. A Lévy process-based framework for the fair valuation of participating life insurance contracts. *Insurance: Mathematics and Economics*, 37(2):173–196.
Barth M.E. 2006. Including estimates of the future in today's financial statements. *Accounting Horizons*, 20(3):271–285.
Basel Committee on Banking Supervision. Revisions to the securitisation framework. *Basel III Document*. December 2014. https://www.bis.org/bcbs/publ/d303.pdf.
Basel Committee on Banking Supervision. 2013. Foundations of the Proposed Modified Supervisory Formula Approach. http://www.bis.org/publ/bcbs_wp22.htm.
Basel Committee. 2005. An Explanatory Note on the Basel II IRB Risk Weight Functions. http://www.bis.org/bcbs/irbriskweight.pdf.

J.-P. Laurent et al. (eds.), *Modelling in Life Insurance – A Management Perspective*,
EAA Series, DOI 10.1007/978-3-319-29776-7

Basel Committee on Banking Supervision. 2014. Foundations of the standardised approach for measuring counterparty credit risk exposures. August, http://www.bis.org/publ/bcbs_wp26.htm.

Bauer D., Bergmann D., Reuss A. 2010. Solvency II and nested simulations—a least-squares Monte Carlo approach. In *Proceedings of the 2010 ICA congress*.

Bauer D., Reuss A., Singer D. 2012. On the calculation of the solvency capital requirement based on nested simulations. *Astin Bulletin*. 42 (02): 453–499.

Benston G.J. 2006. Fair-value accounting: A cautionary tale from Enron. *Journal of Accounting and Public Policy*, 25(4):465–484.

Berdin E., Grundl H. 2015. Effects of a low interest rate environment on life insurers. *The Geneva Papers on Risk and Insurance*—Issues and Practice, 40(3):385–415.

Bernard C., Jiang X. Wang R. 2014. Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, 54, 93–108.

Beutner E., Pelsser A., Schweizer J. 2015. Theory and Validation of Replicating Portfolios in Insurance Risk Management. *Available at SSRN 2557368*.

Black F., Scholes M. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (3): 637–654.

Boekel P., Van Delft L., Hoshino T., Ino R., Reynolds C., Verheugen H. 2009. Replicating portfolios. An introduction: Analysis and illustrations.

Bonnin F., Combes F., Planchet F., Tammar M. 2014. Un modèle de projection pour des contrats de retraite dans le cadre de l'ORSA. *Les cahiers de recherche de l'ISFA*. No. 2014.1.

Bonnin F., Combes F., Planchet F., Tammar M. 2014b. Un modèle de projection pour des contrats de retraite dans le cadre de l'ORSA. *Bulletin Français d'Actuariat*. vol. 14, No. 28.

Bonnin F., Juillard M., Planchet F. 2014a. Best Estimate Calculations of Savings Contracts by Closed Formulas—Application to the ORSA. *European Actuarial Journal*. Vol. 4, Issue 1: 181–196. http://dx.doi.org/10.1007/s13385-014-0086-z.

Bonnin F., Planchet F. 2013. Engagement best estimate d'un contrat d'épargne en Euro.*La Tribune de l'Assurance (rubrique "le mot de l'actuaire")*. N°185 of 01/11/2013.

Bonnin F., Planchet F., Juillard M. 2014. Best estimate calculations of savings contracts by closed formulas: application to the ORSA. *European Actuarial Journal*, **4**(1): 181–196.

Borel-mathurin F., Darpeix P.-E., Guibert Q., Loisel S. 2015. Main Determinants of Profit Sharing Policy in the French Life Insurance Industry.

Botvinnik A., Hoerig M., Ketterer F., Tazov A., & Wechsung F. 2014. Replicating portfolios revisited.

Breeden D.T., Litzenberger R. H. 1978. Prices of state-contingent claims implicit in option prices. *Journal of business:* 621–651.

Brennan M.J., Schwartz E.S. 1976. The pricing of equity-linked life insurance policies with an asset value guarantee. *Journal of Financial Economics*, 3(3):195–213.

