Stochastic Claims Reserving under Consideration of Various Different Sources of Information

Dissertation

**Zur Erlangung der Würde
des Doktors der Wirtschaftswissenschaften
der Universität Hamburg**

vorgelegt von

Sebastian Happ

geb. am 21.01.1983 in Tübingen

Hamburg, Juli 2014

Vorsitzender: Prof. Dr. Bernhard Arnold (Universität Hamburg)

Erstgutachter (Supervisor): Prof. Dr. Michael Merz (Universität Hamburg)

Zweitgutachter (Co-Supervisor): Prof. Dr. Mario V. Wüthrich (ETH Zürich)

Datum der Disputation: 05.06.2014

Acknowledgements

During my diploma studies in mathematics and business administration at the Eberhard Karls Universität Tübingen Prof. Dr. Michael Merz started teaching at university as an assistant professor at the department of business administration. This gave me the chance to attend his lecture series on selected topics in actuarial science. Visiting these lectures I got a fundamental insight in practical problems in quantitative risk management and insurance and how they can be approached by mathematical statistical concepts. This awakened my deep interest in actuarial science. At this point I would like to express my deepest gratitude to my supervisor Michael Merz for giving me the chance to pursue a PhD at the faculty of business administration in Hamburg. Not only did he support me in scientific issues but also in personal matters.

I am deeply grateful to my co-supervisor Prof. Dr. Mario V. Wüthrich from ETH Zürich for his great support and invaluable advice in our joint contributions. He permanently supported me with his vast knowledge and experience in actuarial science.

Moreover, I would like to thank my co-author René Dahms for his valuable collaboration. My thanks also go to the whole team, namely T. Gummersbach, J. Heberle, Nha-Nghi Huynh, A. Johannssen, A. Ruiz-Merino and A. Thomas at the chair of mathematics and statistics in business administration at Universität Hamburg under the administration and supervision of Michael Merz for its support. I would like to thank Marco Breitig for his valuable discussions.

Finally, I thank my wife Svetlana for her confidence and support in all those years.

Sebastian Happ

Contents

List of Figures

List of Tables

1 Introduction

Recent developments in (financial) markets have shown that unexpected negative events may have a tremendous impact on a wide range of financial institutions such as banks, funds, investment and insurance companies. Often such events are followed by serious problems ranging from economic depression with high unemployment rates, a decrease in common wealth and bad medical maintenance to social riots. Governmental authorities and regulatory institutions have been established to adopt and develop regulatory frameworks for the financial industry, in order to reduce negative impact of such events and to avoid collateral damage on other parts of the economy in the future.

In Germany the *Federal Financial Supervisory Authority* (BaFin) supervises banks, financial services provider, insurance companies as well as securities trading. Moreover, in response to the financial crisis 2007–2008 the *European Union* (EU) created the European System of Financial Supervision, which consists of three European Supervisory Authorities:

1. *European Banking Authority* (EBA) for the European banking sector
2. *European Insurance and Occupational Pensions Authority* (EIOPA) for the insurance sector
3. *European Securities and Markets Authority* (ESMA) for securities trading

For the banking sector the corresponding regulatory framework called Basel II was developed by the Basel Committee on Banking Supervision and is currently replaced by its successor Basel III. Insurance companies in Europe are controlled by the regulatory framework called Solvency II. In Switzerland the regulation of all financial institutions including insurance companies is provided by the *Swiss Financial Markets Authority* (FINMA) with the corresponding regulation frameworks Basel II and Basel III for the banking sector and the *Swiss Solvency Test* (SST) for the insurance industry.

For an insurance company there are two ways a regulatory framework can be looked at.

a) **From the perspective of investors and the management**: The function and the existence of the company must be maintained in the mid/long term run to generate earnings for the investors and the management. Moreover, these earnings should be maximized (profit maximization).

b) **From the perspective of regulatory authorities**: Financial liquidity of the insurance company must be provided even in times of extreme financial distress and phases of an extraordinary accumulation of claim compensation payments. The ability of the insurer to pay losses has to be maintained in almost all realistic scenarios to prevent losses for the policyholders and to eliminate wide-ranging negative effects on the whole economy.

Similar to Basel II, the Solvency II regulatory framework is subdivided into three main pillars to incorporate the main ideas of the regulatory authorities' point of view:

- **Pillar I: Minimum Standard and Implementation**
  - Market consistent valuation of assets and liabilities
  - Internal models, best-estimate reserves, technical provisions, solvency capital requirements, target capital and own funds
- **Pillar II: Supervisor Review and Control**
  - Group supervision
  - Supervisory review process
  - Governance
- **Pillar III: Disclosure**
  - Supervisory transparency
  - Accountability
  - Reporting and disclosure

For details on the technical standards, further guidelines and information, see the Website of EIOPA[1]. For the basic structure of the SST we refer to the Website of FINMA[2].

In this thesis we focus on the the first pillar. Moreover, one has to distinguish between life and non-life insurance business, since the contract specifications, risk drivers and payoff patterns and hence the methodologic means of approaching and modeling life and non-life contract liabilities differ substantially. For an illustration of this fact we refer to the examples given in Chapter 7 in WÜTHRICH–MERZ [62]. An introduction on stochastic models in life insurance can be found in GERBER [26] and KOLLER [35]. It is crucial to keep in mind that from now on throughout the thesis we will strictly deal with non-life insurance business.

The first pillar in non-life insurance has been subject to many quantitative scientific studies, see WÜTHRICH–MERZ [63] and [62] for an overview, since it is directly associated with the

[1] `https://eiopa.europa.eu/activities/insurance/solvency-ii`
[2] `http://www.finma.ch/archiv/bpv/e/themen/00506/index.html?lang=en`

problem of the management and quantification of (random) future cash flows. These cash flows typically arise from assets and claims payments, see Wüthrich–Merz [62]. The corresponding field of study to analyze (random) risk outcomes and associated loss liability cash flows in insurance with mathematical and statistical methods is called actuarial science. Actuarial science comprises the following aspects:

1. Evaluation of (random) outstanding loss liability cash flows and setting-up of sufficient reserves to meet these liabilities
2. Evaluation of assets and its associated risk
3. Level of premiums in policies
4. Reinsurance
5. Asset liability management (ALM) comprising all previous aspects

All stated aspects have an impact on the process of future cash flows and are therefore crucial for management purposes in insurance companies. That means that actuarial science is directly associated with the central problem in insurance companies of predicting future cash flows. Therefore, the crucial task and main goal of actuaries is the

prediction of (random) future cash flows.

Among the five aspects stated above we focus in this thesis on the first aspect, i.e. the field of predicting future outstanding loss liabilities. In actuarial science this field is called *claims reserving*. Claims reserving belongs to the main tasks of a non-life actuary, since claims reserves are the biggest position on a balance sheet of a non-life insurance company and must therefore be predicted very precisely.

Therefore, in this doctoral thesis we will focus on the task of predicting future loss liabilities and calculating the corresponding reserves needed to cover these outstanding loss liabilities in non-life insurance companies. For this prediction problem there are often various sources of information available. Most classical claims reserving methods are very limited w.r.t. the sources of information they can incorporate. We present in this thesis two powerful models which can cope with several sources of information in a mathematically consistent way. The first model generalizes most widely used distribution-free claims reserving methods. This provides a new perspective and new possibilities for distribution-free claims modeling and is subject to Part II of this thesis. The second method is an important representative of the class of distributional claims reserving methods which can cope with two different data sources often available in insurance practice. This is subject to Part III. The thesis is closed up by Part IV discussing some central aspects of claims reserving under new solvency requirements like Solvency II or SST.

Outline

This thesis is divided into four parts:

Part I: In the first part (Chapter 2) the classical claims reserving problem is introduced. We consider the associated general prediction problem and point out which data sources have been used in classical as well as in state-of-the-art claims reserving methods for the prediction of future loss liabilities. Moreover we show how the incorporated prediction uncertainty is classically quantified in long term as well as in short term risk considerations.
Part II: In the second part (Chapters 3 and 4) we briefly present widely used classical claims reserving methods. Following Dahms [17] and Dahms–Happ [15] all these methods are then merged in a general state-of-the-art distribution-free claims reserving framework in Chapter 4. This model framework comprises almost all distribution-free claims reserving methods. Moreover, it allows for the incorporation of various sources of information for the prediction process and hence provides a new perspective and possibilities of distribution-free claims reserving.
Part III: In contrary to Part II this part is subject to distributional claims reserving. In the model class of distributional claims reserving methods we consider in Chapters 5 and 6 an important representative, the *paid-incurred chain* (PIC) reserving method presented in Merz–Wüthrich [46]. Following Happ et al. [30] and Happ–Wüthrich [31] we consider for this method the quantification of the one-year reserving risk and generalize the classical PIC method so that dependence structures in the data can be appropriately captured. Moreover, the whole predictive distribution of the claims development result is derived via *Monte-Carlo* (MC) methods.
Part IV: In this part (Chapter 7) we point out central regulatory requirements included in recent solvency frameworks like SST or Solvency II. These solvency requirements are not coherent with most classical claims reserving methods. We point out simplification methods proposed in the SST and show how they make most claims reserving methods accessible for these solvency requirements. We close up this part by presenting an example where reserves are calculated regarding the SST reserving requirements.

2 Reserving Problem

2.1 Insurance Contracts and Process of Claims Settlement

An insurance contract is an agreement of two parties: For a fixed payment (insurance premium) the insurer (insurance company) obliges to pay a financial compensation to the insured (policyholder) in the case of an occurrence of some well defined (random) future event in a well specified time period. In the case of such an event at a certain date (occurrence date) during the insured period, the insured person reports the claim to the insurance company at the so-called reporting (notification) date. The time between the occurrence and the reporting date is called reporting delay. After the reporting of the claim the insurance company verifies whether all insurance contract specifications are fulfilled so that the insurer has to provide coverage of the claim. If this is the case, the insurance company starts payments for the financial compensation of the claim in accordance to the contract specifications. This claims settlement process typically consists of one or more payments to the policyholder. It ends with the closure date where no further claims payments are expected and the claim is (presumably) completely settled and closed. The time line of typical non-life insurance claims from occurrence to the final settlement is illustrated in Figure 2.1. Time delays from occurrence to notification and from settlement process to the

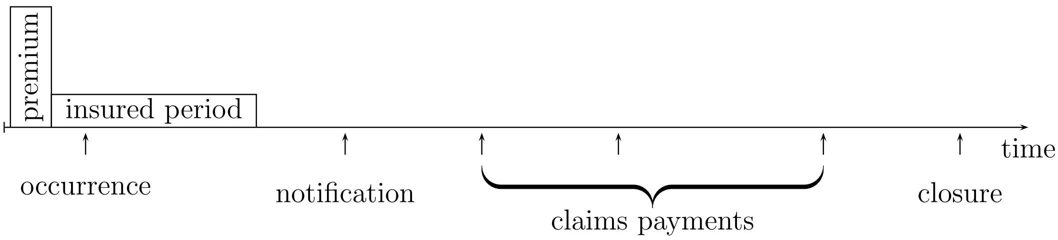

Figure 2.1: Generic time line of the claims settlement process

closure date are typical for non-life insurance claims and can be caused by different reasons:

- Delays when incurred claim events are not immediately reported to the insurance company
- Final claim amounts are determined over a long period of time (up to several decades)
- Juridical inspection of a claim. The liability of the insurance company to pay for the claim

is to be determined

- Court decisions leading to payment adjustments, reverse transactions of already paid compensations or additional claims payments

These time delays often lead to a very slow claims settlement process with claims payments far in the future (up to several decades). This shows that the very nature of insurance business (i.e. underwriting risks through insurance contracts) often causes a very slow settlement process and the prediction of this process becomes a central point of interest. For a more detailed discussion on that topic, see Wüthrich–Merz [63].

General Remark:

In non-life insurance business many claim characteristics (occurrence date, frequency of claims, severity of a claim, claim settlement pattern, claims payments, etc.) are subject to randomness and can not be predicted without uncertainty. Hence, probability theory and statistics provide suitable mathematical tools for dealing with those claim characteristics. Thereby, it is assumed that the very nature and the behavior of these claim characteristics do not change "too fast" over time. This assumption is required to utilize past observations for predicting purposes and to reveal systematic properties (behavior) of the quantities under consideration. For this reason we model all quantities of interest in a stochastic framework as random variables, which are defined on a common probability space $(\Omega, \mathcal{D}, \mathbb{P})$.

2.2 Data Basis in a Non-life Insurance Company

In general, insurance companies group policies (insurance contracts) with similar risk characteristics or comparable contract specifications into sufficiently homogeneous insurance portfolios. This is often done by *Lines of Business* (LoB), but can be subdivided further into smaller units. Typical LoBs are: Motor third party, product liability, private and commercial property, commercial liability, health insurance, etc. An insurance company has to put provisions aside, in order to cover future loss liabilities arising from these grouped insurance portfolios. For this reason an accurate prediction of future loss liabilities and the associated cash flows in the claim settlement process is of central interest. This prediction can be based on various sources of information.

2.2.1 Classical View

In the classical view the prediction of future loss liabilities is often based on the information of the past observed development of the settlement process itself. Classical claims reserving literature often assumes that an insurance company has, after grouping of individual contracts, $M \geq 1$ nearly homogeneous portfolios. All claims, which occur in year i, are called claims in accident year $i \in \{0, \dots, I\}$, where I is the *current* year. The number k of years between accident year and the year of the actual claims payment is called development year $k \in \{0, \dots, J\}$, with J being the total number of development years. It is usually assumed that $I \geq J$ and that all claims are completely settled in development year J, i.e. there are no claims payments beyond development year J. For models considering claims payments beyond development year J by means of so-called tail factors, see MACK [40] and MERZ–WÜTHRICH [42]. We denote all payments for accident year i and development year k in the m-th portfolio ($m \in \{1, \dots, M\}$) by $S_{i,k}^m$ and say that all claims payments $S_{i,k}^m$ with $i + k = n$ and $n \in \{0, \dots, I + J\}$ belong to accounting year n. This notation is called incremental claims representation in the actuarial literature, because we consider claims payments $S_{i,k}^m$ in accident year i and development year k of the m-th portfolio. In the actuarial literature (cf. WÜTHRICH–MERZ [63]) the cumulative claims payments representation of the claim settlement process is also used. In this representation one considers cumulated amounts in accident year i up to development year k defined by

$$C_{i,k}^m := \sum_{j=0}^{k} S_{i,j}^m, \tag{2.1}$$

where all claims payments which belong to accident year i up to development year k in the m-th portfolio are aggregated. At time $n \in \{0, \dots, I + J\}$ all claims payments $S_{i,k}^m$ with $i + k \leq n$ and $1 \leq m \leq M$ are observed and generate the σ-field

$$\begin{aligned} \mathcal{D}^n &:= \sigma \left\{ S_{i,k}^m \middle| i + k \leq n,\ 0 \leq i \leq I,\ 0 \leq k \leq J,\ 1 \leq m \leq M \right\} \\ &= \sigma \left\{ C_{i,k}^m \middle| i + k \leq n,\ 0 \leq i \leq I,\ 0 \leq k \leq J,\ 1 \leq m \leq M \right\}. \end{aligned} \tag{2.2}$$

Moreover, we denote the resulting filtration by $\mathbb{D} := (\mathcal{D}^n)_{0 \leq n \leq I+J}$ leading to the probability space with filtration $(\Omega, \mathcal{D}, \mathbb{D}, \mathbb{P})$. The two representations (incremental or cumulative representation) are commonly used in the claims reserving literature, and it mainly depends on the model choice whether the incremental or the cumulative representation is used. The settlement process of the m-th portfolio in the incremental as well as in the cumulative claims payments representation is illustrated in claims development (run-off) trapezoids where accident years $i \in \{0, \dots, I\}$ and development years $k \in \{0, \dots, J\}$ are given by the rows and the columns, respectively. This means the incremental claims payments in accident year i and development year k of the m-th portfolio are positioned in the i-th row and the k-th column in the m-th development trapezoid, see Figure 2.2.

We will see in Chapter 4 that the incremental claims payments representation is an appropriate choice for almost all distribution-free claims reserving methods. Moreover, the incremental representation is advantageous if one is interested in the valuation of outstanding loss liabili-

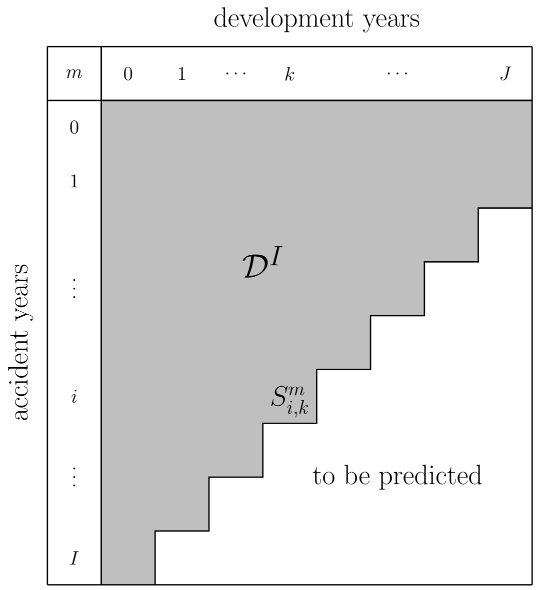

Figure 2.2: Classical view (extended view): Generic run-off trapezoid of the m-th LoB (claim information) for $m \in \{1, \dots, M\}$ and incremental claims payments (claim information) of accident year i and development year k with $i + k = I$

ties via valuation portfolios, see Wüthrich–Merz [62]. However, we switch to the cumulated claims payments representation, if helpful (Chapters 5 and 6).

2.2.2 Extended View

Beside the claim settlement process data there are often other sources of information available for the prediction of loss liabilities:

- Settlement processes of other correlated portfolios
- Data of collectives which may influence the settlement process under consideration
- Incurred losses: Claims payments plus individual case dependent loss reserves
- Prior ultimate claim estimates: This information may include pricing arguments
- Insured volume
- Number and size of contracts
- etc.

Recent publications in actuarial science consider new models which allow for including some of these sources of information in a mathematically consistent way, see for example Dahms [17] and Merz–Wüthrich [46]. In these models $S_{i,k}^m$ and $C_{i,k}^m$ do not necessarily only correspond

anymore to incremental claims payments and cumulative claims payments (i.e. information from the claim settlement process). They may also represent some other sources of information stated above, for example incurred losses data, see the PIC reserving method in Merz–Wüthrich [46], or prior ultimate claim estimates, see the *Bornhuetter–Ferguson* (BF) method in Mack [39]. Therefore, it is necessary to extend the denotation of $S_{i,k}^m$ of the classical view, since we focus in the actuarial contributions of this thesis on such new model classes, see Chapters 4–6. Throughout the thesis $S_{i,k}^m$ denotes the m-th ($m \in \{1, \dots, M\}$) claim information of accident year $i \in \{0, \dots, I\}$ and development year $k \in \{0, \dots, J\}$ and not necessarily only the claims payments as it is convenient in classical claims reserving methods. These claim information may – beside the claims payments process – contain incurred losses, see Merz–Wüthrich [46] and Dahms [16], received premium and the average loss ratio, see Bühlmann [11], prior ultimate claim estimates, see Mack [39] and Arbenz–Salzmann [6], claim volume information, see Dahms [17], or other additional sources of information.
By a slight abuse of notation we will call also $m \in \{1, \dots, M\}$ the m-th claim information by identifying the index m with its associated claim information $S_{i,k}^m$. In the extended view some claim information $S_{i,k}^m$ do not generate any loss liability cash flows in the future and thus do not have to be predicted. Therefore, we define

$$\mathcal{M} := \left\{ m \in M \middle| \; S_{i,k}^m \text{ generates loss liability cash flows} \right\}. \tag{2.3}$$

By definition $\mathcal{M}$ is the set of claim information which generate cash flows, see (2.3), and is therefore of central interest for claims reserving and risk management.

Remarks 2.1 (Set $\mathcal{M}$) *In most classical claims reserving methods, each claim information $m \in \mathcal{M}$ is given by the claims payments of an insurance portfolio of a specific LoB, see Chapter 3 for examples. However, this is not always the case. In Example 1 in* Dahms *[17] there is a claim information $m \in \mathcal{M}$ of subrogation payments. This shows that $\mathcal{M}$ may – beside the claims payments of different LoBs – also contain other claim information which also generate cash flows. That means that the claim information in $\mathcal{M}$ are not explicitly restricted to claims payments of different insurance portfolios. However, for a simpler interpretation of the set $\mathcal{M}$ one may think of each claim information $m \in \mathcal{M}$ as claims payments arising from an insurance portfolio of a certain LoB.*

As a consequence of the definition of $\mathcal{M}$, the set of all claim information $\{1, \dots, M\}$ is divided into disjoint subsets $\mathcal{M} \subseteq \{1, \dots, M\}$ and $\mathcal{M}^c = \{1, \dots, M\} \backslash \mathcal{M}$. The claim information $m \in \mathcal{M}$ have already been discussed above. The set $\mathcal{M}^c$ of claim information is not of central interest for risk management and claims reserving, because it does not generate any loss liability cash flows. However, claim information out of $\mathcal{M}^c$ are utilized in many models for the prediction of claim information $m \in \mathcal{M}$ under consideration, i.e. they contain information, which are required for the prediction of claim information $m \in \mathcal{M}$. To name only a few of them, an ultimate claim estimate (as a claim information $m \in \mathcal{M}^c$) is incorporated for the prediction of claims payments (as a claim information $m \in \mathcal{M}$) in the BF method, see Section 3.5, incurred losses are used for the prediction of claims payments in the *extended complementary loss ratio* (ECLR) method, see Dahms [16], and in the PIC reserving method, see Chapters 5 and 6, or volume measures are included for claims payments predictions in the *additive loss reserving* (ALR) method in Merz–Wüthrich [44]. In analogy to the classical view, claim information $m \in \{1, \dots, M\}$ in the extended view are also illustrated in development (run-off) trapezoids, see Figure 2.2.

Notational Convention:

Unless otherwise indicated we work in this thesis within the extended view, i.e. we assume that a set of $M \geq 1$ claim information (sources of information) is available today, i.e at time I. In this extended view all claim information $m \in \mathcal{M}$ generate loss liability cash flows and hence have to be predicted, whereas claim information $m \in \mathcal{M}^c$ are used only for the prediction of claim information $m \in \mathcal{M}$.

2.3 Prediction Problem

As mentioned in the previous section insurance companies often have various sources of information (claim information) for the prediction of future loss liabilities cash flows $S^m_{i,k}$ with $m \in \mathcal{M}$. We work in the extended view, i.e. we assume that $M \geq 1$ claim information $m \in \{1, \dots, M\}$ (as mentioned above we identify m by its corresponding claim information $S^m_{i,k}$) are available today (at time I). The set of claim information generating cash flows is denoted by $\mathcal{M} \subseteq \{1, \dots, M\}$. A reserving actuary has to predict today (at time I) and at all future times up to the final run-off, i.e. at times $n \in \{I, \dots, I+J-1\}$, the outstanding loss liability cash flows. These are given for claim information $m \in \mathcal{M}$ and accident year $i \in \{I-J+1, \dots, I\}$ at time $n \in \{I, \dots, I+J-1\}$ by (an empty sum is defined by zero)

$$\mathcal{R}_i^{m|n} := \sum_{j=n-i+1}^{J} S^m_{i,j}. \tag{2.4a}$$

By summation of (2.4a) over all claim information of interest, i.e. $m \in \mathcal{M}$, we obtain the aggregated outstanding loss liabilities of accident year i given by

$$\mathcal{R}_i^{n} := \sum_{m \in \mathcal{M}} \mathcal{R}_i^{m|n} = \sum_{m \in \mathcal{M}} \sum_{j=n-i+1}^{J} S^m_{i,j} \tag{2.4b}$$

and the aggregated outstanding loss liabilities for several accident years are given by

$$\mathcal{R}^n := \sum_{i=n-J+1}^{I} \mathcal{R}_i^n = \sum_{i=n-J+1}^{I} \sum_{m\in\mathcal{M}} \sum_{j=n-i+1}^{J} S_{i,j}^m. \tag{2.4c}$$

It mainly depends on the situation which of the quantities in (2.4a)–(2.4c) is of interest. In an accounting view the actuary often considers at time $n \in \{I, \ldots, I+J-1\}$ the aggregated outstanding loss liability cash flows $\mathcal{R}^n$ given by (2.4c). However, in some situations a more detailed analysis at time n of the outstanding loss liabilities of a specific claim information $m \in \mathcal{M}$ and a certain accident year i in (2.4a) is required. One such situation is that most classical claims reserving model frameworks consider claims payments (in these models only claims payments are considered and hence we speak about claims payments instead of loss liability cash flows) on the level of each individual claim information of a specific accident year (cf. Wüthrich–Merz [63]). The aggregated claims payments $\mathcal{R}^n$ are then derived by aggregation over different individual claim information as in (2.4c).

In order to build up sufficient reserves for outstanding loss liabilities an insurance company is obliged to predict precisely all outstanding loss liabilities, based on information $\mathcal{D}^n$ in (2.2) available at time n of prediction. In most well-known claims reserving methods predictors for the incremental claim information $S_{i,k}^m$ for $i+k>n$ and $m \in \mathcal{M}$ are derived. This is described for some well-known claims reserving methods in Chapter 3 and in a more general model framework in Chapter 4. Throughout this thesis we will denote those predictors for $S_{i,k}^m$ based on the data $\mathcal{D}^n$ at time n by $\widehat{S}_{i,k}^{m|n}$. At time $n \in \{I, \ldots, I+J-1\}$ the outstanding loss liabilities in (2.4a)–(2.4c) consist of (sums of) incremental claim information $S_{i,k}^m$ with $i+k>n$ and $m \in \mathcal{M}$. Therefore, the prediction of these loss liabilities is equivalent to the prediction of incremental claim information $S_{i,k}^m$ for $i+k>n$ and $m \in \mathcal{M}$. For the resulting predictors for $\mathcal{R}_i^{m|n}$, $\mathcal{R}_i^n$ and $\mathcal{R}^n$ based on the data $\mathcal{D}^n$ we then obtain

$$\widehat{\mathcal{R}}_i^{m|n} := \sum_{j=n-i+1}^{J} \widehat{S}_{i,j}^{m|n}, \tag{2.5a}$$

$$\widehat{\mathcal{R}}_i^n := \sum_{m\in\mathcal{M}} \widehat{\mathcal{R}}_i^{m|n} = \sum_{m\in\mathcal{M}} \sum_{j=n-i+1}^{J} \widehat{S}_{i,j}^{m|n} \tag{2.5b}$$

and the predictor for aggregated outstanding loss liabilities for all accident years is given by

$$\widehat{\mathcal{R}}^n := \sum_{i=n-J+1}^{I} \widehat{\mathcal{R}}_i^n = \sum_{i=n-J+1}^{I} \sum_{m\in\mathcal{M}} \sum_{j=n-i+1}^{J} \widehat{S}_{i,j}^{m|n}. \tag{2.5c}$$

So far, we do not state any requirements w.r.t. properties of the predictors of loss liabilities $\widehat{S}_{i,k}^{m|n}$ and hence for (2.5a)–(2.5c), except that the predictor $\widehat{S}_{i,k}^{m|n}$ at time n must be $\mathcal{D}^n$-measurable. In Chapter 7 we state a regulatory requirement for these predictors to be so-called *best-estimate valuation of liabilities* (BEL).

Remarks 2.2 (Discounting) *In the definition of outstanding loss liabilities at time* $n \in \{I, \ldots, I+J-1\}$ *in* (2.4a)–(2.4c) *loss liabilities* $S_{i,k}^m$ *occurring at different times* $i+k>n$ *are aggregated. In these aggregations* $S_{i,k}^m$ *are not weighted by a discount factor and hence time value of money (discounting) is not incorporated. This shows that we work on a nominal scale as almost all classical claims reserving methods. The problem of the incorporation of stochastic discounting in claims reserving methods is an important topic in recent actuarial research and leads to the concept of market-consistent valuation via valuation portfolios. Since a detailed discussion is beyond the scope of this thesis we will not discuss this further here and refer to* Wüthrich–Merz *[62].*

Concluding at time $n = I$, an reserving actuary calculates within a certain model framework predictors $\widehat{S}_{i,k}^{m|I}$ for loss liabilities $S_{i,k}^m$ with $m \in \mathcal{M}$ and $i+k>I$. This leads to the predictors of outstanding loss liabilities given by (2.5a)–(2.5c).

2.4 Inflation

For the discussion on inflation we partly follow Taylor [57] and Wüthrich–Merz [63]. Inflation in claims reserving has not been often discussed in classical claims reserving literature. For the time being we assume that each claim information $m \in \mathcal{M}$ corresponds to claims payments of a specific LoB. Most claims reserving methods are based on the assumption that the observed outcome of the claim settlement process of claim information out of $\mathcal{M}$ in the past plus other additional sources of information of claim information out of $\mathcal{M}^c$ can be used to predict future outcomes of a claim information $m \in \mathcal{M}$. Therefore, we have to differentiate between the development of the claim settlement process itself and the “inflation noise” which overlays the claim settlement process. The crucial point is that depending on the claim information $m \in \mathcal{M}$ under consideration the development of claim costs may vary over time. The claims payments $S_{i,k}^m$ in accounting year $i+k=n$ and $m \in \mathcal{M}$ and its development over time may therefore be affected not only by the “pure severity and other characteristics” of the claim but also by claims inflation. In general, this claims inflation does not coincide with (but may be effected by) the classical inflation. Moreover, the impact and the severity of claims inflation may differ in each specific LoB under consideration. Therefore, for each LoB we try to exclude the inflation from the claims payments $S_{i,k}^m$ at time $i+k=n$. Let $\lambda^m(n)$ be an inflation index that measures claims inflation of LoB (claim information) $m \in \mathcal{M}$ at time n relative to time 0. Then the indexed claims payments $S_{i,k}^{m,ind}$ are given by

$$S_{i,k}^{m,ind} := \frac{1}{\varphi_I^m} \cdot \varphi_n^m \cdot S_{i,k}^m, \quad \text{for} \quad i+k=n, \tag{2.6}$$

where $\varphi_n^m := (\lambda^m(n))^{-1}$. Note that φ_n^m play the role of stochastic deflators as discussed in Wüthrich–Merz [62]. The indexed payments $S_{i,k}^{m,ind}$ in (2.6) should be the basis for claims reserving modeling, since they contain the "pure" information of the claims payments development without the inflation "noise". However, it is difficult in practice to model such an inflation process, because $\lambda^m(n)$ is not directly observable. Moreover, significant changes in $\lambda^m(n)$ are mainly caused by new developments, innovations in certain industries, for example in health care, and also by common inflation. Thus, it is difficult to calibrate a time series model for the claims inflation rate by data of the past. We propose two strategies:

1. In the claims reserving model non-indexed claims payments $S_{i,k}^m$ are considered. This is an acceptable assumption as long as there is no period of high claims inflation followed by a period of low claims inflation or vice versa, i.e. as long as there is no regime switch in the claims inflation process.

2. All observations in $\mathcal{D}^I$ are adjusted at time I by the observed (claims) inflation rate and inflation-adjusted claims payments $S_{i,k}^{m,ind}$ are modeled, leading to the predictor of inflation-adjusted claims payments $\widehat{S}_{i,k}^{m,ind|I}$. In this case, the inflation index $\lambda^m(n)$ is modeled independently from the claims payments leading to a predictor of the inflation index $\widehat{\lambda}^m(n)$ for $n > I$. The predicted values of the inflation-adjusted claims payments $\widehat{S}_{i,k}^{m,ind|I}$ and the inflation index $\widehat{\lambda}^m(n)$ are then combined to the predictor of claims payments

$$\widehat{S}_{i,k}^{m|I} := \frac{1}{\varphi_n^m} \, \varphi_I^m \, \widehat{S}_{i,k}^{m,ind|I} \quad \text{for} \quad i+k=n>I, \tag{2.7}$$

where

$$(\varphi_n^m)^{-1} := \begin{cases} \widehat{\lambda}^m(n) & \text{for} \quad n>I \\ \lambda^m(n) & \text{for} \quad n \le I \end{cases}. \tag{2.8}$$

As mentioned above, it is difficult in general to predict the claims inflation process $\lambda^m(n)$, since changes in this process are mainly caused by exogenous shocks. Hence, it is difficult to calculate (2.8) and (2.7). Therefore, Strategy 1. is preferred and non-indexed claims payments are modeled. The restriction in the beginning of this section that all claim information $m \in \mathcal{M}$ correspond to claims payments of a specific LoB can be removed, since the arguments above hold true not only for claims payments, but also for information $m \in \mathcal{M}$. Thus, we model throughout this thesis non (claims) inflation-adjusted quantities. This is in line with almost all classical claims reserving methods.

2.5 Prediction Precision

As outlined in Section 2.3, at time $n \in \{I, \ldots, I+J-1\}$ a reserving actuary has to calculate $\widehat{\mathcal{R}}^n$ to meet outstanding loss liabilities in the run-off portfolios. Depending on the data sources available, see Section 2.2 for an overview of possible sources of information, and the structure of the data the actuary sets up a model framework, see Chapters 4–6. The model is then calibrated to the data and the outstanding loss liabilities are predicted in this model framework. This leads to (model based) reserves.
Of course, there is a risk that the actual outcome of loss liabilities $\mathcal{R}^n$ significantly deviates from the prediction $\widehat{\mathcal{R}}^n$. This may have the following reasons:

1. Model misspecification: The chosen model does not describe the stochastic dynamics of the loss liability process appropriately
2. Parameter uncertainty: Within a given model framework unknown model parameters are replaced by estimates. These estimates may deviate from the "true" values, due to randomness in the parameter estimation
3. Process variance of the stochastic (random) process of loss liabilities: Even if we assume that we have chosen the "right" model and model parameters the realization of the stochastic process of loss liabilities may be far from a "typical" realization (the mean) by pure randomness

The appropriateness of the model under consideration is to be verified before using the model. This can be done in some cases by statistical methods similar to Chapter 11 in Wüthrich–Merz [63]. Of course, more sophisticated models require other strategies for verifying model assumptions. Statistical tests have to be deduced in each individual model under consideration. This is not well developed so far in actuarial science and should be subject to further research.

Having chosen a model framework, one is interested in the quantification of the prediction uncertainty. However, for this we must find an agreement in what sense the "distance" between the prediction and the actual outcomes should be measured. For that reason we have to choose an appropriate risk measure which determines a conception of measuring the quality of prediction. There is a large range of reasonable risk measures (cf. Artzner et al. [7]) which could be used to quantify prediction uncertainty. The choice of a sensible risk measure is not a pure mathematical issue and it mainly depends on the application at hand which risk measure is most appropriate. In actuarial tradition the most important risk measure is the *(conditional) mean squared error of prediction* (MSEP). However, the (conditional) MSEP has some conceptional weaknesses, see Remarks 2.4. Therefore, the MSEP is supplemented by other risk measures like *Value-at-Risk* (VaR) or *Expected Shortfall* (ES) in recent regulatory solvency frameworks like Solvency II and SST, see European Commission [23], FOPI [24] and FOPI [25]. For a proper definition of VaR and ES, see Definitions 7.7 and 7.8. The issue of the incorporation of VaR and ES in solvency considerations is discussed in Chapter 7.

2.5.1 Mean Squared Error of Prediction

As already stated above the most popular risk measure in actuarial science is the (conditional) MSEP.

Definition 2.3 (MSEP) *For a square integrable random variable X and a $\mathcal{D}^I$-measurable predictor $\widehat{X}$ the conditional MSEP is defined by*

$$\mathrm{msep}_{X|\mathcal{D}^I}\left[\widehat{X}\right] := \mathrm{E}\left[\left(X - \widehat{X}\right)^2 \middle| \mathcal{D}^I\right].$$

□

Remarks 2.4 (MSEP)

i) The (conditional) MSEP is very popular in statistics and actuarial science, since it corresponds to the squared norm of the Hilbert space of square integrable random variables $\mathcal{L}^2$ with respect to $\mathbb{P}(\ |\mathcal{D}^I)$. This allows for using basic Hilbert space theory (cf. Kolmogorov–Fomin *[36]) what makes many calculations easier to handle.*

ii) In claims reserving practice one is basically interested in the shortfall risk, i.e. in the risk of not having adequate reserves to meet loss liabilities. The MSEP uses a quadratic loss function and therefore does not reflect this risk potential, because upside as well as downside deviations are taken in the same way into account.

iii) Replacing the MSEP by another more reasonable risk measure requires completely new models in claims reserving with much stronger model assumptions. Moreover, analytical closed form results would mostly be infeasible, because other risk measures are often much harder to handle. Instead simulation methods such as Markov-Chain-Monte-Carlo (MCMC) have to be used in those cases (cf. Scollnik *[55]).*

The (conditional) MSEP has the useful property that it can be decomposed into

$$\mathrm{msep}_{X|\mathcal{D}^I}\left[\widehat{X}\right] = \underbrace{\mathrm{Var}\left[X \,|\, \mathcal{D}^I\right]}_{\text{process variance}} + \underbrace{\left(\widehat{X} - \mathrm{E}\left[X \,|\, \mathcal{D}^I\right]\right)^2}_{\text{estimation error}}. \tag{2.9}$$

This decomposition is a central technique to derive estimates for the MSEP of the outstanding loss liabilities in various claims reserving methods, see Mack [38], Wüthrich–Merz [63], Dahms [17] and Dahms–Happ [15]. Unless otherwise indicated we will use the (conditional) MSEP as an optimality criterion (risk measure) and the term "best" means with the smallest (conditional) MSEP.

