EAA Series

Mario V. Wüthrich

Market-Consistent Actuarial Valuation

Third Edition

european actuarial academy

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

Mario V. Wüthrich

Market-Consistent Actuarial Valuation

Third Edition

Springer

Mario V. Wüthrich
RiskLab, Department of Mathematics
ETH Zurich
Zürich
Switzerland

ISSN 1869-6929 ISSN 1869-6937 (electronic)
EAA Series
ISBN 978-3-319-46635-4 ISBN 978-3-319-46636-1 (eBook)
DOI 10.1007/978-3-319-46636-1

Library of Congress Control Number: 2016954541

Mathematics Subject Classification (2010): 91B30, 97M30, 62P05

The first and the second edition were published under Mario Wüthrich, Hans Bühlmann and Hansjörg Furrer.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface to the Third Edition

Since the first lecture in 2004/2005 at ETH Zurich, market-consistent actuarial valuation has become the standard framework in insurance valuation, risk management and statistical modelling. It is the basis of a consistent economic balance sheet, and it is also the basis of risk-based solvency considerations such as the Swiss Solvency Test and Solvency II. These legal and regulatory developments have also led to changes in our perception of the field. The main improvement of the third edition compared to the previous editions is that we have given an update on recent developments and changes.

In Chap. 4, we elaborate more on different risk measures. Our original solvency definition was rather restrictive. In the light of the Swiss Solvency Test and Solvency II, we relax our restrictive definition to more practical risk measures which are used in the industry. We also provide a comparison between our initial definition and industry practice.

The biggest changes concern Chap. 5 on non-life insurance modelling. Over the last ten years, the viewpoint of non-life insurance reserving risk has fundamentally changed from a static view towards a dynamic view. This change required a rewrite of big parts of this chapter in order to adapt the valuation framework to this dynamic view.

Most importantly, I want to express my most sincere gratitude to my former co-authors of these notes Hans Bühlmann and Hansjörg Furrer. The foundations of these lecture notes go back to Hans and Hansjörg, and their ideas and our conversations have always been a great source of inspiration. I very much regret that they did not have a sufficient amount of time to contribute to this latest edition.

Finally, I would like to kindly thank Philippe Deprez for his very beneficial support in preparing this new edition. His mathematical comments have substantially enhanced this manuscript and his graphical skills have improved the illustrations. My final thanks go to the many students who have attended this lecture.

Zürich
July 2016

Mario Wüthrich

Preface to the Second Edition

The financial crisis of 2007–2010 has shown that the topic of market-consistent valuation and solvency has nothing lost of its topicality. On the contrary, it has shown that we need a much deeper understanding of the models used, their limitations, etc., in order to model real world problems. In this spirit, the first edition of these lecture notes has initiated a very active discussion among academics and practitioners about actuarial modelling and the use of models.

Since the first edition of these notes this course was again held at ETH Zurich in 2008 and 2010. Moreover, we have also presented part of these notes in various European countries, such as Germany, the UK, France, The Netherlands and Sweden. These presentations have stimulated several interesting discussions which we have implemented into the new version. The main new features are:

In Chapter 2, we elaborate on the separation of financial deflators and probability distortions. For the financial deflator, we then give a simple explicit example in terms of the discrete time Vasiček model. Probability distortions on the other hand can be understood in various ways. We give different examples that lead to the Esscher premium, to the cost-of-capital loading for expected shortfall and to first order life tables (in Chapter 3).

In Chapter 3, we introduce the approximate valuation portfolio which is useful in the case where we are not able to construct an exact valuation portfolio. This is done using selected scenarios evaluated with the help of an appropriate distance function. This is in line with the state-of-the-art concepts used in life insurance practice.

Finally, in Chapter 6, we add two sections that discuss losses and gains from insurance technical risks. This is closely related to the actual discussion of the claims development result in non-life insurance, but of course also applies to life insurance problems.

Zürich
May 2010

Mario Wüthrich
Hans Bühlmann
Hansjörg Furrer

Preface to the First Edition

The balance sheet of an insurance company is often difficult to interpret. This derives from the fact that assets and liabilities are measured by different yardsticks. Assets are mostly valued at market prices; liabilities—as far as they relate to contractual obligations to the insured—are measured by established actuarial methods. Since, in general, there does not exist a market for trading insurance policies, the question arises how these actuarial methods need to be changed to give values—as if these markets existed. The answer to this question is "Market-Consistent Actuarial Valuation". These lecture notes explain the logical mathematical framework that leads to market-consistent values for insurance liabilities.

In Chapter 1, we motivate the use of market-consistent values. Solvency requirements by regulators are one major reason for it.

Chapter 2 introduces stochastic discounting, which in a market-consistent actuarial valuation framework replaces discounting with the classical technical interest rate. In this chapter, we introduce the notion of "Financial Variables" (which follow the laws of financial markets) and the notion of "Technical Variables" (which are purely dependent on insurance events).

In Chapter 3, the concept of the "Valuation Portfolio" (VaPo) is introduced and explained in the life insurance context. The basic idea is not to consider liabilities in monetary values but in units, which are appropriately chosen financial instruments. For life insurance products, this choice is quite natural. The risk due to technical variables is included in the protected (against technical risk) VaPo, denoted by VaPo^{prot}.

Financial risk is treated in Chapter 4. It derives from the fact that the actual investment portfolio of the insurance company differs from the VaPo^{prot}. Ways to control the financial risk are—among others—derivative securities such as Margrabe Options and/or (additional) Risk Bearing Capital.

In Chapter 5, the notion of the Valuation Portfolio (VaPo) and the protected (against technical risk) Valuation Portfolio (VaPo^{prot}) is extended to the non-life insurance sector. The basic difference to life insurance derives from the fact that in property and casualty insurance the technical risk is much more important. The discussion of appropriate risk measures (in particular the mean square error of prediction) is therefore a central issue.

The final Chapter 6 contains selected topics. We mention only the treatment of the “Legal Quote” in life insurance.

These lecture notes stem from a course on market-consistent actuarial valuation, so far given twice at ETH Zürich, namely in 2004/2005 by HB and HJF and in 2006 by MW and HJF. MW has greatly improved on the first version of these notes. But obviously this version is not to be considered as final. For this reason, we are grateful that Springer’s newly created EAA Lecture Notes series has given us the opportunity to share these notes with many friends and colleagues, whom we invite to participate in the process of discussion and further improvement of the present text, and to help clarify our way of understanding and modelling.

The authors wish to thank Professor Paul Embrechts for his interest and constant encouragement while they were working on this project. His support has been a great stimulus for us.

It is also a great honour for us that our text appears as the first volume of Springer’s newly founded EAA Lecture Notes series. We are grateful to Peter Diethelm, who as Managing Director has been the driving force in getting this series started.

Zürich
May 2007

Mario Wüthrich
Hans Bühlmann
Hansjörg Furrer

Contents

Chapter 1
Introduction

1.1 The Three Pillar Approach

Recent years have shown that (financial) companies need to have a good executive board, a good business strategy, a good financial strength and a sound risk management practice in order to survive financial distress periods. It is essential that the risks are known, assessed, controlled and, whenever possible, quantified and mitigated by the management and the respective specialized units.

Especially in the past few years, we have observed several failures of financial companies (for example, Barings Bank, HIH Insurance Australia, Lehman Brothers, Washington Mutual, etc.). Many companies have faced severe solvency and liquidity problems. As a consequence, supervision and politics have started several initiatives to analyze these problems and to improve qualitative and quantitative risk management within the companies (Basel II/III, Solvency II and local initiatives like the Swiss Solvency Test [SST06], for an overview we refer to Sandström [Sa06, Sa07]). The financial crisis of 2007–2010 and the European sovereign debt crisis have also shown that there is still a long way to go (leading economists view these last financial crises as the worst ones since the Great Depression of the 1930s). These crises have (again) shown that risks need to be understood and managed properly on the one hand, and on the other hand that mathematical models, their assumptions and their limitations need to be well-understood in order to analyze and solve real world problems.

Concerning insurance companies: the goal of the solvency initiatives is to protect the policyholder and/or the (injured) third party, respectively, from the *consequences* of an insolvency of an insurance company. Henceforth, in most cases it is not primarily the objective of the regulator to avoid insolvencies of insurance companies, but given that an insolvency of an insurance company has occurred, the regulator has to ensure that all (insurance) liabilities are covered with assets and can be fulfilled in an appropriate way (this is not the shareholder's point of view).

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1_1

An early special project was carried out by the “London working group”. The London working group analyzed 21 cases of solvency problems (actual failures and ‘near misses’) in 17 European countries. Their findings can be found in the famous Sharma Report [Sha02]. The main lessons learned are:

- in most cases bad management was the source of the problem;
- there was a lack of a comprehensive risk landscape;
- the company had misspecified business strategies which were not adapted to local situations.

From this perspective, what can we really do?

Sharma says: “Capital is only the second strategy of defense in a company, the first is a good risk management”.

Supervision has started several initiatives to strengthen the financial basis and to improve risk management thinking within the industry and companies (e.g. Basel II/III, Solvency II, Swiss Solvency Test [SST06], Laeven–Valencia [LV08], Besar et al. [BBCMP11]). Most of the new approaches and requirements are formulated in three pillars:

1. Pillar 1: Minimum financial requirements (quantitative requirements)
2. Pillar 2: Supervisory review process, adequate risk management (qualitative requirements)
3. Pillar 3: Market discipline and public transparency

Consequences are that regulators, academics as well as actuaries, mathematicians and risk managers of financial institutions are searching for appropriate solvency regulations. These regulations should be **risk-adjusted** and they should be based on a **market-consistent valuation** of the balance sheet (full balance sheet approach).

From this perspective we derive the valuation portfolio which reflects a market-consistent actuarial valuation of the balance sheet (which is more a static view). Moreover, we describe the uncertainties within that portfolio which corresponds to a risk-adjusted dynamic analysis of the assets and liabilities.

1.2 Solvency

The International Association of Insurance Supervisors IAIS [IAIS05] defines solvency as follows:

“the ability of an insurer to meet its obligations (liabilities) under all contracts at any time. Due to the very nature of insurance business, it is impossible to guarantee solvency with certainty. In order to come to a practicable definition, it is necessary to make clear under which circumstances the appropriateness of the assets to cover claims is to be considered, ...”.

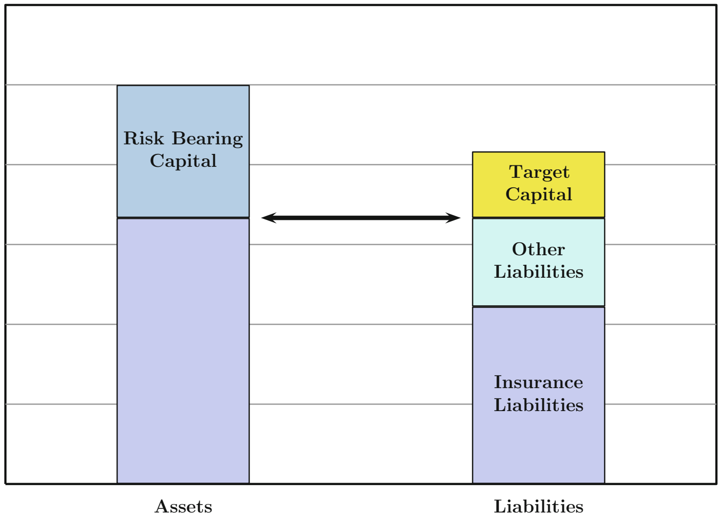


Fig. 1.1 Balance sheet of an insurance company with solvency relation (1.1)

The aim of solvency is to protect the policyholder and/or the (injured) third party, respectively. As it is formulated in Swiss law: it is not the main objective of the regulator to avoid insolvencies of insurance companies, but in case of an insolvency the policyholder's demands must still be met. Avoiding insolvencies must be the main task of the management and the board of an insurance company. Moreover, hopefully, modern risk-based solvency requirements ensure more stability of the financial system.

In this lecture we give a mathematical approach and interpretation to the solvency definition of the IAIS [IAIS05].

Let us start with two definitions (see also Fig. 1.1):

1. **Available Solvency Surplus** (see [IAIS05]) or **Risk Bearing Capital** RBC (see [SST06]) is the difference between the market-consistent value of the assets minus the market-consistent value of the liabilities. This corresponds to the Available Risk Margin, the Available Risk Capacity or the Financial Strength of a company.
2. **Required Solvency Margin** (see [IAIS05]) or **Target Capital** TC (see [SST06]) is the required risk capacity (from a regulatory point of view) in order to be able to run the business such that certain adverse scenarios are also covered (see solvency definition of the IAIS [IAIS05]). This is the Necessary Risk Capacity, Required Risk Capacity or the Minimal Financial Requirement for writing well-specified business.

The general solvency requirement is then given by:

$$\mathrm{TC} \overset{!!!}{\leq} \mathrm{RBC}, \tag{1.1}$$

that is, the underlying risk (measured by TC) needs to be dominated by the available risk capacity (measured by RBC). Otherwise, if (1.1) is not satisfied, the authorities force the company to take certain actions to improve the financial strength or to reduce the risks within the company, such as: write less risky business, purchase protection (like derivative securities or reinsurance), sell part of the business or even close the company, and make sure that another company guarantees the smooth run-off of the liabilities.

1.3 From the Past to the Future

A historical overview of solvency requirements can be found in Sandström [Sa07]. In the past, the evaluation of the Risk Bearing Capital RBC was not based on market-consistent valuation techniques of assets and liabilities nor was the Target Capital TC calculated in a risk-based manner. For example, Solvency I regulation in non-life insurance had the simple form

$$\text{Target Capital TC } = 16\,\% \text{ of premium,} \tag{1.2}$$

and in traditional life insurance it was essentially given by

$$\begin{aligned}\text{Target Capital TC} = {} & 4\,\% \text{ of the mathematical reserves (financial risk)}\\ & + 3\,‰ \text{ of capital at risk (technical risk).}\end{aligned} \tag{1.3}$$

These solvency regulations are very simple and robust, easy to understand and to use. They are rule-based but not risk-based. As such, they are not tailored to the specifics of the written business and neglect the differences between the asset and the liability profiles. Moreover, risk mitigation techniques such as reinsurance are only allowed to a limited extent as eligible elements for the solvency margin.

Our goal in this lecture is to give a mathematical theory to a market-consistent valuation approach. Moreover, our model builds a bridge of understanding between actuaries and asset managers. In the past, actuaries were responsible for the liabilities on the balance sheet and asset managers were concerned with the active side of the balance sheet. But these two parties do not always speak the same language, which makes it difficult to design a successful asset and liability management (ALM) strategy. In this lecture we introduce a language which allows actuaries and asset managers to communicate in a successful way, leading to a canonical risk-adjusted full balance sheet approach to the solvency problem.

1.4 The Full Balance Sheet Approach

A typical balance sheet of an insurance company looks as follows:

Assets	Liabilities
cash and cash equivalents	deposits
debt securities	policyholder deposits
bonds	reinsurance deposits
loans	borrowings
mortgages	money market
real estate	hybrid debt
equity	convertible debt
equity securities	insurance liabilities
private equity	mathematical reserves
investments in associates	(claims) reserves
hedge funds	premium reserves
derivatives	derivatives
futures, swaptions, options	
insurance and other receivables	insurance and other payables
reinsurance assets	reinsurance liabilities
property and equipment	employee benefit plan
intangible assets	provisions
goodwill	
deferred acquisition costs	
income tax assets	income tax liabilities
other assets	other liabilities

It is necessary that assets and liabilities are measured in a consistent way. Market values have no absolute significance, depending on the purpose other values may be better (for example statutorial values). But market values guarantee the switching property (at market prices).

Applications of these lectures are found in the following:

- pricing and reserving of insurance products,
- value-based management tools, dynamical financial analysis tools,
- risk management tools,
- for solvency purposes, which are based on a market-consistent valuation,
- finding prices for trading insurance policies and for loss portfolio transfers.

In the past actuaries mostly used deterministic models for discounting liabilities. As soon as interest rates are assumed to be stochastic, life becomes much more complicated. This is illustrated by the following example. Let $r > 0$ be a (non-trivial) stochastic interest rate, then (by Jensen's inequality applied to the convex function $u(x) = (1+x)^{-1}$)

$$1 = E\left[\frac{1+r}{1+r}\right] \neq E\,[1+r]\;\; E\left[\frac{1}{1+r}\right] > 1, \tag{1.4}$$

that is, in a stochastic environment we cannot simply exchange the expectation of the stochastic return $1 + r$ with the expectation of the stochastic discount $(1 + r)^{-1}$. This problem arises as soon as we work with (non-trivial) random variables. In the next chapter we define a consistent model for the stochastic discounting (deflating) of cash flows.

1.5 Recent Financial Failures and Difficulties

We close this chapter with some recent examples of failures in the industry. This list is far from being complete. For instance, it does not contain companies which were taken over by other companies just before they would have collapsed.

- 1988–1991: Lloyd's London loss of more than USD 3 billion due to asbestos and other health IBNR claims.
- 1991: Executive Life Insurance Company due to junk bonds.
- 1993: Confederation Life Insurance, Canada, loss of USD 1.3 billion due to fatal errors in asset investments.
- 1997: Nissan Mutual Life, Japan, too high guarantees on rates cost 300 billion Yen.
- 2000: Dai-ichi Mutual Fire and Marine Insurance Company, Japan, is liquidated, strategic mismanagement of their insurance merchandise.
- 2001: HIH Insurance Australia is liquidated due to a loss of USD 4 billion.
- 2001: Independent Insurance UK is liquidated due to rapid growth, insufficient reserves and inadequate premiums.
- 2001: Taisei Fire and Marine, Japan, loss of 100 billion Yen due to large reinsurance claims, e.g. World Trade Center on September 11, 2001.
- 2002: Gerling Global Re, Germany, seemed to be under-capitalized and under-reserved for many years. Further problems arose by the acquisition of Constitution Re.
- 2003: Equitable Life Assurance Society UK is liquidated due to concentration and interest rate risks.
- 2003: KBV Krankenkasse, Switzerland, is liquidated due to financial losses caused by fraud.

This was the list presented in the first edition of these lecture notes in 2007. Meanwhile we have many additional bad examples that have failed in the financial crisis 2007–2010. For example, we may mention:

- Yamato Life Insurance, loss of 22 billion Yen.
- AIG, loss of USD 145 billion by August 2009, see Donnelly–Embrechts [DE10].
- In 2008: 25 US banks failed, for example, Washington Mutual with a loss of USD 307 billion.
- In 2009: 140 US banks failed.

- Investment banks have disappeared: Merrill Lynch, Lehman Brothers, Bear Stearns, etc.
- Fraud cases like the Madoff Investment Securities LLC.

A detailed anatomy of the credit crisis 2007–2010 can be found in the Geneva Reports [GR10].

This shows that the topic of the current lecture notes has not lost any of its topicality and there is a strong need for sound quantitative methods, and an understanding of their limitations.

Chapter 2
Stochastic Discounting

In this chapter we define a mathematically consistent model for calculating (time) values of cash flows. The key objects are so-called deflators which play the role of stochastic discount factors. Our definition (via deflators) leads to values which are consistent with the classical financial theory that involves risk neutral valuation. Typically, in financial mathematics the pricing formulas are based on equivalent martingale measures (see, for example, Föllmer–Schied [FS11]), economists use the notion of state price densities (see Malamud et al. [MTW08]) and actuaries use the terminology of deflators under the real world probability measure (see Duffie [Du96] and Bühlmann et al. [BDES98]). In this chapter we introduce and describe these terminologies.

Moreover, we emphasize that in financial mathematics one usually works under risk neutral measures (equivalent martingale measures) for pricing financial assets. In actuarial mathematics, however, one should also analyze the processes under the real world probability measure (physical measure). This makes it necessary for us to understand the connection between these measures as well as the measure transformation techniques.

2.1 The Basic Discrete and Finite Time Model

In this chapter we develop the theoretical foundations of market-consistent valuation. We work in a discrete and finite time setting which has the advantage that the mathematical machinery becomes simpler; for continuous and infinite time horizon models we refer to the standard literature in financial mathematics, see for example Jeanblanc et al. [JYC09] or Elliott–Kopp [EK99].

Fix $n \in \mathbb{N}$; this n denotes the final time horizon. Then, w.l.o.g., we measure time in years and we consider cash flows on the yearly time grid $t = 0, 1, \ldots, n$.

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1_2

We choose a sufficiently rich probability space $(\Omega, \mathcal{F}, P)$ and an increasing sequence of σ-fields $\mathbb{F} = (\mathcal{F}_t)_{t=0,\ldots,n}$ on $(\Omega, \mathcal{F}, P)$ with

$$\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subseteq \mathcal{F}, \tag{2.1}$$

and for simplicity we assume $\mathcal{F}_n = \mathcal{F}$. We call $(\Omega, \mathcal{F}, P, \mathbb{F})$ a filtered probability space with filtration $\mathbb{F}$. The σ-field $\mathcal{F}_t$ plays the role of the information available/known at time t. This may include demographic information, insurance technical information on insurance contracts, financial and economic information and any other information (weather conditions, legal changes, politics, etc.) that is available at time t.

Then, we consider $\mathbb{F}$-adapted random vectors

$$\mathbf{X} = (X_0, X_1, \ldots, X_n) \tag{2.2}$$

on the filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$, implying that every component X_k of $\mathbf{X}$ is $\mathcal{F}_t$-measurable for $k \leq t$.

Interpretation and Aim.
$\mathbf{X}$ models (random) cash flows with single payments at times t described by X_t. If we have information $\mathcal{F}_t$, then X_k is known for all $k \leq t$, and otherwise it reflects a random payment in the future. On the one hand, we aim at predicting future payments $X_k, k > t$, based on the information $\mathcal{F}_t$ available at time t. On the other hand, our goal is to determine the (time) value of such cash flows $\mathbf{X}$ at any time point $t = 0, \ldots, n$, see also Fig. 2.1.

We make the following technical assumption.

Assumption 2.1 Assume that every component of the $\mathbb{F}$-adapted cash flow $\mathbf{X}$ on $(\Omega, \mathcal{F}, P, \mathbb{F})$ is square integrable.

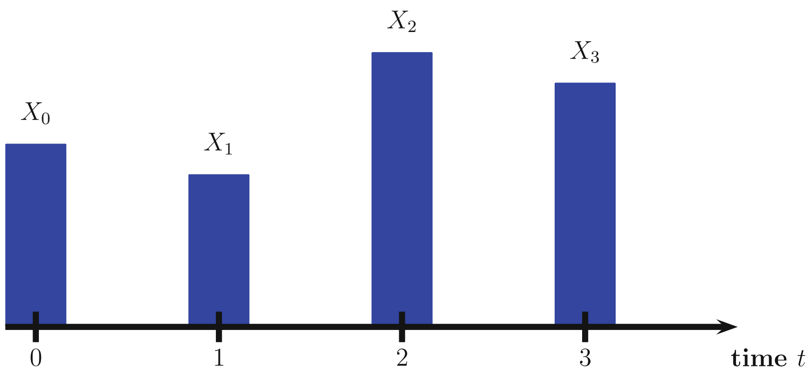


Fig. 2.1 Cash flow $\mathbf{X} = (X_0, X_1, \ldots, X_n)$ for $n = 3$

For the space of all cash flows $\mathbf{X}$ satisfying Assumption 2.1 we write

$$\mathbf{X} = (X_0, X_1, \ldots, X_n) \in L^2_{n+1}(P, \mathbb{F}). \tag{2.3}$$

We remark that $L^2_{n+1}(P, \mathbb{F})$ is a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$:

$$E\left[\sum_{t=0}^{n} X_t^2\right] < \infty \qquad \text{for all } \mathbf{X} \in L^2_{n+1}(P, \mathbb{F}), \tag{2.4}$$

$$\langle \mathbf{X}, \mathbf{Y} \rangle = E\left[\sum_{t=0}^{n} X_t Y_t\right] < \infty \qquad \text{for all } \mathbf{X}, \mathbf{Y} \in L^2_{n+1}(P, \mathbb{F}), \tag{2.5}$$

$$\|\mathbf{X}\| = \langle \mathbf{X}, \mathbf{X} \rangle^{1/2} < \infty \qquad \text{for all } \mathbf{X} \in L^2_{n+1}(P, \mathbb{F}). \tag{2.6}$$

Technical Remark. The equality $\|\mathbf{X} - \mathbf{Y}\| = 0$ implies that $\mathbf{X} = \mathbf{Y}$, P-a.s. As usually done, we identify random variables which are equal, P-a.s.

Example 2.1 (*Life insurance*) We consider a general life insurance policy financed by a regular premium income stream $(\Pi_0, \ldots, \Pi_n)$, where Π_t denotes the premium payment made at time t. Furthermore, cash outflows comprise the expenses and the benefit payments occurring in the time interval $(t-1, t]$. If we map all cash flows occurring in the time interval $(t-1, t]$ to its right end point t, we obtain a discrete time cash flow for $t \in \{0, \ldots, n\}$ given by

$$X_t = -\Pi_t + \text{benefits and expenses paid within } (t-1, t]. \tag{2.7}$$

Henceforth, $\mathbf{X} = (X_0, \ldots, X_n)$ denotes the cash flow generated by this insurance policy. The aim is to find an appropriate stochastic model that is able to describe the key features of $\mathbf{X}$. □

Example 2.2 (*Non-life insurance*) In non-life insurance the insurance company usually receives a (risk) premium at the beginning of a well-defined insurance period (upfront premium). Within this insurance period well-defined potential (random) financial losses are covered. We denote the upfront premium payment by $\Pi = -X_0$. The occurrence of an insured event (covered claim) during the insurance period typically generates a sequence of future cash outflows, the so-called claims payments, until the claim is finally settled. That is, usually the insurance company cannot immediately settle a claim, but it takes time until the ultimate claim amount is known. The delay in the settlement is due to the fact that, for example, it takes time until the total medical expenses are known, until the claim is settled at court, until the damaged building is fixed, until the recovery process is understood, etc. (see also Wüthrich–Merz [WM08, WM15] and Wüthrich [Wü13]).

Since one does not wait with the payments until the ultimate claim amount is known (e.g. medical expenses and salaries are paid when they occur) a claim consists of several single payments X_t which reflect the ongoing recovery process. Hence, the total or ultimate nominal claim amount is given by

$$C_n = \sum_{t=1}^{n} X_t, \tag{2.8}$$

where X_t, $t = 1, \ldots, n-1$, denote the single claims payments in the corresponding accounting years and X_n denotes the final payment when the claim is finally settled. Typically, at time t we have information $\mathcal{F}_t$ and the payments X_k, $k \leq t$, are already made, whereas the future payments X_k, $k > t$, need to be predicted based on the information $\mathcal{F}_t$ available at time t. These predictions of future payments determine the claims reserves in the balance sheet of the insurance company.

The underwriting loss (nominal loss) can then be written as

$$UL = \sum_{t=0}^{n} X_t = -\Pi + C_n. \tag{2.9}$$

Remark. UL does not necessarily need to be negative to run this non-life insurance business successfully. The nominal underwriting loss UL does not consider the financial income during the settlement of the claim. That is, the delay in the claims payments allows us to discount these payments, which in the profit and loss statement is considered similarly to investment income on financial assets at the insurance company (see the next sections). $\square$

2.2 Market-Consistent Valuation in the Basic Discrete Time Model

We now value the stochastic cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. We proceed as in Bühlmann [Bü92, Bü95] using a linear, positive and continuous (valuation) functional on the space of cash flows $L^2_{n+1}(P, \mathbb{F})$. Since we are dealing with random vectors $\mathbf{X}$ the definition of positivity needs some care. This we are going to introduce first.

Definition 2.2 (*Positivity*) Choose $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. We define the following notions of positivity for $\mathbf{X}$.

- $\mathbf{X} \geq 0 \iff X_t \geq 0$, P-a.s., for all $t = 0, \ldots, n$.
- $\mathbf{X} > 0 \iff \mathbf{X} \geq 0$ and there exists a $k \in \{0, \ldots, n\}$ such that $X_k > 0$ with positive probability.
- $\mathbf{X} \gg 0 \iff X_t > 0$, P-a.s., for all $t = 0, \ldots, n$.

Assumption 2.3 (*Valuation functional*) Assume that $Q : L^2_{n+1}(P, \mathbb{F}) \to \mathbb{R}$ is a (1) linear, (2) positive, (3) continuous, and (4) normalized functional on $L^2_{n+1}(P, \mathbb{F})$.

This means that the functional Q satisfies the following four properties:

(1) Linearity: For all $\mathbf{X}, \mathbf{Y} \in L^2_{n+1}(P, \mathbb{F})$ and $a, b \in \mathbb{R}$ we have

$$Q\left[a\mathbf{X} + b\mathbf{Y}\right] = aQ\left[\mathbf{X}\right] + bQ\left[\mathbf{Y}\right]. \tag{2.10}$$

(2) Positivity: For any $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ with $\mathbf{X} > 0$ we have $Q[\mathbf{X}] > 0$.
(3) Continuity: For any sequence $(\mathbf{X}^{(k)})_k \subset L^2_{n+1}(P, \mathbb{F})$ with $\mathbf{X}^{(k)} \to \mathbf{X}$ in $L^2_{n+1}(P, \mathbb{F})$ as $k \to \infty$, we have $Q\left[\mathbf{X}^{(k)}\right] \to Q\left[\mathbf{X}\right]$ in $\mathbb{R}$ as $k \to \infty$.
(4) Normalization: For all $\mathbf{X}_0 = (x_0, 0, \ldots, 0) \in L^2_{n+1}(P, \mathbb{F})$ we have $Q\left[\mathbf{X}_0\right] = x_0$.

Interpretation.
The mapping $\mathbf{X} \mapsto Q[\mathbf{X}]$ assigns a monetary value $Q[\mathbf{X}] \in \mathbb{R}$ to every cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$, which can be seen as the price of $\mathbf{X}$ at time 0. As we will see below, this valuation can be extended in a consistent way to future time points which leads to a risk neutral valuation scheme on $L^2_{n+1}(P, \mathbb{F})$. We will call Q satisfying Assumption 2.3 a *valuation functional*.

Remark. Assumptions (1) and (2) ensure that one can develop an arbitrage-free pricing system (see Lemma 2.9 and Remark 2.15, below).

Lemma 2.4 *Assumptions (1) and (2) imply (3) in Assumption 2.3.*

Proof Choose a sequence $(\mathbf{X}^{(k)})_k \subset L^2_{n+1}(P, \mathbb{F})$ with $\mathbf{X}^{(k)} \to \mathbf{X}$ in $L^2_{n+1}(P, \mathbb{F})$ as $k \to \infty$. Define $\mathbf{Y}^{(k)} = \mathbf{X}^{(k)} - \mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. Due to the linearity of Q it suffices to prove that $\mathbf{Y}^{(k)} \to 0$ in $L^2_{n+1}(P, \mathbb{F})$ implies that $Q\left[\mathbf{Y}^{(k)}\right] \to 0$ in $\mathbb{R}$.

In the first step we assume that $\mathbf{Y}^{(k)} \geq 0$ for all k. Then we claim

$$\mathbf{Y}^{(k)} \to 0 \text{ in } L^2_{n+1}(P, \mathbb{F}) \text{ implies } Q\left[\mathbf{Y}^{(k)}\right] \to 0 \text{ as } k \to \infty. \tag{2.11}$$

Assume (2.11) does not hold true, hence (using the positivity of the linear functional Q) there exists $\varepsilon > 0$ and an infinite subsequence k' of k such that for all k'

$$Q\left[\mathbf{Y}^{(k')}\right] \geq \varepsilon. \tag{2.12}$$

Choose an infinite subsequence k'' of k' with

$$\sum_{k''} \left\| \mathbf{Y}^{(k'')} \right\| < \infty. \tag{2.13}$$

We define

$$\mathbf{Y} = \sum_{k''} \mathbf{Y}^{(k'')}. \tag{2.14}$$

Due to the completeness of $L^2_{n+1}(P, \mathbb{F})$ we know that $\mathbf{Y} \in L^2_{n+1}(P, \mathbb{F})$. But linearity and positivity implies

$$Q\left[\mathbf{Y}\right] \geq Q\left[\sum_{k''=1}^{K} \mathbf{Y}^{(k'')}\right] \geq K\,\varepsilon \qquad \text{for every } K. \tag{2.15}$$

This implies that $Q\left[\mathbf{Y}\right] = \infty$ is not finite, which is a contradiction.

Second step: Decompose $\mathbf{Y}^{(k)} = \mathbf{Y}_+^{(k)} - \mathbf{Y}_-^{(k)}$ into a positive and a so-called negative part. Since $\|\mathbf{Y}_+^{(k)}\| \le \|\mathbf{Y}^{(k)}\| \to 0$ and $\|\mathbf{Y}_-^{(k)}\| \le \|\mathbf{Y}^{(k)}\| \to 0$ we see that both $\mathbf{Y}_+^{(k)}$ and $\mathbf{Y}_-^{(k)}$ tend to 0. Because $\mathbf{Y}_+^{(k)} \ge 0$ and $\mathbf{Y}_-^{(k)} \ge 0$ we have – as proved in the first step –

$$Q\left[\mathbf{Y}_+^{(k)}\right] \to 0 \quad \text{and} \quad Q\left[\mathbf{Y}_-^{(k)}\right] \to 0. \tag{2.16}$$

Using once more the linearity of Q completes the proof. □

Theorem 2.5 (Riesz' representation) *Assume that the functional* $Q : L^2_{n+1}(P, \mathbb{F}) \to \mathbb{R}$ *fulfills Assumption 2.3. There exists a* $\varphi \in L^2_{n+1}(P, \mathbb{F})$ *such that for all* $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ *we have*

$$Q[\mathbf{X}] \;=\; \langle \varphi, \mathbf{X} \rangle \;=\; E\left[\sum_{t=0}^{n} \varphi_t X_t\right]. \tag{2.17}$$

This $\varphi \in L^2_{n+1}(P, \mathbb{F})$ *has the following properties:*

- φ *is* $\mathbb{F}$*-adapted;*
- φ *has square integrable components* φ_t *for* $t = 0, \ldots, n$*;*
- φ *is unique;*
- $\varphi \gg 0$*; and*
- $\varphi_0 = 1$.

Proof of Theorem 2.5. The classical Riesz' representation theorem provides for every linear and continuous functional Q the existence of a $\varphi \in L^2_{n+1}(P, \mathbb{F})$ such that (2.17) holds for every $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. Thus, in view of Assumption 2.3 and Lemma 2.4 we may relax the assumptions to apply the classical Riesz' representation theorem.

There remain the proofs of the properties of φ: $\mathbb{F}$-adaptedness and square integrability are immediately clear since $\varphi \in L^2_{n+1}(P, \mathbb{F})$. Next we prove uniqueness. Assume that there are two random vectors φ and φ^* in $L^2_{n+1}(P, \mathbb{F})$ satisfying for all $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$

$$Q[\mathbf{X}] = \langle \varphi, \mathbf{X} \rangle = \langle \varphi^*, \mathbf{X} \rangle. \tag{2.18}$$

But then we may choose $\mathbf{X} = \varphi - \varphi^* \in L^2_{n+1}(P, \mathbb{F})$. This and (2.18) imply

$$0 = \langle \varphi - \varphi^*, \mathbf{X} \rangle = \|\varphi - \varphi^*\|^2, \tag{2.19}$$

which immediately gives $\varphi = \varphi^*$, P-a.s. This proves uniqueness.

Next we prove $\varphi \gg 0$. Assume that the latter does not hold true. Then there exists a $k \in \{0, \ldots, n\}$ such that $P[\varphi_k \le 0] > 0$. We define the cash flow

$$\mathbf{X} = (0, \ldots, 0, 1_{\{\varphi_k \le 0\}}, 0, \ldots, 0) \in L^2_{n+1}(P, \mathbb{F}), \tag{2.20}$$

with the non-zero entry being in the $(k+1)$-st component of $\mathbf{X}$. Note that this cash flow satisfies $\mathbf{X} > 0$. Positivity of Q then implies

$$0 < Q[\mathbf{X}] = \langle \varphi, \mathbf{X} \rangle = E\left[\varphi_k 1_{\{\varphi_k \le 0\}}\right] \le 0, \tag{2.21}$$

which is the desired contradiction. Finally, normalization follows from the assumption $\mathcal{F}_0 = \{\emptyset, \Omega\}$ which implies for every non-zero cash flow $\mathbf{X}_0 = (x_0, 0, \ldots, 0) \in L^2_{n+1}(P, \mathbb{F})$

$$x_0 = Q[\mathbf{X}_0] = \langle \varphi, \mathbf{X}_0 \rangle = E[\varphi_0 x_0] = \varphi_0 x_0. \tag{2.22}$$

This proves the claim. □

Definition 2.6 The vector φ (and its single components φ_t) satisfying the properties in Theorem 2.5 is called a (state price) deflator.

The terminology (state price) deflator was introduced by Duffie [Du96] and Bühlmann et al. [BDES98]. In economic theory deflators are called "state price densities" and in financial mathematics "financial pricing kernels", "stochastic interest rates" or "stochastic discount factors".

Remarks.

- We have assumed that $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ in order to find the state price deflator φ. This can be generalized to cash flows $\mathbf{X} \in L^p_{n+1}(P, \mathbb{F})$, $1 \le p \le \infty$, and then the deflator φ would be in $L^q_{n+1}(P, \mathbb{F})$ with $1/p + 1/q = 1$. Or, even more generally, we can take a fixed deflator $\varphi \in L^1_{n+1}(P, \mathbb{F})$ with $\varphi \gg 0$ and then define the set of cash flows that can be priced by

$$\mathcal{L}_\varphi = \left\{ \mathbf{X} \in L^1_{n+1}(P, \mathbb{F}) : \; E\left[\sum_{t=0}^{n} \varphi_t |X_t| \right] < \infty \right\}. \tag{2.23}$$

  For these cash flows we then define the valuation functional Q on $\mathcal{L}_\varphi$ by $Q[\mathbf{X}] = \langle \varphi, \mathbf{X} \rangle$. For more details we refer to Wüthrich–Merz [WM13].
- There is a one-to-one correspondence between the valuation functional Q of Assumption 2.3 and the state price deflator φ according to Definition 2.6. Theorem 2.5 proves one direction and the other direction is immediate.

2.2.1 *The Task of Modelling*

Find the appropriate valuation functional Q or equivalently find the appropriate $\mathbb{F}$-adapted state price deflator φ!

In the more general setup, one would define/choose $\varphi \in L^1_{n+1}(P, \mathbb{F})$ and then value the cash flows $\mathbf{X} \in \mathcal{L}_\varphi$, see (2.23). The choice of φ will include market risk aversion as well as individual risk aversion, this will be described in the following chapters, and we will also describe the connection between the state price deflators and the risk neutral martingale measures.

The $\mathbb{F}$-adaptedness will be crucial in the sequel. It essentially means that the deflator φ_t (stochastic discount factor) is known at time t and, hence, allows us to make a direct connection between the $\mathcal{F}_t$-measurable cash flow X_t and the behaviour of the financial market at time t described by φ_t. In particular, this implies that φ_t will allow us to model embedded options and guarantees in X_t that depend on economic and financial scenarios.

Examples of state price deflators can be found in Bühlmann [Bü95], for example the Ehrenfest Urn with the limit Ornstein–Uhlenbeck model, in Filipović–Zabczyk [FZ02] or one can easily discretize, for example, the Vasiček model, see Brigo–Mercurio [BM06] and Exercise 2.3. For more discrete time examples we refer to Wüthrich–Merz [WM13].

Exercise 2.3 (*Discrete time Vasiček* [Va77] *model*) Choose a filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ and assume that $(\varepsilon_t)_{t=0,\ldots,n}$ is $\mathbb{F}$-adapted, that ε_t is independent of $\mathcal{F}_{t-1}$ for all $t = 1, \ldots, n$ and standard Gaussian distributed w.r.t. P. Then, we define the stochastic process $(r_t)_{t=0,\ldots,n}$ by $r_0 > 0$ (fixed) and for $t \geq 1$

$$r_t = b + \beta r_{t-1} + \rho \varepsilon_t, \tag{2.24}$$

for given parameters b, β, $\rho > 0$. This $(r_t)_{t=0,\ldots,n}$ describes the spot rate dynamics of the Vasiček model under the (real world) probability measure P, see Wüthrich–Merz [WM13], Sect. 3.3.

Next, we choose $\lambda \in \mathbb{R}$ (market price of risk) and define the deflator in the Vasiček model by

$$\varphi_t = \exp\left\{ -\sum_{k=1}^{t} \left[r_{k-1} + \frac{\lambda^2}{2} r_{k-1}^2 \right] - \sum_{k=1}^{t} \lambda\, r_{k-1}\, \varepsilon_k \right\}, \tag{2.25}$$

for $t = 1, \ldots, n$ and $\varphi_0 = 1$. Prove that $\boldsymbol{\varphi} = (\varphi_0, \ldots, \varphi_n) \in L^1_{n+1}(P, \mathbb{F})$ is a deflator. Moreover, prove that the cash flow $\mathbf{X} = (0, \ldots, 0, 1, 0, \ldots, 0)$ is in $\mathcal{L}_{\varphi}$, see (2.23). □

2.2.2 Understanding Deflators and Zero Coupon Bonds

A *deflator* φ_t transports the cash amount at time t to its value at time 0, see Fig. 2.2. This transportation is a stochastic transportation (stochastic discounting). A cash flow $\mathbf{X}_t = (0, \ldots, 0, X_t, 0, \ldots, 0)$ does not necessarily need to be independent of (or uncorrelated with) φ_t, which, in general, gives

$$Q\left[\mathbf{X}_t\right] = E\left[X_t\, \varphi_t\right] \neq E\left[X_t\right] E\left[\varphi_t\right]. \tag{2.26}$$

$Q\left[\mathbf{X}_t\right]$ describes the value/price of $\mathbf{X}_t$ at time 0, where X_t is stochastically discounted with the $\mathcal{F}_t$-measurable deflator φ_t.

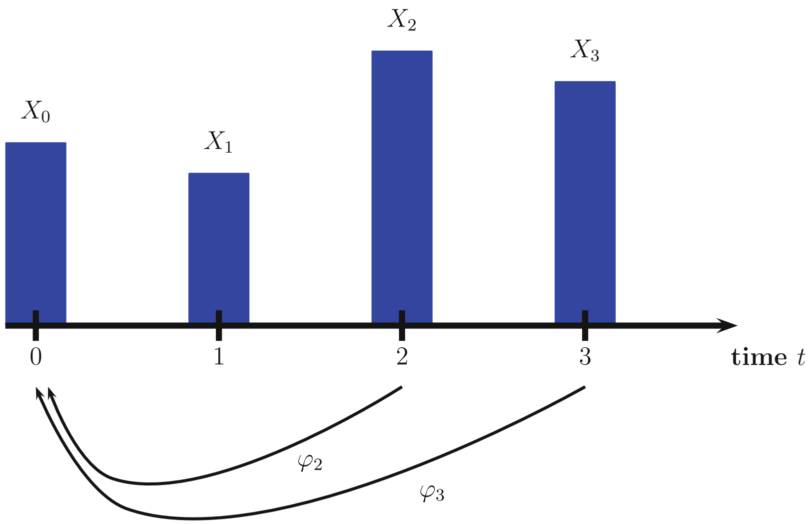


Fig. 2.2 Deflator φ and cash flow $\mathbf{X}$

We decompose the deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$ into its *span-deflators*. Since $\varphi \gg 0$ we can build the following ratios for all $t > 0$, P-a.s.:

$$Y_t = \frac{\varphi_t}{\varphi_{t-1}}. \tag{2.27}$$

Moreover, we define $Y_0 = 1$. Thus, $\mathbf{Y} = (Y_t)_{t=0,\ldots,n}$ is $\mathbb{F}$-adapted and satisfies

$$\varphi_t = Y_0 \, Y_1 \, \cdots \, Y_t = \prod_{k=0}^{t} Y_k. \tag{2.28}$$

$\mathbf{Y} = (Y_t)_{t=0,\ldots,n}$ is called a span-deflator. Span-deflators Y_t, $t \geq 1$, transport the cash amount at time t to its value at time $t-1$ (one-year deflating), see Fig. 2.3. For more information we refer to p. 28.

Question. How is the deflator φ related to zero coupon bonds and classical financial discounting?

Definition 2.7 A (default-free) zero coupon bond is a financial instrument that pays one unit of currency at a fixed maturity date $t \in \{0, \ldots, n\}$. Its cash flow is denoted by $\mathbf{Z}^{(t)} = (0, \ldots, 0, 1, 0, \ldots, 0) \in L^2_{n+1}(P, \mathbb{F})$.

For a given state price deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$, the value at time 0 of this zero coupon bond is given by

$$d_{0,t} = Q\left[\mathbf{Z}^{(t)}\right] = E\left[\varphi_t\right]. \tag{2.29}$$

In the financial literature $d_{0,t}$ is often denoted by $P(0,t)$. We come back to this in (2.50).