Brigo D., Mercurio F. 2007. Interest rate models-theory and practice: with smile, inflation and credit. Springer.

Briys E., de Varenne F. 1997. On the risk of insurance liabilities: Debunking some common pitfalls. *The Journal of Risk and Insurance*, 64(4):673–694.

Broadie M., Du Y., Moallemi C.C. 2011. Efficient risk estimation via nested sequential simulation. *Management Science*, *57*(6): 1172–1194.

Cairns A. 2000. A discussion of parameter and model uncertainty in insurance. *Insurance: Mathematics and Economics*, 27, 313–330.

Cairns A.J.G., Blake D., Dowd K. 2008. Modelling and management of mortality risk: an overview. *Scandinavian Actuarial Journal*, 70–113.

Caja A., Planchet F. 2010. La mesure du prix de marché du risque: quels outils pour une utilisation dans les modèles en assurance ? *Insurance and Risk Management*. Vol.78 (3/4).

Carr P., Madan D. 2001. Towards a theory of volatility trading. *Option Pricing, Interest Rates and Risk Management, Handbooks in Mathematical Finance:* 458–476.

CEIOPS. 2009. Advice for Level 2 Implementing Measures on Solvency II: Article 86f. Standards for Data Quality, CEIOPS-DOC-37/09.

Chen X., Nelson B.L., & Kim K.K. 2012. Stochastic kriging for conditional value-at-risk and its sensitivities. In *Simulation Conference (WSC), Proceedings of the 2012 Winter*: 1–12. IEEE.

Chen X., Kim K.K. 2013. Building metamodels for quantile-based measures using sectioning. In *Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World*: 521–532. IEEE Press.

Clot D., Ingram D., Loisel S., Olympio A. 2015a. Risk attitudes in the US and in Europe, pre-preprint.

Clot D., Ingram D., Loisel S., Olympio A. 2015b. Do risk attitudes determine attitudes with respect to internal models? pre-preprint.

Cont R. 2006. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16, 519–554.

Dal Moro E., Faulkner J. 2014. The computational and timing challenge of quarterly nonlife (re) insurance liability evaluation under IFRS 4 phase 2. *The Journal of Financial Perspectives*, *2*(3).

Danielsson J., James K., Valenzuela M. Zer I. 2014. Model risk of risk models. Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington, D.C.

Darolles S., Laurent J.P. 2000. Approximating payoffs and pricing formulas. *Journal of Economic Dynamics and Control*, *24*(11): 1721–1746.

Derman E. 1996. Model risk. Goldman Sachs Quantitative Strategies Research Notes.

Derman E. 2001. The principles and practice of verifying derivatives prices.

Derman E. 2011. Models. Behaving. Badly: why confusing illusion with reality can lead to disaster, on Wall Street and in life. Wiley.

Derman E., Ergener D., Kani I. 1995. Static options replication. *The Journal of Derivatives*, *2*(4): 78–95.

Devineau L., Loisel S. 2009. Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula? *Bulletin Français d'Actuariat*, 9(18): 107–145.

Djehiche B., Hörfelt P. 2005. Standard approaches to asset & liability risk. Supported by the Swedish Insurance Federation. *Scandinavian Actuarial Journal*. 2005(5): 377–400.

Dupire B. 1994. Pricing with a smile. *Risk*, *7*(1): 18–20.

Eiopa-BoS-13/25. 2013. Explanatory text on the proposal for guidelines on forward looking assessment of the undertaking's own risks (based on the ORSA principles).

Eiopa -14-322. 2014. The underlying assumptions in the standard formula for the Solvency Capital Requirement calculation.

Eiopa. 2014. Technical Specification for the Preparatory Phase. No. EIOPA-14/209. Frankfurt. EIOPA.

Eling M., Holder S. 2012. Maximum Technical Interest Rates in Life Insurance in Europe and the United States: An Overview and Comparison. *The Geneva Papers on Risk and Insurance Issues and Practice*.