In the context of our claims reserving problem at time I the (conditional) MSEP of the aggregated outstanding loss liabilities $\mathcal{R}^I$ in (2.4c) is given by

$$\operatorname{msep}_{\mathcal{R}^I|\mathcal{D}^I}\left[\widehat{\mathcal{R}}^I\right] = \mathrm{E}\left[\left(\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+1}^{J}\left(S_{i,j}^m - \widehat{S}_{i,j}^{m|I}\right)\right)^2\middle|\mathcal{D}^I\right]. \tag{2.10}$$

The conditional MSEP (2.10) measures the mean quadratic deviation of the aggregated liability predictors $\widehat{S}_{i,j}^{m|I}$ and the actual loss liability outcomes $S_{i,j}^m$ up to the final settlement of the run-off in development year J. The conditional MSEP in (2.10) is called long term or ultimate claim view of the prediction uncertainty. In new solvency regulation frameworks such as Solvency II and SST the so-called one-year view is of central interest, see European Commission [23] and FOPI [24], which is quite different from the ultimate claim view. This short term view focuses on the changes in the loss liability predictions, i.e. the change of the prediction in an one-year horizon (from time I to time $I+1$). The stochastic quantity which describes these changes in an one-year horizon is the so-called *claims development result* (CDR).

2.6 Claims Development Result

We recapitulate the prediction problem of a reserving actuary at times I and $I+1$.

Accounting year I:

The information $\mathcal{D}^I$ is available. Based on this information the actuary determines the (model dependent) predictor of aggregated outstanding loss liabilities

$$\widehat{\mathcal{R}}^I = \sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+1}^{J}\widehat{S}_{i,j}^{m|I}.$$

Accounting year $I+1$:

All loss liabilities

$$S_{I+1}^{\mathcal{M}} := \sum_{m\in\mathcal{M}}\sum_{i=I-J+1}^{I} S_{i,I-i+1}^m \tag{2.11}$$

for accounting year $I+1$ are paid out to the policyholder (or paid to the insurance company in the case of subrogation payments). Since new updated information $\mathcal{D}^{I+1}$ is available at time $I+1$, updated predictors $\widehat{\mathcal{R}}^{I+1}$ are calculated based on $\mathcal{D}^{I+1}$, see Figure 2.3. The resulting updated loss liability predictor at time $I+1$ is then given by

$$\widehat{\mathcal{R}}^{I+1} = \sum_{i=I-J+2}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+2}^{J}\widehat{S}_{i,j}^{m|I+1}.$$

The CDR describes the one-year change of predictions of aggregated outstanding loss liabilities for several accident years $\widehat{\mathcal{R}}^I$ in the time step from accounting year I to $I+1$, adjusted by the loss liability payments $S_{I+1}^{\mathcal{M}}$ in (2.11) at time $I+1$:

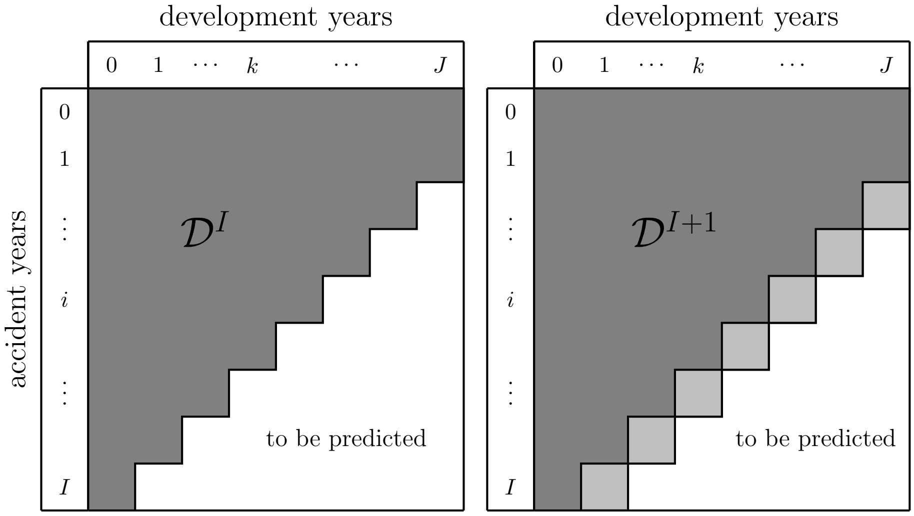

Figure 2.3: Data set $\mathcal{D}^I$ observable at time I and data set $\mathcal{D}^{I+1}$ observable at time $I+1$

Definition 2.5 (Claims Development Result) *The claims development result (CDR) at time $I+1$ of the predictor $\widehat{\mathcal{R}}^I$ of aggregated outstanding loss liabilities for several accident years is defined by*

$$\mathrm{CDR}^{\mathcal{M},I+1} := \widehat{\mathcal{R}}^I - \left(\widehat{\mathcal{R}}^{I+1} + S^{\mathcal{M}}_{I+1}\right). \tag{2.12}$$

□

In many claims reserving methods the CDR is often considered on the level of single accident years $i \in \{I-J+1, \dots, I\}$ and claim information $m \in \mathcal{M}$. The CDR at time $I+1$ for the predictor of outstanding loss liabilities $\widehat{\mathcal{R}}_i^{m|I}$ of accident year i and claim information $m \in \mathcal{M}$ is given by

$$\mathrm{CDR}_i^{m,I+1} := \widehat{\mathcal{R}}_i^{m|I} - \left(\widehat{\mathcal{R}}_i^{m|I+1} + S^m_{i,I-i+1}\right).$$

Moreover, the CDR at time $I+1$ for the predictor of aggregated loss liabilities $\widehat{\mathcal{R}}_i^I$ of single accident years $i \in \{I-J+1, \dots, I\}$ is defined by

$$\mathrm{CDR}_i^{\mathcal{M},I+1} := \sum_{m\in\mathcal{M}} \mathrm{CDR}_i^{m,I+1} = \widehat{\mathcal{R}}_i^I - \left(\widehat{\mathcal{R}}_i^{I+1} + S^{\mathcal{M}}_{i,I-i+1}\right),$$

with the aggregated loss liabilities of accident year i and development year k

$$S^{\mathcal{M}}_{i,k} := \sum_{m\in\mathcal{M}} S^m_{i,k}.$$

By Definition 2.5 the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ exactly corresponds to the change between i) the predictor $\widehat{\mathcal{R}}^I$ at time I and ii) the predictor $\widehat{\mathcal{R}}^{I+1}$ at time $I+1$ plus the loss liabilities $S^{\mathcal{M}}_{I+1}$ paid out at time $I+1$, see (2.12). A negative claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$

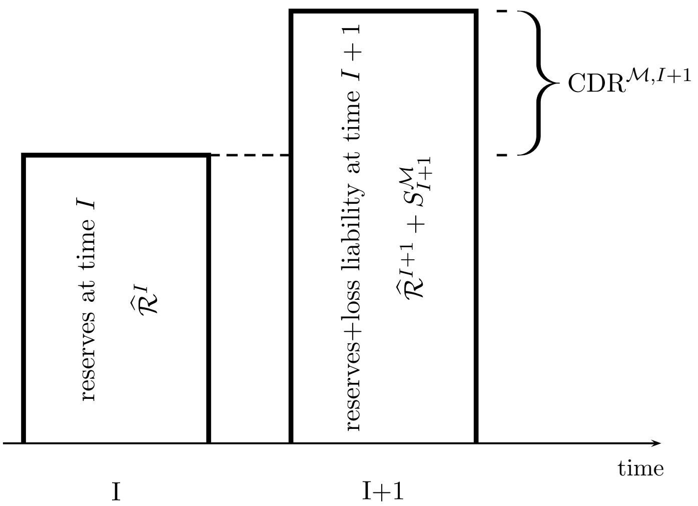

Figure 2.4: Reserves $\widehat{\mathcal{R}}^I$ based on $\mathcal{D}^I$ at time I, updated reserves $\widehat{\mathcal{R}}^{I+1}$ based on $\mathcal{D}^{I+1}$ at time $I+1$ and the resulting claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$

results in a loss in the position "loss experience prior accident years" on the balance sheet of an insurance company, whereas a positive one leads to a profit in this position (cf. Merz–Wüthrich [45]). Hence the CDR directly effects the profit and loss statement in the balance sheet of an insurance company. This reveals the direct link of the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ to re-adjustments of the predictor $\widehat{\mathcal{R}}^{I+1}$ in accounting year $I+1$, see Figure 2.4. We analyze the CDR in more detail.

Properties of the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$

In accounting year I the best $\mathcal{D}^I$-measurable estimator for $S_{i,k}^m$ is given by $\mathrm{E}\left[S_{i,k}^m \middle| \mathcal{D}^I\right]$. If for the predictor holds $\widehat{S}_{i,k}^{m|I} = \mathrm{E}\left[S_{i,k}^m \middle| \mathcal{D}^I\right]$, the linearity and the tower property of conditional expectations (cf. Williams [59]) imply for the expected claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ at time I

$$
\begin{aligned}
\mathrm{E}\left[\mathrm{CDR}^{\mathcal{M},I+1}\middle|\mathcal{D}^I\right] &= \mathrm{E}\left[\widehat{\mathcal{R}}^I - \left(\widehat{\mathcal{R}}^{I+1} + S_{I+1}^{\mathcal{M}}\right)\middle|\mathcal{D}^I\right] \\
&= \mathrm{E}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+1}^{J}\widehat{S}_{i,j}^{m|I} - \left(\sum_{i=I-J+2}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+2}^{J}\widehat{S}_{i,j}^{m|I+1} + S_{I+1}^{\mathcal{M}}\right)\middle|\mathcal{D}^I\right] \\
&= \mathrm{E}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+1}^{J}\widehat{S}_{i,j}^{m|I} - \sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+1}^{J}\widehat{S}_{i,j}^{m|I+1}\middle|\mathcal{D}^I\right] \\
&= \sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{j=I-i+1}^{J}\left(\mathrm{E}\left[\widehat{S}_{i,j}^{m|I}\middle|\mathcal{D}^I\right] - \mathrm{E}\left[\widehat{S}_{i,j}^{m|I+1}\middle|\mathcal{D}^I\right]\right) \\
&= 0.
\end{aligned}
$$

The interpretation of this result is as follows: Assume that the data generating process of loss liabilities is given by the model used by the reserving actuary for claims reserving and the model allows for the calculation of $\mathrm{E}\left[S_{i,k}^m \middle| \mathcal{D}^I\right]$. Then the prediction at time I of aggregated outstanding loss liabilities of several accident years $\widehat{\mathcal{R}}^I$ equals in the average the sum of $\widehat{\mathcal{R}}^{I+1}$ and $S_{I+1}^{\mathcal{M}}$, viewed from time I. That means that the amount of $\widehat{\mathcal{R}}^I$ is such that, viewed from time I, in the average the prediction $\widehat{\mathcal{R}}^{I+1}$ as well as the loss liability payments $S_{I+1}^{\mathcal{M}}$ can be financed by $\widehat{\mathcal{R}}^I$. This property is often called self-financing property. In most (classical) distribution-free claims reserving methods estimates of $\mathrm{E}\left[S_{i,k}^m \middle| \mathcal{D}^I\right]$ are used as a predictor for outstanding loss liabilities, see the *chain ladder* (CL) method in MACK [38] or the *linear stochastic reserving methods* (LSRMs) in DAHMS [17] among others. This motivates the fact that the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ is often predicted by 0 in the profit and loss statement of the balance sheet in a non-life insurance company.

Similar to the quantification of the prediction uncertainty of the aggregated outstanding loss liabilities in terms of the (conditional) MSEP, see (2.10), we measure the prediction uncertainty of the CDR by means of the (conditional) MSEP given by

$$\mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0] := \mathrm{E}\left[\left(\left(\mathrm{CDR}^{\mathcal{M},I+1}\right) - 0\right)^2 \middle| \mathcal{D}^I\right]. \tag{2.13}$$

Sometimes the (conditional) MSEP for the claims development result $\mathrm{CDR}_i^{\mathcal{M},I+1}$ for single accident years $i \in \{I-J+1, \dots, J\}$ is considered

$$\mathrm{msep}_{\mathrm{CDR}_i^{\mathcal{M},I+1}|\mathcal{D}^I}[0] := \mathrm{E}\left[\left(\left(\mathrm{CDR}_i^{\mathcal{M},I+1}\right) - 0\right)^2 \middle| \mathcal{D}^I\right].$$

For more information on the CDR see OHLSSON–LAUZENINGKS [49].

Remarks 2.6 (CDR)

i) The CDR is the risk driver in the one-year reserving risk (for the multi-year reserving risk, see DIERS–LINDE *[20]). Therefore, the CDR is the central quantity in current regulatory solvency frameworks, see Chapter 7 for details.*

ii) Regulatory solvency rules aim to protect against shortfalls in the CDR, see EUROPEAN UNION *[23] or* FOPI *[24]. In these rules the (conditional) MSEP* (2.13) *is utilized to calibrate a log-normal distribution by the method of moments, see Chapter 7. Therefore, it is questionable, whether the choice of the MSEP as a risk measure is appropriate, since many distributional properties can not be captured by the MSEP.*

Risk Characteristics in Classical Claims Reserving Methods

In this chapter we introduced predictors for aggregated outstanding loss liabilities for several accident years $\widehat{\mathcal{R}}^I$ and the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ as the central stochastic quantities under consideration for claims reserving at time I (today). In a first step a predictor

$\widehat{\mathcal{R}}^I$ for aggregated outstanding loss liabilities $\mathcal{R}^I$ is calculated. In classical claims reserving the prediction uncertainty of the predictor $\widehat{\mathcal{R}}^I$ for the outstanding loss liabilities $\mathcal{R}^I$, see (2.5c), as well as the predictor 0 for the $\mathrm{CDR}^{\mathcal{M},I+1}$, see (2.13), is usually measured by the (conditional) MSEP, see Definition 2.3. The CDR became the central quantity in recent solvency frameworks, see FOPI [24], since it reflects the adjustments in the predictions of outstanding loss liabilities that will (possibly) be necessary in the time step from I to $I+1$. As motivated in the last section the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ is predicted by 0 at time I, i.e. no adjustment at time $I+1$ are to be expected. As a conclusion, we note that in classical claims reserving the aggregated outstanding loss liabilities $\mathcal{R}^I$ and the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ are predicted by the predictors $\widehat{\mathcal{R}}^I$ and 0, respectively. The corresponding prediction uncertainty is measured by the MSEP, see Table 2.1 for an overview. These quantities can be calculated (or

quantity	predictor	prediction uncertainty
$\mathcal{R}^I$	$\widehat{\mathcal{R}}^I$	$\mathrm{msep}_{\mathcal{R}^I \mid \mathcal{D}^I}\left[\widehat{\mathcal{R}}^I\right]$
$\mathrm{CDR}^{\mathcal{M},I+1}$	0	$\mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1} \mid \mathcal{D}^I}[0]$

Table 2.1: Classical risk characteristics: Reserves and CDR and the corresponding first two moments

estimated) for many claims reserving methods. Especially the claims reserving methods which belong to the class of LSRMs or Bayesian LSRMs allow for the derivation of these quantities, see Chapter 4.

In recent regulatory solvency frameworks more sophisticated risk measures like higher moments or quantile based risk measures such as VaR or ES are required. The calculation of such risk measures overcharges the possibilities of distribution-free claims reserving methods. They are only accessible under simplifications or approximations, which are subject to Chapter 7. These different accessibility levels of claims reserving methods w.r.t. the MSEP (basic risk measure) and VaR or ES (more sophisticated risk measures) are important in new solvency frameworks, see European Commission [23] and Fopi [24].

Therefore, we divide the set of stochastic claims reserving methods into two groups:

1. **(Bayesian) distribution-free claims reserving methods**:
   (Bayesian) distribution-free claims reserving methods comprise many classical claims reserving methods used in actuarial practice. We present some important representatives of this class in Chapter 3. These methods can be essentially summarized in the wide class of (Bayesian) LSRMs introduced in Dahms [17] and Dahms–Happ [15]. This unification and generalization of distribution-free claims reserving methods is subject to Chapter 4.

In this class only the risk characteristics given in Table 2.1 can be derived. More sophisticated risk measures are accessible in this model class only under simplifications and approximations, see Simplifications I–II in Chapter 7.

2. **Distributional methods**:
   In actuarial science various distributional claims reserving methods have been introduced. Among others *generalized linear model* (GLM) techniques were used in England–Verrall [21] and [22], Haberman–Renshaw [29], Taylor–McGuire [58] and Alai et al. [3], *generalized linear mixed models* (GLMM) were applied in Antonio–Beirlant [5], Frees–Shi [56] and De Jong [19] used copula based models and distributional Bayesian models were applied in Salzmann–Wüthrich [54], Wüthrich [60], Merz–Wüthrich [46], Happ–Wüthrich [31] and others. These methods often allow for the derivation of the whole predictive distribution of outstanding loss liabilities and not only of the first two (conditional) moments given in Table 2.1.

   In Chapter 5 we highlight one important representative of the class of distributional methods, namely the *paid-incurred chain* (PIC) reserving method by Merz–Wüthrich [46]. This distributional claims reserving method combines in an elegant way claims payments and incurred losses data. Following Happ et al. [15] we recapitulate the PIC reserving method and show how the MSEP of the CDR can be derived analytically. Moreover, we derive the whole predictive distribution of the CDR via *Monte-Carlo* (MC) simulations. A generalization of the PIC model which allows to model dependence structures in the data presented in Happ–Wüthrich [30] is subject to Chapter 6.

3 Classical Distribution-Free Claims Reserving Methods

In this chapter we present classical distribution-free claims reserving methods commonly used in actuarial practice. We state model assumptions underlying each method and present predictors $\widehat{\mathcal{R}}^I$ for outstanding aggregated loss liabilities (claims payments) $\mathcal{R}^I$. Prediction uncertainty is not considered in this chapter. This has the following reason: As will be shown in Chapter 4 almost all classical distribution-free claims reserving methods can be embedded in the general (Bayesian) LSRM framework. For this model class the (conditional) MSEP of aggregated outstanding loss liabilities for several accident years $\widehat{\mathcal{R}}^I$ and the CDR, see Table 2.1, are derived in Dahms [17] for LSRMs and in Chapter 4 for Bayesian LSRMs. In the sequel of each method presented we state some remarks on advantages and disadvantages and point out to what extend the disadvantages could be tackled in state-of-the-art claims reserving methods.

3.1 General Notation

In this chapter we work under the extended view given in Chapter 2. For the general formulation of the stochastic dynamics of classical distribution-free claims reserving methods we define the linear subspaces $\mathbb{L}^n$ and $\mathbb{L}_k$ denoting the linear spaces generated by all increments $S_{i,j}^m$ up to accounting year n and development year k, respectively. Furthermore, the linear subspace generated by $\mathbb{L}^n$ and $\mathbb{L}_k$ is denoted by $\mathbb{L}_k^n$, i.e.

$$\mathbb{L}^n := \left\{ \sum_{m=1}^{M} \sum_{i=0}^{I} \sum_{j=0}^{(n-i)\wedge J} x_{i,j}^m S_{i,j}^m \;:\; x_{i,j}^m \in \mathbb{R} \right\},$$

$$\mathbb{L}_k := \left\{ \sum_{m=1}^{M} \sum_{i=0}^{I} \sum_{j=0}^{k} x_{i,j}^m S_{i,j}^m \;:\; x_{i,j}^m \in \mathbb{R} \right\},$$

$$\mathbb{L}_k^n := \left\{ \sum_{m=1}^{M} \sum_{i=0}^{I} \sum_{j=0}^{((n-i)\wedge J)\vee k} x_{i,j}^m S_{i,j}^m \;:\; x_{i,j}^m \in \mathbb{R} \right\},$$

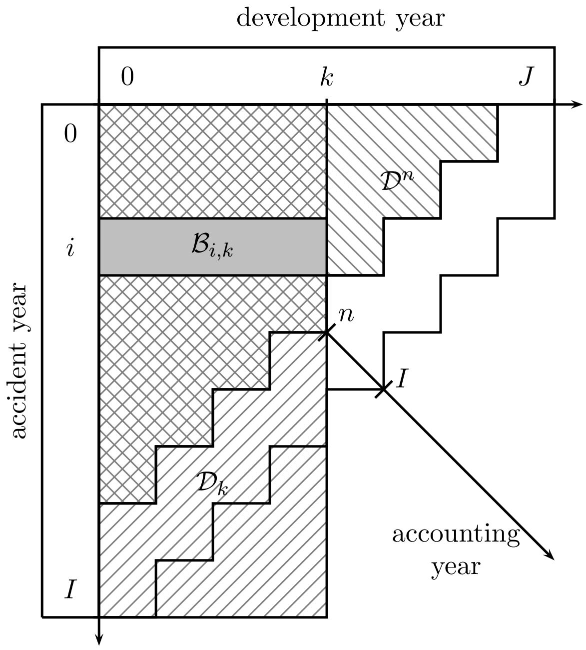

Figure 3.1: σ-fields (sets of observations): $\mathcal{B}_{i,k}$ - all claim information in accident year i up to development year k, $\mathcal{D}_k$ - all claim information up to development year k, $\mathcal{D}^n$ - all claim information up to accounting year n and $\mathcal{D}_k^n$ - the union of all information in $\mathcal{D}_k$ and $\mathcal{D}^n$

where $a \wedge b$ and $a \vee b$ denote the minimum and maximum of the real numbers a and b, respectively. Moreover, the corresponding σ-fields (sets of observations) are defined by

$$\mathcal{B}_{i,k} := \sigma\left(S_{i,j}^m \,:\, 1 \leq m \leq M,\ 0 \leq j \leq k\right), \qquad \mathcal{D}_k := \sigma\left(\mathbb{L}_k\right) = \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,k}\right),$$

$$\mathcal{D}^n := \sigma\left(\mathbb{L}^n\right) = \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,(n-i)\wedge J}\right), \qquad \mathcal{D}_k^n := \sigma\left(\mathbb{L}_k^n\right) = \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,((n-i)\wedge J)\vee k}\right)$$

and illustrated in Figure 3.1.

3.2 Chain Ladder Method

The CL method is the most popular and widest used claims reserving method. The original version of the CL method was a purely deterministic procedure for calculating reserves without considering the claims reserving problem in a stochastic framework. In 1993, Mack [38] was the first to formulate a distribution-free stochastic framework where the reserves resulting from the original deterministic method are given a meaningful stochastic foundation. Within the stochastic framework, Mack [38] was able to present estimates for the prediction uncertainty in terms of the (conditional) MSEP. The classical model assumptions of Mack [38] for the CL method are formulated in the cumulative claims payments representation, i.e. we consider $C_{i,k}^m$ instead of $S_{i,k}^m$, and the method uses its own past claims settlement process as the only source of information. That means in the CL method we consider one cumulative claim information $C_{i,k}$, i.e. $M = 1$, $\mathcal{M} = \{1\}$ (the superscript 1 in the exponent will be omitted, because there is only one claim information).

Model Assumptions 3.1 (Distribution-free CL model of Mack) *There exist constant factors* $g_0, \ldots, g_{J-1}$ *and variance parameters* $\sigma_0^2, \ldots, \sigma_{J-1}^2 > 0$ *such that*

a) Cumulative claims payments $\{C_{i,k}|\, 0 \le k \le J\}$ *from different accident years* $i \in \{0, \ldots, I\}$ *are independent.*

b) For all $0 \le i \le I$ *and* $0 \le k \le J-1$ *holds*

$$\begin{aligned} \mathrm{E}[C_{i,k+1}|\, \mathcal{B}_{i,k}] &= \mathrm{E}[C_{i,k+1}|\, C_{i,k}] = g_k C_{i,k}, \\ \mathrm{Var}[C_{i,k+1}|\, \mathcal{B}_{i,k}] &= \mathrm{Var}[C_{i,k+1}|\, C_{i,k}] = \sigma_k^2 C_{i,k}. \end{aligned}$$

□

The factors g_k are called development or age-to-age factors, because they describe the expected one step transition of $(C_{i,k})_{k \ge 0}$. The conditional expectation $\mathrm{E}\big[C_{i,k+1}|\, \mathcal{D}^I\big]$ is the best $\mathcal{D}^I$-measurable predictor for $C_{i,k+1}$ ($k \in \{0, \ldots, J-1\}$), see Williams [59, p. 86], page 86. In accounting year I (i.e. given $\mathcal{D}^I$) the conditional expected cumulative claims payments for accident year $I - J + 1 \le i \le I$ up to development year $I - i < k + 1 \le J$ are given under Model Assumptions 3.1 by

$$\mathrm{E}\big[C_{i,k+1}|\, \mathcal{D}^I\big] = C_{i,I-i} \prod_{j=I-i}^{k} g_j,$$

see Wüthrich–Merz [63] for a proof. For the ultimate claim, i.e. $k = J-1$, we obtain

$$\mathrm{E}\big[C_{i,J}|\, \mathcal{D}^I\big] = C_{i,I-i} \prod_{j=I-i}^{J-1} g_j.$$

In practice the development factors $\{g_k|\, 0 \le k \le J-1\}$ are unknown and have to be estimated from the data $\mathcal{D}^I$ available in accounting year I. A (conditional) unbiased estimator with minimum (conditional) variance in the class of linear unbiased estimators is given by

$$\widehat{g}_k^{I,CL} := \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k}}, \qquad k \in \{0, \ldots, J-1\}, \tag{3.1}$$

see Wüthrich–Merz [63] for a proof. The claim predictors are then given by

$$\widehat{C}_{i,k+1}^{I,CL} := C_{i,I-i} \prod_{j=I-i}^{k} \widehat{g}_j^{I,CL}. \tag{3.2}$$

For these predictors in (3.2) we obtain by the (conditional) unbiasedness of the estimators $\widehat{g}_k^{I,CL}$ in (3.1)

$$\begin{aligned}
\mathrm{E}\left[\widehat{C}_{i,k+1}^{I,CL} \middle| C_{i,I-i}\right] &= \mathrm{E}\left[C_{i,I-i} \prod_{j=I-i}^{k} \widehat{g}_j^{I,CL} \middle| C_{i,I-i}\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[C_{i,I-i} \prod_{j=I-i}^{k} \widehat{g}_j^{I,CL} \middle| \mathcal{D}_k\right] \middle| C_{i,I-i}\right] \\
&= \mathrm{E}\left[C_{i,I-i} g_k \prod_{j=I-i}^{k-1} \widehat{g}_j^{I,CL} \middle| C_{i,I-i}\right] \\
&\vdots \\
&= C_{i,I-i} \prod_{j=I-i}^{k} g_j \\
&= \mathrm{E}\left[C_{i,k+1} \middle| \mathcal{D}^I\right].
\end{aligned}$$

This shows that the predictor (3.2) is conditionally given $C_{i,I-i}$ unbiased for $\mathrm{E}\left[C_{i,k+1} \middle| \mathcal{D}^I\right]$. This motivates the explicit claim predictor (3.2). Finally, we obtain the CL predictor of aggregated outstanding loss liabilities of accident year $i \geq I-J+1$ at time I (note that $M=1$)

$$\widehat{\mathcal{R}}_i^{I,CL} = \widehat{C}_{i,J}^{I,CL} - C_{i,I-i} = C_{i,I-i} \prod_{j=I-i}^{J-1} \widehat{g}_j^{I,CL} - C_{i,I-i} = C_{i,I-i} \left(\prod_{j=I-i}^{J-1} \widehat{g}_j^{I,CL} - 1 \right). \tag{3.3}$$

The CL predictor of aggregated loss liabilities for several accident years follows by summation of (3.3) over accident years $i \in \{I-J+1, \ldots, J\}$

$$\widehat{\mathcal{R}}^{I,CL} := \sum_{i=I-J+1}^{I} \widehat{\mathcal{R}}_i^{I,CL}. \tag{3.4}$$

The prediction uncertainty in terms of the conditional MSEP of the CL predictors $\widehat{\mathcal{R}}^{I,CL}$ is given under slightly varying approaches in Mack [38], Murphy [48] and Buchwalder et al. [10]. For the quantification of the CDR uncertainty by means of the (conditional) mean squared error of prediction $\mathrm{msep}_{\mathrm{CDR}^{1,I+1}|\mathcal{D}^I}[0]$ we refer to Merz–Wüthrich [47].

Remarks 3.2 (Distribution-free CL model of Mack)

i) The CL method is by far the most popular claims reserving method, since it is easy to use, very intuitive and simple to implement.

ii) The CL method is based on only one source of information (the settlement process itself) and does not respect other sources of information available. In insurance practice additional sources of information, for example incurred losses data, are often available. The problem of considering cumulative payments and incurred losses simultaneously is addressed in Chapters 5 and 6.

iii) According to (3.4) *and* (3.3) *the reserves* $\widehat{\mathcal{R}}^I$ *in the CL method completely rely on the last observable entries* $\{C_{I-J+1,J-1}, \ldots, C_{I,0}\}$*. This makes the method very sensitive with respect to outliers or zeros on the diagonal leading to nonsense reserve estimates. Such scenarios are not unusual in excess-of-loss reinsurance.*

iv) The classical CL method does not allow for the incorporation of expert knowledge or information from industry-wide data in the development factors g_k*. This problem is addressed in the Bayes CL method, see Section 3.3. In Chapter 4 the general case is considered of incorporating such information in all models which belong to the wide class of LSRMs. This includes the CL method and the class of Bayesian LSRMs is therefore a generalization of the Bayes CL method.*

3.3 Bayes Chain Ladder Method

Model Assumptions 3.1 require unknown development (age-to-age) factors $\{g_k \mid 0 \le k \le J-1\}$ which have to be estimated appropriately from the data. If there is additional portfolio data available, for instance industry-wide data or expert knowledge concerning the development factors of the claims reserving method, new models are needed to cope with those new information sources. The *Bayes chain ladder* (Bayes CL) model in Gisler–Wüthrich [27] allows for the incorporation of additional information on the development factors in the CL method within a credibility based approach. In the Bayes CL model it is assumed that the development factors g_k with $0 \le k \le J-1$ in the classical CL model are random variables denoted by G_k with known (conditional) mean and variance. The model assumptions are formulated for the development ratios $Y_{i,k} := \frac{C_{i,k+1}}{C_{i,k}}$ as follows:

Model Assumptions 3.3 (Bayes CL model)

a) Conditionally, given $\mathbf{G} := (G_0, \ldots, G_{J-1})$*, the cumulative claims payments* $\{C_{i,k} \mid 0 \le k \le J\}$ *from different accident years* $i \in \{0, \ldots, I\}$ *are independent.*

b) Conditionally, given $\mathbf{G}$ *and* $\mathcal{B}_{i,k}$*, the distribution of* $Y_{i,k}$ *depends only on* $C_{i,k}$ *and it holds*

$$\mathrm{E}[Y_{i,k} \mid \mathbf{G}, \mathcal{B}_{i,k}] = G_k$$

$$\mathrm{Var}[Y_{i,k} \mid \mathbf{G}, \mathcal{B}_{i,k}] = \frac{\sigma_k^2(G_k)}{C_{i,k}},$$

for $i \in \{0, \dots, I\}$ *and* $k \in \{0, \dots, J-1\}$.

c) $\{G_0, G_1, \dots, G_{J-1}\}$ *are independent.*

□

In the Bayes CL method a credibility theory based approach is used for the prediction of the development factors G_k, see Bühlmann–Gisler [14]. The resulting credibility predictor $\widehat{G}_k^{I,Cred}$ for the development factor G_k is then given by a credibility weighted average of the classical CL estimate $\widehat{g}_k^{I,CL}$ in (3.1) and the prior mean $\mathrm{E}[G_k | \mathcal{D}_k]$

$$\widehat{G}_k^{I,Cred} := \alpha_k \widehat{g}_k^{I,CL} + (1-\alpha_k) g_k, \tag{3.5}$$

where $\alpha_k \in [0,1]$ is given by

$$\alpha_k := \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k} + \frac{\sigma_k^2}{\tau_k^2}}, \tag{3.6}$$

with the prior structural parameters

$$g_k := \mathrm{E}[G_k | \mathcal{D}_k] \qquad \sigma_k^2 := \mathrm{E}\left[\sigma_k^2(G_k) \middle| \mathcal{D}_k\right] \qquad \tau_k^2 := \mathrm{Var}[G_k | \mathcal{D}_k]. \tag{3.7}$$

This leads to credibility based claim predictors for accident year $i \in \{I-J+1, \dots, I\}$ and development year $(k+1) \in \{I-i+1, \dots, J\}$ given by

$$\widehat{C}_{i,k+1}^{I,Cred} := C_{i,I-i} \prod_{j=I-i}^{k} \widehat{G}_j^{I,Cred}. \tag{3.8}$$

This is exactly the predictor in (3.2) in the classical CL model, but with the estimate $\widehat{g}_j^{I,CL}$ replaced by the credibility predictor $\widehat{G}_j^{I,Cred}$. By Equation (3.8) we obtain the predictor for aggregated outstanding loss liabilities for accident year $i \in \{I-J+1, \dots, J\}$

$$\widehat{\mathcal{R}}_i^{I,Cred} = \widehat{C}_{i,J}^{I,Cred} - C_{i,I-i} = C_{i,I-i} \left(\prod_{j=I-i}^{J-1} \widehat{G}_j^{I,Cred} - 1 \right). \tag{3.9}$$

Finally, by summation of (3.9) over all accident years $i \in \{I-J+1, \dots, I\}$ we get the credibility based CL predictor for aggregated outstanding loss liabilities given by

$$\widehat{\mathcal{R}}^{I,Cred} = \sum_{i=I-J+1}^{I} \widehat{\mathcal{R}}_i^{I,Cred} = \sum_{i=I-J+1}^{I} C_{i,I-i} \left(\prod_{j=I-i}^{J-1} \widehat{G}_j^{I,Cred} - 1 \right).$$

For the derivation of the prediction uncertainty $\mathrm{msep}_{\mathcal{R}^I | \mathcal{D}^I}\left[\widehat{\mathcal{R}}^{I,Cred}\right]$ in terms of the (conditional) MSEP we refer to Gisler–Wüthrich [27]. For the prediction uncertainty of the CDR, see Chapter 4, where the Bayes CL method is looked at as a Bayesian LSRM. In this general model framework the (conditional) MSEP of the CDR as well as the MSEP of the predictor $\widehat{\mathcal{R}}^{I,Cred}$ for outstanding loss liabilities are derived.

Remarks 3.4 (Bayes CL model)

i) The credibility predictor (3.5) *is a credibility weighted average of the classical CL estimate* $\widehat{g}_k^{I,CL}$ *purely based on the data and the structural parameter* g_k*. The structural parameters* (3.7)*, which are required for the calculation of the credibility weights* (3.6)*, can be estimated from data, see* Bühlmann–Gisler *[14], or deduced from external sources of information.*

ii) Since the CL method belongs to the class of LSRMs, see Chapter 4, one can ask whether this incorporation of prior knowledge of the development pattern by means of credibility theory does not only work in the CL method but also for all other methods which belong to the class of LSRMs. This question is answered in Chapter 4. Moreover, estimates of (conditional) mean squared error of prediction $\text{msep}_{\mathcal{R}^I|\mathcal{D}^I}\left[\widehat{\mathcal{R}}^I\right]$ *and* $\text{msep}_{\text{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]$ *are derived.*

3.4 Complementary Loss Ratio Method

In the *complementary loss ratio* (CLR) method, see Bühlmann [12], incremental claims payments $S_{i,k}^1$ are considered as the first claim information. An external given exposure P_i (e.g. earned premium, volume measure, ultimate claim prediction, number and size of contracts, etc.) defined by

$$S_{i,k}^2 := \begin{cases} P_i & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

and independent of time is used as a second source of information. That means in this model there are $M = 2$ claim information $S_{i,k}^1$ and $S_{i,k}^2$ and we are interested in the prediction of the first claim information $S_{i,k}^1$. The second claim information $S_{i,k}^2$ is used for the prediction of $S_{i,k}^1$ but is not predicted itself, i.e. $\mathcal{M} = \{1\}$ and $\mathcal{M}^c = \{2\}$.