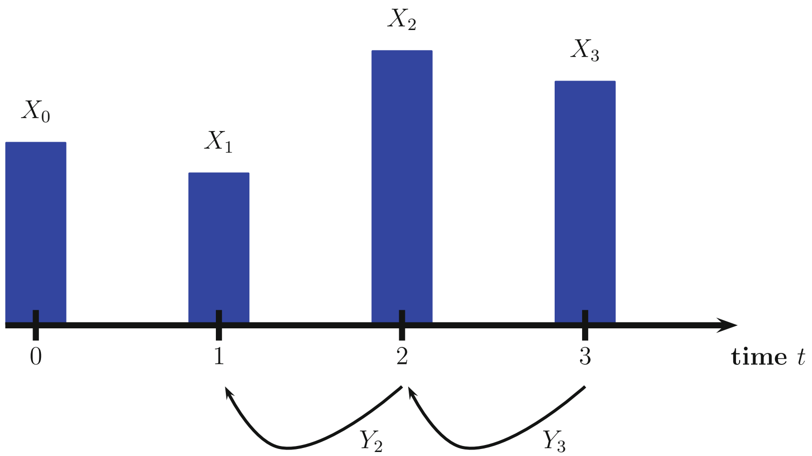


Fig. 2.3 Span-deflators Y_t and cash flow $\mathbf{X}$

Henceforth, $d_{0,t}$ transports the cash amounts at time t to their value at 0. But $d_{0,t}$ is $\mathcal{F}_0$-measurable, whereas φ_t is a $\mathcal{F}_t$-measurable random variable. This means that the deterministic discount factor $d_{0,t}$ is known at the beginning of the time period $(0, t]$, whereas φ_t is only known at the end of the time period $(0, t]$. As long as we are dealing with deterministic cash flows $\mathbf{X}$, we can either work with zero coupon bond prices $d_{0,t}$ or with deflators φ_t to determine the value of $\mathbf{X}$ at time 0. But as soon as the cash flows $\mathbf{X}$ are stochastic we need to work with deflators (see (2.26)) since X_t and φ_t may be influenced by the same risk factors (are dependent). An easy example of this dependence is constructed by choosing X_t as an option that depends on the actual realization of φ_t. Various life insurance companies hold such embedded options and financial guarantees in their insurance portfolios, and henceforth need to use the deflator framework for valuation.

Classical actuarial discounting is taking a constant interest rate $i > 0$. That is, in classical actuarial models φ_t has the following form

$$\varphi_t = (1 + i)^{-t}. \tag{2.30}$$

This deflator gives a consistent theory but its behaviour is far from the economic observations found in practice. This indicates that we have to be very careful with this deterministic modelling choice in a total balance sheet approach, since it implies that we will obtain values which are far away from financial market observations.

Exercise 2.4 (*Zero coupon bond price in the Vasiček model*) We revisit the discrete time Vasiček model presented in Exercise 2.3. Calculate for this model the zero coupon bond prices $d_{0,t}$. We claim that these prices are given by

$$d_{0,t} = \exp\{A(0,t) - r_0\, B(0,t)\}, \tag{2.31}$$

for appropriate functions $A(0,\cdot)$ and $B(0,\cdot)$.

Hint: the claim is proved by induction using properties of log-normal distributions.

Give an interpretation of r_0 in terms of $d_{0,1}$. For more background information we refer to Sect. 3.3 of Wüthrich–Merz [WM13]. □

Before we define prices at arbitrary time points $t = 0, \ldots, n$ we give a finite probability space example which shows the relation between deflators and replication. This is done in the next subsection.

2.2.3 A Toy Example for Deflators

In this subsection we give a toy example of a deflator construction that is based on a finite probability space. In a first step we introduce a statistical model that is calibrated. In a second step we construct the deflator. Our example is borrowed from Jarvis et al. [JSV01].

We consider a one-period model and we assume that there are two possible states at time 1 called ω_1 and ω_2. This setup can be described by the measurable space $(\Omega, \mathcal{F})$ with $\Omega = \{\omega_1, \omega_2\}$ and $\mathcal{F} = 2^{\Omega}$ being the corresponding power set. Finding deflators on finite probability spaces is essentially an exercise in linear algebra. We should also mention that models on finite probability spaces often have the advantage that the crucial mathematical and economic structures are easier to detect (see Malamud et al. [MTW08]).

Step 1. In a first step we construct the *state space securities* SS_1 and SS_2 for the two states ω_1 and ω_2, respectively. A state space security SS_i for state ω_i pays one unit of currency if state ω_i occurs at time 1. These state space securities are used to construct a consistent pricing model. That is, we aim at calibrating the following table.

	SS_1	SS_2
Market price $Q[\cdot]$ at time 0	?	?
Payout if state ω_1 occurs at time 1	1	0
Payout if state ω_2 occurs at time 1	0	1

Since we have two possible states ω_1 and ω_2 we need two linearly independent assets A and B to calibrate the model. Assume that assets A and B have the following prices and payout structures:

	Asset A	Asset B
Market price $Q[\cdot]$ at time 0	1.65	1
Payout if state ω_1 occurs at time 1	3	2
Payout if state ω_2 occurs at time 1	1	0.5

With this information we can construct the two state space securities SS_1 and SS_2 and calculate their (consistent) prices at time 0. For this purpose, say for SS_1, we construct a portfolio that consists of x_1 units of asset A and y_1 units of asset B. The goal is to determine x_1 and y_1 such that the resulting portfolio pays 1 if state ω_1 occurs at time 1 and 0 otherwise. That is, this portfolio exactly replicates the state space security SS_1. Mathematically speaking we need to solve the equation $SS_1 = x_1 A + y_1 B$ on Ω for SS_1, and a similar equation for SS_2. The solution to these two equations provides the following table:

	Units x_i of A	Units y_i of B	Market price Q
State space security SS_1	-1	2	0.35
State space security SS_2	4	-6	0.60

The last column of this table is obtained by requiring that asset prices are linear, in the sense that

$$\text{price of } SS_i = x_i \cdot \text{price of asset } A + y_i \cdot \text{price of asset } B. \tag{2.32}$$

Note that this is similar to the derivation of the Arbitrage Pricing Theory framework (see Ingersoll [Ing87], Chap. 7). Basically, we need that asset A and asset B are linearly independent and that the valuation functional Q is linear. Based on these assumptions the consistent market prices of all other assets and cash flows are calculated using replication arguments. Thus, if we have another asset $\mathbf{X}$ which pays 2 in state ω_1 and 1 in state ω_2, its consistent price at time 0 is given by

$$Q[\mathbf{X}] = 2 \cdot 0.35 + 1 \cdot 0.60 = 1.3. \tag{2.33}$$

Consider the zero coupon bond $\mathbf{Z}^{(1)}$ that pays in both states ω_1 and ω_2 the amount 1:

$$d_{0,1} = Q\left[\mathbf{Z}^{(1)}\right] = 1 \cdot 0.35 + 1 \cdot 0.60 = 0.95, \tag{2.34}$$

which leads to a risk-free return of $(0.95)^{-1} - 1 = 5.26\,\%$. The pricing model is now calibrated so that we have consistent market prices.

Step 2. Next we construct the deflators. By a slight abuse of notation, we denote by $Q(\omega_i)$ the market price of the state space security SS_i at time 0, i.e. $Q(\omega_1) = 0.35$ and $Q(\omega_2) = 0.60$. Moreover, let $X_1(\omega_i)$ denote the payout of the asset $\mathbf{X} = (0, X_1)$ if state ω_i occurs at time 1. The consistent market price of $\mathbf{X}$ at time 0 is given by, see (2.33),

$$Q[\mathbf{X}] = \sum_{i=1}^{2} X_1(\omega_i)\; Q(\omega_i). \tag{2.35}$$

Note: So far we have not used any probabilities! All arguments are based on linearity and replication only!

Now we complete $(\Omega, \mathcal{F})$ to a probability space by assuming that state ω_1 occurs with probability $p(\omega_1) \in (0, 1)$ and state ω_2 with probability $p(\omega_2) = 1 - p(\omega_1) \in (0, 1)$. Identity (2.35) can be rewritten as follows

$$Q[\mathbf{X}] = \sum_{i=1}^{2} X_1(\omega_i)\ Q(\omega_i) = \sum_{i=1}^{2} p(\omega_i)\ \frac{Q(\omega_i)}{p(\omega_i)}\ X_1(\omega_i) = E\left[\frac{Q}{p}\ X_1\right], \quad (2.36)$$

where E denotes the expected value induced by the probabilities $p(\cdot)$. Henceforth, we may define the random variable

$$\varphi_1 = \frac{Q}{p}, \quad (2.37)$$

which immediately implies the pricing formula

$$Q[\mathbf{X}] = E[\varphi_1 X_1]. \quad (2.38)$$

For an explicit choice of probabilities $p(\omega_i)$, the deflator φ_1 takes the following values:

	Value of deflator φ_1	Probability $p(\omega_i)$
If state ω_1 Occurs at time 1	0.7	0.5
If state ω_2 Occurs at time 1	1.2	0.5

Alternatively to (2.34) this provides the value of the zero coupon bond

$$Q\left[\mathbf{Z}^{(1)}\right] = E[\varphi_1] = \sum_{i=1}^{2} \varphi_1(\omega_i)\ p(\omega_i) = 0.7 \cdot 0.5 + 1.2 \cdot 0.5 = 0.95. \quad (2.39)$$

Note that in our example the deflator φ_1 is not necessarily smaller than 1. With probability 1/2 we will observe that the deflator has a value of 1.2. This may be counter-intuitive from an economic point of view but it makes perfect sense in our model world. Henceforth, the model and parameters need to be specified carefully in order to get economically meaningful models. Moreover, for strict positivity of the deflator, we need that the state space securities have strictly positive prices (i.e. that Q is positive) which is a natural property in consistent pricing systems.

2.3 Valuation at Time $t > 0$

Postulate: Correct prices should eliminate the possibility of "playing games" with cash flows. This should be interpreted as the non-existence of strategies that provide expected gains without downside risks (see also Remark 2.15).

Assume a fixed deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$ is given. We define the price process $(Q_t[\mathbf{X}])_{t=0,\dots,n}$ of the random vector $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ w.r.t. φ as follows: for $t = 0, \dots, n$ we set

$$Q_t[\mathbf{X}] = \frac{1}{\varphi_t} E\left[\sum_{k=0}^{n} \varphi_k X_k \middle| \mathcal{F}_t \right]. \tag{2.40}$$

Observe that $\varphi \gg 0$ and $\mathbf{X}, \varphi \in L^2_{n+1}(P, \mathbb{F})$ imply that $Q_t[\mathbf{X}]$ is well-defined. From definition (2.40) it is also obvious that price $Q_t[\mathbf{X}]$ is $\mathcal{F}_t$-measurable for all t, and henceforth $(Q_t[\mathbf{X}])_{t=0,\dots,n}$ is $\mathbb{F}$-adapted. Note that the right-hand side of (2.40) can be decoupled because the payments X_k (and the deflators φ_k) are $\mathcal{F}_t$-measurable for $k \le t$.

Interpretation.
The mapping $\mathbf{X} \mapsto Q_t[\mathbf{X}]$ assigns a monetary value $Q_t[\mathbf{X}]$ at time t to the cash flow $\mathbf{X}$, i.e. it attaches an $\mathcal{F}_t$-measurable price to the cash flow $\mathbf{X}$. Of course, in general, this price is stochastic seen from time 0 and depends on $\mathcal{F}_t$. We will see below that these price processes $(Q_t[\mathbf{X}])_{t=0,\dots,n}$ lead to a consistent pricing system for a given state price deflator φ. Moreover, this consistent pricing system is equivalent to a risk neutral (and arbitrage-free) valuation scheme (see Lemma 2.9 and Remark 2.15).

By our assumptions we have $Q[\mathbf{X}] = Q_0[\mathbf{X}]$. This is implied by normalization $\varphi_0 = 1$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

The justification of our price process definition $(Q_t[\mathbf{X}])_{t=0,\dots,n}$ may use an equilibrium argument or alternatively a no-arbitrage argument. The latter (more technical) argument will be provided in Lemma 2.9 and Remark 2.15. At the current stage we provide the equilibrium argument. Assume we purchase cash flow $\mathbf{X}$ at time t at price $Q_t[\mathbf{X}]$. Hence, we generate the following payment cash flow by this acquisition

$$Q_t[\mathbf{X}]\ \mathbf{Z}^{(t)} = (0, \dots, 0, Q_t[\mathbf{X}], 0, \dots, 0), \tag{2.41}$$

if we pay the price for $\mathbf{X}$ at time t. From today's point of view this payment stream has value

$$Q_0\left[Q_t[\mathbf{X}]\ \mathbf{Z}^{(t)}\right], \tag{2.42}$$

since we have only information $\mathcal{F}_0$ at time 0 about the price $Q_t[\mathbf{X}]$ of $\mathbf{X}$ at time t. Equilibrium requires that

$$Q_0[\mathbf{X}] = Q_0\left[Q_t[\mathbf{X}]\ \mathbf{Z}^{(t)}\right], \tag{2.43}$$

since (based on today's information $\mathcal{F}_0$) the two payment streams should have the same value. That is, we agree *today* to either purchase and pay $\mathbf{X}$ today or to purchase and pay $\mathbf{X}$ at time t (at its current price $Q_t[\mathbf{X}]$ at that time). Since we use the same information $\mathcal{F}_0$ for these two contracts and they provide the same asset $\mathbf{X}$ the two contracts should have the same price.

Suppose now that we play the following game: We decide to purchase and pay cash flow $\mathbf{X}$ if and only if an event $F_t \in \mathcal{F}_t$ occurs. Since from today's point of view we do not know whether the event F_t will occur, we should have the following price equilibrium, see also (2.43),

$$Q_0\left[\mathbf{X}\, 1_{F_t}\right] = Q_0\left[Q_t\left[\mathbf{X}\right]\, \mathbf{Z}^{(t)}\, 1_{F_t}\right]. \tag{2.44}$$

Note that strictly speaking (2.44) is not well-defined because the cash flow $\mathbf{X}\, 1_{F_t}$ on the left-hand side is not $\mathbb{F}$-adapted. To make this argument rigorous one should restrict to the outstanding cash flow of $\mathbf{X}$ at time $t-1$. This is similar to the considerations (2.52) and (2.54). Nevertheless, we may rewrite (2.44) for the given deflator φ as follows

$$E\left[\sum_{k=0}^{n} \varphi_k\, X_k\, 1_{F_t}\right] = E\left[\varphi_t\, Q_t\left[\mathbf{X}\right]\, 1_{F_t}\right]. \tag{2.45}$$

Since $\varphi_t\, Q_t\left[\mathbf{X}\right]$ is $\mathcal{F}_t$-measurable and Eq. (2.45) should hold true for all $F_t \in \mathcal{F}_t$, this is exactly the definition of the conditional expectation given the σ-field $\mathcal{F}_t$. Henceforth, (2.45) implies (2.40), P-a.s., and justifies that (2.40) is an economically meaningful definition. A more financial mathematically based argumentation would say that deflated price processes need to be $(P, \mathbb{F})$-martingales in order to have an arbitrage-free pricing scheme, see Lemma 2.9 and Remark 2.15.

Remark. We would like to emphasize that we *first* choose the state price deflator φ and *then* calculate the price processes $(Q_t[\mathbf{X}])_{t=0,\ldots,n}$ of all cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ under the *same* state price deflator φ. This provides a consistent pricing system (that naturally depends on the choice of φ).

We close this section with remarks on discounting and forward rates. We have defined the discount factors at time 0 by

$$d_{0,m} = Q_0\left[\mathbf{Z}^{(m)}\right] = E\left[\varphi_m\right], \qquad \text{for } m = 1, \ldots, n, \tag{2.46}$$

calculated from zero coupon bond cash flows $\mathbf{Z}^{(m)}$ with maturity dates m. For $t < m$, let $d_{t,m}$ be the *forward discount factor fixed at time 0* for discounting back from time m to time t. The terminology 'forward' refers to this fixing at an earlier time point. In this deterministic ($\mathcal{F}_0$-measurable) setup consistency requires

$$d_{0,t}\, d_{t,m} = d_{0,m}. \tag{2.47}$$

The left-hand side of (2.47) is the price at time 0 for receiving amount $d_{t,m}$ at time t, and an immediate reinvestment of this amount into 1 forward discount factor contract at time t at price $d_{t,m}$ (fixed at time 0) provides 1 unit of currency at maturity date m. The right-hand side of (2.47) is the price at time 0 of a zero coupon bond contract for directly receiving 1 unit of currency at time m. Hence, both strategies provide the same cash flow and are based on the same information $\mathcal{F}_0$, therefore consistency requires identity (2.47). Thus, the (consistent) forward discount factors for $t < m$ are determined by

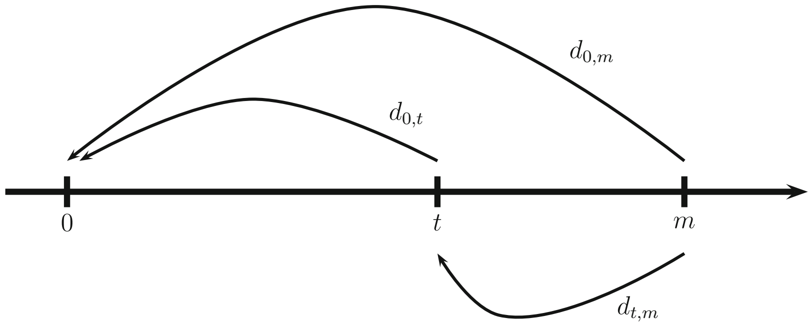


Fig. 2.4 $\mathcal{F}_0$-measurable forward discount factor $d_{t,m}$ for $t < m$

$$d_{t,m} = \frac{d_{0,m}}{d_{0,t}}. \tag{2.48}$$

This is the forward price of a zero coupon bond with maturity date m fixed at time 0, i.e. $\mathcal{F}_0$-measurable, to be paid at time t (Fig. 2.4).

On the other hand, the $\mathcal{F}_t$-measurable price at time t of a zero coupon bond with maturity date m is given by

$$Q_t\left[\mathbf{Z}^{(m)}\right] = \frac{1}{\varphi_t}\, E\left[\varphi_m \,|\, \mathcal{F}_t\right] = E\left[\left.\frac{\varphi_m}{\varphi_t}\right| \mathcal{F}_t\right]. \tag{2.49}$$

This price is in line with definition (2.40) for a single deterministic payment of size 1 at time m, see also Definition 2.7.

Notation. In financial mathematics one uses the following notation and identities for zero coupon bond prices, $t < m$,

$$P(t, m) = Q_t\left[\mathbf{Z}^{(m)}\right] = E\left[\left.\frac{\varphi_m}{\varphi_t}\right| \mathcal{F}_t\right] = E^*\left[\left.\exp\left\{-\sum_{s=t}^{m-1} r_s\right\}\right| \mathcal{F}_t\right], \tag{2.50}$$

where $(r_t)_{t=0,\ldots,n}$ is the spot rate process and E^* is the expectation under an equivalent martingale measure $P^* \sim P$. The first identity in (2.50) is a definition, the second one is obtained from (2.40) and the last one is going to be derived in Sect. 2.5. Note that for $t = 0$ we have $d_{0,m} = P(0, m) = Q_0\left[\mathbf{Z}^{(m)}\right]$. In these notes we use $d_{t,m}$ for $\mathcal{F}_0$-measurable forward discount factors and $P(t, m)$ for $\mathcal{F}_t$-measurable zero coupon bond prices.

Exercise 2.5 We revisit the discrete time Vasiček model presented in Exercise 2.3. Calculate for this model the zero coupon bond prices $P(t, m)$ at times $t < m$. We claim that these prices are given by

$$P(t,m) = Q_t\left[\mathbf{Z}^{(m)}\right] = \exp\left\{A(t,m) - r_t\; B(t,m)\right\}, \tag{2.51}$$

for appropriate functions $A(\cdot,\cdot)$ and $B(\cdot,\cdot)$ and $\mathcal{F}_t$-measurable spot rates r_t, see also (2.31).

Give an interpretation of r_t in terms of $P(t, t+1)$.

Remark. A zero coupon bond price representation of the form (2.51) is called an affine term structure, because its logarithm is an affine function of the observed spot rate r_t for all $t = 0, \ldots, m-1$, see also Filipović [Fi09] and Wüthrich–Merz [WM13]. □

2.4 The Meaning of Reserves

In the previous section we have considered the valuation of cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ at any time point $t = 0, \ldots, n$. However, in an insurance context we are mainly interested in the valuation of *future* cash flows $(0, \ldots, 0, X_{t+1}, \ldots, X_n)$ if we are currently at time t. For these future cash flows we need to build reserves on the liability side of the balance sheet, because they refer to the outstanding (loss) liabilities we still need to meet. This means that we need to predict and assign (market-)consistent values to X_k, $k > t$, based on the available information $\mathcal{F}_t$ at time t.

Note that from an economic point of view the terminology *reserves* is not completely correct (because reserves rather refers to shareholder value) and one should call the reserves instead *provisions* because they belong to the insured and the policyholder, respectively.

Postulate: Correct reserves should eliminate the possibility of “playing games” with insurance liabilities.

Throughout, we assume a fixed deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$ is given.

Assume that the insurance contract is represented by the (stochastic) cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. We define for $1 \le k \le n$ the *outstanding liabilities* of $\mathbf{X}$ at time $k-1$ by

$$\mathbf{X}_{(k)} = (0, \ldots, 0, X_k, \ldots, X_n) \in L^2_{n+1}(P, \mathbb{F}). \tag{2.52}$$

This is the remaining cash flow of $\mathbf{X}$ after time $k-1$. $\mathbf{X}_{(k)}$ represents the amounts for which we have to build reserves at time $k-1$, such that we are able to meet all future payments arising from this contract. The *reserves at time* $0 \le t \le k-1$ for the outstanding liabilities $\mathbf{X}_{(k)}$ are defined as

$$R_t^{(k)} = Q_t[\mathbf{X}_{(k)}] = \frac{1}{\varphi_t}\; E\left[\sum_{s=k}^{n} \varphi_s\; X_s \middle|\, \mathcal{F}_t\right]. \tag{2.53}$$

On the one hand, $R_t^{(k)}$ corresponds to the conditionally expected monetary value of the cash flow $\mathbf{X}_{(k)}$ viewed from time t. On the other hand, $R_t^{(k)}$ is used to predict the monetary value of the random variable $\mathbf{X}_{(k)}$. We will comment more on this prediction below.

We justify that (2.53) is a reasonable definition of the reserves. We argue for $R_t^{(k)}$ in a similar fashion as in the last section using an equilibrium argument: we want to avoid that we can play games with insurance contracts. In particular, we consider the following game: assume we have two insurance companies A and B that have exactly the same liability $\mathbf{X}$ and the following two business strategies.

- Company A keeps the contract until all payments are met.
- Company B decides (at time 0) to sell the run-off of the outstanding liabilities at time $t-1$ at price $R_{t-1}^{(t)}$ if an event $F_{t-1} \in \mathcal{F}_{t-1}$ occurs.

This implies that the two strategies generate the following cash flows:

	0	...	$t-1$	t	...	n
$\mathbf{X}^{(A)} =$	$(X_0,$	...,	$X_{t-1},$	$X_t,$	...,	$X_n)$
$\mathbf{X}^{(B)} =$	$(X_0,$	...,	$X_{t-1} + R_{t-1}^{(t)} 1_{F_{t-1}},$	$X_t 1_{F_{t-1}^c},$	...,	$X_n 1_{F_{t-1}^c})$

Hence, the price difference at time 0 of these two strategies is given by

$$Q_0\left[\mathbf{X}^{(A)} - \mathbf{X}^{(B)}\right] = E\left[-\varphi_{t-1} R_{t-1}^{(t)} 1_{F_{t-1}}\right] + E\left[\sum_{s=t}^{n} \varphi_s X_s 1_{F_{t-1}}\right]. \tag{2.54}$$

As in (2.44), we have that the two strategies based on information $\mathcal{F}_0$ should have the same initial value, because they are based on the same information and they face the same liability $\mathbf{X}$. This implies requirement $Q_0\left[\mathbf{X}^{(A)} - \mathbf{X}^{(B)}\right] = 0$, and thus we should have for any event $F_{t-1} \in \mathcal{F}_{t-1}$ the following equality

$$E\left[\varphi_{t-1} R_{t-1}^{(t)} 1_{F_{t-1}}\right] = E\left[\sum_{s=t}^{n} \varphi_s X_s 1_{F_{t-1}}\right]. \tag{2.55}$$

Using the definition of conditional expectations, this gives the following definition of the reserves:

$$R_{t-1}^{(t)} = \frac{1}{\varphi_{t-1}} E\left[\left.\sum_{s=t}^{n} \varphi_s\, X_s \right| \mathcal{F}_{t-1}\right] = Q_{t-1}\left[\mathbf{X}_{(t)}\right], \tag{2.56}$$

which justifies (2.53) for $k = t$. The case $k > t$ is then easily obtained by iteration. Observe that the argument in (2.44) was not completely correct because of measurability issues. This is now solved in (2.54) by only considering outstanding liabilities after time $t-1$. However, there is still a minor issue in (2.54) because we did not prove that $R_k^{(t)}$, $k \leq t-1$, is square integrable. As a consequence, our consistent pricing framework may allow for price processes that are not square integrable and the sale (or purchase) of such financial instruments may generate cash flows that are not square integrable. However, importantly, consistency is still preserved and the cash flow valuation framework may need to be extended to a bigger space, see also (2.23) and Lemma 2.9.

Observe that we have the following mean self-financing property:

Corollary 2.8 (Mean self-financing property) *The following recursion holds for all* $t = 1, \ldots, n-1$

$$E\left[\varphi_t \left(R_t^{(t+1)} + X_t\right)\middle|\, \mathcal{F}_{t-1}\right] = \varphi_{t-1}\, R_{t-1}^{(t)}. \tag{2.57}$$

Remark.

- The classical actuarial theory with $\varphi_t = (1+i)^{-t}$ for some positive interest rate i (see (2.30)) forms a consistent theory but the deflators are not market-consistent, because they are often far from observed economic behaviours at traded markets.
- Corollary 2.8 basically says that if we want to avoid arbitrage opportunities for reserves then we need to define them as conditional expectations of the stochastically discounted random cash flows.

Proof of Corollary 2.8. We have the following identity (using the $\mathcal{F}_t$-measurability of X_t and the tower property of conditional expectations, see Williams [Wi91], Chap. 9)

$$E\left[\varphi_t \left(R_t^{(t+1)} + X_t\right)\middle|\, \mathcal{F}_{t-1}\right] = E\left[\sum_{k=t}^{n} \varphi_k X_k \middle|\, \mathcal{F}_{t-1}\right] = \varphi_{t-1} R_{t-1}^{(t)}. \tag{2.58}$$

This completes the proof of the corollary. □

2.5 Equivalent Martingale Measures

Throughout, we assume a fixed deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$ is given.

Every deflated price process defined by (2.40) gives a $(P, \mathbb{F})$-martingale according to Lemma 2.9. As a consequence the pricing system implied by the given deflator φ satisfies the efficient market hypothesis and is consistent as described in Remark 2.15.

Lemma 2.9 *For every cash flow* $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ *the deflated price process defined by (2.40) satisfies:*

$$(\varphi_t Q_t [\mathbf{X}])_{t=0,\ldots,n} \quad \textit{is a } (P, \mathbb{F})\textit{-martingale.} \tag{2.59}$$

Proof Integrability of the components of $(\varphi_t Q_t[\mathbf{X}])_{t=0,\dots,n}$ immediately follows from the assumed square integrability of the cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. Since $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ we have, using the tower property of conditional expectations, see Williams [Wi91], Chap. 9,

$$E\left[\varphi_{t+1} Q_{t+1}[\mathbf{X}] \middle| \mathcal{F}_t\right] = E\left[E\left[\sum_{k=0}^{n} \varphi_k X_k \middle| \mathcal{F}_{t+1} \right] \middle| \mathcal{F}_t \right] \tag{2.60}$$
$$= E\left[\sum_{k=0}^{n} \varphi_k X_k \middle| \mathcal{F}_t \right] = \varphi_t Q_t[\mathbf{X}].$$

This finishes the proof of the lemma. □

Remarks on Deflating and Discounting.

- From the martingale property derived in Lemma 2.9 we immediately have

$$Q_t[\mathbf{X}] = \frac{1}{\varphi_t} E\left[\varphi_{t+1} Q_{t+1}[\mathbf{X}] \middle| \mathcal{F}_t\right] = E\left[\frac{\varphi_{t+1}}{\varphi_t} Q_{t+1}[\mathbf{X}] \middle| \mathcal{F}_t \right]. \tag{2.61}$$

  This implies for the span-deflated price process the identity

$$Q_t[\mathbf{X}] = E\left[Y_{t+1} Q_{t+1}[\mathbf{X}] \middle| \mathcal{F}_t\right], \tag{2.62}$$

  with *span-deflator* Y_{t+1} defined in (2.27). The span-deflator Y_{t+1} is $\mathcal{F}_{t+1}$-measurable, i.e. it is known only at the end of the time period $(t, t+1]$, and not at the beginning of that time period.
- We define the *span-discount* known at the beginning of the time period $(t, t+1]$, i.e. which is observable at time t:

$$D(\mathcal{F}_t) = E\left[Y_{t+1} | \mathcal{F}_t\right] = E\left[\frac{\varphi_{t+1}}{\varphi_t} \middle| \mathcal{F}_t \right] = P(t, t+1). \tag{2.63}$$

  It is often convenient to rewrite (2.62) in terms of the span-discount $D(\mathcal{F}_t)$ instead of the span-deflator Y_{t+1}. The reason is that the span-discount is previsible (known a priori and observable at the market) whereas the span-deflator is always a "hidden variable" that cannot directly be extracted from the actual market information. The basic idea is to change the probability measure P to P^* so that we can change from span-deflators Y_{t+1} to previsible span-discounts $D(\mathcal{F}_t)$.
- If the time interval $(t, t+1]$ is one year, then $D(\mathcal{F}_t) = P(t, t+1)$ is exactly the price of the zero coupon bond with a time to maturity of 1 year at time t, i.e., on this yearly time grid this corresponds to the one-year risk-free investment at time t. Therefore, $D(\mathcal{F}_t)^{-1}$ describes the development of the value of the bank account from time t to time $t+1$. That is, if we invest $B_0 = 1$ units of currency into the bank account at time 0, then the value of this investment at time $t \geq 1$ is given by the *annually risk-free roll-over*:

$$B_t = \prod_{s=0}^{t-1} D(\mathcal{F}_s)^{-1} = \prod_{s=0}^{t-1} E\left[Y_{s+1}|\mathcal{F}_s\right]^{-1} = \exp\left\{\sum_{s=0}^{t-1} r_s\right\}, \tag{2.64}$$

where we have defined

$$r_t = -\log E\left[Y_{t+1}|\mathcal{F}_t\right] = -\log D(\mathcal{F}_t) = -\log P(t, t+1). \tag{2.65}$$

The process $(r_t)_{t=0,\dots,n-1}$ is called the *spot rate process* in discrete time and we have already met it in Exercise 2.3. The process $(B_t)_{t=0,\dots,n}$ is called the value process of the *bank account*.

- The change of probability measure from P to P^* mentioned above will then allow us to change from deflators φ to discounting with the bank account numeraire $(B_t)_{t=0,\dots,n}$. This is going to be derived next.

We define the process $\boldsymbol{\xi} = (\xi_s)_{s=0,\dots,n}$ by $\xi_0 = 1$ and for $s = 1, \dots, n$ we set

$$\xi_s = \prod_{t=0}^{s-1} \frac{Y_{t+1}}{D(\mathcal{F}_t)} = \varphi_s\, B_s. \tag{2.66}$$

Corollary 2.10 $\boldsymbol{\xi} \gg 0$ *is a (normalized) density process w.r.t.* P *and* $\mathbb{F}$*.*

Proof Strict positivity $\gg$ is immediately clear. Moreover, $\boldsymbol{\xi}$ is a $(P, \mathbb{F})$-martingale (which immediately follows from Lemma 2.9 because $(B_t)_{t=0,\dots,n}$ is the price process of the bank account) with normalization $E[\xi_n] = 1$. This proves the claim. □

Since $\boldsymbol{\xi}$ is a density process w.r.t. P and $\mathbb{F}$ we can use it for the following change of measure. For $A \in \mathcal{F}_n = \mathcal{F}$ we define

$$P^*[A] = \int_A \xi_n \, dP = E\left[\xi_n 1_A\right]. \tag{2.67}$$

Lemma 2.11 *The following statements hold:*

(1) *P^* is a probability measure on $(\Omega, \mathcal{F})$ equivalent to P.*
(2) *We have a Radon–Nikodým derivative for $s \le n$*

$$\left.\frac{dP^*}{dP}\right|_{\mathcal{F}_s} = \xi_s, \qquad P\text{-a.s.} \tag{2.68}$$

(3) *Moreover, for $s \le t$ and $A \in \mathcal{F}_t$*

$$P^*\left[A|\,\mathcal{F}_s\right] = \frac{1}{\xi_s} E\left[\xi_t\, 1_A|\,\mathcal{F}_s\right], \qquad P\text{-a.s.} \tag{2.69}$$

Proof The proof of statement (1) follows from Corollary 2.10. The normalization implies that $P^*[\Omega] = E[\xi_n] = 1$, which says that P^* is a probability measure on

$(\Omega, \mathcal{F})$. Moreover, $\xi_n > 0$, P-a.s., implies that $P^* \sim P$, i.e. they are equivalent measures.

Next we prove statement (2). Note that for any set $C \in \mathcal{F}_s$

$$P^*[C] = E\left[\xi_n 1_C\right] = E\left[E\left[\xi_n \middle| \mathcal{F}_s\right] 1_C\right] = E\left[\xi_s 1_C\right], \tag{2.70}$$

using the martingale property of $\boldsymbol{\xi}$ in the last step. Therefore, ξ_s is the resulting density on $\mathcal{F}_s$.

Finally we prove (3). Note that we have for any set $C \in \mathcal{F}_s$, $s \le t \le n$, and using (2.70) in the 4th step

$$\begin{aligned} E^*\left[1_C\, 1_A\right] = E\left[1_C\, \xi_n\, 1_A\right] &= E\left[1_C\, E\left[\xi_n\, 1_A \middle| \mathcal{F}_s\right]\right] \\ &= E\left[\xi_s \left(1_C\, \frac{1}{\xi_s}\, E\left[\xi_n\, 1_A \middle| \mathcal{F}_s\right]\right)\right] \\ &= E^*\left[1_C\, \frac{1}{\xi_s}\, E\left[\xi_n\, 1_A \middle| \mathcal{F}_s\right]\right] \\ &= E^*\left[1_C\, \frac{1}{\xi_s}\, E\left[1_A\, E\left[\xi_n \middle| \mathcal{F}_t\right] \middle| \mathcal{F}_s\right]\right] \\ &= E^*\left[1_C\, \frac{1}{\xi_s}\, E\left[\xi_t\, 1_A \middle| \mathcal{F}_s\right]\right]. \end{aligned} \tag{2.71}$$

Since this holds for all $C \in \mathcal{F}_s$ the claim follows by the definition of the conditional expectation w.r.t. P^*. □

Item (3) of Lemma 2.11 immediately provides the next corollary:

Corollary 2.12 *For $s \le t$ we have*

$$E^*\left[B_t^{-1} Q_t[\mathbf{X}] \middle| \mathcal{F}_s\right] = \frac{1}{\xi_s}\, E\left[\xi_t B_t^{-1} Q_t[\mathbf{X}] \middle| \mathcal{F}_s\right]. \tag{2.72}$$

If we apply Corollary 2.12 for $s = t - 1$ and use Lemma 2.9 we obtain

$$\begin{aligned} E^*\left[B_t^{-1} Q_t[\mathbf{X}] \middle| \mathcal{F}_{t-1}\right] &= \frac{1}{\xi_{t-1}}\, E\left[\xi_t B_t^{-1} Q_t[\mathbf{X}] \middle| \mathcal{F}_{t-1}\right] \\ &= \frac{1}{\xi_{t-1}}\, E\left[\varphi_t Q_t[\mathbf{X}] \middle| \mathcal{F}_{t-1}\right] \\ &= \frac{1}{\xi_{t-1}}\, \varphi_{t-1} Q_{t-1}[\mathbf{X}] \;=\; B_{t-1}^{-1} Q_{t-1}[\mathbf{X}]. \end{aligned} \tag{2.73}$$

We have just proved the following corollary, compare to Lemma 2.9.

Corollary 2.13 *For every cash flow* $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ *the bank account numeraire discounted price process satisfies:*

$$\left(B_t^{-1} Q_t[\mathbf{X}]\right)_{t=0,\dots,n} \quad \textit{is a } (P^*, \mathbb{F})\textit{-martingale.} \tag{2.74}$$

Remark 2.14 (**Real world and equivalent martingale measures**)

- Every price process $(Q_t[\mathbf{X}])_{t=0,\dots,n}$ constructed by (2.40) is a $(P, \mathbb{F})$-martingale if deflated with the fixed chosen deflator φ, see Lemma 2.9. This price process is also a $(P^*, \mathbb{F})$-martingale if it is discounted with the bank account numeraire $(B_t)_{t=0,\dots,n}$ and under the change of measure (2.67), see Corollary 2.13.
- The probability measure P is called the *real world probability measure*, *objective measure* or *physical measure* and it describes the probability law as it can be observed in real world. If, for pricing purposes, we switch to the bank account numeraire discounting we also need to transform the underlying probability measure, and the resulting (equivalent) probability measure P^* is called the *equivalent martingale measure*, *pricing measure*, or *risk neutral measure*.
- If we work with financial instruments only, then it is often easier to work under P^*. If we additionally have insurance products, then one usually works under P. Therefore, actuaries need to fully understand the connection between these two measures.
- For the equivalent martingale measure P^* we always choose the bank account numeraire $(B_t)_{t=0,\dots,n}$ for discounting. In general, if $(A_t)_{t=0,\dots,n}$ is any strictly positive and normalized price process, then we could also choose this price process as a numeraire and find the appropriate equivalent measure $P^A \sim P$ such that all price processes $(A_t^{-1} Q_t[\mathbf{X}])_{t=0,\dots,n}$ are $(P^A, \mathbb{F})$-martingales. For more on this subject we refer to Sect. 4.3, and to Wüthrich–Merz [WM13], Sect. 11.2.

In the one-period model we obtain identity

$$Q_0[\mathbf{X}] = D(\mathcal{F}_0)\, E^*[\,Q_1[\mathbf{X}]\,] = E[\,Y_1\, Q_1[\mathbf{X}]\,]. \tag{2.75}$$

This shows the difference between discounting with $\mathcal{F}_0$-measurable span-discount $D(\mathcal{F}_0)$ under P^* and deflating with $\mathcal{F}_1$-measurable span-deflator Y_1 under P.

Remark 2.15 (**Fundamental theorem of asset pricing (FTAP)**)

- The *efficient market hypothesis* in its strong form assumes that all deflated price processes
$$\widetilde{Q}_t = \varphi_t\, Q_t[\mathbf{X}], \qquad t = 0, \dots, n, \tag{2.76}$$
form $(P, \mathbb{F})$-martingales. This implies for the expected net gains, $t > s$,
$$E\left[\widetilde{Q}_t - \widetilde{Q}_s \middle| \mathcal{F}_s\right] = 0, \tag{2.77}$$
which means that there cannot be strictly positive expected net gains without any downside risks.

- The *efficient market hypothesis* in its weak form assumes that “there is no free lunch”, i.e. there do not exist (appropriately defined) self-financing trading strategies with positive expected gains and without any downside risks. In a finite and discrete time model this is equivalent to the existence of an equivalent martingale measure for the bank account numeraire discounted price processes (see e.g. Theorem 2.6 in Lamberton–Lapeyre [LL91]). The proof for a finite probability space is essentially an exercise in linear algebra (see the toy model in Sect. 2.2.3); in more general settings the characterization is more delicate and typically referred to as the fundamental theorem of asset pricing (FTAP), see Delbaen–Schachermayer [DS94, DS06] and Föllmer–Schied [FS11].
- Summarizing: the existence of an equivalent martingale measure rules out appropriately defined arbitrage (which is the easier direction). The opposite direction that no-arbitrage defined in the right way implies the existence of an equivalent martingale measure is more delicate. In this lecture we will always identify no-arbitrage (defined in the right way) with the existence of an equivalent martingale measure for the bank account numeraire in the sense of Corollary 2.13 and Lemma 2.9, respectively.
- In complete markets, the equivalent martingale measure is unique and we can perfectly replicate any cash flow by traded instruments (for an example see Sect. 2.2.3). The uniqueness of the equivalent martingale measure also implies uniqueness of the state price deflator.
- In incomplete markets, where we have cash flows that cannot be perfectly replicated by traded instruments, we typically have more than one equivalent martingale measure (and state price deflator), and we need an economic model to decide which measure is appropriate for calculating prices of non-traded instruments, see Föllmer–Schied [FS11] or Malamud et al. [MTW08]. This in particular applies to insurance products.

Toy Example (Revisited).
We revisit the toy example of Sect. 2.2.3. We transform our probability measure according to Lemma 2.11 (here we work in a one-period model with $Q_0 = Q$): the equivalent martingale measure is given by

$$p^*(\omega_i) = \xi_1(\omega_i)\ p(\omega_i) = \frac{\varphi_1(\omega_i)}{E[\varphi_1]}\ p(\omega_i) = \frac{Q(\omega_i)}{Q\left[\mathbf{Z}^{(1)}\right]}. \tag{2.78}$$

Hence, from (2.35) and (2.38) we obtain

$$Q[\mathbf{X}] = E[\varphi_1\ X_1] = \sum_{i=1}^{2} Q(\omega_i)\ X_1(\omega_i), \tag{2.79}$$

$$Q[\mathbf{X}] = B_1^{-1}\ E^*[X_1] = \sum_{i=1}^{2} Q(\omega_i)\ X_1(\omega_i), \tag{2.80}$$

with

$$B_1^{-1} = E\left[\varphi_1\right] = Q\left[\mathbf{Z}^{(1)}\right], \tag{2.81}$$

which is deterministic at time 0. Hence under P^* we have

$$Q\left[\mathbf{X}\right] = B_1^{-1}\, E^*\left[X_1\right] = Q\left[\mathbf{Z}^{(1)}\right] E^*\left[X_1\right]. \tag{2.82}$$

This leads to the following table with $p^*(\omega_1) = 0.368$:

	$\mathbf{Z}^{(1)}$	Asset A	Asset B
Market price $Q_0[\cdot]$ at time 0	0.95	1.65	1.00
Payout if state ω_1 occurs	1	3	2
Payout if state ω_2 occurs	1	1	0.5
P^* expected payout	1	1.737	1.053
P^* expected return	5.26 %	5.26 %	5.26 %

which is the martingale property of the discounted cash flow $Q\left[\mathbf{Z}^{(1)}\right] X_1$ w.r.t. the equivalent martingale measure P^*. □

Exercise 2.6 We revisit the discrete time Vasiček model given in Exercise 2.3. The spot rate dynamics $(r_t)_{t=0,\ldots,n}$ was given by $r_0 > 0$ (fixed) and for $t \geq 1$

$$r_t = b + \beta r_{t-1} + \rho \varepsilon_t, \tag{2.83}$$

for given $b, \beta, \rho > 0$, and $(\varepsilon_t)_{t=0,\ldots,n}$ is $\mathbb{F}$-adapted with ε_t independent of $\mathcal{F}_{t-1}$ for all $t = 1, \ldots, n$ and standard Gaussian distributed under the *real world probability measure* P.

The deflator $\boldsymbol{\varphi}$ was then defined by

$$\varphi_t = \exp\left\{ -\sum_{k=1}^{t} \left[r_{k-1} + \frac{\lambda^2}{2} r_{k-1}^2 \right] - \sum_{k=1}^{t} \lambda r_{k-1} \varepsilon_k \right\}, \tag{2.84}$$

for given $\lambda \in \mathbb{R}$.

- Calculate the span-discount $D(\mathcal{F}_t) = P(t, t+1)$ from the span-deflator

$$Y_{t+1} = \frac{\varphi_{t+1}}{\varphi_t} = \exp\left\{ -\left[r_t + \frac{\lambda^2}{2} r_t^2 \right] - \lambda r_t \varepsilon_{t+1} \right\}, \tag{2.85}$$

  and show that the model is well-defined.
- Prove that the density process $\boldsymbol{\xi} = (\xi_t)_{t=0,\ldots,n}$ is given by

$$\xi_t = \exp\left\{ -\sum_{k=1}^{t} \frac{\lambda^2}{2} r_{k-1}^2 - \sum_{k=1}^{t} \lambda r_{k-1} \varepsilon_k \right\}, \tag{2.86}$$

  where an empty sum is set equal to zero.

- Prove that
$$\varepsilon_t^* = \varepsilon_t + \lambda r_{t-1} \tag{2.87}$$
has, conditionally given $\mathcal{F}_{t-1}$, a standard Gaussian distribution under the *equivalent martingale measure* $P^* \sim P$, obtained from the density process $\boldsymbol{\xi}$ as in Lemma 2.11.
Hint: use the moment generating function and Lemma 2.11.
- Prove that (2.87) implies for the spot rate process $(r_t)_{t=0,\dots,n}$: $r_0 > 0$ (fixed) and for $t \geq 1$
$$r_t = b + (\beta - \lambda\rho) r_{t-1} + \rho \varepsilon_t^*, \tag{2.88}$$
where $(\varepsilon_t^*)_{t=0,\dots,n}$ is $\mathbb{F}$-adapted with ε_t^* independent of $\mathcal{F}_{t-1}$ for all $t = 1, \dots, n$ and standard Gaussian distributed under the equivalent martingale measure P^*.
- Calculate the zero coupon bond prices $t < m$ (see also Exercise 2.5)
$$P(t,m) = E^*\left[\exp\left\{ -\sum_{s=t}^{m-1} r_s \right\} \middle| \mathcal{F}_t \right] = \exp\left\{ A(t,m) - r_t\, B(t,m) \right\}. \tag{2.89}$$

$\square$

Remark on Exercise 2.6. In (2.86) we calculate the density process $(\xi_t)_{t=0,\dots,n}$ for the discrete time Vasiček model. It depends on the parameter $\lambda \in \mathbb{R}$. We see that if $\lambda = 0$, the density process is identically equal to 1, and henceforth $P^* = P$. Therefore, λ models the difference between the real world probability measure P and the equivalent martingale measure P^* which is in economic theory explained through the market risk aversion. Therefore, λ is often called the *market price of risk* parameter and explains the aggregate market risk aversion (in our Vasiček model). In general, a higher (market) risk aversion explains lower prices because the more risk averse someone is, the less he is willing to accept risky positions.