Eling M., Schmeiser H., Schmit J.T. 2007. The Solvency II process: Overview and critical analysis. *Risk Management and Insurance Review*, *10*(1): 69–85.

Ellsberg D. 1961. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643–669.

Embrechts P., Puccetti G., Rüschendorf L. 2013. Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, 37, 2750–2764.

Embrechts P., Wang B., Wang R. 2014. Aggregation-robustness and model uncertainty of regulatory risk measures. To appear in *Finance and Stochastics*.

Emerson D.J., Karim K.E., Rutledge R.W. 2010. Fair Value Accounting: A Historical Review Of The Most Controversial. *Journal of Business & Economics Research*, 8(4):77–86.

Faleh A., Planchet F., Rullière D. 2012. Scénarios économiques et techniques d'allocation d'actif —application aux assurances et aux fonds de pension. Paris*: Economica*.

Félix J.P., Planchet F. 2015 « Calcul des engagements en assurance-vie : quel calibrage 'cohérent avec des valeurs de marché' ? », L'Actuariel, n°16 du 01/03/2015.

Föllmer H., Schied A. 2011. Stochastic Finance, 3rd edition. De Gruyter, Berlin.

Galbraith, J.K. 2007. The New Industrial State. Princeton University Press. (Orignally published 1967).

Gauthier T. 2011. Développement d'un générateur de scénarii économiques au sein d'une compagnie d'assurance-vie. *Mémoire d'actuaire.* ISFA.

Glasserman P., Yu B. 2004. Simulation for American options: regression now or regression later? In *Monte Carlo and Quasi-Monte Carlo Methods 2002*: 213–226. Springer Berlin Heidelberg.

Glasserman P., Xu X. 2014. Robust risk measurement and model risk. To appear *in Quantitative Finance.*

Gordy M., Jones D. 2003. Credit portfolio risk: Random tranches. *Risk-London-Risk magazine limited*, *16*(3): 78–83.

Gourieroux C., Laurent J.P. 1996. Estimation of a Dynamic Hedge. INSEE.

Grigorov V. 2012. ORSA and model risk. Focus on Life insurance. Conférence-débat de l'Institut des Actuaires. Paris, November 30, 2012.

Grosen A., Jorgensen P.L. 2000. Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options and bonus policies. *Insurance: Mathematics and Economics*, 26(1):37–57.

Guibert Q., Juillard M., Nteukam T.O., Planchet F. 2014. Solvabilité Prospective en Assurance. *Méthodes quantitatives pour l'ORSA*. Paris: *Economica.*

Guibert Q., Planchet F. 2014. L'utilisation des actions du management en assurance dépendance. *L'Actuariel*. N°11 du 01/01/2014.

Guibert Q., Juillard M., Planchet F. 2012. Measuring Uncertainty of Solvency Coverage Ratio in ORSA for Non-Life Insurance. *European Actuarial Journal*. 2:205–226. doi:10.1007/s13385-012-0051-7.

Gwilliam D., Jackson, R.H. 2008. Fair value in financial reporting: problems and pitfalls in practice – a case study analysis of the use of fair valuation at Enron. *Accounting Forum*, 32 (3):240–259.

Hagan P.S., West G. 2006. Interpolation methods for curve construction. *Applied Mathematical Finance*, 13(2): 89–129.

Haldane A.G., Madouros V. 2012. The dog and the frisbee. *Revista de Economía Institucional*, *14* (27), 13–56.

Haldeman Jr, R.G. 2006. Fact, Fiction, and Fair Value Accounting at Enron. *CPA Journal*, 76 (11):14–21.

Hansen L.P., Sargent, T.J. 2007. Robustness. Princeton University Press.

Hansen L.P., Jagannathan R. 1997. Assessing specification errors in stochastic discount factor models. *The Journal of Finance*, *52*(2): 557–590.

Heath D., Jarrow R., Morton A. 1992. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*: 77–105.