Model Assumptions 3.5 (CLR model) *There exist constant weights* $g_0, \ldots, g_{J-1}$ *and variance parameters* $\sigma_0^2, \ldots, \sigma_{J-1}^2 > 0$ *such that*

a) Incremental claims payments $\{S_{i,k}^1 |\, 0 \le i \le I, 0 \le k \le J\}$ *are independent.*

b) For $0 \le i \le I$ *and* $0 \le k \le J-1$ *holds*

$$\begin{aligned} \text{E}\left[S_{i,k+1}^1 \middle| \mathcal{B}_{i,k} \right] &= g_k P_i, \\ \text{Var}\left[S_{i,k+1}^1 \middle| \mathcal{B}_{i,k} \right] &= \sigma_k^2 P_i. \end{aligned}$$

□

In the CLR method estimates of the unknown model parameters g_k are given by

$$\widehat{g}_k^{I,CLR} := \frac{1}{\sum_{i=0}^{I-k-1} P_i} \sum_{i=0}^{I-k-1} P_i \frac{S_{i,k+1}^1}{P_i}.$$

This leads to the CLR predictor

$$\widehat{S}_{i,k+1}^1 := \widehat{g}_k^{I,CLR} P_i \tag{3.10}$$

for $i + k \geq I$. We then obtain for the predictor of aggregated outstanding loss liabilities of accident year $i \in \{I - J + 1, \ldots, J\}$

$$\widehat{\mathcal{R}}_i^{I,CLR} := \sum_{k=I-i}^{J-1} \widehat{S}_{i,k+1}^1 \tag{3.11}$$

and the corresponding predictor for several accident years is given by

$$\widehat{\mathcal{R}}^{I,CLR} := \sum_{i=I-J+1}^{I} \widehat{\mathcal{R}}_i^{I,CLR} = \sum_{i=I-J+1}^{I} \sum_{k=I-i}^{J-1} \widehat{S}_{i,k+1}^1. \tag{3.12}$$

Note that if we choose in Model Assumptions 3.5 $P_i := V_i$ where V_i is a volume measure we obtain the ALR method in Merz-Wüthrich [44]. This shows that the ALR method is a special case of the CLR method. Moreover, in Chapter 4 we will see that the CLR method belongs to the class of LSRMs. Therefore, estimates of the (conditional) mean squared error of prediction $\mathrm{msep}_{\mathcal{R}^I|\mathcal{D}^I}\left[\widehat{\mathcal{R}}^I\right]$ and $\mathrm{msep}_{\mathrm{CDR}^{1,I+1}|\mathcal{D}^I}[0]$ for the CLR method are given in Dahms [17].

3.5 Bornhuetter–Ferguson Method

Beside the CL method the *Bornhuetter–Ferguson* (BF) method presented in Bornhuetter–Ferguson [8] is one of the most popular claims reserving methods. In 1972, Bornhuetter and Ferguson introduced the BF method in order to solve the main problem of the CL method that the reserve of accident year $i \in \{I - J + 1, \ldots, J\}$ in (3.3) is proportional to the last diagonal entry $C_{i,I-i}$. The central idea of the BF method is that for each accident year i the incremental claims payments $S_{i,k}^1$ correspond to a fixed percentage y_k of a prior ultimate claim estimate x_i. We consider the BF method in its incremental representation. A cumulative version is given in Wüthrich–Merz [63]. The incremental claims payments $S_{i,k}^1$ constitute the first claim information and the second claim information is given by

$$S_{i,k}^2 := \begin{cases} x_i & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases},$$

where x_i is a prior ultimate claim estimate. That means that in the BF method there are $M = 2$ claim information whereas the set of claim information generating cash flows is given by $\mathcal{M} = \{1\}$. There are many slightly varying model assumptions underlying the BF method. The assumptions in Mack [39] are given by:

Model Assumptions 3.6 (BF model of Mack) *There exist constant weights $y_0, \ldots, y_J$ and variance parameters $s_0^2, \ldots, s_J^2 > 0$ such that*

a) Incremental claims payments $\{S^1_{i,k} | \, 0 \le i \le I, \; 0 \le k \le J\}$ are independent.

b) For $0 \le i \le I$ and $0 \le k \le J$ holds

$$\mathrm{E}\big[S^1_{i,k}\big] = y_k x_i \quad \text{and} \quad y_0 + \ldots + y_J = 1$$
$$\mathrm{Var}\big[S^1_{i,k}\big] = s_k^2 x_i.$$

□

By setting

$$g_k := y_{k+1}, \quad P_i := x_i \quad \text{and} \quad \sigma_k^2 := s_{k+1}^2 \quad \text{for} \qquad 0 \le k \le J-1$$

we obtain for the BF model

$$\mathrm{E}\big[S^1_{i,k+1} \big| \mathcal{B}_{i,k}\big] = \mathrm{E}\big[S^1_{i,k+1}\big] = y_{k+1} x_i = g_k P_i,$$
$$\mathrm{Var}\big[S^1_{i,k+1} \big| \mathcal{B}_{i,k}\big] = \mathrm{Var}\big[S^1_{i,k+1}\big] = s_{k+1}^2 x_i = \sigma_k^2 P_i.$$

This shows that the BF method can be looked at as a special case of the CLR method. Therefore, the predictor of outstanding loss liabilities of accident year $i \in \{I-J+1, \ldots, I\}$ in the BF method is given by (3.11), with $\widehat{g}_k$ and P_i replaced by $\widehat{y}_{k+1}$ and x_i, respectively. For the discussion of the associated prediction uncertainty in the BF method we refer to the discussion of the prediction uncertainty in the CLR method given above. As already mentioned above there are various versions of the BF model with slightly varying model assumptions. For the corresponding reserves and prediction uncertainty in these cases we refer to Wüthrich–Merz [63], Alai et al. [3], Alai et al. [4] and Saluz et al. [52].

Remarks 3.7 (CLR and BF method)

i) In the CLR method and therefore in the BF method too the second claim information $S^2_{i,k}$ is used for the prediction of the claims payments process $S^1_{i,k}$, see Model Assumptions 3.5 and 3.6. This is the classical case, where $S^2_{i,k}$ is used for predicting $S^1_{i,k}$ but not predicted itself, i.e. $M = 2$ and $\mathcal{M} = \{1\}$.

ii) The BF method requires and allows for a prior estimate of the ultimate claim x_i for each accident year i. These prior estimates are often deduced from pricing arguments. In the CLR method a prior exposure P_i is required.

iii) The CLR and BF methods solve the basic problem of the CL method that the reserves strongly depend on the observations on the last observed diagonal. This becomes obvious by comparing the CL reserves in (3.3) *and the CLR and BF reserves in* (3.11)*.*

iv) The classical BF method does not allow for including other sources of information, for example incurred losses data, and is therefore to some extend inflexible in insurance practice. This problem is addressed in Chapters 5 and 6.

v) Prior expert knowledge or industry wide data can not be included in the estimation of the parameters y_k in the BF method and g_k in the CLR method. This task is considered for the wide class of LSRMs in Chapter 4.

All so far presented claims reserving methods (and much more) belong to the class of (Bayesian) LSRMs. Therefore, best-estimate predictors for outstanding loss liabilities, their prediction uncertainty and the one-year reserving risk can be calculated in the general framework of (Bayesian) LSRMs in Dahms [17] and Dahms–Happ [15] presented in Chapter 4.

A distribution-free claims reserving method which does not fit into the LSRM model framework is briefly presented in the following section.

3.6 Munich Chain Ladder Method

Beside the claims payments data, incurred losses as a second data source are often available in insurance companies. Applying the CL method, see Section 3.2, to the claims payments data leads to CL reserves based on claims payments data. On the other side, if one applies the CL method to the incurred losses data we obtain CL reserves based on incurred losses data. However, this strategy leads to differing ultimate claim predictions which generally do not coincide and there remains a gap between the prediction based on the cumulative payments and the prediction based on incurred losses. This gap is reduced by applying the *Munich chain ladder* (MCL) method introduced by Quarg–Mack [50] in 2004. However, this method does not completely close the gap between the predictions. Let $C^1_{i,k}$ be the cumulative claims payments and $C^2_{i,k}$ the cumulative incurred losses of accident year $i \in \{0, \ldots, I\}$ and development year $k \in \{0, \ldots, J\}$, i.e. $M = 2$ and $\mathcal{M} = \{1\}$. As shown in Merz–Wüthrich [43] the model assumptions for the MCL method can be formulated as follows:

Model Assumptions 3.8 (MCL model)

a) The sets $\{C^1_{i,j}, C^2_{i,j} |\ 0 \leq j \leq J\}$ are independent for different accident years $i \in \{0, \ldots, I\}$.

b) There exist $g^m_0, \ldots, g^m_{J-1} > 0$ and $\sigma^m_0, \ldots, \sigma^m_{J-1} > 0$ for $m \in \{1, 2\}$ such that for all

$0 \le i \le I$ *and* $0 \le k \le J-1$ *holds*

$$\mathrm{E}\big[C^1_{i,k+1}\big|\,\mathcal{C}_k\big] = g^1_k C^1_{i,k} \qquad \mathrm{Var}\big[C^1_{i,k+1}\big|\,\mathcal{C}_k\big] = \big(\sigma^1_k\big)^2 C^1_{i,k}$$
$$\mathrm{E}\big[C^2_{i,k+1}\big|\,\mathcal{I}_k\big] = g^2_k C^2_{i,k} \qquad \mathrm{Var}\big[C^2_{i,k+1}\big|\,\mathcal{I}_k\big] = \big(\sigma^2_k\big)^2 C^2_{i,k}$$

with

$$\mathcal{C}_k := \sigma(C^1_{i,j} : 0 \le i \le I,\ 0 \le j \le k) \qquad \mathcal{I}_k := \sigma(C^2_{i,j} : 0 \le i \le I,\ 0 \le j \le k).$$

c) There are constants $\lambda^1, \lambda^2 \in \mathbb{R}$ *such that*

$$\mathrm{E}\left[\frac{C^1_{i,k+1}}{C^1_{i,k}}\middle|\,\mathcal{C}_k, \mathcal{I}_k\right] = g^1_k + \lambda^1 \mathrm{Var}\left[\frac{C^1_{i,k+1}}{C^1_{i,k}}\middle|\,\mathcal{C}_k\right]^{1/2} \frac{\frac{C^2_{i,k}}{C^1_{i,k}} - \mathrm{E}\left[\frac{C^2_{i,k}}{C^1_{i,k}}\middle|\,\mathcal{C}_k\right]}{\mathrm{Var}\left[\frac{C^2_{i,k}}{C^1_{i,k}}\middle|\,\mathcal{C}_k\right]^{1/2}}$$

and

$$\mathrm{E}\left[\frac{C^2_{i,k+1}}{C^2_{i,k}}\middle|\,\mathcal{C}_k, \mathcal{I}_k\right] = g^2_k + \lambda^2 \mathrm{Var}\left[\frac{C^2_{i,k+1}}{C^2_{i,k}}\middle|\,\mathcal{I}_k\right]^{1/2} \frac{\frac{C^1_{i,k}}{C^2_{i,k}} - \mathrm{E}\left[\frac{C^1_{i,k}}{C^2_{i,k}}\middle|\,\mathcal{I}_k\right]}{\mathrm{Var}\left[\frac{C^1_{i,k}}{C^2_{i,k}}\middle|\,\mathcal{I}_k\right]^{1/2}}$$

for $0 \le i \le I$ *and* $0 \le k \le J-1$.

□

Remarks 3.9 (MCL method)

i) To the best of our knowledge, estimates for the prediction uncertainty as well as for the one-year reserving risk CDR *in terms of the (conditional) MSEP could not be derived, yet.*

ii) On the contrary to the MCL method, the extended complementary loss ratio (ECLR) method by Dahms *[16] provides one unified predictor for outstanding claims payments based on the claims payments and incurred losses data simultaneously. The corresponding prediction uncertainty and the one-year reserving risk is derived in* Dahms et al. *[18].*

iii) Another claims reserving method being able to incorporate cumulative claims payments and incurred losses data leading to one unified ultimate claim prediction is the PIC reserving method by Merz–Wüthrich *[46]. In the PIC reserving method unified ultimate claim predictors as well as the prediction uncertainty in terms of the (conditional) MSEP are derived. Moreover, the MSEP of the one-year reserving risk* CDR *and the whole predictive distribution of the* CDR *can be derived, see Chapter 5.*

Summary

In this chapter we gave a brief introduction in classical distribution-free claims reserving methods often used in actuarial practice. The CL method can cope with only one source of information (i.e. the claims payments process itself). The Bayes CL method additionally includes prior knowledge of the development factors using credibility theory. However, the main problem of the CL method remains that predictors for the outstanding loss liabilities are very sensitive w.r.t. outliers on the last observed diagonal, see (3.3) and (3.8). Thus, we proceeded with the CLR method. This method allows to incorporate prior external knowledge, for example ultimate claim estimates, as a second source of information and is more robust than the CL method w.r.t. outliers on the last observed diagonal. We saw that the well-known BF method is a special case of the CLR method. As shown in Dahms [17] all these methods can be embedded into the class of LSRMs, see Chapter 4 for details. This new general class of reserving methods provides a very flexible and powerful framework for distribution-free claims reserving modeling. Many different sources of information can be included for the prediction of outstanding loss liabilities. However, claims reserving methods which belong to the class of LSRMs do not allow for the incorporation of prior expert knowledge or industry wide data in the claims reserving process. In Gisler–Wüthrich [27] this problem is solved for the classical CL method resulting in the Bayes CL method. We consider this problem for the whole class of LSRMs and generalize the approach from the Bayes CL method to all LSRMs. This leads to the new class of Bayesian LSRMs presented in Dahms–Happ [15]. This very general and powerful class of reserving methods is subject to Chapter 4.

4 (Bayesian) Linear Stochastic Reserving Methods

4.1 Linear Stochastic Reserving Methods

As outlined in the previous chapter, many classical claims reserving methods are formulated in distribution-free model frameworks using various sources of information for the description of the stochastic dynamics of the underlying model. The most popular representatives among them are the well-known CL, BF and CLR methods. In actuarial science it was not completely understood what all these distribution-free claims reserving methods have in common. In 2012, Dahms [17] introduced the class of *linear stochastic reserving methods* (LSRMs) and pointed out that many of the well-known distribution-free claims reserving methods can be looked at as LSRMs. Among them are the CL, BF, and CLR methods presented in Chapter 3, but also the *hybrid chain ladder* (HCL) method by Arbenz–Salzmann [6] and the ECLR method by Dahms [16]. This means that from a mathematical point of view LSRMs are a state-of-the-art generalization of the models mentioned above. This provides a complete new viewing angle on the large class of distribution-free claims reserving methods. Benefits for practitioners from this generalization are listed in Conclusions 4.3.

Notational convention:
For reasons of notational consistency with Dahms [17] the counting for m starts with 0, $M \geq 0$ and the number of claim information is $M + 1$ in this chapter.

In this section we follow Dahms [17]. In the LSRM framework incremental claim information $S_{i,k}^m$ are considered for $m \in \{0, \ldots, M\}$, $i \in \{0, \ldots, I\}$ and $k \in \{0, \ldots, J\}$, i.e. we work under the extended view.

Model Assumptions 4.1 (LSRM)

a) There exist $f_k^m \in \mathbb{R}$ such that for all i, m and k the expectation of the incremental claim information $S_{i,k+1}^m$ under the condition of all information of its past $\mathcal{D}_k^{i+k}$ is proportional to an exposure $R_{i,k}^m$ contained in $\mathbb{L}^{i+k} \cap \mathbb{L}_k$, i.e.

$$\mathrm{E}\left[S_{i,k+1}^m \middle| \mathcal{D}_k^{i+k} \right] = f_k^m R_{i,k}^m \in \mathbb{L}^{i+k} \cap \mathbb{L}_k.$$

b) There exist $\sigma_k^{m_1,m_2} > 0$ such that for all i, m_1, m_2 and k the covariance of the incremental claim information $S_{i,k+1}^{m_1}$ and $S_{i,k+1}^{m_2}$ under the condition of all information of their past $\mathcal{D}_k^{i+k}$ is proportional to an exposure $R_{i,k}^{m_1,m_2}$ contained in $\mathbb{L}^{i+k} \cap \mathbb{L}_k$, i.e.

$$\mathrm{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \middle| \mathcal{D}_k^{i+k} \right] = \sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k.$$

□

For an illustration of the linear spaces (or the generated σ-fields, respectively) used in Model Assumptions 4.1, see Figure 4.1.

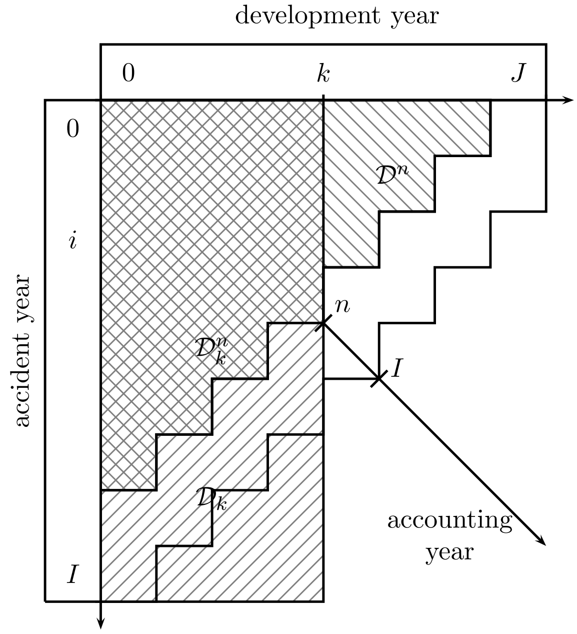

Figure 4.1: σ-fields (sets of observations): $\mathcal{D}_k$ - all claim information up to development year k, $\mathcal{D}^n$ - all claim information up to accounting year n and $\mathcal{D}_k^n$ - the union of all information in $\mathcal{D}_k$ and $\mathcal{D}^n$

Remarks 4.2 (LSRM)

i) Beside the claims payments process LSRMs can include various sources of information, for example external given exposures like prior ultimate claim estimates (BF method) or incurred losses as a second source of information (ECLR method), see Dahms *[16].*

ii) In Model Assumptions 4.1 a) and b) $R_{i,k}^m$ and $R_{i,k}^{m_1,m_2}$ are assumed to be elements out of $\mathbb{L}^{i+k} \cap \mathbb{L}_k$. That means it is implicitly assumed that there exist (unique) exposure parameters $\gamma_{i,k,h,j}^{m,l} \in \mathbb{R}$ and $\gamma_{i,k,h,j}^{m_1,m_2,l} \in \mathbb{R}$ such that

$$R_{i,k}^m = \sum_{l=1}^{M} \sum_{h=0}^{I} \sum_{j=0}^{(i+k-h)\wedge k} \gamma_{i,k,h,j}^{m,l} S_{h,j}^l \quad \text{and} \quad R_{i,k}^{m_1,m_2} = \sum_{l=1}^{M} \sum_{h=0}^{I} \sum_{j=0}^{(i+k-h)\wedge k} \gamma_{i,k,h,j}^{m_1,m_2,l} S_{h,j}^l, \tag{4.1}$$

respectively. These exposure parameters are called LSRM defining parameters, since they define the stochastic dynamics of the LSRM.

iii) *The LSRM defining parameters $\gamma_{i,k,h,j}^{m,l}$ and $\gamma_{i,k,h,j}^{m_1,m_2,l}$ in (4.1) have to be determined. There are often heuristic and/or LoB based information that motivate a specific parameter choice. An excellent example for such a scenario is given in Example 1 in* Dahms *[17]. In other cases one can use backtesting techniques for verifying whether the LSRM would have worked well in the past.*

4.1.1 Classical Claims Reserving Methods as LSRMs

To get an idea of the flexibility of the class of LSRMs we first analyze which of the classical claims reserving methods belong to the class of LSRMs.

CL Method

Model Assumptions 3.1 for the distribution-free CL method are:
There exist constant factors $g_0, \dots, g_{J-1}$ and variance parameters $\sigma_0^2, \dots, \sigma_{J-1}^2 > 0$ such that

a) Cumulative claims payments $\{C_{i,k} | \ 0 \le k \le J\}$ from different accident years $i \in \{0, \dots, I\}$ are independent.

b) For all $0 \le i \le I$ and $0 \le k \le J-1$ holds

$$\begin{aligned} \mathrm{E}[C_{i,k+1} | \mathcal{B}_{i,k}] &= \mathrm{E}[C_{i,k+1} | C_{i,k}] = g_k C_{i,k}, \\ \mathrm{Var}[C_{i,k+1} | \mathcal{B}_{i,k}] &= \mathrm{Var}[C_{i,k+1} | C_{i,k}] = \sigma_k^2 C_{i,k}. \end{aligned}$$

Since in the CL method only one source of information, namely the claims payments process itself, is incorporated we have that $M = 0$. The incremental claims payments in the CL method are given by

$$S_{i,k+1}^0 := C_{i,k+1} - C_{i,k}.$$

With

$$R_{i,k}^0 := C_{i,k} = \sum_{j=0}^{k} S_{i,j}^0 =: R_{i,k}^{0,0} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k \tag{4.2}$$

and with Model Assumptions 3.1 a) and b) we obtain

$$\begin{aligned} \mathrm{E}\Big[S_{i,k+1}^0 \Big| \mathcal{D}_k^{i+k} \Big] &= \mathrm{E}\big[S_{i,k+1}^0 \big| \mathcal{B}_{i,k} \big] = \underbrace{(g_k - 1)}_{=: f_k^0} \underbrace{C_{i,k}}_{= R_{i,k}^0} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k, \\ \mathrm{Var}\Big[S_{i,k+1}^0 \Big| \mathcal{D}_k^{i,k} \Big] &= \mathrm{Var}\big[S_{i,k+1}^0 \big| \mathcal{B}_{i,k} \big] = \underbrace{\sigma_k^2}_{=: \sigma_k^{0,0}} \underbrace{C_{i,k}}_{= R_{i,k}^{0,0}} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k. \end{aligned}$$

This shows that the CL method belongs to the class of LSRMs.

□

CLR Method

Model Assumptions 3.5 for the CLR method are given by:

a) Incremental claims payments $\{S^0_{i,k} | \, 0 \le i \le I, 0 \le k \le J\}$ are independent.

b) For $0 \le i \le I$ and $0 \le k \le J-1$ holds

$$\begin{aligned} \mathrm{E}\big[S^0_{i,k+1} \big| \, \mathcal{B}_{i,k} \big] &= g_k P_i, \\ \mathrm{Var}\big[S^0_{i,k+1} \big| \, \mathcal{B}_{i,k} \big] &= \sigma_k^2 P_i. \end{aligned}$$

If we take

$$S^1_{i,k} := \begin{cases} P_i & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

we see that the CLR method belongs to the class of LSRMs. Since the BF method is a special case of the CLR method, this implies that the BF method also belongs to the class of LSRMs. As already mentioned above the ECLR method can also be looked at as a LSRM, see Dahms [17].

4.1.2 Parameter Estimation for LSRMs

In the LSRMs defined above the model parameters f^m_k and $\sigma_k^{m_1,m_2}$ are unknown and have to be estimated from the data. Note that the LSRM defining exposure parameters $\gamma^{m,l}_{i,k,h,j} \in \mathbb{R}$ and $\gamma^{m_1,m_2,l}_{i,k,h,j} \in \mathbb{R}$ in (4.1) are required by the model to provide a well-defined LSRM, see Remarks 4.2. An ($\mathcal{D}_k$ conditional) unbiased estimator for f^m_k is given by (we set $\frac{0}{0} := 0$)

$$\widehat{f}^m_k := \sum_{i=0}^{I-1-k} w^m_{i,k} \frac{S^m_{i,k+1}}{R^m_{i,k}}, \tag{4.3}$$

where $w^m_{i,k} \ge 0$ are $\mathcal{D}^n \cap \mathcal{D}_k$-measurable weights with

i) $R^m_{i,k} = 0$ implies $w^m_{i,k} = 0$ and

ii) $\sum\limits_{i=0}^{I-1-k} w^m_{i,k} = 1$ if at least one $R^m_{i,k} \ne 0$.

With the choice of explicit weights

$$w^m_{i,k} := \frac{\left(R^m_{i,k}\right)^2}{R^{m,m}_{i,k}} \left(\sum_{h=0}^{I-1-k} \frac{\left(R^m_{h,k}\right)^2}{R^{m,m}_{h,k}} \right)^{-1} \tag{4.4}$$

the estimators (4.3) have minimum variance in the class of all linear estimators of the form (4.3), i.e. they are (homogeneous) credibility estimators, see Bühlmann–Gisler [14] for a definition. In the case of the CL method, the CLR method or the ECLR method these minimum variance estimators are the well-known standard estimators, see for example Mack [38] and Dahms [16]. Estimators for the second model parameter $\sigma_k^{m_1,m_2}$ are not required for the prediction of claim information $S_{i,k}^m$ in LSRMs, but for the quantification of the prediction uncertainty. An unbiased estimator for $\sigma_k^{m_1,m_2}$ is given by

$$\widehat{\sigma}_k^{m_1,m_2} := \frac{1}{Z_k^{m_1,m_2}} \sum_{i=0}^{I-1-k} \frac{R_{i,k}^{m_1} R_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}} \left(\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \widehat{f}_k^{m_1} \right) \left(\frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \widehat{f}_k^{m_2} \right), \tag{4.5}$$

where

$$Z_k^{m_1,m_2} := \sum_{i=0}^{I-1-k} \left(1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + w_{i,k}^{m_1} w_{i,k}^{m_2} \frac{R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}} \sum_{h=0}^{I-1-k} \frac{R_{h,k}^{m_1} R_{h,k}^{m_2}}{R_{h,k}^{m_1,m_2}} \right)$$

with

$$w_{i,j}^m := \frac{(R_{i,j}^m)^2}{R_{i,j}^{m,m}} \left(\sum_{h=0}^{I-1-j} \frac{(R_{h,j}^m)^2}{R_{h,j}^{m,m}} \right)^{-1}.$$

For a proof of the stated unbiasedness of the estimators (4.3) and (4.5), see Dahms [17].

4.1.3 Prediction of Future Claim Information

By (4.1) it becomes evident that $R_{i,k}^m$ is a linear combination of $S_{h,j}^l \in \mathbb{L}^{i+k} \cap \mathbb{L}_k$. This implies that the linear projection

$$\mathbf{P}^n \colon \mathbb{L}^n \longrightarrow \mathbb{L}^{n+1},\ \mathbf{x} \longmapsto (\mathbf{P}^n \mathbf{x})_{i,k}^m := \begin{cases} x_{i,k}^m & \text{for } i+k \le n \\ P_{i,k-1}^m \mathbf{x} & \text{for } i+k = n+1 \end{cases},$$

where

$$P_{i,k}^m \colon \mathbb{L}^{i+k} \longrightarrow \mathbb{R},\ \mathbf{x} \longmapsto P_{i,k}^m \mathbf{x} := f_k^m \mathbf{\Gamma}_{i,k}^m \mathbf{x}$$

with

$$\mathbf{\Gamma}_{i,k}^m \colon \mathbb{L}^{i+k} \longrightarrow \mathbb{R},\ \mathbf{x} \longmapsto \mathbf{\Gamma}_{i,k}^m \mathbf{x} := \sum_{l=1}^{M} \sum_{h=0}^{I} \sum_{j=0}^{(n-h)\wedge(k)} \gamma_{i,k,h,j}^{m,l} x_{h,j}^l$$

generates, based on all claim information in $\mathcal{D}^n$, the next diagonal (accounting year) $n+1$ and projects all claim information out of $\mathcal{D}^n$ on this next diagonal. The concatenation of these linear projections fills up several diagonals at once and is given by

$$\mathbf{P}^{n_2 \leftarrow n_1} \colon \mathbb{L}^{n_1} \longrightarrow \mathbb{L}^{n_2+1},\ \mathbf{x} \longmapsto \mathbf{P}^{n_2 \leftarrow n_1} \mathbf{x} := \begin{cases} \Pi_{\mathbb{L}^{n_2+1}} \mathbf{x} & \text{for } n_2 < n_1 \\ \mathbf{P}^{n_2} \mathbf{P}^{n_2-1} \cdots \mathbf{P}^{n_1} \mathbf{x} & \text{for } n_2 \ge n_1 \end{cases}$$

with $\Pi_{\mathbb{L}^n}$ denoting the projection on the first n diagonals and we define

$$\mathbf{P}_{i,k}^{m|n} \colon \mathbb{L}^n \longrightarrow \mathbb{R},\ \mathbf{x} \longmapsto \mathbf{P}_{i,k}^{m|n}\mathbf{x} := \left(\mathbf{P}^{i+k \leftarrow n}\mathbf{x}\right)_{i,k+1}^{m}. \tag{4.6}$$

With

$$\mathbf{S}^n := \left(S_{i,k}^m\right)_{i+k\leq n}^{0\leq m\leq M}$$

we obtain for $1 \leq i \leq I$ and $k+n+1 \leq J$ the best predictor for $S_{i,k+n+1}^m$ given by

$$\mathrm{E}\left[S_{i,k+n+1}^m \,\middle|\, \mathcal{D}_k^{i+k} \right] = \mathbf{P}_{i,k+n}^{m|i+k}\mathbf{S}^{i+k}. \tag{4.7}$$

Replacing the unknown development factors f_k^m in (4.6) by their estimates $\widehat{f}_k^m$ we obtain with (4.7) for $I-i \leq k < J$ the LSRM predictor

$$\widehat{S}_{i,k+1}^{m|I} := \widehat{\mathbf{P}}_{i,k}^{m|I}\mathbf{S}^I. \tag{4.8}$$

Beside the LSRM predictor (4.8) the prediction uncertainty of $\widehat{S}_{i,k+1}^{m|I}$ and $\mathrm{CDR}^{\mathcal{M},I+1}$ quantified by means of the MSEP are of interest. However, this requires cumbersome notation and long calculations and we refer therefore to Dahms [17].

Conclusion 4.3

i) In most distribution-free claims reserving methods accident year independence is a central model assumption. This is not required in LSRMs.

ii) LSRMs possess (conditional) uncorrelated diagonals, see Lemma 2.3 c) in Dahms *[17]. This implies that calendar year effects like inflation which have an impact on the whole diagonal can not be captured by LSRMs. This issue should be subject to further research.*

iii) A regime change can be modeled by an exchange of external given exposures leading to a much faster calibration of the model to the new regime than in the CL method. One example for such an exposure change is illustrated in Example 1 in Dahms *[17].*

In LSRMs there is no mathematically consistent way to incorporate prior expert knowledge or industry-wide data knowledge into the development factors f_k^m. This is exactly the same issue that is considered in the Bayes CL method in Gisler–Wüthrich [27] for the CL method. We generalize this aspect to the whole class of LSRMs what is subject to the next section.

4.2 Bayesian Linear Stochastic Reserving Methods

In insurance practice some LoBs show large deviations and irregularities in the claims development. Therefore, it is difficult for a reserving actuary to find reliable estimates for the development factors for his model. In classical LSRMs the development factors f_k^m are estimated in

(4.3) based on data of the trapezoids only. However information about "typical" development factors in similar LoBs or from industry-wide data could be helpful and included in the estimation of the development factors. To solve this problem one has to ask the following question: What can we learn from the collective (industry-wide data) for the development pattern of the LoB under consideration? This question has its origin in credibility theory where knowledge of the collective and individual loss records are combined to improve estimates, see Bühlmann–Gisler [14]. In our reserving context we combine experience of the industry-wide data and the individual LoB's claims record to calculate credibility development factors. This leads to the class of Bayesian LSRMs. In this section we follow Dahms-Happ [15].

In the Bayesian LSRM setup we assume that the development factor f_k^m of claim information $m \in \{0, \dots M\}$ in development year $k \in \{0, \dots J-1\}$ is a realization of a random variable F_k^m. We denote the random matrix containing all development factors F_k^m by

$$\mathbf{F} := (F_k^m)_{0 \le k \le J-1}^{0 \le m \le M} . \tag{4.9}$$

The vector collecting all development factors of development year $k \in \{0, \dots, J-1\}$ is denoted by

$$\mathbf{F}_k := (F_k^0, \dots, F_k^M)'$$

and we define in an analog way

$$\mathbf{f} := (f_k^m)_{0 \le k \le J-1}^{0 \le m \le M} \in \mathbb{R}^{(J) \times (M+1)}$$

and

$$\mathbf{f}_k := (f_k^0, \dots, f_k^M)' \in \mathbb{R}^{M+1}.$$

Moreover, for an arbitrary mapping

$$h_{\mathbf{a}} : A \longrightarrow B, \; x \longmapsto h_{\mathbf{a}}(x),$$

(for arbitrary sets A and B) depending on a fixed parameter (vector) $\mathbf{a}$ we denote by

$$h|_{\mathbf{a}=\mathbf{b}}(x) := h_{\mathbf{b}}(x)$$

the function h, but with $\mathbf{a}$ replaced by $\mathbf{b}$.

For the construction of Bayesian LSRMs we assume that conditionally, given $\mathbf{F}$, Model Assumptions 4.1 for classical LSRMs are fulfilled.

Model Assumptions 4.4 (Bayesian LSRM)

a) Conditionally, given $\mathbf{F} = (F_k^m)_{0\le k\le J-1}^{0\le m\le M}$*, for all* i*,* m *and* k *the expectation of the incremental claim information* $S_{i,k+1}^m$ *under the condition of all information of its past* $\mathcal{D}_k^{i+k}$ *is proportional to an exposure* $R_{i,k}^m \in \mathbb{L}^{i+k} \cap \mathbb{L}_k$*, i.e.*

$$\mathrm{E}\left[S_{i,k+1}^m \middle| \mathcal{D}_k^{i+k}, \mathbf{F} \right] = F_k^m R_{i,k}^m.$$

b) Conditionally, given $\mathbf{F}$*, the covariance of the incremental claim information* $S_{i,k+1}^{m_1}$ *and* $S_{i,k+1}^{m_2}$ *under the condition of all information of their past* $\mathcal{D}_k^{i+k}$ *is proportional to an exposure* $R_{i,k}^{m_1,m_2} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k$*, i.e.*

$$\mathrm{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \middle| \mathcal{D}_k^{i+k}, \mathbf{F} \right] = \sigma_k^{m_1,m_2}(\mathbf{F}) R_{i,k}^{m_1,m_2} \quad \text{with} \quad R_{i,k}^{m_1,m_2} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k.$$

c) For all $n \in \{I, \ldots, I+J\}$*,* $j \le J-1$ *and* $0 \le k_0 < k_1 < \ldots < k_j \le J-1$ *it holds*

$$\mathrm{E}\left[\prod_{i=0}^{j} \Omega_{k_i} \middle| \mathcal{D}^n \right] = \prod_{i=0}^{j} \mathrm{E}[\Omega_{k_i} | \mathcal{D}^n],$$

where, for $k \in \{0, \ldots, J-1\}$*,* $\Omega_k \in \left\{ F_k^m, \sigma_k^{m_1,m_2}(\mathbf{F}), F_k^{m_1} F_k^{m_2} \;\middle|\; 0 \le m, m_1, m_2 \le M \right\}$*.*

□

Remarks 4.5 (Bayesian LSRM)

i) The covariance coefficients $\sigma_k^{m_1,m_2}(\mathbf{F})$ *in Model Assumptions 4.4 b) have to be chosen, so that* $\left(\sigma_k^{m_1,m_2}(\mathbf{F}) R_{i,k}^{m_1,m_2} \right)_{m_1,m_2}$ *is positive semidefinite almost sure for all* i *and* k*.*

ii) Model Assumption 4.4 c) is a kind of uncorrelatedness assumption, which is actually less restrictive than a priori independence of $\{\mathbf{F}_k : 0 \le k \le J-1\}$*. Indeed, one can show that unconditional independence of the development factors* $\{\mathbf{F}_k : 0 \le k \le J-1\}$ *together with Model Assumption 4.4 a) satisfy c), if in b)* $\sigma_k^{m_1,m_2}(\mathbf{F})$ *depends on* F_k *only and in a) the slightly stronger assumption holds that not only the expected value but the distribution of* $S_{i,k+1}^m$ *depends on* $\mathbf{F}$ *and* $\mathcal{D}_k^{i+k}$ *via* F_k^m *and* $R_{i,k}^m$ *only (compare this with the Model Assumptions 3.3 for the Bayes CL method).*

iii) The Bayes CL method presented in Gisler-Wüthrich *[27] belongs to the class of Bayesian LSRMs, see the next subsection.*

4.2.1 Classical Bayesian Claims Reserving Methods as Bayesian LSRMs

We first analyze which of the classical Bayesian claims reserving methods belong to the class of Bayesian LSRMs.