Conclusions:

- We have found three different ways to value cash flows $\mathbf{X}$:

1. via linear, positive and normalized functionals Q,
2. via deflators $\boldsymbol{\varphi}$ under the real world measure P,
3. via the bank account numeraire $(B_t)_t$ under equivalent martingale measures P^*.

- The advantage of using equivalent martingale measures is that the discount factor is a priori known (previsible), which means that we have a state independent discount factor (for the one-period risk-free roll-over). The main disadvantage of using equivalent martingale measures is that the concept is not straightforward for the calibration of real world events and prediction cannot be done under equivalent martingale measures.

- By contrast, deflators are calculated using the real world probability measure (expressing market risk aversion). Moreover, as shown below, they clearly describe the dependence structures (also between deflators and cash flows). From a practical point of view, deflators allow us to model embedded (financial) options and guarantees in insurance policies, and they are therefore preferred especially by actuaries who value life insurance products that contain both financial and insurance technical risk factors.

2.6 Insurance Technical and Financial Variables

2.6.1 Choice of Numeraire

Choose a cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. For practical purposes in insurance applications it makes sense to factorize the payments X_k into an appropriate financial basis $\mathcal{U}_k$, $k = 0, \ldots, n$, and the number of units Λ_k of this basis needed. Assume that we can split the payments X_k as follows

$$X_k = \Lambda_k \, U_k^{(k)}, \quad k = 0, \ldots, n, \tag{2.90}$$

where the variable $U_t^{(k)}$ denotes the price of one unit of the financial instrument $\mathcal{U}_k$ at time $t = 0, \ldots, n$, and (for non-zero X_k and $U_k^{(k)}$, respectively)

$$\Lambda_k = \frac{X_k}{U_k^{(k)}}, \quad k = 0, \ldots, n, \tag{2.91}$$

gives the number of units (insurance technical variable) that we need to hold in order to replicate X_k with financial instrument $\mathcal{U}_k$. This means that we measure insurance liabilities in units $\mathcal{U}_k$ (financial instruments) which have price processes $(U_t^{(k)})_{t=0,\ldots,n}$, and in insurance technical variables Λ_k.

We emphasize that one should clearly distinguish between the financial instrument $\mathcal{U}_k$ (which can be understood as a contract) and its price process given by

$$(U_t^{(k)})_{t=0,\ldots,n} = \left(U_0^{(k)}, U_1^{(k)}, \ldots, U_k^{(k)}, \ldots, U_n^{(k)} \right). \tag{2.92}$$

The financial instrument $\mathcal{U}_k$ can either be an asset or a liability and it can be purchased or sold at given prices $U_t^{(k)}$ in time points t.

Assume that the price process $(U_t^{(k)})_{t=0,\ldots,k} \gg 0$ is strictly positive (up to time k). Then $(U_t^{(k)})_{t=0,\ldots,k}$ should be used as a *numeraire* to study the liability X_k, thus, every payment X_k should be studied in its appropriate unit (and numeraire). We have already met this idea in Remark 2.14.

Examples of Units for Numeraires.

- cash in different currencies like CHF, USD, EUR
- indexed cash typically described by an inflation index, salary index, claims inflation index, medical expenses index, etc.
- company share, stock, private equity, real estate, etc. (be careful with strict positivity of price processes)
- asset portfolio (with strictly positive price process)

Examples of Insurance Technical Events.

- death, survival, disability
- car accident, fire event, burglary, nuclear power accident
- medical expenses, workman's compensation

We would like to factorize the filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ into a product space such that we get an independent decoupling into:

$$\mathbb{T} = (\mathcal{T}_t)_{t=0,\ldots,n} \quad \text{filtration of insurance technical events,} \tag{2.93}$$

$$\mathbb{G} = (\mathcal{G}_t)_{t=0,\ldots,n} \quad \text{filtration of financial events,} \tag{2.94}$$

with for all $t = 0, \ldots, n$

$$\mathcal{F}_t = \sigma(\mathcal{T}_t, \mathcal{G}_t) = \text{smallest } \sigma\text{-field containing all events of } \mathcal{T}_t \text{ and } \mathcal{G}_t. \tag{2.95}$$

We assume that under P the two filtrations $\mathbb{T}$ and $\mathbb{G}$ are independent, thus, $\mathbb{F}$ can be decoupled into two independent filtrations, one describing insurance technical events $\mathbb{T}$ and one describing financial events $\mathbb{G}$. That is, the real world probability measure P admits a product representation (by a slight abuse of notation)

$$P = P_{\mathbb{T}} \times P_{\mathbb{G}}, \tag{2.96}$$

with $P_{\mathbb{T}}$ describing insurance technical risks $\mathbf{\Lambda} = (\Lambda_0, \ldots, \Lambda_n)$ which will be $\mathbb{T}$-adapted and with $P_{\mathbb{G}}$ describing financial price processes $(U_t^{(k)})_{t=0,\ldots,n}$ which will be $\mathbb{G}$-adapted. This decoupling will be crucial in the sequel of this manuscript and is explained in the next assumption.

Assumption 2.16 Assume that $\mathbb{T}$ and $\mathbb{G}$ are two independent filtrations on the probability space $(\Omega, \mathcal{F}, P)$ that generate filtration $\mathbb{F}$ according to (2.95). Assume that the cash flows $\mathbf{X}$ of interest are of the form

$$\mathbf{X} = (\Lambda_0 \, U_0^{(0)}, \ldots, \Lambda_n \, U_n^{(n)}), \tag{2.97}$$

with $\mathbf{\Lambda} \in L^2_{n+1}(P, \mathbb{T})$ and $(U_t^{(k)})_{t=0,\ldots,n} \in L^2_{n+1}(P, \mathbb{G})$ for all $k = 0, \ldots, n$. Moreover, assume that the chosen (fixed) deflator $\boldsymbol{\varphi} \in L^2_{n+1}(P, \mathbb{F})$ factorizes $\varphi_k = \varphi_k^{\mathbb{T}} \varphi_k^{\mathbb{G}}$ for all $k = 0, \ldots, n$ such that $\boldsymbol{\varphi}^{\mathbb{T}} = (\varphi_k^{\mathbb{T}})_{k=0,\ldots,n}$ is $\mathbb{T}$-adapted and $\boldsymbol{\varphi}^{\mathbb{G}} = (\varphi_k^{\mathbb{G}})_{k=0,\ldots,n}$ is $\mathbb{G}$-adapted.

The valuation of these cash flows $\mathbf{X} = (\Lambda_0\, U_0^{(0)}, \ldots, \Lambda_n\, U_n^{(n)}) \in L^2_{n+1}(P, \mathbb{F})$ is then under Assumption 2.16 given by

$$\begin{aligned} \varphi_t\; Q_t\,[\mathbf{X}] &= E\left[\sum_{k=0}^{n} \varphi_k\; X_k \middle| \mathcal{F}_t\right] \qquad (2.98)\\ &= E\left[\sum_{k=0}^{n} \varphi_k^{\mathbb{T}}\; \Lambda_k\; \varphi_k^{\mathbb{G}}\; U_k^{(k)} \middle| \mathcal{T}_t, \mathcal{G}_t\right]\\ &= \sum_{k=0}^{n} E\left[\varphi_k^{\mathbb{T}} \Lambda_k \middle| \mathcal{T}_t\right]\; E\left[\varphi_k^{\mathbb{G}} U_k^{(k)} \middle| \mathcal{G}_t\right]. \end{aligned}$$

Remarks.

- The term $E[\varphi_k^{\mathbb{T}} \Lambda_k | \mathcal{T}_t]$ describes the price of the insurance technical cover in units of the corresponding numeraire instrument $\mathcal{U}_k$ at time t. $\boldsymbol{\varphi}^{\mathbb{T}}$ defines the *insurance technical loading*, the so-called *probability distortion*, of the insurance technical price. This is further outlined in Sect. 2.6.2.
- The term $E[\varphi_k^{\mathbb{G}} U_k^{(k)} | \mathcal{G}_t]$ relates to the price of one unit of the financial instrument $\mathcal{U}_k$ at time t, see also Sect. 2.6.2 on probability distortions below. From this we conclude that $\boldsymbol{\varphi}^{\mathbb{G}}$ should be obtained from financial market data, because it should reflect asset prices at (traded) financial markets appropriately. For example, we can use the Vasiček model, proposed in Exercise 2.3, and calibrate the model to financial market data, see Wüthrich–Bühlmann [WB08] and Wüthrich–Merz [WM13].
- We have separated the pricing problem into two independent pricing problems, one for pricing insurance technical cover in units of a numeraire instrument and one for pricing units of financial instruments. This split looks very natural, but in practice one needs to be careful with its applications. Especially in non-life insurance, it is very difficult to find such an "orthogonal" (independent) split, since the severities of the claims often depend on the financial market and the split is far from non-trivial. For example, if we consider workman's compensation (which pays the salary when someone is injured or sick), it is very difficult to describe the dependence structure between (1) the salary height, (2) the length of the sickness (which may have a mental cause), (3) the state of the job market, (4) the state of the financial market, and (5) the political environment.
- The financial economy including insurance products could also be defined in other ways that would allow for similar splits. For an example we refer to Malamud et al. [MTW08]. There one starts with a complete financial market model described by the financial filtration. Then one introduces insurance products that enlarge the underlying financial filtration. This enlargement in general makes the market incomplete (but still arbitrage-free) and adds idiosyncratic risks to the economic model. Finally, one defines the "hedgeable" filtration that exactly describes the part of the insurance claims that can be described via financial market movements. The remaining parts are then the insurance technical risks. For an analysis of this split in terms of projections we also refer to Happ et al. [HMW15].

2.6.2 Probability Distortion

In this section we discuss the factorization of the deflator $\varphi_k = \varphi_k^{\mathbb{T}} \varphi_k^{\mathbb{G}}$ given in Assumption 2.16. The choice of the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$ needs some care in order to obtain a reasonable model, as we will see shortly.

(1) Firstly, we choose $\boldsymbol{\varphi}^{\mathbb{T}} \gg 0$ and $\boldsymbol{\varphi}^{\mathbb{G}} \gg 0$ which is in line with $\boldsymbol{\varphi} \gg 0$. Moreover, $\boldsymbol{\varphi}^{\mathbb{T}} \in L^2_{n+1}(P, \mathbb{T})$ is a necessary assumption which follows from $\boldsymbol{\varphi} \in L^2_{n+1}(P, \mathbb{F})$ and the independence and $\mathbb{T}$-adaptedness in Assumption 2.16.

(2) Secondly, to avoid ambiguity, we set for all $t = 0, \ldots, n$

$$E\left[\varphi_t^{\mathbb{T}}\right] = 1. \tag{2.99}$$

Otherwise, the decoupling into a product $\varphi_t = \varphi_t^{\mathbb{T}} \varphi_t^{\mathbb{G}}$ is not unique, which can easily be seen by multiplying and dividing both terms by the same positive constant.

(3) Thirdly, we assume that the sequence $(\varphi_t^{\mathbb{T}})_{t=0,\ldots,n}$ is a $(P, \mathbb{T})$-martingale, i.e.

$$E\left[\varphi_{t+1}^{\mathbb{T}} \middle| \mathcal{T}_t\right] = \varphi_t^{\mathbb{T}}. \tag{2.100}$$

Of course, the normalization (2.99) is then an easy consequence of the requirement

$$\varphi_0^{\mathbb{T}} = 1. \tag{2.101}$$

Under Assumption 2.16 and assuming (1)–(3) for the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$ we see that

$$(\varphi_t^{\mathbb{T}})_{t=0,\ldots,n} \text{ is a density process w.r.t. } \mathbb{T} \text{ and } P. \tag{2.102}$$

This allows us to define an equivalent probability measure $P_{\mathbb{T}}^* \sim P$ on $(\Omega, \mathcal{T}_n, P)$ via the Radon–Nikodým derivative

$$\left.\frac{dP_{\mathbb{T}}^*}{dP}\right|_{\mathcal{T}_n} = \varphi_n^{\mathbb{T}}. \tag{2.103}$$

Moreover, we define the price process for the insurance technical variable Λ_k as follows: for $t \leq k$

$$\Lambda_{t,k} = \frac{1}{\varphi_t^{\mathbb{T}}} E\left[\varphi_k^{\mathbb{T}} \Lambda_k \middle| \mathcal{T}_t\right]. \tag{2.104}$$

Lemma 2.17 *Assume Assumption 2.16 and (2.102) hold true. The probability distorted process*

$$\left(\varphi_t^{\mathbb{T}} \Lambda_{t,k}\right)_{t=0,\ldots,k} \quad \textit{forms a } (P, \mathbb{T})\textit{-martingale.} \tag{2.105}$$

The process

$$\left(\Lambda_{t,k}\right)_{t=0,\ldots,k} \quad \textit{forms a } (P^*_{\mathbb{T}}, \mathbb{T})\textit{-martingale.} \tag{2.106}$$

Proof of Lemma 2.17. The first claim follows similarly to Lemma 2.9 and uses the tower property of conditional expectations, see Williams [Wi91]. The second claim follows similarly to Corollary 2.13 and equality (2.73). Note that here the numeraire is equal to 1 (due to our choice of the density process). □

An immediate consequence of Lemma 2.17 is the following corollary:

Corollary 2.18 *Under the assumptions of Lemma 2.17 we have*

$$\Lambda_{t,k} = \frac{1}{\varphi_t^{\mathbb{T}}}\, E\left[\varphi_k^{\mathbb{T}} \Lambda_k \middle| \mathcal{T}_t\right] = E^*_{\mathbb{T}}\left[\Lambda_k \middle| \mathcal{T}_t\right]. \tag{2.107}$$

This has further consequences:

Theorem 2.19 *Under the assumptions of Lemma 2.17 and (2.59) we obtain that the price process $(U_t^{(k)})_{t=0,\ldots,k}$ of the financial instrument $\mathcal{U}_k$ satisfies for $t < k$*

$$U_t^{(k)} = \frac{1}{\varphi_t^{\mathbb{G}}}\, E\left[\varphi_{t+1}^{\mathbb{G}} U_{t+1}^{(k)} \middle| \mathcal{G}_t\right]. \tag{2.108}$$

Proof of Theorem 2.19. We define the cash flow $\mathbf{X} = U_k^{(k)} \mathbf{Z}^{(k)} = (0, \ldots, 0, U_k^{(k)}, 0, \ldots, 0) \in L^2_{n+1}(P, \mathbb{F})$. Note that in fact the cash flow $\mathbf{X}$ is in $L^2_{n+1}(P, \mathbb{G})$. The martingale property (2.59), Assumption 2.16 and Corollary 2.18 imply for $t < s \le k$

$$\begin{aligned} \varphi_t Q_t[\mathbf{X}] &= E\left[\varphi_s Q_s[\mathbf{X}] \middle| \mathcal{F}_t\right] \;=\; E\left[\varphi_s U_s^{(k)} \middle| \mathcal{F}_t\right] \\ &= E\left[\varphi_s^{\mathbb{T}} \varphi_s^{\mathbb{G}} U_s^{(k)} \middle| \mathcal{F}_t\right] \\ &= E\left[\varphi_s^{\mathbb{T}} \middle| \mathcal{T}_t\right] E\left[\varphi_s^{\mathbb{G}} U_s^{(k)} \middle| \mathcal{G}_t\right] \\ &= \varphi_t^{\mathbb{T}}\, E\left[\varphi_s^{\mathbb{G}} U_s^{(k)} \middle| \mathcal{G}_t\right]. \end{aligned} \tag{2.109}$$

This implies that

$$U_t^{(k)} = Q_t[\mathbf{X}] = \frac{1}{\varphi_t^{\mathbb{G}}} E\left[\varphi_s^{\mathbb{G}} U_s^{(k)} \middle| \mathcal{G}_t\right]. \tag{2.110}$$

Henceforth $(\varphi_t^{\mathbb{G}} U_t^{(k)})_{t=0,\ldots,k}$ is a $(P, \mathbb{G})$-martingale. □

Of course, Theorem 2.19 is not really a surprise because it (only) says that a pure financial price process needs to have the martingale property w.r.t. the financial market model $(P, \mathbb{G})$ in order to be free of arbitrage.

Corollary 2.18 and Theorem 2.19 imply that we can study the insurance technical variables $\mathbf{\Lambda}$ and the price processes of the financial instruments $\mathcal{U}_k$ independently. The valuation of the outstanding loss liabilities

$$\mathbf{X}_{(k)} = (0, \ldots, 0, \Lambda_k U_k^{(k)}, \ldots, \Lambda_n U_n^{(n)}) \in L^2_{n+1}(P, \mathbb{F}) \tag{2.111}$$

at time $t \le k$ can then easily be done, and the reserves are given by

$$R_t^{(k)} = Q_t\left[\mathbf{X}_{(k)}\right] = \sum_{s=k}^{n} \frac{1}{\varphi_t^{\mathbb{T}}} E\left[\varphi_s^{\mathbb{T}} \Lambda_s \middle| \mathcal{T}_t\right] \frac{1}{\varphi_t^{\mathbb{G}}} E\left[\varphi_s^{\mathbb{G}} U_s^{(s)} \middle| \mathcal{G}_t\right] \tag{2.112}$$
$$= \sum_{s=k}^{n} \Lambda_{t,s}\, U_t^{(s)}.$$

Conclusions. Under the product space Assumption 2.16, the assumption (2.102) that the insurance technical deflator is a density process w.r.t. $\mathbb{T}$ and P, and under the no-arbitrage assumption (2.59) we obtain that we can separate the valuation problem into two independent valuation problems:

(1) the insurance technical processes $(\Lambda_{t,k})_{t=0,\dots,k}$, $k = 0, \dots, n$, describe the probability distorted developments of the predictions of the insurance technical variables Λ_k if we increase the information $\mathcal{T}_t \to \mathcal{T}_{t+1}$;
(2) the financial processes $(U_t^{(k)})_{t=0,\dots,k}$, $k = 0, \dots, n$, describe the price processes of the financial instruments $\mathcal{U}_k$ in the financial market model $(\Omega, \mathcal{G}_n, P, \mathbb{G})$.

Example 2.7 (*Best-estimate predictions and reserves*) Choose $\boldsymbol{\varphi}^{\mathbb{T}} \equiv 1$. Hence, $\boldsymbol{\varphi}^{\mathbb{T}}$ gives an admissible probability distortion (normalized martingale). This choice implies for the insurance technical process at time $t \le k$

$$\Lambda_{t,k} = E\left[\Lambda_k \middle| \mathcal{T}_t\right], \tag{2.113}$$

i.e. $\Lambda_{t,k}$ is simply the "best-estimate" prediction of Λ_k based on the information $\mathcal{T}_t$ (conditional expectation which has minimal conditional prediction variance). If we use this probability distortion in the reserves definition (2.112), then we call $R_t^{(k)}$ best-estimate reserves at time $t < k$ for the outstanding liabilities $\mathbf{X}_{(k)}$. □

Exercise 2.8 (*Esscher premium*) We choose a positive random variable Y on the underlying filtered probability space $(\Omega, \mathcal{T}_n, P, \mathbb{T})$ such that for some $\alpha > 0$ the following moment generating function exists

$$M_Y(2\alpha) = E\left[\exp\{2\alpha Y\}\right] < \infty. \tag{2.114}$$

Then we define the probability distortion

$$\varphi_t^{\mathbb{T}} = \frac{E\left[\exp\{\alpha Y\} \middle| \mathcal{T}_t\right]}{E\left[\exp\{\alpha Y\}\right]} = \frac{E\left[\exp\{\alpha Y\} \middle| \mathcal{T}_t\right]}{M_Y(\alpha)}. \tag{2.115}$$

(1) Prove that $\boldsymbol{\varphi}^{\mathbb{T}} \gg 0$. Moreover, prove that $\boldsymbol{\varphi}^{\mathbb{T}} \in L_{n+1}^2(P, \mathbb{T})$.
(2) Show that $(\varphi_t^{\mathbb{T}})_{t=0,\dots,n}$ is a density process w.r.t. $\mathbb{T}$ and P.

Assume that $\mathbf{X}_k = (0, \ldots, 0, X_k, 0, \ldots, 0)$ with $X_k = \Lambda_k U_k^{(k)}$. Choose $Y = \Lambda_k$ and $t < k$. Prove under Assumption 2.16 and (2.59) that

$$Q_t[\mathbf{X}_k] = \frac{1}{E\left[\exp\{\alpha\Lambda_k\} \middle| \mathcal{T}_t\right]} E\left[\Lambda_k e^{\alpha\Lambda_k} \middle| \mathcal{T}_t\right] U_t^{(k)}. \tag{2.116}$$

If we define the conditional moment generating function by

$$M_{\Lambda_k|\mathcal{T}_t}(\alpha) = E\left[\exp\{\alpha\Lambda_k\} \middle| \mathcal{T}_t\right], \tag{2.117}$$

then the term

$$\Lambda_{t,k} = \frac{d}{dr} \log M_{\Lambda_k|\mathcal{T}_t}(r)\Big|_{r=\alpha} = M_{\Lambda_k|\mathcal{T}_t}(\alpha)^{-1} E\left[\Lambda_k e^{\alpha\Lambda_k} \middle| \mathcal{T}_t\right] \tag{2.118}$$

describes the Esscher premium of Λ_k at time $t < k$, see Bühlmann [Bü80] and Gerber–Pafumi [GP98].

(3) Prove that the Esscher premium (2.118) is strictly increasing in α.

Remark: α plays the role of the risk aversion. □

Exercise 2.9 (*Expected shortfall*) Choose an absolutely continuous and integrable random variable Y on the filtered probability space $(\Omega, \mathcal{T}_n, P, \mathbb{T})$. Denote the distribution function of Y by $F_Y(x) = P[Y \le x]$ and the generalized inverse by $F_Y^{\leftarrow}$, where $F_Y^{\leftarrow}(u) = \inf\{x | F_Y(x) \ge u\}$. Henceforth, the Value-at-Risk of Y at level $1 - \alpha \in (0, 1)$ is given by

$$\mathrm{VaR}_{1-\alpha}(Y) = F_Y^{\leftarrow}(1-\alpha). \tag{2.119}$$

We obtain, see also Sect. 1.2.1 in Wüthrich [Wü13],

$$\begin{aligned} P\left[Y > \mathrm{VaR}_{1-\alpha}(Y)\right] &= 1 - P\left[Y \le \mathrm{VaR}_{1-\alpha}(Y)\right] \\ &= 1 - F_Y\left(\mathrm{VaR}_{1-\alpha}(Y)\right) \\ &= 1 - F_Y\left(F_Y^{\leftarrow}(1-\alpha)\right) = \alpha. \end{aligned} \tag{2.120}$$

Choose $c \in (0, 1)$ and define (note that Y is $\mathcal{T}_n$-measurable)

$$\varphi_n^{\mathbb{T}} = (1-c) + \frac{c}{\alpha} 1_{\{Y > \mathrm{VaR}_{1-\alpha}(Y)\}}, \tag{2.121}$$

and for $t < n$

$$\varphi_t^{\mathbb{T}} = E\left[\varphi_n^{\mathbb{T}} \middle| \mathcal{T}_t\right]. \tag{2.122}$$

(1) Prove that $\boldsymbol{\varphi}^{\mathbb{T}} \gg 0$. Moreover, prove that $\boldsymbol{\varphi}^{\mathbb{T}} \in L^2_{n+1}(P, \mathbb{T})$.
(2) Show that $(\varphi_t^{\mathbb{T}})_{t=0,\ldots,n}$ is a density process w.r.t. $\mathbb{T}$ and P.

(3) Assume that $\mathbf{X}_k = (0, \ldots, 0, X_k, 0, \ldots, 0)$ with $X_k = \Lambda_k U_k^{(k)}$. Choose $Y = \Lambda_k$ and $t < k$. Under Assumption 2.16 and (2.59) show that

$$Q_t[\mathbf{X}_k] = \left\{ \beta_t E[\Lambda_k | \mathcal{T}_t] + (1-\beta_t) \frac{E\left[\Lambda_k 1_{\{\Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k)\}} \middle| \mathcal{T}_t\right]}{P\left[\Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k) \middle| \mathcal{T}_t\right]} \right\} U_t^{(k)}, \tag{2.123}$$

with so-called credibility weights

$$\beta_t = \frac{1-c}{(1-c) + c\, \frac{P\left[\Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k) | \mathcal{T}_t\right]}{\alpha}}. \tag{2.124}$$

We define the conditional probability

$$\alpha_t = P\left[\Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k) \middle| \mathcal{T}_t\right], \tag{2.125}$$

which says

$$\mathrm{VaR}_{1-\alpha}(\Lambda_k) = \mathrm{VaR}_{1-\alpha_t}(\Lambda_k | \mathcal{T}_t), \tag{2.126}$$

where $\mathrm{VaR}_{1-\alpha_t}(\Lambda_k|\mathcal{T}_t)$ denotes the Value-at-Risk of $\Lambda_k|\mathcal{T}_t$ at level $1-\alpha_t$. Henceforth, the credibility weight is given by

$$\beta_t = \frac{1-c}{(1-c) + c\, \frac{\alpha_t}{\alpha}}. \tag{2.127}$$

The last term in the bracket of (2.123) can be interpreted as the expected shortfall of $\Lambda_k|\mathcal{T}_t$ at level $1-\alpha_t$. We highlight this for $t = 0$. Then we have $\alpha_0 = \alpha$ (note that $\mathcal{T}_0 = \{\emptyset, \Omega\}$), which implies $\beta_t = 1 - c$ and

$$\begin{aligned} Q_0[\mathbf{X}_k] &= \left\{(1-c)E[\Lambda_k] + cE\left[\Lambda_k | \Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k)\right]\right\} U_0^{(k)} \\ &= \left\{ E[\Lambda_k] + c\left(E\left[\Lambda_k | \Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k)\right] - E[\Lambda_k] \right) \right\} U_0^{(k)}. \end{aligned} \tag{2.128}$$

Henceforth, the reserve for Λ_k at time 0 is given by its expected value $E[\Lambda_k]$ plus a loading where $c \in (0, 1)$ plays the role of the cost-of-capital rate and

$$E\left[\Lambda_k | \Lambda_k > \mathrm{VaR}_{1-\alpha}(\Lambda_k)\right] - E[\Lambda_k] \tag{2.129}$$

is the capital-at-risk (unexpected loss) measured by the expected shortfall on the security level $1-\alpha$. This is in line with the actual solvency considerations, see for example SST [SST06], Pelsser [Pe10], Salzmann–Wüthrich [SW10] and Sect. 4.5. □

2.7 Conclusions on Chapter 2

We have developed the theoretical foundations of *market-consistent actuarial valuation* based on (potentially) *distorted expected values*, see (2.98) and (2.112). The distorted probabilities lead to the *price of risk*. The framework as developed above is not yet the full story, since it only gives the price of risk, the so-called probability distorted risk premium of assets and insurance liabilities. However, it does not provide sufficient information about the risk bearing and risk mitigation. That is, we have not described how the risk bearing should be organized in order to protect against insolvencies in adverse scenarios, but we have only calculated its market-consistent price.

An insurance company can take the following measures to protect itself against the financial impacts of adverse scenarios:

1. buying options and reinsurance, if available,
2. hedging options internally,
3. setting up sufficient risk bearing capital (solvency margin).

In practice, one has to be very careful in each application whether the price of risk resulting from a mathematical model is already sufficient to finance adverse scenarios, in particular, model risk is often not considered appropriately.

Remark on the Existing Literature. There is a wide range of literature on the definition of market-consistent values. Usually these definitions are not mathematically rigorous and they often (slightly) differ from each other, e.g. they state that market-consistent values should be realistic values or that they should serve for the exchange of two portfolios, etc. One has to be prudent with these definitions in a modelling context because often they are not sufficiently precise.

Our approach gives a mathematical framework for market-consistent valuation that is fully consistent and that respects the theory of classical financial mathematics. Charges for the risk bearing are integrated via distorted probabilities, however (as mentioned above) this does not solve the question of the organization and mitigation of risk, yet. In the next couple of chapters we mainly focus on risk mitigation techniques which leads us to the description of the valuation portfolio and of asset and liability management (ALM) techniques in a full balance sheet approach.

Chapter 3
The Valuation Portfolio in Life Insurance

In this chapter we define the valuation portfolio (VaPo) for a life insurance liability cash flow. The construction is explained with an explicit example. We proceed in two steps: First, we assume that the cash flows have deterministic insurance technical risk, i.e. we have a deterministic mortality table, and only the prices of the financial instruments are described by stochastic processes. Under this assumption, we select appropriate financial instruments and map the cash flows onto these financial instruments. In the second step, we introduce stochastic mortality tables yielding stochastic insurance technical risks. We then follow the construction in the first step, but we add loadings for the insurance technical risks. This construction gives us a replicating portfolio protected against insurance technical risks in terms of financial instruments.

3.1 Deterministic Life Insurance Models: An Example

To introduce the VaPo we start with a deterministic life insurance example without insurance technical risk involved (see Baumgartner et al. [BBK04]). We assume that we have a deterministic mortality table (second order life table) giving the mortality *without loadings*. Let l_x denote the number of insured lives aged x at the beginning of the accounting year and d_x the number of these insured lives who die within that accounting year:

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1_3

$$\begin{array}{lll} l_x & & \\ \downarrow & \longrightarrow & d_x = l_x - l_{x+1} \geq 0 \\ l_{x+1} & & \\ \downarrow & \longrightarrow & d_{x+1} = l_{x+1} - l_{x+2} \geq 0 \\ l_{x+2} & & \\ \downarrow & \longrightarrow & d_{x+2} = l_{x+2} - l_{x+3} \geq 0 \\ \vdots & & \vdots \end{array}$$

Example 3.1 (Endowment insurance policy) We assume that the initial sum insured is CHF 1, the age at policy inception is $x = 50$ and the contract term is $n = 5$. Moreover, we assume that:

- the annual premium $\Pi_t = \Pi$, $t = 0, \dots, 4$, is due in non-indexed CHF at the beginning of each accounting year $(t, t+1]$;
- the benefits are indexed by a well-defined index $\mathcal{I}$ having price process $(I_t)_{t=0,\dots,5}$ with initial value $I_0 = 1$:
  - death benefit at time t is the maximum of I_t and $(1+i)^t$ for some fixed guaranteed interest rate $i \geq 0$,
  - survival benefit at time $t = n = 5$ is the index value I_5, i.e. no minimal guarantee in the case of survival,

  the benefits are always paid at the end of the accounting years.

This means that the survival benefit is given by an index $\mathcal{I}$ whose price is described by a process $(I_t)_{t=0,\dots,5}$ with initial value $I_0 = 1$. This index can be any financial instrument like a stock, a fund, etc. To replicate the survival benefit we need to purchase one unit of the financial instrument $\mathcal{I}$ at price $I_0 = 1$ and it generates the (random) survival benefit I_5 at time $t = n = 5$. The death benefits are more sophisticated as we will see shortly.

The above endowment contract gives us the following Cash Flow Scheme $\mathbf{X} = (X_0, \dots, X_5)$ for initially l_{50} persons alive:

time t	cash flow	premium	death benefit	survival benefit
0	X_0	$-l_{50}\,\Pi$		
1	X_1	$-l_{51}\,\Pi$	$d_{50}\left(I_1 \vee (1+i)^1\right)$	
2	X_2	$-l_{52}\,\Pi$	$d_{51}\left(I_2 \vee (1+i)^2\right)$	
3	X_3	$-l_{53}\,\Pi$	$d_{52}\left(I_3 \vee (1+i)^3\right)$	
4	X_4	$-l_{54}\,\Pi$	$d_{53}\left(I_4 \vee (1+i)^4\right)$	
5	X_5		$d_{54}\left(I_5 \vee (1+i)^5\right)$	$l_{55}\,I_5$

Cash inflows (premium) have a negative sign, cash outflows have a positive sign in our terminology, and $x \vee y = \max\{x, y\}$. □

Task: Value and replicate these endowment policies at the inception of the contract and at every successive accounting year!

3.2 The Valuation Portfolio for the Deterministic Life Insurance Model

We construct the VaPo for the life insurance contracts considered in Example 3.1 (with deterministic mortality table). Roughly speaking the VaPo is a portfolio of financial instruments whose price process replicates the insurance liabilities. The procedure to replicate the insurance cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$ is the following: In a first step we specify the set of financial instruments that will be used for the replication. Secondly, the appropriate number of units of each financial instrument must be determined, this gives us the VaPo for $\mathbf{X}$. Thirdly, we define the market-consistent value of the cash flow $\mathbf{X}$ to be equal to the value of the VaPo. This convention is consistent with the well-known "law of one price"-principle which says that in an arbitrage-free economy two instruments that generate the same cash flows must have the same price.

$\triangleright$ We emphasize that our construction distinguishes between the *financial instrument* $\mathcal{U}$ (which is considered to be a contract) and its *price process* $(U_t)_{t=0,\ldots,n}$ (at which this contract is traded), see also Sect. 2.6.

Step 1. Define the units, i.e. choose a financial basis.

- The premium Π is due at times $t = 0, \ldots, 4$ in non-indexed CHF. Hence, as units we choose the zero coupon bonds $\mathcal{Z}^{(0)}, \ldots, \mathcal{Z}^{(4)}$. The units are denoted by $\mathcal{Z}^{(t)}$ whereas its cash flows are denoted by $\mathbf{Z}^{(t)}$, see also Definition 2.7. Note that $\mathcal{Z}^{(0)}$ is simply cash at time $t = 0$.
- Survival benefit: As unit we choose the financial instrument $\mathcal{I}$ with price process $(I_t)_{t=0,\ldots,5}$.
- The death benefit $I_t \vee (1+i)^t$ can be modelled by financial instrument $\mathcal{I}$ plus a *put option* on $\mathcal{I}$ with strike $(1+i)^t$ at strike time t. We denote this put option by $\mathrm{Put}^{(t)}(\mathcal{I}, (1+i)^t)$.

Hence, we have the following units (financial instruments)

$$(\mathcal{U}_1, \ldots, \mathcal{U}_{11}) = \left(\mathcal{Z}^{(0)}, \ldots, \mathcal{Z}^{(4)}, \mathcal{I}, \mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right), \ldots, \mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right)\right), \tag{3.1}$$

i.e. the total number of different units equals 11. These units play the role of the basis (financial instruments) in which we measure the insurance liabilities.

Step 2. Determine the number (amount) of each unit $\mathcal{U}_i$ needed.
At the beginning of the contracts we need:

<u>Valuation Scheme A:</u> *Cash Flow Representation* (for l_{50} persons)

time t	premium	death benefit	survival benefit
0	$-l_{50}\ \Pi\ \mathcal{Z}^{(0)}$		
1	$-l_{51}\ \Pi\ \mathcal{Z}^{(1)}$	$d_{50}\ \left(\mathcal{I} + \mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right)\right)$	
2	$-l_{52}\ \Pi\ \mathcal{Z}^{(2)}$	$d_{51}\ \left(\mathcal{I} + \mathrm{Put}^{(2)}\left(\mathcal{I}, (1+i)^2\right)\right)$	
3	$-l_{53}\ \Pi\ \mathcal{Z}^{(3)}$	$d_{52}\ \left(\mathcal{I} + \mathrm{Put}^{(3)}\left(\mathcal{I}, (1+i)^3\right)\right)$	
4	$-l_{54}\ \Pi\ \mathcal{Z}^{(4)}$	$d_{53}\ \left(\mathcal{I} + \mathrm{Put}^{(4)}\left(\mathcal{I}, (1+i)^4\right)\right)$	
5		$d_{54}\ \left(\mathcal{I} + \mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right)\right)$	$l_{55}\ \mathcal{I}$

This immediately leads to the summary of units:

<u>*Valuation Scheme B*</u>: *Instrument Representation* (for l_{50} persons)

unit $\mathcal{U}_i$	number of units
$\mathcal{Z}^{(0)}$	$-l_{50}\ \Pi$
$\mathcal{Z}^{(1)}$	$-l_{51}\ \Pi$
$\mathcal{Z}^{(2)}$	$-l_{52}\ \Pi$
$\mathcal{Z}^{(3)}$	$-l_{53}\ \Pi$
$\mathcal{Z}^{(4)}$	$-l_{54}\ \Pi$
$\mathcal{I}$	$d_{50} + d_{51} + d_{52} + d_{53} + d_{54} + l_{55} = l_{50}$
$\mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right)$	d_{50}
$\mathrm{Put}^{(2)}\left(\mathcal{I}, (1+i)^2\right)$	d_{51}
$\mathrm{Put}^{(3)}\left(\mathcal{I}, (1+i)^3\right)$	d_{52}
$\mathrm{Put}^{(4)}\left(\mathcal{I}, (1+i)^4\right)$	d_{53}
$\mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right)$	d_{54}

Observe that the number of units of $\mathcal{I}$ needed is exactly l_{50} because every person insured receives one financial instrument $\mathcal{I}$, no matter whether he dies during the term of the contract or not.

Our valuation portfolio VaPo($\mathbf{X}$) for cash flow $\mathbf{X}$ is a point in an 11-dimensional vector space (see (3.1) and (3.2) in Sect. 3.3 below) where we have specified a basis of financial instruments $\mathcal{U}_i$ (basis of vector space) and the number of instruments we need to hold (point in vector space) to replicate the insurance liabilities.

Step 3. To obtain the (monetary) value of our cash flow we need to apply an accounting principle to this VaPo($\mathbf{X}$), see Sect. 3.3 below. □

Conclusion. In a first and second step, we decompose the liability cash flow $\mathbf{X} = (X_0, \ldots, X_5)$ into an 11-dimensional vector VaPo($\mathbf{X}$), whose basis consists of the financial instruments $\mathcal{U}_1, \ldots, \mathcal{U}_{11}$. The VaPo($\mathbf{X}$) then provides a replicating portfolio *in terms of financial instruments*. Only in a third step do we calculate the monetary value of the cash flow $\mathbf{X}$ by applying an accounting principle to the units $\mathcal{U}_i$, and thus to VaPo($\mathbf{X}$).
We have found the general valuation procedure described in the next section.

3.3 The General VaPo Construction for Deterministic Insurance Technical Risks

We use the notation introduced in Sect. 2.6.

1. For every insurance liability cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$ having deterministic insurance technical risk we construct the VaPo($\mathbf{X}$) as follows: Choose appropriate units $\mathcal{U}_i$ (basis of a multidimensional vector space) and determine the (deterministic) number $\lambda_i(\mathbf{X}) \in \mathbb{R}$ of each unit $\mathcal{U}_i$:

$$\mathbf{X} \mapsto \text{VaPo}(\mathbf{X}) = \sum_i \lambda_i(\mathbf{X})\, \mathcal{U}_i. \tag{3.2}$$

   From a theoretical point of view the VaPo mapping needs to be a linear and continuous (under an appropriate norm) function that maps the insurance liabilities $\mathbf{X}$ to a valuation portfolio VaPo($\mathbf{X}$) which replicates the insurance liabilities in terms of financial instruments.
2. Apply an accounting principle $\mathcal{A}_t$ to the VaPo to obtain a monetary value at time $t \geq 0$

$$\text{VaPo}(\mathbf{X}) \mapsto \mathcal{A}_t\left(\text{VaPo}(\mathbf{X})\right) = Q_t\left[\mathbf{X}\right]. \tag{3.3}$$

   This mapping should be linear and continuous.
   Moreover, the sequence of accounting principles $(\mathcal{A}_t)_{t=0,\ldots,n}$ must satisfy additional consistency properties in order to have an arbitrage-free pricing system. In fact, we require the existence of a deflator φ such that martingale property (2.59) holds for all deflated price processes. This is further discussed below.

The construction of the VaPo adds enormously to the understanding and communication between actuaries and asset managers and investors, respectively. In a first step the actuary decomposes the insurance liabilities into a portfolio of financial instruments, and in a second step the asset manager evaluates the financial instruments. Indeed, it is the key step to a successful *asset and liability management* (ALM), and it clearly highlights the sources of uncertainties involved in the process. It also allocates the responsibilities for the uncertainties to the different parties involved.

Remark 1 For a cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$ with no insurance technical risk involved we obtain for the value at time 0

$$\mathbf{X} \mapsto Q_0\left[\mathbf{X}\right] = \mathcal{A}_0\left(\text{VaPo}(\mathbf{X})\right) = \sum_i \lambda_i(\mathbf{X})\, \mathcal{A}_0\left(\mathcal{U}_i\right) \in \mathbb{R}, \tag{3.4}$$

which should be a linear, positive, continuous and normalized functional on $L^2_{n+1}(P, \mathbb{G})$, see Assumption 2.3. Linearity and continuity follow from the assumptions made above. Normalization requires for cash $\mathcal{A}_0\left(\mathcal{Z}^{(0)}\right) = 1$. Only positivity is a bit sophisticated. We will come back to this.

Remark 2 By linearity the individual policies can be added to a portfolio, i.e. individual cash flows $\mathbf{X}^{(\ell)} \in L^2_{n+1}(P, \mathbb{G})$ easily merge to $\sum_\ell \mathbf{X}^{(\ell)}$ which has value

$$Q_t\left[\sum_\ell \mathbf{X}^{(\ell)}\right] = \sum_\ell \mathcal{A}_t\left(\mathrm{VaPo}(\mathbf{X}^{(\ell)})\right). \tag{3.5}$$

This means that we can value portfolios of a single contract as well as of the whole insurance company. Note that this aggregation needs to be done carefully when insurance technical risk is also involved.

Examples of accounting principles $\mathcal{A}_t$. An accounting principle $\mathcal{A}_t$ attaches a value to the instruments. There are different ways to choose an appropriate accounting principle. In fact, choosing an appropriate accounting principle very much depends on the problem under consideration. We give two examples.

- Classical actuarial discounting. In many situations, for example in (traditional) communication with regulators, the value of the financial instruments are determined by a mathematical model (such as amortized costs, etc.). If we choose the model where we discount with a fixed constant interest rate we denote the (actuarial) accounting principle by $\mathcal{D}_t$.
- In modern actuarial valuation, the instruments are often valued at an economic value, a market value or a value according to the IASB accounting rules. In general, this means that the value of the asset is essentially the price at which it can be exchanged at a (traded) market. When we adopt such an economic accounting principle we use the symbol $\mathcal{E}_t$.

Both principles $\mathcal{D}_t$ and $\mathcal{E}_t$ need to fulfill the consistency property in order to have an arbitrage-free pricing system. That is, assume we choose the economic accounting principles $\mathcal{E}_t, t = 0, \ldots, n$. Then, for cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$ (with deterministic insurance technical risk) we have the following price at time 0

$$Q_0[\mathbf{X}] = \mathcal{E}_0(\mathrm{VaPo}(\mathbf{X})) = \sum_i \lambda_i(\mathbf{X})\, \mathcal{E}_0(\mathcal{U}_i). \tag{3.6}$$

Using Riesz' representation (Theorem 2.5) we find a state price deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$ such that for all cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$

$$\langle \varphi, \mathbf{X} \rangle \;=\; Q_0[\mathbf{X}] \;=\; \mathcal{E}_0(\mathrm{VaPo}(\mathbf{X})). \tag{3.7}$$

Using price definition (2.40), Lemma 2.9 then implies that

$$(\varphi_t \mathcal{E}_t(\mathrm{VaPo}(\mathbf{X})))_{t=0,\ldots,n} \qquad \text{needs to be a } (P, \mathbb{F})\text{-martingale.} \tag{3.8}$$

Note that, in fact, we could restrict ourselves to the Hilbert space $L^2_{n+1}(P, \mathbb{G})$ under deterministic insurance technical risk. In that case we could find a $\mathbb{G}$-adapted financial deflator $\varphi^{\mathbb{G}} \in L^2_{n+1}(P, \mathbb{G})$ such that we obtain $(P, \mathbb{G})$-martingales in (3.8) for $\varphi^{\mathbb{G}}$ deflated price processes (hint: define the financial deflator as $\varphi^{\mathbb{G}}_t = \mathbb{E}[\varphi_t|\mathcal{G}_t]$).

3.4 Linearity of the VaPo for Deterministic Insurance Technical Risk

In (2.52) we have defined $\mathbf{X}_{(k)} \in L^2_{n+1}(P, \mathbb{G})$ as the remaining cash flow after time $k-1$ (outstanding liabilities). Moreover, define the cash flow

$$\mathbf{X}_k = X_k\, \mathbf{Z}^{(k)} = (0, \ldots, 0, X_k, 0 \ldots, 0) \in L^2_{n+1}(P, \mathbb{G}). \tag{3.9}$$

Linearity of cash flows implies

$$\mathbf{X}_{(k)} = \mathbf{X}_{(k+1)} + \mathbf{X}_k, \tag{3.10}$$

and using the linearity of the VaPo construction (3.2) implies the following lemma.

Lemma 3.1 *For* $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$ *we have*

$$\mathrm{VaPo}\left(\mathbf{X}_{(k)}\right) = \mathrm{VaPo}\left(\mathbf{X}_{(k+1)}\right) + \mathrm{VaPo}\left(\mathbf{X}_k\right). \tag{3.11}$$

Of course, in this lemma we assume that the vector space is spanned by the financial instruments determined by $\mathbf{X}_{(k)}$.

Remark. At time k, the last term in (3.11) provides simply cash value, which we may abbreviate by

$$\mathrm{VaPo}\left(\mathbf{X}_k\right) = X_k \mathcal{Z}^{(k)} \qquad \text{at time } k. \tag{3.12}$$

Note that, in general, this *only* holds at time k and we may take advantage of $\mathcal{A}_k(\mathcal{Z}^{(k)}) = 1$. At earlier time points the appropriate financial instrument may differ, we come back to this in Sect. 3.6, below.

Studying now the values given by the accounting principle $\mathcal{A}_t$, we have by the linearity of $\mathcal{A}_t$ the following lemma.