Hibbert J., Mowbray P., Turnbull C. 2001. A Stochastic Asset Model & Calibration for Long-Term Financial Planning Purposes. Barrie & Hibbert Limited.

Holthausen, R.W., Watts, R.L. 2001. The relevance of the value-relevance literature for financial accounting standard setting. *Journal of Accounting and Economics*, 31(1–3):3–75.

Hull J., White A. 1990. Pricing interest-rate-derivative securities. *Review of financial studies*, 3(4): 573–592.

Ingram, D., Underwood, A. 2012. The ERM Rainbow, *Bulletin Français d'Actuariat*, Vol. 12 (24), 5–14.

Institut des Actuaires. 2014. LORSA : Quelques Exemples de Pratiques actuarielles. Document de travail. Institut des Actuaires.

Internal Models and Solvency II. 2014. From Regulation to Implementation. *Risk books*. Caldoni, P. (ed.).

Jaggi B., Winder J.P., Lee C.-F. 2010. Is There a Future for Fair Value Accounting After the 2008–2009 Financial Crisis ? *Review of Pacific Basin Financial Markets and Policies*, 13 (03):469–493.

Kablau A., Weiss M. 2014. How is the low interest-rate environment affecting the solvency of German life insurers ? Technical Report 27, Bundesbank.

Kamega A., Planchet F., Thérond P.E. 2009. Scénarios économiques en assurance—Modélisation et simulation. Paris: *Economica.*

Kijima M. 1997. Markov processes for stochastic modelling. Vol. 6. CRC Press.

Kling A., Richter A., Ruÿ J. 2007. The interaction of guarantees, surplus distribution, and asset allocation in with-profit life insurance policies. *Insurance: Mathematics and Economics*, 40(1):164–178.

Koivu M., Pennanen T., Ranne A. 2005. Modeling assets and liabilities of a Finnish pension insurance company: a VEqC approach. *Scandinavian Actuarial Journal*, 2005(1): 46–76.

Krugman P. 2008. Return of Depression Economics and the Crisis of 2008. W. W. Norton & Company.

Lagerås A. 2014. How to hedge extrapolated yield curves. arXiv preprint arXiv:1406.6142.

Laïdi Y., Planchet F. 2014. Calibrating LMN Model to Compute Best Estimates in Life Insurance. *The working papers of the ISFA.* No. 2014.13.

Laidi Y. 2013. Problématiques de calibration en vue de l'évaluation des risques de taux, de défaut et de liquidité. *Mémoire d'actuaire*. CNAM.

Laux C., Leuz C. 2009. The crisis of fair-value accounting: Making sense of the recent debate. *Accounting, Organizations and Society*, 34(6–7):826–834.

Laux C., Leuz C. 2010. Did Fair-Value Accounting Contribute to the Financial Crisis ? *Journal of Economic Perspectives*, 24(1):93–118.

Lenhard J. 2004. Scepticism and mathematization: Pascal and Peirce on mathematical epistemology. *Philosophica-Gent-*, *74*, 85.

Leroy G., Planchet F. 2011. Solvabilité 2 : Quels standards pour le risque de marché ? *La Tribune de l'Assurance (section "le mot de l'actuaire").* No. 156 of 01/03/2011.

Li H. 2014. Orthogonal Polynomials and Their Applications in Financial and Actuarial Models. *Doctoral dissertation.* University of Alberta.

Liu M. 2010. Efficient simulation in financial risk management. *Doctoral dissertation*. Northwestern University. http://users.iems.northwestern.edu/~staum/Ming_Liu_Thesis.pdf.

Liu M., Nelson B.L., Staum J. 2010. An efficient simulation procedure for point estimation of expected shortfall. In *Simulation Conference (WSC), Proceedings of the 2010 Winter*: 2821–2831. IEEE.

Loisel S., Milhaud X. 2011. From deterministic to stochastic surrender risk models: Impact of correlation crises on economic capital. *European Journal of Operational Research*, 214 (2):348–357.