Bayes CL Method

Model Assumptions 3.3 for the Bayes CL method are:

a) *Conditionally, given* $\mathbf{G} := (G_0, \ldots, G_{J-1})$, *the cumulative claims payments* $\{C_{i,k} | \, 0 \le k \le J\}$ *from different accident years* $i \in \{0, \ldots, I\}$ *are independent.*

b) *Conditionally, given* $\mathbf{G}$ *and* $\mathcal{B}_k$, *the distribution of* $Y_{i,k}$ *depends only on* $C_{i,k}$ *and it holds*

$$\begin{aligned} \mathrm{E}[Y_{i,k} | \, \mathbf{G}, \mathcal{B}_{i,k}] &= G_k \\ \mathrm{Var}[Y_{i,k} | \, \mathbf{G}, \mathcal{B}_{i,k}] &= \frac{\sigma_k^2(G_k)}{C_{i,k}}, \end{aligned}$$

for $i \in \{0, \ldots, I\}$ *and* $k \in \{0, \ldots, J-1\}$.

c) $\{G_0, G_1, \ldots, G_{J-1}\}$ *are independent.*

In the Bayes CL method we have that $M = 0$. The incremental claims payments are given by

$$S_{i,k+1}^0 := C_{i,k+1} - C_{i,k}.$$

We define componentwise $\mathbf{F} := \mathbf{G} - 1$ and

$$R_{i,k}^0 := C_{i,k} = \sum_{j=0}^{k} S_{i,j}^0 =: R_{i,k}^{0,0} \in \mathbb{L}^{i+k} \cap \mathbb{L}_k. \tag{4.10}$$

With Model Assumptions 3.3 a) and b) we obtain

$$\begin{aligned} \mathrm{E}\Big[S_{i,k+1}^0 \Big| \, \mathcal{D}_k^{i+k}, \mathbf{F} \Big] &= \mathrm{E}\Big[S_{i,k+1}^0 \Big| \, \mathcal{D}_k^{i+k}, \mathbf{G} \Big] = \mathrm{E}\big[S_{i,k+1}^0 \big| \, \mathcal{B}_{i,k}, \mathbf{G} \big] = (G_k - 1) C_{i,k} = F_k^0 R_{i,k}^0 \\ \mathrm{Var}\Big[S_{i,k+1}^0 \Big| \, \mathcal{D}_k^{i,k}, \mathbf{F} \Big] &= \mathrm{Var}\big[S_{i,k+1}^0 \big| \, \mathcal{B}_{i,k}, \mathbf{G} \big] = \sigma_k^2(G_k) C_{i,k} = \underbrace{\sigma_k^2(F_k^0 + 1)}_{=: \, \sigma_k^{0,0}(F_k^0)} C_{i,k} = \sigma_k^{0,0}(F_k^0) R_{i,k}^{0,0}. \end{aligned}$$

This shows that Model Assumption 4.4 a) and b) are fulfilled and c) follows by Theorem 3.2 in GISLER–WÜTHRICH [27]. It follows that the Bayes CL method belongs to the class of Bayesian LSRMs.

□

Credibility-Based Additive Loss Reserving Method

The credibility-based ALR method in MERZ–WÜTHRICH [44] is based on the classical ALR method, which is a special case of the CLR method, see Section 3.4. Since the CLR method belongs to the class of LSRMs we expect that the credibility-based ALR method also belongs to class of Bayesian LSRMs. The model assumptions of the credibility-based ALR method are given by:

Model Assumptions 4.6 (Credibility-based ALR model) *There exist constant positive volume measures* $V_0, \ldots, V_I$ *such that*

a) Conditionally, given $\mathbf{G} = (G_0, \ldots, G_J)$ *the incremental claim payments* $\{S^0_{i,k} |\ 0 \le i \le I, 0 \le k \le J\}$ *are independent.*

b) Conditionally, given $\mathbf{G}$*, the distribution of* $S^0_{i,k+1}$ *only depends on* G_k *and the constant* V_i*, and for all* $i = 0, \ldots, I$ *and* $k = 0, \ldots, J$ *holds*

$$\begin{aligned} \mathrm{E}\big[S^0_{i,k+1}\big|\, \mathbf{G}\big] &= G_{k+1} V_i \\ \mathrm{Var}\big[S^0_{i,k+1}\big|\, \mathbf{G}\big] &= \sigma^2_{k+1}(G_{k+1}) V_i. \end{aligned}$$

c) $\{G_0, \ldots, G_J\}$ *are independent with prior distributions* $U(g)$*.*

□

In the credibility-based ALR method we have $M = 1$ and $\mathcal{M} = \{0\}$. We define

$$S^1_{i,k} := \begin{cases} V_i & \text{for } k = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Setting $F^0_k := G_{k+1}$ for $k = 0 \ldots, J-1$ and $\sigma^{0,0}_k(F_k) := \sigma^2_{k+1}(G_{k+1})$ we obtain

$$\begin{aligned} \mathrm{E}\Big[S^0_{i,k+1}\Big|\, \mathcal{D}^{i+k}_k, \mathbf{F}\Big] &= \mathrm{E}\Big[S^0_{i,k+1}\Big|\, \mathcal{D}^{i+k}_k, \mathbf{G}\Big] = G_{k+1} V_i = F^0_k V_i \\ \mathrm{Var}\Big[S^0_{i,k+1}\Big|\, \mathcal{D}^{i+k}_k, \mathbf{F}\Big] &= \mathrm{Var}\Big[S^0_{i,k+1}\Big|\, \mathcal{D}^{i+k}_k, \mathbf{G}\Big] = \sigma^2_{k+1}(G_{k+1}) V_i = \sigma^{0,0}_k(F^0_k) V_i. \end{aligned}$$

Model Assumptions 4.4 a) and b) are satisfied, since $V_i \in \mathbb{L}^{i+k} \cap \mathbb{L}_k$. Model Assumption 4.4 c) follows by Theorem 3.3 in Merz–Wüthrich [44]. This shows that the credibility-based ALR method also belongs to the class of Bayesian LSRMs.

□

4.2.2 Prediction of Future Claim Information

The Bayesian LSRM is constructed in such a way that conditionally, given $\mathbf{F}$, Model Assumptions 4.1 of the classical LSRM are fulfilled. This implies that, in analogy to (4.7) for classical LSRMs, we obtain for Bayesian LSRMs for $i \in \{0, \ldots, I\}$, $i + k \ge I$, $k + n + 1 \le J$ and $m \in \{0, \ldots, M\}$

$$\widehat{S}^{m|i+k,\mathbf{F}}_{i,k+n+1} := \mathrm{E}\Big[S^m_{i,k+n+1}\Big|\, \mathcal{D}^{i+k}, \mathbf{F}\Big] = \mathbf{P}^{m|i+k,\mathbf{F}}_{i,k+n} \mathbf{S}^{i+k}, \tag{4.11}$$

where $\mathbf{P}^{m|n,\mathbf{F}}_{i,k} := \mathbf{P}^{m|n}_{i,k}\Big|_{\mathbf{f}=\mathbf{F}}$. An immediate consequence is the following result:

Predictor 4.7 (Bayesian predictor of $S^m_{i,k+1}$ at time I) *Under Model Assumptions 4.4 we obtain for $i \in \{0, \dots, I\}$, $i + k \in \{I, \dots, I + J - 1\}$, $k + n + 1 \le J$ and $m \in \{0, \dots, M\}$*

$$\mathrm{E}\Big[S^m_{i,k+n+1} \Big| \mathcal{D}^{i+k} \Big] = \mathbf{P}^{m|i+k,Bayes}_{i,k+n} \mathbf{S}^{i+k},$$

where $\mathbf{P}^{m|n,Bayes}_{i,k} := \mathbf{P}^{m|n}_{i,k} \Big|_{\mathbf{f} = \mathrm{E}[\mathbf{F} | \mathcal{D}^n]}$.

Proof: Conditionally, given $\mathcal{D}^{i+k}$ and $\mathbf{F}$, we obtain with (4.11)

$$\mathrm{E}\Big[S^m_{i,k+n+1} \Big| \mathcal{D}^{i+k}, \mathbf{F} \Big] = \mathbf{P}^{m|i+k,\mathbf{F}}_{i,k+n} \mathbf{S}^{i+k}.$$

Because each mapping $P^m_{i,k}$ is linear in f^m_j, we get with the second equation in (4.11) and Model Assumption 4.4 c)

$$\begin{aligned}\mathrm{E}\Big[S^m_{i,k+n+1} \Big| \mathcal{D}^{i+k} \Big] &= \mathrm{E}\Big[\mathrm{E}\Big[S^m_{i,k+n+1} \Big| \mathcal{D}^{i+k}, \mathbf{F} \Big] \Big| \mathcal{D}^{i+k} \Big] \\ &= \mathrm{E}\Big[\mathbf{P}^{m|i+k,\mathbf{F}}_{i,k+n} \mathbf{S}^{i+k} \Big| \mathcal{D}^{i+k} \Big] \\ &= \mathbf{P}^{m|i+k,Bayes}_{i,k+n} \mathbf{S}^{i+k}.\end{aligned}$$

□

4.2.3 Credibility for Linear Stochastic Reserving Methods

Our goal in this section is to derive a predictor for the unknown incremental claim information $S^m_{i,k+1}$ for $i+k \ge I$. It is a well known result in probability theory that the best square-integrable predictor for the incremental claim information $S^m_{i,k+1}$, given the data $\mathcal{D}^I$ at time I, is given by

$$\widehat{S}^{m|I,Bayes}_{i,k+1} := \mathrm{E}\big[S^m_{i,k+1} \big| \mathcal{D}^I \big] = \mathbf{P}^{m|I,Bayes}_{i,k} \mathbf{S}^I. \tag{4.12}$$

In order to calculate $\mathbf{P}^{m|I,Bayes}_{i,k}$ one needs to know the joint prior distribution of $\mathbf{F}$ and the explicit form of the conditional joint distribution of $S^m_{i,k}$ for $i + k \le I$, given $\mathbf{F}$. These distributions are often unknown in practice and it is not obvious, how reasonable estimates of these distributions can be derived. Thus, we choose a so-called credibility based approach, where only the first two moments (or appropriate estimates) of the conditional distribution of $\mathbf{F}$, given $\mathcal{D}_k$, are required. That means instead of calculating $\mathrm{E}\big[\mathbf{F} \big| \mathcal{D}^I\big]$ contained in $\mathbf{P}^{m|I,Bayes}_{i,k}$ (cf. (4.12)) we use a so-called credibility predictor for $\mathbf{F}_k$ for $k \in \{0, \dots, J-1\}$. That means that for all development years $k \in \{0, \dots, J-1\}$ we replace $\big(\mathrm{E}\big[\mathbf{F} \big| \mathcal{D}^I\big]\big)_k = \mathrm{E}\big[\mathbf{F}_k \big| \mathcal{D}^I\big]$ by best predictors, which are affine-linear in the observations

$$\mathbf{Y}_{i,k} := \big(Y^0_{i,k}, Y^1_{i,k}, \dots, Y^M_{i,k}\big)' \quad \text{with} \quad Y^m_{i,k} := \frac{S^m_{i,k+1}}{R^m_{i,k}} \quad \text{for} \quad i = 0, \dots, I-k-1.$$

This implies that at time I we use best predictors from the linear subspace

$$\mathbb{L}^{ind}(\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k}) := \left\{ \widehat{\mathbf{F}}_k \middle| \widehat{\mathbf{F}}_k = \mathbf{a}_0^k + \sum_{i=0}^{I-k-1} \mathbf{A}_i \mathbf{Y}_{i,k},\ \mathbf{a}_0^k \in \mathbb{R}^{(M+1)}, \mathbf{A}_i \in \mathbb{R}^{(M+1)\times(M+1)} \right\},$$

which contains all affine-linear combinations of the observations $\mathbf{Y}_{i,k}$ with $i \in \{0,\dots,I-k-1\}$ at time I. The best predictor from this subspace is given by

Definition 4.8 (Credibility predictor of $\mathbf{F}_k$ at time I) *The credibility predictor at time I for the development factor $\mathbf{F}_k$ for $k \in \{0,\dots,J-1\}$ is defined by*

$$\widehat{\mathbf{F}}_k^{I,Cred} := \mathrm{Pro}(\mathbf{F}_k |\ \mathbb{L}^{ind}(\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k})),$$

where $\mathrm{Pro}(\mathbf{F}_k |\ \mathbb{L}^{ind}(\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k}))$ *denotes the orthogonal projection operator on the linear subspace* $\mathbb{L}^{ind}(\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k})$*.*

□

For many derivations it is easier to work with the following so-called "normal equations" which are necessary and sufficient conditions for orthogonal projections.

Lemma 4.9 (Normal equations) *A predictor* $\widehat{\mathbf{F}}_k \in \mathbb{L}^{ind}(\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k})$ *is the credibility predictor for* $\mathbf{F}_k$*, i.e.* $\widehat{\mathbf{F}}_k = \widehat{\mathbf{F}}_k^{I,Cred}$*, if and only if for* $i \in \{0,\dots,I-k-1\}$ *the following equations hold true:*

1. $\mathrm{E}\left[\left(\widehat{\mathbf{F}}_k - \mathbf{F}_k\right)\mathbf{Y}'_{i,k}\right] = \mathbf{0} \in \mathbb{R}^{(M+1)\times(M+1)} \quad \left(\textit{denoted by: } \mathbf{F}_k - \widehat{\mathbf{F}}_k^{I,Cred} \perp \mathbf{Y}_{i,k}\right)$,
2. $\mathrm{E}\left[\left(\widehat{\mathbf{F}}_k - \mathbf{F}_k\right)' \cdot \mathbf{1}\right] = \mathbf{0} \in \mathbb{R}^{(M+1)\times(M+1)}$.

Proof: For a proof of Lemma 4.9 we refer to Brockwell-Davis [9]. □

Lemma 4.9 is often called Hilbert projection theorem. We will use the normal equations for the derivation of the credibility predictor $\widehat{\mathbf{F}}_k^{I,\mathrm{Cred}}$ in Theorems 4.13 and 4.15 below.

In the same way as in (4.9) we collect all credibility predictors $\widehat{F}_k^{m|I,Cred} := \left(\widehat{\mathbf{F}}_k^{I,Cred}\right)_m$ for claim information $m \in \{0,\dots,M\}$ and development year $k \in \{0,\dots,J-1\}$ in $\widehat{\mathbf{F}}^{I,Cred}$ defined by

$$\widehat{\mathbf{F}}^{I,Cred} := \left(\widehat{F}_k^{m|I,Cred}\right)_{0\le k\le J-1}^{0\le m\le M}.$$

Definition 4.10 (Credibility based predictor of $S_{i,k+1}^m$ at time I) *The credibility based predictor of the incremental claim information* $S_{i,k+1}^m$ *at time I is given by*

$$\widehat{S}_{i,k+1}^{m|I,(Cred)} := \mathbf{P}_{i,k}^{m|I,Cred}\mathbf{S}^I,$$

where $\mathbf{P}_{i,k}^{m|I,Cred} := \mathbf{P}_{i,k}^{m|I}\Big|_{\mathbf{f}=\widehat{\mathbf{F}}^{I,Cred}}$. □

Remarks 4.11 (Credibility based predictor)

i) Credibility predictors are the best predictors, which are affine-linear in the observations. We base the prediction of F_k^m *at time* I *on the normalized observations* $\{\mathbf{Y}_{i,k} | i = 0, 1, \dots, I - k - 1\}$*, since they are the only observations containing information on* $\mathbf{F}_k$*.*

ii) For the credibility predictor it holds

$$\widehat{\mathbf{F}}_k^{I,Cred} := \underset{\widehat{\mathbf{F}}_k \in \mathbb{L}^{ind}(\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k})}{\arg\min} \mathrm{E}\left[\left(\widehat{\mathbf{F}}_k - \mathbf{F}_k \right)' \left(\widehat{\mathbf{F}}_k - \mathbf{F}_k \right) \middle| \mathcal{D}_k \right],$$

i.e. $\widehat{\mathbf{F}}_k^{I,Cred}$ *minimizes the conditional, given* $\mathcal{D}_k$*, MSEP in the linear subspace* $\mathbb{L}^{ind}(\mathbf{Y}_{0,k}, \dots, \mathbf{Y}_{I-k-1,k})$*.*

iii) The superscript "Cred" in brackets in Definition 4.10 means that $S_{i,k+1}^{m|I,(Cred)}$ *is a predictor based on the credibility predictors* $\widehat{\mathbf{F}}^{I,Cred}$ *and not a credibility predictor itself. This follows from the definition, because credibility predictors are affine-linear functions of the observations. In our multiplicative model structure, it would not make sense to restrict to affine-linear predictors of* $S_{i,k}^m$*.*

In order to calculate the credibility based predictor $\widehat{S}_{i,k+1}^{m|I,(Cred)}$ stated in Definition 4.10 we have to derive the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ for $\mathbf{F}_k$ given in Definition 4.8.
We start our calculations with some basic results on $\mathbf{Y}_{i,k}$:

Lemma 4.12 *Under Model Assumptions 4.4 it holds for* $i \in \{0, \dots, I\}$*,* $k \in \{0, \dots, J-1\}$ *and* $m_1, m_2 \in \{0, \dots, M\}$*:*

i) $\mathbf{Y}_{i,k}$ *are conditionally, given* $\mathcal{D}_k$ *and* $\mathbf{F}$*, unbiased predictors for* $\mathbf{F}_k$*, i.e.*

$$\mathrm{E}[\mathbf{Y}_{i,k} | \mathcal{D}_k, \mathbf{F}] = \mathbf{F}_k.$$

ii) $\mathbf{Y}_{i,k}$ *are conditionally, given* $\mathcal{D}_k$ *and* $\mathbf{F}$*, uncorrelated for different accident years, i.e.*

$$\mathrm{Cov}[\mathbf{Y}_{i_1,k}, \mathbf{Y}_{i_2,k} | \mathcal{D}_k, \mathbf{F}] = \mathbf{0} \in \mathbb{R}^{(M+1)\times(M+1)} \quad \text{for} \quad i_1 \neq i_2.$$

iii) We have

$$\mathrm{Cov}\left[Y_{i,k}^{m_1}, Y_{i,k}^{m_2} \middle| \mathcal{D}_k, \mathbf{F} \right] = \frac{R_{i,k}^{m_1,m_2} \sigma_k^{m_1,m_2}(\mathbf{F})}{R_{i,k}^{m_1} R_{i,k}^{m_2}} =: (\mathbf{\Sigma}_{i,k}(\mathbf{F}))_{m_1,m_2}. \tag{4.13}$$

Proof: The first and third claim is a direct consequence of Model Assumptions 4.4. For the second claim we assume without loss of generality that $i_1 < i_2$. Then $\mathbf{Y}_{i_1,k}$ is $\mathcal{D}_k^{i_2+k}$-measurable

and

$$\begin{aligned}\mathrm{Cov}[\mathbf{Y}_{i_1,k},\mathbf{Y}_{i_2,k}|\,\mathcal{D}_k,\mathbf{F}] &= \mathrm{E}\Big[\mathrm{Cov}\Big[\mathbf{Y}_{i_1,k},\mathbf{Y}_{i_2,k}|\,\mathcal{D}_k^{i_2+k},\mathbf{F}\Big]\Big|\,\mathcal{D}_k,\mathbf{F}\Big] \\ &\quad + \mathrm{Cov}\Big[\mathrm{E}\Big[\mathbf{Y}_{i_1,k}|\,\mathcal{D}_k^{i_2+k},\mathbf{F}\Big],\mathrm{E}\Big[\mathbf{Y}_{i_2,k}|\,\mathcal{D}_k^{i_2+k},\mathbf{F}\Big]\Big|\,\mathcal{D}_k,\mathbf{F}\Big] \\ &= \mathbf{0} + \mathrm{Cov}[\mathbf{Y}_{i_1,k},\mathbf{F}_k|\,\mathcal{D}_k,\mathbf{F}] = \mathbf{0}.\end{aligned}$$

□

By Definition 4.8 the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ for the development factor $\mathbf{F}_k$ is an affine-linear predictor in the normalized observations $\mathbf{Y}_{i,k}$. We compress the data $\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k}$ linearly to one $(M+1)$-dimensional vector $\mathbf{C}_k$ (see Theorem 4.13) and show that the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ depends on the data $\mathbf{Y}_{0,k},\dots,\mathbf{Y}_{I-k-1,k}$ only via $\mathbf{C}_k$, i.e. $\mathbf{C}_k$ is a sufficient statistics (see Theorem 4.15). For that reason, we define

$$\mathbf{C}_k := \mathbf{W}_{\bullet,k}^{-1}\sum_{i=0}^{I-k-1}\mathbf{W}_{i,k}\mathbf{Y}_{i,k}, \tag{4.14}$$

where

$$\mathbf{W}_{\bullet,k}^{-1} := \left(\sum_{i=0}^{I-k-1}\mathbf{W}_{i,k}\right)^{-1} \quad\text{and}\quad \mathbf{W}_{i,k} := \mathrm{diag}\left(\frac{R_{i,k}^0R_{i,k}^0}{R_{i,k}^{0,0}},\dots,\frac{R_{i,k}^MR_{i,k}^M}{R_{i,k}^{M,M}}\right). \tag{4.15}$$

For this compressed data vector $\mathbf{C}_k$ we obtain with Lemma 4.12 and (4.13) that

$$\mathrm{E}[\mathbf{C}_k|\,\mathcal{D}_k,\mathbf{F}] = \mathbf{F}_k \tag{4.16}$$

$$\begin{aligned}\mathrm{Cov}[\mathbf{C}_k,\mathbf{C}_k|\,\mathcal{D}_k,\mathbf{F}] &= \mathbf{W}_{\bullet,k}^{-1}\left(\sum_{i=0}^{I-k-1}\mathbf{W}_{i,k}\mathrm{Cov}[\mathbf{Y}_{i,k},\mathbf{Y}_{i,k}|\,\mathcal{D}_k,\mathbf{F}]\,\mathbf{W}_{i,k}\right)\mathbf{W}_{\bullet,k}^{-1} \\ &= \mathbf{W}_{\bullet,k}^{-1}\left(\sum_{i=0}^{I-k-1}\mathbf{W}_{i,k}\boldsymbol{\Sigma}_{i,k}(\mathbf{F})\mathbf{W}_{i,k}\right)\mathbf{W}_{\bullet,k}^{-1}.\end{aligned} \tag{4.17}$$

Theorem 4.13 (Credibility predictor of $\mathbf{F}_k$ at time I based on compressed data)
Under Model Assumptions 4.4 the credibility predictor for $\mathbf{F}_k$ based on the compressed data vector $\mathbf{C}_k$ is given by

$$\widehat{\mathbf{F}}_k^{I,Cred} = \mathbf{A}_k^I\mathbf{C}_k + (\mathbf{I}-\mathbf{A}_k^I)\boldsymbol{\mu}_k,$$

with the identity matrix $\mathbf{I}$, the structural parameter vector (prior mean)

$$\boldsymbol{\mu}_k := \mathrm{E}[\mathbf{F}_k|\,\mathcal{D}_k]$$

and the credibility weight

$$\mathbf{A}_k^I := \mathbf{T}_k(\mathbf{T}_k+\mathbf{U}_k^I)^{-1}, \tag{4.18}$$

where

$$\mathbf{U}_k^I := \mathbf{W}_{\bullet,k}^{-1}\left(\sum_{i=0}^{I-k-1} \mathbf{W}_{i,k}\mathrm{E}[\mathbf{\Sigma}_{i,k}(\mathbf{F})|\,\mathcal{D}_k]\,\mathbf{W}_{i,k}\right)\mathbf{W}_{\bullet,k}^{-1}$$

and the structural parameter matrix

$$\mathbf{T}_k := \mathrm{Cov}[\mathbf{F}_k,\ \mathbf{F}_k|\,\mathcal{D}_k]\,.$$

Proof: Conditionally on $\mathcal{D}_k$ the random variable $\mathbf{C}_k$ fulfills Model Assumptions 7.1 in Bühlmann-Gisler [14] and is therefore the credibility predictor for $\mathbf{F}_k$ at time I based on the compressed data vector $\mathbf{C}_k$. □

Remarks 4.14 (Credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$)

i) The vector $\mathbf{C}_k$ is the estimator for $\mathbf{f}_k$ in the classical (non-Bayesian) LSRM framework in Section 4.1. This becomes clear by comparing $\mathbf{C}_k$ in (4.14) *and associated weights* (4.15) *to $\widehat{\mathbf{f}}_k$ in* (4.3) *with corresponding weights* (4.4) *componentwise.*

ii) The credibility predictor is a credibility weighted average of the prior mean $\boldsymbol{\mu}_k$ and the compressed data vector $\mathbf{C}_k$ consisting of credibility weighted observations.

iii) Note, that we only need estimators for the first two moments of $\mathbf{F}_k$, given $\mathcal{D}_k$, and not the full joint distribution. This is the great advantage of credibility theory.

iv) The credibility predictor in Theorem 4.13 is based on the three structural parameters: $\boldsymbol{\mu}_k$, $\mathbf{T}_k$ and $\mathbf{U}_k^I$. These parameters can either be estimated by including prior expert knowledge or by portfolio data (see Section 7.3.5 in Bühlmann-Gisler *[14]). Note that the estimation of $\mathbf{U}_k^I$ can be reduced to an estimator $\widehat{\sigma}_k^{m_1,m_2}$ of $\mathrm{E}\big[\sigma_k^{m_1,m_2}(\mathbf{F})\big|\,\mathcal{D}_k\big]$, because all remaining terms in $\mathbf{U}_k^I$ are $\mathcal{D}_k^I$-measurable and can be observed. An unbiased estimator for $\mathrm{E}\big[\sigma_k^{m_1,m_2}(\mathbf{F})\big|\,\mathcal{D}_k\big]$ is given by* (4.5)*. The (conditional) unbiasedness of the estimator follows by*

$$\mathrm{E}\big[\widehat{\sigma}_k^{m_1,m_2}\big|\,\mathcal{D}_k\big] = \mathrm{E}\big[\mathrm{E}\big[\widehat{\sigma}_k^{m_1,m_2}\big|\,\mathbf{F},\mathcal{D}_k\big]\big|\,\mathcal{D}_k\big] = \mathrm{E}\big[\sigma_k^{m_1,m_2}(\mathbf{F})\big|\,\mathcal{D}_k\big]\,.$$

Now we will prove that the data compression $\mathbf{C}_k$ is an admissible compression, i.e. that the credibility predictor $\widehat{\mathbf{F}}_k^{\mathrm{Cred}}$ for $\mathbf{F}_k$ based on $\mathbf{C}_k$ in Theorem 4.13 is also the credibility predictor for $\mathbf{F}_k$ based on all data $\mathbf{Y}_{0,k},\ldots,\mathbf{Y}_{I-k-1,k}$.

Theorem 4.15 (Credibility predictor of $\mathbf{F}_k$ at time I based on all data) *Under Model Assumptions 4.4 the credibility predictor for $\mathbf{F}_k$ based on $\mathbf{Y}_{0,k},\ldots,\mathbf{Y}_{I-k-1,k}$ is given by*

$$\widehat{\mathbf{F}}_k^{I,Cred} = \mathbf{A}_k^I\mathbf{C}_k + (\mathbf{I}-\mathbf{A}_k^I)\boldsymbol{\mu}_k,$$

where $\mathbf{A}_k^I$, $\mathbf{C}_k$ and $\boldsymbol{\mu}_k$ are defined as in Theorem 4.13.

Proof: Choose $i \in \{0, \ldots, I-k-1\}$. All following calculations are done conditionally given $\mathcal{D}_k$. Since $\widehat{\mathbf{F}}_k^{I,Cred} = \mathbf{A}_k^I \mathbf{C}_k + (\mathbf{I} - \mathbf{A}_k^I)\boldsymbol{\mu}_k$, it follows straightforward that $\widehat{\mathbf{F}}_k^{I,Cred} \in \mathbb{L}^{ind}(\mathbf{Y}_{0,k}, \ldots, \mathbf{Y}_{I-k-1,k})$. Moreover, using (4.16) we get

$$\mathrm{E}[\mathbf{C}_k] = \mathrm{E}[\mathrm{E}[\mathbf{C}_k | \mathbf{F}]] = \mathrm{E}[\mathbf{F}_k] = \boldsymbol{\mu}_k.$$

This implies

$$\mathrm{E}\Big[\widehat{\mathbf{F}}_k^{I,Cred}\Big] = \mathrm{E}\big[\mathbf{A}_k^I \mathbf{C}_k + (\mathbf{I} - \mathbf{A}_k^I)\boldsymbol{\mu}_k\big] = \boldsymbol{\mu}_k = \mathrm{E}[\mathbf{F}_k],$$

i.e. Condition 2. in Lemma 4.9 is fulfilled. It remains to show that

$$\mathbf{F}_k - \widehat{\mathbf{F}}_k^{I,Cred} \quad \perp \quad \mathbf{Y}_{i,k}.$$

It holds

$$\begin{aligned} \mathrm{E}\Big[\Big(\mathbf{F}_k - \widehat{\mathbf{F}}_k^{I,Cred}\Big)\,\mathbf{Y}_{i,k}'\Big] &= \mathrm{E}\big[\big(\mathbf{F}_k - \mathbf{A}_k^I \mathbf{C}_k - \big(\mathbf{I} - \mathbf{A}_k^I\big)\,\boldsymbol{\mu}_k\big)\,\mathbf{Y}_{i,k}'\big] \\ &= \mathbf{A}_k^I \mathrm{E}\big[(\mathbf{F}_k - \mathbf{C}_k)\,\mathbf{Y}_{i,k}'\big] + \big(\mathbf{I} - \mathbf{A}_k^I\big)\,\mathrm{E}\big[(\mathbf{F}_k - \boldsymbol{\mu}_k)\,\mathbf{Y}_{i,k}'\big]. \end{aligned} \tag{4.19}$$

We have (see Theorem A.3 of Appendix A in Bühlmann-Gisler [14] for a proof) that $\mathbf{C}_k$ is the orthogonal projection of $\mathbf{F}_k$ on the affine-linear subspace

$$\begin{aligned} &\mathbb{L}_e^{ind}(\mathbf{Y}_{0,k}, \ldots, \mathbf{Y}_{I-k-1,k}) \\ &:= \left\{ \widehat{\mathbf{F}}_k \middle| \widehat{\mathbf{F}}_k = \sum_{i=0}^{I-k-1} \mathbf{A}_{i,k} \mathbf{Y}_{i,k},\ \mathbf{A}_{i,k} \in \mathbb{R}^{(M+1)\times(M+1)}, \mathrm{E}\Big[\widehat{\mathbf{F}}_k \Big| \mathbf{F}_k\Big] = \mathbf{F}_k \right\}, \end{aligned}$$

of all conditionally, given $\mathbf{F}_k$, unbiased estimators for $\mathbf{F}_k$, which are linear in $\mathbf{Y}_{i,k},\ i = 0, \ldots, I-k-1$. This implies

$$\mathrm{E}\big[(\mathbf{F}_k - \mathbf{C}_k)(\mathbf{Y}_{i,k} - \mathbf{C}_k)'\big] = \mathbf{0} \in \mathbb{R}^{(M+1)\times(M+1)}$$

and together with (4.17) we obtain

$$\begin{aligned} \mathrm{E}\big[(\mathbf{F}_k - \mathbf{C}_k)\mathbf{Y}_{i,k}'\big] &= \mathrm{E}\big[(\mathbf{F}_k - \mathbf{C}_k)\mathbf{C}_k'\big] \\ &= \mathrm{E}\big[\mathrm{E}\big[(\mathbf{F}_k - \mathbf{C}_k)\mathbf{C}_k' \big| \mathbf{F}\big]\big] \\ &= \mathrm{E}\big[\mathrm{E}\big[(\mathbf{F}_k - \mathbf{C}_k)(\mathbf{C}_k - \mathbf{F}_k)' \big| \mathbf{F}\big]\big] \\ &= -\mathrm{E}\left[\mathbf{W}_{\bullet,k}^{-1}\left(\sum_{i=0}^{I-k-1} \mathbf{W}_{i,k}\boldsymbol{\Sigma}_{i,k}(\mathbf{F}_k)\mathbf{W}_{i,k}\right)\mathbf{W}_{\bullet,k}^{-1}\right] \\ &= -\mathbf{U}_k^I. \end{aligned}$$

Using the same arguments, we get

$$\mathrm{E}\big[(\mathbf{F}_k - \boldsymbol{\mu}_k)\mathbf{Y}_{i,k}'\big] = \mathbf{T}_k.$$

By putting this into (4.19) and using the definition of $\mathbf{A}_k^I$ we get

$$\begin{aligned}\mathbf{A}_k^I \mathrm{E}\left[(\mathbf{F}_k - \mathbf{C}_k)\mathbf{Y}'_{i,k}\right] + (\mathbf{I} - \mathbf{A}_k^I)\mathrm{E}\left[(\mathbf{F}_k - \boldsymbol{\mu}_k)\mathbf{Y}'_{i,k}\right] &= -\mathbf{A}_k^I \mathbf{U}_k^I + (\mathbf{I} - \mathbf{A}_k^I)\mathbf{T}_k \\ &= -\mathbf{A}_k^I(\mathbf{T}_k + \mathbf{U}_k^I) + \mathbf{T}_k \\ &= -\mathbf{T}_k(\mathbf{T}_k + \mathbf{U}_k^I)^{-1}(\mathbf{T}_k + \mathbf{U}_k^I) + \mathbf{T}_k \\ &= \mathbf{T}_k - \mathbf{T}_k = \mathbf{0},\end{aligned}$$

i.e. $\mathrm{E}\left[\left(\mathbf{F}_k - \widehat{\mathbf{F}}_k^{I,Cred}\right)\mathbf{Y}'_{i,k}\right] = \mathbf{0}$. This completes the proof. □

Theorem 4.15 provides an explicit formula for the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ and allows for a direct calculation of the credibility based predictor $\widehat{S}_{i,k+1}^{m|I,(Cred)}$ stated in Definition 4.10. In the following subsection our goal is to quantify the MSEP of this credibility based predictor. We will base several definitions in the following subsections on $\widehat{S}^{n|n_1} \in \mathbb{L}^n$ with $n_1 \in \{I, I+1\}$ defined by

$$\left(\widehat{\mathbf{S}}^{n|n_1}\right)_{i,k}^m := \begin{cases} \widehat{S}_{i,k}^{m|I} & \text{for} \quad n_1 < i+k \le n \\ S_{i,k}^m & \text{for} \quad 0 \le i+k \le n_1 \end{cases}, \tag{4.20}$$

for an arbitrary $\sigma(\mathcal{D}^{n_1}, \mathbf{F})$-measurable predictor $\widehat{S}_{i,k}^{m|I}$ for $S_{i,k}^{m|I}$. That means that the matrix $\mathbf{S}^I$ containing all observations up to time I is extended by additional predicted diagonals up to accounting year n.

4.2.4 Mean Squared Error of Prediction

One is often interested in weighted sums of incremental claim information of the form $\sum_{k=I-i}^{J-1} \alpha_i^m S_{i,k+1}^m$ for fixed $m \in \mathcal{M}$, $i \in \{I-J+1, \dots, I\}$ and $\alpha_i^m \in \mathbb{R}$. The credibility based predictor for these sums at time I is given by $\sum_{k=I-i}^{J-1} \alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}$. For the prediction uncertainty, we consider the (conditional) MSEP given by

$$\begin{aligned}&\mathrm{msep}_{\sum\limits_{m\in\mathcal{M}}\sum\limits_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \Big| \mathcal{D}^I}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right] \\ &\qquad = \mathrm{E}\left[\left(\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m\left(S_{i,k+1}^m - \widehat{S}_{i,k+1}^{m,I,(Cred)}\right)\right)^2 \middle| \mathcal{D}^I\right].\end{aligned}$$

We decompose the (conditional) MSEP into

$$
\begin{aligned}
&\operatorname{msep}_{\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \mid \mathcal{D}^I}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right]\\
&= \underbrace{\mathrm{E}\left[\operatorname{Var}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \,\middle|\, \mathbf{F},\mathcal{D}^I\right]\middle|\,\mathcal{D}^I\right]}_{\text{“average” process variance}} \qquad (4.21)\\
&+ \underbrace{\mathrm{E}\left[\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \,\middle|\, \mathbf{F},\mathcal{D}^I\right] - \sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m,I,(Cred)}\right)^2\middle|\,\mathcal{D}^I\right]}_{\text{“average” estimation error}}. \qquad (4.22)
\end{aligned}
$$

At first we analyze the “average” estimation error (4.22).