Lemma 3.2 *For* $\mathbf{X} \in L^2_{n+1}(P, \mathbb{G})$ *and* $t \le k$ *we have*

$$\mathcal{A}_t\left(\mathrm{VaPo}\left(\mathbf{X}_{(k)}\right)\right) = \mathcal{A}_t\left(\mathrm{VaPo}\left(\mathbf{X}_{(k+1)}\right)\right) + \mathcal{A}_t\left(\mathrm{VaPo}\left(\mathbf{X}_k\right)\right). \tag{3.13}$$

In particular, if the VaPo of $\mathbf{X}_k$ is evaluated at time k then

$$\mathcal{A}_k\left(\mathrm{VaPo}(\mathbf{X}_k)\right) = Q_k\left[\mathbf{X}_k\right] = \frac{1}{\varphi_k}\, E\left[\varphi_k\, X_k | \mathcal{F}_k\right] = X_k, \tag{3.14}$$

hence

$$\mathcal{A}_k\left(\mathrm{VaPo}\left(\mathbf{X}_{(k)}\right)\right) = \mathcal{A}_k\left(\mathrm{VaPo}\left(\mathbf{X}_{(k+1)}\right)\right) + X_k, \tag{3.15}$$

which again tells us that the VaPo for $\mathbf{X}_k$ at time k is simply cash. This observation is fundamental and should hold independently of the value assigned to the VaPo by the accounting principle $\mathcal{A}_t$. Lemma 3.2 can also be interpreted as a static counterpart of Corollary 2.8.

3.5 The VaPo Protected Against Insurance Technical Risks

So far we have considered an ideal situation which is an important point of reference to measure deviations.

ideal		realistic		deviation
deterministic mortality	$\leftrightarrow$	stochastic mortality	$\Rightarrow$	technical risk
VaPo	$\leftrightarrow$	real asset portfolio $\mathcal{S}$	$\Rightarrow$	financial risk

The ideal situation is often called the base scenario and one then studies deviations from this base scenario. In this section we want to consider deviations induced by insurance technical risks. They come from the fact that the insurance liabilities are not deterministic, i.e. in general, $\mathbf{X}$ is not $\mathbb{G}$-adapted. This means in the present example that we have a stochastic mortality table in our example.

For the potential deviations from the deterministic model (which in our context are expectations and best-estimates for the insurance liabilities) we add a protection. Such a protection can be obtained e.g. via reinsurance products, risk loadings or risk bearing capital. The VaPo with this additional protection will be called *VaPo protected against insurance technical risks*.

3.5.1 Construction on Example 3.1 with Stochastic Mortality

Let us return to our Example 3.1. The stochastic mortality table reads as:

$$\begin{array}{lcl}
l_x & & \\
\downarrow & \longrightarrow & D_x = l_x - L_{x+1} \geq 0 \\
L_{x+1} & & \\
\downarrow & \longrightarrow & D_{x+1} = L_{x+1} - L_{x+2} \geq 0 \\
L_{x+2} & & \\
\downarrow & \longrightarrow & D_{x+2} = L_{x+2} - L_{x+3} \geq 0 \\
\vdots & & \vdots
\end{array}$$

where L_{x+k} and D_{x+k-1} are $\mathcal{T}_k$-measurable random variables for $k \geq 1$. In Example 3.1 we have for $x = 50$

$$D_{50} = l_{50} - L_{51}, \tag{3.16}$$

$$d_{50} = l_{50} - l_{51}, \tag{3.17}$$

which implies

$$D_{50} - d_{50} = l_{51} - L_{51}. \tag{3.18}$$

This describes the deviations of D_{50} and L_{51} from their expected values d_{50} and l_{51}, respectively. In fact, in a first step we use the expected value d_{50} as a predictor for the random variable D_{50}, and in a second step we need to study the prediction uncertainty or the (potential) deviation of the random variable D_{50} around its predictor d_{50}.

The Valuation Scheme A (cash flow representation) then reads as follows for the stochastic mortality table:

time t	premium	death benefit	survival benefit
0	$-l_{50}\ \Pi\ \mathcal{Z}^{(0)}$		
1	$-L_{51}\ \Pi\ \mathcal{Z}^{(1)}$	$D_{50}\left(\mathcal{I} + \mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right)\right)$	
2	$-L_{52}\ \Pi\ \mathcal{Z}^{(2)}$	$D_{51}\left(\mathcal{I} + \mathrm{Put}^{(2)}\left(\mathcal{I}, (1+i)^2\right)\right)$	
3	$-L_{53}\ \Pi\ \mathcal{Z}^{(3)}$	$D_{52}\left(\mathcal{I} + \mathrm{Put}^{(3)}\left(\mathcal{I}, (1+i)^3\right)\right)$	
4	$-L_{54}\ \Pi\ \mathcal{Z}^{(4)}$	$D_{53}\left(\mathcal{I} + \mathrm{Put}^{(4)}\left(\mathcal{I}, (1+i)^4\right)\right)$	
5		$D_{54}\left(\mathcal{I} + \mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right)\right)$	$L_{55}\ \mathcal{I}$

The Valuation Scheme B (instrument representation) for stochastic mortality table is given by:

unit $\mathcal{U}_i$	number of units
$\mathcal{Z}^{(0)}$	$-l_{50}\ \Pi$
$\mathcal{Z}^{(1)}$	$-L_{51}\ \Pi$
$\mathcal{Z}^{(2)}$	$-L_{52}\ \Pi$
$\mathcal{Z}^{(3)}$	$-L_{53}\ \Pi$
$\mathcal{Z}^{(4)}$	$-L_{54}\ \Pi$
$\mathcal{I}$	$D_{50} + D_{51} + D_{52} + D_{53} + D_{54} + L_{55} = l_{50}$
$\mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right)$	D_{50}
$\mathrm{Put}^{(2)}\left(\mathcal{I}, (1+i)^2\right)$	D_{51}
$\mathrm{Put}^{(3)}\left(\mathcal{I}, (1+i)^3\right)$	D_{52}
$\mathrm{Put}^{(4)}\left(\mathcal{I}, (1+i)^4\right)$	D_{53}
$\mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right)$	D_{54}

The important observation is that there is *no randomness* in the number of financial instruments $\mathcal{I}$ because every person insured receives exactly one financial instrument $\mathcal{I}$.

Let us define the *expected* survival probabilities and the *expected* death probabilities (second order life table) for $s \geq x$:

$$p_s = \frac{l_{s+1}}{l_s} \stackrel{!}{=} \frac{1}{L_s} E[L_{s+1}|\mathcal{T}_{s-x}] \quad \text{and} \quad q_s = 1 - p_s = \frac{d_s}{l_s} \stackrel{!}{=} \frac{1}{L_s} E[D_s|\mathcal{T}_{s-x}], \tag{3.19}$$

assuming that $L_s > 0$ and L_s is $\mathcal{T}_{s-x}$-measurable.

Denote by VaPo$(\mathbf{X}_{(t+1)})$ the VaPo for the cash flows after time t with deterministic insurance technical risks (deterministic mortality table as defined in Sect. 3.2). That is, VaPo$(\mathbf{X}_{(t+1)})$ denotes the VaPo based on the *expected mortality* cash flows (with L_{x+k} being replaced by its mean l_{x+k}).

If we allow for a stochastic survival in period (0, 1] we have the following deviations from the (expected) VaPo (deterministic insurance technical risks):
For $t = 1$ and $x = 50$ we obtain the following deviations in Example 3.1

$$(D_{50} - d_{50}) \; \left(\mathcal{I} + \mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right)\right), \tag{3.20}$$

$$(l_{51} - L_{51}) \; \Pi \; \mathcal{Z}^{(1)}, \tag{3.21}$$

$$(L_{51} - l_{51}) \; \frac{\mathrm{VaPo}(\mathbf{X}_{(2)})}{l_{51}}, \tag{3.22}$$

if VaPo$(\mathbf{X}_{(2)})$ denotes the VaPo for deterministic (expected) mortality after time $t = 1$ (according to Sect. 3.2). This means that we have deviations in the payments at time $t = 1$, due to the stochastic mortality, and then at time $t = 1$, we start with a new basis of L_{51} insured lives (instead of l_{51}), which gives a new expected VaPo after time $t = 1$ of (use the linearity of the VaPo)

$$L_{51} \; \frac{\mathrm{VaPo}(\mathbf{X}_{(2)})}{l_{51}}. \tag{3.23}$$

Using (3.18) and equations (3.20)–(3.22) we see that we need additional reserves (in terms of a portfolio) of size

$$(D_{50} - d_{50}) \; \left(\mathcal{I} + \mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right) + \Pi \; \mathcal{Z}^{(1)} - \frac{\mathrm{VaPo}(\mathbf{X}_{(2)})}{l_{51}}\right) \tag{3.24}$$

for the deviations from the expected mortality table within (0, 1]. Note that this deviation is stochastic seen from time $t = 0$. Hence the portfolio at risk is

$$\mathcal{I} + \mathrm{Put}^{(1)}\left(\mathcal{I}, (1+i)^1\right) + \Pi \; \mathcal{Z}^{(1)} - \frac{\mathrm{VaPo}(\mathbf{X}_{(2)})}{l_{51}}. \tag{3.25}$$

We can now iterate this procedure:
For $t = 2$ we have the following deviation from the expected VaPo. The expected VaPo starts after $t = 1$ with the new basis of L_{51} insured lives (see (3.24)). Note that conditionally, given L_{51}, we expect $q_{51} L_{51}$ persons to die within the time interval

$(1, 2]$ and we observe D_{51} deaths at time $t = 2$. This gives the following deviations, conditional on L_{51},

$$(D_{51} - q_{51}\ L_{51})\ \left(\mathcal{I} + \mathrm{Put}^{(2)}\left(\mathcal{I}, (1+i)^2\right)\right), \tag{3.26}$$

$$(p_{51}\ L_{51} - L_{52})\ \Pi\ \mathcal{Z}^{(2)}, \tag{3.27}$$

$$(L_{52} - p_{51}\ L_{51})\ \frac{\mathrm{VaPo}(\mathbf{X}_{(3)})}{p_{51}\ L_{51}}\ \frac{L_{51}}{l_{51}}, \tag{3.28}$$

where the last term can be simplified to

$$\frac{\mathrm{VaPo}(\mathbf{X}_{(3)})}{p_{51}\ L_{51}}\ \frac{L_{51}}{l_{51}} = \frac{\mathrm{VaPo}(\mathbf{X}_{(3)})}{l_{52}}. \tag{3.29}$$

Hence we need for the deviation in $(1, 2]$ an additional portfolio, conditional on L_{51} and using $D_{51} - q_{51}L_{51} = p_{51}L_{51} - L_{52}$,

$$(D_{51} - q_{51}\ L_{51})\ \left(\mathcal{I} + \mathrm{Put}^{(2)}\left(\mathcal{I}, (1+i)^2\right) + \Pi\ \mathcal{Z}^{(2)} - \frac{\mathrm{VaPo}(\mathbf{X}_{(3)})}{l_{52}}\right). \tag{3.30}$$

And analogously for $t = 3, 4, 5$ we obtain the deviations, conditional on L_{50+t-1} and using $D_{50+t-1} - q_{50+t-1}L_{50+t-1} = p_{50+t-1}L_{50+t-1} - L_{50+t}$,

$$\begin{aligned}
&(D_{52} - q_{52}\ L_{52})\ \left(\mathcal{I} + \mathrm{Put}^{(3)}\left(\mathcal{I}, (1+i)^3\right) + \Pi\ \mathcal{Z}^{(3)} - \frac{\mathrm{VaPo}(\mathbf{X}_{(4)})}{l_{53}}\right), \\
&(D_{53} - q_{53}\ L_{53})\ \left(\mathcal{I} + \mathrm{Put}^{(4)}\left(\mathcal{I}, (1+i)^4\right) + \Pi\ \mathcal{Z}^{(4)} - \frac{\mathrm{VaPo}(\mathbf{X}_{(5)})}{l_{54}}\right), \\
&(D_{54} - q_{54}\ L_{54})\ \left(\mathcal{I} + \mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right) - \mathcal{I}\right).
\end{aligned} \tag{3.31}$$

Remark. One can see that when adding up the terms inside the brackets of (3.24) and (3.30)–(3.31), the unit $\mathcal{I}$ cancels since $\mathrm{VaPo}(\mathbf{X}_{(t+1)})$ contains exactly l_{x+t} units of $\mathcal{I}$ at time $t = 0$. This is immediately clear because the number of financial instruments $\mathcal{I}$ we need to purchase at the beginning of the policy does not depend on the mortality table (see Valuation Scheme B on p. 53), i.e. no matter whether a person dies or stays alive it receives $\mathcal{I}$.

Hence, we find the following *portfolios at risk* (where $\mathcal{I}$ cancels and we could re-define cash flow $\mathbf{X}$ accordingly):

$$\begin{aligned}
t = 1, 2, 3, 4:\ \ \mathrm{P@R}_t &= \mathcal{I} + \mathrm{Put}^{(t)}\left(\mathcal{I}, (1+i)^t\right) + \Pi\ \mathcal{Z}^{(t)} - \frac{\mathrm{VaPo}(\mathbf{X}_{(t+1)})}{l_{50+t}}, \\
t = 5:\ \ \mathrm{P@R}_5 &= \mathrm{Put}^{(5)}\left(\mathcal{I}, (1+i)^5\right).
\end{aligned} \tag{3.32}$$

The interpretation of (3.32) is the following. Consider for example the period $(2, 3]$, if more people die than expected ($D_{52} > q_{52}\ L_{52}$) we have to pay an additional death benefit (in terms of a portfolio) of

$$(D_{52} - q_{52}\ L_{52})\ \left(\mathcal{I} + \text{Put}^{(3)}\left(\mathcal{I}, (1+i)^3\right)\right). \tag{3.33}$$

On the other hand for all these people the contracts are terminated which means that our liabilities are reduced by

$$(D_{52} - q_{52}\ L_{52})\ \left(-\Pi\ \mathcal{Z}^{(3)} + \frac{\text{VaPo}(\mathbf{X}_{(4)})}{l_{53}}\right). \tag{3.34}$$

These insurance technical risks are now protected against adverse mortality developments by adding a security loading. This gives us the following *reinsurance premium loadings as a portfolio* seen from time 0: for $t = 0, \ldots, 4$

$$\text{RPP}_t = l_{50+t}\ \left(q^*_{50+t} - q_{50+t}\right) \text{P@R}_{t+1}, \tag{3.35}$$

where $q^*_{x+t} - q_{x+t}$, $x = 50$, denote the loadings charged by the reinsurer against insurance technical risks, and l_{x+t} is the number of units we need to purchase. Here, q^*_{x+t} can be interpreted as the yearly renewable term rates charged by the reinsurer, and later on we will see that it may also correspond to a first order life table.
The **VaPo protected against insurance technical risks** is defined as

$$\text{VaPo}^{prot}\,(\mathbf{X}) = \text{VaPo}\,(\mathbf{X}) + \sum_{t=0}^{4} \text{RPP}_t, \tag{3.36}$$

where VaPo $(\mathbf{X})$ is obtained from the deterministic (expected) mortality table.

Remarks.

- For a monetary reinsurance premium we need to apply an accounting principle $\mathcal{A}_s$, $s \le t$, to the reinsurance premium portfolio RPP_t (yearly renewable term): set $x = 50$

$$\begin{aligned} \Pi_s^{R(t)} &= \mathcal{A}_s\,(\text{RPP}_t) \\ &= l_{x+t}\left(q^*_{x+t} - q_{x+t}\right)\ \mathcal{A}_s\,(\text{P@R}_{t+1})\,. \end{aligned} \tag{3.37}$$

- The last term in (3.37) highlights that the choice of the loadings $q^*_{x+t} - q_{x+t}$ needs some care. The monetary values of the portfolios at risk (3.32) may have both signs. Therefore the signs of the loadings may depend on the monetary values of the portfolios at risk. For example, for death benefits we decrease the survival probabilities p_{x+t}, whereas for annuities we increase these survival probabilities.
- There are different ways to determine the premium: We could choose an actuarial accounting principle $\mathcal{D}_s$ or an economic accounting principle $\mathcal{E}_s$ (which gives an economic yearly renewable term, see also p. 50). This idea opens interesting new reinsurance products: offer a reinsurance cover against insurance technical risks in terms of a VaPo.

- A static hedging strategy is to invest the reinsurance premium into the VaPos of the reinsurer.
- At the moment, the choice of the first order life table q^*_{x+t} looks artificial. In the next section we are going to explain how this fits into our probability distortion modelling framework.
- The construction (3.36) is seen from time 0 because L_{x+t} in (3.30)–(3.31) are replaced by $l_{x+t} = E[L_{x+t}|\mathcal{T}_0]$ in (3.35). Below, we will see that this is not fully consistent and we may ask whether l_{x+t} is sufficiently prudent.

3.5.2 *Probability Distortion of Life Tables*

The choice of the death probabilities q^*_{x+t} in (3.35) may look artificial at first sight. They often come from a *first order life table*. A first order life table refers to survival or death probabilities that are chosen prudently (i.e. with a safety margin), whereas the *second order life table* refers to best-estimate survival and death probabilities (expected values). The choice of a first order life table fits perfectly into our modelling framework. Indeed, the first order life tables can be explained by probability distortions: in (2.104) we have considered the terms $\Lambda_{t,k} = \frac{1}{\varphi^{\mathbb{T}}_t} E[\varphi^{\mathbb{T}}_k \Lambda_k|\mathcal{T}_t]$, $k > t$, referring to the price of the insurance cover in units. We will show that the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$ may serve to construct appropriate first order life tables.

To explain this we revisit Example 3.1 with a stochastic mortality table. For illustrative purposes we choose $t = 2$ and we set $x = 0$ (the latter simplifies the notation). We assume that the number L_2 of persons alive at time $t = 2$ is $\mathcal{T}_2$-measurable and that $\Lambda_3 = D_2$ models the (insurance technical) death benefit at time $t = 3$. Then, we need to study

$$\Lambda_{2,3} = \frac{1}{\varphi^{\mathbb{T}}_2} E\left[\varphi^{\mathbb{T}}_3 \Lambda_3 \middle| \mathcal{T}_2\right] = \frac{1}{\varphi^{\mathbb{T}}_2} E\left[\varphi^{\mathbb{T}}_3 D_2 \middle| \mathcal{T}_2\right], \tag{3.38}$$

which describes for how many financial units we build insurance technical reserves.

In a first step we choose $\varphi^{\mathbb{T}}_3 = \varphi^{\mathbb{T}}_2$, P-a.s. (and $\mathcal{T}_2$-measurable). Then we obtain

$$\Lambda_{2,3} = \frac{1}{\varphi^{\mathbb{T}}_2} E\left[\varphi^{\mathbb{T}}_3 \Lambda_3 \middle| \mathcal{T}_2\right] = E\left[D_2 \middle| \mathcal{T}_2\right] = q_2\, L_2, \tag{3.39}$$

where q_2 describes the expected death probability within accounting year (2, 3] for a person aged $x = 2$ at the beginning of this accounting year. Formula (3.39) provides the expected mortalities; similarly to Example 2.7, the assumption $\varphi^{\mathbb{T}}_3 = \varphi^{\mathbb{T}}_2$ means that we do not distort the death probabilities in period (2, 3].

We now model a non-trivial probability distortion $\varphi^{\mathbb{T}}_3$ for period (2, 3] so that we obtain a first order life table q^*_2. Note that

$$\Lambda_{2,3} = \frac{1}{\varphi_2^{\mathbb{T}}} E\left[\varphi_3^{\mathbb{T}} \Lambda_3 \middle| \mathcal{T}_2\right] = E\left[\frac{\varphi_3^{\mathbb{T}}}{\varphi_2^{\mathbb{T}}} D_2 \middle| \mathcal{T}_2\right] = \sum_{i=1}^{L_2} E\left[\frac{\varphi_3^{\mathbb{T}}}{\varphi_2^{\mathbb{T}}} I_i \middle| \mathcal{T}_2\right], \tag{3.40}$$

where I_i is the indicator whether person i, $i = 1, \ldots, L_2$, dies within $(2, 3]$.

We assume that single life times (of persons all of the same age) are i.i.d., conditionally given $\mathcal{T}_2$. Then we choose the probability distortion $\varphi_3^{\mathbb{T}}$ of the following form

$$\varphi_3^{\mathbb{T}} = \varphi_2^{\mathbb{T}} \prod_{j=1}^{L_2} \widetilde{\varphi}_3^{\mathbb{T}}(I_j), \tag{3.41}$$

such that each factor $\widetilde{\varphi}_3^{\mathbb{T}}(I_j)$ of the product has conditional expectation 1. Henceforth, we can write in this probability distorted case

$$\Lambda_{2,3} = \sum_{i=1}^{L_2} E\left[\frac{\varphi_3^{\mathbb{T}}}{\varphi_2^{\mathbb{T}}} I_i \middle| \mathcal{T}_2\right] = \sum_{i=1}^{L_2} E\left[\widetilde{\varphi}_3^{\mathbb{T}}(I_i)\, I_i \middle| \mathcal{T}_2\right], \tag{3.42}$$

where we have used mutual conditional independence of $\widetilde{\varphi}_3^{\mathbb{T}}(I_j)$, $j = 1, \ldots, L_2$, and (conditional) normalization to 1. The factors of the probability distortions are now chosen as follows: Take $q_2^* \in (0, 1)$ and define

$$\widetilde{\varphi}_3^{\mathbb{T}}(1) = \frac{q_2^*}{q_2}, \tag{3.43}$$

$$\widetilde{\varphi}_3^{\mathbb{T}}(0) = \frac{1 - q_2^*}{1 - q_2}. \tag{3.44}$$

We then obtain the required normalization to 1

$$E\left[\widetilde{\varphi}_3^{\mathbb{T}}(I_i) \middle| \mathcal{T}_2\right] = q_2 \,\frac{q_2^*}{q_2} + p_2 \,\frac{1 - q_2^*}{1 - q_2} = 1, \tag{3.45}$$

and the first order life table

$$E\left[\widetilde{\varphi}_3^{\mathbb{T}}(I_i) I_i \middle| \mathcal{T}_2\right] = q_2 \,\frac{q_2^*}{q_2} = q_2^*. \tag{3.46}$$

This implies in the probability distorted case

$$\Lambda_{2,3} = \frac{1}{\varphi_2^{\mathbb{T}}} E\left[\varphi_3^{\mathbb{T}} \Lambda_3 \middle| \mathcal{T}_2\right] = q_2^* \, L_2. \tag{3.47}$$

In other words, the transition from the second order life table $(p_t)_t$ to the first order life table $(p_t^*)_t$ may stem precisely from a probability distortion induced by $(\varphi_{t+1}^{\mathbb{T}})_t$.

Exercise 3.2 (*Life-time annuity*) Consider a life-time annuity for a man aged x at time 0. We assume that the life-time annuity contract is paid by a single premium installment π_0 at the beginning of the insurance period (initial lump sum, upfront premium) and that the insured receives an annual payment of size M until he dies.

- Determine the VaPo based on the second order life table $(p_t)_t$.
- Calculate the portfolios at risk and the VaPo protected against insurance technical risks.
- Determine the sign of the loadings $p_t^* - p_t$ for receiving a positive loading.
- Express the second order life table $(p_t^*)_t$ with the help of probability distortions $\varphi^{\mathbb{T}}$.

□

3.6 Back to the Basic Model: Formal Construction

As in Chap. 2 we choose a fixed deflator $\varphi \in L^2_{n+1}(P, \mathbb{F})$ to value the cash flows $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$. Under Assumption 2.16, we assume that we can decompose our filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ so that the filtration $\mathbb{F}$ decouples into two independent filtrations $\mathbb{T}$ and $\mathbb{G}$ w.r.t. P, and

$$X_k = \Lambda_k \, U_k^{(k)} \qquad \text{for all } k = 0, \ldots, n, \tag{3.48}$$

$$\varphi = \varphi^{\mathbb{T}} \, \varphi^{\mathbb{G}} \qquad \text{(component-wise product)}, \tag{3.49}$$

such that $\boldsymbol{\Lambda} = (\Lambda_t)_{t=0,\ldots,n}, \varphi^{\mathbb{T}} \in L^2_{n+1}(P, \mathbb{T})$ as well as $(U_t^{(k)})_{t=0,\ldots,n}, \varphi^{\mathbb{G}} \in L^2_{n+1}(P, \mathbb{G})$ for all $k = 0, \ldots, n$. This implies that we can split the valuation problem into two independent problems, one measuring insurance technical risks w.r.t. $\mathbb{T}$ and one describing the financial price processes w.r.t. $\mathbb{G}$. Moreover, we assume that $\varphi^{\mathbb{T}}$ is a density process w.r.t. $\mathbb{T}$ and P, see (2.102).

In Example 3.1 we have seen that it is not always possible to represent a cash flow X_k by one single financial instrument $\mathcal{U}_k$. For instance, in period $k = 5$ we use the index $\mathcal{I}$ as well as the put option $\text{Put}^{(5)}(\mathcal{I}, (1+i)^5)$ for representing X_5. This may require that cash flow $\mathbf{X}$ is split into cash flows $\mathbf{X}^{(\ell)}$ such that $\mathbf{X} = \sum_\ell \mathbf{X}^{(\ell)}$, and then the VaPo is constructed for these cash flows $\mathbf{X}^{(\ell)}$, we also refer to Remark 2 on p. 50.

To not over-complicate this construction we assume that for each cash flow $\mathbf{X}$ of interest we can find financial instruments $\mathcal{U}_1, \ldots, \mathcal{U}_p$ such that we obtain VaPo

$$\mathbf{X} \mapsto \sum_{i=1}^{p} \Lambda_i(\mathbf{X}) \, \mathcal{U}_i, \tag{3.50}$$

with $\Lambda_i(\mathbf{X})$ describing the (random) number of units $\mathcal{U}_i$ needed to replicate $\mathbf{X}$. There are two important remarks: (1) lower index $i = 1, \ldots, p$ is not linked to a time index in (3.50); and (2) replication of the cash flow $\mathbf{X}$ is only achieved by selling the appropriate number of instruments at the right time points. Under the assumption of linearity in (3.50), the second item (2) is better described by

$$\mathbf{X} = \begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_n \end{pmatrix} \mapsto \sum_{i=1}^{p} \begin{pmatrix} \Lambda_i(\mathbf{X}_0) \\ \Lambda_i(\mathbf{X}_1) \\ \vdots \\ \Lambda_i(\mathbf{X}_n) \end{pmatrix}' \begin{pmatrix} \mathcal{U}_i \\ \mathcal{U}_i \\ \vdots \\ \mathcal{U}_i \end{pmatrix}, \tag{3.51}$$

where $\mathbf{X}_k = (0, \ldots, 0, X_k, 0, \ldots, 0)$ denotes the cash flow at time k. Thus, $\Lambda_i(\mathbf{X}_k)$ describes how many instruments $\mathcal{U}_i$ we need to sell at time k to replicate X_k.

For the *VaPo construction seen from time 0* we then replace the random numbers $\Lambda_i(\mathbf{X}_k)$ of instruments $\mathcal{U}_i$ by their expected values at time 0:

$$\Lambda_i(\mathbf{X}_k) \;\mapsto\; l_{i,k}^{(0)} = E\left[\Lambda_i(\mathbf{X}_k)\middle|\mathcal{T}_0\right]. \tag{3.52}$$

If $\Lambda_i(\mathbf{X}_k)$ is deterministic as in Sect. 3.1, we have $\Lambda_i(\mathbf{X}_k) = \lambda_i(\mathbf{X}_k) = l_{i,k}^{(0)}$ (see (3.2)). Mapping (3.52) corresponds to a probability distortion choice of $\boldsymbol{\varphi}^{\mathbb{T}} = 1$ (best-estimates, see Example 2.7).

For the *VaPo protected against insurance technical risks seen from time 0* we replace $\Lambda_i(\mathbf{X}_k)$ by the following distorted expected values at time 0:

$$\Lambda_i(\mathbf{X}_k) \;\mapsto\; l_{i,k}^{*,0} = \frac{1}{\varphi_0^{\mathbb{T}}} E\left[\varphi_k^{\mathbb{T}}\, \Lambda_i(\mathbf{X}_k)\middle|\mathcal{T}_0\right], \tag{3.53}$$

which adds a loading to $l_{i,k}^{(0)}$ for insurance technical risks. This loading depends on the choice of the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$. If $\Lambda_i(\mathbf{X}_k)$ is deterministic as in Sect. 3.1, i.e. $\mathbf{X} \in L_{n+1}^2(P, \mathbb{G})$, we have $\Lambda_i(\mathbf{X}_k) = \lambda_i(\mathbf{X}_k) = l_{i,k}^{(0)} = l_{i,k}^{*,0}$ due to the martingale property (2.100), i.e. we do not need a loading for insurance technical "risks" in this case.

This now gives for the *VaPo at time 0*

$$\mathrm{VaPo}_0(\mathbf{X}) = \sum_{i=1}^{p} \begin{pmatrix} l_{i,0}^{(0)} \\ \vdots \\ l_{i,n}^{(0)} \end{pmatrix}' \begin{pmatrix} \mathcal{U}_i \\ \vdots \\ \mathcal{U}_i \end{pmatrix}. \tag{3.54}$$

This can also be written as

$$\mathrm{VaPo}_0(\mathbf{X}) = \sum_{i=1}^{p} l_i^{(0)}\, \mathcal{U}_i, \tag{3.55}$$

with the expected numbers of units

$$l_i^{(0)} = \sum_{k=0}^{n} l_{i,k}^{(0)}. \tag{3.56}$$

The *VaPo protected against insurance technical risks at time 0* is given by

$$\mathrm{VaPo}_0^{prot}(\mathbf{X}) = \sum_{i=1}^{p} \begin{pmatrix} l_{i,0}^{*,0} \\ \vdots \\ l_{i,n}^{*,0} \end{pmatrix}' \begin{pmatrix} \mathcal{U}_i \\ \vdots \\ \mathcal{U}_i \end{pmatrix}, \tag{3.57}$$

or equivalently

$$\mathrm{VaPo}_0^{prot}(\mathbf{X}) = \sum_{i=1}^{p} l_i^{*,0}\, \mathcal{U}_i, \tag{3.58}$$

with distorted expected numbers of units

$$l_i^{*,0} = \sum_{k=0}^{n} l_{i,k}^{*,0}. \tag{3.59}$$

Remark. Observe that (3.54) and (3.55) provide two representations for $\mathrm{VaPo}_0(\mathbf{X})$. Firstly, we have the *cash flow representation*, which corresponds to the Valuation Scheme A in Sect. 3.2. That is, (3.54) implies

$$\mathrm{VaPo}_0(\mathbf{X}) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}' \begin{pmatrix} \sum_{i=1}^{p} l_{i,0}^{(0)}\, \mathcal{U}_i \\ \vdots \\ \sum_{i=1}^{p} l_{i,n}^{(0)}\, \mathcal{U}_i \end{pmatrix}. \tag{3.60}$$

Secondly, we have the *instrument representation* (3.55) which corresponds to the Valuation Scheme B in Sect. 3.2. Analogously, we have the two representations (3.57) and (3.58) for the VaPo protected against insurance technical risks $\mathrm{VaPo}_0^{prot}(\mathbf{X})$.

For many purposes the instrument representation (3.55) of the VaPo suffices, in particular, in situations with static hedging. Sometimes, however, it may be necessary to work with the cash flow representation (3.60), in particular, if we want to study the cash flows generated, see for example Sect. 3.8 below.

Applying an accounting principle $\mathcal{A}_0$ to the VaPo (or equivalently to the financial instruments $\mathcal{U}_i$) gives a monetary value for the cash flow $\mathbf{X}$ at time 0.

Remark. It is important to realize that the VaPo construction in (3.52) is seen from time 0. If the cash flows have no insurance technical risks (as in Sect. 3.3) there are no deviations in $\Lambda_i(\mathbf{X})$ over time, which means that $l_{i,k}^{(0)}$ remains constant over time. But if insurance technical risks are involved, the following

$$l_{i,k}^{(t)} = E\left[\Lambda_i(\mathbf{X}_k) \middle| \mathcal{T}_t\right], \tag{3.61}$$

$$l_{i,k}^{*,t} = \frac{1}{\varphi_t^{\mathbb{T}}} E\left[\varphi_k^{\mathbb{T}} \Lambda_i(\mathbf{X}_k) \middle| \mathcal{T}_t\right] \tag{3.62}$$

are functions of time $t = 0, \ldots, n$. This then leads to time-dependent VaPos

$$\mathrm{VaPo}_t(\mathbf{X}) = \sum_{i=1}^{p} \sum_{k=0}^{n} l_{i,k}^{(t)} \mathcal{U}_i, \tag{3.63}$$

$$\mathrm{VaPo}_t^{prot}(\mathbf{X}) = \sum_{i=1}^{p} \sum_{k=0}^{n} l_{i,k}^{*,t} \mathcal{U}_i. \tag{3.64}$$

We then also need to study the changes in these VaPos over time, i.e. the developments

$$\mathrm{VaPo}_t(\mathbf{X}) - \mathrm{VaPo}_{t-1}(\mathbf{X}), \tag{3.65}$$

and

$$\mathrm{VaPo}_t^{prot}(\mathbf{X}) - \mathrm{VaPo}_{t-1}^{prot}(\mathbf{X}), \tag{3.66}$$

respectively, which consider the update of information $\mathcal{T}_{t-1} \mapsto \mathcal{T}_t$ similar to the *claims development result* (CDR) in non-life insurance, see Chap. 5, below, and Merz–Wüthrich [MW08, MW14] and Salzmann–Wüthrich [SW10]. Moreover, (3.66) also describes the expected release of the margin generated by the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$, see Sect. 8.2 in Wüthrich–Merz [WM13].

3.7 A Numerical Unit-Linked Insurance Example

In this section we give two numerical examples of Example 3.1 (endowment insurance policy). Note that the mathematical details for the evaluation of the accounting principles are given in Chap. 4, below.

For the deterministic mortality table we choose Table 3.1.

Example 3.3 (Equity-linked endowment insurance policy) We revisit Example 3.1 of the equity-linked life insurance contracts (also called unit-linked insurance or variable life insurance). Assume that $(I_t)_{t=0,\ldots,n}$ denotes the price process of the equity index $\mathcal{I}$ in the economic world. That is, we choose an economic accounting principle that describes the price processes at the financial market. Assume that this accounting principle can be described by the financial deflator $\boldsymbol{\varphi}^{\mathbb{G}}$. The no-arbitrage property of Theorem 2.19 implies that we need to have the corresponding martingale property, which gives for all $t = 0, \ldots, n-1$

$$\varphi_t^{\mathbb{G}} I_t = E\left[\varphi_{t+1}^{\mathbb{G}} I_{t+1} \middle| \mathcal{G}_t\right]. \tag{3.67}$$

Table 3.1 Deterministic mortality table, portfolio of 1,000 insured lives

Time t	Survivals	Deaths
0	$l_{50} = 1{,}000$	
1	$l_{51} = 996$	$d_{50} = 4$
2	$l_{52} = 991$	$d_{51} = 5$
3	$l_{53} = 986$	$d_{52} = 5$
4	$l_{54} = 981$	$d_{53} = 5$
5	$l_{55} = 975$	$d_{54} = 6$

The price of the zero coupon bond $\mathcal{Z}^{(m)}$ with maturity date m is at time $t \le m$ obtained by

$$P(t,m) = Q_t\left[\mathbf{Z}^{(m)}\right] = \frac{1}{\varphi_t^{\mathbb{G}}} E\left[\varphi_m^{\mathbb{G}} \middle| \mathcal{G}_t\right], \tag{3.68}$$

where $\mathbf{Z}^{(m)}$ is the cash flow of the zero coupon bond $\mathcal{Z}^{(m)}$. The zero coupon bond yield rates $Y(t,m)$ (continuously-compounded spot rates) at time $t < m$ are given by

$$Y(t,m) = -\frac{1}{m-t} \log P(t,m) \quad \Longleftrightarrow \quad P(t,m) = \exp\left\{-(m-t)\, Y(t,m)\right\}. \tag{3.69}$$

Considering historical data we observe (the source of the zero coupon bond yield curves is the Schweizerische Nationalbank [SNB] webpage): see Table 3.2.

We assume that our endowment insurance policy starts in year 2000, i.e., we identify the starting point at age $x = 50$ with the year $t = 0$.

Assume that the guaranteed interest rate is $i = 2\,\%$.

To adapt the option pricing formula to the case of non-constant interest rates we will consider the price process $(I_t)_{t=0,\dots,n}$ under changes of numeraires. Assume we would like to evaluate the put option $\mathrm{Put}^{(m)}\left(\mathcal{I}, (1+i)^m\right)$ at time $t \le m$, this put option is reflected by the cash flow

$$\mathbf{X} = \Big(0, \dots, 0, \left((1+i)^m - I_m\right)_+, 0, \dots, 0\Big), \tag{3.70}$$

where the potentially non-zero cash flow is at the maturity date m. The price of this put option at time $t \le m$ is given by

$$\mathcal{A}_t^{(m)} = \mathcal{A}_t\left(\mathrm{Put}^{(m)}\left(\mathcal{I}, (1+i)^m\right)\right) = \frac{1}{\varphi_t^{\mathbb{G}}} E\left[\varphi_m^{\mathbb{G}}\left((1+i)^m - I_m\right)_+ \middle| \mathcal{G}_t\right]. \tag{3.71}$$

We define a density process $\boldsymbol{\zeta} = (\zeta_0, \dots, \zeta_m)$ by

$$\zeta_t \;=\; \varphi_t^{\mathbb{G}} P(t,m), \tag{3.72}$$

for $t \le m$. Note that $\zeta \gg 0$ is a normalized $(P, \mathbb{G})$-martingale, this is in complete analogy to Corollary 2.10. This allows us to define the m-forward measure $P^{(m)} \sim P$ via the Radon–Nikodým derivative

$$\left.\frac{dP^{(m)}}{dP}\right|_{\mathcal{G}_m} = \zeta_m = \varphi_m^{\mathbb{G}} P(m,m) = \varphi_m^{\mathbb{G}}. \tag{3.73}$$

Under this m-forward measure $P^{(m)}$ the price of the put option at time $t \le m$ is given by, see also Lemma 2.11,

$$\begin{aligned} \mathcal{A}_t^{(m)} &= \frac{P(t,m)}{\varphi_t^{\mathbb{G}} P(t,m)} E\left[\left. \varphi_m^{\mathbb{G}} \left((1+i)^m - I_m\right)_+ \right| \mathcal{G}_t\right] \\ &= \frac{P(t,m)}{\zeta_t} E\left[\left. \zeta_m \left((1+i)^m - I_m\right)_+ \right| \mathcal{G}_t\right] \\ &= P(t,m)\; E^{(m)}\left[\left. \left((1+i)^m - I_m\right)_+ \right| \mathcal{G}_t\right], \end{aligned} \tag{3.74}$$

where $E^{(m)}$ denotes expectation under the m-forward measure $P^{(m)}$. Define for $t \le m$ the equity price process under the numeraire $\mathcal{Z}^{(m)}$ given by

$$\widetilde{I}_t = P(t,m)^{-1} I_t. \tag{3.75}$$

Then, our pricing framework implies that $(\widetilde{I}_t)_{t=0,\dots,m}$ is a $(P^{(m)}, \mathbb{G})$-martingale, see also Corollary 2.13.

To calculate the price process of the put option we need to have an explicit distributional model for the price process $(\widetilde{I}_t)_{t=0,\dots,m}$: we assume that it is a log-normal process under the m-forward measure $P^{(m)}$ with independent log-increments having constant volatility parameter $\sigma > 0$. More formally, we assume for $t = 1, \dots, m$

$$\widetilde{I}_t = \widetilde{I}_{t-1} \exp\left\{-\sigma^2/2 + \sigma\varepsilon_t\right\}, \tag{3.76}$$

with $(\varepsilon_t)_{t=0,\dots,m}$ being $\mathbb{G}$-adapted and ε_t being independent of $\mathcal{G}_{t-1}$, having a standard Gaussian distribution under $P^{(m)}$. Note that the drift term $-\sigma^2/2$ is naturally given because of the $(P^{(m)}, \mathbb{G})$-martingale property. This implies that, conditionally given $\mathcal{G}_t$,

$$I_m = P(m,m)^{-1} I_m = \widetilde{I}_m = \widetilde{I}_t \exp\left\{-(m-t)\sigma^2/2 + \sigma \sum_{s=t+1}^{m} \varepsilon_s\right\} \tag{3.77}$$

has a log-normal distribution under $P^{(m)}$ with mean parameter

$$\log \widetilde{I}_t - \sigma^2(m-t)/2 = \log\left(\frac{I_t}{P(t,m)}\right) - \sigma^2(m-t)/2, \tag{3.78}$$

and variance parameter $(m-t)\sigma^2$.

Under these assumptions we can explicitly calculate the prices (3.74). They are for $t \leq m$ given by the following (for more details we refer to Example 4.2 below)

$$\mathcal{A}_t^{(m)} = P(t,m)(1+i)^m \, \Phi\left(d_+(t,m)\right) - I_t \, \Phi\left(d_-(t,m)\right), \tag{3.79}$$

where Φ denotes the standard Gaussian distribution and

$$d_+(t,m) = \frac{\log\left(P(t,m)(1+i)^m/I_t\right) + \sigma^2(m-t)/2}{\sigma\sqrt{m-t}}, \tag{3.80}$$

$$d_-(t,m) = d_+(t,m) - \sigma\sqrt{m-t}. \tag{3.81}$$

Remark. For $P(t,m) = \exp\{-(m-t)r\}$ with $r > 0$ constant, (3.79) is the well-known Black–Scholes formula.

We choose the observations of the price process $(I_t)_t$ of $\mathcal{I}$ and of the price processes $(P(t,m))_t$ of the zero coupon bonds $\mathcal{Z}^{(m)}$ according to Table 3.2, with normalization $I_0 = 1$. The volatility parameter is chosen as $\sigma = 15\,\%$. This leads to the prices for the put options displayed in Table 3.3 (observe that in year $t = 0$ (year 2000) we have rather high yield rates $Y(0,m)$, which gives us low prices for the put options).

In practical applications it is probably doubtful to choose a constant volatility parameter σ because the ratio in (3.75) is likely to have a time to maturity $m-t$ dependent volatility parameter. This is not really a difficulty in the log-normal modelling approach described above, but it was skipped for simplicity.

Next, we calculate the monetary value of the VaPo of cash flow $\mathbf{X}$. Assume that the survival and death benefits (before index linking) are equal to 100,000. Using the premium equivalence principle we require the identity

$$\mathcal{A}_0\left(\mathrm{VaPo}_0(\mathbf{X})\right) = Q_0[\mathbf{X}] \overset{!}{=} 0. \tag{3.82}$$

Table 3.2 Observed returns of the equity index $\mathcal{I}$ and yield rates $Y(t,m)$ of the zero coupon bonds as a function of time to maturity $m-t$ in calendar years 1996–2005

t	$\frac{I_t}{I_{t-1}} - 1$ (%)	$Y(t,m)$ for time to maturity $m-t=$				
		1 (%)	2 (%)	3 (%)	4 (%)	5 (%)
1996	12.99	1.94	2.42	2.79	3.12	3.42
1997	13.35	1.82	1.92	2.20	2.48	2.74
1998	22.11	1.71	1.81	1.95	2.10	2.27
1999	5.41	2.21	2.06	2.21	2.31	2.42
2000	2.02	3.37	3.52	3.53	3.56	3.60
2001	8.60	2.00	2.85	2.90	2.96	3.02
2002	−12.41	0.69	1.84	2.14	2.38	2.57
2003	−14.83	0.58	0.79	1.14	1.46	1.72
2004	15.87	0.99	1.11	1.42	1.70	1.94
2005	1.83	1.41	1.14	1.32	1.48	1.62
Average	5.49	1.67	1.95	2.16	2.35	2.53

Table 3.3 Prices of put options $\mathcal{A}_t^{(m)} = \mathcal{A}_t \left(\mathrm{Put}^{(m)} \left(\mathcal{I}, (1+i)^m \right) \right)$ given in (3.79)

	Time to maturity $m - t =$				
	1	2	3	4	5
$t = 0$	0.053	0.069	0.080	0.088	0.093
$t = 1$	0.034	0.051	0.066	0.076	
$t = 2$	0.117	0.131	0.144		
$t = 3$	0.249	0.267			
$t = 4$	0.140				

This provides the following identity (see also Valuation Scheme B (instrument representation) on p. 48)

$$\sum_{t=0}^{4} l_{50+t} \ \Pi \ P(0,t) = l_{50} \ I_0 + \sum_{t=0}^{4} d_{50+t} \ \mathcal{A}_0^{(t+1)}. \tag{3.83}$$

Solving this identity provides the market-consistent pure risk premium $\Pi = 21{,}667$ (per policy). In fact, the guarantee turns out to be rather cheap in our calibration, namely its price is 42 per policy.

Now we consider the VaPos at different time points $0 \le t \le n-1$. Denote by $\mathbf{X}_{(t+1)} = (0, \ldots, 0, X_{t+1}, \ldots, X_n)$ the cash flow of the outstanding liabilities at time t. The reserves before receiving the premium payments of the people alive at time t are given by

$$\begin{aligned} R_t^{(t+1,-)} &= \mathcal{A}_t \left(\mathrm{VaPo}_t \left(\mathbf{X}_{(t+1)} \right) - l_{50+t} \ \Pi \ \mathcal{Z}^{(t)} \right) \\ &= \mathcal{A}_t \left(\mathrm{VaPo}_t \left(\mathbf{X}_{(t+1)} \right) \right) - l_{50+t} \ \Pi \\ &= Q_t \left[\mathbf{X}_{(t+1)} \right] - l_{50+t} \ \Pi = R_t^{(t+1)} - l_{50+t} \ \Pi. \end{aligned} \tag{3.84}$$

The reserves after premium payments at time t are given by

$$R_t^{(t+1)} = \mathcal{A}_t \left(\mathrm{VaPo}_t \left(\mathbf{X}_{(t+1)} \right) \right) = Q_t \left[\mathbf{X}_{(t+1)} \right]. \tag{3.85}$$

This gives the run-off behaviour (savings process) given in Table 3.4. Note that the death benefits have a minimal interest rate guarantee of $i = 2\,\%$, whereas the survival benefit (by assumption) does not have a minimal guarantee. What we derive from Table 3.4 is that the performance of the index $(I_t)_t$ was below $2\,\%$, in fact, it was negative with $I_4 = 97.26\,\%$. Therefore, the reserves $R_4^{(5)} < 100{,}000 \cdot (1 + 2\,\%)^4$ and survivals would also have benefited if they would have bought a minimal guarantee (at least if the last return $I_5/I_4 - 1$ is not excessive). This can also be seen from Table 3.3 because the put option is in the money for later periods t.