Longstaff F.A., Mithal S., Neis, E. 2005. Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance*. 60 (5): 2213–2253.

Longstaff F. A., Schwartz E.S. 2001. Valuing American options by simulation: A simple least-squares approach. *Review of Financial studies*, *14*(1): 113–147.

Madan D.B., Milne F. 1994. Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance*, *4*(3): 223–245.

Maurer R., Rogalla R., Siegelin I. 2013. Participating payout life annuities: lessons from Germany. ASTIN Bulletin, 43(3):159–187.

Merton R.C. 1973. Theory of rational option pricing. *The Bell Journal of economics and management science*: 141–183.

Monin C. 2014. Optimisation multi objectifs de l'allocation d'actifs par un algorithme génétique. *Mémoire d'actuaire.* ISFA.

Morini M. 2011. Understanding and managing model risk: A practical guide for quants, traders and validators. *Wiley Finance*.

Nachman D.C. 1988. Spanning and completeness with options. *Review of Financial Studies*, *1*(3): 311–328.

Natolski J., Werner R. 2014. Mathematical analysis of different approaches for replicating portfolios. *European Actuarial Journal*, *4*(2): 411–435.

Norberg R. 2015. Life insurance mathematics. *Wiley StatsRef: Statistics Reference Online*. 1–19. http://onlinelibrary.wiley.com/doi/10.1002/9781118445112.stat04351.pub2/full.

Nteukam T.O., Planchet F., Ren J. 2014. Internal Model in Life insurance: Application of Least Square Monte-Carlo in Risk Assessment. *Les cahiers de recherche de l'ISFA*. N°2014.12.

Oechslin J., Aubry O., Aellig M., Käppeli A., Brönnimann D., Tandonnet A., Valois G. 2007. Replicating embedded options. *Life & Pension Risk February*.

Orléan A. 2011. L'empire de la valeur. Seuil.

Parkinson P.M. 2011. Guidance on Model Risk Management. Division of Banking Supervision and Regulation for the Board of Governors of the Federal Reserve System.

Pascal B., Krailsheimer A. J. 1968. Pensees: Translated with an Introduction by AJ Krailsheimer. Penguin.

Pelsser A. 2003. Pricing and hedging guaranteed annuity options via static option replication. *Insurance: Mathematics and Economics*, *33*(2):283–296.

Pelsser A., Schweizer J. 2015. The Difference between LSMC and Replicating Portfolio in Insurance Liability Modeling. *Available at SSRN 2557383*.

Piketty T. 2014. *Capital in the Twenty-First Century*. President and Fellows of Harvard College.

Piterbarg V. 2006. Markovian projection method for volatility calibration. *Available at SSRN 906473*.

Planchet F., Bonnin F. 2013. Engagement best estimate d'un contrat d'épargne en Euro. *La Tribune de l'Assurance (section "le mot de l'actuaire")*. No. 185 of 01/11/2013.

Planchet F., Leroy G. 2013. Risque de taux, spread et garanties de long terme. *La Tribune de l'Assurance (section "le mot de l'actuaire")*. No. 178 of 01/03/2013.

Planchet F., Tomas J. 2014. Uncertainty on Survival Probabilities and Solvency Capital Requirement: Application to LTC Insurance. *Scandinavian Actuarial Journal*. Doi: 10.1080/03461238.2014.925496.

Planchet F. 2011. Paramètres de générateur de scénarios économiques, comment choisir? *La Tribune de l'Assurance (section "le mot de l'actuaire")*. No. 164 of 01/12/2011.

Planchet F. 2009b. Provisionnement et couverture des garanties financières : deux notions indissociables. *La Tribune de l'Assurance (section "le mot de l'actuaire")*. No. 138 of 01/07/2009.

Planchet F. 2009a. Quel modèle d'actifs en assurance? *La Tribune de l'Assurance (section "lemot de l'actuaire")*. No. 136 of 01/05/2009.

Planchet F., Thérond P. 2007. Mesure et gestion des risques d'assurance. *Economica*.