Estimation Error for Single Accident Years

For the “average” estimation error (4.22) of a single accident year $i \in \{I-J+1,\dots,I\}$ we find (using $\simeq$ to indicate that the equation is approximately fulfilled)

$$
\begin{aligned}
&\mathrm{E}\left[\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \,\middle|\, \mathbf{F},\mathcal{D}^I\right] - \sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right)^2\middle|\,\mathcal{D}^I\right]\\
&\simeq \mathrm{E}\left[\left(\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,\mathbf{F}} - \sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,Bayes}\right)^2\middle|\,\mathcal{D}^I\right]\\
&= \operatorname{Var}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right] \qquad (4.23)
\end{aligned}
$$

The conditional variance (4.23) can further be decomposed into

$$
\begin{aligned}
&\operatorname{Var}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right] \qquad (4.24)\\
&= \mathrm{E}\left[\left(\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,\mathbf{F}}\right)^2\middle|\,\mathcal{D}^I\right] - \mathrm{E}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right]^2\\
&= \sum_{m_1,m_2\in\mathcal{M}}\sum_{k_1,k_2=I-i}^{J-1}\alpha_i^{m_1}\alpha_i^{m_2}\mathrm{E}\left[\widehat{S}_{i,k_1+1}^{m_1|I,\mathbf{F}}\widehat{S}_{i,k_2+1}^{m_2|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right] \qquad (4.25)\\
&\quad - \sum_{m_1,m_2\in\mathcal{M}}\sum_{k_1,k_2=I-i}^{J-1}\alpha_i^{m_1}\alpha_i^{m_2}\mathrm{E}\left[\widehat{S}_{i,k_1+1}^{m_1|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right]\mathrm{E}\left[\widehat{S}_{i,k_2+1}^{m_2|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right].
\end{aligned}
$$

In order to find an estimator for (4.25), we have to make some computations for expectations of products of $\widehat{S}_{i,k+1}^{m|I,\mathbf{F}}$. Let $i_2+k_2 \geq I$ and $k_2 > k_1$. Then we get with Model Assumption 4.4 c)

$$
\mathrm{E}\left[\widehat{S}_{i_1,k_1+1}^{m_1|I,\mathbf{F}}\widehat{S}_{i_2,k_2+1}^{m_2|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right] = \mathrm{E}\left[\widehat{S}_{i_1,k_1+1}^{m_1|I,\mathbf{F}}\mathbf{P}_{i_2,k_2}^{m_2|i_2+k_2,Bayes}\widehat{\mathbf{S}}^{i_2+k_2|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right],
$$

where (cf. (4.20))

$$\widehat{\mathbf{S}}^{n|I,\mathbf{F}} := \widehat{\mathbf{S}}^{n|I}\Big|_{\widehat{S}^{m|I}_{i,k}=\widehat{S}^{m|I,\mathbf{F}}_{i,k}}.$$

For $k_1 = k_2 =: k$ and $i_1 + k,\ i_2 + k \geq I$ we calculate again with Model Assumption 4.4 c)

$$\begin{aligned}\mathrm{E}\Big[\widehat{S}^{m_1|I,\mathbf{F}}_{i_1,k+1}\,\widehat{S}^{m_2|I,\mathbf{F}}_{i_2,k+1}\Big|\,\mathcal{D}^I\Big] &= \mathrm{E}\big[F^{m_1}_k F^{m_2}_k\big|\,\mathcal{D}^I\big]\,\mathrm{E}\Big[\mathbf{\Gamma}^{m_1}_{i_1,k}\widehat{\mathbf{S}}^{i_1+k|I,\mathbf{F}}\mathbf{\Gamma}^{m_2}_{i_2,k}\widehat{\mathbf{S}}^{i_2+k|I,\mathbf{F}}\Big|\,\mathcal{D}^I\Big]\\ &= \big(\mathrm{E}\big[F^{m_1}_k\big|\,\mathcal{D}^I\big]\,\mathrm{E}\big[F^{m_2}_k\big|\,\mathcal{D}^I\big] + \varrho^{*\,m_1,m_2}_k\big)\mathrm{E}\Big[\mathbf{\Gamma}^{m_1}_{i_1,k}\widehat{\mathbf{S}}^{i_1+k|I,\mathbf{F}}\mathbf{\Gamma}^{m_2}_{i_2,k}\widehat{\mathbf{S}}^{i_2+k|I,\mathbf{F}}\Big|\,\mathcal{D}^I\Big],\end{aligned}$$

with

$$\varrho^{*\,m_1,m_2}_k := \big(\mathrm{Cov}\big[\mathbf{F}_k,\mathbf{F}_k\big|\,\mathcal{D}^I\big]\big)_{m_1,m_2}.$$

To simplify notation of the terms above, we define for $n \in \{I,\dots,I+J-1\}$ and $k \in \{0,\dots,J-1\}$ the linear operators (the symbol $\otimes$ denotes the tensor product, see Lang [37])

$$\mathbf{H}^n_k(\boldsymbol{\tau})\colon \mathbb{L}^n_k \otimes \mathbb{L}^n_k \longrightarrow \mathbb{L}^n_{k+1} \otimes \mathbb{L}^n_{k+1},\ \mathbf{xy} \longmapsto \mathbf{H}^n_k(\boldsymbol{\tau})\,\mathbf{xy}$$

with

$$\begin{aligned}&\Big(\mathbf{H}^n_k(\boldsymbol{\tau})\,\mathbf{xy}\Big)^{m_1,m_2}_{i_1,k_1,i_2,k_2}\\ &\quad:= \begin{cases}\mathbf{P}^{m_1|n\vee(i_1+k)}_{i_1,k_1-1}\mathbf{x}\,\mathbf{P}^{m_2|n\vee(i_2+k)}_{i_2,k_2-1}\mathbf{y} & \text{for } i_1\wedge i_2 \leq n-k-1 \text{ or } k_1\wedge k_2 \leq k\\ \Big(f^{m_1}_k f^{m_2}_k + \tau^{m_1,m_2}_{i_1,i_2,k}\Big)\,\mathrm{E}\Big[\mathbf{\Gamma}^{m_1}_{i_1,k}\mathbf{x}\mathbf{\Gamma}^{m_2}_{i_2,k}\mathbf{y}\Big|\,\mathcal{D}^I\Big] & \text{otherwise}\end{cases},\end{aligned}$$

where $\boldsymbol{\tau}$ is a $(M+1)\times(M+1)\times I\times I\times J$ tensor and $\tau^{m_1,m_2}_{i_1,i_2,k}$ are the entries of $\boldsymbol{\tau}$. Concatenations of these operators will be denoted by

$$\begin{aligned}\mathbf{H}^n_{k_2\leftarrow k_1}(\boldsymbol{\tau}) &:= \begin{cases}\mathbf{H}^n_{k_2}(\boldsymbol{\tau})\mathbf{H}^n_{k_2-1}(\boldsymbol{\tau})\cdots\mathbf{H}^n_{k_1}(\boldsymbol{\tau}) & \text{for } k_2 \geq k_1\\ \Pi_{\mathbb{L}^n_{k_2+1}\otimes\mathbb{L}^n_{k_2+1}} & \text{otherwise}\end{cases},\\ \mathbf{H}^{m_1,m_2|n}_{i_1,k_1,i_2,k_2}(\boldsymbol{\tau})\mathbf{xy} &:= \Big(\mathbf{H}^n_{(k_1\vee k_2)\leftarrow 0}(\boldsymbol{\tau})\mathbf{xy}\Big)^{m_1,m_2}_{i_1,k_1+1,i_2,k_2+1},\end{aligned}\tag{4.26}$$

where $\Pi_{\mathbb{L}^n_{k_2+1}\otimes\mathbb{L}^n_{k_2+1}}$ denotes the projection onto $\mathbb{L}^n_{k_2+1}\otimes\mathbb{L}^n_{k_2+1}$.
By replacing in (4.26) $\mathbf{f}$ by $\mathrm{E}\big[\mathbf{F}\big|\,\mathcal{D}^I\big]$ and $\tau^{m_1,m_2}_{i_1,i_2,k} := \varrho^{*\,m_1,m_2}_k$ we obtain

$$\mathbf{H}^{m_1,m_2|n,Bayes}_{i_1,k_1,i_2,k_2}(\boldsymbol{\varrho}^*) := \mathbf{H}^{m_1,m_2|n}_{i_1,k_1,i_2,k_2}(\boldsymbol{\tau})\Big|_{\mathbf{f}=\mathrm{E}[\mathbf{F}|\,\mathcal{D}^I]\ \wedge\ \tau^{m_1,m_2}_{i_1,i_2,k}=\varrho^{*\,m_1,m_2}_k}.\tag{4.27}$$

With $n = I$ in (4.27) we get for each summand in (4.25)

$$\mathrm{E}\Big[\widehat{S}^{m_1|I,\mathbf{F}}_{i_1,k_1+1}\widehat{S}^{m_2|I,\mathbf{F}}_{i_2,k_2+1}\Big|\,\mathcal{D}^I\Big] = \mathbf{H}^{m_1,m_2|I,Bayes}_{i_1,k_1,i_2,k_2}(\boldsymbol{\varrho}^*)\mathbf{S}^I\mathbf{S}^I.\tag{4.28}$$

However, $\mathbf{H}^{m_1,m_2|I,Bayes}_{i_1,k_1,i_2,k_2}(\boldsymbol{\varrho}^*)$ in (4.28) still depends on $\mathrm{Cov}\big[\mathbf{F}_k,\mathbf{F}_k\big|\,\mathcal{D}^I\big]$ and $\mathrm{E}\big[\mathbf{F}\big|\,\mathcal{D}^I\big]$, see (4.27). In a first step we estimate $\mathrm{E}\big[\mathbf{F}\big|\,\mathcal{D}^I\big]$ by $\widehat{\mathbf{F}}^{I,Cred}$, i.e.

$$\widehat{\mathrm{E}}\big[\mathbf{F}\big|\,\mathcal{D}^I\big] := \widehat{\mathbf{F}}^{I,Cred}.\tag{4.29}$$

Using the approximation

$$\begin{aligned}\mathrm{Cov}\big[\mathbf{F}_k,\mathbf{F}_k\big|\,\mathcal{D}^I\big] &= \mathrm{E}\Big[\big(\mathbf{F}_k-\mathrm{E}\big[\mathbf{F}_k\big|\,\mathcal{D}^I\big]\big)\big(\mathbf{F}_k-\mathrm{E}\big[\mathbf{F}_k\big|\,\mathcal{D}^I\big]\big)'\Big|\,\mathcal{D}^I\Big]\\ &\simeq \mathrm{E}\bigg[\Big(\mathbf{F}_k-\widehat{\mathbf{F}}_k^{I,Cred}\Big)\Big(\mathbf{F}_k-\widehat{\mathbf{F}}_k^{I,Cred}\Big)'\bigg|\,\mathcal{D}^I\bigg] \\ &\simeq \mathbf{A}_k^I\mathbf{U}_k^I,\end{aligned} \tag{4.30}$$

where in the last approximation we used the loss matrix of the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ (see Theorem 7.5 in Bühlmann-Gisler [14]), we find the following estimator for $\varrho^{*\,m_1,m_2}_{i_1,i_2,k}$

$$\widehat{\varrho}^{*\,m_1,m_2}_{i_1,i_2,k} := \big(\mathbf{A}_k^I\mathbf{U}_k^I\big)_{m_1,m_2}. \tag{4.31}$$

Putting the estimates (4.29) and (4.31) into (4.27) we obtain for $n\in\{I,\dots,I+J-1\}$

$$\widehat{\mathbf{H}}^{m_1,m_2|n,Cred}_{i_1,k_1,i_2,k_2}(\widehat{\boldsymbol{\varrho}}^*) := \mathbf{H}^{m_1,m_2|n}_{i_1,k_1,i_2,k_2}(\boldsymbol{\tau})\Big|_{\mathbf{f}=\widehat{\mathbf{F}}^{I,Cred}\ \wedge\ \boldsymbol{\tau}=\widehat{\boldsymbol{\varrho}}^*}.$$

For $n=I$ this leads to the following estimator of the estimation error (4.22).

Estimator 4.16 (Estimation error for single accident years) *Under Model Assumptions 4.4 at time I an estimator for the estimation error* (4.22) *of accident year $i\in\{I-J+1,\dots,I\}$ is given by*

$$\begin{aligned}&\widehat{\mathrm{E}}\left[\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m\middle|\,\mathbf{F},\mathcal{D}^I\right]-\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m\widehat{S}_{i,k+1}^{m|I,(Cred)}\right)^2\middle|\,\mathcal{D}^I\right]\\ &\qquad := \sum_{m_1,m_2\in\mathcal{M}}\sum_{k_1,k_2=I-i}^{J-1}\alpha_i^{m_1}\alpha_i^{m_2}\left(\widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^*)-\widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\mathbf{0})\right)\mathbf{S}^I\mathbf{S}^I,\end{aligned}$$

where $\mathbf{0}$ *denotes the* 0*-tensor.*

Process Variance for Single Accident Years

For the process variance (4.21) we see that, conditionally given $\mathbf{F}$, Model Assumptions 4.1 for the classical LSRM are fulfilled and with Lemma 4.2 in Dahms [17] follows for $I\le i_1+k_1, i_2+k_2$

$$\begin{aligned}&\mathrm{Cov}\Big[S^{m_1}_{i_1,k_1+1},S^{m_2}_{i_2,k_2+1}\Big|\,\mathbf{F},\mathcal{D}^I\Big]\\ &=\sum_{n=I+1}^{(i_1+k_1)\wedge(i_2+k_2)+1}\sum_{l_1,l_2=0}^{M}\sum_{j=n-I}^{J}\sigma^{l_1,l_2}_{j-1}(\mathbf{F})\,\mathrm{E}\Big[\boldsymbol{\Gamma}^{l_1,l_2}_{n-j,j-1}\mathbf{S}^{n-1}\Big|\,\mathbf{F},\mathcal{D}^I\Big]\Big(\mathbf{P}^{m_1|n,\mathbf{F}}_{i_1,k_1}\Big)^{l_1}_{n-j,j}\Big(\mathbf{P}^{m_2|n,\mathbf{F}}_{i_2,k_2}\Big)^{l_2}_{n-j,j}.\end{aligned}$$

with the coupling exposure

$$\boldsymbol{\Gamma}^{m_1,m_2}_{i,k}\mathbf{S}^{i+k} := \sum_{l=1}^{M}\sum_{h=0}^{I}\sum_{j=0}^{(i+k-h)\wedge k}\gamma^{m_1,m_2,l}_{i,k,h,j}S^l_{h,j} = R^{m_1,m_2}_{i,k},$$

see (4.1). Thus, with Model Assumptions 4.4 c) we obtain for a fixed accident year $i \in \{I-J+1, \ldots, I\}$ for the process variance (4.21)

$$\mathrm{E}\left[\mathrm{Var}\left[\sum_{m=0}^{M}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \middle| \mathbf{F}, \mathcal{D}^I\right]\middle| \mathcal{D}^I\right] \tag{4.32}$$

$$= \sum_{m_1,m_2\in\mathcal{M}} \alpha_i^{m_1}\alpha_i^{m_2}\left(\sum_{k_1,k_2=I-i}^{J-1}\sum_{n=I+1}^{i+(k_1\wedge k_2)+1}\sum_{l_1,l_2=0}^{M}\sum_{j=n-I}^{J}\right. \tag{4.33}$$

$$\left.\mathrm{E}\left[\sigma_{j-1}^{l_1,l_2}(\mathbf{F})\middle| \mathcal{D}^I\right]\boldsymbol{\Gamma}_{n-j,j-1}^{l_1,l_2}\widehat{\mathbf{S}}^{n-1|I,Bayes}\mathrm{E}\left[\left(\mathbf{P}_{i,k_1}^{m_1|n,\mathbf{F}}\right)_{n-j,j}^{l_1}\left(\mathbf{P}_{i,k_2}^{m_2|n,\mathbf{F}}\right)_{n-j,j}^{l_2}\middle| \mathcal{D}^I\right]\right),$$

where

$$\widehat{\mathbf{S}}^{n|I,Bayes} := \widehat{\mathbf{S}}^{n|I}\Big|_{\widehat{S}_{i,k}^{m|I}=\widehat{S}_{i,k}^{m|I,Bayes}}.$$

Since in the derivation of (4.28) no special property of $\mathbf{S}^I$ is used, except that it is contained in $\mathbb{L}^I$, we get with the same arguments for $i_1+k_1,\ i_2+k_2 \geq I$

$$\begin{aligned}\mathrm{E}\left[\left(\mathbf{P}_{i_1,k_1}^{m_1|n,\mathbf{F}}\right)_{n-j,j}^{l_1}\left(\mathbf{P}_{i_2,k_2}^{m_2|n,\mathbf{F}}\right)_{n-j,j}^{l_2}\middle| \mathcal{D}^I\right] &= \mathrm{E}\left[\mathbf{P}_{i_1,k_1}^{m_1|n,\mathbf{F}}\mathbf{e}_{n-j,j}^{l_1}\mathbf{P}_{i_2,k_2}^{m_2|n,\mathbf{F}}\mathbf{e}_{n-j,j}^{l_2}\middle| \mathcal{D}^I\right]\\ &= \mathbf{H}_{i_1,k_1,i_2,k_2}^{m_1,m_2|n,Bayes}(\boldsymbol{\varrho}^*)\mathbf{e}_{n-j,j}^{l_1}\mathbf{e}_{n-j,j}^{l_2},\end{aligned} \tag{4.34}$$

where $\mathbf{e}_{n-j,j}^{l_1} \in \mathbb{L}^n$ with $\mathbf{e}_{n-j,j}^{l_1} = 1$ in the entry $(n-j, j, l_1)$, 0 otherwise. Using the approximations (estimates)

$$\begin{aligned}\widehat{\mathbf{S}}^{i+k|I,(Cred)} &\simeq \widehat{\mathbf{S}}^{i+k|I,Bayes}\\ \widehat{\sigma}_{j-1}^{l_1,l_2} &:= \widehat{\mathrm{E}}\left[\sigma_{j-1}^{l_1,l_2}(\mathbf{F})\middle| \mathcal{D}^I\right]\\ \widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|n,Cred}(\widehat{\boldsymbol{\varrho}}^*) &\simeq \mathbf{H}_{i,k_1,i,k_2}^{m_1,m_2|n,Bayes}(\boldsymbol{\varrho}^*),\end{aligned} \tag{4.35}$$

see Remark 4.14 for the estimate $\widehat{\sigma}_{j-1}^{l_1,l_2}$, and putting the estimates (4.35) into (4.33) and (4.34), respectively, we get the following estimator:

Estimator 4.17 (Process variance for single accident years) *Under Model Assumptions 4.4 at time I an estimator for the process variance* (4.21) *of accident year $i \in \{I-J+1, \ldots, I\}$ is given by*

$$\widehat{\mathrm{E}}\left[\mathrm{Var}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \middle| \mathbf{F}, \mathcal{D}^I\right]\middle| \mathcal{D}^I\right]$$

$$:= \sum_{m_1,m_2\in\mathcal{M}} \alpha_i^{m_1}\alpha_i^{m_2}\left(\sum_{k_1,k_2=I-i}^{J-1}\sum_{n=I+1}^{i+(k_1\wedge k_2)+1}\sum_{l_1,l_2=0}^{M}\sum_{j=n-I}^{J}\right.$$

$$\left.\widehat{\sigma}_{j-1}^{l_1,l_2}\boldsymbol{\Gamma}_{n-j,j-1}^{l_1,l_2}\widehat{\mathbf{S}}^{n-1|I,(Cred)}\widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|n,Cred}(\widehat{\boldsymbol{\varrho}}^*)\mathbf{e}_{n-j,j}^{l_1}\mathbf{e}_{n-j,j}^{l_2}\right).$$

Mean Squared Error of Prediction for Single Accident Years

Combining the Estimators 4.16 and 4.17 we get an estimator for the (conditional) MSEP for single accident years.

Estimator 4.18 (MSEP for single accident years) *Under Model Assumptions 4.4 at time I an estimator for the (conditional) MSEP of accident year $i \in \{I-J+1, \dots, I\}$ is given by*

$$\widehat{\text{msep}}_{\sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m S_{i,k+1}^m \big| \mathcal{D}^I} \left[\sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)} \right]$$
$$:= \sum_{m_1, m_2 \in \mathcal{M}} \alpha_i^{m_1} \alpha_i^{m_2} \left(\sum_{k_1,k_2=I-i}^{J-1} \left[\left(\widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}^*) - \widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|I,Cred}(\mathbf{0}) \right) \mathbf{S}^I \mathbf{S}^I \right.\right.$$
$$\left.\left. + \sum_{n=I+1}^{i+(k_1 \wedge k_2)+1} \sum_{l_1,l_2=0}^{M} \sum_{j=n-I}^{J} \widehat{\sigma}_{j-1}^{l_1,l_2} \boldsymbol{\Gamma}_{n-j,j-1}^{l_1,l_2} \widehat{\mathbf{S}}^{n-1|I,(Cred)} \widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|n,Cred}(\widehat{\boldsymbol{\varrho}}^*) \mathbf{e}_{n-j,j}^{l_1} \mathbf{e}_{n-j,j}^{l_2} \right] \right).$$

Mean Squared Error of Prediction for Aggregated Accident Years

Now we take a closer look at the prediction uncertainty of sums of credibility based predictors $\widehat{S}_{k+1}^{m|I,(Cred)}$ for different accident years. Since these predictors depend on data of all accident years they are not independent. Again, we decompose the (conditional) MSEP into

$$\text{msep}_{\sum_{i=I-J+1}^{I} \sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m S_{i,k+1}^m \big| \mathcal{D}^I} \left[\sum_{i=I-J+1}^{I} \sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)} \right] \tag{4.36}$$
$$= \text{E}\left[\text{Var}\left[\sum_{i=I-J+1}^{I} \sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m S_{i,k+1}^m \middle| \mathbf{F}, \mathcal{D}^I \right] \middle| \mathcal{D}^I \right] \tag{4.37}$$
$$+ \text{E}\left[\left(\text{E}\left[\sum_{i=I-J+1}^{I} \sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m S_{i,k+1}^m \middle| \mathbf{F}, \mathcal{D}^I \right] - \sum_{i=I-J+1}^{I} \sum_{m \in \mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)} \right)^2 \middle| \mathcal{D}^I \right], \tag{4.38}$$

where (4.37) corresponds to the process variance and (4.38) to the estimation error, respectively. Using the same techniques as in the previous section leads to the following estimator:

Estimator 4.19 (MSEP for aggregated accident years) *Under Model Assumptions 4.4 at time I an estimator for the (conditional) MSEP for aggregated accident years* (4.36) *is given by*

$$\widehat{\text{msep}}_{\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \Big| \mathcal{D}^I}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right]$$

$$:= \sum_{i_1,i_2=I-J+1}^{I}\sum_{m_1,m_2\in\mathcal{M}}\alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2}\left(\sum_{k_1=I-i_1}^{J-1}\sum_{k_2=I-i_2}^{J-1}\left[\left(\widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}^*) - \widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\mathbf{0})\right)\mathbf{S}^I\mathbf{S}^I\right.\right.$$

$$\left.\left. + \sum_{n=I+1}^{(i_1+k_1)\wedge(i_2+k_2)+1}\sum_{l_1,l_2=0}^{M}\sum_{j=n-I}^{J}\widehat{\sigma}_{j-1}^{l_1,l_2}\boldsymbol{\Gamma}_{n-j,j-1}^{l_1,l_2}\widehat{\mathbf{S}}^{n-1|I,(Cred)}\widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|n,Cred}(\widehat{\boldsymbol{\varrho}}^*)\mathbf{e}_{n-j,j}^{l_1}\mathbf{e}_{n-j,j}^{l_2}\right]\right).$$

4.2.5 Special Case: Mean Squared Error of Prediction for the Bayes CL Method

We saw in Subsection 4.2.1 that the Bayes CL method in GISLER–WÜTHRICH [27] belongs to the class of Bayesian LSRMs. In the following we show that if the Bayesian LSRM is the Bayes CL method the LSRM estimate for the (conditional) MSEP given by Estimator 4.18 coincides with the estimate in the Bayes CL method given in Theorem 4.4 in GISLER–WÜTHRICH [27]. Again we use the identity

$$\mathbf{G} := (G_0,\dots,G_{J-1}) = \left(F_0^0+1,\dots,F_{J-1}^0+1\right) = \mathbf{F}+1.$$

We first study the average process variance (4.21) for single accident years i given by formula (4.33).

Process Variance

For the Bayes CL method looked at as a LSRM, see Subsection 4.2.1, we first analyze the inner part of the average process variance (4.32)

$$\text{Var}\left[\sum_{m=0}^{M}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \,\middle|\, \mathbf{F},\mathcal{D}^I\right] \tag{4.39}$$

$$= \sum_{k_1,k_2=I-i}^{J-1}\sum_{n=I+1}^{i+(k_1\wedge k_2)+1}\sum_{j=n-I}^{J}\sigma_{j-1}^{0,0}(\mathbf{F})\boldsymbol{\Gamma}_{n-j,j-1}^{0,0}\widehat{\mathbf{S}}^{n-1|I,\mathbf{F}}\left(\mathbf{P}_{i,k_1}^{0|n,\mathbf{F}}\right)_{n-j,j}^{0}\left(\mathbf{P}_{i,k_2}^{0|n,\mathbf{F}}\right)_{n-j,j}^{0}. \tag{4.40}$$

A short calculation yields (the empty product is set to 0)

$$\left(\mathbf{P}_{i,k_1}^{0|n,\mathbf{F}}\right)_{n-j,j}^{0} = \begin{cases} 0 & \text{for } i \neq n-j \\ 1 & \text{for } i = n-j \wedge k_1 = j-1 \\ \prod_{l=n-i}^{k_1} G_l - \prod_{l=n-i}^{k_1-1} G_l & \text{for } i = n-j \wedge k_1 \geq j \end{cases} \tag{4.41}$$

and

$$\mathbf{\Gamma}^{0,0}_{n-j,j-1}\widehat{\mathbf{S}}^{n-1|I,\mathbf{F}} = C_{n-j,I-n+j} \prod_{l=I-n+j}^{j-2} G_l. \tag{4.42}$$

With (4.41) and (4.42) we obtain for (4.40)

$$\begin{aligned}
&\sum_{k_1,k_2=I-i}^{J-1} \sum_{n=I+1}^{i+(k_1\wedge k_2)+1} \sigma^{0,0}_{n-i-1}(\mathbf{F})\mathbf{\Gamma}^{0,0}_{i,n-i-1}\widehat{\mathbf{S}}^{n-1|I,\mathbf{F}} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_1}\right)^0_{i,n-i} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_2}\right)^0_{i,n-i} &(4.43)\\
&= C_{i,I-i} \sum_{n=I+1}^{J+i} \sigma^2_{n-i-1}(G_{n-i-1}) \prod_{l=I-i}^{n-i-2} G_l \sum_{k_1=n-i-1}^{J-1} \sum_{k_2=n-i-1}^{J-1} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_1}\right)^0_{i,n-i} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_2}\right)^0_{i,n-i}\\
&= C_{i,I-i} \sum_{n=I+1}^{J+i} \sigma^2_{n-i-1}(G_{n-i-1}) \prod_{l=I-i}^{n-i-2} G_l \sum_{k_1=n-i-1}^{J-1} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_1}\right)^0_{i,n-i} \sum_{k_2=n-i-1}^{J-1} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_2}\right)^0_{i,n-i}\\
&= C_{i,I-i} \sum_{n=I+1}^{J+i} \sigma^2_{n-i-1}(G_{n-i-1}) \prod_{l=I-i}^{n-i-2} G_l \left(\sum_{k_1=n-i-1}^{J-1} \left(\mathbf{P}^{0|n,\mathbf{F}}_{i,k_1}\right)^0_{i,n-i} \right)^2\\
&= C_{i,I-i} \sum_{n=I+1}^{J+i} \sigma^2_{n-i-1}(G_{n-i-1}) \prod_{l_1=I-i}^{n-i-2} G_{l_1} \prod_{l_2=n-i}^{J-1} (G_{l_2})^2\\
&= C_{i,I-i} \sum_{n=I-i}^{J-1} \sigma^2_{n}(G_{n}) \prod_{l_1=I-i}^{n-1} G_{l_1} \prod_{l_2=n+1}^{J-1} (G_{l_2})^2 . &(4.44)
\end{aligned}$$

Taking the conditional (on $\mathcal{D}^I$) expectations of (4.43) and (4.44) and using Model Assumptions 4.4 c) implies for the process variance

$$\begin{aligned}
&\mathrm{E}\left[\mathrm{Var}\left[\sum_{m=0}^{M} \sum_{k=I-i}^{J-1} \alpha_i^m S^m_{i,k+1} \middle| \mathbf{F}, \mathcal{D}^I \right] \middle| \mathcal{D}^I \right] &(4.45)\\
&= C_{i,I-i} \sum_{n=I-i}^{J-1} \mathrm{E}\big[\sigma^2_n(G_n)\big|\, \mathcal{D}^I\big] \prod_{l_1=I-i}^{n-1} \mathrm{E}\big[G_{l_1}\big|\, \mathcal{D}^I\big] \prod_{l_2=n+1}^{J-1} \mathrm{E}\Big[(G_{l_2})^2\Big|\, \mathcal{D}^I\Big] &(4.46)\\
&= C_{i,I-i} \sum_{n=I-i}^{J-1} \mathrm{E}\big[\sigma^2_n(F^0_n+1)\big|\, \mathcal{D}^I\big] \prod_{l_1=I-i}^{n-1} \mathrm{E}\big[F^0_{l_1}+1\big|\, \mathcal{D}^I\big] \prod_{l_2=n+1}^{J-1} \mathrm{E}\big[(F^0_{l_2}+1)^2\big|\, \mathcal{D}^I\big] . &(4.47)
\end{aligned}$$

Formula (4.46) for the process variance in the LSRM is identical with Formula (4.13) in Gisler–Wüthrich [27] for the process variance in the Bayes CL method. Hence, we have to check that in the Bayesian LSRM and in the Bayes CL method each component of (4.46) is estimated in the same way. In the Bayesian LSRM the following estimates are used, see (4.29)–(4.31),

$$\widehat{\mathrm{E}}\big[F_k^0+1\big|\,\mathcal{D}^I\big] := \widehat{F}_k^{0|I,Cred}+1 \tag{4.48}$$

$$\begin{aligned}\mathrm{E}\Big[\big(F_k^0+1\big)^2\Big|\,\mathcal{D}^I\Big] &= \mathrm{E}\Big[\big(F_k^0-\mathrm{E}\big[F_k^0\big|\,\mathcal{D}^I\big]\big)^2\Big|\,\mathcal{D}^I\Big]+\mathrm{E}\big[F_k^0+1\big|\,\mathcal{D}^I\big]^2\\ &\simeq \mathrm{E}\Big[\Big(F_k^0-F_k^{0|I,Cred}\Big)^2\Big|\,\mathcal{D}_k\Big]+\Big(F_k^{0|I,Cred}+1\Big)^2\\ &= \mathbf{A}_k^I\mathbf{U}_k^I+\Big(F_k^{0|I,Cred}+1\Big)^2\\ &= \alpha_k\frac{\mathrm{E}\Big[\sigma_k^{0,0}(F_k^0)\Big|\,\mathcal{D}_k\Big]}{\sum_{i=0}^{I-k-1}C_{i,k}}+\Big(F_k^{0|I,Cred}+1\Big)^2,\end{aligned} \tag{4.49}$$

with

$$\alpha_k := \frac{\sum_{i=0}^{I-k-1}C_{i,k}}{\sum_{i=0}^{I-k-1}C_{i,k}+\frac{\mathrm{E}\left[\sigma_k^{0,0}(F_k^0)\middle|\mathcal{D}_k\right]}{\mathrm{Var}\left[F_k^0\middle|\mathcal{D}_k\right]}}. \tag{4.50}$$

The approximations (4.48) and (4.49) for the LSRM coincide with the approximations (4.15) and (4.16) in Gisler–Wüthrich [27] for these quantities for the Bayes CL method. By putting the approximations (4.48) and (4.49) into (4.47) and using the fact that

$$F_k^{0|I,Cred}+1 = G_k^{I,Cred} \text{ and } \sigma_k^{0,0}(F_k^0)=\sigma_k^2(G_k)$$

we obtain the Bayesian LSRM estimate for the process variance (cf. Estimator 4.17) in the case of the Bayes CL method

$$\begin{aligned}&\widehat{\mathrm{E}}\left[\mathrm{Var}\left[\sum_{m=0}^{M}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m\middle|\,\mathbf{F},\mathcal{D}^I\right]\middle|\,\mathcal{D}^I\right]\\ &= C_{i,I-i}\sum_{n=I-i}^{J-1}\prod_{l_1=I-i}^{n-1}G_{l_1}^{I,Cred}\widehat{\mathrm{E}}\big[\sigma_n^2(G_n)\big|\,\mathcal{D}^I\big]\prod_{l_2=n+1}^{J-1}\left(\Big(G_{l_2}^{I,Cred}\Big)^2+\widehat{\alpha}_{l_2}\frac{\widehat{\mathrm{E}}\big[\sigma_{l_2}^2(G_{l_2})\big|\,\mathcal{D}_{l_2}\big]}{\sum_{i=0}^{I-l_2-1}C_{i,l_2}}\right),\end{aligned} \tag{4.51}$$

where $\widehat{\mathrm{E}}\big[\sigma_k^2(G_k)\big|\,\mathcal{D}^I\big]$ and $\widehat{\mathrm{Var}}[G_k|\,\mathcal{D}_k]$ are appropriate estimates for $\mathrm{E}\big[\sigma_k^2(G_k)\big|\,\mathcal{D}^I\big]$ and $\mathrm{E}\big[\sigma_k^2(G_k)\big|\,\mathcal{D}_k\big]$ as well as $\mathrm{Var}[G_k|\,\mathcal{D}_k]$ and

$$\widehat{\alpha}_k := \alpha_k\big|_{\mathrm{E}\left[\sigma_k^2(G_k)\middle|\mathcal{D}^I\right]=\widehat{\mathrm{E}}\left[\sigma_k^2(G_k)\middle|\mathcal{D}^I\right]\wedge\mathrm{Var}[G_k|\mathcal{D}_k]=\widehat{\mathrm{Var}}[G_k|\mathcal{D}_k]}.$$

If we use in the Bayesian LSRM as well as in the Bayes CL method the same estimators $\widehat{\mathrm{E}}\big[\sigma_k^2(G_k)\big|\,\mathcal{D}^I\big]$ and $\widehat{\mathrm{Var}}[G_k|\,\mathcal{D}_k]$ for the structural parameters $\mathrm{E}\big[\sigma_k^2(G_k)\big|\,\mathcal{D}^I\big]$ and $\mathrm{Var}[G_k|\,\mathcal{D}_k]$ (this is the case, since in both methods the Bühlmann-Straub estimators, see Bühlmann-Gisler [14], are proposed) the estimator (4.51) is the well-known estimator for the process variance in the Bayes CL method given in Theorem 4.4 in Gisler–Wüthrich [27]. For non-informative (vague) priors for the development factors F_k^0, i.e.

$$\mathrm{Var}[G_k|\,\mathcal{D}_k] = \mathrm{Var}\left[F_k^0\middle|\,\mathcal{D}_k\right] \longrightarrow \infty,$$

we have that $\alpha_k \longrightarrow 1$ and for the credibility predictor $F_k^{0|I,Cred}$ holds in this case

$$G_k^{I,Cred} = F_k^{0|I,Cred} + 1 \quad \underset{\alpha_k \to 1}{\longrightarrow} \quad \widehat{f}_k^0 + 1 = \widehat{g}_k^{I,CL},$$

i.e. for non-informative priors the credibility predictor $G_k^{I,Cred}$ coincides with the classical CL estimator $\widehat{g}_k^{I,CL}$, see (3.1). This then results in the estimate for the process variance with non-informative priors given by

$$\widehat{\mathrm{E}}\left[\mathrm{Var}\left[\sum_{m=0}^{M}\sum_{k=I-i}^{J-1} \alpha_i^m S_{i,k+1}^m \middle| \mathbf{F}, \mathcal{D}^I\right]\middle| \mathcal{D}^I\right] \tag{4.52}$$

$$= C_{i,I-i} \sum_{n=I-i}^{J-1} \prod_{l_1=I-i}^{n-1} \widehat{g}_{l_1}^{I,CL} \widehat{\mathrm{E}}\left[\sigma_n^2(G_n)\middle|\,\mathcal{D}^I\right] \prod_{l_2=n+1}^{J-1} \left(\left(\widehat{g}_{l_2}^{I,CL}\right)^2 + \frac{\widehat{\mathrm{E}}\left[\sigma_{l_2}^2(G_{l_2})\middle|\,\mathcal{D}_{l_2}\right]}{\sum_{i=0}^{I-l_2-1} C_{i,l_2}} \right). \tag{4.53}$$

The estimator (4.53) is exactly the estimator for the process variance for non-informative priors in the Bayes CL method, see Theorem 4.4 in Gisler–Wüthrich [27]. The estimator (4.53) is slightly higher then the estimator given in Mack [38] and Buchwalder et al. [10]. For a detailed comparison of the different estimators for the process variance in the (Bayes) CL method we refer to Gisler–Wüthrich [27] and Wüthrich–Merz [63].