For the VaPo protected against insurance technical risks, we proceed as follows: we define p_t and q_t as in (3.19). Moreover, we choose $q_t^* = 1.5 \cdot q_t$ (first order life table). Hence, we consider the premium for the yearly renewable term $\Pi_0^{R(k)} =$

Table 3.4 Development of the savings process

	$R_t^{(t+1,-)}$	$R_t^{(t+1)}$
$t = 0$	0	21,666,637
$t = 1$	26,370,714	47,950,684
$t = 2$	32,423,186	53,894,823
$t = 3$	39,619,061	60,982,365
$t = 4$	74,244,766	95,499,737

Table 3.5 Monetary yearly renewable term premiums

	$\Pi_0^{R(k)}$
$k = 0$	167,885
$k = 1$	162,340
$k = 2$	115,180
$k = 3$	68,723
$k = 4$	27,818

$\mathcal{A}_0$ (RPP$_k$) defined in (3.37) for our accounting principle $\mathcal{A}_0$. This gives the monetary reinsurance loadings at time $t = 0$ presented in Table 3.5. □

Example 3.4 (Wage index) In non-life insurance the products are rather linked to other indexes like the inflation index, wage index, the consumer price index or a medical expenses index. As index we choose the wage index (source Schweizerische Nationalbank [SNB] webpage): see Table 3.6.

This time we choose a minimal interest rate guarantee of $i = 1.5\,\%$. For the volatility parameter we choose $\sigma = 1\,\%$. Under these parameters we consider the same log-normal model as in the previous example. This implies that the market-consistent pure risk premium Π equals $\Pi = 21{,}625$ (per policy) and the prices for the put options can be found in Table 3.7.

The premium Π and the put option prices are smaller in the wage index example than in the equity-linked example. This comes from the fact that the choice of σ is much smaller in the second example. In fact, the put option prices are very small in this example. This gives the savings process provided in Table 3.8. Here, we see that $R_4^{(5)} \approx 100{,}000 \cdot (1 + 1.5\,\%)^4$ which means that the put options are probably not needed (the numbers are not directly comparable because the reserves also account for people that consume the (early) death benefits). The reinsurance loadings are given in Table 3.9. The reinsurance premium looks rather small compared to the pure risk premium $l_t\ \Pi\ P(0,t)$. This comes from the following facts: σ is rather small; the minimal guarantee $i = 1.5\,\%$ is rather low compared to the yield rates $Y(0,\cdot)$ in year $t = 0$; and the randomness of D_t is rather small compared to the total volume l_t. □

Table 3.6 Observed returns of the wage inflation index $\mathcal{I}$ and yield rates $Y(t,m)$ of the zero coupon bond as a function of time to maturity $m-t$ in calendar years 1996–2005

t	$\frac{I_t}{I_{t-1}} - 1$ (%)	$Y(t,m)$ for time to maturity $m-t=$				
		1 (%)	2 (%)	3 (%)	4 (%)	5 (%)
1996	1.30	1.94	2.42	2.79	3.12	3.42
1997	1.26	1.82	1.92	2.20	2.48	2.74
1998	0.47	1.71	1.81	1.95	2.10	2.27
1999	0.69	2.21	2.06	2.21	2.31	2.42
2000	0.29	3.37	3.52	3.53	3.56	3.60
2001	1.26	2.00	2.85	2.90	2.96	3.02
2002	2.48	0.69	1.84	2.14	2.38	2.57
2003	1.79	0.58	0.79	1.14	1.46	1.72
2004	1.40	0.99	1.11	1.42	1.70	1.94
2005	0.93	1.41	1.14	1.32	1.48	1.62
Average	1.19	1.67	1.95	2.16	2.35	2.53

Table 3.7 Put option prices $\mathcal{A}_t^{(m)} = \mathcal{A}_t\left(\text{Put}^{(m)}\left(\mathcal{I}, (1+i)^m\right)\right)$ given in (3.79)

	Time to maturity $m-t=$				
	1	2	3	4	5
$t=0$	$1.16 \cdot 10^{-4}$	$8.26 \cdot 10^{-6}$	$8.42 \cdot 10^{-7}$	$7.36 \cdot 10^{-8}$	$4.74 \cdot 10^{-9}$
$t=1$	$2.81 \cdot 10^{-3}$	$2.27 \cdot 10^{-4}$	$6.12 \cdot 10^{-5}$	$1.39 \cdot 10^{-5}$	
$t=2$	$4.57 \cdot 10^{-3}$	$1.20 \cdot 10^{-3}$	$4.72 \cdot 10^{-4}$		
$t=3$	$3.70 \cdot 10^{-3}$	$8.25 \cdot 10^{-3}$			
$t=4$	$2.42 \cdot 10^{-3}$				

Table 3.8 Development of the savings process

	$R_t^{(t+1,-)}$	$R_t^{(t+1)}$
$t=0$	0	21,624,505
$t=1$	18,725,275	40,263,282
$t=2$	39,786,604	61,216,489
$t=3$	61,744,032	83,065,794
$t=4$	83,860,324	105,073,964

Table 3.9 Monetary yearly renewable term premiums

	$\Pi_0^{R(k)}$
$k=0$	157,404
$k=1$	145,186
$k=2$	95,278
$k=3$	46,890
$k=4$	0.0014

3.8 The Approximate Valuation Portfolio

In Sect. 3.2 we have constructed the VaPo for a rather simple example. We have considered a small homogeneous portfolio and its liabilities were easily described by financial instruments. In practice the situation is often more complicated. Life insurance companies have high-dimensional portfolios which usually involve embedded options and guarantees as well as management decisions, etc. Such VaPos become (highly) path dependent and the determination of the liability cash flows and the appropriate financial instruments is far from being trivial. In such situations one often tries to *approximate* the VaPo by a financial portfolio. Here, we will define the approximate VaPo (denoted by $\text{VaPo}^{\text{approx}}$) which plays the role of an approximate replicating portfolio.

We choose a filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ and we assume that we have an insurance liability cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ that we would like to describe.

In order to construct an approximate VaPo we choose a set of basic traded (or tradable) financial instruments $\mathcal{U}_1, \ldots, \mathcal{U}_q$ from which we believe that they can replicate the liabilities in an appropriate way and for which we can *easily* describe their price processes

$$(U_t^{(i)})_{t=0,\ldots,n} = (\mathcal{A}_t(\mathcal{U}_i))_{t=0,\ldots,n}, \qquad \text{for } i = 1, \ldots, q, \tag{3.86}$$

i.e. we want to choose q financial instruments of which we have a good understanding, both on the financial market side but also from the modelling perspective.

We now want to approximate the cash flow representation (3.60) of

$$\text{VaPo}_0(\mathbf{X}) = \sum_{k=0}^{n} \text{VaPo}_0(\mathbf{X}_k), \tag{3.87}$$

at time 0. That is, for all single cash flows X_k, $k = 0, \ldots, n$, our goal is to choose $\mathbf{y}_k^{(0)} \in \mathbb{R}^q$ such that the VaPo defined by

$$\text{VaPo}_0(\mathbf{Y}_k) = \sum_{i=1}^{q} y_{i,k}^{(0)}\, \mathcal{U}_i \tag{3.88}$$

approximates $\text{VaPo}_0(\mathbf{X}_k)$ as well as possible. Or in vector notation, we choose $\mathbf{y}^{(0)} \in \mathbb{R}^{q \times (n+1)}$ such that

$$\text{VaPo}_0(\mathbf{Y}) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}' \begin{pmatrix} \sum_{i=1}^q y_{i,0}^{(0)}\, \mathcal{U}_i \\ \vdots \\ \sum_{i=1}^q y_{i,n}^{(0)}\, \mathcal{U}_i \end{pmatrix} \tag{3.89}$$

approximates $\text{VaPo}_0(\mathbf{X})$ as well as possible, see (3.60). That is, our aim is to choose $\mathbf{y}^{(0)} \in \mathbb{R}^{q \times (n+1)}$ such that $\mathbf{X}$ and $\mathbf{Y}$ are “close”. Of course, “close” will depend on some distance function and, importantly, a cash flow $\mathbf{Y}$ is only obtained by selling the $\text{VaPo}_0(\mathbf{Y})$ at the right time points indicated by the second lower index in $y_{i,k}^{(0)}$.

If there does not exist insurance technical risk (i.e. everything is $\mathbb{G}$-adapted) and if $\mathcal{U}_1, \ldots, \mathcal{U}_q$ is a complete financial basis for the liabilities we can achieve

$$\mathbf{X} = \mathbf{Y}, \qquad P\text{-a.s.} \tag{3.90}$$

In general, we are not able to achieve (3.90) nor is it possible to evaluate the random vectors $\mathbf{X}$ and $\mathbf{Y}$ for all sample points $\omega \in \Omega$. Therefore, one chooses a finite set of so-called scenarios $\Omega_K = \{\omega_1, \ldots, \omega_K\} \subset \Omega$ and one evaluates the random vectors $\mathbf{X}$ and $\mathbf{Y}$ in these scenarios. We introduce a distance function

$$\text{dist}\left(\mathbf{X}(\cdot), \mathbf{Y}(\cdot), \Omega_K\right) \in \mathbb{R}, \tag{3.91}$$

and determine the allocation

$$\mathbf{y}^{\diamond(0)} = \arg \min_{\mathbf{y}^{(0)} \in \mathbb{R}^{q\times(n+1)}} \text{dist}\left(\mathbf{X}(\cdot), \mathbf{Y}(\cdot), \Omega_K\right). \tag{3.92}$$

Then, we define the approximate VaPo at time 0 for $k = 0, \ldots, n$

$$\text{VaPo}_0^{\text{approx}}(\mathbf{X}_k) = \sum_{i=1}^{q} y_{i,k}^{\diamond(0)}\, \mathcal{U}_i, \tag{3.93}$$

or, respectively,

$$\text{VaPo}_0^{\text{approx}}(\mathbf{X}) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}' \begin{pmatrix} \sum_{i=1}^{q} y_{i,0}^{\diamond(0)}\, \mathcal{U}_i \\ \vdots \\ \sum_{i=1}^{q} y_{i,n}^{\diamond(0)}\, \mathcal{U}_i \end{pmatrix}. \tag{3.94}$$

Remark. It is important to realize that the approximate VaPo generated by $\mathbf{y}^{\diamond(0)}$ depends on

(a) the choice of the financial instruments $\mathcal{U}_1, \ldots, \mathcal{U}_q$,
(b) the choice of the scenarios Ω_K,
(c) the choice of the distance function, and
(d) the choice of the accounting principles $(\mathcal{A}_t)_t$.

Based on the purpose (e.g. profit testing, solvency testing, extremal behaviour) these choices will vary and so will the approximate VaPo, and there does *not* exist an obvious best choice.

Example 3.5 (Cash flow matching) We assume that we want to match the entire cash flow $\mathbf{X}$ as well as possible and we use an L^2-distance measure. Choose positive deterministic weight functions $\chi_t : \Omega_K \to \mathbb{R}_+$ for $t = 0, \ldots, n$. Our first distance function is defined by

$$\operatorname{dist}\left(\mathbf{X}(\cdot), \mathbf{Y}(\cdot), \Omega_K\right) = \sum_{k=1}^{K} \sum_{t=0}^{n} \chi_t(\omega_k) \left(X_t(\omega_k) - Y_t(\omega_k)\right)^2, \tag{3.95}$$

where $Y_t(\omega_k)$ is obtained by the cash flow generated at time t of portfolio $\sum_{i=1}^{q} y_{i,t}^{(0)} \mathcal{U}_i$ under scenario $\omega_k \in \Omega_K$. For $\chi_t(\cdot)$ we can make different choices. Often one wants to account for time values, therefore one chooses the financial deflator $\varphi^{\mathbb{G}}$ (see Assumption 2.16) and a normalized positive deterministic weight function $p : \Omega_K \to \mathbb{R}_+$ with $\sum_{k=1}^{K} p(\omega_k) = 1$ and defines for $t = 0, \ldots, n$

$$\chi_t(\omega_k) = p(\omega_k) \; \left(\varphi_t^{\mathbb{G}}(\omega_k)\right)^2. \tag{3.96}$$

The distance function is then rewritten as

$$\begin{aligned} \operatorname{dist}\left(\mathbf{X}(\cdot), \mathbf{Y}(\cdot), \Omega_K\right) &= \sum_{k=1}^{K} p(\omega_k) \sum_{t=0}^{n} \left(\varphi_t^{\mathbb{G}}(\omega_k)\right)^2 \left(X_t(\omega_k) - Y_t(\omega_k)\right)^2 \\ &= E_K\left[\sum_{t=0}^{n} \left(\varphi_t^{\mathbb{G}} X_t - \varphi_t^{\mathbb{G}} Y_t\right)^2\right], \end{aligned} \tag{3.97}$$

where E_K denotes the expected value under the discrete probability measure P_K which assigns probability weights $p(\omega_k)$ to the scenarios in Ω_K. Henceforth, we aim to minimize the L^2-distance on the finite probability space $(\Omega_K, 2^{\Omega_K}, P_K)$.

The distance function defined in (3.97) tries to match pointwise in time the values of the cash flows $\mathbf{X}$ and $\mathbf{Y}$ as well as possible. Other approaches often work under equivalent probability measures (equivalent martingale measures or forward measures) so that the discount factors become measurable at the beginning of the corresponding periods. □

Example 3.6 (Time value matching)

We assume that we want to match the time value of $\mathbf{X}$ as well as possible and we use an L^2-distance measure. For a positive deterministic weight function χ_t we define similarly to (3.96) the distance function

$$\begin{aligned} \operatorname{dist}\left(\mathbf{X}(\cdot), \mathbf{Y}(\cdot), \Omega_K\right) &= \sum_{k=1}^{K} p(\omega_k) \left\{\sum_{t=0}^{n} \varphi_t^{\mathbb{G}}(\omega_k) \left(X_t(\omega_k) - Y_t(\omega_k)\right)\right\}^2 \\ &= E_K\left[\left(\sum_{t=0}^{n} \varphi_t^{\mathbb{G}} X_t - \varphi_t^{\mathbb{G}} Y_t\right)^2\right], \end{aligned} \tag{3.98}$$

where E_K denotes the expected value under the discrete probability measure P_K which assigns probability weights $p(\omega_k)$ to the scenarios in Ω_K.

The distance function defined in (3.98) tries to match the time value of the entire cash flows $\mathbf{X}$ and $\mathbf{Y}$ as well as possible. Note that the difference is that we match the entire time value of $\mathbf{X}$ in (3.98) whereas in (3.97) we match each cash flow X_k individually. □

Exercise 3.7 Calculate the approximate VaPos explicitly under distance functions (3.97) and (3.98), respectively.

Hints: Note that we have a quadratic form in (3.98). Set the gradient equal to zero and calculate the Hessian matrix (see Ingersoll [Ing87], formula (37) on page 8, and Wüthrich–Merz [WM13], Sect. 7.4). □

3.9 Conclusions on Chapter 3

We have decomposed the cash flow $\mathbf{X}$ in a two-step procedure:

1. Choose a multidimensional vector space whose basis consists of financial instruments $\mathcal{U}_1, \ldots, \mathcal{U}_p$.
2. Express the cash flow $\mathbf{X}$ as a vector in this vector space. The number of each unit is determined by the expected number of units (where the expectation is calculated with possibly distorted probabilities).

Calculating the monetary value of the VaPo is then the third step where we use an accounting principle to give values to the vectors in the multidimensional vector space.

We should mention that we have constructed our VaPo for a very basic example. In practice the VaPo construction is much more difficult because, for example, (a) modelling embedded options and guarantees can become very difficult, see Sect. 3.8; (b) often one does not have the necessary information on the single policies in the portfolio (e.g. collective policies). Moreover, in practice one faces a lot of problems concerning data storage and data management since the volume of the data can become very large.

Finally, we mention that we can also construct the VaPo if the financial instruments do not exist at the financial market, e.g. a 41-year zero coupon bond. The VaPo construction still works. However, calculating the monetary value of the VaPo is not straightforward if the instruments do not exist at the financial market, and in particular, this means that the financial market is not complete. In an incomplete market we choose a deflator φ that (a) matches prices of traded instruments, and (b) which allows us to extend these prices to non-traded instruments in a consistent (arbitrage-free) way. In this spirit we can value *any* cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$ under a given deflator φ, no matter whether these cash flows stem from traded or non-traded instruments.

Chapter 4
Financial Risks and Solvency

In the previous chapter we have constructed the VaPo for a portfolio of life insurance contracts. This VaPo can be viewed as a replicating portfolio of the insurance liabilities in terms of financial instruments. In this chapter we analyze the financial risks which derive from the fact that the VaPo and the real existing asset portfolio $\mathcal{S}$ (that the insurance company holds on the asset side of its balance sheet) typically differ.

4.1 Asset and Liability Management

We assume that the VaPo, the VaPo^{prot} and the VaPo^{approx}, respectively, describe the outstanding insurance liabilities in terms of financial instruments $\mathcal{U}_i$. We need to compare these VaPos to the real existing asset portfolio $\mathcal{S}$ that the insurance company holds on the asset side of its balance sheet, see Fig. 4.1.

In the sequel we drop the upper indexes "$prot$" and "$approx$" and, in general, we assume that we have constructed the appropriate VaPo. If we purchase this VaPo in terms of financial instruments, the resulting portfolio is called the *replicating portfolio* for the outstanding insurance liabilities. Therefore, it is convenient to use the VaPo for both: (1) the portfolio of liabilities, and (2) the replicating portfolio of assets, since they are physically the same portfolio.

▷ *Financial risks derive from the fact that the real existing asset portfolio $\mathcal{S}$ and the replicating portfolio* VaPo *typically differ.*

Financial risk management and *asset and liability management* (ALM) is concerned with maximizing the financial returns under the constraint that one has to cover the given liabilities VaPo. The ultimate goal is to obtain solvency at any time, which means that the chosen asset strategy should be able to cover the liabilities at any time. Since the VaPo and the asset portfolio $\mathcal{S}$ in general differ, solvency needs to be defined relative to an accounting principle, and this means that the chosen ALM

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1_4

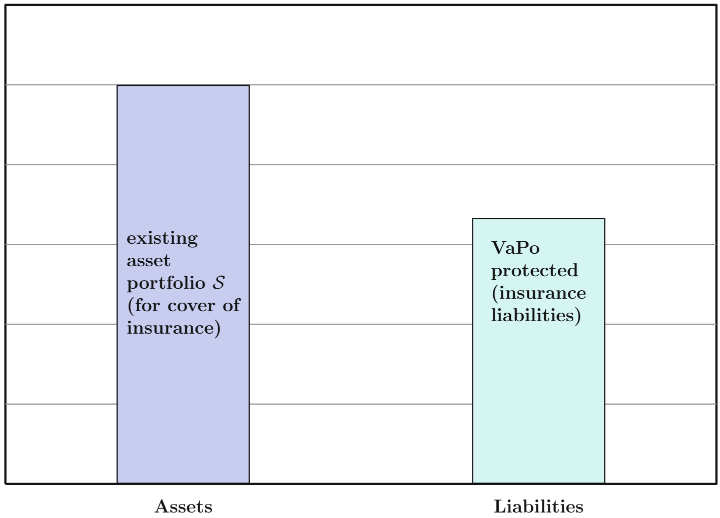


Fig. 4.1 Real existing asset portfolio $\mathcal{S}$ for the cover of outstanding insurance liabilities (on the *left-hand side*) and the VaPo protected against insurance technical risks VaPo^{prot} (on the *right-hand side*)

strategy needs to be able to cover the costs of the liabilities measured by the chosen accounting principle.

In actual solvency considerations one chooses an economic accounting principle $\mathcal{A}_t = \mathcal{E}_t$ (we refer to Sect. 3.3) because this allows us to transfer assets and liabilities at market(-consistent) prices. In our interpretation this means that we use market prices for traded instruments and market-consistent prices for non-traded instruments so that the chosen sequence of accounting principles $(\mathcal{A}_t)_{t=0,\dots,n}$ is consistent for all instruments and the resulting pricing system is free of arbitrage. In summary, we require the martingale property (3.8) for $(\varphi_t \mathcal{A}_t(\cdot))_{t=0,\dots,n}$ under the chosen (resulting) deflator $\boldsymbol{\varphi} \in L^2_{n+1}(P, \mathbb{F})$.

Definition 4.1 Choose $t_0 \in \{0, \dots, n-1\}$ and assume that the insurance company has outstanding liabilities $\text{VaPo} = \text{VaPo}_{t_0}$ and asset portfolio $\mathcal{S} = \mathcal{S}_{t_0}$ at time t_0. This insurance company is solvent at time t_0 if the following two conditions hold:

$$\mathcal{A}_{t_0}(\mathcal{S}) \geq \mathcal{A}_{t_0}(\text{VaPo}), \tag{4.1}$$

this is the accounting condition on the actual balance sheet, and

$$\mathcal{A}_t(\mathcal{S}) \geq \mathcal{A}_t(\text{VaPo}) \qquad \text{for all } t = t_0 + 1, \dots, n, \text{ a.s.}, \tag{4.2}$$

this is the insurance contract condition on the future balance sheets.

Remarks.

- Definition 4.1 is *our first* definition of solvency. Pay attention to the fact that there is not a unique definition of solvency, indeed the solvency rules in industry slightly differ from country to country. In particular, they depend on the risk classes considered, the risk measure used, the security level chosen, the stochastic model used, etc. Moreover, importantly, we could replace the economic accounting principle $\mathcal{E}_t$ by any other appropriate accounting principle $\mathcal{A}_t$ and we would obtain different solvency results, for instance, statutorial ones.
- The accounting condition (4.1) is necessary but not sufficient for solvency. It says that the (market-consistent) value of the outstanding liabilities is covered by today's asset value. If the market-consistent value refers to a transfer value, then the liabilities could be transferred to another company at an appropriate price.
- Either there is no insurance technical risk involved (i.e. the VaPo is deterministic w.r.t. insurance technical risks) or, if there is insurance technical risk involved, we consider the VaPo protected against insurance technical risks. In both situations the resulting VaPos are deterministic w.r.t. insurance technical risks, given information $\mathcal{F}_{t_0}$. Therefore, the insurance contract condition (4.2) *only* considers financial risks after time t_0.
- Viewed from time t_0, the future values $\mathcal{A}_t\,(\mathcal{S})$ and $\mathcal{A}_t$ (VaPo) are random variables for $t = t_0 + 1, \ldots, n$. Therefore, we require the insurance contract condition (4.2) to hold with $P[\cdot|\mathcal{F}_{t_0}]$-probability equal to 1 (we write in short 'a.s.' in (4.2)). At first sight this seems rather restrictive and in practice one often relaxes (4.2) to hold with high probability. This is further discussed in Sect. 4.5.
- In many solvency calculations the time interval under consideration is 1 year. Therefore one often assumes that the accounting condition is fulfilled at time t_0 and that the insurance contract condition should hold at time $t_0 + 1$ (this is the so-called one-year solvency view). At time $t_0 + 1$ one iterates this one-year procedure with a new accounting condition at time $t_0 + 1$ and so on (until the run-off of all outstanding liabilities is done). In the sequel we will also take this point of view. Then the problem of solvency decouples into one-period problems (that need to be calculated recursively and involve multiperiod risk measures, see also Salzmann–Wüthrich [SW10]).

  Considering successive *one-period solvency problems* also highlights the fact that we can restructure and update the asset portfolio $\mathcal{S} = \mathcal{S}_t$ at any future time point $t > t_0$ (dynamic asset allocation). This also holds true for the outstanding liabilities VaPo_t because they will also evolve over time (insurance technical risks are $\mathbb{T}$-adapted). These updates are not considered in Definition 4.1, yet.

The task of financial risk management. If (4.1) is satisfied: how do we need to choose the asset portfolio $\mathcal{S} = \mathcal{S}_{t_0}$ at time t_0 such that the company with outstanding liabilities $\text{VaPo} = \text{VaPo}_{t_0}$ is solvent at time t_0?

(a) Prudent solution. Choose $\mathcal{S}$ at time t_0 as follows

$$\mathcal{S} = \text{VaPo} + F, \tag{4.3}$$

where VaPo is the replicating portfolio of the outstanding liabilities and F are the *free reserves* which need to satisfy $\mathcal{A}_t(F) \geq 0$ for all $t \geq t_0$, a.s. Henceforth, solvency is guaranteed with this asset strategy which, from a mathematical point of view, shows that solvency is possible.

(b) Realistic situation.

- $\mathcal{S}$ does not (entirely) contain VaPo.
- ALM mismatch between $\mathcal{S}$ and VaPo is often desired, because taking additional financial risks on the asset side of the balance sheet opens the possibility of receiving higher investment returns.
- This mismatch asks for an additional protection against financial risks to achieve solvency. In fact, regulators ask for a substantially increased risk capacity for the protection against financial risks. It turns out in the Swiss Solvency Test [SST06] that the financial risk is the dominant risk driver for life insurance companies, whereas in a typical non-life insurance company the risk capital for financial risks has about the same size as the one for insurance technical risks.

4.2 The Procedure to Control Financial Risks

As described above, we decouple the solvency problem into *one-period solvency problems*, and we replace Definition 4.1 by the following one.

Definition 4.2 Choose $t_0 \in \{0, \ldots, n-1\}$ and assume that the insurance company has outstanding liabilities $\text{VaPo} = \text{VaPo}_{t_0}$ and asset portfolio $\mathcal{S} = \mathcal{S}_{t_0}$ at time t_0. This insurance company is solvent at time t_0 if the following two conditions hold:

$$\mathcal{A}_{t_0}(\mathcal{S}) \geq \mathcal{A}_{t_0}(\text{VaPo}), \tag{4.4}$$

this is the accounting condition on the actual balance sheet, and

$$\mathcal{A}_{t_0+1}(\mathcal{S}) \geq \mathcal{A}_{t_0+1}(\text{VaPo}), \qquad \text{a.s.}, \tag{4.5}$$

this is the insurance contract condition on the next balance sheet.

We decompose the asset portfolio $\mathcal{S}$ at the beginning t_0 of the calendar year $(t_0, t_0+1]$ into three parts:

$$\mathcal{S} = \widetilde{\mathcal{S}} + \mathcal{M} + F, \tag{4.6}$$

where $\widetilde{\mathcal{S}}$ is any asset portfolio which satisfies the accounting condition (4.4), $\mathcal{M}$ is a margin which is determined below and F denotes the remainder, i.e. we consider an asset allocation (4.6) for which

$$\begin{array}{ll} \mathcal{A}_{t_0}(\widetilde{\mathcal{S}}) = \mathcal{A}_{t_0}(\text{VaPo}) & \text{accounting condition,} \\ \mathcal{M} & \text{margin (further determined below),} \\ F & \text{free reserves.} \end{array} \tag{4.7}$$

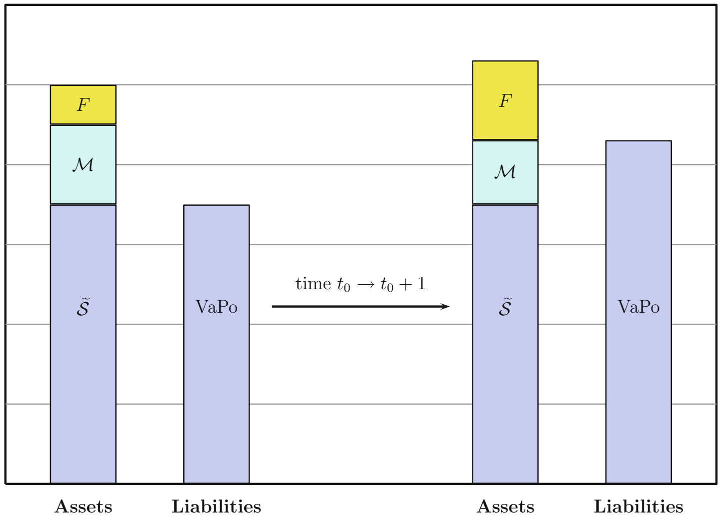


Fig. 4.2 Time evolution of the asset and liability portfolios

At the end $t_0 + 1$ of the calendar year $(t_0, t_0 + 1]$ we should be able to

(a) switch from $\widetilde{\mathcal{S}} + \mathcal{M}$ to VaPo, if necessary, and
(b) $\mathcal{A}_t(F)$ is not allowed to become negative for $t = t_0, t_0 + 1$, a.s.

This implies that the margin $\mathcal{M}$ is determined such that we receive *the right* to switch from asset portfolio $\widetilde{\mathcal{S}} + \mathcal{M}$ to VaPo at the end of the calendar year $(t_0, t_0 + 1]$, if necessary (Fig. 4.2).

Formalizing (a). A Margrabe option gives us the right to exchange one asset for another at a given maturity date. It is named after William Margrabe [Ma78].

In terms of financial instruments, $\mathcal{M}$ is chosen to be a Margrabe option that allows us to switch from the asset portfolio $\widetilde{\mathcal{S}}$ to the asset portfolio VaPo whenever

$$\mathcal{A}_{t_0+1}\,(\text{VaPo}) > \mathcal{A}_{t_0+1}(\widetilde{\mathcal{S}}). \tag{4.8}$$

This means that the decomposition (4.6) is chosen such that

(1) $\mathcal{A}_{t_0}(\widetilde{\mathcal{S}}) = \mathcal{A}_{t_0}\,(\text{VaPo})$ satisfies the accounting condition;
(2) $\mathcal{M}$ allows us to switch from $\widetilde{\mathcal{S}} + \mathcal{M}$ to VaPo at time $t_0 + 1$ whenever (4.8) occurs at time $t_0 + 1$. Note that the Margrabe option is a right (and not an obligation), which implies that $\mathcal{A}_t(\mathcal{M}) \geq 0$ for $t = t_0, t_0 + 1$, a.s.;
(3) $\mathcal{A}_t(F) \geq 0$ for $t = t_0, t_0 + 1$, a.s.

In the next section we calculate the price of the Margrabe option $\mathcal{M}$. We therefore consider the two price processes generated by $\widetilde{\mathcal{S}}$ and VaPo:

$$Y_t = \mathcal{A}_t(\widetilde{\mathcal{S}}), \tag{4.9}$$
$$V_t = \mathcal{A}_t\,(\text{VaPo})\,. \tag{4.10}$$

For explicit calculations it may be useful to consider a continuous time model $t \in [t_0, t_0 + 1]$ because this allows us to apply classical financial mathematics like the geometric Brownian motion framework (when we discuss hedging).

4.3 Financial Modelling of the Margrabe Option

4.3.1 Stochastic Discounting: Repetition

We briefly refresh the valuation techniques introduced in Sect. 2.6. To model the financial market described by $\mathbb{G}$, throughout we set Assumption 2.16 and we assume that the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$ is a density process.

We choose a fixed financial deflator $\boldsymbol{\varphi}^{\mathbb{G}} \in L^2_{n+1}(P, \mathbb{G})$ on the (financial) filtered probability space $(\Omega, \mathcal{G}_n, P, \mathbb{G})$. The price processes of all financial instruments $\mathcal{U}$ should satisfy $(U_t)_{t=0,\ldots,n} \in L^2_{n+1}(P, \mathbb{G})$ and the deflated price processes $(\varphi_t^{\mathbb{G}} U_t)_{t=0,\ldots,n}$ need to be $(P, \mathbb{G})$-martingales, see Theorem 2.19 and Remarks 2.15. In particular, for $s \le t$ we have identity

$$\varphi_s^{\mathbb{G}}\, U_s = E\left[\varphi_t^{\mathbb{G}}\, U_t \middle| \mathcal{G}_s\right]. \tag{4.11}$$

The vector $\mathbf{Z}^{(t)} = (0, \ldots, 0, 1, 0, \ldots, 0) \in L^2_{n+1}(P, \mathbb{G})$ denotes the cash flow of the zero coupon bond $\mathcal{Z}^{(t)}$ paying 1 at maturity date t. Its price $P(s, t)$ at time $s \le t$ is given by

$$P(s, t) = Q_s\left[\mathbf{Z}^{(t)}\right] = \frac{1}{\varphi_s^{\mathbb{G}}}\, E\left[\varphi_t^{\mathbb{G}} \middle| \mathcal{G}_s\right]. \tag{4.12}$$

If we consider the equivalent martingale measure $P^* \sim P$ for the bank account numeraire $(B_t)_{t=0,\ldots,n}$, see Lemma 2.11 and Corollary 2.13, then we get discount factors that are known at the beginning of the period under consideration (previsible). In particular, we have the martingale property

$$B_{t-1}^{-1} U_{t-1} = B_t^{-1} E^*\left[U_t \middle| \mathcal{G}_{t-1}\right], \tag{4.13}$$

for all financial instruments $\mathcal{U}$, compare to (4.11).

Exercise 4.1 *(Pricing of financial assets)*
We revisit the discrete time Vasiček model of Exercise 2.3. In all the corresponding assumptions and statements we (may) replace the filtration $\mathbb{F}$ by the financial filtration $\mathbb{G}$. We assume that we have a two-dimensional process $(\varepsilon_t, \delta_t)_{t=0,\dots,n}$ that is $\mathbb{G}$-adapted, and that $(\varepsilon_t, \delta_t)$ is independent of $\mathcal{G}_{t-1}$ with a two-dimensional Gaussian distribution with means 0, variances 1 and correlation ρ for $t = 1, \dots, n$, under P.

Assume that the financial asset $\mathcal{U}$ has a price process given by $U_0 > 0$ (fixed) and for $t = 1, \dots, n$

$$U_t = U_{t-1} \ \exp\{\mu_t - \sigma \delta_t\}, \tag{4.14}$$

for given constant parameters $\mu_t \in \mathbb{R}$ and $\sigma > 0$.

Determine the necessary properties of $\mu_t \in \mathbb{R}$ so that the price process

$$(\varphi_t^{\mathbb{G}} U_t)_{t=0,\dots,n} \quad \text{is a } (P, \mathbb{G})\text{-martingale}, \tag{4.15}$$

for the financial deflator $\boldsymbol{\varphi}^{\mathbb{G}}$ given in (2.25).

Hint: use the properties of log-normal distributions.

□

4.3.2 *Modelling Margrabe Options*

Recall definitions (4.9) and (4.10) of Y_t and V_t, respectively. Our goal is to price the Margrabe option $\mathcal{M}$ introduced in Sect. 4.2. The Margrabe option is exercised at time $t_0 + 1$ whenever (4.8) occurs. Therefore, it generates the cash flow

$$\mathbf{X} = \big(0, \dots, 0, (V_{t_0+1} - Y_{t_0+1})_+, 0, \dots, 0\big) \in L^2_{n+1}(P, \mathbb{F}), \tag{4.16}$$

where the potentially non-zero entry is at the component $t_0 + 2$. Observe that this cash flow is $\mathcal{H}_{t_0+1} = \sigma(\mathcal{F}_{t_0}, \mathcal{G}_{t_0+1})$-measurable. $\mathcal{H}_{t_0+1}$ has the interpretation of the hedgeable filtration from t_0 to $t_0 + 1$, see also Malamud et al. [MTW08], with

$$\mathcal{F}_{t_0} \subset \mathcal{H}_{t_0+1} = \sigma(\mathcal{T}_{t_0}, \mathcal{G}_{t_0+1}) \subset \mathcal{F}_{t_0+1}. \tag{4.17}$$

Thus, we can identify the Margrabe option $\mathcal{M}$ with the cash flow that it produces at time $t_0 + 1$ and its price is at time t_0 given by

$$M_{t_0} = Q_{t_0}[\mathbf{X}] = \frac{1}{\varphi_{t_0}^{\mathbb{G}}} \, E\left[\varphi_{t_0+1}^{\mathbb{G}} \left(V_{t_0+1} - Y_{t_0+1}\right)_+ \middle| \mathcal{F}_{t_0}\right] \geq 0. \tag{4.18}$$

It is convenient to make a change of numeraire. Assume that $(V_t)_{t=t_0,t_0+1} = (\mathcal{A}_t(\text{VaPo}))_{t=t_0,t_0+1} \gg 0$ is strictly positive. In this case, the price process $(V_t)_{t=t_0,t_0+1}$ may serve as a numeraire as follows. We define the density process $\boldsymbol{\zeta} = (\zeta_t)_{t=t_0,t_0+1}$ for $t = t_0, t_0 + 1$ by

$$\zeta_t = \varphi_t^{\mathbb{G}} V_t / (\varphi_{t_0}^{\mathbb{G}} V_{t_0}). \tag{4.19}$$

This is a strictly positive $(P, (\mathcal{F}_{t_0}, \mathcal{H}_{t_0+1}))$-martingale that is normalized, and hence it is a density process w.r.t. the filtration $(\mathcal{F}_{t_0}, \mathcal{H}_{t_0+1})$ and P. This allows us to define the equivalent probability measure $P^Y \sim P$ via the Radon–Nikodým derivative

$$\left.\frac{dP^Y}{dP}\right|_{\mathcal{H}_{t_0+1}} = \zeta_{t_0+1} = \varphi^{\mathbb{G}}_{t_0+1} V_{t_0+1} / (\varphi^{\mathbb{G}}_{t_0} V_{t_0}). \tag{4.20}$$

Under this equivalent probability measure P^Y the price of the Margrabe option $\mathcal{M}$ at time t_0 is given by, see also Lemma 2.11,

$$\begin{aligned} M_{t_0} \;=\; Q_{t_0}[\mathbf{X}] &= \frac{1}{\varphi^{\mathbb{G}}_{t_0}} \, E\left[\varphi^{\mathbb{G}}_{t_0+1} \left(V_{t_0+1} - Y_{t_0+1} \right)_+ \middle| \mathcal{F}_{t_0} \right] \\ &= \frac{1}{\varphi^{\mathbb{G}}_{t_0}} \, E\left[\varphi^{\mathbb{G}}_{t_0+1} V_{t_0+1} \left(1 - \frac{Y_{t_0+1}}{V_{t_0+1}} \right)_+ \middle| \mathcal{F}_{t_0} \right] \\ &= V_{t_0} \, E\left[\zeta_{t_0+1} \left(1 - \frac{Y_{t_0+1}}{V_{t_0+1}} \right)_+ \middle| \mathcal{F}_{t_0} \right] \\ &= V_{t_0} \, E^Y\left[\left(1 - \widetilde{Y}_{t_0+1} \right)_+ \middle| \mathcal{F}_{t_0} \right], \end{aligned} \tag{4.21}$$

where E^Y denotes expectation under the measure $P^Y \sim P$ and where we set for $t = t_0, t_0 + 1$

$$\widetilde{Y}_t = \frac{Y_t}{V_t} = \frac{\mathcal{A}_t(\widetilde{\mathcal{S}})}{\mathcal{A}_t\,(\mathrm{VaPo})}. \tag{4.22}$$

The advantage of using $(\widetilde{Y}_t)_t$ under the probability measure P^Y is that the deflator disappears since both expressions Y_t and V_t have the same time value. Growth of $\widetilde{Y}_t$ means that we have an extensive growth of the assets Y_t relative to the liabilities V_t. There only remains the calculation of the right-hand side of (4.21) under the probability measure P^Y in order to get the price $M_{t_0} = Q_{t_0}[\mathbf{X}]$ of the Margrabe option $\mathcal{M}$ at time t_0. An example is provided in Example 4.2.

Observe that we have for the price process $(\widetilde{Y}_t)_t$ under the probability measure P^Y

$$\begin{aligned} E^Y\left[\widetilde{Y}_{t_0+1} \middle| \mathcal{F}_{t_0} \right] &= E\left[\zeta_{t_0+1} \widetilde{Y}_{t_0+1} \middle| \mathcal{F}_{t_0} \right] \\ &= \frac{1}{\varphi^{\mathbb{G}}_{t_0} V_{t_0}} \, E\left[\varphi^{\mathbb{G}}_{t_0+1} V_{t_0+1} \frac{Y_{t_0+1}}{V_{t_0+1}} \middle| \mathcal{F}_{t_0} \right] \;=\; \widetilde{Y}_{t_0} \;=\; 1, \end{aligned} \tag{4.23}$$

where we have also used (3.8). From this we see that $(\widetilde{Y}_t)_t$ is a $(P^Y, (\mathcal{F}_{t_0}, \mathcal{H}_{t_0+1}))$-martingale which is needed for the calibration of $(\widetilde{Y}_t)_t$ under P^Y.

Example 4.2 (*Margrabe option, log-normal example*)
We assume that $\widetilde{Y}_{t_0+1} = \exp\{W\}$ has a log-normal distribution with parameters μ^Y and σ^2, conditionally given $\mathcal{F}_{t_0}$ and w.r.t. P^Y. Since deflated price processes are

martingales, see (4.23), we obtain

$$\widetilde{Y}_{t_0} = E^Y\left[\widetilde{Y}_{t_0+1}\middle|\mathcal{F}_{t_0}\right] = \exp\left\{\mu^Y + \sigma^2/2\right\}. \tag{4.24}$$

The accounting condition $Y_{t_0} = V_{t_0}$ implies $\widetilde{Y}_{t_0} = 1$ and provides the drift condition $\mu^Y = -\sigma^2/2$. Henceforth, the right-hand side of (4.21) is simply the price of a European put option for log-normal prices. We calculate

$$\begin{aligned} E^Y\left[\left(1-\widetilde{Y}_{t_0+1}\right)_+\middle|\mathcal{F}_{t_0}\right] &= E^Y\left[\left(1-\widetilde{Y}_{t_0+1}\right)1_{\{\widetilde{Y}_{t_0+1}\le 1\}}\middle|\mathcal{F}_{t_0}\right] \\ &= P^Y\left[\widetilde{Y}_{t_0+1}\le 1\middle|\mathcal{F}_{t_0}\right] - E^Y\left[\widetilde{Y}_{t_0+1}1_{\{\widetilde{Y}_{t_0+1}\le 1\}}\middle|\mathcal{F}_{t_0}\right] \\ &= P^Y\left[W\le 0|\mathcal{F}_{t_0}\right] - \int_0^1 \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{y}\, y\, \exp\left\{-\frac{1}{2}\frac{\left(\log y+\sigma^2/2\right)^2}{\sigma^2}\right\}dy \\ &= P^Y\left[W\le 0|\mathcal{F}_{t_0}\right] - \int_0^1 \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{y}\exp\left\{-\frac{1}{2}\frac{\left(\log y-\sigma^2/2\right)^2}{\sigma^2}\right\}dy \\ &= P^Y\left[W\le 0|\mathcal{F}_{t_0}\right] - P^Y\left[\widetilde{W}\le 0\middle|\mathcal{F}_{t_0}\right], \end{aligned} \tag{4.25}$$

with $W|\mathcal{F}_{t_0} \overset{P^Y}{\sim} \mathcal{N}(-\sigma^2/2,\ \sigma^2)$ and $\widetilde{W}|\mathcal{F}_{t_0} \overset{P^Y}{\sim} \mathcal{N}(\sigma^2/2,\ \sigma^2)$. This implies for the price of the Margrabe option $\mathcal{M}$

$$M_{t_0} = V_{t_0}E^Y\left[\left(1-\widetilde{Y}_{t_0+1}\right)_+\middle|\mathcal{F}_{t_0}\right] = V_{t_0}\left(\Phi(\sigma/2)-\Phi(-\sigma/2)\right), \tag{4.26}$$

where $\Phi(\cdot)$ denotes the standard Gaussian distribution function. Note that (4.26) is the Black–Scholes price for a European put option, see Lamberton–Lapeyre [LL91], Sect. 3.2.

We find the following relative loadings (depending on the volatility of the assets relative to the liabilities):

σ	Price relative to V_{t_0} (%)
0.05	1.99
0.10	3.99
0.20	7.97
0.30	11.92

□

Conclusions.
We have decoupled the solvency problem into recursive one-period solvency problems. To protect against financial risks one has to invest each year the price of the Margrabe option. This price measures the ALM mismatch between the real existing asset portfolio $\widetilde{\mathcal{S}}$ and the outstanding liability portfolio VaPo.

The agents who are entitled to receive the earnings beyond the VaPo should also finance this option:

- With-profit policies share the price between the policyholder and the shareholder according to their participation.
- For non-participating policies the shareholder has to pay the full price for financial risks.

As the price of the Margrabe option is relative to the VaPo, we can easily make a similar calculation for any version of the VaPo. If the VaPo (protected against insurance technical risks) and the Margrabe option $\mathcal{M}$ cannot be financed we need to (a) have more capital, (b) do a better ALM, and/or (c) reduce insurance technical risks.

4.4 Hedging Margrabe Options

In the previous section we calculated the price M_{t_0} of the Margrabe option $\mathcal{M}$ at time t_0 in Example 4.2 under a log-normal assumption. The difficulty in the real world is that in most cases the Margrabe option for specific portfolios $\widetilde{\mathcal{S}}$ and VaPo is not traded at the financial market. Therefore, there does not exist a seller of such an option. This implies that the insurance company needs to hedge this option internally, using the price of the Margrabe option. At this point the discrete time modelling framework breaks down because hedging typically needs to be done in continuous time. We therefore extend our modelling framework to continuous time, but only in this section.