Planchet F., Thérond P., Juillard M. 2011. Modèles financiers en assurance. Analyses de risque dynamiques. *Economica*, 2nd edition.

Plantin G., Sapra H., Shin H. S. 2008a. Fair value accounting and financial stability. *Financial Stability Review*, (12):85–94.

Plantin G., Sapra H., Shin H. S. 2008b. Marking-to-market: Panacea or pandora's box? *Journal of Accounting Research*, 46(2):435–460.

Power M. 2010. Fair value accounting, financial economics and the transformation of reliability. *Accounting and Business Research*, 40(3):197–210.

PwC's Financial Services Regulatory practice. 2013. The evolution of model risk management.

Rebonat, R. 2002. Theory and practice of model risk management. Quantitative Research Centre (QUARC) of the Royal Bank of Scotland. Oxford Financial Research Centre.

Revelen J. 2009. Replicating Portfolio et capital économique en assurance vie. *Mémoire d'ingénieur*. ISFA, École Centrale de Lyon.

Richard J., Bensadon D., Collette C. 2014. Comptabilité financière. IFRS versus normes françaises. Dunod, 10th edition.

Robert C., Therond P. 2014. Distortion risk measures, ambiguity aversion and optimal effort. *Astin Bulletin*, 44, 277–302.

Ross S.A. 1976. Options and efficiency. *The Quarterly Journal of Economics:* 75–89.

Ross S.A. 1978. A simple approach to the valuation of risky streams. *Journal of business*: 453–475.

Schrager D. 2008. Replicating portfolios for insurance liabilities. *Actuarial Sciences*.

Séjourné B. 2006. Pourquoi le comportement des épargnants français est-il si peu conforme à la théorie traditionnelle du portefeuille ? Les cahiers scientifiques 1, Autorité des Marchés Financiers.

Séjourné B. 2007. Comment les épargnants français intègrent-ils le paramètre temps dans la gestion de leur portefeuille ? Les cahiers scientifiques 4, Autorité des Marchés Financiers.

Sepp A. 2003. Fourier transform for option pricing under affine jump-diffusions: An overview. Unpublished Manuscript, available at http://www.hot.ee/seppar.

Sibbertsen P., Stahl G., Luedtke C. 2008. Measuring model risk. *Journal of Risk Model Validation*, 2, 65–81.

Smith A., Wilson T. 2001. Fitting yield curves with long term constraints. *Technical report.*

Stentoft L. 2004. Convergence of the least squares Monte Carlo approach to American option valuation. *Management Science*, *50*(9): 1193–1203.

Stricker M., Wang S., Strommen S.J. 2014. Model validation for insurance enterprise risk and capital models. Sponsored by CAS, CIA, SOA Joint Risk Management Section.

Tchundjang Pouemi J. 1979. Monnaie, servitude et liberté. La répression monétaire de l’Afrique. Yaoundé. Editions Menaibuc.

Thérond P. 2008. Ifrs, solvabilité 2, embedded value : quel traitement du risque ? *Bulletin Français d'Actuariat*, 8(15):64–94.

Thompson, M., Ellis, R., Wildavsky, A. 1990. Cultural theory. Political cultures. *Boulder, CO, US: Westview Press Cultural theory.* xvi 296 pp.

Turing A.M. 1950. Computing machinery and intelligence. *Mind*, 433–460.

Vedani J., Ramaharobandro F. 2013. Continuous compliance: a proxy-based monitoring framework. *arXiv preprint arXiv*: *1309.7222*.

Walter C., Brian E. 2008. Critique de la valeur fondamentale. Springer.

Wilkie A.D. 1986. A Stochastic Investment Model for Actuarial Use. T. F. A. 39.

Ziemba W., Mulvey J. 1998. The Towers Perrin Global Capital Market Scenario Generation System: CAP:Link. *World Wide Asset and Liability Modeling (eds.)*. Cambridge University Press. 1998 (with A. Eric Thorlacius).