Estimation Error

With the approximation

$$F_k^{m|I,Cred} \simeq F_k^{m|I,Bayes} \tag{4.54}$$

the estimation error (4.22) can be rewritten by

$$
\begin{aligned}
&\mathrm{E}\left[\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m\,\middle|\,\mathbf{F},\mathcal{D}^I\right]-\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right)^2\middle|\,\mathcal{D}^I\right]\\
&\simeq \sum_{k_1,k_2=I-i}^{J-1}\mathbf{H}_{i,k_1,i,k_2}^{0,0|I,Bayes}(\boldsymbol{\varrho}^*)\mathbf{S}^I\mathbf{S}^I-\sum_{k_1,k_2=I-i}^{J-1}\mathbf{H}_{i,k_1,i,k_2}^{0,0|I,Bayes}(\mathbf{0})\mathbf{S}^I\mathbf{S}^I\\
&=\sum_{k_1,k_2=I-i}^{J-1}\mathrm{E}\left[\widehat{S}_{i,k_1+1}^{0|I,\mathbf{F}}\widehat{S}_{i,k_2+1}^{0|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right]-\sum_{k_1,k_2=I-i}^{J-1}\mathrm{E}\left[\widehat{S}_{i,k_1+1}^{0|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right]\mathrm{E}\left[\widehat{S}_{i,k_2+1}^{0|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right]\\
&=\mathrm{E}\left[\left(\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{0|I,\mathbf{F}}\right)^2\middle|\,\mathcal{D}^I\right]-\left(\sum_{k=I-i}^{J-1}\mathrm{E}\left[\widehat{S}_{i,k+1}^{0|I,\mathbf{F}}\,\middle|\,\mathcal{D}^I\right]\right)^2\\
&=C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\mathrm{E}\left[(G_j)^2\,\middle|\,\mathcal{D}^I\right]-2\prod_{j=I-i}^{J-1}\mathrm{E}\left[G_j\,\middle|\,\mathcal{D}^I\right]+1\right)\\
&\qquad-C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\mathrm{E}\left[G_j\,\middle|\,\mathcal{D}^I\right]^2-2\prod_{j=I-i}^{J-1}\mathrm{E}\left[G_j\,\middle|\,\mathcal{D}^I\right]+1\right)\\
&=C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\mathrm{E}\left[(G_j)^2\,\middle|\,\mathcal{D}^I\right]-\prod_{j=I-i}^{J-1}\mathrm{E}\left[G_j\,\middle|\,\mathcal{D}^I\right]^2\right)\\
&=C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\left(\left(F_j^{0|I,Cred}+1\right)^2+\alpha_j\frac{\mathrm{E}\left[\sigma_j^{0,0}(F_j^0)\,\middle|\,\mathcal{D}_j\right]}{\sum_{i=0}^{I-j-1}C_{i,j}}\right)-\prod_{j=I-i}^{J-1}\left(F_j^{0|I,Cred}+1\right)^2\right)\\
&=C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\left(\left(G_j^{I,Cred}\right)^2+\alpha_j\frac{\mathrm{E}\left[\sigma_j^2(G_j)\,\middle|\,\mathcal{D}_j\right]}{\sum_{i=0}^{I-j-1}C_{i,j}}\right)-\prod_{j=I-i}^{J-1}\left(G_j^{I,Cred}\right)^2\right) &(4.55)\\
&\simeq C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\left(\left(G_j^{I,Cred}\right)^2+\widehat{\alpha}_j\frac{\widehat{\mathrm{E}}\left[\sigma_j^2(G_j)\,\middle|\,\mathcal{D}_j\right]}{\sum_{i=0}^{I-j-1}C_{i,j}}\right)-\prod_{j=I-i}^{J-1}\left(G_j^{I,Cred}\right)^2\right). &(4.56)
\end{aligned}
$$

Formula (4.55) is exactly the formula for the estimation error of the Bayes CL method in Theorem 4.4 in Gisler–Wüthrich [27]. In the same way as for the process variance, if we use in the Bayesian LSRM and in the Bayes CL method the same estimators $\widehat{\mathrm{E}}\left[\sigma_k^2(G_k)\,\middle|\,\mathcal{D}_k\right]$ and $\widehat{\mathrm{Var}}[G_k|\,\mathcal{D}_k]$ for the structural parameters $\mathrm{E}\left[\sigma_k^2(G_k)\,\middle|\,\mathcal{D}_k\right]$ and $\mathrm{Var}[G_k|\,\mathcal{D}_k]$ (this is the case, since in both methods the Bühlmann-Straub estimators, see Bühlmann-Gisler [14], are proposed) the estimator (4.56) is the well-known estimator for the process variance in the Bayes CL method given in Theorem 4.4 Gisler–Wüthrich [27].

For non-informative priors we have that $\alpha_k \longrightarrow 1$ and for the estimation error we obtain

$$\widehat{\mathrm{E}}\left[\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m \middle| \mathbf{F},\mathcal{D}^I\right] - \sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right)^2 \middle| \mathcal{D}^I\right]$$

$$:= C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\left(\left(\widehat{g}_j^{I,CL}\right)^2 + \frac{\widehat{\mathrm{E}}\left[\sigma_{l_2}^2(G_j)\middle|\mathcal{D}_j\right]}{\sum_{i=0}^{I-j-1} C_{i,j}}\right) - \prod_{j=I-i}^{J-1}\left(\widehat{g}_j^{I,CL}\right)^2\right).$$

This estimator for the estimation error is the same as in Buchwalder et al. [10] for the classical (non-Bayesian) CL method, but different from the one in Mack [38]. For details see Chapter 3 in Wüthrich–Merz [63].

With the same techniques as for single accident years, we also obtain for several accident years the estimates for the MSEP in the Bayes CL method in Gisler–Wüthrich [27] as a special case of CL method looked at as a Bayesian LSRM.

4.2.6 Claims Development Result

Now we turn back to general Bayesian LSRMs and consider the one-year prediction uncertainty in terms of the CDR. In the current business period I, we observe the data $\mathcal{D}^I$ and use the credibility based predictors $\widehat{S}_{i,k+1}^{m|I,(Cred)}$ for the prediction of outstanding incremental claim information. In the next business period (i.e. at time $I+1$ with observed data $\mathcal{D}^{I+1}$) we calculate the credibility based predictors $\widehat{S}_{i,k+1}^{m|I+1,(Cred)}$ for outstanding incremental claim information. The at time $I+1$ observed CDR for accident year $i \in \{I-J+1,\dots,I\}$ and $\mathcal{M} \subseteq \{0,\dots,M\}$ measures the difference between these two predictions:

$$\mathrm{CDR}_i^{\mathcal{M},I+1} = \sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\left(\widehat{S}_{i,k+1}^{m|I,(Cred)} - \widehat{S}_{i,k+1}^{m|I+1,(Cred)}\right).$$

For the estimation of the at time I expected claims development result $\mathrm{CDR}_i^{\mathcal{M},I+1}$ and its (conditional) MSEP we state a theorem, which provides an updating-formula for the credibility predictors $\widehat{\mathbf{F}}_k^{I,Cred}$ that will be crucial in further calculations.

Theorem 4.20 (Updating-formula for $\widehat{\mathbf{F}}_k^{I,Cred}$) *Under Model Assumptions 4.4 for the credibility predictors $\left(\widehat{\mathbf{F}}_k^{I,Cred}\right)_I$ the following updating-formula holds:*

$$\widehat{\mathbf{F}}_k^{I+1,Cred} = \widehat{\mathbf{F}}_k^{I,Cred} + \mathbf{Z}_k^I\left(\mathbf{Y}_{I-k,k} - \widehat{\mathbf{F}}_k^{I,Cred}\right) = \mathbf{Z}_k^I\mathbf{Y}_{I-k,k} + \left(\mathbf{I}-\mathbf{Z}_k^I\right)\widehat{\mathbf{F}}_k^{I,Cred},$$

with

$$\mathbf{Z}_k^I := \mathbf{A}_k^I\mathbf{U}_k^I\left[\mathbf{A}_k^I\mathbf{U}_k^I + \mathrm{E}\left[\,\boldsymbol{\Sigma}_{I-k,k}(\mathbf{F})\middle|\mathcal{D}_k\right]\right]^{-1},$$

where $\mathbf{A}_k^I$ and $\mathbf{U}_k^I$ are defined in Theorem 4.13 and $\boldsymbol{\Sigma}_{i,k}(\mathbf{F})$ is given in Lemma 4.12. $[\mathbf{B}]^{-1}$ denotes a generalized inverse of the matrix $\mathbf{B}$.

Proof: For the loss matrix of the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ we obtain by Theorem 7.5 in Bühlmann-Gisler [14]:

$$\mathrm{E}\left[\left(\mathbf{F}_k - \widehat{\mathbf{F}}_k^{I,Cred}\right)\left(\mathbf{F}_k - \widehat{\mathbf{F}}_k^{I,Cred}\right)' \middle| \mathcal{D}_k\right] = \mathbf{A}_k^I \mathbf{U}_k^I.$$

Now the claim follows by the Kalman-Filter Algorithm in Theorem 10.1 in Bühlmann-Gisler [14] or by Proposition 12.2.2 in Brockwell-Davis [9]. □

Remarks 4.21 (Updating-formula for $\widehat{\mathbf{F}}_k^{I,Cred}$)

i) The credibility predictor $\widehat{\mathbf{F}}_k^{I+1,Cred}$ based on the data $\mathcal{D}^{I+1}$ at time $I+1$ is a credibility weighted average of the new observation $\mathbf{Y}_{I-k,k}$ at time $I+1$ and the previous credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ at time I.

ii) Theorem 4.20 is central for the derivation of the CDR *uncertainty, because it allows to separate the credibility predictor $\widehat{\mathbf{F}}_k^{I,Cred}$ at time I and the new observation $\mathbf{Y}_{I-k,k}$ at time $I+1$, together leading to the new credibility predictor $\widehat{\mathbf{F}}_k^{I+1,Cred}$ at time $I+1$.*

Now we consider what we can say about $\widehat{\mathbf{F}}_k^{I+1,Cred}$ conditionally given $\mathcal{D}^I$ at time I. With Theorem 4.20, Model Assumption 4.4 a) and the fact that $\widehat{\mathbf{F}}_k^{I,Cred}$ is $\mathcal{D}_k^I$-measurable we get

$$\begin{aligned}\bar{\mathbf{F}}_k := \mathrm{E}\left[\widehat{\mathbf{F}}_k^{I+1,Cred} \middle| \mathcal{D}_k^I\right] &= \mathbf{Z}_k^I \mathrm{E}\left[\mathbf{Y}_{I-k,k} \middle| \mathcal{D}_k^I\right] + \left(\mathbf{I} - \mathbf{Z}_k^I\right) \widehat{\mathbf{F}}_k^{I,Cred} \\ &= \mathbf{Z}_k^I \mathrm{E}\left[\mathrm{E}\left[\mathbf{Y}_{I-k,k} \middle| \mathcal{D}_k^I, \mathbf{F}_k\right] \middle| \mathcal{D}_k^I\right] + \left(\mathbf{I} - \mathbf{Z}_k^I\right) \widehat{\mathbf{F}}_k^{I,Cred} \\ &= \mathbf{Z}_k^I \mathrm{E}\left[\mathbf{F}_k \middle| \mathcal{D}_k^I\right] + \left(\mathbf{I} - \mathbf{Z}_k^I\right) \widehat{\mathbf{F}}_k^{I,Cred}.\end{aligned} \tag{4.57}$$

In the same way as in (4.9) we define with $\bar{F}_k^m := \left(\bar{\mathbf{F}}_k\right)_m$

$$\bar{\mathbf{F}} := \left(\bar{F}_k^m\right)_{0 \le k \le J-1}^{0 \le m \le M}.$$

For further calculations we use

$$\begin{aligned}\widetilde{\mathbf{F}}_k &:= \mathrm{E}\left[\bar{\mathbf{F}}_k \middle| \mathcal{D}^I\right] = \mathbf{Z}_k^I \mathrm{E}\left[\mathbf{F}_k \middle| \mathcal{D}^I\right] + \left(\mathbf{I} - \mathbf{Z}_k^I\right) \widehat{\mathbf{F}}_k^{I,Cred} \simeq \widehat{\mathbf{F}}_k^{I,Cred}, \\ \widetilde{\mathbf{F}}_k^* &:= \mathbf{Z}_k^I \mathrm{E}\left[\mathbf{F}_k \middle| \mathcal{D}^I\right] + \left(\mathbf{I} - \mathbf{Z}_k^I\right) \mathrm{E}\left[\mathbf{F}_k \middle| \mathcal{D}_k\right], \\ \mathbf{F}_k^* &:= \mathrm{E}\left[\mathbf{F}_k \middle| \mathcal{D}_k\right].\end{aligned} \tag{4.58}$$

Note that by (4.58) follows that $\bar{\mathbf{F}}_k$ can be estimated at time I by $\widehat{\mathbf{F}}_k^{I,Cred}$. The approximation

$$\widetilde{\mathbf{F}}_k = \mathrm{E}\left[\widehat{\mathbf{F}}_k^{I+1,(Cred)} \middle| \mathcal{D}^I\right] \simeq \widehat{\mathbf{F}}_k^{I,Cred},$$

see (4.58), motivates for the expected claims development result $\mathrm{CDR}_i^{\mathcal{M},I+1}$ at time I the estimate

$$\widehat{\mathrm{E}}\left[\mathrm{CDR}_i^{\mathcal{M},I+1} \middle| \mathcal{D}^I\right] := 0. \tag{4.59}$$

For the derivation of an estimator for the (conditional) MSEP of the claims development result $\mathrm{CDR}_i^{\mathcal{M},I+1}$ for single accident years, we use the predictor in (4.59). This implies

$$\begin{aligned}
\mathrm{msep}_{\mathrm{CDR}_i^{\mathcal{M},I+1}|\mathcal{D}^I}[0] := {} & \mathrm{E}\left[\left(\mathrm{CDR}_i^{\mathcal{M},I+1}-0\right)^2\middle|\mathcal{D}^I\right] \\
= {} & \mathrm{Var}\left[\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\mathcal{D}^I\right] \qquad (4.60)\\
& +\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\left(\widehat{S}_{i,k+1}^{m|I+1,(Cred)}-\widehat{S}_{i,k+1}^{m|I,(Cred)}\right)\middle|\mathcal{D}^I\right]\right)^2, \qquad (4.61)
\end{aligned}$$

where we used that $\widehat{S}_{i,k+1}^{m|I,(Cred)}$ is $\mathcal{D}^I$-measurable. The first term (4.60) corresponds to the process variance, whereas the second term (4.61) is a kind of estimation error.

Process Variance for Single Accident Years

We decompose the process variance (4.60) as follows

$$\begin{aligned}
& \mathrm{Var}\left[\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\mathcal{D}^I\right] \\
& = \mathrm{E}\left[\left(\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\right)^2\middle|\mathcal{D}^I\right]-\left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\mathcal{D}^I\right]\right)^2.
\end{aligned} \qquad (4.62)$$

We begin with the calculation of (4.62) by computing products of conditional expectations of $\widehat{S}_{i,k+1}^{m|I+1,(Cred)}$. For $k_1 < k_2$ and $i_2 + k_2 \geq I$ we get with (4.57)

$$\begin{aligned}
\mathrm{E}\left[\widehat{S}_{i_1,k_1+1}^{m_1|I+1,(Cred)}\widehat{S}_{i_2,k_2+1}^{m_2|I+1,(Cred)}\middle|\mathcal{D}^I\right] &= \mathrm{E}\left[\mathrm{E}\left[\widehat{S}_{i_1,k_1+1}^{m_1|I+1,(Cred)}\widehat{S}_{i_2,k_2+1}^{m_2|I+1,(Cred)}\middle|\mathcal{D}_{k_2}^I\right]\middle|\mathcal{D}^I\right] \\
&= \mathrm{E}\left[\widehat{S}_{i_1,k_1+1}^{m_1|I+1,(Cred)}\bar{P}_{i_2,k_2}^{m_2,\bar{\mathbf{F}}}\widehat{\mathbf{S}}^{i_2+k_2|I+1,(Cred)}\middle|\mathcal{D}^I\right]
\end{aligned} \qquad (4.63)$$

with

$$\bar{P}_{i,k}^{m,\bar{\mathbf{F}}} := \begin{cases} P_{i,k}^m\Big|_{\mathbf{f}_k=\bar{\mathbf{F}}_k} & i+k>I \\ P_{i,k}^m\Big|_{\mathbf{f}_k=\mathrm{E}\left[\mathbf{F}_k\middle|\mathcal{D}_k^I\right]} & i+k\leq I \end{cases}.$$

For

$$\widehat{R}_{i,k}^{m|I+1,(Cred)} := \mathbf{\Gamma}_{i,k}^m\widehat{\mathbf{S}}^{i+k|I+1,(Cred)} := \sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=0}^{(i+k-h)\wedge k}\gamma_{i,k,h,j}^{m,l}\widehat{S}_{h,j}^{l|I+1,(Cred)}$$

we compute in the case of $k_1 = k_2 =: k$ for $i_1, i_2 \geq I-k$

$$
\begin{aligned}
&\mathrm{E}\Big[\,\widehat{S}^{m_1|I+1,(Cred)}_{i_1,k+1}\,\widehat{S}^{m_2|I+1,(Cred)}_{i_2,k+1}\Big|\,\mathcal{D}^I\Big] \\
&= \mathrm{E}\Big[\mathrm{E}\Big[\,\widehat{S}^{m_1|I+1,(Cred)}_{i_1,k+1}\,\widehat{S}^{m_2|I+1,(Cred)}_{i_2,k+1}\Big|\,\mathcal{D}^I_k\Big]\Big|\,\mathcal{D}^I\Big] \\
&= \begin{cases}
\mathrm{E}\Big[\Big(\bar{F}^{m_1}_k\bar{F}^{m_2}_k + \bar{\varrho}^{m_1,m_2}_{i_1,i_2,k}\Big)\,\widehat{R}^{m_1|I+1,(Cred)}_{i_1,k}\,\widehat{R}^{m_2|I+1,(Cred)}_{i_2,k}\Big|\,\mathcal{D}^I\Big] & \text{for } i_1, i_2 > I-k \\
\mathrm{E}\Big[\Big(\mathrm{E}\big[F^{m_1}_k\big|\,\mathcal{D}^I_k\big]\,\bar{F}^{m_2}_k + \bar{\varrho}^{m_1,m_2}_{i_1,i_2,k}\Big)\,\widehat{R}^{m_1|I+1,(Cred)}_{i_1,k}\,\widehat{R}^{m_2|I+1,(Cred)}_{i_2,k}\Big|\,\mathcal{D}^I\Big] & \text{for } i_2 > i_1 = I-k \\
\mathrm{E}\Big[\Big(\bar{F}^{m_1}_k\mathrm{E}\big[F^{m_2}_k\big|\,\mathcal{D}^I_k\big] + \bar{\varrho}^{m_1,m_2}_{i_1,i_2,k}\Big)\,\widehat{R}^{m_1|I+1,(Cred)}_{i_1,k}\,\widehat{R}^{m_2|I+1,(Cred)}_{i_2,k}\Big|\,\mathcal{D}^I\Big] & \text{for } i_1 > i_2 = I-k \\
\mathrm{E}\Big[\Big(\mathrm{E}\big[F^{m_1}_k\big|\,\mathcal{D}^I_k\big]\,\mathrm{E}\big[F^{m_2}_k\big|\,\mathcal{D}^I_k\big] + \bar{\varrho}^{m_1,m_2}_{i_1,i_2,k}\Big)\,\widehat{R}^{m_1|I+1,(Cred)}_{i_1,k}\,\widehat{R}^{m_2|I+1,(Cred)}_{i_2,k}\Big|\,\mathcal{D}^I\Big] & \text{for } i_1 = i_2 = I-k
\end{cases}
\end{aligned}
\tag{4.64}
$$

with

$$
\bar{\varrho}^{m_1,m_2}_{i_1,i_2,k} := \begin{cases}
\mathrm{Cov}\Big[\widehat{F}^{m_1|I+1,Cred}_k, \widehat{F}^{m_2|I+1,Cred}_k\Big|\,\mathcal{D}^I_k\Big] & \text{for } i_1, i_2 > I-k \\
\dfrac{\mathrm{Cov}\Big[S^{m_1}_{i_1,k+1}, \widehat{F}^{m_2|I+1,Cred}_k\Big|\,\mathcal{D}^I_k\Big]}{R^{m_1}_{i_1,k}} & \text{for } i_2 > i_1 = I-k \\
\dfrac{\mathrm{Cov}\Big[\widehat{F}^{m_1|I+1,Cred}_k, S^{m_2}_{i_2,k+1}\Big|\,\mathcal{D}^I_k\Big]}{R^{m_2}_{i_2,k}} & \text{for } i_1 > i_2 = I-k \\
\dfrac{\mathrm{Cov}\Big[S^{m_1}_{i_1,k+1}, S^{m_2}_{i_2,k+1}\Big|\,\mathcal{D}^I_k\Big]}{R^{m_1}_{i_1,k}R^{m_2}_{i_2,k}} & \text{for } i_1 = i_2 = I-k \\
0 & \text{otherwise or denominator equals zero}
\end{cases}.
$$

A short calculation yields

$$
\bar{\varrho}^{m_1,m_2}_{i_1,i_2,k} = \begin{cases}
\Big(\mathbf{Z}^I_k\left(\mathrm{E}\big[\mathbf{\Sigma}_{I-k,k}(\mathbf{F})\big|\,\mathcal{D}^I_k\big] + \mathrm{Cov}\big[\mathbf{F}_k, \mathbf{F}_k\big|\,\mathcal{D}^I_k\big]\right)\mathbf{Z}^{I'}_k\Big)_{m_1,m_2} & \text{for } i_1, i_2 > I-k \\
\Big(\left(\mathrm{E}\big[\mathbf{\Sigma}_{I-k,k}(\mathbf{F})\big|\,\mathcal{D}^I_k\big] + \mathrm{Cov}\big[\mathbf{F}_k, \mathbf{F}_k\big|\,\mathcal{D}^I_k\big]\right)\mathbf{Z}^{I'}_k\Big)_{m_1,m_2} & \text{for } i_2 > i_1 = I-k \\
\left(\mathbf{Z}^I_k\left(\mathrm{E}\big[\mathbf{\Sigma}_{I-k,k}(\mathbf{F})\big|\,\mathcal{D}^I_k\big] + \mathrm{Cov}\big[\mathbf{F}_k, \mathbf{F}_k\big|\,\mathcal{D}^I_k\big]\right)\right)_{m_1,m_2} & \text{for } i_1 > i_2 = I-k \\
\left(\mathrm{E}\big[\mathbf{\Sigma}_{I-k,k}(\mathbf{F})\big|\,\mathcal{D}^I_k\big] + \mathrm{Cov}\big[\mathbf{F}_k, \mathbf{F}_k\big|\,\mathcal{D}^I_k\big]\right)_{m_1,m_2} & \text{for } i_1 = i_2 = I-k \\
0, & \text{otherwise or denominator equals zero}
\end{cases}.
\tag{4.65}
$$

Replacing all unknown components in (4.65) (note that by Remark 4.14 iv) we already have an estimate for the structural matrix $\mathrm{E}[\mathbf{\Sigma}_{i,k}(\mathbf{F})|\,\mathcal{D}_k]$) by the estimates

$$
\begin{aligned}
\widehat{\mathrm{E}}\big[\mathbf{\Sigma}_{I-k,k}(\mathbf{F})\big|\,\mathcal{D}^I_k\big] &:= \mathrm{E}[\mathbf{\Sigma}_{I-k,k}(\mathbf{F})|\,\mathcal{D}_k] \\
\widehat{\mathrm{Cov}}\big[\mathbf{F}_k, \mathbf{F}_k\big|\,\mathcal{D}^I_k\big] &:= \mathbf{A}^I_k\mathbf{U}^I_k
\end{aligned}
\tag{4.66}
$$

leads to an estimator $\widehat{\bar{\boldsymbol{\varrho}}}$ of $\bar{\boldsymbol{\varrho}}$. Replacing $\bar{\boldsymbol{\varrho}}$ by its estimate $\widehat{\bar{\boldsymbol{\varrho}}}$ and using the estimate

$$\widehat{\mathrm{E}}\left[\mathbf{F}_k\middle|\,\mathcal{D}_k^I\right] := \mathbf{F}_k^{I,Cred} \tag{4.67}$$

in (4.64) and (4.63) we obtain the following estimator for each summand in the first addend of (4.62)

$$\widehat{\mathrm{E}}\left[\widehat{S}_{i_1,k_1+1}^{m_1|I+1,(Cred)}\widehat{S}_{i_2,k_2+1}^{m_2|I+1,(Cred)}\middle|\,\mathcal{D}^I\right] = \widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|I,Cred}(\widehat{\bar{\boldsymbol{\varrho}}})\mathbf{S}^I\mathbf{S}^I.$$

Estimator 4.22 (Process variance for single accident years) *Under Model Assumptions 4.4 at time I an estimator for the process variance* (4.60) *for single accident years $i \in \{I-J+1,\ldots,I\}$ is given by*

$$\begin{aligned}&\widehat{\mathrm{Var}}\left[\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\,\mathcal{D}^I\right]\\&\qquad := \sum_{m_1,m_2\in\mathcal{M}}\alpha_i^{m_1}\alpha_i^{m_2}\left(\sum_{k_1,k_2=I-i}^{J-1}\left(\widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|I,Cred}(\widehat{\bar{\boldsymbol{\varrho}}})-\widehat{\mathbf{H}}_{i,k_1,i,k_2}^{m_1,m_2|I,Cred}(\mathbf{0})\right)\mathbf{S}^I\mathbf{S}^I\right).\end{aligned}$$

Estimation Error for Single Accident Years

We get with (4.58), Model Assumptions 4.4 c) and Lemma 4.12 i)

$$\widetilde{S}_{i,k+1}^m := \mathrm{E}\left[\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\,\mathcal{D}^I\right] = \bar{P}_{i,k}^{m|I,\widetilde{\mathbf{F}}}\mathbf{S}^I,$$

where

$$\bar{P}_{i,k}^{m|I,\widetilde{\mathbf{F}}} := P_{i,k}^{m|I}\Big|_{P_{i_1,k_1}^{m_1}=\bar{P}_{i_1,k_1}^{m_1,\widetilde{\mathbf{F}}}} \qquad \text{and} \qquad \bar{P}_{i,k}^{m,\widetilde{\mathbf{F}}} := \begin{cases} P_{i,k}^m\Big|_{\mathbf{f}_k=\widetilde{F}_k} & i+k>I \\ P_{i,k}^m\Big|_{\mathbf{f}_k=\mathrm{E}[\mathbf{F}_k|\,\mathcal{D}^I]} & i+k\leq I\end{cases}$$

and we define

$$\widetilde{\mathbf{S}}^n := \left(\widetilde{S}_{i,k}^m\right)_{i+k\leq n}^{0\leq m\leq M}.$$

Having this notation the estimation error (4.61) can be decomposed into

$$\begin{aligned}\bar{\Delta}^{\mathcal{M}} :&= \left(\mathrm{E}\left[\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\left(\widehat{S}_{i,k+1}^{m|I+1,(Cred)}-\widehat{S}_{i,k+1}^{m|I,(Cred)}\right)\middle|\,\mathcal{D}^I\right]\right)^2\\&= \sum_{m_1,m_2\in\mathcal{M}}\sum_{k_1,k_2=I-i}^{J-1}\Big[\widetilde{S}_{i,k_1+1}^m\widetilde{S}_{i,k_2+1}^m-\widetilde{S}_{i,k_1+1}^m\widehat{S}_{i,k_2+1}^{m_2|I,(Cred)}\\&\qquad -\widehat{S}_{i,k_1+1}^{m_1|I,(Cred)}\widetilde{S}_{i,k_2+1}^m+\widehat{S}_{i,k_1+1}^{m_1|I,(Cred)}\widehat{S}_{i,k_2+1}^{m_2|I,(Cred)}\Big].\end{aligned} \tag{4.68}$$

For the estimation of (4.68) we apply the resampling approach, see Wüthrich–Merz [63], i.e. we define the resampling probability measure as the product measure of

$$\mathrm{P}^*\Big(\widehat{F}_k^{I,Cred} \in A\Big) := \mathrm{P}\Big(\widehat{F}_k^{I,Cred} \in A\Big|\mathcal{D}_k\Big) \tag{4.69}$$

and estimate the estimation error in (4.68) by its expectation under this resampling probability measure. We denote the corresponding expected value under P^* by E^* and the covariance by Cov^*, respectively. For $k_1 > k_2$ and $i_1 + k_1 \geq I$ we obtain for the terms in (4.68)

$$\begin{aligned}
\mathrm{E}^*\Big[\widetilde{S}^{m_1}_{i_1,k_1+1}\widetilde{S}^{m_2}_{i_2,k_2+1}\Big] &= \mathrm{E}^*\Big[\bar{P}^{m_1,\widetilde{\mathbf{F}}^*}_{i_1,k_1}\widetilde{\mathbf{S}}^{i_1+k_1}\widetilde{S}^{m_2}_{i_2,k_2+1}\Big]\\
\mathrm{E}^*\Big[\widehat{S}^{m_1|I,Cred}_{i_1,k_1+1}\widetilde{S}^{m_2}_{i_2,k_2+1}\Big] &= \mathrm{E}^*\Big[P^{m_1,\mathbf{F}^*}_{i_1,k_1}\widehat{\mathbf{S}}^{i_1+k_1|I,Cred}\widetilde{S}^{m_2}_{i_2,k_2+1}\Big]\\
\mathrm{E}^*\Big[\widetilde{S}^{m_1}_{i_1,k_1+1}\widehat{S}^{m_2|I,Cred}_{i_2,k_2+1}\Big] &= \mathrm{E}^*\Big[\bar{P}^{m_1,\widetilde{\mathbf{F}}^*}_{i_1,k_1}\widetilde{\mathbf{S}}^{i_1+k_1}\widehat{S}^{m_2|I,Cred}_{i_2,k_2+1}\Big]\\
\mathrm{E}^*\Big[\widehat{S}^{m_1|I,(Cred)}_{i_1,k_1+1}\widehat{S}^{m_2|I,(Cred)}_{i_2,k_2+1}\Big] &= \mathrm{E}^*\Big[P^{m_1,\mathbf{F}^*}_{i_1,k_1}\widehat{\mathbf{S}}^{i_1+k_1|I,Cred}\widehat{S}^{m_2|I,(Cred)}_{i_2,k_2+1}\Big]
\end{aligned} \tag{4.70}$$

with

$$P^{m,\mathbf{F}^*}_{i,k} := P^m_{i,k}\Big|_{\mathbf{f}_k=\mathrm{E}[\mathbf{F}_k|\mathcal{D}_k]} \qquad \text{and} \qquad \bar{P}^{m,\widetilde{\mathbf{F}}^*}_{i,k} := \begin{cases} P^m_{i,k}\Big|_{\mathbf{f}_k=\widetilde{F}^*_k} & i+k>I \\ P^m_{i,k}\Big|_{\mathbf{f}_k=\mathrm{E}[\mathbf{F}_k|\mathcal{D}^I]} & i+k\leq I \end{cases}.$$

Moreover, in the case of $k_1 = k_2 =: k$ the first two identities in (4.70) still hold if at least one claim property lies on or above the diagonal $I+1$. Otherwise, for $k_1 = k_2 =: k$, we get after short calculations

$$\begin{aligned}
\mathrm{E}^*\Big[\widetilde{S}^{m_1}_{i_1,k+1}\widetilde{S}^{m_2}_{i_2,k+1}\Big] &= \mathrm{E}^*\Big[(\widetilde{F}^{*m_1}_k\widetilde{F}^{*m_2}_k + \varrho^{*11\,m_1,m_2}_{i_1,i_2,k})\widetilde{R}^{m_1}_{i_1,k}\widetilde{R}^{m_1}_{i_2,k}\Big]\\
\mathrm{E}^*\Big[\widehat{S}^{m_1|I,(Cred)}_{i_1,k+1}\widetilde{S}^{m_2}_{i_2,k+1}\Big] &= \mathrm{E}^*\Big[(\mathrm{E}\big[F^{m_1}_k\big|\mathcal{D}_k\big]\widetilde{F}^{*m_2}_k + \varrho^{*12\,m_1,m_2}_{i_1,i_2,k})\widehat{R}^{m_1|I,(Cred)}_{i_1,k}\widetilde{R}^{m_2}_{i_2,k}\Big]\\
\mathrm{E}^*\Big[\widetilde{S}^{m_1}_{i_1,k+1}\widehat{S}^{m_2|I,(Cred)}_{i_2,k+1}\Big] &= \mathrm{E}^*\Big[(\widetilde{F}^{*m_1}_k\mathrm{E}\big[F^{m_2}_k\big|\mathcal{D}_k\big] + \varrho^{*21\,m_1,m_2}_{i_1,i_2,k})\widetilde{R}^{m_1}_{i_1,k}\widehat{R}^{m_2|I,(Cred)}_{i_2,k}\Big]\\
\mathrm{E}^*\Big[\widehat{S}^{m_1|I,(Cred)}_{i_1,k+1}\widehat{S}^{m_2|I,(Cred)}_{i_2,k+1}\Big] &= \mathrm{E}^*\Big[(\mathrm{E}\big[F^{m_1}_k\big|\mathcal{D}_k\big]\,\mathrm{E}\big[F^{m_2}_k\big|\mathcal{D}_k\big] + \varrho^{*22\,m_1,m_2}_{i_1,i_2,k})\widehat{R}^{m_1|I,(Cred)}_{i_1,k}\widehat{R}^{m_2|I,(Cred)}_{i_2,k}\Big]
\end{aligned} \tag{4.71}$$

where

$$\widehat{R}^{m|I,(Cred)}_{i,k} := \mathbf{\Gamma}^m_{i,k}\widehat{\mathbf{S}}^{i+k|I,(Cred)} \qquad \text{and} \qquad \widetilde{R}^m_{i,k} := \mathbf{\Gamma}^m_{i,k}\widetilde{\mathbf{S}}^{i+k}$$

and

$$
\begin{aligned}
\varrho^{*11\, m_1,m_2}_{i_1,i_2,k} &:= \begin{cases} \left(\left(\mathbf{I} - \mathbf{Z}^I_k\right) \mathbf{A}^I_k \left(\mathbf{T}_k + \mathbf{U}^I_k\right) \left(\mathbf{A}^I_k\right)' \left(\mathbf{I} - \mathbf{Z}^I_k\right)' \right)_{m_1,m_2} & \text{for } i_1 + k > I \text{ and } i_2 + k > I \\ 0 & \text{otherwise} \end{cases} \\
\varrho^{*12\, m_1,m_2}_{i_1,i_2,k} &:= \begin{cases} \left(\mathbf{A}^I_k \left(\mathbf{T}_k + \mathbf{U}^I_k\right) \left(\mathbf{A}^I_k\right)' \left(\mathbf{I} - \mathbf{Z}^I_k\right)' \right)_{m_1,m_2} & \text{for } i_1 + k \geq I,\ i_2 + k > I \\ 0 & \text{otherwise} \end{cases} \\
\varrho^{*21\, m_1,m_2}_{i_1,i_2,k} &:= \begin{cases} \left(\left(\mathbf{I} - \mathbf{Z}^I_k\right) \mathbf{A}^I_k \left(\mathbf{T}_k + \mathbf{U}^I_k\right) \left(\mathbf{A}^I_k\right)' \right)_{m_1,m_2} & \text{for } i_1 + k > I,\ i_2 + k \geq I \\ 0 & \text{otherwise} \end{cases} \\
\varrho^{*22\, m_1,m_2}_{i_1,i_2,k} &:= \begin{cases} \left(\mathbf{A}^I_k \left(\mathbf{T}_k + \mathbf{U}^I_k\right) \left(\mathbf{A}^I_k\right)' \right)_{m_1,m_2} & \text{for } i_1 + k \geq I \text{ and } i_2 + k \geq I \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\quad . \tag{4.72}
$$