We choose a filtered probability space $(\Omega, \mathcal{H}_{t_0+1}, P^Y, \mathbb{H})$ with filtration $\mathbb{H} = (\mathcal{H}_t)_{t\in[t_0,t_0+1]}$ satisfying the usual conditions (completeness and right-continuity, see for example Filipović [Fi09], p. 59), $\mathcal{H}_{t_0} = \mathcal{F}_{t_0}$ and $\mathcal{H}_{t_0+1} = \sigma(\mathcal{F}_{t_0}, \mathcal{G}_{t_0+1})$. That is, we fix insurance technical risk at time t_0 through $\mathcal{F}_{t_0}$ and then consider the hedgeable part within $(t_0, t_0 + 1]$.

Assume that $(\widetilde{Y}_t)_t$ is an $\mathbb{H}$-adapted geometric Brownian motion on that filtered probability space having zero drift under P^Y. That is, for $t \in [t_0, t_0 + 1]$ it is of the form

$$\widetilde{Y}_t = \widetilde{Y}_{t_0} \exp\left\{-(t - t_0)\sigma^2/2 + \sigma W_{t-t_0}\right\}, \tag{4.27}$$

where $(W_{t-t_0})_{t\geq t_0}$ is a $(P^Y, \mathbb{H})$-standard Brownian motion with $W_{t_0} = 0$, a.s. Note that the random variable $\widetilde{Y}_{t_0+1}$ has a log-normal distribution under P^Y, conditionally given $\mathcal{H}_{t_0}$, with mean $\widetilde{Y}_{t_0}$ and variance parameter σ^2. Therefore, we obtain *a* continuous-time extension of (4.24) if we start at $\widetilde{Y}_{t_0} = 1$, a.s., at time t_0. This continuous-time extension possesses the necessary $(P^Y, \mathbb{H})$-martingale property which allows us to model *an* entire price process in the time interval $[t_0, t_0 + 1]$. Under this continuous-time extension of (4.21) we aim to calculate the price of the Margrabe option $\mathcal{M}$ at time $t \in [t_0, t_0 + 1]$

$$M_t = V_t E^Y \left[\left(1 - \widetilde{Y}_{t_0+1}\right)_+ \middle| \mathcal{H}_t \right]. \tag{4.28}$$

This provides under (4.27) the well-known Black–Scholes price for $t \in [t_0, t_0 + 1)$, see for instance Chap. 4 in Lamberton–Lapeyre [LL91],

$$M_t = V_t \Phi \left(\frac{-\log \widetilde{Y}_t}{\sigma_t} + \sigma_t/2 \right) - Y_t \, \Phi \left(\frac{-\log \widetilde{Y}_t}{\sigma_t} - \sigma_t/2 \right), \tag{4.29}$$

for $\sigma_t^2 = \sigma^2 \ (t_0 + 1 - t)$. Note that for $t = t_0$ we have $\sigma_{t_0}^2 = \sigma^2$, $\log \widetilde{Y}_{t_0} = 0$ and $V_{t_0} = Y_{t_0}$ which implies that M_{t_0} is equal to (4.26).

Define the function

$$H(t,x) = x \; \Phi \left(\frac{\log x + \sigma_t^2/2}{\sigma_t} \right) - \Phi \left(\frac{\log x - \sigma_t^2/2}{\sigma_t} \right), \tag{4.30}$$

and note that we have

$$M_t = Y_t \; H\left(t, \widetilde{Y}_t^{-1}\right). \tag{4.31}$$

By extending our model to the full Black–Scholes model, one considers (based on Itô calculus) the following candidate for the hedging strategy

$$\begin{aligned} &\frac{\partial}{\partial x} H(t,x) \qquad (4.32) \\ &= \Phi \left(\frac{\log x + \sigma_t^2/2}{\sigma_t} \right) + \varphi \left(\frac{\log x + \sigma_t^2/2}{\sigma_t} \right) / \sigma_t - \varphi \left(\frac{\log x - \sigma_t^2/2}{\sigma_t} \right) / (x \sigma_t) \\ &= \Phi \left(\frac{\log x + \sigma_t^2/2}{\sigma_t} \right), \end{aligned}$$

which is a well-known expression for the European call option in the Black–Scholes model (see e.g. Remarque 3.6 in Lamberton–Lapeyre [LL91] on p. 79). This motivates the hedging strategy $\psi = (\widetilde{\lambda}_t, \lambda_t)_t$ given by the following natural candidate (see e.g. Sect. 3.3 in Lamberton–Lapeyre [LL91]): Invest

$$\widetilde{\lambda}_t = \left. \frac{\partial}{\partial x} H(t,x) \right|_{x = \widetilde{Y}_t^{-1}} = \Phi \left(\frac{-\log \widetilde{Y}_t + \sigma_t^2/2}{\sigma_t} \right) \tag{4.33}$$

into the asset V_t and

$$\lambda_t = 1 - \Phi \left(\frac{-\log \widetilde{Y}_t - \sigma_t^2/2}{\sigma_t} \right) \tag{4.34}$$

into asset Y_t. This provides the value of the hedging portfolio at any time t

Table 4.1 One realization of $(\widetilde{Y}_t)_{t\in[t_0,t_0+1]}$ with its associated value process $(Y_t + M_t)_{t\in[t_0,t_0+1]}$ for $\sigma = 0.05$ and initialization $V_t \equiv 1$ for the Black–Scholes model

$\widetilde{Y}_t$	$\lambda_{t-1}\ Y_t$	$\widetilde{\lambda}_{t-1}\ V_t$	$Y_t + M_t$ (%)	λ_t (%)	$\widetilde{\lambda}_t$ (%)
1.0000	–	–	101.99	51.0	51.0
1.0077	51.4 %	51.0 %	102.39	57.3	44.6
1.0215	58.5 %	44.6 %	103.12	68.8	32.9
1.0295	70.8 %	32.9 %	103.66	75.6	25.8
1.0100	76.3 %	25.8 %	102.13	60.4	41.2
1.0423	63.0 %	41.2 %	104.14	86.5	14.3
1.0443	90.4 %	14.3 %	104.68	89.3	11.3
1.0192	91.0 %	11.3 %	102.38	72.8	28.3
1.0252	74.6 %	28.3 %	102.92	81.0	19.8
1.0062	81.5 %	19.8 %	101.30	60.2	40.7
0.9901	59.7 %	40.7 %	100.37	31.7	69.1
0.9941	31.5 %	69.1 %	100.54	34.4	66.2
0.9845	33.8 %	66.2 %	100.00	0.0	100.0

$$\begin{aligned} \widetilde{\lambda}_t\ V_t + \lambda_t\ Y_t &= V_t\widetilde{\lambda}_t + Y_t\lambda_t \\ &= V_t\Phi\left(\frac{-\log \widetilde{Y}_t + \sigma_t^2/2}{\sigma_t}\right) + Y_t\left(1 - \Phi\left(\frac{-\log \widetilde{Y}_t - \sigma_t^2/2}{\sigma_t}\right)\right) \\ &= Y_t + M_t, \end{aligned} \tag{4.35}$$

which means that we can switch to the VaPo at any time $t \in [t_0, t_0 + 1]$. For a detailed and more concise treatment we refer to Lamberton–Lapeyre [LL91].

Example 4.3
We choose $\sigma = 0.05$ as in Example 4.2. The price of the Margrabe option is under the log-normal model $M_{t_0} = 1.99\,\% \cdot V_{t_0}$. We then choose a specific realization of $(\widetilde{Y}_t)_t$, see the first column in Table 4.1, and for simplicity we set $V_t \equiv 1$. This leads to the developments of the price processes given in Table 4.1. If we plot the processes for three different realizations of $(\widetilde{Y}_t)_t$ we obtain a picture as shown in Fig. 4.3. Observe that the path of $(Y_t + M_t)_t$ never falls below $(V_t)_t \equiv 1$, i.e. we have full financial coverage of all liabilities during the whole investment period. □

Conclusions. If we cannot buy the Margrabe option we need to hedge it for switching from the asset portfolio $\widetilde{S}$ to the VaPo at the end of each calendar year. The hedging strategy can be made cheaper if we only hedge for switching at the point when the money is needed (because the price of the Margrabe option becomes cheaper if we only hedge the necessary maturity date).

(a) Money is needed in n years. Switch at the end of every year. Henceforth, the price is
$$n \cdot \text{price Margrabe option}(\sigma). \tag{4.36}$$

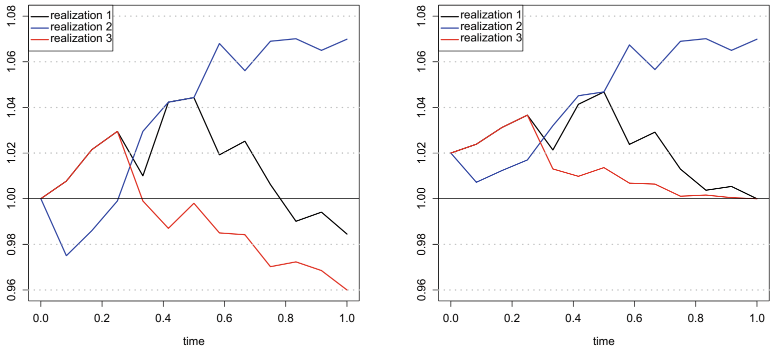


Fig. 4.3 (lhs) Three different realizations of $(\widetilde{Y}_t)_t$, and (rhs) the corresponding three value processes $(Y_t + M_t)_t$ in the Black–Scholes model for initialization $V_t \equiv 1$

(b) Money is needed in n years. Switch at the end of that period. Henceforth, the price is

$$\text{price Margrabe option}(\sqrt{n}\,\sigma). \tag{4.37}$$

Roughly speaking: approach (a) corresponds to a yearly guarantee whereas approach (b) only corresponds to a final wealth guarantee. Therefore it is clear that approach (a) needs to be more expensive. Observe also that in case (a) we may release part of the potential gain $(Y_t - V_t)_+$ every year (as long as future ALM strategies can be financed), whereas in case (b) these potential profits must be left in the risk and value process.

4.5 The Risk Measure Approach

Solvency Definitions 4.1 and 4.2 are very restrictive because they require an a.s. condition to be fulfilled. In real life situations it is not possible to fulfill this a.s. condition because the Margrabe option is neither sold at the market nor can it be perfectly hedged in real life. That is, Definition 4.2 only works in a sufficiently nice model world. For this reason we need to relax Definition 4.2 in order to get a practical solvency definition. We therefore introduce the general notion of risk measures, see McNeil et al. [MFE15].

To keep things simple we set $t_0 = 0$ for which $\mathcal{G}_{t_0} = \mathcal{F}_{t_0} = \{\emptyset, \Omega\}$. A risk measure ϱ on $L^1(P)$ is a function

$$\varrho : L^1(P) \to \overline{\mathbb{R}}; \qquad Y \mapsto \varrho(Y). \tag{4.38}$$

That is, a risk measure ϱ attaches to every risky position $Y \in L^1(P)$ a value $\varrho(Y) \in \overline{\mathbb{R}}$. This value is going to be used to check the solvency condition. In order to have a useful risk measure we assume that ϱ fulfills additional properties. Let us start with a list of possible properties, where in the following $Y, Z \in L^1(P)$, $c \in \mathbb{R}$ and $\lambda > 0$, then:

(a) normalization: $\varrho(0) = 0$;
(b) monotonicity: for all Y, Z with $Y \leq Z$, P-a.s., we have $\varrho(Y) \leq \varrho(Z)$;
(c) translation invariance: for all Y and c we have $\varrho(Y + c) = \varrho(Y) + c$;
(d) positive homogeneity: for all Y and $\lambda > 0$ we have $\varrho(\lambda Y) = \lambda\varrho(Y)$;
(e) subadditivity: for all Y, Z we have $\varrho(Y + Z) \leq \varrho(Y) + \varrho(Z)$.

A risk measure ϱ that fulfills all these properties (a)–(e) is called a coherent risk measure which, in general, is considered to be a "good" risk measure. Expected shortfall used in Exercise 2.9 is a coherent risk measure (note that the distribution function F_Y of Y was assumed to be continuous), whereas Value-at-Risk is not a coherent risk measure.

For the present outline we do not require coherence, in general.

Definition 4.3 Assume that the risk measure ϱ on $L^1(P)$ is (a) normalized and (b) monotone. Assume that the insurance company has outstanding liabilities $\mathrm{VaPo} = \mathrm{VaPo}_0$ and asset portfolio $\mathcal{S} = \mathcal{S}_0$ at time 0. This insurance company is solvent at time 0 w.r.t. ϱ if the following two conditions hold:

$$\mathcal{A}_0\left(\mathcal{S}\right) \geq \mathcal{A}_0\left(\mathrm{VaPo}\right), \tag{4.39}$$

this is the accounting condition on the actual balance sheet, and

$$\varrho\left(\mathcal{A}_1\left(\mathrm{VaPo}\right) - \mathcal{A}_1\left(\mathcal{S}\right)\right) \leq 0, \tag{4.40}$$

this is the insurance contract condition on the next balance sheet.

Lemma 4.4 *Set $t_0 = 0$. Solvency according to Definition 4.2 implies solvency of Definition 4.3.*

Proof of Lemma 4.4. We only need to verify the insurance contract condition (4.40). From (4.5) we obtain

$$\mathcal{A}_1\left(\mathrm{VaPo}\right) - \mathcal{A}_1\left(\mathcal{S}\right) \leq 0, \qquad P\text{-a.s.} \tag{4.41}$$

Monotonicity and normalization of ϱ prove the claim. □

Lemma 4.4 implies that solvency Definition 4.3 is weaker than solvency Definition 4.2. In practical applications one now chooses an appropriate risk measure ϱ in order to answer the solvency question. Solvency II uses the Value-at-Risk risk measure on a 99.5 % security level and the Swiss Solvency Test uses the expected shortfall risk measure on a 99 % security level. We briefly study these two risk measures and compare them to the Margrabe option approach of Example 4.2.

Example 4.5 (*Value-at-Risk*)
We set for $t \in \{0, 1\}$: $V_t = \mathcal{A}_t$ (VaPo) and $Y_t = \mathcal{A}_t(\widetilde{\mathcal{S}})$ with initial condition $V_0 = Y_0$. The aim is to find a minimal loading constant $q_0 \in \mathbb{R}_+$ such that

$$\varrho\left(V_1 - (1+q_0)Y_1\right) \leq 0. \tag{4.42}$$

Note that the existence of such a loading constant implies that the insurance company is solvent w.r.t. the risk measure ϱ if it holds asset portfolio $\mathcal{S} = (1+q_0)\widetilde{\mathcal{S}}$ at time 0. Moreover, this also specifies the asset allocation on the risk margin $q_0 Y_0 = q_0 \mathcal{A}_0(\widetilde{\mathcal{S}})$ (the so-called excess capital above $Y_0 = V_0$).

We consider the Value-at-Risk risk measure for ϱ. Denote the distribution function of $V_1 - (1+q_0)Y_1$ by F_{q_0}. The Value-at-Risk of $V_1 - (1+q_0)Y_1$ on the security level $1-\alpha \in (0,1)$ is given by

$$\begin{aligned} \varrho\left(V_1 - (1+q_0)Y_1\right) &= \mathrm{VaR}_{1-\alpha}\left(V_1 - (1+q_0)Y_1\right) \\ &= F_{q_0}^{\leftarrow}(1-\alpha) \; = \; \inf\left\{x \middle| \, F_{q_0}(x) \geq 1-\alpha\right\}. \end{aligned} \tag{4.43}$$

Observe that the Value-at-Risk risk measure is normalized and monotone. Choose $q_0 \in \mathbb{R}_+$ such that

$$F_{q_0}(0) = P\left[V_1 - (1+q_0)Y_1 \leq 0\right] \geq 1-\alpha. \tag{4.44}$$

Therefore, for any such $q_0 \in \mathbb{R}_+$ we find

$$\mathrm{VaR}_{1-\alpha}(V_1 - (1+q_0)Y_1) \leq 0, \tag{4.45}$$

thus, for any $q_0 \in \mathbb{R}_+$ that satisfies (4.44) the company is solvent w.r.t. the Value-at-Risk risk measure on security level $1-\alpha$. Therefore, we choose $q_0 \in \mathbb{R}_+$ minimal such that

$$P\left[(1+q_0)\, Y_1 \geq V_1\right] = P\left[(1+q_0)\, \widetilde{Y}_1 \geq 1\right] \; \geq \; 1-\alpha, \tag{4.46}$$

where we have again defined $\widetilde{Y}_1 = Y_1/V_1$ (under the assumption that $V_1 > 0$, P-a.s.). In (4.24) we have studied $\widetilde{Y}_1$ under the equivalent probability measure $P^Y \sim P$, now we need to study $\widetilde{Y}_1$ under the real world probability measure P.

Assume that $\widetilde{Y}_1$ also has a log-normal distribution under the real world probability measure P with parameters μ and σ^2. For σ^2 we choose the same numerical value as in (4.24). Note that under P^Y the mean parameter $\mu^Y = -\sigma^2/2$ was naturally given by the $(P^Y, (\mathcal{F}_0, \mathcal{H}_1))$-martingale property, here under P the mean parameter is not given and needs to be calibrated.

For the log-normal distribution we need to study

$$P\left[\log \widetilde{Y}_1 \geq -\log\left(1+q_0\right)\right] = 1-\alpha. \tag{4.47}$$

Table 4.2 Loading constant q_0 depending on the choice of α, μ and σ compared to the Margrabe option price $M_0 = \mathcal{A}_0(\mathcal{M})$, we initialize $V_0 = 1$

Rating	Default probab. α	Normal quantile $\Phi^{-1}(\alpha)$	$\sigma =$	0.05	0.10	0.15	0.20
			$\mu =$	0.012	0.024	0.036	0.048
AAA	0.01 %	−3.72		19.0 %	41.6 %	68.5 %	100.5 %
AA	0.03 %	−3.43		17.3 %	37.6 %	61.4 %	89.3 %
A	0.07 %	−3.19		15.9 %	34.4 %	55.8 %	80.6 %
BBB	0.18 %	−2.91		14.3 %	30.6 %	49.3 %	70.6 %
BB	1.08 %	−2.30		10.8 %	22.8 %	36.2 %	50.9 %
B	6.41 %	−1.52		6.6 %	13.7 %	21.2 %	29.2 %
B−	11.61 %	−1.19		4.9 %	10.0 %	15.4 %	21.0 %
$M_0 = \mathcal{A}_0(\mathcal{M})$	0 %			2.0 %	4.0 %	6.0 %	8.0 %

Henceforth, q_0 is given by

$$-\log(1 + q_0) = \sigma\, \Phi^{-1}(\alpha) + \mu. \tag{4.48}$$

We choose the following example: choose $\sigma \in \{0.05, 0.10, 0.15, 0.20\}$, see also Example 4.2. For the security level we choose different $\alpha > 0$ according to Standard and Poors ratings and for the drift we choose $\mu = \sigma \cdot 24\,\%$. The results are shown in Table 4.2.

Observe that for the solvency calculation we also need to specify the mean parameter μ in (4.48). In discussions with economists and in the developments of the Swiss Solvency Test it has turned out that it is highly non-trivial to estimate μ for the different asset classes. For example, for the Swiss Solvency Test 2005, even experts have had such different opinions on how to estimate μ that finally the expected investment return has been put equal to the risk-free rate. But Example 4.5 shows that for this risk measure approach it only makes sense to consider the expected return and the volatility simultaneously. Higher expected returns will increase the uncertainty because one needs to invest in more risky assets. This example also shows that the choice of the asset portfolio is more crucial than the choice of the security level α.

Note that the risk measure approach with q_0 is less safe than solvency Definition 4.2 because we allow for default in the former (with a small positive probability). A reason why the risk measure approach is much more expensive is that this approach reflects a buy and hold strategy (passive strategy of defense). On the other hand, as seen in Sect. 4.4, the Margrabe option approach continuously optimizes the asset portfolio in the background, i.e. this is an active strategy of defense against an insolvency.

Finally, we remark that the Margrabe option approach only considers financial risks. The risk measure approach is more general in this context because it would allow us to study any potential default factor in (4.40), for instance, insurance technical risk driven risk factors. □

Example 4.6 (*Expected shortfall*)
We revisit Example 4.5 for the expected shortfall risk measure. Assume that the random variable $Z_1 = V_1 - (1 + q_0)Y_1$ has a continuous distribution function. In this case the expected shortfall risk measure on the security level $1 - \alpha \in (0, 1)$ is given by

$$\varrho\,(Z_1) = \mathbb{E}\left[\, Z_1 \middle|\, Z_1 > \mathrm{VaR}_{1-\alpha}(Z_1)\right]. \tag{4.49}$$

From this we immediately see that there is no simple decoupling as in the Value-at-Risk risk measure case. For solvency testing we need to analyze the sign of $\varrho(Z_1) \leq 0$. This is equivalent to analyzing

$$\mathbb{E}\left[\, V_1 \middle|\, Z_1 > \mathrm{VaR}_{1-\alpha}(Z_1)\right] \leq (1 + q_0)\mathbb{E}\left[\, Y_1 \middle|\, Z_1 > \mathrm{VaR}_{1-\alpha}(Z_1)\right]. \tag{4.50}$$

In general, there is no simple solution to this latter requirement and numerical solutions need to be found. Note that Z_1 also depends on q_0. □

Chapter 5
The Valuation Portfolio in Non-life Insurance

In this chapter we consider the VaPo construction for a non-life insurance run-off portfolio. In essence, the VaPo construction in non-life insurance is similar to the one in life insurance, but we would like to highlight a few difficulties that may occur in a non-life insurance modelling context.

5.1 Introduction to Claims Reserving

To illustrate the non-life insurance claims run-off problem we consider a non-life insurance contract that protects a policyholder against claims that occur within a fixed calender year, see Fig. 5.1. The policyholder exchanges a deterministic premium Π (upfront premium) against a contract that gives him a cover against well-specified random events (claims) occurring within the fixed time period. For an extended introduction to non-life insurance we refer to Wüthrich [Wü13].

Assume that a claim occurs within this fixed time period. In that case the insurance company replaces the financial damage caused by that claim according to the insurance contract terms. In general, the insurance company is not able to assess the total claim amount immediately at claims occurrence due to the following reasons.

1. Usually, there is a *reporting delay* (time gap between the claims accident date and the claims reporting to the insurance company). This time gap can be small (a few days), for example, in motor hull insurance, but it can also be quite large (months or years). Especially, in general liability insurance we can have large reporting delays, typical examples are asbestos claims that were caused several years ago but the outbreak of the disease is only noticed and reported today.
2. Usually, it takes quite some time to settle a claim (time difference between reporting date and settlement date). This *settlement delay* is due to several different reasons, for example, for bodily injury claims we first have to monitor the recovery process before finally deciding on the claim and on the compensation, or other claims can only be settled at court which usually takes quite some time until the

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1_5

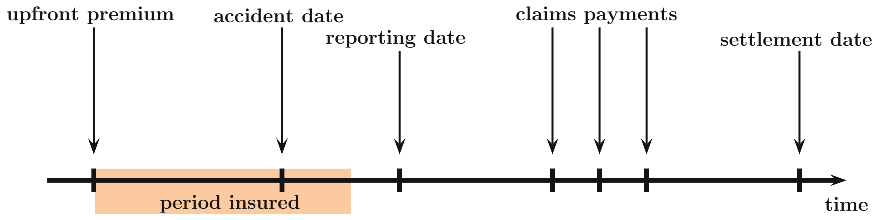


Fig. 5.1 Claims development process in non-life insurance

final settlement takes place. In most cases a (more complex) claim is settled by several claims payments: whenever a justified bill for that specific claim arrives it is paid by the insurance company.

Assume that a contract (or a portfolio of contracts) generates a cash flow

$$\mathbf{X} = (X_0, \dots, X_N), \tag{5.1}$$

where X_k denotes the payments at time $k \in \mathbb{N}$ (we set $X_k = 0$ if there is no payment at time k) and N is the (random) settlement delay, i.e. the last payment takes place at time N.

Remarks.

- In general, non-life insurance payments are done continuously over time. For modelling purposes however, we choose a yearly time grid $k = 0, 1, 2, \dots,$ and we map all payments done within accounting year $(k, k + 1]$ to its right endpoint $k+1$, that is, X_{k+1} will denote all payments made within accounting year $(k, k+1]$.
- The settlement date is random for non-life insurance claims and therefore the time point N of the last payment is also random. In our case, we assume that $n \in \mathbb{N}$ is sufficiently large such that $N \le n$, a.s.

With these remarks we set

$$X_0 = -\Pi, \qquad \text{upfront premium}, \tag{5.2}$$

$$X_k, \quad k \in \{1, \dots, n\}, \text{ nominal claims payments in period } (k-1, k]. \tag{5.3}$$

Denote the *nominal cumulative claims payments* until time $k \in \{1, \dots, n\}$ by

$$C_k = \sum_{t=1}^{k} X_t, \tag{5.4}$$

and the *nominal ultimate claim* is given by $C_N = C_n$.

We choose a filtered probability space $(\Omega, \mathcal{F}, P, \mathbb{F})$ and assume that the insurance liability cash flow satisfies $\mathbf{X} \in L^2_{n+1}(P, \mathbb{F})$.

The approach used in practice.

(1) Predict the nominal ultimate claim C_n for given information $\mathcal{F}_k$ at time $k < n$. This prediction problem is known as the *claims reserving problem*. This is not discussed here in great detail, but we refer to the vast literature on the claims reserving problem. For a reference, see Wüthrich–Merz [WM08, WM15].

(2) Split the nominal ultimate claim C_n into the different (annual) payments $X_1, \ldots, X_n$, i.e. predict a cash flow pattern for C_n, so that one can value the resulting insurance cash flows $X_t = C_t - C_{t-1}$.

This approach can be considered as a top-down approach and many claims reserving methods are based on it (because they are quite robust). A bottom-up approach would directly predict the single future cash flows X_t, $t \in \{k+1, \ldots, n\}$, given $\mathcal{F}_k$, which can then be aggregated to the nominal ultimate claim C_n.

A rather difficult task in non-life insurance is the modelling of the flow of information $\mathbb{F}$. In general, $\mathbb{F}$ should capture any relevant information. This often turns out to be too complex in statistical modelling. Therefore, one restricts the available information so that one obtains tractable statistical models. In these notes we make our lives rather simple in the sense that we assume that the insurance technical filtration $\mathbb{T}$ is generated by the claims payments $\mathbf{X}$, after a possible inflation adjustment.

Under Assumption 2.16, this then immediately implies the following for the VaPo construction described in Chap. 3:

Step 1. The appropriate basis of financial instruments $\mathcal{U}_1, \mathcal{U}_2, \ldots,$ is composed of either

- the zero coupon bonds $\mathcal{Z}^{(t)}$, $t \in \{0, \ldots, n\}$, or
- inflation protected zero coupon bonds,

depending on whether we consider nominal payments $\mathbf{X}$ or an inflation adjusted version thereof.

Step 2. Determine the number of units $\Lambda_i(\mathbf{X}_k)$, $l_{i,k}^{(t)}$ and $l_{i,k}^{*,t}$, respectively, of the financial instruments $\mathcal{U}_i$ we need to purchase in order to meet all our future obligations (which are covered by past premium), see Sect. 3.6.

In this chapter we concentrate on the prediction problem of $\mathbf{X}$, conditionally given $\mathcal{F}_t$. We will give special attention to the changes in these predictions when information increases from $\mathcal{F}_t$ to $\mathcal{F}_{t+1}$.

Assumption 5.1 We set Assumption 2.16 with $\varphi^{\mathbb{T}}$ being a density process w.r.t. $\mathbb{T}$ and P. We assume that the appropriate financial basis for the cash flow $\mathbf{X}$ under consideration is given by the zero coupon bonds $\mathcal{Z}^{(t)}$, $t = 0, \ldots, n$, that is, $\mathbf{X} \in L^2_{n+1}(P, \mathbb{T})$.

Assumption 5.1 implies that we can work on a product probability space with

1. the financial market model $L^2_{n+1}(P, \mathbb{G})$ modelling the price processes of the zero coupon bonds $\mathcal{Z}^{(t)}$, $t = 0, \ldots, n$;
2. the insurance technical parts being modelled by $\boldsymbol{\Lambda} = \mathbf{X} \in L^2_{n+1}(P, \mathbb{T})$;
3. $\mathbb{T}$ and $\mathbb{G}$ being independent w.r.t. P generating $\mathbb{F}$ according to (2.95) and the chosen deflator $\boldsymbol{\varphi}$ being assumed to split into two independent parts, the financial deflator $\boldsymbol{\varphi}^{\mathbb{G}} \in L^2_{n+1}(P, \mathbb{G})$ and the probability distortion $\boldsymbol{\varphi}^{\mathbb{T}} \in L^2_{n+1}(P, \mathbb{T})$, the latter being a density process w.r.t. $\mathbb{T}$ and P.

Thus, the zero coupon bond $\mathcal{Z}^{(k)}$ is the right financial instrument under Assumption 5.1 for replicating the cash flow X_k and we may identify

$$\mathbf{X} = \boldsymbol{\Lambda} = (\Lambda_0, \ldots, \Lambda_n) \in L^2_{n+1}(P, \mathbb{T}). \tag{5.5}$$

<u>Valuation Scheme</u>: Cash flow and instrument representation after time $t < n$

Time	Unit $\mathcal{U}_i$	Cash flow		Number of units
$t+1$	$\mathcal{Z}^{(t+1)}$	X_{t+1}	$\longrightarrow$	$\Lambda_{t+1} = X_{t+1}$
$t+2$	$\mathcal{Z}^{(t+2)}$	X_{t+2}	$\longrightarrow$	$\Lambda_{t+2} = X_{t+2}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$t+k$	$\mathcal{Z}^{(t+k)}$	X_{t+k}	$\longrightarrow$	$\Lambda_{t+k} = X_{t+k}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
n	$\mathcal{Z}^{(n)}$	X_n	$\longrightarrow$	$\Lambda_n = X_n$

In this particular case we do not distinguish between Valuation Scheme A (cash flow representation) and Valuation Scheme B (instrument representation) because they coincide, see also page 53.

Finally, we replace the numbers of units by $\mathcal{F}_t$-measurable variables, see Sect. 3.6. For the VaPo construction we choose best-estimates for $k \geq 1$ (we may drop one index in this case because there is only one instrument per cash flow)

$$l^{(t)}_{t+k} = E\left[\Lambda_{t+k}(\mathbf{X}_{t+k})|\mathcal{T}_t\right] = E\left[X_{t+k}|\mathcal{T}_t\right], \tag{5.6}$$

and similarly for the VaPo protected against insurance technical risks we have

$$l^{*,t}_{t+k} = \frac{1}{\varphi^{\mathbb{T}}_t} E\left[\varphi^{\mathbb{T}}_{t+k} X_{t+k}\middle|\mathcal{T}_t\right]. \tag{5.7}$$

This provides the VaPos for the outstanding liabilities at time t

$$\mathrm{VaPo}_t\left(\mathbf{X}_{(t+1)}\right) = \sum_{s=t+1}^{n} l_s^{(t)}\, \mathcal{Z}^{(s)} = \sum_{s=t+1}^{n} E\left[X_s \middle| \mathcal{T}_t\right] \mathcal{Z}^{(s)}, \tag{5.8}$$

$$\mathrm{VaPo}_t^{prot}\left(\mathbf{X}_{(t+1)}\right) = \sum_{s=t+1}^{n} l_s^{*,t}\, \mathcal{Z}^{(s)} = \sum_{s=t+1}^{n} \frac{1}{\varphi_t^{\mathbb{T}}} E\left[\varphi_s^{\mathbb{T}} X_s \middle| \mathcal{T}_t\right] \mathcal{Z}^{(s)}. \tag{5.9}$$

Remark 5.2

- The zero coupon bonds $\mathcal{Z}^{(s)}$ are chosen as financial instruments to represent the cash flows X_s, $s \geq t+1$. The choice of the financial instruments was rather obvious in the life insurance example. In non-life insurance this is one of the crucial and non-trivial steps: find a decoupling such that the price processes of the units $\mathcal{U}_i$ and the number of units Λ_i are independent. In particular, find the (financial) instruments that replicate claims inflation in an appropriate way.
- The mapping $X_{t+k} \mapsto l_{t+k}^{(t)}$ incorporates the actual (latest) information $\mathcal{T}_t$ considered. In this chapter we put special emphasis on changes in these predictions as information increases to $\mathcal{T}_{t+1}$. This will lead to the notion of the claims development result (CDR). This was already briefly mentioned in (3.65)–(3.66).
- The financial risk is treated in exactly the same way as in Chap. 4.

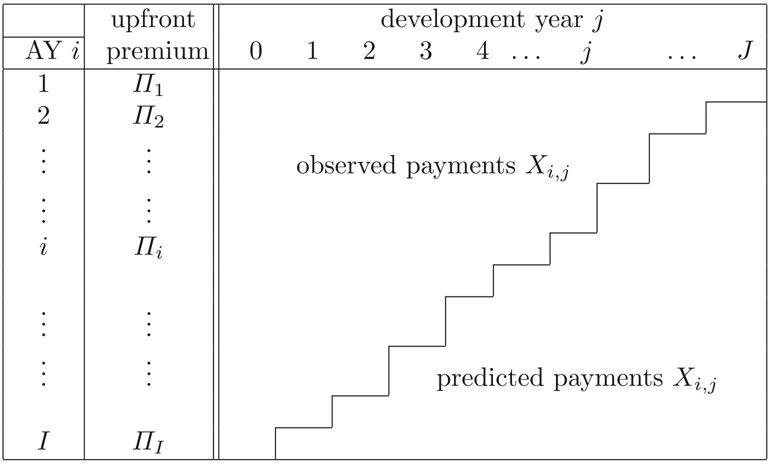
5.2 Construction of the VaPo in Non-life Insurance

In this section we derive more explicitly the construction of the VaPo under Assumption 5.1. This can be done by concentrating on the modelling of the individual cash flows $X_k = \Lambda_k$, $k = 1, \ldots, n$. We introduce the gamma-gamma Bayesian chain-ladder (BCL) model to predict and model these cash flows. For an extended treatment of the gamma-gamma BCL model we refer to Wüthrich–Merz [WM15].

5.2.1 Loss Development Triangles

Usually, in non-life insurance, data is pooled so that one obtains homogeneous subgroups. In claims reserving one typically builds different sub-portfolios according to different lines of business, claims types, etc. These sub-portfolios are then further structured by a time component like the accident year. This is motivated by the fact that claims that occur in the same period are affected by the same external factors (like weather conditions, natural hazards, etc.).

Then, claims data is typically structured in a triangular form with the vertical axis labelling accident years $i \in \{1, \ldots, I\}$ (year of claims occurrence), and the horizontal axis labelling development years $j \in \{0, \ldots, J\}$ (settlement/payment delay), see Wüthrich–Merz [WM08, WM15] for more details. I denotes the last accident (exposure) year considered and we assume that all claims are settled after J development years.



Let $X_{i,j}$ denote the (incremental) payments for accident year $i \in \{1, \dots, I\}$ in development year $j \in \{0, \dots, J\}$. These are the payments for accident year i made in calendar year $k = i + j$. We initialize $X_{i,-1} = -\Pi_i$ for the upfront premium received for accident year i. Nominal *cumulative claims payments* for accident year i within the first j development years are given by

$$C_{i,j} = \sum_{s=0}^{j} X_{i,s}. \tag{5.10}$$

The payments made in a fixed accounting year k are given by

$$X_k = \sum_{i+j=k} X_{i,j}, \tag{5.11}$$

these are the diagonals of our loss development squares. This implies (if we neglect the premium payments $X_{i,-1} = -\Pi_i$) for accounting year k

$$X_k = \sum_{i=1 \vee (k-J)}^{I \wedge k} X_{i,k-i}. \tag{5.12}$$

Example 5.1 (*Non-life insurance loss development triangles*)
We use the Taylor–Ashe [TA83] data (neglecting premium payments $X_{i,-1} = \Pi_i$, $1 \le i \le I$, because they will not influence our modelling after time $t = I$).

Observed (incremental) payments $\{X_{i,j};\ 1 \le i+j \le I,\ 1 \le i \le I,\ 0 \le j \le J\}$ at time I

	0	1	2	3	4	5	6	7	8	9
1	357,848	766,940	610,542	482,940	527,326	574,398	146,342	139,950	227,229	67,948
2	352,118	884,021	933,894	1,183,289	445,745	320,996	527,804	266,172	425,046	
3	290,507	1,001,799	926,219	1,016,654	750,816	146,923	495,992	280,405		
4	310,608	1,108,250	776,189	1,562,400	272,482	352,053	206,286			
5	443,160	693,190	991,983	769,488	504,851	470,639				
6	396,132	937,085	847,498	805,037	705,960					
7	440,832	847,631	1,131,398	1,063,269						
8	359,480	1,061,648	1,443,370							
9	376,686	986,608								
10	344,014									

Observed cumulative payments $\{C_{i,j};\ 1 \le i+j \le I,\ 1 \le i \le I,\ 0 \le j \le J\}$ at time I

	0	1	2	3	4	5	6	7	8	9
1	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
2	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
3	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
4	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
5	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
6	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
7	440,832	1,288,463	2,419,861	3,483,130						
8	359,480	1,421,128	2,864,498							
9	376,686	1,363,294								
10	344,014									

Observed accounting year payments $\{X_k;\ 1 \le k \le I\}$ at time I

1	2	3	4	5	6	7	8	9	10
357,848	1,119,058	1,785,070	2,729,241	4,188,244	3,902,308	5,150,454	3,911,256	5,221,066	5,993,545

The run-off situation.
We fix the last accident (exposure) year I and consider the *run-off situation* after time I, that is, we predict and value the outstanding liabilities $\mathbf{X}_{(t)}$ for $t = I+1, \ldots, I+J$ without adding new business (new exposure) to the portfolio after time I.
The relevant data (observations) at time t are then given by

$$\mathcal{D}_t = \left\{X_{i,j};\ 1 \le i+j \le t,\ 1 \le i \le I,\ 0 \le j \le J\right\}, \tag{5.13}$$

and, respectively, by

$$\mathcal{D}_t = \left\{ C_{i,j};\ 1 \le i+j \le t,\ 1 \le i \le I,\ 0 \le j \le J \right\}. \tag{5.14}$$

Note that because the payments $X_{i,j}$ and the cumulative payments $C_{i,j}$ generate the same σ-field $\sigma(\mathcal{D}_t)$, we have adopted a slight abuse of notation here, using $\mathcal{D}_t$ for both cases above.

5.2.2 *The Gamma-Gamma Bayesian Chain-Ladder Model*

Probably the most popular method of predicting future claims payments in non-life insurance is the so-called chain-ladder (CL) method. Originally, the CL method was not based on a stochastic model but it was introduced as a (deterministic) algorithm to calculate the claims reserves. Only much later have several stochastic models been proposed that provide the CL reserves as claims predictors. Here, we use the gamma-gamma Bayesian chain-ladder (BCL) model for the modelling of the nominal cumulative payments $C_{i,j}$ for $1 \le i \le I$ and $0 \le j \le J$. This model provides a stochastic representation of the CL method (for non-informative priors).

Since we will only consider the run-off situation after time I we will completely neglect the premium payments Π_i modelling in the following assumptions. This simplifies some of the technicalities.

Assumption 5.3 (*Gamma-gamma BCL model*) Assume $I > J$ and set $n = I + J$;

- the σ-fields $\mathcal{T}_t$ are generated by $\mathcal{D}_t$, given in (5.14), for $0 \le t \le n$;
- there are given fixed constants $\sigma_j > 0$ for $0 \le j \le J-1$; and

(a) conditionally, given vector $\boldsymbol{\Theta} = (\Theta_0, \ldots, \Theta_{J-1})$, $(C_{i,j})_{j=0,\ldots,J}$ are independent (in i) Markov processes (in j) with conditional distributions

$$C_{i,j+1} \big|_{\mathcal{T}_{i+j}, \boldsymbol{\Theta}} \sim \Gamma\left(C_{i,j}\sigma_j^{-2}, \Theta_j \sigma_j^{-2} \right),$$

for $1 \le i \le I$ and $0 \le j \le J-1$;

(b) the components Θ_j of $\boldsymbol{\Theta}$ are mutually independent and $\Gamma(\gamma_j, f_j(\gamma_j - 1))$-distributed with given prior parameters $f_j > 0$ and $\gamma_j > 1$ for $0 \le j \le J-1$;

(c) $\boldsymbol{\Theta}$ and $(C_{i,0})_{1 \le i \le I}$ are independent and $C_{i,0} > 0$, P-a.s., for all $1 \le i \le I$.

Remarks.

- There is a vast literature on the CL method. Besides the gamma-gamma BCL model, the most popular stochastic representations of the CL method are Mack's distribution-free CL model [Ma93] and England–Verrall's over-dispersed Poisson

(ODP) model [EV02]. We use a Bayesian representation here because then parameter uncertainty is naturally integrated. This will facilitate several considerations.
- Since we only consider the run-off situation after time I, we neglect the explicit modelling of the premium payment Π_i and we assume that the knowledge of $C_{i,0}$ is sufficient to model all future payments $C_{i,j}, j \geq 1$.

Lemma 5.4 *Under Assumption 5.3 we have for $t \geq I \geq i > t-j$*

$$\widehat{C}_{i,j}^{BCL(t)} = E\left[\, C_{i,j} \middle|\, \mathcal{T}_t \right] = C_{i,t-i} \prod_{\ell=t-i}^{j-1} \widehat{f}_{\ell}^{BCL(t)}, \tag{5.15}$$

with BCL factors $\widehat{f}_{\ell}^{BCL(t)}$ for $0 \leq \ell \leq J-1$ given by

$$\widehat{f}_{\ell}^{BCL(t)} = E\left[\, \Theta_{\ell}^{-1} \middle|\, \mathcal{T}_t \right] = \omega_{\ell}^{(t)} \widehat{f}_{\ell}^{CL(t)} + (1-\omega_{\ell}^{(t)}) f_{\ell}, \tag{5.16}$$

and credibility weights and CL factors, respectively,

$$\omega_{\ell}^{(t)} = \frac{\sum_{k=1}^{(t-\ell-1)\wedge I} C_{k,\ell}}{\sum_{k=1}^{(t-\ell-1)\wedge I} C_{k,\ell} + \sigma_{\ell}^2\,(\gamma_{\ell}-1)} \;\in\; (0,1), \tag{5.17}$$

$$\widehat{f}_{\ell}^{CL(t)} = \frac{\sum_{k=1}^{(t-\ell-1)\wedge I} C_{k,\ell+1}}{\sum_{k=1}^{(t-\ell-1)\wedge I} C_{k,\ell}}. \tag{5.18}$$

Proof The proof applies Bayes' theorem to find the posterior distributions of Θ_{ℓ}, $0 \leq \ell \leq J-1$, given $\mathcal{T}_t$. Since we have a Bayesian model with conjugate priors, these posterior distributions turn out to be independent and gamma distributed with changed shape and scale parameters. These posterior distributions are then used for the calculation of the posterior means of $C_{i,j}$, this essentially uses the tower property of conditional expectations, conditional independence of observations in different accident years and their conditional Markov property. For details, see Theorem 9.5 in Wüthrich [Wü13]. □

The non-informative prior limit of Lemma 5.4 is obtained by considering the limit $\gamma = (\gamma_0, \ldots, \gamma_{J-1}) \to \mathbf{1}$ which provides the classical CL predictors at time $t \geq I$

$$\widehat{C}_{i,j}^{CL(t)} = \lim_{\gamma \to \mathbf{1}} \widehat{C}_{i,j}^{BCL(t)} = C_{i,t-i} \prod_{\ell=t-i}^{j-1} \widehat{f}_{\ell}^{CL(t)}. \tag{5.19}$$

This exactly justifies the use of the non-informative prior limit gamma-gamma BCL model for analyzing the prediction uncertainty of the CL method. The (incremental) claims payments $X_{i,j} = C_{i,j} - C_{i,j-1}$, for $i+j > t \geq I$, are in the non-informative prior limit gamma-gamma BCL model predicted by

$$\widehat{X}_{i,j}^{CL(t)} = \lim_{\gamma \to \mathbf{1}} E\left[X_{i,j} \middle| \mathcal{T}_t\right] = C_{i,t-i} \prod_{\ell=t-i}^{j-2} \widehat{f}_\ell^{\,CL(t)} \left(\widehat{f}_{j-1}^{\,CL(t)} - 1\right), \tag{5.20}$$

where an empty product is set equal to 1.

5.2.3 The VaPo Construction with CL in the Run-Off Situation

Formula (5.20) gives the classical CL predictors for the future claims payments. We use these predictors for the construction of the VaPo at time $t \geq I > J$ for the study of the run-off of the outstanding liabilities $\mathbf{X}_{(t+1)}$ given by

$$\begin{aligned} \mathbf{X}_{(t+1)} &= \left(0, \ldots, 0, \sum_{i+j=t+1} X_{i,j}, \ldots, \sum_{i+j=I+J} X_{i,j}\right) \\ &= \left(0, \ldots, 0, \sum_{i=t+1-J}^{I} X_{i,t+1-i}, \ldots, \sum_{i=I}^{I} X_{i,I+J-i}\right). \end{aligned} \tag{5.21}$$

Using (5.20) we predict these future cash flows in the non-informative prior limit gamma-gamma BCL model of Assumption 5.3 at time $t \geq I$ for $k \geq 1$ by

$$l_{t+k}^{(t)} \stackrel{\text{def.}}{=} \lim_{\gamma \to \mathbf{1}} E\left[X_{t+k} \middle| \mathcal{T}_t\right] = \sum_{i=t+k-J}^{I} C_{i,t-i} \prod_{\ell=t-i}^{t+k-i-2} \widehat{f}_\ell^{\,CL(t)} \left(\widehat{f}_{t+k-i-1}^{\,CL(t)} - 1\right). \tag{5.22}$$

Under Assumptions 5.1 and 5.3 we obtain VaPo at time $I \leq t < I + J$

$$\mathrm{VaPo}_t(\mathbf{X}_{(t+1)}) = \sum_{s=t+1}^{I+J} \left[\sum_{i=s-J}^{I} C_{i,t-i} \prod_{\ell=t-i}^{s-i-2} \widehat{f}_\ell^{\,CL(t)} \left(\widehat{f}_{s-i-1}^{\,CL(t)} - 1\right)\right] \mathcal{Z}^{(s)}. \tag{5.23}$$

This corresponds to the $\mathcal{T}_t$-measurable VaPo of the outstanding liabilities $\mathbf{X}_{(t+1)}$ represented in terms of the zero coupon bonds $\mathcal{Z}^{(s)}$, $s \geq t+1$.

Example 5.1 (*revisited*).
We revisit the data given on page 97 and construct the VaPo at time $t = I$.

The lower triangle is completed using predictions $\widehat{C}_{i,j}^{CL(I)}$ for nominal cumulative payments based on $\mathcal{T}_I$:

	0	1	2	3	4	5	6	7	8	9
1	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
2	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	5,433,719
3	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315	5,285,148	5,378,826
4	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268	4,835,458	5,205,637	5,297,906
5	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311	4,207,459	4,434,133	4,773,589	4,858,200
6	396,132	1,333,217	2,180,715	2,985,752	3,691,712	4,074,999	4,426,546	4,665,023	5,022,155	5,111,171
7	440,832	1,288,463	2,419,861	3,483,130	4,088,678	4,513,179	4,902,528	5,166,649	5,562,182	5,660,771
8	359,480	1,421,128	2,864,498	4,174,756	4,900,545	5,409,337	5,875,997	6,192,562	6,666,635	6,784,799
9	376,686	1,363,294	2,382,128	3,471,744	4,075,313	4,498,426	4,886,502	5,149,760	5,544,000	5,642,266
10	344,014	1,200,818	2,098,228	3,057,984	3,589,620	3,962,307	4,304,132	4,536,015	4,883,270	4,969,825
$\widehat{f}_j^{CL(I)}$	3.4906	1.7473	1.4574	1.1739	1.1038	1.0863	1.0539	1.0766	1.0177	

This provides predicted claims payments $\widehat{X}_{i,j}^{CL(I)}$ for $i + j > I$, see (5.20):

	0	1	2	3	4	5	6	7	8	9
1										
2										94,634
3									375,833	93,678
4								247,190	370,179	92,268
5							334,148	226,674	339,456	84,611
6						383,287	351,548	238,477	357,132	89,016
7					605,548	424,501	389,349	264,121	395,534	98,588
8				1,310,258	725,788	508,792	466,660	316,566	474,073	118,164
9			1,018,834	1,089,616	603,569	423,113	388,076	263,257	394,241	98,266
10		856,804	897,410	959,756	531,636	372,687	341,826	231,882	347,255	86,555

From these predictions we construct the VaPo, see (5.23), for the outstanding liabilities $\mathbf{X}_{(I+1)}$ at time $t = I$. This gives (see Table 5.1)

$$\text{VaPo}_I(\mathbf{X}_{(I+1)}) = \sum_{s=I+1}^{I+J} l_s^{(I)} \, \mathcal{Z}^{(s)}. \tag{5.24}$$

Applying an accounting principle $\mathcal{A}_I$ to the zero coupon bonds $\mathcal{Z}^{(s)}$ we receive the best-estimate reserves at time I, see also Example 2.7. These are given by

$$R_I^{(I+1)} = \sum_{s=I+1}^{I+J} l_s^{(I)} \, \mathcal{A}_I(\mathcal{Z}^{(s)}) = \sum_{s=I+1}^{I+J} l_s^{(I)} \, P(I, s). \tag{5.25}$$

There remains an explicit choice of the zero coupon bond prices $P(I, s)$ at time I. These may be illustrated by the yield rates $(Y(I, s))_{s>I}$ at time I, see also (3.69),

$$P(I, s) = \exp\left\{-(s - I) \, Y(I, s)\right\}. \tag{5.26}$$

We choose three different examples: (1) nominal values which correspond to the choices $Y(I, s) \equiv 0$; (2) constant yield rates $Y(I, s) \equiv 1.5\,\%$; and (3) risk-free yield rates as used in the Swiss Solvency Field-Test 2005: see Table 5.2.

Table 5.1 Predicted cash flows $l_{I+k}^{(I)}$, for $k = 1, \ldots, J$, based on information $\mathcal{T}_I$, see (5.22); figures are shown in 1,000

$I+k$	$I+1$	$I+2$	$I+3$	$I+4$	$I+5$	$I+6$	$I+7$	$I+8$	$I+9$
$l_{I+k}^{(I)}$	5,227	4,179	3,132	2,127	1,562	1,178	744	446	87

Table 5.2 Risk-free yield rates $Y(I, s)$ as a function of time to maturity $s - I$

$s-I$	1	2	3	4	5	6	7	8	9
$Y(I,s)$	0.88 %	1.14 %	1.36 %	1.57 %	1.75 %	1.91 %	2.05 %	2.18 %	2.29 %

Table 5.3 Best-estimate reserves for the outstanding liabilities at time I using the three different accounting principles (1) nominal values; (2) constant yield rates of 1.5 %; (3) variable yield rates given in Table 5.2

	Best-estimate reserves $R_I^{(I+1)}$	Difference to nominal in %	
(1) Nominal values	18,680,856		
(2) Constant yield rates of 1.50 %	17,868,119	812,737	4.35 %
(3) Variable yield rates	17,840,966	839,889	4.50 %

This gives the best-estimate reserves presented in Table 5.3. We observe that discounting with non-zero yield rates leads to a substantial reduction in the best-estimate reserves of size 4 % to 5 %. This insight is particularly useful in view of statutorial accounting which considers nominal values. □

This closes the section on the construction of the VaPo. Next we discuss the VaPo protected against insurance technical risks construction in the gamma-gamma BCL model.

5.3 The Protected VaPo for a Non-life Insurance Run-Off

To construct the VaPo protected against insurance technical risks we need to choose an appropriate probability distortion $\varphi^{\mathbb{T}} \in L^2_{n+1}(P, \mathbb{T})$. This probability distortion is used to calculate the distorted expected claims payments

$$l_{t+k}^{*,t} = \frac{1}{\varphi_t^{\mathbb{T}}} E\left[\varphi_{t+k}^{\mathbb{T}} \sum_{i+j=t+k} X_{i,j} \middle| \mathcal{T}_t \right]. \tag{5.27}$$

Throughout we make Assumptions 5.1 and 5.3. A necessary requirement for obtaining a positive margin for insurance technical risks is that $\varphi_{t+k}^{\mathbb{T}}$ and $\sum_{i+j=t+k} X_{i,j}$ are positively correlated w.r.t. $P[\cdot|\mathcal{T}_t]$. This implies

$$l_{t+k}^{*,t} \geq \frac{1}{\varphi_t^{\mathbb{T}}} E\left[\varphi_{t+k}^{\mathbb{T}} \middle| \mathcal{T}_t \right] E\left[\sum_{i+j=t+k} X_{i,j} \middle| \mathcal{T}_t \right] = l_{t+k}^{(t)}, \tag{5.28}$$

where we have used the $(P, \mathbb{T})$-martingale property of the probability distortion $\varphi^{\mathbb{T}}$. This specifies reasonable choices of the probability distortion $\varphi^{\mathbb{T}}$. For more on this we refer to Sect. 5.3.7.

There is no special difficulty in applying this concept of probability distortions, but we refrain from doing so. The main reason why other methods are typically used is that the probability distortion approach is difficult to calibrate and it often lacks a good interpretation. We refer to Wüthrich et al. [WET11] where an explicit calculation and calibration of a probability distortion is presented within a Bayesian multivariate log-normal model. The concept we are going to present in the following sections may be ad-hoc, but it is closer to industry practice and it allows us to give a more explicit interpretation. This concept will be based on the claims development result (CDR) and on the cost-of-capital method.

5.3.1 The Claims Development Result

We derive a loading for insurance technical run-off risks that is based on the cost-of-capital method. The cost-of-capital method quantifies the costs of holding target capital against downside risks. The target capital will be determined by means of a risk measure. Before we are able to evaluate a risk measure we need to describe the position at risk. In non-life insurance this is done by considering the sequence of the CDRs of the run-off risks. The sequence of the CDRs describes the dynamic behaviour of the claims reserving uncertainty. This dynamic behaviour has been studied in the course of developing risk-adjusted solvency requirements for non-life insurance. This was done for the CL method simultaneously by several actuaries who were all studying the same problem, we refer to De Felice–Moriconi [dFM03, dFM06], Böhm–Glaab [BG06], England [En08], Merz–Wüthrich [MW07, MW08] and Wüthrich et al. [WML09]. The picture has recently been completed by Röhr [Rö16], Merz–Wüthrich [MW14] and Diers et al. [DLH16].

The CDR was originally introduced for nominal claims reserves, see Merz–Wüthrich [MW08]. We introduce it here first in terms of VaPos, before we are going to study it for nominal claims reserves. Under Assumption 5.1, the VaPo of the outstanding liabilities at time t is given by

$$\mathrm{VaPo}_t\left(\mathbf{X}_{(t+1)}\right) = \sum_{s=t+1}^{n} E\left[X_s \middle| \mathcal{T}_t\right] \mathcal{Z}^{(s)}. \tag{5.29}$$

We now consider the time series of these VaPos when we increase information $\mathcal{T}_t$. The CDR at time $t+1$ is then in terms of a portfolio defined by

$$\begin{aligned}\mathcal{CDR}_{t+1}(\mathbf{X}) &= \mathrm{VaPo}_t\left(\mathbf{X}_{(t+1)}\right) - \mathrm{VaPo}_{t+1}\left(\mathbf{X}_{(t+1)}\right) \qquad (5.30)\\ &= \sum_{s=t+1}^{n} E\left[X_s \middle| \mathcal{T}_t\right] \mathcal{Z}^{(s)} - \sum_{s=t+1}^{n} E\left[X_s \middle| \mathcal{T}_{t+1}\right] \mathcal{Z}^{(s)},\end{aligned}$$

or, equivalently, by

$$\begin{aligned} \mathcal{CDR}_{t+1}(\mathbf{X}) = & \left(E\left[X_{t+1}|\,\mathcal{T}_t\right] - X_{t+1}\right)\mathcal{Z}^{(t+1)} \\ & + \sum_{s=t+2}^{n}\left(E\left[X_s|\,\mathcal{T}_t\right] - E\left[X_s|\,\mathcal{T}_{t+1}\right]\right)\mathcal{Z}^{(s)}. \end{aligned} \tag{5.31}$$

This describes the portfolio at risk in accounting year $t+1$ induced by increasing the information, see also Sect. 3.5.1. It shows which assets $\mathcal{Z}^{(s)}$, $s \geq t+1$, need to be sold and purchased, respectively, to convert the VaPo at time t into the VaPo at time $t+1$. The aim here is to study the financial impact of these fluctuations. A risk measure evaluated on these fluctuations determines the necessary risk capacity (target capital) in accounting year $t+1$ against the potential shortfalls in the CDR. The financial impact at time $t+1$ of the CDR is given by

$$\begin{aligned} \mathcal{A}_{t+1}\left(\mathcal{CDR}_{t+1}(\mathbf{X})\right) = & \, E\left[X_{t+1}|\,\mathcal{T}_t\right] - X_{t+1} \\ & + \sum_{s=t+2}^{n}\left(E\left[X_s|\,\mathcal{T}_t\right] - E\left[X_s|\,\mathcal{T}_{t+1}\right]\right)P(t+1,s), \end{aligned} \tag{5.32}$$

where $P(t+1,s)$ is the price of the zero coupon bond $\mathcal{Z}^{(s)}$ at time $t+1$ under the accounting principle $\mathcal{A}_{t+1}$.

Lemma 5.5 *Set Assumption 5.1 and assume that the zero coupon bond price processes are $\mathbb{G}$-adapted. The expected CDR viewed from time t is given by*

$$E\left[\mathcal{A}_{t+1}\left(\mathcal{CDR}_{t+1}(\mathbf{X})\right)|\,\mathcal{T}_t, \mathcal{G}_{t+1}\right] = 0. \tag{5.33}$$

Proof The zero coupon bond prices $P(t+1,s)$, $s \geq t+1$, are $\mathcal{G}_{t+1}$-measurable, but then the claim is a direct consequence of the tower property of conditional expectations. □

Remarks.

- Lemma 5.5 reveals that best-estimate reserving provides neither conditionally expected losses nor conditionally expected gains, but reserving is done in an unbiased way. Thus, at time t we predict the future CDRs to be 0 (either as portfolios or in money value). The probability distortion $\varphi^{\mathbb{T}}_{t+1}$ for accounting year $t+1$ should protect against potential shortfalls in this CDR prediction.
- Note that $\mathcal{H}_{t+1} = \sigma(\mathcal{T}_t, \mathcal{G}_{t+1})$ describes the hedgeable filtration that we have already met in Sect. 4.3.2. It shows that financial risks can be hedged using the right asset allocation and the insurance technical part can only be predicted in an optimal (unbiased) way.

5.3.2 The Nominal Claims Development Result

In order to keep this text simple we restrict ourselves to nominal payments, meaning that we set $P(t,m) = 1$ for all $0 \le t \le m \le n$. Therefore, under Assumption 5.1 we can restrict ourselves to the filtered probability space $(\Omega, \mathcal{T}_n, P, \mathbb{T})$. The nominal CDR at time $t+1$ of cash flow $\mathbf{X} \in L^2_{n+1}(P, \mathbb{T})$ is then given by

$$\mathrm{CDR}_{t+1}(\mathbf{X}) = \sum_{s=t+1}^{n} E\left[X_s \middle| \mathcal{T}_t\right] - E\left[X_s \middle| \mathcal{T}_{t+1}\right]. \tag{5.34}$$

The nominal ultimate claim of cash flow $\mathbf{X}$ is given by

$$C = \sum_{s=1}^{n} X_s. \tag{5.35}$$

This allows us to rewrite the nominal CDR in accounting year $t+1$ as follows

$$\mathrm{CDR}_{t+1}(\mathbf{X}) = E\left[C \middle| \mathcal{T}_t\right] - E\left[C \middle| \mathcal{T}_{t+1}\right]. \tag{5.36}$$

The nominal CDRs exactly describe the successive (best-estimate) innovations under increasing information. We define the sequence $(\widehat{C}^{(t)})_{t=0,\ldots,n}$ by

$$\widehat{C}^{(t)} = E\left[C \middle| \mathcal{T}_t\right]. \tag{5.37}$$

A simple consequence is that this sequence $(\widehat{C}^{(t)})_{t=0,\ldots,n}$ is a $(P, \mathbb{T})$-martingale and (5.36) simply describes the innovation process of this martingale.

5.3.3 The Conditional Mean Square Error of Prediction

In this section we choose as (conditional) risk measure $\varrho^{(t)}(\cdot) = \varrho(\cdot|\mathcal{T}_t)$ the conditional mean square error of prediction (MSEP). Note that this (slightly) differs from Sect. 4.5 because, here, our risk measure is $\mathcal{T}_t$-measurable as it should quantify potential losses seen from time t. It is defined by

$$\mathrm{msep}_{C|\mathcal{T}_t}\left(\widehat{C}^{(t)}\right) = E\left[\left(C - \widehat{C}^{(t)}\right)^2 \middle| \mathcal{T}_t\right] = \mathrm{Var}\left(C \middle| \mathcal{T}_t\right), \tag{5.38}$$

where for the second identity we explicitly use (best-estimate) prediction (5.37). The conditional MSEP (5.38) measures the conditional L^2-distance between the true (random) claim C and its best predictor $\widehat{C}^{(t)} = E[C|\mathcal{T}_t]$ at time t. Using the best predictor we see that the conditional MSEP is equal to the conditional prediction variance. If we use a different $\mathcal{T}_t$-measurable predictor, there is an additional error term coming from the estimation error, see Sect. 3.1 in Wüthrich–Merz [WM08].

Definition (5.38) considers a *static view* of prediction uncertainty because the uncertainty is quantified over the entire life time of the claim C. We call this static view the *long-term view* or *total uncertainty view*. This total uncertainty view was the state-of-the-art in claims reserving before risk-adjusted solvency requirements were developed. These new developments have added the *dynamic view* of prediction uncertainty that is concerned with the *run-off of risk profile*. The dynamic view can be integrated by considering the martingale description of successive predictions. Note that $\widehat{C}^{(n)} = C$. This allows us to rewrite the conditional MSEP, see (5.36),

$$\begin{aligned}
\mathrm{msep}_{C|\mathcal{T}_t}\left(\widehat{C}^{(t)}\right) &= E\left[\left(\widehat{C}^{(n)} - \widehat{C}^{(t)}\right)^2 \middle| \mathcal{T}_t\right] \\
&= E\left[\left(\sum_{s=t+1}^{n} \mathrm{CDR}_s(\mathbf{X})\right)^2 \middle| \mathcal{T}_t\right] \\
&= \mathrm{Var}\left(\sum_{s=t+1}^{n} \mathrm{CDR}_s(\mathbf{X}) \middle| \mathcal{T}_t\right),
\end{aligned} \tag{5.39}$$

where in the last step we have used that the nominal CDRs have conditional means equal to zero, see also Lemma 5.5.

Lemma 5.6 *Set Assumption 5.1. We have*

$$\mathrm{msep}_{C|\mathcal{T}_t}\left(\widehat{C}^{(t)}\right) = \sum_{s=t+1}^{n} \mathrm{Var}\left(\mathrm{CDR}_s(\mathbf{X})\middle| \mathcal{T}_t\right). \tag{5.40}$$

Proof Note that $(\mathrm{CDR}_s(\mathbf{X}))_{s\geq t+1}$ is a sequence of martingale innovations. Therefore, we have for $u > s > t$, using the tower property of conditional expectations,

$$E\left[\mathrm{CDR}_s(\mathbf{X})\mathrm{CDR}_u(\mathbf{X})\middle| \mathcal{T}_t\right] = E\left[\mathrm{CDR}_s(\mathbf{X})E\left[\mathrm{CDR}_u(\mathbf{X})\middle| \mathcal{T}_s\right]\middle| \mathcal{T}_t\right] = 0. \tag{5.41}$$

This proves that martingale innovations are uncorrelated and then the claim follows from (5.39). □

Lemma 5.6 explains how the total (static) prediction uncertainty needs to be allocated to the future accounting years. This allocation may be described by the conditional MSEPs of the nominal CDR predictions. We define

$$\begin{aligned}
\mathrm{msep}_{\mathrm{CDR}_{t+1}(\mathbf{X})|\mathcal{T}_t}(0) &= E\left[\left(\mathrm{CDR}_{t+1}(\mathbf{X}) - 0\right)^2\middle| \mathcal{T}_t\right] \\
&= \mathrm{Var}\left(\mathrm{CDR}_{t+1}(\mathbf{X})\middle| \mathcal{T}_t\right) \;=\; \mathrm{Var}\left(\widehat{C}^{(t+1)}\middle| \mathcal{T}_t\right).
\end{aligned} \tag{5.42}$$

In industry, the conditional MSEP (5.42) of the CDR prediction is also called *one-year uncertainty* or short-term uncertainty, in contrast to the long-term uncertainty defined by (5.38).

Using Lemma 5.6 we find the following corollary.

Corollary 5.7 *Set Assumption 5.1. We have*

$$\mathrm{msep}_{C|\mathcal{T}_t}\left(\widehat{C}^{(t)}\right) = \sum_{s=t}^{n-1} E\left[\mathrm{msep}_{\mathrm{CDR}_{s+1}(\mathbf{X})|\mathcal{T}_s}(0)\middle|\,\mathcal{T}_t\right]. \tag{5.43}$$

Corollary 5.7 is the key result to allocate the total prediction uncertainty measured by the conditional MSEP to the different future accounting years. We may interpret this as a run-off of risk profile if we measure risk by the conditional MSEP. Note that we use the uncorrelatedness of martingale innovations to obtain Corollary 5.7 and, in general, there are no similar results available for other conditional risk measures $\varrho^{(t)}$. This may be considered as an advantage of the conditional MSEP because it will allow us to allocate risk, but it may also be considered as a weakness because the conditional MSEP is not able to capture dependence beyond linear correlations.

5.3.4 The Conditional Long-Term MSEP in the CL Method

Our aim is to determine the run-off of risk profile of Corollary 5.7 for the gamma-gamma BCL model of Assumption 5.3. The main results that we state in the next sections are proved in full detail in Merz–Wüthrich [MW14]. Since these proofs are laborious (but not too difficult) we only sketch them here. We define (subject to being well-defined)

$$\Psi_j^{(t)} = \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} + \sigma_j^2(\gamma_j - 2)}. \tag{5.44}$$

Note that $\Psi_j^{(t)}$ is $\mathcal{T}_{t-1}$-measurable, i.e. previsible w.r.t. $\mathcal{T}_t$.

Theorem 5.8 *Under Assumption 5.3 the BCL predictor satisfies for $t \geq I \geq i > t-J$*

$$\mathrm{msep}_{C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{BCL(t)}\right) = \widehat{C}_{i,J}^{BCL(t)} \sum_{j=t-i}^{J-1} \sigma_j^2 \prod_{m=j}^{J-1} \left(\widehat{f}_m^{BCL(t)}\left(1+\Psi_m^{(t)}\right)\right)$$
$$+ \left(\widehat{C}_{i,J}^{BCL(t)}\right)^2 \left(\prod_{j=t-i}^{J-1}\left(1+\Psi_j^{(t)}\right) - 1\right), \tag{5.45}$$

under the additional assumption that $\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}/\sigma_j^2 + \gamma_j > 2$ for all $t-i \leq j \leq J-1$; otherwise the second moment is infinite. The aggregated conditional MSEP is given by

$$\mathrm{msep}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\sum_i \widehat{C}_{i,J}^{BCL(t)}\right) = \sum_i \mathrm{msep}_{C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{BCL(t)}\right) + 2\sum_{i<m} \widehat{C}_{i,J}^{BCL(t)}\, \widehat{C}_{m,J}^{BCL(t)} \left(\prod_{j=t-i}^{J-1}\left(1+\Psi_j^{(t)}\right) - 1\right), \tag{5.46}$$

where the summations run over $t - J + 1 \leq i \leq I$ *and* $t - J + 1 \leq i < m \leq I$, *respectively.*

Sketch of proof. The proof uses the fact that, under Assumption 5.3, we have a Bayesian model with conjugate priors. The resulting posteriors are still independent and gamma distributed with changed scale and shape parameters. The remaining calculations then use the conditional independence and Markov property of cumulative payments as well as the tower property of conditional expectations. For more details we refer to Merz–Wüthrich [MW14]. □

Theorem 5.8 fully specifies the total (long-term) prediction uncertainty (5.38) in the gamma-gamma BCL model under Assumptions 5.1 and 5.3 (and setting $P(t, m) = 1$ for obtaining nominal values). Because we would like to study prediction uncertainty for the classical CL method, we consider the non-informative prior limit in Theorem 5.8 by letting $\boldsymbol{\gamma} \to \mathbf{1}$. Additionally, we will consider a linear approximation to the full Bayesian uncertainty formula of Theorem 5.8. This linear approximation has the advantage that it simplifies calculations and, moreover, we rediscover Mack's uncertainty formula [Ma93].

For this purpose, we need to analyze the terms $\Psi_j^{(t)}$ for $0 \leq j \leq J - 1$. We make the following assumptions for the observed nominal cumulative claims: assume for $t \geq I$ and all $0 \leq j \leq J - 1$

$$\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} \geq \sum_{\ell=1}^{I-j-1} C_{\ell,j} \gg \sigma_j^2, \tag{5.47}$$

i.e. the variance parameters σ_j^2 are much smaller than the column-wise observations at time I. Real data often satisfies this assumption. It also guarantees that the conditional MSEPs provided in Theorem 5.8 are finite for all $\gamma_j > 1$, thus we may consider their non-informative prior limit $\boldsymbol{\gamma} \to \mathbf{1}$. Moreover, it makes the following linear approximations appropriate.

Assumption (5.47) implies uniformly for $\gamma_j > 1$

$$0 < \Psi_j^{(t)} < \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} - \sigma_j^2} \leq \frac{\sigma_j^2}{\sum_{\ell=1}^{I-j-1} C_{\ell,j} - \sigma_j^2} \ll 1. \tag{5.48}$$

This gives an approximation for the first term on the right-hand side of (5.45)

$$\widehat{C}_{i,J}^{BCL(t)} \sum_{j=t-i}^{J-1} \sigma_j^2 \prod_{m=j}^{J-1} \left(\widehat{f}_m^{BCL(t)} \left(1 + \Psi_m^{(t)}\right)\right) \tag{5.49}$$

$$\approx \widehat{C}_{i,J}^{BCL(t)} \sum_{j=t-i}^{J-1} \sigma_j^2 \prod_{m=j}^{J-1} \widehat{f}_m^{BCL(t)} = \left(\widehat{C}_{i,J}^{BCL(t)}\right)^2 \sum_{j=t-i}^{J-1} \frac{\sigma_j^2}{\widehat{C}_{i,j}^{BCL(t)}}.$$

In fact, the right-hand side is a *lower bound* for the left-hand side for *any* $\gamma_j > 1$ (where the second posterior moment exists). For the second term on the right-hand side of (5.45) we have an approximation under (5.47)

$$\left(\widehat{C}_{i,J}^{BCL(t)}\right)^2 \left(\prod_{j=t-i}^{J-1} \left(1 + \Psi_j^{(t)}\right) - 1\right) \approx \left(\widehat{C}_{i,J}^{BCL(t)}\right)^2 \sum_{j=t-i}^{J-1} \Psi_j^{(t)}. \tag{5.50}$$

The right-hand side is again a *lower bound* for the left-hand side for *any* $\gamma_j > 1$.

This implies under assumption (5.47) for all $t - i \le j \le J - 1$ that we receive approximation

$$\text{msep}_{C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{BCL(t)}\right) \approx \left(\widehat{C}_{i,J}^{BCL(t)}\right)^2 \sum_{j=t-i}^{J-1} \left[\frac{\sigma_j^2}{\widehat{C}_{i,j}^{BCL(t)}} + \Psi_j^{(t)}\right], \tag{5.51}$$

where the right-hand side is a *lower bound* for the left-hand side for *any* $\gamma_j > 1$.

In the non-informative prior limit $\boldsymbol{\gamma} \to \mathbf{1}$ we get under (5.47) approximation

$$\lim_{\gamma_j \to 1} \Psi_j^{(t)} = \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} - \sigma_j^2} \approx \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}}, \tag{5.52}$$

where the right-hand side is a *lower bound* for the left-hand side. This provides the following conditional MSEP estimator in the non-informative prior limit.

Estimator 5.9 *Under Assumption 5.3 and under assumption (5.47) for all $t - i \le j \le J - 1$ the following estimator is appropriate*

$$\lim_{\boldsymbol{\gamma} \to \mathbf{1}} \text{msep}_{C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{BCL(t)}\right) \approx \widehat{\text{msep}}_{C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) \tag{5.53}$$

$$\stackrel{\text{def.}}{=} \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \sum_{j=t-i}^{J-1} \left[\frac{\sigma_j^2}{\widehat{C}_{i,j}^{CL(t)}} + \frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}}\right],$$

and the right-hand side is a lower bound for the left-hand side.

The right-hand side of (5.53) is Mack's famous uncertainty formula [Ma93]. There are two important remarks: Firstly, Mack [Ma93] derived his uncertainty formula in a different stochastic model. What we have just shown is that under assumption (5.47) Mack's distribution-free CL model approach and our non-informative prior limit gamma-gamma BCL model come to a similar conclusion about the prediction uncertainty measured by the conditional MSEP. Secondly, Mack [Ma93] uses a different parametrization for the variance parameters which may be obtained by setting $\widehat{s}_j^{(t)} = \sigma_j \widehat{f}_j^{CL(t)}$. This is justified by the fact that we have for the first two conditional moments in the gamma-gamma BCL model

$$\mathbb{E}\left[C_{i,j+1} \middle| \mathcal{T}_{i+j}, \boldsymbol{\Theta} \right] = \Theta_j^{-1}\, C_{i,j}, \tag{5.54}$$

$$\mathrm{Var}\left(C_{i,j+1} \middle| \mathcal{T}_{i+j}, \boldsymbol{\Theta} \right) = \sigma_j^2 \Theta_j^{-2}\, C_{i,j}. \tag{5.55}$$

Parameter Θ_j^{-1} plays the role of the CL factor in Mack's distribution-free CL model and $\sigma_j^2 \Theta_j^{-2}$ plays the role of the variance parameter (which can be denoted by s_j^2).

For aggregated accident years we obtain the following estimator in the non-informative prior limit case.

Estimator 5.10 *Under Assumption 5.3 and under assumption (5.47) for all $t - I \leq j \leq J-1$ the following estimator is appropriate*

$$\lim_{\gamma \to \mathbf{1}} \mathrm{msep}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\sum_i \widehat{C}_{i,J}^{BCL(t)} \right) \approx \widehat{\mathrm{msep}}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\sum_i \widehat{C}_{i,J}^{CL(t)} \right) \tag{5.56}$$

$$\stackrel{\text{def.}}{=} \sum_i \widehat{\mathrm{msep}}_{C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{CL(t)} \right) + 2 \sum_{i<m} \widehat{C}_{i,J}^{CL(t)}\, \widehat{C}_{m,J}^{CL(t)} \sum_{j=t-i}^{J-1} \frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}},$$

where the summations run over $t - J + 1 \leq i \leq I$ and $t - J + 1 \leq i < m \leq I$, respectively. The right-hand side is a lower bound for the left-hand side.

Example 5.1 (*revisited*).
We revisit the data given on page 97. We analyze the non-informative prior limit $\gamma \to \mathbf{1}$ in Theorem 5.8 and its approximation (and lower bound) given in Estimators 5.9 and 5.10 at time $t = I$. Therefore, we still need to calibrate the variance parameters σ_j^2. We first estimate $(\widehat{s}_j^{(t)})^2$ with the classical estimators used in Mack [Ma93] and from these estimates we calibrate σ_j^2. We use as estimators for $0 \leq j \leq J-1$ and $t \geq I$

$$(\widehat{s}_j^{(t)})^2 = \frac{1}{((t-j-1) \wedge I) - 1} \sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} \left(\frac{C_{\ell,j+1}}{C_{\ell,j}} - \widehat{f}_j^{CL(t)} \right)^2. \tag{5.57}$$

If $t = I = J+1$, i.e. if we have a triangle of observations, we cannot estimate $(\widehat{s}_{J-1}^{(t)})^2$ as in (5.57), because we only have one observation for the last developmentperiod

$J-1$. In that case, we use for the last development period the estimator proposed by Mack [Ma93] given by

$$(\widehat{s}_{J-1}^{(t)})^2 = \min\left\{(\widehat{s}_{J-3}^{(t)})^2, (\widehat{s}_{J-2}^{(t)})^2, (\widehat{s}_{J-2}^{(t)})^4/(\widehat{s}_{J-3}^{(t)})^2\right\}. \tag{5.58}$$

The numerical values for our data of these estimates are also given after Table 1 in Mack [Ma93]. We then use the following estimates for σ_j^2, $0 \le j \le J-1$, at time $t \ge I$

$$(\widehat{\sigma}_j^{(t)})^2 = (\widehat{s}_j^{(t)})^2/(\widehat{f}_j^{CL(t)})^2. \tag{5.59}$$

This provides the results in Table 5.4. The table shows the rooted versions of the non-informative prior limits of the conditional MSEP formulas given in Theorem 5.8 and its approximations (and lower bounds) given in Estimators 5.9 and 5.10. Note that Estimators 5.9 and 5.10 correspond to Mack's formula [Ma93] and are identical to the column 'Chain ladder' in Table 3 of Mack [Ma93]. The first observation is that the approximations (and lower bounds) are sufficiently close to the non-informative prior limits of Theorem 5.8. This justifies under (5.47) the use of the approximations, and in the sequel we directly work with Mack's formula (5.53), (5.56), see also (5.77) below. From a quantitative point of view we find that the prediction uncertainty has a substantial size and confidence bounds of one standard deviation would roughly have a size of 13 % of the total nominal best-estimate reserves. This indicates that prediction uncertainty may substantially influence the VaPo and we should add a significant loading against downside risks in this example. In order to do so we need to understand the dynamic behaviour of this uncertainty, this is exactly what we are going to derive next. □

Table 5.4 Nominal best-estimate reserves $R_I^{(I+1)}$ and the rooted conditional MSEPs of the non-informative prior limit in the gamma-gamma BCL model (5.45) and (5.46) and of Mack's formula (5.53) and (5.56).

Accident year i	Best-estimate reserves $R_I^{(I+1)}$	Bayes' (5.46) $\lim_{\gamma \to \mathbf{1}} \text{msep}^{1/2}$	Mack's (5.56) $\widehat{\text{msep}}^{1/2}$	In % reserves (%)
1	0	0	0	–
2	94,634	75,539	75,535	80
3	469,511	121,710	121,699	26
4	709,638	133,562	133,549	19
5	984,889	261,482	261,406	27
6	1,419,459	411,216	411,010	29
7	2,177,641	558,696	558,317	26
8	3,920,301	876,326	875,328	22
9	4,278,972	972,733	971,258	23
10	4,625,811	1,367,285	1,363,155	29
Total	18,680,856	2,450,978	2,447,095	13

5.3.5 *The Claims Development Result Uncertainty in the CL Method*

In this section we study the one-year uncertainties of the CDRs, see (5.42). Corollary 5.7 then explains how these one-year uncertainties need to be allocated and aggregated. We define for $1 \leq t-j \leq I$ the (credibility) weights

$$\alpha_j^{(t)} = \frac{C_{t-j,j}}{\sum_{\ell=1}^{t-j} C_{\ell,j} + \sigma_j^2\,(\gamma_j - 1)} \;\in\; (0,1). \tag{5.60}$$

These weights are $\mathcal{T}_t$-measurable. We define the nominal CDR of accident year i at time $t \geq I \geq i > t-J$ by, see (5.36),

$$\mathrm{CDR}_{i,t+1} = \widehat{C}_{i,J}^{BCL(t)} - \widehat{C}_{i,J}^{BCL(t+1)}. \tag{5.61}$$

Theorem 5.11 *Under Assumption 5.3 the BCL predictor satisfies for $t \geq I \geq i > t-J$*

$$\begin{aligned} &\mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) \\ &\quad = (\widehat{C}_{i,J}^{BCL(t)})^2 \left[\left(1 + (\alpha_{t-i}^{(t)})^{-1} \Psi_{t-i}^{(t)}\right) \prod_{j=t-i+1}^{J-1} \left(1 + \alpha_j^{(t)} \Psi_j^{(t)}\right) - 1 \right], \end{aligned} \tag{5.62}$$

where we assume that $\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}/\sigma_j^2 + \gamma_j > 2$ for all $t-i \leq j \leq J-1$, otherwise the second moment is infinite. The aggregated conditional MSEP is given by

$$\begin{aligned} \mathrm{msep}_{\sum_i \mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) &= \sum_i \mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) \\ &+ 2 \sum_{i<m} \widehat{C}_{i,J}^{BCL(t)} \widehat{C}_{m,J}^{BCL(t)} \left[\left(1 + \Psi_{t-i}^{(t)}\right) \prod_{j=t-i+1}^{J-1} \left(1 + \alpha_j^{(t)} \Psi_j^{(t)}\right) - 1 \right], \end{aligned} \tag{5.63}$$

the summations run over $t-J+1 \leq i \leq I$ and $t-J+1 \leq i < m \leq I$, respectively.

Sketch of proof. The proof uses in a careful manner the conditional independence and Markov property of cumulative payments: note that going from information $\mathcal{T}_t$ to information $\mathcal{T}_{t+1}$ adds one new diagonal of observations to the loss development trapezoid $\mathcal{T}_t$. This new diagonal only contains cumulative payments that belong to different accident years and different development years. This is crucial in the calculations and the proof fails when studying several accounting years simultaneously. For details we refer to Merz–Wüthrich [MW14]. □

We give approximations to the non-informative prior limits of the results of Theorem 5.11. This is similar to the derivations in Sect. 5.3.4 and as a result we obtain the Merz–Wüthrich (MW) formula [MW08]. We use (5.47) and obtain for $1 \leq t-j \leq I$

$$0 < \alpha_j^{(t)} \Psi_j^{(t)} < \frac{\sigma_j^2}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j} - \sigma_j^2} \le \frac{\sigma_j^2}{\sum_{\ell=1}^{I-j-1} C_{\ell,j} - \sigma_j^2} \ll 1. \tag{5.64}$$

This allows us to treat the terms under the product in (5.62). The first term in the square bracket of (5.62) is more sophisticated. From Corollary A.4 in [MW14] we have

$$(\alpha_{t-i}^{(t)})^{-1} \Psi_{t-i}^{(t)} = \left(\frac{\sigma_{t-i}^2}{C_{i,t-i}} + 1 \right) \left(1 + \Psi_{t-i}^{(t)} \right) - 1. \tag{5.65}$$

We make an additional assumption. For $0 \le t - i \le J - 1$ we assume

$$C_{i,t-i} \gg \sigma_{t-i}^2. \tag{5.66}$$

Assumptions (5.47) and (5.66) imply that

$$(\alpha_{t-i}^{(t)})^{-1} \Psi_{t-i}^{(t)} \approx \frac{\sigma_{t-i}^2}{C_{i,t-i}} + \Psi_{t-i}^{(t)} \ll 1, \tag{5.67}$$

where the right-hand side of the approximation is a lower bound for the left-hand side. This provides an approximation and lower bound to (5.62)

$$\mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) \approx (\widehat{C}_{i,J}^{BCL(t)})^2 \left[\frac{\sigma_{t-i}^2}{C_{i,t-i}} + \Psi_{t-i}^{(t)} + \sum_{j=t-i+1}^{J-1} \alpha_j^{(t)} \Psi_j^{(t)} \right]. \tag{5.68}$$

Finally, we consider the non-informative prior limit $\boldsymbol{\gamma} \to \mathbf{1}$ which provides the following estimator.

Estimator 5.12 *Under Assumption 5.3 and under assumptions* (5.47) *and* (5.66) *for all* $t - i \le j \le J - 1$ *the following estimator is appropriate*

$$\lim_{\boldsymbol{\gamma}\to\mathbf{1}} \mathrm{msep}_{\mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) \approx \widehat{\mathrm{msep}}_{\mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0), \tag{5.69}$$

$$\stackrel{\text{def.}}{=} (\widehat{C}_{i,J}^{CL(t)})^2 \left[\frac{\sigma_{t-i}^2}{C_{i,t-i}} + \frac{\sigma_{t-i}^2}{\sum_{\ell=1}^{i-1} C_{\ell,t-i}} + \sum_{j=t-i+1}^{J-1} \overline{\alpha}_j^{(t)} \frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right],$$

with weights for $1 \le t - j \le I$

$$\overline{\alpha}_j^{(t)} = \lim_{\gamma_j \to 1} \alpha_j^{(t)} = \frac{C_{t-j,j}}{\sum_{\ell=1}^{t-j} C_{\ell,j}} \in (0, 1]. \tag{5.70}$$

The right-hand side in (5.69) *is a lower bound for the left-hand side.*

This is exactly the MW formula [MW08] if we use the re-parametrization described after (5.53), that is, set $\widehat{s}_j^{(t)} = \sigma_j \widehat{f}_j^{CL(t)}$. Compare Mack's formula (5.53) to the MW formula (5.69): the first term of the summation of Mack's formula (5.53) with index $j = t - i$ is the same in the MW formula (5.69); from the remaining terms of the summation of Mack's formula with indexes $j \geq t - i + 1$ only the second part is considered in the MW formula with the additional scaling $\overline{\alpha}_j^{(t)} \in (0, 1]$, for more information we refer to Merz–Wüthrich [MW14].

Estimator 5.13 *Define the estimator*

$$\widehat{\mathrm{msep}}_{\sum_i \mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) = \sum_i \widehat{\mathrm{msep}}_{\mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0) \tag{5.71}$$

$$+ 2 \sum_{i<m} \widehat{C}_{i,J}^{CL(t)} \widehat{C}_{m,J}^{CL(t)} \left[\frac{\sigma_{t-i}^2}{\sum_{\ell=1}^{i-1} C_{\ell,t-i}} + \sum_{j=t-i+1}^{J-1} \overline{\alpha}_j^{(t)} \frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}} \right],$$

where the summations run over $t - J + 1 \leq i \leq I$ and $t - J + 1 \leq i < m \leq I$, respectively. Under Assumption 5.3 and under assumptions (5.47) and (5.66) for all $t - I \leq j \leq J - 1$ this is an appropriate estimator and a lower bound for the non-informative prior limit of (5.63).

Example 5.1 (*revisited*).
We revisit the data given on page 97. We compare Mack's total uncertainty formula (5.56) to the MW one-period uncertainty formula (5.71) at time $t = I$.

This gives the results presented in Table 5.5. We observe that the one-year uncertainty measured in terms of the rooted conditional MSEP (5.71) accounts for roughly 75 % of the total uncertainty. This indicates that the considered business is rather

Table 5.5 Nominal best-estimate reserves $R_I^{(I+1)}$ and the rooted conditional MSEPs of the total prediction uncertainty (5.53) and (5.56) and of the one-year CDR uncertainty (5.69) and (5.71)

Accident year i	Best-estimate reserves $R_I^{(I+1)}$	Total $\widehat{\mathrm{msep}}^{1/2}$ see (5.56)	CDR $\widehat{\mathrm{msep}}^{1/2}$ see (5.71)	CDR in % of total (%)
1	0	0	0	–
2	94,634	75,535	75,535	100
3	469,511	121,699	105,309	87
4	709,638	133,549	79,846	60
5	984,889	261,406	235,115	90
6	1,419,459	411,010	318,427	77
7	2,177,641	558,317	361,089	65
8	3,920,301	875,328	629,681	72
9	4,278,972	971,258	588,662	61
10	4,625,811	1,363,155	1,029,925	76
Total	18,680,856	2,447,095	1,778,968	73

long-tailed and we expect a slow run-off of risk profile. This is common for liability insurance and for personal injury claims which may have a ratio in the range of 60 %. In property insurance this ratio is in the range of 90 % which indicates a fast run-off of risk profile, we also refer to Table 11 in the field study of AISAM-ACME [AISAM07]. □

5.3.6 The Run-Off of Risk Profile in the CL Method

In Theorems 5.8 and 5.11 we have provided the conditional MSEP formulas for the total prediction uncertainty and for the one-year CDR uncertainty, respectively. Corollary 5.7 describes the splitting property, that is, for $t \geq I$

$$\mathrm{msep}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\sum_i \widehat{C}_{i,J}^{BCL(t)}\right) = \sum_{s=t}^{I+J-1} E\left[\mathrm{msep}_{\sum_i \mathrm{CDR}_{i,s+1}|\mathcal{T}_s}(0)\,\Big|\,\mathcal{T}_t\right]. \tag{5.72}$$

The remaining difficulty is that the terms on the right-hand side cannot be easily calculated for $s > t$, i.e. the conditional expectations of the corresponding CDR uncertainties may be non-trivial. Merz–Wüthrich [MW14] derive an approximation and lower bound for these conditional expectations in the non-informative prior limit case $\boldsymbol{\gamma} \to \mathbf{1}$ and under assumptions (5.47) and (5.66). We therefore define for $t \geq I \geq i \geq t - J + k + 1$

$$\begin{aligned}\chi_{i,t+k+1}^{(t)} = &\left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \left[\frac{\sigma_{t+k-i}^2}{\widehat{C}_{i,t+k-i}^{CL(t)}} + \prod_{s=1}^{k}\left(1 - \overline{\alpha}_{t+s-i}^{(t)}\right)\frac{\sigma_{t+k-i}^2}{\sum_{\ell=1}^{i-k-1} C_{\ell,t+k-i}}\right] \\ &+ \left(\widehat{C}_{i,J}^{CL(t)}\right)^2 \sum_{j=t+k-i+1}^{J-1}\left[\overline{\alpha}_{j-k}^{(t)}\prod_{s=0}^{k-1}\left(1 - \overline{\alpha}_{j-s}^{(t)}\right)\frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}}\right],\end{aligned} \tag{5.73}$$

and for aggregated accident years

$$\begin{aligned}\chi_{t+k+1}^{(t)} = &\sum_i \chi_{i,t+k+1}^{(t)} \\ &+ 2\sum_{i<m} \widehat{C}_{i,J}^{CL(t)}\widehat{C}_{m,J}^{CL(t)}\prod_{s=1}^{k}\left(1 - \overline{\alpha}_{t+s-i}^{(t)}\right)\frac{\sigma_{t+k-i}^2}{\sum_{\ell=1}^{i-k-1} C_{\ell,t+k-i}} \\ &+ 2\sum_{i<m} \widehat{C}_{i,J}^{CL(t)}\widehat{C}_{m,J}^{CL(t)}\sum_{j=t+k-i+1}^{J-1}\left[\overline{\alpha}_{j-k}^{(t)}\prod_{s=0}^{k-1}\left(1 - \overline{\alpha}_{j-s}^{(t)}\right)\frac{\sigma_j^2}{\sum_{\ell=1}^{t-j-1} C_{\ell,j}}\right],\end{aligned} \tag{5.74}$$

where the sums run over $t - J + k + 1 \le i \le I$ and $t - J + k + 1 \le i < m \le I$. For the special case $k = 0$ we have

$$\chi_{t+1}^{(t)} = \widehat{\mathrm{msep}}_{\sum_i \mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0). \tag{5.75}$$

The crucial observation is that $\chi_{t+k+1}^{(t)}$ is under assumptions (5.47) and (5.66) in the non-informative prior limit case $\boldsymbol{\gamma} \to \mathbf{1}$ an appropriate approximation to the corresponding term on the right-hand side of (5.72). This is the main result derived in Sect. 6 of Merz–Wüthrich [MW14], moreover their Corollary 6.5 gives the following allocation formula:

Corollary 5.14 *Choose $t \geq I$. We have the total prediction uncertainty splitting property*

$$\widehat{\mathrm{msep}}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\sum_i \widehat{C}_{i,J}^{CL(t)}\right) = \sum_{k=0}^{I+J-t-1} \chi_{t+k+1}^{(t)}. \tag{5.76}$$

Assume $\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}/\sigma_j^2 + \gamma_j > 2$ for all $t-I \le j \le J-1$. The total approximation error under Assumption 5.3 in the non-informative prior limit $\boldsymbol{\gamma} \to \mathbf{1}$ is given by

$$\begin{aligned} 0 \geq & \sum_{k=0}^{I+J-t-1} \chi_{t+k+1}^{(t)} - \sum_{k=0}^{I+J-t-1} \lim_{\gamma \to \mathbf{1}} \mathbb{E}\left[\mathrm{msep}_{\sum_i \mathrm{CDR}_{i,t+k+1}|\mathcal{T}_{t+k}}(0) \middle| \mathcal{T}_t \right] \\ &= \widehat{\mathrm{msep}}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\sum_i \widehat{C}_{i,J}^{CL(t)}\right) - \lim_{\gamma \to \mathbf{1}} \mathrm{msep}_{\sum_i C_{i,J}|\mathcal{T}_t}\left(\widehat{C}_{i,J}^{BCL(t)}\right), \end{aligned} \tag{5.77}$$

and the non-informative prior limits $\boldsymbol{\gamma} \to \mathbf{1}$ are obtained from Theorems 5.8 and 5.11.