Summarizing all parts and replacing all unknown parameters in (4.70), (4.71) and (4.72) by their estimates leads to the following estimator:

Estimator 4.23 (Estimation error for single accident years) *Under Model Assumptions 4.4 at time I an estimator for the estimation error* (4.61) *for single accident years $i \in \{I - J + 1, \ldots, I\}$ is given by*

$$
\begin{aligned}
\widehat{\widehat{\Delta}}^{\mathcal{M}} := \sum_{m_1,m_2 \in \mathcal{M}} \alpha_i^{m_1} \alpha_i^{m_2} \Bigg(\sum_{k_1,k_2=I-i}^{J-1} \Big(& \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*11}) - \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*12}) \\
& - \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*21}) + \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*22}) \Big) \mathbf{S}^I \mathbf{S}^I \Bigg).
\end{aligned}
$$

Mean Squared Error of Prediction for Single Accident Years

Combining the Estimators 4.22 and 4.23 implies

Estimator 4.24 (MSEP for single accident years) *Under Model Assumptions 4.4 at time I an estimator for the (conditional) mean squared error of prediction for single accident years $i \in \{I - J + 1, \ldots, I\}$ is given by*

$$
\begin{aligned}
\operatorname{msep}_{\mathrm{CDR}_i^{\mathcal{M},I+1}|\mathcal{D}^I}[0] := \sum_{m_1,m_2 \in \mathcal{M}} \alpha_i^{m_1} \alpha_i^{m_2} \Bigg(\sum_{k_1,k_2=I-i}^{J-1} \Big(& \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}) - \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\mathbf{0}) \\
& + \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*11}) - \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*12}) \\
& - \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*21}) + \widehat{\mathbf{H}}^{m_1,m_2|I,Cred}_{i,k_1,i,k_2}(\widehat{\boldsymbol{\varrho}}^{*22}) \Big) \mathbf{S}^I \mathbf{S}^I \Bigg).
\end{aligned}
$$

Mean Squared Error of Prediction for Aggregated Accident Years

In the same way as for single accident years we decompose the MSEP into

$$\operatorname{msep}_{\sum_{i=I-J+1}^{I} \mathrm{CDR}_i^{\mathcal{M},I+1}\big|\mathcal{D}^I}[0] \tag{4.73}$$

$$:= \mathrm{E}\left[\left(\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\left(\widehat{S}_{i,k+1}^{m|I+1,(Cred)}-\widehat{S}_{i,k+1}^{m|I,(Cred)}\right)-0\right)^2\middle|\mathcal{D}^I\right]$$

$$= \operatorname{Var}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\mathcal{D}^I\right] \tag{4.74}$$

$$+\left(\mathrm{E}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\left(\widehat{S}_{i,k+1}^{m|I+1,(Cred)}-\widehat{S}_{i,k+1}^{m|I,(Cred)}\right)\middle|\mathcal{D}^I\right]\right)^2. \tag{4.75}$$

We decompose the process variance (4.74) by

$$\operatorname{Var}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\mathcal{D}^I\right]$$
$$= \mathrm{E}\left[\left(\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\right)^2\middle|\mathcal{D}^I\right]$$
$$-\left(\mathrm{E}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^{m|I+1,(Cred)}\middle|\mathcal{D}^I\right]\right)^2,$$

and the estimation error (4.75) into

$$\left(\mathrm{E}\left[\sum_{i=I-J+1}^{I}\sum_{m\in\mathcal{M}}\alpha_i^m\sum_{k=I-i}^{J-1}\left(\widehat{S}_{i,k+1}^{m|I+1,(Cred)}-\widehat{S}_{i,k+1}^{m|I,(Cred)}\right)\middle|\mathcal{D}^I\right]\right)^2$$
$$= \sum_{i_1,i_2=I-J+1}^{I}\sum_{m_1,m_2\in\mathcal{M}}\alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2}\left(\sum_{k_1,k_2=I-i}^{J-1}\left[\widetilde{S}_{i_1,k_1+1}^{m}\widetilde{S}_{i_2,k_2+1}^{m}-\widetilde{S}_{i_2,k_1+1}^{m}\widehat{S}_{i,k_2+1}^{m_2|I,(Cred)}\right.\right.$$
$$\left.\left.-\widehat{S}_{i_1,k_1+1}^{m_1|I,(Cred)}\widetilde{S}_{i_2,k_2+1}^{m}+\widehat{S}_{i_1,k_1+1}^{m_1|I,(Cred)}\widehat{S}_{i_2,k_2+1}^{m_2|I,(Cred)}\right]\right).$$

We obtain in the same way as for single accident years the following result:

Estimator 4.25 (MSEP for aggregated accident years) *Under Model Assumptions 4.4 at time I the mean squared error of prediction for aggregated accident years can be estimated by*

$$\widehat{\text{msep}}_{\sum_{i=I-J+1}^{I} \text{CDR}_i^{\mathcal{M},I+1} \Big| \mathcal{D}^I}[0]$$
$$:= \sum_{i_1,i_2=I-J+1}^{I} \sum_{m_1,m_2\in\mathcal{M}} \alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2} \left(\sum_{k_1=I-i_1}^{J-1} \sum_{k_2=I-i_2}^{J-1} \Big(\widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}) - \widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\mathbf{0}) \right.$$
$$+ \widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}^{*11}) - \widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}^{*12})$$
$$\left. - \widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}^{*21}) + \widehat{\mathbf{H}}_{i_1,k_1,i_2,k_2}^{m_1,m_2|I,Cred}(\widehat{\boldsymbol{\varrho}}^{*22})\Big)\mathbf{S}^I\mathbf{S}^I \right).$$

4.2.7 Special Case: Claims Development Result for the Bayes CL Method

We saw in Subsection 4.2.1 that the Bayes CL method in Gisler–Wüthrich [27] belongs to the class of Bayesian LSRMs. In Subsection 4.2.5 we showed that in the case of the Bayes CL method the Bayesian LSRM estimate for the MSEP coincides with the estimate derived in Gisler–Wüthrich [27] for the Bayes CL method. Now we consider whether this also holds true for the MSEP of the CDR. The MSEP for the CDR in the Bayesian CL method was derived in Bühlmann et al. [13]. In this derivation only the process variance is taken into account, whereas the estimation error is set to 0. This becomes clear by Formula (4.5) in Bühlmann et al. [13], where in the first step the approximation $G_k^{n,Cred} \simeq \mathrm{E}[G_k|\,\mathcal{D}^n]$ for $n \in \{I, I+1\}$ (exact credibility case) is used leading to

$$\begin{aligned} \text{msep}_{\text{CDR}_i^{0,I+1}|\mathcal{D}^I}[0] &= \mathrm{E}\Big[\Big(\text{CDR}_i^{0,I+1} - 0\Big)^2\Big|\,\mathcal{D}^I\Big] \\ &= \mathrm{Var}\Big[\text{CDR}_i^{0,I+1}\Big|\,\mathcal{D}^I\Big] + \mathrm{E}\Big[\text{CDR}_i^{0,I+1}\Big|\,\mathcal{D}^I\Big]^2 \\ &\simeq \mathrm{Var}\Big[\text{CDR}_i^{0,I+1}\Big|\,\mathcal{D}^I\Big], \end{aligned} \tag{4.76}$$

i.e. the calculation of the MSEP of the CDR is reduced to the calculation of the conditional variance (4.76). In the Bayesian LSRM we additionally quantify the estimation error, see Estimator 4.23. Thus, Estimator 4.24 for the MSEP of the CDR in the Bayesian LSRM and the estimator in Bayes CL method in Result 4.1 in Bühlmann et al. [13] do not coincide. However, we show that the Estimator 4.22 for the process variance in the Bayesian LSRM is identical with the estimator of the MSEP in the Bayes CL method in Result 4.1 in Bühlmann et al. [13]. In order to prove this equality we have to verify that in the derivation of the process variance (4.76) in both methods the same approximations are used. In the Bayes CL method in Bühlmann et al. [13] Lemmata 4.5 and 4.6 use the following approximations (with slightly other notation)

$$\begin{aligned}
\mathrm{E}\big[\,\sigma_k^2(G_k)\big|\,\mathcal{D}^I\big] &\simeq \widehat{\mathrm{E}}\big[\,\sigma_k^2(G_k)\big|\,\mathcal{D}^I\big] \\
\mathrm{Var}\big[\,G_k\big|\,\mathcal{D}^I\big] &\simeq \mathrm{E}\Big[\Big(G_k^{I,Cred} - G_k\Big)\Big|\,\mathcal{D}^I \cap \mathcal{D}_k\Big] \\
\mathrm{E}\big[\,G_k\big|\,\mathcal{D}^I\big] &\simeq G_k^{I,Cred} \\
\mathrm{Var}\bigg[\Big(G_k^{I+1,Cred}\Big)^2\bigg|\,\mathcal{D}^I\bigg] &\simeq \mathrm{E}\bigg[\Big(G_k^{I+1,Cred} - G_k^{I,Cred}\Big)^2\bigg|\,\mathcal{D}^I \cap \mathcal{D}_k\bigg] + \Big(G_k^{I,Cred}\Big)^2 .
\end{aligned}$$

Recalling the identities

$$\mathbf{F} = \mathbf{G} - 1 \quad \text{and} \quad \sigma_k^2(G_k) = \sigma_k^{0,0}(F_k^0)$$

(cf. Subsection 4.2.1) we see that exactly the same approximations are used in the Bayesian LSRM, see (4.66) and (4.67), for the estimation of the process variance of the CDR. Consequently, the Estimator 4.22 and the estimator in Result 4.1 in Bühlmann et al. [13] coincide. By the same arguments as above also follows that the MSEP of the CDR for several accident years in the Bayes CL method, see Result 4.7 in Bühlmann et al. [13], coincides with the Bayesian LSRM estimator for the process variance for several accident years, given by the first line of Estimator 4.25. For a discussion of the case that non-informative priors are used and the link to the MSEP of the CDR in the classical CL method we refer to Bühlmann et al. [13], Merz–Wüthrich [45] and Wüthrich et al. [64].

4.3 Example Bayesian LSRM

For a detailed knowledge of profitability and a better understanding of pricing for different business units (BU) we have to calculate best-estimate reserves and its corresponding prediction uncertainty in terms of the MSEP for each BU. Furthermore, we consider the variability (MSEP) of the CDR as a measure for the one-year reserving risk, what is required under Solvency 2 and SST.

For our example we revisit the building engineering data set of Winterthur Insurance Company presented by Gisler-Wüthrich [27]. It contains trapezoids of incremental claims payments of six BUs multiplied by a constant due to confidentiality reasons. For simplicity and illustration purposes we pick out BU 1–3 (of totally 6 BU) and apply the Bayesian LSRM. The data used is provided in Tables 7.5–7.7.

For the detailed specification of the Bayesian LSRM we choose the exposure $R_{i,k}^m$ to be the sum over all payments from all BUs in accident year i up to development year k. In a similar way we use for the coupling $R_{i,k}^{m_1,m_2}$ of these three BUs all payments from all BUs in accident year i up to development year k. For the structural parameter $\sigma_k^{m_1,m_2}(\mathbf{F})$ we use the unbiased estimator given in (4.5).

Method	Reserves			$\text{MSEP}^{1/2}$		
	1	2	3	1	2	3
CL	486	235	701	657 (135%)	288 (122%)	411 (58%)
Cred CL	504	244	517	498 (99%)	402 (164%)	520 (100%)
LSRM	245	340	598	774 (316%)	273 (80%)	346 (58%)
Bayesian LSRM	267	349	594	520 (195%)	222 (64%)	248 (42%)

Table 4.1: Reserves and prediction uncertainty

The prior covariance matrix $\mathbf{T}_k$ of the development factors $\mathbf{F}_k$ is, see (4.18), a component of the weights $\mathbf{A}_k^I$ given to the prior mean $\boldsymbol{\mu}_k$ and the observation $\mathbf{C}_k$. That means that $\mathbf{T}_k$ can be interpreted as an input parameter reflecting the actuaries confidence in the data in comparison to the believe in prior expert knowledge (see Theorem 4.13). In our example the confidence in the data of development year k is the higher, the more observations are available in development year k and thus we choose $\mathbf{T}_k = 0.0001(10-k)\cdot\mathbf{I}$. In order to get a smoothing effect for the development pattern, we choose the prior mean $\boldsymbol{\mu}_k$ to be the mean over all BUs of development factor estimates resulting from the classical LSRM, i.e.

$$\boldsymbol{\mu}_k^m := \frac{1}{3}(\widehat{f}_k^0 + \widehat{f}_k^1 + \widehat{f}_k^2) \qquad \text{for} \qquad m \in \{0,1,2\}.$$

Applying the Bayesian LSRM for this parameter constellation leads to reserves and corresponding MSEP given in the last line of Table 4.1. Compared to the classical LSRM results, we observe a smoothing effect (lower fluctuations in the reserves of different BUs) for the reserves in the credibility case. This is a direct consequence of the smoothing effect of credibility for each individual development factor. The incorporation of prior knowledge for the mean of the development factors has a common influence on the development factors of all BUs, i.e. the credibility development factors are inbetween the classical LSRM development factors and the prior means. This smoothing effect of prior information on the credibility development factors is illustrated for BU 1 in Figure 4.2. For BU 2 and 3 we obtain similar results (not shown here). Similar smoothing effects were observed in the Bayes CL method in Gisler-Wüthrich [27]. For the prediction uncertainty (MSEP) we obtain slightly lower values as in the classical LSRM in all BUs. This is not always the case, because the prediction uncertainty depends directly on the prior (co)variance for the development factors (see Estimator 4.16 for that). Now we take a look on the one-year reserving risk uncertainty in the CDR presented in Table 4.2. The one-year prediction uncertainty in the Bayesian LSRM only slightly differs from the uncertainty in classical LSRM and the estimates are quite robust with respect to different prior choices for the development factor covariance matrix.

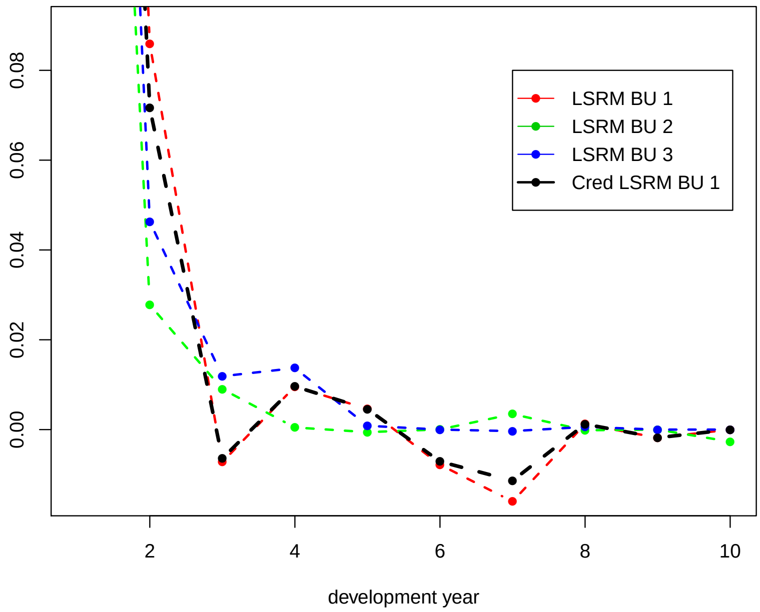

Figure 4.2: Development factors for BUs 1–3 in the classical LSRM and credibility development factor $\widehat{F}_k^{0|I,Cred}$ $k \in \{0, \ldots, 10\}$ for BU 1

Method	$\mathrm{MSEP}_{\mathrm{CDR}}^{1/2}$				$\mathrm{MSEP}_{\mathrm{CDR}}^{1/2}/\mathrm{MSEP}^{1/2}$		
	1	2	3	$\sum$	1	2	3
LSRM	494	207	240	659	64%	76%	69%
Bayesian LSRM	480	207	240	646	92%	93%	96%

Table 4.2: Individual LoB and overall CDR uncertainty

4.4 Conclusions

The classical LSRMs presented in Dahms [17] constitute a wide class of distribution-free stochastic claims reserving methods covering many popular distribution-free stochastic claims reserving methods such as the CL, BF and (E)CLR method. As already mentioned in Dahms [17] for setting up an adequate claims reserving method it is crucial to identify appropriate exposures for the stochastic dynamics in order to specify a LSRM. If one is interested in using prior expert knowledge or information from industry-wide data in the LSRM framework the Bayesian LSRMs presented in this chapter provide an appropriate mathematically consistent basis for the incorporation of such information. Conservative prior means for the development factors can generate risk-margins in the resulting credibility predictors and hence in the corresponding reserves. Moreover, Bayesian LSRMs provide the welcome effect of smoother development factors as shown in Figure 4.2.

Main results of (Bayesian) LSRMs:

For solvency considerations in Chapter 7 we summarize all quantities of interest derived in the LSRM framework:

1. The predictor $\widehat{\mathcal{R}}^I$ for outstanding loss liabilities $\mathcal{R}^I$, see (2.5c) and (2.4c), given by

$$\widehat{\mathcal{R}}^I = \sum_{i=I-J+1}^{I} \sum_{m\in\mathcal{M}} \sum_{k=I-i}^{J-1} \alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)} \tag{4.77}$$

2. The estimator for the prediction uncertainty in terms of the (conditional) MSEP

$$\widehat{\mathrm{msep}}_{\sum\limits_{i=I-J+1}^{I}\sum\limits_{m\in\mathcal{M}}\sum\limits_{k=I-i}^{J-1}\alpha_i^m S_{i,k+1}^m\big|\mathcal{D}^I}\left[\sum_{i=0}^{I}\sum_{m\in\mathcal{M}}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m|I,(Cred)}\right] \tag{4.78}$$

given by Estimator 4.19.

3. The estimator for the CDR uncertainty in terms of the (conditional) MSEP

$$\widehat{\mathrm{msep}}_{\sum\limits_{i=I-J+1}^{I}\mathrm{CDR}_i^{\mathcal{M},I+1}\big|\mathcal{D}^I}[0] \tag{4.79}$$

given by Estimator 4.25.

5 Paid-Incurred Chain Reserving Method

In the previous chapter we considered the class of LSRMs which covers many popular distribution-free claims reserving methods and gives a new perspective on these methods. In a second step we extended the LSRMs to the class of Bayesian LSRMs that allows for the incorporation of prior knowledge of the development pattern. In this Bayesian LSRM framework we stated explicit predictors for the outstanding loss liabilities $\mathcal{R}^I$ and estimates for its associated prediction and CDR uncertainty. This shows that all classical risk characteristics in Table 2.1 can be calculated in the Bayesian LSRM framework. However, with respect to recent solvency regulation in Solvency II and SST insurance liabilities have often to be evaluated by risk measures like VaR and ES, see AISAM–ACME [2] and FOPI [24], and not only by the risk characteristics given in Table 2.1. The knowledge of the predictive distribution of outstanding loss liabilities an the CDR allows for the calculation of such risk measures, see Robert [51], Merz–Wüthrich [46] and Happ et al. [30]. That means that we are interested in claims reserving modeling where the predictive distribution of outstanding claims payments and the distribution of the CDR can be derived (analytically or simulatively).

A second important aspect which is assigned to the choice of a claims reserving method is the data which can be incorporated in the method. In insurance practice cumulative claims payments and incurred losses data are often available and should therefore be utilized for the prediction of outstanding claims payments. Thus, we are looking for a flexible model i) which is able to cope with these two data sources and ii) allows for the derivation of the predictive distribution of outstanding claims payments and the CDR. At first, we consider the task of modeling cumulative claims payments and incurred losses simultaneously.

The MCL method in Quarg–Mack [50] addresses this problem, see also Section 3.6. This method reduces the gap between the CL ultimate claim predictor based purely on cumulative claims payments data and the CL ultimate claim predictor based on incurred losses, respectively, but does not close this gap. However, since the MCL method is distribution-free no predictive distribution can be derived. Moreover, to the best of our knowledge, even estimates for the (conditional) MSEP of the ultimate claim and the CDR have not been found up to now.

Another distribution-free approach to the problem of the incorporation of cumulative claims payments and incurred losses in claims reserving is presented in Dahms [16] with the ECLR

method. The ECLR method was the first claims reserving method which can cope with cumulative claims payments and incurred losses simultaneously leading to one unified ultimate claim prediction. It allows for the derivation of predictors for the ultimate claim and the CDR and estimates for their corresponding (conditional) MSEP. Unfortunately, the ECLR method also does not allow for the derivation of a predictive distribution.
An alternative is presented in Merz–Wüthrich [46] by a distributional approach. Based on Hertig's log-normal claims reserving method (cf. Hertig [32]) and Gogol's Bayesian claims reserving method for incurred losses (cf. Gogol [28]), Merz and Wüthrich introduced the PIC reserving method and provided an ultimate claim predictor as well as the corresponding prediction uncertainty. In this chapter we derive the uncertainty of the CDR for the PIC reserving method and calculate the predictive distribution of the CDR. This is crucial for new solvency considerations, see Chapter 7 as well as AISAM–ACME [2] and FOPI [24], [25]. In this chapter we follow Happ et al. [30].
Notational convention:
In this and in the following chapter we use the notation employed in the PIC reserving method in Merz–Wüthrich [46] for consistency reasons. Therefore, we recapitulate the notation used in this paper.

5.1 Notation and Model Assumptions

The PIC reserving method combines two channels of information: i) claims payments, which correspond to the payments for reported claims; ii) incurred losses, which refer to the reported claim amounts. In the following, we assume $I = J$ for notational simplicity, but all results hold true also in the case $I > J$. Cumulative claims payments in accident year i after j development years are denoted by $P_{i,j}$ and the corresponding incurred losses by $I_{i,j}$. The crucial observation is that the claims payments and incurred losses time series must reach the same ultimate value, because these two time series both converge to the total ultimate claim. Therefore, we assume that all claims are settled and closed after development year J, i.e. $P_{i,J} = I_{i,J}$ holds with probability 1 for all $i \in \{0, \dots, J\}$, see Model Assumptions 5.1. After accounting year $t = J$ we have observations in the paid and incurred triangles given by (see Figure 5.1)

$$\mathcal{D}_J := \{P_{i,j}, I_{i,j};\ 0 \le i \le J,\ 0 \le j \le J,\ 0 \le i+j \le J\}.$$

After accounting year $t = J + 1$ we have observations in the paid and incurred trapezoids given by (see Figure 5.2)

$$\mathcal{D}_{J+1} = \{P_{i,j}, I_{i,j};\ 0 \le i \le J,\ 0 \le j \le J,\ 0 \le i+j \le J+1\}.$$

This means the updating of information $\mathcal{D}_J \mapsto \mathcal{D}_{J+1}$ adds a new diagonal to the observations.

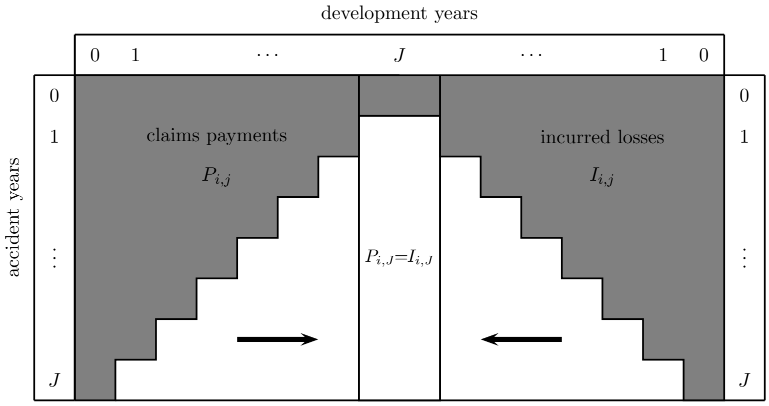

Figure 5.1: Cumulative claims payments $P_{i,j}$ and incurred losses $I_{i,j}$ observed at time $t = J$ both leading to the ultimate loss $P_{i,J} = I_{i,J}$

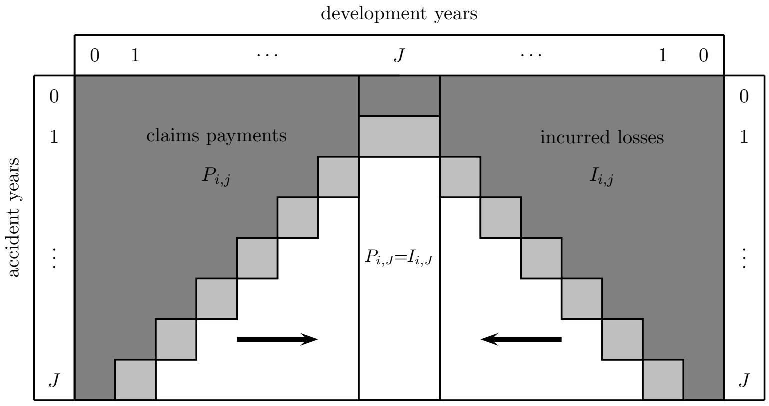

Figure 5.2: Updated cumulative claims payments $P_{i,j}$ and incurred losses $I_{i,j}$ observed at time $t = J + 1$

Our goal is to predict the ultimate losses $P_{i,J} = I_{i,J}$, $i = 1, \ldots, J$, based on the information $\mathcal{D}_J$ and $\mathcal{D}_{J+1}$, respectively. We state the PIC model, which combines both cumulative payments and incurred losses information:

Model Assumptions 5.1 (PIC model)

a) Conditionally, given the parameter vector $\mathbf{\Theta} := (\Phi_0; \Phi_1, \Psi_1, \Phi_2, \Psi_2, \ldots, \Phi_J, \Psi_J)'$, *we assume:*

- *the random vectors* $\mathbf{\Xi}_i := (\xi_{i,0}; \xi_{i,1}, \zeta_{i,1}, \xi_{i,2}, \zeta_{i,2}, \ldots, \xi_{i,J}, \zeta_{i,J})'$ *are i.i.d. with multivariate Gaussian distribution*

$$\mathbf{\Xi}_i \sim \mathcal{N}(\mathbf{\Theta}, \mathbf{V}) \qquad \textit{for} \quad i \in \{0, 1, \ldots, J\}$$

and positive definite covariance matrix $\mathbf{V}$ *as well as individual development factors*

$$\xi_{i,j} := \log \frac{P_{i,j}}{P_{i,j-1}} \qquad \textit{and} \qquad \zeta_{i,l} := \log \frac{I_{i,l}}{I_{i,l-1}},$$

for $j \in \{0, 1, \dots, J\}$ and $l \in \{1, 2, \dots, J\}$, where we have set $P_{i,-1} := 1$;

- $P_{i,J} = I_{i,J}$, *$\mathbb{P}$-a.s., for all $i \in \{0, 1, \dots, J\}$.*

b) The components of $\boldsymbol{\Theta}$ are independent with prior distributions

$$\Phi_j \sim \mathcal{N}(\phi_j, s_j^2) \quad \textit{for} \quad j \in \{0, \dots, J\} \qquad \textit{and} \qquad \Psi_l \sim \mathcal{N}(\psi_l, t_l^2) \quad \textit{for} \quad l \in \{1, \dots, J\}$$

with prior parameters ϕ_j, $\psi_l \in \mathbb{R}$ and $s_j^2 > 0$, $t_l^2 > 0$.

□

Remarks 5.2 (PIC reserving method)

i) In Model Assumptions 5.1 we can choose any arbitrary positive definite covariance matrix $\mathbf{V}$. This allows for modeling dependence structures between claims payments ratios $\frac{P_{i,j}}{P_{i,j-1}}$ and incurred losses ratios $\frac{I_{i,l}}{I_{i,l-1}}$.

ii) Expert opinion should be included to structure the covariance matrix $\mathbf{V}$. For a more detailed discussion on this topic and suitable choices for $\mathbf{V}$ we refer to Happ–Wüthrich *[31]. However, the problem of finding statistically optimal estimators should be subject to further statistical research.*

iii) We define prior distributions for the components of the mean vector $\boldsymbol{\Theta}$ and assume $\mathbf{V}$ to be a given covariance matrix. This Bayesian approach guarantees closed form results. If we also put a prior on $\mathbf{V}$ we have to use Markov-Chain-Monte-Carlo (MCMC) methods for the calculation of the posterior distribution (see Merz–Wüthrich *[46]).*

5.2 One-year Claims Development Result

We consider the short term (one-year) run-off risk introduced in Section 2.6. This means, we study the uncertainty in the one-year CDR for accounting year $J+1$ given by

$$\mathrm{CDR}_i^{J+1} = \mathrm{E}[\, P_{i,J} |\, \mathcal{D}_J] - \mathrm{E}[\, P_{i,J} |\, \mathcal{D}_{J+1}] \,, \quad i = 1, \dots, J,$$

between the best estimates for the ultimate claim $P_{i,J}$ at times J and $J+1$. The one-year CDR in accounting year $J+1$ measures the change in the prediction by updating the information from $\mathcal{D}_J$ to $\mathcal{D}_{J+1}$. With the tower property of the conditional expectation we obtain for the expected one-year CDR for accident year i, viewed from time J,

$$\mathrm{E}\Big[\, \mathrm{CDR}_i^{J+1} \Big|\, \mathcal{D}_J \Big] = 0,$$

which is the martingale property of successive predictions. This justifies the fact that, in the budget statement, the one-year CDR is usually predicted by 0 at time J. In the following we study the uncertainty in this prediction by means of the conditional MSEP, given the observations $\mathcal{D}_J$. In other words we calculate, see Wüthrich–Merz [63], Section 3.1,

$$\text{msep}_{\text{CDR}_i^{J+1}|\mathcal{D}_J}[0] = \text{E}\left[\left(\text{CDR}_i^{J+1} - 0\right)^2 \middle| \mathcal{D}_J\right] = \text{Var}\left[\text{CDR}_i^{J+1}\middle| \mathcal{D}_J\right] = \text{Var}[\text{E}[P_{i,J}|\,\mathcal{D}_{J+1}]|\,\mathcal{D}_J]\,. \tag{5.1}$$

The conditional MSEP is probably the most popular uncertainty measure in claims reserving practice and has the advantage that it can be derived analytically in the PIC model. Moreover, we also present the full predictive distribution below, which also allows to evaluate other risk measures.

5.3 Expected Ultimate Claim at Time $J+1$

In this section we derive the conditional expected ultimate claim $\text{E}[P_{i,J}|\,\mathcal{D}_k]$ for $k \in \{J, J+1\}$ in two steps. In the first step we derive $\text{E}[P_{i,J}|\,\boldsymbol{\Theta}, \mathcal{D}_k]$ and in the second step we calculate $\text{E}[P_{i,J}|\,\mathcal{D}_k]$, see Corollary 5.7.

In the following we can either work with the random vector $\boldsymbol{\Xi}_i \in \mathbb{R}^{2J+1}$ (see Model Assumptions 5.1) or with the logarithmized observations of accident year i, namely,

$$\mathbf{X}_i := (\log P_{i,0}, \log I_{i,0}, \log P_{i,1}, \ldots, \log P_{i,J-1}, \log I_{i,J-1}, \log P_{i,J})' \in \mathbb{R}^{2J+1}.$$

This is possible, since there exist an invertible matrix $\mathbf{B} \in \mathbb{R}^{(2J+1)\times(2J+1)}$ such that $\mathbf{X}_i = \mathbf{B}\ \boldsymbol{\Xi}_i$, i.e. there is a one-to-one correspondence between $\mathbf{X}_i$ and $\boldsymbol{\Xi}_i$. This implies

$$\mathbf{X}_i|_{\boldsymbol{\Theta}} = \mathbf{B}\ \boldsymbol{\Xi}_i|_{\boldsymbol{\Theta}} \sim \mathcal{N}(\boldsymbol{\mu} := \mathbf{B}\boldsymbol{\Theta}, \boldsymbol{\Sigma} := \mathbf{B}\mathbf{V}\mathbf{B}'). \tag{5.2}$$

Let $k \in \{J, J+1\}$ and define $n := 2J+1$ and $q := q_k(i) := 2(k-i+1)$. To simplify notation we define:

$$\mathbf{X}_{i,k}^{(1)} := \begin{cases} (\log P_{i,0}, \log I_{i,0}, \log P_{i,1}, \log I_{i,1} \ldots, \log P_{i,k-i}, \log I_{i,k-i})' \in \mathbb{R}^q & \text{for } k-i < J, \\ \mathbf{X}_i & \text{otherwise;} \end{cases}$$

$$\mathbf{X}_{i,k}^{(2)} := \begin{cases} (\log P_{i,k-i+1}, \log I_{i,k-i+1}, \ldots, \log P_{i,J-1}, \log I_{i,J-1}, \log P_{i,J})' \in \mathbb{R}^{n-q} & \text{for } k-i < J, \\ (\log P_{i,J}) & \text{otherwise.} \end{cases}$$

$\mathbf{X}_{i,k}^{(1)}$ describes the observations at time $k \in \{J, J+1\}$, i.e. it corresponds to the σ-field generated by $\mathcal{D}_k$, see Figures 5.1 and 5.2. $\mathbf{X}_{i,k}^{(2)}$ is the part of claims development that needs to be predicted at time k for $i > k - J$.

For $k - i < J$ we decompose the transformation matrix $\mathbf{B}$ in a similar way into

$$\mathbf{B} := \begin{pmatrix} \mathbf{B}_{i,k}^{(1)} \\ \mathbf{B}_{i,k}^{(2)} \end{pmatrix}, \tag{5.3}$$

where $\mathbf{B}_{i,k}^{(1)} \in \mathbb{R}^{q\times n}$. For $k-i \geq J$ we set $\mathbf{B}_{i,k}^{(1)} := \mathbf{B}$ and $\mathbf{B}_{i,k}^{(2)} := \mathbf{B}_{1,J}^{(2)}$, but (5.3) does not hold in this case. We obtain, for $k-i<J$, a decomposition

$$\boldsymbol{\mu} = \mathbf{B}\boldsymbol{\Theta} = \left(\boldsymbol{\mu}_{i,k}^{(1)}, \boldsymbol{\mu}_{i,k}^{(2)}\right)' \in \mathbb{R}^n$$

of the mean vector, where

$$\boldsymbol{\mu}_{i,k}^{(1)} := \mathrm{E}\left[\mathbf{X}_{i,k}^{(1)}\middle|\,\boldsymbol{\Theta}\right] = \mathbf{B}_{i,k}^{(1)}\boldsymbol{\Theta} \quad \text{and} \quad \boldsymbol{\mu}_{i,k}^{(2)} := \mathrm{E}\left[\mathbf{X}_{i,k}^{(2)}\middle|\,\boldsymbol{\Theta}\right] = \mathbf{B}_{i,k}^{(2)}\boldsymbol{\Theta}.$$

For $k-i<J$ the covariance matrix is decomposed in a similar way such that

$$\boldsymbol{\Sigma} = \mathbf{B}\mathbf{V}\mathbf{B}' = \begin{pmatrix} \boldsymbol{\Sigma}_{i,k}^{(11)} & \boldsymbol{\Sigma}_{i,k}^{(12)} \\ \boldsymbol{\Sigma}_{i,k}^{(21)} & \boldsymbol{\Sigma}_{i,k}^{(22)} \end{pmatrix} \in \mathbb{R}^{n\times n}, \tag{5.4}$$

with $\boldsymbol{\Sigma}_{i,k}^{(11)} \in \mathbb{R}^{q\times q}$. For $k-i\geq J$ we set $\boldsymbol{\Sigma}_{i,k}^{(11)} = \boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_{i,k}^{(12)} = \boldsymbol{\Sigma}_{1,J}^{(12)}$ and $\boldsymbol{\Sigma}_{i,k}^{(22)} = \boldsymbol{\Sigma}_{1,J}^{(22)}$, but (5.4) does not hold in this case. Now having this notation we provide the following lemma:

Lemma 5.3 (Conditional distribution) *Choose $k \in \{J, J+1\}$ and $i > k-J$. Under Model Assumptions 5.1 we obtain for the conditional distribution of $\mathbf{X}_{i,k}^{(2)}$, given $\{\boldsymbol{\Theta}, \mathcal{D}_k\}$,*

$$\mathbf{X}_{i,k}^{(2)}|_{\{\boldsymbol{\Theta},\mathcal{D}_k\}} = \mathbf{X}_{i,k}^{(2)}|_{\{\boldsymbol{\Theta},\mathbf{X}_{i,k}^{(1)}\}} \sim \mathcal{N}\left(\tilde{\boldsymbol{\mu}}_{i,k}^{(2)}, \tilde{\boldsymbol{\Sigma}}_{i,k}^{(22)}\right),$$

where

$$\begin{aligned} \tilde{\boldsymbol{\mu}}_{i,k}^{(2)} &:= \boldsymbol{\mu}_{i,k}^{(2)} + \boldsymbol{\Sigma}_{i,k}^{(21)}(\boldsymbol{\Sigma}_{i,k}^{(11)})^{-1}\left(\mathbf{X}_{i,k}^{(1)} - \boldsymbol{\mu}_{i,k}^{(1)}\right) \in \mathbb{R}^{n-q}, \\ \tilde{\boldsymbol{\Sigma}}_{i,k}^{(22)} &:= \boldsymbol{\Sigma}_{i,k}^{(22)} - \boldsymbol{\Sigma}_{i,k}^{(21)}(\boldsymbol{\Sigma}_{i,k}^{(11)})^{-1}\boldsymbol{\Sigma}_{i,k}^{(12)}. \end{aligned}$$

For $k=J$ we obtain

$$(\log P_{i,J-i+1}, \log I_{i,J-i+1})'|_{\{\boldsymbol{\Theta},\mathcal{D}_J\}} \sim \mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad \textit{for} \quad i \in \{2,\dots,J\},$$

with

$$\boldsymbol{\mu}_i := \begin{pmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \end{pmatrix} \tilde{\boldsymbol{\mu}}_{i,J}^{(2)} \qquad \textit{and} \qquad \boldsymbol{\Sigma}_i := \begin{pmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \end{pmatrix} \tilde{\boldsymbol{\Sigma}}_{i,J}^{(22)} \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 \end{pmatrix}, \tag{5.5}$$

where $\mathbf{e}_k := \mathbf{e}_k(i) \in \mathbb{R}^{n-q}$ is the k-th canonical basis vector of dimension $n-q$. Moreover, for $i=1$ we have

$$\log P_{1,J}|_{\{\boldsymbol{\Theta},\mathcal{D}_J\}} \sim \mathcal{N}\left(\boldsymbol{\mu}_1 := \tilde{\boldsymbol{\mu}}_{1,J}^{(2)}, \boldsymbol{\Sigma}_1 := \tilde{\boldsymbol{\Sigma}}_{(1,J)}^{(22)}\right).$$

Proof: Conditionally given the parameter vector $\boldsymbol{\Theta}$, the random vectors $\mathbf{X}_i$ are independent for different accident years. Therefore, the conditional distribution of $\mathbf{X}_{i,k}^{(2)}$ depends on $\mathcal{D}_k$ only through $\mathbf{X}_{i,k}^{(1)}$. This shows the first equality in the first claim. The distributional claim is a well-known result for multivariate normal distributions using the Schur complement for the calculation of the conditional covariance matrix. The second claim is a direct consequence of the first claim. This proves the lemma. □

Remarks 5.4 (Conditional distribution)

i) The second claim in Lemma 5.3 is used to derive the distribution of the elements in the next diagonal $\mathcal{D}_{J+1}\backslash\mathcal{D}_J$. This is needed for the calculation of the full predictive distribution of the CDR via Monte-Carlo methods. For details see Section 5.5.