Conclusions.

- Formula (5.76) gives an explicit allocation of the total prediction uncertainty to estimated one-year prediction uncertainties $\chi_{t+k+1}^{(t)}$, $k \geq 0$. The resulting split is an approximation in the non-informative prior limit gamma-gamma BCL model. The advantage of having this approximation is that it is analytically tractable, i.e. (5.74) can easily be calculated. The quality of this approximation can be evaluated explicitly with (5.77).
- We interpret the sequence $(\chi_{t+k+1}^{(t)})_{k=0,\ldots,I+J-t-1}$ as the *run-off of risk profile* at time t because it explains how the total prediction uncertainty (measured by the conditional MSEP) is released over time. This is in completion to the run-off of the claims reserves that describes the release of the reserves (conditional expectations).

Example 5.1 (*revisited*).
We revisit the data given on page 97 and provide the run-off of risk profile. We mention that (5.74) is implemented in the CRAN R package `ChainLadder`, see Gesmann et al. [Getal15].

Table 5.6 Run-off of the expected best-estimate reserves and of the corresponding run-off of risk profile

Accounting years $t \geq I$	Run-off exp. best-estimate reserves $R_I^{(t+1)}$	Run-off rooted exp. MSEP $\eta_I^{(t+1)}$	In % reserves (%)	Expected cash flow $E[X_t \vert \mathcal{T}_I]$	Rooted expected CDR MSEP $\sqrt{\chi_t^{(I)}}$
10	18,680,856	2,447,095	13	–	–
11	13,454,320	1,680,341	12	5,226,536	1,778,968
12	9,274,925	1,198,543	13	4,179,394	1,177,727
13	6,143,258	808,063	13	3,131,668	885,178
14	4,015,986	532,562	13	2,127,272	607,736
15	2,454,107	315,998	13	1,561,879	428,681
16	1,276,363	168,216	13	1,177,744	267,503
17	532,076	108,489	20	744,287	128,557
18	86,555	49,055	57	445,521	96,764
19	0	0	–	86,555	49,055

Table 5.6 provides the run-off of the expected nominal best-estimate reserves in the non-informative prior limit gamma-gamma BCL model for $t \geq I$, given by

$$R_I^{(t+1)} = \sum_{s=t+1}^{I+J} E\left[X_s \middle| \mathcal{T}_I\right] = \sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(I)} - \widehat{C}_{i,t-i}^{CL(I)}. \tag{5.78}$$

The corresponding cash flow predictions are given by

$$E\left[X_{t+1} \middle| \mathcal{T}_I\right] = R_I^{(t+1)} - R_I^{(t+2)}, \tag{5.79}$$

see also Table 5.1. These are complemented by the rooted uncertainty estimates $\sqrt{\chi_{t+1}^{(I)}}$ of formula (5.74) and the *remaining run-off uncertainty* at time $t \geq I$ defined by

$$\eta_I^{(t+1)} = \sqrt{\sum_{\ell \geq t} \chi_{\ell+1}^{(I)}} \approx \left(\sum_{\ell \geq t} E\left[\widehat{\mathrm{msep}}_{\sum_i \mathrm{CDR}_{i,\ell+1} | \mathcal{T}_\ell}(0) \middle| \mathcal{T}_I \right] \right)^{1/2}. \tag{5.80}$$

Thus, we can now study two run-offs, the first one describing the expected *release of best-estimate reserves* and the second the *run-off of risk profile*:

$$\left(R_I^{(t+1)}\right)_{t=I,\dots,I+J-1} \quad \text{and} \quad \left(\eta_I^{(t+1)}\right)_{t=I,\dots,I+J-1}. \tag{5.81}$$

From Table 5.6 we see that in the actual example the behaviour of these two run-offs is very similar because the ratio $\eta_I^{(t+1)}/R_I^{(t+1)}$ is more or less constant in t. This behaviour is not a typical one, and in many other examples the run-off risk profile $\eta_I^{(t+1)}$ is slower than the expected run-off of reserves $R_I^{(t+1)}$ implied pattern, we refer to Merz–Wüthrich [MW14] for more examples. □

5.3.7 *The Cost-of-Capital Loading*

We now use the uncertainty estimates $\sqrt{\chi_{t+1}^{(I)}}$, $t = I, \ldots, I + J - 1$, for the construction of a cost-of-capital loading in the CL method. Before doing so, we would like to present the idea behind the cost-of-capital loading. The VaPo protected against insurance technical risks is given by

$$\mathrm{VaPo}_t^{prot}\left(\mathbf{X}_{(t+1)}\right) = \sum_{k=1}^{n-t} l_{t+k}^{*,t}\, \mathcal{Z}^{(t+k)}, \tag{5.82}$$

for a given probability distortion $\boldsymbol{\varphi}^{\mathbb{T}}$ and distorted means

$$l_{t+k}^{*,t} = \frac{1}{\varphi_t^{\mathbb{T}}}\, E\left[\left.\varphi_{t+k}^{\mathbb{T}} X_{t+k}\right| \mathcal{T}_t\right]. \tag{5.83}$$

Reasonable choices of probability distortions $\boldsymbol{\varphi}^{\mathbb{T}}$ provide positive loadings for (non-hedgeable) insurance technical risks, see (3.37). Since zero coupon bond prices are positive, this asks for a positive correlation between $\varphi_{t+k}^{\mathbb{T}}$ and X_{t+k} under P, conditionally given $\mathcal{T}_t$. This positive correlation then implies, also using the martingale property of probability distortions,

$$l_{t+k}^{*,t} = \frac{1}{\varphi_t^{\mathbb{T}}}\, E\left[\left.\varphi_{t+k}^{\mathbb{T}} X_{t+k}\right| \mathcal{T}_t\right] \geq \frac{1}{\varphi_t^{\mathbb{T}}}\, E\left[\left.\varphi_{t+k}^{\mathbb{T}}\right| \mathcal{T}_t\right] E\left[\left.X_{t+k}\right| \mathcal{T}_t\right] = l_{t+k}^{(t)}. \tag{5.84}$$

Thus, we obtain a positive loading on the liability side of the balance sheet for non-hedgeable insurance technical risks. In order to have a reasonable (expected) release of this loading over time we need the additional requirement that $(l_{t+k}^{*,s})_{s \leq t+k}$ is a $(P, \mathbb{T})$-super-martingale. This super-martingale assumption implies

$$E\left[\left. l_{t+k}^{*,t+1} \right| \mathcal{T}_t\right] \leq l_{t+k}^{*,t}. \tag{5.85}$$

Thus, if we purchase $l_{t+k}^{*,t}$ zero coupon bonds $\mathcal{Z}^{(t+k)}$ at time t we expect (based on information $\mathcal{T}_t$) to be able to sell at time $t + 1$ the portfolio

$$\left(l_{t+k}^{*,t} - E\left[\left. l_{t+k}^{*,t+1} \right| \mathcal{T}_t\right]\right) \mathcal{Z}^{(t+k)}, \tag{5.86}$$

to re-adjust $\mathrm{VaPo}_t^{prot}\left(\mathbf{X}_{(t+1)}\right)$ to $\mathrm{VaPo}_{t+1}^{prot}\left(\mathbf{X}_{(t+1)}\right)$ at time $t+1$. This expected release of financial instruments generates expected profits which should be considered as a reward for risk bearing in accounting year $(t, t+1]$. We have just explained that the (market-)consistent construction of a loading can be seen as a reward for risk bearing (of insurance technical risks). The crucial assumption is that the probability distortion $\varphi^{\mathbb{T}}$ needs to induce the super-martingale property (5.85) for obtaining a dynamically reasonable economic model.

In industry practice, the super-martingale property (5.85) is constructed in an ad-hoc manner. We follow this ad-hoc construction here because calibrating probability distortions $\varphi^{\mathbb{T}}$ is rather difficult. The argument is as follows: In accounting year $(t, t+1]$ we face downside risks described by $\mathcal{CDR}_{t+1}(\mathbf{X})$, see (5.30). To protect against these downside risks we need to hold risk bearing capital denoted by the conditional $\mathcal{T}_t$-measurable risk measure $\varrho^{(t)}$. If this risk bearing capital $\varrho^{(t)}$ is provided by a shareholder at time t he expects an expected return in the form of a cost-of-capital reward at time $t+1$. If the rate of return, called the *cost-of-capital rate*, is denoted by $r_{\mathrm{CoC}} > 0$ then he expects a reward at time $t+1$ (usually in the form of dividends) of

$$r_{\mathrm{CoC}} \cdot \varrho^{(t)} \tag{5.87}$$

on his investment $\varrho^{(t)}$. Note that we have deliberately dropped the time dependence in the cost-of-capital rate r_{CoC} because the following case study will only consider nominal claims.

We set Assumptions 5.1 and 5.3 and we work with the non-informative prior limit of the gamma-gamma BCL model. Under the additional assumption of nominal values $P(t, m) \equiv 1$, we choose conditional risk measures

$$\varrho^{(t)} = \kappa \left(\widehat{\mathrm{msep}}_{\sum_i \mathrm{CDR}_{i,t+1}|\mathcal{T}_t}(0)\right)^{1/2} = \kappa \sqrt{\chi_{t+1}^{(t)}}, \tag{5.88}$$

given by (5.71) and (5.75), and $\kappa > 0$ being chosen typically in the range of $[2, 3]$ (determining the width of the confidence bounds). Formula (5.87) then gives the margin that we need to hold for accounting period $(t, t+1]$ in order to finance this risk measure (5.88) provided by the shareholders.

The risk-adjusted reserves stemming from the cost-of-capital loading are then for the non-informative prior limit gamma-gamma BCL model defined by

$$\begin{aligned} R_I^{*,t+1} &= \sum_{i=t-J+1}^{I} \widehat{C}_{i,J}^{CL(I)} - \widehat{C}_{i,t-i}^{CL(I)} + r_{\mathrm{CoC}} \sum_{s=t}^{I+J-1} \kappa \sqrt{\chi_{s+1}^{(I)}} \\ &= R_I^{(t+1)} + r_{\mathrm{CoC}} \sum_{s=t}^{I+J-1} \kappa \sqrt{\chi_{s+1}^{(I)}}. \end{aligned} \tag{5.89}$$

Remarks.

- $R_I^{(t+1)}$ denotes the best-estimate reserves at time I for the outstanding liabilities $\mathbf{X}_{(t+1)}$. These are the predicted cash flows without any loading, see (5.78).
- The second term on the right-hand side of (5.89) denotes the cost-of-capital margin (similar to a loading induced by a probability distortion). As risk measure we choose (5.88) and a constant cost-of-capital rate $r_{\mathrm{CoC}} > 0$. Observe that representation (5.89) uses an approximation and neglects Jensen's inequality, because we have for $s > I$

$$E\left[\varrho^{(s)}\middle|\mathcal{T}_I\right] = \kappa\, E\left[\left(\widehat{\mathrm{msep}}_{\sum_i \mathrm{CDR}_{i,s+1}|\mathcal{T}_s}(0)\right)^{1/2}\middle|\mathcal{T}_I\right]$$
$$\leq \kappa\, E\left[\widehat{\mathrm{msep}}_{\sum_i \mathrm{CDR}_{i,s+1}|\mathcal{T}_s}(0)\middle|\mathcal{T}_I\right]^{1/2} \tag{5.90}$$
$$\approx \kappa\sqrt{\chi_{s+1}^{(I)}},$$

where in the last step we have used the same approximation as in (5.80).

Example 5.1 (*revisited, non-life insurance run-off*).
We revisit the data given on page 97 and compare the expected run-off of the risk-adjusted reserves $R_I^{*,t+1}$, $t \geq I$, to the expected run-off of the best-estimate reserves $R_I^{(t+1)}$ in the CL method, see (5.89).

In Fig. 5.2 we provide the expected run-offs for loading constant $\kappa = 2$ (which corresponds to confidence bounds of two standard deviations) and with cost-of-capital rate $r_{\mathrm{CoC}} = 6\,\%$ (which is the standard value used in Solvency II). The resulting margin at time $t = I$ is comparably small

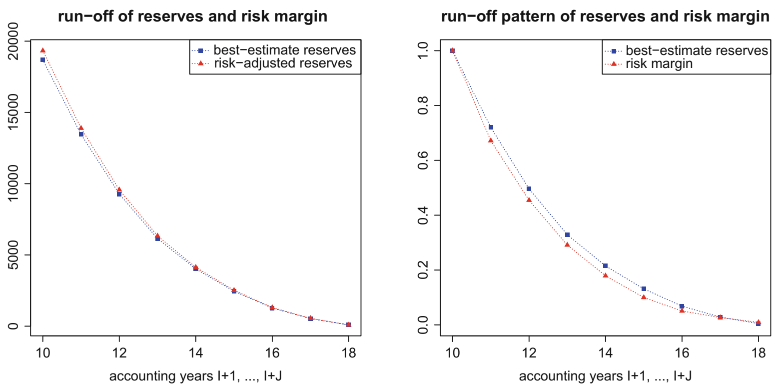


Fig. 5.2 (*lhs*) Expected run-off of the best-estimate reserves $R_I^{(t+1)}$ and of the risk-adjusted reserves $R_I^{*,t+1}$ for $t = I, \ldots, I+J-1$; and (*rhs*) corresponding run-off patterns of the best-estimate reserves and the cost-of-capital margin (normalized to 1 at time I) given in (5.89)

$$r_{\mathrm{CoC}} \sum_{s=I}^{I+J-1} \kappa \sqrt{\chi_{s+1}^{(I)}} = 650,420, \tag{5.91}$$

in relation to the best-estimate reserves $R_I^{(I+1)} = 18,680,856$. In Fig. 5.2 (rhs) we plot the expected run-off pattern of the best-estimate reserves and of the risk margin (last term in (5.89)). We observe that in this example the margin is released faster than the best-estimate reserves (in expectation). In many applied examples this behaviour is just reversed, which may indicate that the Taylor–Ashe [TA83] data is rather special. This closes the non-life insurance run-off example. □

5.4 Unallocated Loss Adjustment Expenses

5.4.1 Motivation

In this section we describe the "New York"-method for the prediction of unallocated loss adjustment expenses (ULAE). The New York-method for predicting ULAE is found in the literature e.g. as footnotes in Feldblum [Fe03], in the CAS notes [CAS90], and in more detail in Buchwalder et al. [BBMW06]. Sometimes this method is also called the paid-to-paid method.

In non-life insurance there are usually two different kinds of claims handling costs: external costs and internal costs. External costs like costs for an external lawyer or for an external expertise are usually allocated to single claims (in the claims files) and are therefore contained in the usual claims payments and loss development figures. These payments are called allocated loss adjustment expenses (ALAE). Typically, internal loss adjustment expenses (income of claims handling department, maintenance of claims handling system, internal lawyers, management fees, etc.) are not contained in the claims figures because they are accounted under administrative expenses. These internal costs can usually not be allocated to single claims. We call these costs therefore unallocated loss adjustment expenses (ULAE). From a regulatory point of view, we should also build reserves for these expenses because they are part of the claims handling process which guarantees that an insurance company is able to meet all its obligations. Thus, ULAE reserves should be built to guarantee the smooth run-off of the old insurance liabilities without a pay-as-you-go system from new business.

In conclusion, this means that ULAE reserves should also be part of the VaPo, if we want to have a self-financing run-off of the insurance liabilities.

5.4.2 Pure Claims Payments

Usually, claims development figures only consist of "pure" claims payments not containing ULAE charges. In this section we denote by $X_{i,j}$ the "pure" incremental payments for accident year $i \in \{1, \dots, I\}$ in development year $j \in \{0, \dots, J\}$. "Pure" always means that these quantities do not contain ULAE. These are exactly the quantities studied in Sect. 5.2. The nominal cumulative pure payments for accident year i after development period j are again denoted by (see (5.10))

$$C_{i,j} = \sum_{s=0}^{j} X_{i,s}. \tag{5.92}$$

We assume that $X_{i,j} = 0$ for all $j > J$, i.e. the ultimate pure claim is given by $C_{i,J}$.

For the New York-method we need a second type of development trapezoids, namely a *reporting trapezoid*: For accident year i, let $Y_{i,j}$ denote the ultimate pure claim amount for all those claims, which are reported up to (and including) development year j. Hence, the vector

$$\left(Y_{i,0}, Y_{i,1}, \dots, Y_{i,J}\right) \tag{5.93}$$

describes how the ultimate pure claim $C_{i,J} = Y_{i,J}$ is reported over time at the insurance company. Of course, this reporting pattern is much more delicate, because claims which are reported in the upper trapezoid

$$\widetilde{\mathcal{D}}_t = \left\{Y_{i,j};\ 1 \le i+j \le t,\ 1 \le i \le I,\ 0 \le j \le J\right\} \tag{5.94}$$

are still developing, since they are not necessarily finally settled yet. Henceforth, the final claim severities may still be random at time t and, in general, $\widetilde{\mathcal{D}}_t$ still contains random variables. However, they correspond to claims reported at time t.

Remark. The New York-method has to be understood as an algorithm used to predict ULAE payments. This algorithm is not based on a stochastic model. Therefore, we assume in this section that all variables are deterministic numbers. Random variables are replaced by their best-estimate predictions at time t. We believe that for the current presentation (to explain the New York-method) it is not helpful to work in a stochastic framework. Moreover, the volume of the ULAE payments is usually comparably small compared to the volume of pure claims payments. Therefore, often, the main risk drivers come from the pure claims payments.

5.4.3 ULAE Charges

The cumulative ULAE payments for accident year i until development period j are denoted by $C_{i,j}^{(ULAE)}$. And finally, the cumulative total payments (pure and ULAE payments) are denoted by

$$C_{i,j}^{+} = C_{i,j} + C_{i,j}^{(ULAE)}. \tag{5.95}$$

The cumulative ULAE payments $C_{i,j}^{(ULAE)}$ and the incremental ULAE charges

$$X_{i,j}^{(ULAE)} = C_{i,j}^{(ULAE)} - C_{i,j-1}^{(ULAE)} \tag{5.96}$$

need to be estimated *and* predicted: the main difficulty in practice is that for each accounting year $k \le I$ (at time I) we have only *one* aggregated observation

$$X_{k}^{(ULAE)} = \sum_{i+j=k} X_{i,j}^{(ULAE)}. \tag{5.97}$$

That is, ULAE payments are usually not available for single accident years but rather we have a position "Total ULAE Expenses" for each accounting year k which is aggregated over all accident years i of ULAE payments done in that accounting year. Usually, ULAE charges are contained in the position "Administrative Expenses" of the annual profit-and-loss statement.

The reason for having only aggregated observations per accounting year is that, in general, the claims handling department treats several claims from different accident years simultaneously. Only an activity-based cost allocation split then allocates these expenses to different accident years. Hence, for the prediction of future ULAE payments we first need to define an appropriate model in order to split the aggregated observations $X_k^{(ULAE)}$ into the different accident years $X_{i,j}^{(ULAE)}$, i.e. to estimate ULAE payments for single past accident years.

5.4.4 *The New York-Method*

The New York-method assumes that one part of the ULAE charge is proportional to the claims registration (denote this proportion by $r \in [0, 1]$) and the other part is proportional to the settlement of the claims (proportion $1 - r$).

Assumption 5.15 There exist two development patterns $(\gamma_j)_{j=0,\dots,J}$ and $(\beta_j)_{j=0,\dots,J}$ with $\gamma_j \ge 0$, $\beta_j \ge 0$, for all $0 \le j \le J$, and $\sum_{j=0}^{J} \gamma_j = \sum_{j=0}^{J} \beta_j = 1$ such that for all $1 \le i \le I$ and $0 \le j \le J$ (cash flow or payout pattern)

$$X_{i,j} = \gamma_j \, C_{i,J}, \tag{5.98}$$

and (reporting pattern)

$$Y_{i,j} = \sum_{\ell=0}^{j} \beta_\ell \, C_{i,J}. \tag{5.99}$$

Remarks.

- Equation (5.98) describes how the ultimate pure claim $C_{i,J}$ is paid over time. In fact, $(\gamma_j)_j$ gives the (expected) cash flow pattern for the ultimate pure claim $C_{i,J}$.

We assume that this cash flow pattern is at time I estimated by the CL factor estimates $(\widehat{f}_\ell^{CL(I)})_\ell$, see (5.18) and Sect. 9.2.2 in Wüthrich [Wü13],

$$\widehat{\gamma}_j^{(I)} = \left(1 - \frac{1}{\widehat{f}_{j-1}^{CL(I)}}\right) \prod_{\ell=j}^{J-1} \frac{1}{\widehat{f}_\ell^{CL(I)}}. \tag{5.100}$$

- The estimation of the claims reporting pattern $(\beta_j)_j$ in (5.99) is more delicate. There are not many claims reserving methods which give a reporting pattern. Such a pattern can only be obtained if one separates the claims estimates for reported but not settled (RBNS) claims from incurred but not yet reported (IBNYR) claims, see Sect. 9.1 in Wüthrich [Wü13].

Model 5.16 *Assume that there exists an* $r \in [0, 1]$ *such that the incremental ULAE payments satisfy for all* $1 \le i \le I$ *and* $0 \le j \le J$

$$X_{i,j}^{(ULAE)} = \left(r\, \beta_j + (1-r)\, \gamma_j\right) C_{i,J}^{(ULAE)}. \tag{5.101}$$

Henceforth, we assume that one part (corresponding to r) of the ULAE charge is proportional to the reporting pattern (one has loss adjustment expenses at the registration of the claim), and the other part (corresponding to $1-r$) of the ULAE charge is proportional to the claims settlement (measured by the payout pattern).

Definition 5.17 (*Paid-to-paid ratio*) We define for $1 \le k \le I + J$

$$\pi_k = \frac{X_k^{(ULAE)}}{X_k} = \frac{\sum_{i+j=k} X_{i,j}^{(ULAE)}}{\sum_{i+j=k} X_{i,j}}. \tag{5.102}$$

π_k is called paid-to-paid ratio and it measures the ULAE payments relative to the pure claims payments in each accounting year k.

Lemma 5.18 *Assume there exists a* $\pi > 0$ *such that for all accident years* $1 \le i \le I$ *we have*

$$\frac{C_{i,J}^{(ULAE)}}{C_{i,J}} = \pi. \tag{5.103}$$

Under Assumption 5.15 and Model 5.16 we have for all accounting years $J < k \le I + J$

$$\pi_k = \pi, \tag{5.104}$$

whenever $C_{i,J}$ *is constant in* i.

Proof of Lemma 5.18. We have for all $J < k \le I + J$ (we also use $I > J$)

$$\pi_k = \frac{\sum_{i+j=k} X_{i,j}^{(ULAE)}}{\sum_{i+j=k} X_{i,j}} = \frac{\sum_{j=0}^{J} \big(r\,\beta_j + (1-r)\,\gamma_j\big)\, C_{k-j,J}^{(ULAE)}}{\sum_{j=0}^{J} \gamma_j C_{k-j,J}}$$
$$= \pi\, \frac{\sum_{j=0}^{J} \big(r\,\beta_j + (1-r)\,\gamma_j\big)\, C_{k-j,J}}{\sum_{j=0}^{J} \gamma_j C_{k-j,J}} = \pi. \tag{5.105}$$

This finishes the proof. □

We define the following split of the outstanding liabilities for accident year $1 \le i \le I$ at development period $0 \le j \le J-1$:

$$R_{i,j} = \sum_{\ell=j+1}^{J} X_{i,\ell} = \sum_{\ell=j+1}^{J} \gamma_\ell\, C_{i,J} \qquad \text{(total future pure payments)},$$

$$R_{i,j}^{(IBNYR)} = \sum_{\ell=j+1}^{J} \beta_\ell\, C_{i,J} \qquad \text{(future pure payments for IBNYR claims)},$$

$$R_{i,j}^{(RBNS)} = R_{i,j} - R_{i,j}^{(IBNYR)} \qquad \text{(future pure payments for RBNS claims)}.$$

Result 5.19 *(***New York-method***) Under Model 5.16 and the assumptions of Lemma 5.18 we determine π from accounting year observations π_k, $k > J$, see* (5.104). *The reserves for ULAE charges for accident year i after development year j, $R_{i,j}^{(ULAE)} = \sum_{\ell=j+1}^{J} X_{i,\ell}^{(ULAE)}$, are determined by*

$$R_{i,j}^{(ULAE)} = \pi\, r\, R_{i,j}^{(IBNYR)} + \pi\,(1-r)\, R_{i,j}$$
$$= \pi\, R_{i,j}^{(IBNYR)} + \pi\,(1-r)\, R_{i,j}^{(RBNS)}. \tag{5.106}$$

Explanation of Result 5.19. We have under Model 5.16 and the assumptions of Lemma 5.18 for all $1 \le i \le I$ and $0 \le j \le J$

$$R_{i,j}^{(ULAE)} = \sum_{\ell=j+1}^{J} \big(r\,\beta_\ell + (1-r)\,\gamma_\ell\big)\, C_{i,J}^{(ULAE)} \tag{5.107}$$
$$= \pi \sum_{\ell=j+1}^{J} \big(r\,\beta_\ell + (1-r)\,\gamma_\ell\big) C_{i,J}$$
$$= \pi\, r\, R_{i,j}^{(IBNYR)} + \pi\,(1-r)\, R_{i,j}. \qquad \square$$

Remarks.

- In practice, one assumes stationarity condition $\pi_k = \pi$ for all $k > J$. This implies that π can be determined from the accounting data of the annual profit-and-loss statements, see (5.102). Pure claims payments are directly contained in the profit-and-loss statements, whereas ULAE payments are often contained in the administrative expenses position. Hence, one needs to divide this position into further subpositions (e.g. with the help of an activity-based cost allocation split).

- Result 5.19 gives an easy formula for determining ULAE reserves. If we are interested in the total ULAE reserves after accounting year $k > J$ we simply have

$$R_k^{(ULAE)} = \sum_{i+j=k} R_{i,j}^{(ULAE)} = \pi \sum_{i+j=k} R_{i,j}^{(IBNYR)} + \pi\,(1-r) \sum_{i+j=k} R_{i,j}^{(RBNS)}, \tag{5.108}$$

  i.e. all we need to know is, how to split the total pure claims reserves into reserves for IBNYR claims and reserves for RBNS claims.
- The assumptions for the New York-method are rather restrictive in the sense that the ultimate pure claim $C_{i,J}$ must be constant in i (see Lemma 5.18). Otherwise the paid-to-paid ratios π_k for different accounting years k are not the same as the ratio $C_{i,J}^{(ULAE)}/C_{i,J}$ even if the latter is assumed to be constant. Of course, in practice, the assumption of equal ultimate pure claim is never fulfilled. If we relax this condition we obtain the following lemma.

Lemma 5.20 *Assume there exists a $\pi > 0$ such that for all accident years $1 \le i \le I$ we have*

$$\frac{C_{i,J}^{(ULAE)}}{C_{i,J}} = \pi \left(r\,\frac{\overline{\beta}}{\overline{\gamma}} + (1-r) \right)^{-1}, \tag{5.109}$$

with, for $k > J$,

$$\overline{\gamma} = \frac{\sum_{j=0}^{J} \gamma_j\, C_{k-j,J}}{\sum_{j=0}^{J} C_{k-j,J}} \qquad \text{and} \qquad \overline{\beta} = \frac{\sum_{j=0}^{J} \beta_j\, C_{k-j,J}}{\sum_{j=0}^{J} C_{k-j,J}}. \tag{5.110}$$

Under Assumption 5.15 and Model 5.16 we have for all accounting years $k > J$

$$\pi_k = \pi. \tag{5.111}$$

Proof of Lemma 5.20. As in Lemma 5.18 we obtain for $k > J$

$$\pi_k = \pi \left(r\,\frac{\overline{\beta}}{\overline{\gamma}} + (1-r) \right)^{-1} \frac{\sum_{j=0}^{J} \big(r\,\beta_j + (1-r)\,\gamma_j\big)\, C_{k-j,J}}{\sum_{j=0}^{J} \gamma_j\, C_{k-j,J}} = \pi. \tag{5.112}$$

This finishes the proof. □

Remarks.

- If all ultimate pure claims are equal then we obtain $\overline{\gamma} = \overline{\beta} = 1/J$ (and we apply Lemma 5.18).
- Assume that there exists a constant $\kappa > 0$ such that for all $i \geq 1$ we have $C_{i+1,J} = (1+\kappa)\, C_{i,J}$, i.e. we assume a constant growth rate κ. If we blindly apply (5.104) of Lemma 5.18 (i.e. if we do not apply the correction factor in (5.109)) and determine the incremental ULAE payments by (5.106) and (5.108) we obtain for $k > J$ (we indicate application of (5.104) by $\stackrel{\triangle}{=}$)

$$\begin{aligned}
\sum_{i+j=k} X_{i,j}^{(ULAE)} &= \sum_{j=0}^{J} \big(r\beta_j + (1-r)\gamma_j\big) C_{k-j,J}^{(ULAE)} \\
&\stackrel{(5.103)}{=} \pi \sum_{j=0}^{J} \big(r\beta_j + (1-r)\gamma_j\big) C_{k-j,J} \\
&\stackrel{\triangle}{=} \frac{X_k^{(ULAE)}}{X_k} \sum_{j=0}^{J} \big(r\, \beta_j + (1-r)\, \gamma_j\big)\, C_{k-j,J} \qquad (5.113)\\
&= \sum_{i+j=k} X_{i,j}^{(ULAE)} \left(r\, \frac{\overline{\beta}}{\overline{\gamma}} + (1-r) \right) \\
&= \sum_{i+j=k} X_{i,j}^{(ULAE)} \left(r\, \frac{\sum_{j=0}^{J} \beta_j\, (1+\kappa)^{k-j}}{\sum_{j=0}^{J} \gamma_j\, (1+\kappa)^{k-j}} + (1-r) \right) \\
&> \sum_{i+j=k} X_{i,j}^{(ULAE)},
\end{aligned}$$

where the last inequality typically holds true for $\kappa > 0$, since usually the reporting pattern $(\beta_j)_j$ develops faster than the cash flow pattern $(\gamma_j)_j$, i.e. we usually have for $J > 1$

$$\sum_{\ell=0}^{j} \beta_\ell > \sum_{\ell=0}^{j} \gamma_\ell \quad \text{for } j = 0, \ldots, J-1. \qquad (5.114)$$

This comes from the fact that the claims are reported before they are paid. Thus, if we blindly apply the New York-method for a constant positive growth rate κ then the ULAE reserves are too high because $\stackrel{\triangle}{=}$ in (5.113) is not an identity in that case (for a constant negative growth rate we obtain the opposite sign). This implies that we always have a positive loss experience on ULAE reserves for constant positive growth rates under (5.113).

Example 5.2 We assume that the observations for π_k are generated by i.i.d. random variables $X_k^{(ULAE)}/X_k$. Hence, we can estimate π from this sequence. Assume

$\pi = 10\,\%$. Moreover, $\kappa = 0$ and set $r = 50\,\%$ (this is the usual choice made in practice). Furthermore, we assume that we have the following reporting and cash flow patterns ($J = 4$):

$$(\beta_0, \ldots, \beta_4) = (90\,\%, 10\,\%, 0\,\%, 0\,\%, 0\,\%), \tag{5.115}$$

$$(\gamma_0, \ldots, \gamma_4) = (30\,\%, 20\,\%, 20\,\%, 20\,\%, 10\,\%). \tag{5.116}$$

Assume that $C_{i,J} = 1,000$. Then the ULAE reserves for accident year i are given by (note that we need to choose $j = -1$ to describe the total ULAE)

$$\left(R_{i,-1}^{(ULAE)}, \ldots, R_{i,3}^{(ULAE)}\right) = (100, 40, 25, 15, 5), \tag{5.117}$$

which implies for the incremental ULAE payments

$$\left(X_{i,0}^{(ULAE)}, \ldots, X_{i,4}^{(ULAE)}\right) = (60, 15, 10, 10, 5). \tag{5.118}$$

Hence for the total incremental payments $X_{i,j}^{+} = X_{i,j} + X_{i,j}^{(ULAE)}$ we have

$$\left(X_{i,0}^{+}, \ldots, X_{i,4}^{+}\right) = (360, 215, 210, 210, 105). \tag{5.119}$$

□

5.5 Conclusions on the Non-life VaPo

We have constructed both the VaPo and the VaPo protected against insurance technical risks for a run-off portfolio of a non-life insurance company. Our approach should be considered as a first approximation to a fully consistent construction.
Open problems are:

- Make an appropriate choice of the financial basis such that we have an independent decoupling into insurance technical risks and financial risks. In fact, this is a very difficult task because claims inflation and accounting year effects may substantially increase the uncertainties, and induce undesirable calendar year dependence.
- Make an appropriate choice of the risk measure which also takes into account the tails of the CDRs. Moreover, we have not considered time value of money in the CDR analysis.
- Make an appropriate economic choice for the cost-of-capital rate r_{CoC}. All choices used in practice are rather ad-hoc.
- Choose an appropriate stochastic claims reserving model in order to determine claims reserves, cash flow patterns, uncertainties in the estimates and predictions, etc. In general, one has different sources of information, e.g. claims payments, claims incurred information, other internal information, expert knowledge, etc.

Most claims reserving methods are not able to cope with all of these different information channels simultaneously. The PIC model is one approach that treats claims payments, claims incurred and prior knowledge information simultaneously, see Merz–Wüthrich [MW10].

- Here we have only treated the run-off situation of a non-life insurance portfolio. The premium liability risk could, theoretically, also be put into our framework, by assuming that $C_{I+1,-1} = -\Pi_{I+1}$ and then applying the CL method to this extended model. However, this approach often does not lead to good estimates for the premium liabilities and premium liability risks. We therefore recommend to rather treat the premium liability risk in a separate model (as is done in almost all risk-adjusted solvency calculations).

  In this separate model, premium liability risks are often split into two categories: (i) small (attritional) claims, (ii) large claims and cumulative events (such as hailstorms, flood events, etc.) (see e.g. SST [SST06] and Gisler [Gi09]). The main risk driver in (i) is the calibration of future model parameters which may have large uncertainties. This can be modelled assuming that the true future model parameters are latent variables which we try to predict (see e.g. Wüthrich [Wü06]).

  The risk drivers in (ii) are often modelled using a compound model (for low frequencies and high severities). One main difficulty in the calibration here is that one usually has only little information. We propose to use internal data, external data and expert opinion for the calibration of these models, for instance, using Bayesian and credibility theory.

  If one models premium liability risks and claims run-off risks separately one should, however, keep in mind that there might be inconsistencies over time because one switches from one model to another. This is highlighted in Ohlsson–Lauzeningks [OL09].

References

[AISAM07] AISAM-ACME (2007). AISAM-ACME study on non-life long tail liabilities. Reserve risk and risk margin assessment under Solvency II. October 17, 2007.

[BBK04] Baumgartner, G., Bühlmann, H., Koller, M. (2004). Multidimensional valuation of life insurance policies and fair value. *Bulletin of the Swiss Association of Actuaries* **2004/1**, 27–64.

[BBCMP11] Besar, D., Booth, P., Chan, K.K., Milne, A.K.L., Pickles, J. (2011). Systemic risk in financial services. *British Actuarial Journal* **16/2**, 195–300.

[BG06] Böhm, H., Glaab, H. (2006). Modellierung des Kalenderjahr-Risikos im additiven und multiplikativen Schadenreservierungsmodell. Talk at the German ASTIN Colloquium 2006.

[BM06] Brigo, D., Mercurio, F. (2006). *Interest Rate Models – Theory and Practice*. 2nd Edition, Springer.

[BBMW06] Buchwalder, M., Bühlmann, H., Merz, M., Wüthrich, M.V. (2006). Estimation of unallocated loss adjustment expenses. *Bulletin of the Swiss Association of Actuaries* **2006/1**, 43–53.

[Bü80] Bühlmann, H. (1980). An economic premium principle. *ASTIN Bulletin* **11/1**, 52–60.

[Bü92] Bühlmann, H. (1992). Stochastic discounting. *Insurance: Mathematics and Economics* **11/2**, 113–127.

[Bü95] Bühlmann, H. (1995). Life insurance with stochastic interest rates. In: *Financial Risk in Insurance*, G. Ottaviani (Ed.), Springer.

[BDES98] Bühlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A. N. (1998). On the Esscher transform in discrete finance models. *ASTIN Bulletin* **28/2**, 171–186.

[CAS90] Casualty Actuarial Society (CAS) (1990). Foundations of Casualty Actuarial Science. 4th Edition.

[dFM03] De Felice, M., Moriconi, F. (2003). Risk based capital in P&C loss reserving or stressing the triangle. Research Group on "Insurance Companies and Pension Funds", Working Paper n. 1, Rome, December 2003.

[dFM06] De Felice, M., Moriconi, F. (2006). Process error and estimation error of year-end reserve estimation in the distribution free chain-ladder model. *Alef Working Paper*.

[DS94] Delbaen, F., Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen* **300**, 463–520.

[DS06] Delbaen, F., Schachermayer, W. (2006). *The Mathematics of Arbitrage*. Springer.

[DLH16] Diers, D., Linde, M., Hahn, L. (2016). Quantification of multi-year non-life insurance risk in chain ladder reserving models. *Insurance: Mathematics and Economics* **67**, 187–199.

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1

[DE10] Donnelly, C., Embrechts, P. (2010). The devil is in the tails: actuarial mathematics and the subprime mortgage crisis. *ASTIN Bulletin* **40/1**, 1–33.

[Du96] Duffie, D. (1996). *Dynamic Asset Pricing Theory*. 2nd Edition. Princeton University Press.

[EK99] Elliott, R.J., Kopp, P.E. (1999). *Mathematics of Financial Markets*. Springer.

[En08] England, P.D. (2008). Private communication.

[EV02] England, P.D., Verrall, R.J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal* **8/3**, 443–518.

[Fe03] Feldblum, S. (2003). Completing and using schedule P. Casualty Actuarial Society (CAS).

[Fi09] Filipović, D. (2009). *Term-Structure Models. A Graduate Course*. Springer.

[FZ02] Filipović, D., Zabczyk, J. (2002). Markovian term structure models in discrete time. *Annals of Applied Probability* **12/2**, 710–729.

[FS11] Föllmer, H., Schied, A. (2011). *Stochastic Finance: An Introduction in Discrete Time*. 3rd Edition. De Gruyter.

[GR10] Geneva Reports (2010). Anatomy of the credit crisis. In: *Risk and Insurance Research* **3**, P.M. Liedtke (Ed.), January 2010.

[GP98] Gerber, H.U., Pafumi, G. (1998). Utility functions: from risk theory to finance. *North American Actuarial Journal* **2/3**, 74–100.

[Getal15] Gesmann, M., Carrato, A., Concina, F., Murphy, D., Wüthrich, M.V., Zhang, W. (2015). Claims reserving with R: ChainLadder-0.2.2 package vignette. Version August 31, 2015.

[Gi09] Gisler, A. (2009). The insurance risk in the SST and in Solvency II: modelling and parameter estimation. *SSRN Manuscript* ID 2704364, Version June 4, 2009.

[HMW15] Happ, S., Merz, M., Wüthrich, M.V. (2015). Best-estimate claims reserves in incomplete markets. *European Actuarial Journal* **5/1**, 55–77.

[IAA04] IAA Insurer Solvency Assessment Working Party (2004). A global framework for insurer solvency assessment. International Actuarial Association IAA, Draft January 9, 2004.

[Ing87] Ingersoll, J.E. (1987). *Theory of Financial Decision Making*. Rowman and Littlefield Publishers.

[IAIS05] International Association of Insurance Supervisors (IAIS). Glossary of terms, February 2005.

[JSV01] Jarvis, S., Southall, F., Varnell, E. (2001). Modern valuation techniques. Presented at the Staple Inn Actuarial Society, February 6, 2001.

[JYC09] Jeanblanc, M., Yor, M., Chesney, M. (2009). *Mathematical Methods for Financial Markets*. Springer.

[LV08] Laeven, L., Valencia, F. (2008). Systemic banking crisis: a new database. *IMF Working Paper* WP/08/224.

[LL91] Lamberton, D., Lapeyre, B. (1991). *Introduction au Calcul Stochastique Appliqué à la Finance*. Mathématiques et Applications. SMAI no. 9, Ellipses-Edition.

[Ma93] Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin* **23/2**, 213–225.

[MTW08] Malamud, S., Trubowitz, E., Wüthrich, M.V. (2008). Market consistent pricing of insurance products. *ASTIN Bulletin* **38/2**, 483–526.

[Ma78] Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance* **33/1**, 177–186.

[MFE15] McNeil A.J., Frey, R., Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton University Press.

[MW07] Merz, M., Wüthrich, M.V. (2007). Prediction error of the expected claims development result in the chain ladder method. *Bulletin of the Swiss Association of Actuaries* **2007/1**, 117–137.

[MW08] Merz, M., Wüthrich, M.V. (2008). Modelling the claims development result for solvency purposes. *CAS E-Forum* **Fall 2008**, 542–568.

[MW10] Merz, M., Wüthrich, M.V. (2010). Paid-incurred chain claims reserving method. *Insurance: Mathematics and Economics* **46/3**, 568–579.

[MW14] Merz, M., Wüthrich, M.V. (2014). Claims run-off uncertainty: the full picture. *SSRN Manuscript* ID 2524352, Version July 3, 2015.

[OL09] Ohlsson, E., Lauzeningks, J. (2009). The one-year non-life insurance risk. *Insurance: Mathematics and Economics* **45/2**, 203–208.

[Pe10] Pelsser, A. (2010). Time-consistent and market-consistent actuarial valuations. *SSRN Manuscript* ID 1551323, Version February 11, 2010.

[Rö16] Röhr, A. (2016). Chain-ladder and error propagation. *ASTIN Bulletin* **46/2**, 293–330.

[SW10] Salzmann, R., Wüthrich, M.V. (2010). Cost-of-capital margin for a non-life insurance liability runoff. *ASTIN Bulletin* **40/2**, 415–451.

[Sa06] Sandström, A. (2006). *Solvency: Models, Assessment and Regulation*. Chapman & Hall/CRC.

[Sa07] Sandström, A. (2007). Solvency – a historical review and some pragmatic solutions. *Bulletin of the Swiss Association of Actuaries* **2007/1**, 11–34.

[Sha02] Sharma, P. (2002). Report on "Prudential supervision of insurance undertakings". Conference of the Insurance Supervisory Services of the Member States of the European Union, Version December, 2002.

[SNB] Statistisches Monatsheft der Schweizerischen Nationalbank SNB. Available under http://www.snb.ch

[SST06] Swiss Solvency Test (2006). FINMA SST Technisches Dokument, Version 2. October 2006.

[TA83] Taylor, G.C., Ashe, F.R. (1983). Second moments of estimates of outstanding claims. *Journal of Econometrics* **23/1**, 37–61.

[Va77] Vasiček, O. (1977). An equilibrium characterization of the term structure. *Journal Financial Economics* **5/2**, 177–188.

[Wi91] Williams, D. (1991). *Probability with Martingales*. Cambridge.

[Wü06] Wüthrich, M.V. (2006). Premium liability risks: modeling small claims. *Bulletin of the Swiss Association of Actuaries* **2006/1**, 27–38.

[Wü13] Wüthrich, M.V. (2013). Non-Life Insurance: Mathematics & Statistics. *SSRN Manuscript* ID 2319328, Version April 14, 2016.

[WB08] Wüthrich, M.V., Bühlmann, H. (2008). The one-year runoff uncertainty for discounted claims reserves. *Giornale dell Istituto Italiano degli Attuari* **LXXI**, 1–37.

[WET11] Wüthrich, M.V., Embrechts, P., Tsanakas, A. (2011). Risk margin for a non-life insurance run-off. *Statistics & Risk Modeling* **28/4**, 299–317.

[WM08] Wüthrich, M.V., Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. Wiley.

[WM13] Wüthrich, M.V., Merz, M. (2013). *Financial Modeling, Actuarial Valuation and Solvency in Insurance*. Springer.

[WM15] Wüthrich, M.V., Merz, M. (2015). Stochastic Claims Reserving Manual: Advances in Dynamic Modeling. *SSRN Manuscript* ID 2649057, Version August 21, 2015.

[WML09] Wüthrich, M.V., Merz, M., Lysenko, N. (2009). Uncertainty in the claims development result in the chain ladder method. *Scandinavian Actuarial Journal* **109/1**, 63–84.

Index

M.V. Wüthrich, *Market-Consistent Actuarial Valuation*, EAA Series,
DOI 10.1007/978-3-319-46636-1

V

W

Y

Z