As a direct consequence of the first claim in Lemma 5.3 we get for the ultimate claim, $i > k-J$,

$$\log I_{i,J}|_{\{\boldsymbol{\Theta},\mathcal{D}_k\}} = \log P_{i,J}|_{\{\boldsymbol{\Theta},\mathcal{D}_k\}} \sim \mathcal{N}\left(\mathbf{e}'_{n-q}\tilde{\boldsymbol{\mu}}^{(2)}_{i,k}, \mathbf{e}'_{n-q}\tilde{\boldsymbol{\Sigma}}^{(22)}_{i,k}\mathbf{e}_{n-q}\right). \tag{5.6}$$

This immediately implies the following corollary:

Corollary 5.5 (Conditional distribution) *For the predictor of the ultimate claim $P_{i,J}$, given $\{\boldsymbol{\Theta},\mathcal{D}_J\}$, we obtain for $i > k-J$*

$$\mathrm{E}[P_{i,J}|\,\boldsymbol{\Theta},\mathcal{D}_k] = \exp\left\{\mathbf{e}'_{n-q}\tilde{\boldsymbol{\mu}}^{(2)}_{i,k} + \mathbf{e}'_{n-q}\tilde{\boldsymbol{\Sigma}}^{(22)}_{i,k}\mathbf{e}_{n-q}/2\right\}.$$

Proof: The claim is a direct consequence of Lemma 5.3 and (5.6). □

We see that the ultimate claim predictor in Corollary 5.5 still depends on $\boldsymbol{\Theta}$, namely through

$$\begin{aligned}\mathbf{e}'_{n-q}\tilde{\boldsymbol{\mu}}^{(2)}_{i,k} &= \mathbf{e}'_{n-q}\left(\boldsymbol{\mu}^{(2)}_{i,k} + \boldsymbol{\Sigma}^{(21)}_{i,k}(\boldsymbol{\Sigma}^{(11)}_{i,k})^{-1}\left(\mathbf{X}^{(1)}_{i,k} - \boldsymbol{\mu}^{(1)}_{i,k}\right)\right)\\ &= \mathbf{e}'_{n-q}\left(\mathbf{B}^{(2)}_{i,k}\boldsymbol{\Theta} + \boldsymbol{\Sigma}^{(21)}_{i,k}(\boldsymbol{\Sigma}^{(11)}_{i,k})^{-1}\left(\mathbf{X}^{(1)}_{i,k} - \mathbf{B}^{(1)}_{i,k}\boldsymbol{\Theta}\right)\right)\\ &= \boldsymbol{\Gamma}_{i,k}\boldsymbol{\Theta} + \mathbf{e}'_{n-q}\boldsymbol{\Sigma}^{(21)}_{i,k}(\boldsymbol{\Sigma}^{(11)}_{i,k})^{-1}\mathbf{X}^{(1)}_{i,k},\end{aligned} \tag{5.7}$$

where $\boldsymbol{\Gamma}_{i,k}$ is given by

$$\boldsymbol{\Gamma}_{i,k} := \mathbf{e}'_{n-q}\left(\mathbf{B}^{(2)}_{i,k} - \boldsymbol{\Sigma}^{(21)}_{i,k}(\boldsymbol{\Sigma}^{(11)}_{i,k})^{-1}\mathbf{B}^{(1)}_{i,k}\right).$$

Our aim now is to calculate the posterior distribution of $\boldsymbol{\Theta}$, conditionally given observations $\mathcal{D}_k$ for $k \in \{J, J+1\}$. The likelihood of the logarithmized observations at time k, given $\boldsymbol{\Theta}$, is given by

$$l_{\mathcal{D}_k}(\boldsymbol{\Theta}) \propto \prod_{i=0}^{J}\exp\left\{-\frac{1}{2}\left(\mathbf{X}^{(1)}_{i,k} - \mathbf{B}^{(1)}_{i,k}\boldsymbol{\Theta}\right)'(\boldsymbol{\Sigma}^{(11)}_{i,k})^{-1}\left(\mathbf{X}^{(1)}_{i,k} - \mathbf{B}^{(1)}_{i,k}\boldsymbol{\Theta}\right)\right\}. \tag{5.8}$$

With Model Assumptions 5.1 and Bayes' theorem follows that the posterior distribution $u(\boldsymbol{\Theta}|\mathcal{D}_k)$ has the form

$$u(\boldsymbol{\Theta}|\mathcal{D}_k) \propto l_{\mathcal{D}_k}(\boldsymbol{\Theta})\exp\left\{-\frac{1}{2}(\boldsymbol{\Theta}-\boldsymbol{\vartheta})'\mathbf{T}^{-1}(\boldsymbol{\Theta}-\boldsymbol{\vartheta})\right\}, \tag{5.9}$$

with prior mean

$$\boldsymbol{\vartheta} := (\phi_0;\phi_1,\psi_1,\phi_2,\psi_2,\ldots,\phi_J,\psi_J)' \in \mathbb{R}^n$$

and prior covariance matrix

$$\mathbf{T} := \operatorname{diag}(s_0^2; s_1^2, t_1^2, s_2^2, t_2^2, \ldots, s_J^2, t_J^2) \in \mathbb{R}^{n \times n}. \tag{5.10}$$

Theorem 5.6 (Posterior distribution of $\boldsymbol{\Theta}$) *Under Model Assumptions 5.1 the posterior distribution $u(\boldsymbol{\Theta}|\mathcal{D}_k)$ is a multivariate Gaussian distribution with posterior mean*

$$\boldsymbol{\vartheta}(\mathcal{D}_k) := \mathbf{T}(\mathcal{D}_k) \left[\mathbf{T}^{-1}\boldsymbol{\vartheta} + \sum_{i=0}^{J} (\mathbf{B}_{i,k}^{(1)})' (\boldsymbol{\Sigma}_{i,k}^{(11)})^{-1} \mathbf{X}_{i,k}^{(1)} \right].$$

and posterior covariance matrix

$$\mathbf{T}(\mathcal{D}_k) := \left(\mathbf{T}^{-1} + \sum_{i=0}^{J} (\mathbf{B}_{i,k}^{(1)})' (\boldsymbol{\Sigma}_{i,k}^{(11)})^{-1} \mathbf{B}_{i,k}^{(1)} \right)^{-1},$$

Proof: From (5.8) immediately follows that the posterior distribution $u(\boldsymbol{\Theta}|\mathcal{D}_k)$ is a multivariate Gaussian distribution. Therefore, it remains to calculate the first two moments of $u(\boldsymbol{\Theta}|\mathcal{D}_k)$. This is done by squaring out all terms and analyzing quadratic and linear terms. □

From (5.7) we see that the exponent of the predictor given in Corollary 5.5 is a affine-linear function of $\boldsymbol{\Theta}$. Using Theorem 5.6 this implies the following corollary:

Corollary 5.7 (Ultimate claim predictor) *The predictor for the ultimate claim for accident year $i > k - J$ and $k \in \{J, J+1\}$, given $\mathcal{D}_k$, is given by*

$$\mathrm{E}[P_{i,J}|\, \mathcal{D}_k] = \exp\left\{ \boldsymbol{\Gamma}_{i,k}\boldsymbol{\vartheta}(\mathcal{D}_k) + \boldsymbol{\Gamma}_{i,k}\mathbf{T}(\mathcal{D}_k)(\boldsymbol{\Gamma}_{i,k})'/2 + \mathbf{e}_{n-q}' \boldsymbol{\Sigma}_{i,k}^{(21)} (\boldsymbol{\Sigma}_{i,k}^{(11)})^{-1} \mathbf{X}_{i,k}^{(1)} + \mathbf{e}_{n-q}' \tilde{\boldsymbol{\Sigma}}_{i,k}^{(22)} \mathbf{e}_{n-q}/2 \right\}.$$

Proof: The proof is a direct consequence of Corollary 5.5 and Theorem 5.6. □

Remarks 5.8 (Ultimate claim predictor)

i) *For $k = J$ and diagonal covariance matrix $\mathbf{V}$ we obtain the same ultimate claim predictor as in* Merz–Wüthrich *[46].*

ii) *For $k = J + 1$ we get a closed formula for the ultimate claim predictor in the case that information $\mathcal{D}_{J+1}$ is available at time $J + 1$. This allows for the simulation of the full predictive distribution of the CDR. This is done in detail in Section 5.5.*

iii) *For other choices of prior distributions MCMC methods can be applied to calculate the posterior distribution in Theorem 5.6. For details see* Merz–Wüthrich *[46].*

5.4 Mean Squared Error of Prediction of the Claims Development Result

5.4.1 Single Accident Years

In the last section we have calculated the expected ultimate claim in the PIC reserving model, given the observations $\mathcal{D}_k$ for $k \in \{J, J+1\}$. Our aim now is to calculate the prediction uncertainty of the CDR in terms of the conditional MSEP. From (5.1) we see that the problem to derive the conditional MSEP for the one-year CDR is solved by calculating $\mathrm{Var}[\mathrm{E}[P_{i,J}|\mathcal{D}_{J+1}]|\mathcal{D}_J]$. Since $(\mathrm{E}[P_{i,J}|\mathcal{D}_J])^2$ is given by Corollary 5.7 for $k = J$, this conditional variance can be derived by calculating $\mathrm{E}\Big[(\mathrm{E}[P_{i,J}|\mathcal{D}_{J+1}])^2\Big|\mathcal{D}_J\Big]$. We see that for $k = J+1$ the exponential term from Corollary 5.7, namely,

$$\boldsymbol{\Gamma}_{i,J+1}\boldsymbol{\vartheta}(\mathcal{D}_{J+1}) + \boldsymbol{\Gamma}_{i,J+1}\mathbf{T}(\mathcal{D}_{J+1})(\boldsymbol{\Gamma}_{i,J+1})'/2 \\ + \mathbf{e}'_{n-q}\boldsymbol{\Sigma}^{(21)}_{i,J+1}\left(\boldsymbol{\Sigma}^{(11)}_{i,J+1}\right)^{-1}\mathbf{X}^{(1)}_{i,J+1} + \mathbf{e}'_{n-q}\tilde{\boldsymbol{\Sigma}}^{(22)}_{i,J+1}\mathbf{e}_{n-q}/2,$$

is affine-linear in the observations $\mathcal{D}_{J+1}\backslash\mathcal{D}_J$ given by

$$\mathbf{Y} := (\log P_{1,J}, \log P_{2,J-1}, \log I_{2,J-1}, \ldots, \log P_{J,1}, \log I_{J,1})'.$$

That means that for all $i > 1$ there exist a matrix $\mathbf{L}_i$ and a $\mathcal{D}_J$-measurable random variable $g_i(\mathcal{D}_J)$ such that

$$\mathbf{L}_i\mathbf{Y} + g_i(\mathcal{D}_J) = \boldsymbol{\Gamma}_{i,J+1}\boldsymbol{\vartheta}(\mathcal{D}_{J+1}) + \boldsymbol{\Gamma}_{i,J+1}\mathbf{T}(\mathcal{D}_{J+1})(\boldsymbol{\Gamma}_{i,J+1})'/2 \\ + \mathbf{e}'_{n-q}\boldsymbol{\Sigma}^{(21)}_{i,J+1}\left(\boldsymbol{\Sigma}^{(11)}_{i,J+1}\right)^{-1}\mathbf{X}^{(1)}_{i,J+1} + \mathbf{e}'_{n-q}\tilde{\boldsymbol{\Sigma}}^{(22)}_{i,J+1}\mathbf{e}_{n-q}/2.$$

For $i = 1$ we set $\mathbf{L}_1$ to be the projection on the first component, i.e. $\mathbf{L}_1\mathbf{Y} := \log P_{1,J}$ and $g_1(\mathcal{D}_J) = 0$. This implies for the ultimate claim predictor in Corollary 5.7

$$\mathrm{E}[P_{i,J}|\mathcal{D}_{J+1}] = \exp\{\mathbf{L}_i\mathbf{Y} + g_i(\mathcal{D}_J)\} \quad \text{for } i = 1, \ldots, J. \tag{5.11}$$

Different accident years are independent, given $\boldsymbol{\Theta}$. Thus, Lemma 5.3 leads to the joint distribution of $\mathbf{Y}$, given $\{\mathcal{D}_J, \boldsymbol{\Theta}\}$:

Lemma 5.9 (Conditional distribution of Y) *Under Model Assumptions 5.1 we have*

$$\mathbf{Y}|_{\{\mathcal{D}_J,\boldsymbol{\Theta}\}} = (\log P_{1,J}, \log P_{2,J-1}, \log I_{2,J-1}, \ldots, \log P_{J,1}, \log I_{J,1})'|_{\{\mathcal{D}_J,\boldsymbol{\Theta}\}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\mu} := \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_J \end{pmatrix} \in \mathbb{R}^{2J-1} \quad \textit{and} \quad \boldsymbol{\Sigma} := \begin{pmatrix} \boldsymbol{\Sigma}_1 & 0 & 0 & \cdots & 0 \\ 0 & \boldsymbol{\Sigma}_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \boldsymbol{\Sigma}_J \end{pmatrix} \in \mathbb{R}^{(2J-1)\times(2J-1)},$$

with $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ defined in Lemma 5.3.

In Lemma 5.9 the distribution of $\mathbf{Y}|_{\{\mathcal{D}_J,\boldsymbol{\Theta}\}}$ still depends on $\boldsymbol{\Theta}$ via

$$\boldsymbol{\mu} = \left(\boldsymbol{\mu}_1', \boldsymbol{\mu}_2', \ldots, \boldsymbol{\mu}_J'\right)' \in \mathbb{R}^{2J-1}$$

and recalling the definition of $\boldsymbol{\mu}_i$ (see Lemma 5.3) we obtain, for $k = J$,

$$\boldsymbol{\gamma}_i := \begin{cases} \begin{pmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \end{pmatrix} \boldsymbol{\Sigma}_{i,J}^{(21)} (\boldsymbol{\Sigma}_{i,J}^{(11)})^{-1} \mathbf{X}_{i,J}^{(1)} & \text{for} \quad i \geq 2 \\ \mathbf{e}_{n-q}' \boldsymbol{\Sigma}_{1,J}^{21} (\boldsymbol{\Sigma}_{1,J}^{(11)})^{-1} \mathbf{X}_{1,J}^{(1)} & \text{for} \quad i = 1 \end{cases}$$

$$\boldsymbol{\mu}_i = \begin{cases} \tilde{\boldsymbol{\Gamma}}_{i,J} \boldsymbol{\Theta} + \boldsymbol{\gamma}_i & \text{for} \quad i \geq 2 \\ \tilde{\boldsymbol{\Gamma}}_{1,J} \boldsymbol{\Theta} + \boldsymbol{\gamma}_1 & \text{for} \quad i = 1 \end{cases}$$

and

$$\tilde{\boldsymbol{\Gamma}}_{i,J} := \begin{cases} \begin{pmatrix} \mathbf{e}_1' \\ \mathbf{e}_2' \end{pmatrix} \left(\mathbf{B}_{i,J}^{(2)} - \boldsymbol{\Sigma}_{i,J}^{(21)} (\boldsymbol{\Sigma}_{i,J}^{(11)})^{-1} \mathbf{B}_{i,J}^{(1)} \right) & \text{for} \quad i \geq 2 \\ \mathbf{e}_{n-q}' \left(\mathbf{B}_{1,J}^{(2)} - \boldsymbol{\Sigma}_{1,J}^{(21)} (\boldsymbol{\Sigma}_{1,J}^{(11)})^{-1} \mathbf{B}_{1,J}^{(1)} \right) & \text{for} \quad i = 1 \end{cases}.$$

Next, we define the matrix $\boldsymbol{\Gamma}$ with rows $\tilde{\boldsymbol{\Gamma}}_{i,J}$, i.e.

$$\boldsymbol{\Gamma} := \begin{pmatrix} \tilde{\boldsymbol{\Gamma}}_{1,J}' & \tilde{\boldsymbol{\Gamma}}_{2,J}' & \ldots & \tilde{\boldsymbol{\Gamma}}_{J,J}' \end{pmatrix}' \in \mathbb{R}^{(2J-1)\times n}.$$

and

$$\boldsymbol{\gamma} := (\boldsymbol{\gamma}_1', \ldots, \boldsymbol{\gamma}_J')' \in \mathbb{R}^{(2J-1)}.$$

This shows that $\boldsymbol{\mu} = \boldsymbol{\Gamma}\boldsymbol{\Theta}' + \boldsymbol{\gamma}$ is a affine-linear function of $\boldsymbol{\Theta}$. This implies together with (5.11) the following theorem.

Theorem 5.10 (Conditional expectation) *Under Model Assumptions 5.1 we obtain for* $i, l \in \{1, \ldots, J\}$

$$\begin{aligned} &\mathrm{E}[\mathrm{E}[P_{i,J}|\,\mathcal{D}_{J+1}]\,\mathrm{E}[P_{l,J}|\,\mathcal{D}_{J+1}]|\,\boldsymbol{\Theta},\mathcal{D}_J] \\ &\qquad = \exp\left\{(\mathbf{L}_i+\mathbf{L}_l)\boldsymbol{\mu}+(\mathbf{L}_i+\mathbf{L}_l)\boldsymbol{\Sigma}(\mathbf{L}_i+\mathbf{L}_l)'/2+g_i(\mathcal{D}_J)+g_l(\mathcal{D}_J)\right\}, \end{aligned}$$

and

$$\mathrm{E}[\mathrm{E}[P_{i,J}|\,\mathcal{D}_{J+1}]\,\mathrm{E}[P_{l,J}|\,\mathcal{D}_{J+1}]|\,\mathcal{D}_J] = \mathrm{E}[P_{i,J}|\,\mathcal{D}_J]\,\mathrm{E}[P_{l,J}|\,\mathcal{D}_J]\exp\{\mathbf{L}_i\boldsymbol{\Gamma}\mathbf{T}(\mathcal{D}_J)\boldsymbol{\Gamma}'\mathbf{L}_l' + \mathbf{L}_i\boldsymbol{\Sigma}\mathbf{L}_l'\}.$$

Proof: Using standard properties of log-normal distribution, the first claim immediately follows by Lemma 5.9 and (5.11). The second claim follows with the identity $\boldsymbol{\mu} = \boldsymbol{\Gamma}\boldsymbol{\Theta}' + \boldsymbol{\gamma}$ and Theorem 5.6. □

By means of this relationship between $\mathrm{E}\Big[(\mathrm{E}[P_{i,J}|\,\mathcal{D}_{J+1}])^2\Big|\,\mathcal{D}_J\Big]$ and $\mathrm{E}[P_{i,J}|\,\mathcal{D}_J]^2$ it is straightforward to derive the (conditional) MSEP of the one-year CDR for single accident years, which is given in the next theorem:

Theorem 5.11 (Conditional MSEP for single accident years) *Under Model Assumptions 5.1 the conditional MSEP, given $\mathcal{D}_J$, of the one-year CDR for single accident years $i \in \{1, \ldots, J\}$ is given by*

$$\mathrm{msep}_{\mathrm{CDR}_i^{J+1}|\mathcal{D}_J}[0] = \left(\mathrm{E}[P_{i,J}|\,\mathcal{D}_J]\right)^2 \left(\exp\{\mathbf{L}_i \boldsymbol{\Gamma} \mathbf{T}(\mathcal{D}_J)\boldsymbol{\Gamma}'\mathbf{L}_i' + \mathbf{L}_i \boldsymbol{\Sigma} \mathbf{L}_i'\} - 1\right).$$

In the following section we consider the conditional MSEP for aggregated accident years.

5.4.2 Aggregated Accident Years

We study the conditional MSEP of the one-year CDR for aggregated accident years:

$$\begin{aligned}\mathrm{msep}_{\sum_{i=1}^J \mathrm{CDR}_i^{J+1}\Big|\mathcal{D}_J}[0] &= \mathrm{E}\left[\left(\sum_{i=1}^J \mathrm{CDR}_i^{J+1} - 0\right)^2 \Bigg|\, \mathcal{D}_J\right] \\ &= \mathrm{Var}\left(\sum_{i=1}^J \mathrm{CDR}_i^{J+1}\Bigg|\, \mathcal{D}_J\right) = \mathrm{Var}\left(\sum_{i=1}^J \mathrm{E}\left[P_{i,J}|\mathcal{D}_{J+1}\right]\Bigg|\, \mathcal{D}_J\right).\end{aligned} \tag{5.12}$$

Using the tower property of conditional expectations and Theorem 5.10 we obtain for (5.12):

Theorem 5.12 (Conditional MSEP for aggregated accident years) *Under Model Assumptions 5.1 the conditional MSEP, given $\mathcal{D}_J$, of the one-year CDR for aggregated accident years is given by*

$$\begin{aligned}\mathrm{msep}_{\sum_{i=1}^J \mathrm{CDR}_i^{J+1}\Big|\mathcal{D}_J}[0] =& \sum_{i=1}^J \mathrm{msep}_{\mathrm{CDR}_i^{J+1}|\mathcal{D}_J}[0] \\ &+ 2\sum_{l>i} \mathrm{E}[P_{i,J}|\,\mathcal{D}_J]\,\mathrm{E}[P_{l,J}|\,\mathcal{D}_J]\left(\exp\left\{\mathbf{L}_i \boldsymbol{\Gamma} \mathbf{T}(\mathcal{D}_J)\boldsymbol{\Gamma}'\mathbf{L}_l' + \mathbf{L}_i \boldsymbol{\Sigma} \mathbf{L}_l'\right\} - 1\right).\end{aligned}$$

Proof: With (5.12) and Theorem 5.10 we obtain

$$\begin{aligned}\mathrm{msep}_{\sum_{i=1}^J \mathrm{CDR}_i^{J+1}\Big|\mathcal{D}_J}[0] =& \sum_{i=1}^J \mathrm{Var}\left[\mathrm{E}[P_{i,J}|\mathcal{D}_{J+1}]|\mathcal{D}_J\right] + 2\sum_{l>i}^J \mathrm{E}\left[\mathrm{E}[P_{i,J}|\,\mathcal{D}_{J+1}]\,\mathrm{E}[P_{l,J}|\,\mathcal{D}_{J+1}]|\,\mathcal{D}_J\right] \\ &- 2\sum_{l>i}^J \mathrm{E}[P_{i,J}|\,\mathcal{D}_J]\;\mathrm{E}[P_{l,J}|\,\mathcal{D}_J] \\ =& \sum_{i=1}^J \mathrm{msep}_{\mathrm{CDR}_i^{J+1}|\mathcal{D}_J}[0] + 2\sum_{l>i}^J \mathrm{E}[P_{i,J}|\,\mathcal{D}_J]\,\mathrm{E}[P_{l,J}|\,\mathcal{D}_J]\left(\exp\left\{\mathbf{L}_i \boldsymbol{\Gamma} \mathbf{T}(\mathcal{D}_J)\boldsymbol{\Gamma}'\mathbf{L}_l' + \mathbf{L}_i \boldsymbol{\Sigma} \mathbf{L}_l'\right\} - 1\right).\end{aligned}$$

□

5.5 Example PIC Reserving Method

We revisit the data given in Dahms [16]. In Model Assumptions 5.1 we can choose any covariance matrices $\mathbf{V}$ as long as it is positive definite. This allows for modeling dependence between paid and incurred data. The task of structuring a suitable covariance matrix $\mathbf{V}$ based on expert opinion and data is discussed in detail in Happ-Wüthrich [31], see Chapter also 6. In this example we choose $\mathbf{V}$ as a diagonal matrix and estimate the variances on the diagonal with standard sample estimates. This is in-line with the choice of $\mathbf{V}$ in Merz–Wüthrich [46]. More detailed, for the estimation of $\mathbf{V}$ we use for $j \in \{0, \ldots, J-1\}$ and $k \in \{1, \ldots, J-1\}$

$$\widehat{\Phi}_j := \frac{1}{J-j+1} \sum_{i=0}^{J-j} \xi_{i,j}, \qquad \widehat{\Psi}_k := \frac{1}{J-k+1} \sum_{i=0}^{J-k} \zeta_{i,k},$$

$$\widehat{\sigma}^2_{\xi_j} := \frac{1}{J-j} \sum_{i=0}^{J-j} \left(\xi_{i,j} - \widehat{\Phi}_j\right)^2 \qquad \text{and} \qquad \widehat{\sigma}^2_{\zeta_k} := \frac{1}{J-k} \sum_{i=0}^{J-k} \left(\zeta_{i,k} - \widehat{\Psi}_k\right)^2.$$

Since we have for the estimation of the two parameters $\sigma^2_{\xi_J}$ and $\sigma^2_{\zeta_J}$ only one observation we use the extrapolation formula, see Wüthrich–Merz [63],

$$\widehat{\sigma}^2_{\xi_J} := \min\{\hat{\sigma}^2_{\xi_{J-2}}, \hat{\sigma}^2_{\xi_{J-1}}, \hat{\sigma}^4_{\xi_{J-2}}/\hat{\sigma}^2_{\xi_{J-1}}\} \quad \text{and} \quad \widehat{\sigma}^2_{\zeta_j} := \min\{\hat{\sigma}^2_{\zeta_{J-2}}, \hat{\sigma}^2_{\zeta_{J-1}}, \hat{\sigma}^4_{\zeta_{J-2}}/\hat{\sigma}^2_{\zeta_{J-1}}\}$$

and set

$$\mathbf{V} := \operatorname{diag}\left(\widehat{\sigma}^2_{\xi_0}, \widehat{\sigma}^2_{\xi_1}, \widehat{\sigma}^2_{\zeta_1}, \ldots, \widehat{\sigma}^2_{\xi_J}, \widehat{\sigma}^2_{\zeta_J}\right).$$

Because we do not have any prior knowledge of the prior distribution parameters ϕ_l and ψ_j we choose non-informative priors, i.e. we let $s_j^2 \to \infty$ and $t_l^2 \to \infty$. This implies that in Theorem 5.6 the matrix $\mathbf{T}^{-1}$, see (5.10), is the matrix consisting of zeros and no prior information is used in our calculations. In Table 5.1 we compare the prediction uncertainty measured by the square root of the conditional MSEP for the one-year CDR calculated by the PIC method and the ECLR method (cf. Dahms [16]). Under Model Assumptions 5.1, these values for the PIC method are calculated analytically with Theorem 5.11 for single accident years and with Theorem 5.12 for aggregated accident years. Note that for the ECLR method we obtain two different values for the (conditional) MSEP because we can estimate the variance in two ways, namely based on paid data or based on incurred data, respectively. We observe in the PIC method for most single accident years and aggregated accident years a lower prediction uncertainty for the CDR than in the ECLR method based on paid or incurred data (see Table 5.1). This can partly be explained by the fact that in the ECLR method we have to estimate 44 parameters (cf. Dahms [16]) whereas in the Bayesian PIC model only 19 variance parameters have to be estimated leading to a lower standard error. Moreover, we observe that in the PIC method it is not unlikely that the total claims reserves increase about 3% in the one-year horizon. This is similar to the findings for the CDR uncertainty for the ECLR method in Dahms et al. [18].

accident year i	claims reserves ECLR	$\text{msep}_{\text{CDR}}^{1/2}$ ECLR method Paid	$\text{msep}_{\text{CDR}}^{1/2}$ ECLR method Incurred	claims reserves PIC	$\text{msep}_{\text{CDR}}^{1/2}$ PIC method	in % reserves PIC
1	314.902	194	14.639	337.799	2.637	0,78%
2	66.994	4.557	4.678	31.686	4.597	14,51%
3	359.384	5.597	6.628	331.890	7.656	2,31%
4	981.883	33.675	34.258	1.018.308	6.606	0,65%
5	1.115.768	30.574	30.997	1.104.816	31.594	2,86%
6	1.786.947	42.598	43.074	1.842.669	43.168	2,34%
7	1.942.518	166.154	166.255	1.953.767	139.352	7,13%
8	1.569.657	138.685	138.740	1.602.229	127.053	7,93%
9	2.590.718	210.899	210.979	2.402.946	173.721	7,23%
Total	10.728.771	346.576	350.534	10.626.108	292.879	2,76%

Table 5.1: Ultimate claim prediction and prediction uncertainty for the one-year CDR calculated by the ECLR method for claims payments and incurred losses (cf. Dahms [16] and Dahms et al. [18]) and by the PIC method, respectively

Table 5.2 provides the ratios of the square root of the conditional MSEP for the one-year CDR and the square root of the conditional MSEP for the ultimate claim. We observe that for later accident years (i.e. $i \geq 7$) and aggregated accident years the values for the ECLR method and for the PIC method only slightly differ. Moreover, we see that for aggregated accident years the one-year uncertainty is about 75% of the uncertainty of the ultimate claim prediction. This result is in line with the field study conducted by AISAM–ACME [2].

accident year	$\text{msep}_{\text{CDR}}^{1/2}/\text{msep}_{\text{Ultimate}}^{1/2}$ ECLR method Incurred	$\text{msep}_{\text{CDR}}^{1/2}/\text{msep}_{\text{Ultimate}}^{1/2}$ ECLR method Paid	$\text{msep}_{\text{CDR}}^{1/2}/\text{msep}_{\text{Ultimate}}^{1/2}$ PIC method Paid & Incurred
1	100.0%	100.0%	100.0%
2	100.0%	84.5%	87.6%
3	53.1%	52.7%	83.7%
4	91.5%	89.6%	62.4%
5	69.6%	69.1%	94.3%
6	65.5%	65.4%	80.8%
7	94.0%	93.9%	93.1%
8	70.1%	70.1%	70.3%
9	65.3%	65.3%	66.4%
Total	74.1%	74.3%	75.2%

Table 5.2: Ratios $\text{msep}_{\text{CDR}}^{1/2}/\text{msep}_{\text{Ultimate}}^{1/2}$ calculated by the ECLR method for claims payments and incurred losses (cf. Dahms et al. [18]) and calculated by the PIC method, respectively

As already mentioned in the PIC method we can not only calculate the conditional MSEP for the one-year CDR but also the full predictive distribution of the one-year CDR by means of MC simulations. Firstly, we apply Theorem 5.6 to $u(\boldsymbol{\Theta}|\mathcal{D}_J)$ to generate Gaussian samples $\boldsymbol{\Theta}^{(n)}$ with

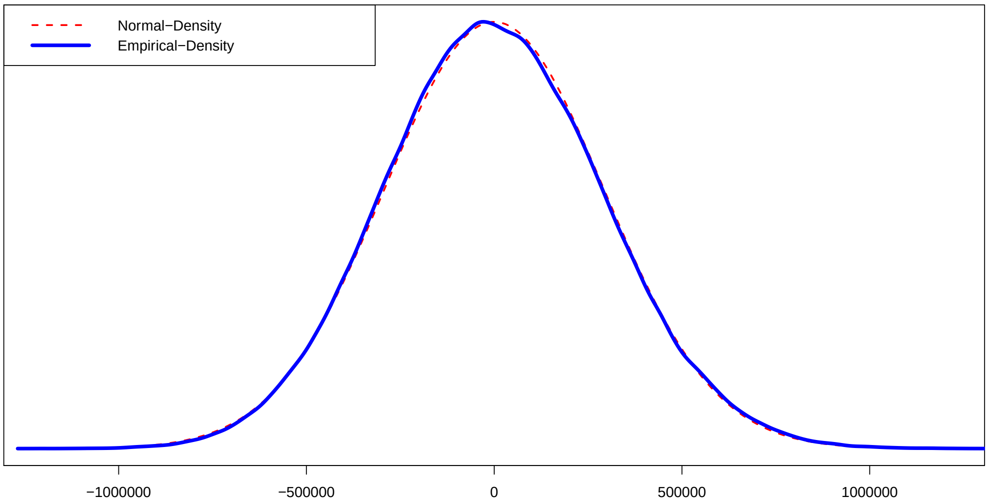

Figure 5.3: Empirical density for the one-year CDR (blue line) from 100.000 simulations and fitted Gaussian density with mean 0 and standard deviation 292.879 (dotted red line)

mean $\boldsymbol{\vartheta}(\mathcal{D}_J)$ and covariance matrix $\mathbf{T}(\mathcal{D}_J)$. Secondly, we generate independent two-dimensional Gaussian samples $(\log P_{i,J-i+1}, \log I_{i,J-i+1})_{\{\mathcal{D}_J,\boldsymbol{\Theta}\}}$ and fill up the off-diagonal entries in the paid and incurred trapezoids (see Lemma 5.3). This way we obtain the data available at time $J+1$, i.e. $\mathcal{D}_{J+1}$, and can calculate $\mathrm{E}[P_{i,J}|\mathcal{D}_{J+1}]$ by means of Corollary 5.7. This provides Figure 5.3, where we compare the empirical density from 100.000 simulations (blue line) to the Gaussian density with mean $\mu = 0$ and standard deviation $\sigma = 292.879$ (dotted red line), see Table 5.1. We observe that these two densities look quite similar. To get a closer look on the left tail of the empirical density for the one-year CDR we show a QQ-plot for quantiles $q \in (0, 0.1)$. We observe that the tail behaviour of the empirical density of the one-year CDR and the fitted Gaussian density with mean 0 and standard deviation 292.879 only slightly differ (see Figure 5.4). This is similar to the findings for the distribution of the ultimate claim in Merz–Wüthrich [46]. This means that using a Gaussian approximation for the density of the one-year CDR provides within the PIC method and for the given data a good approximation for the shortfall risk of the one-year CDR.

5.6 Conclusions

The PIC reserving method provides a framework, where unified ultimate claim predictions can be calculated based on cumulative payments and incurred losses data simultaneously. It allows for the derivation of the (conditional) MSEP for the ultimate claim in the long run as well as for

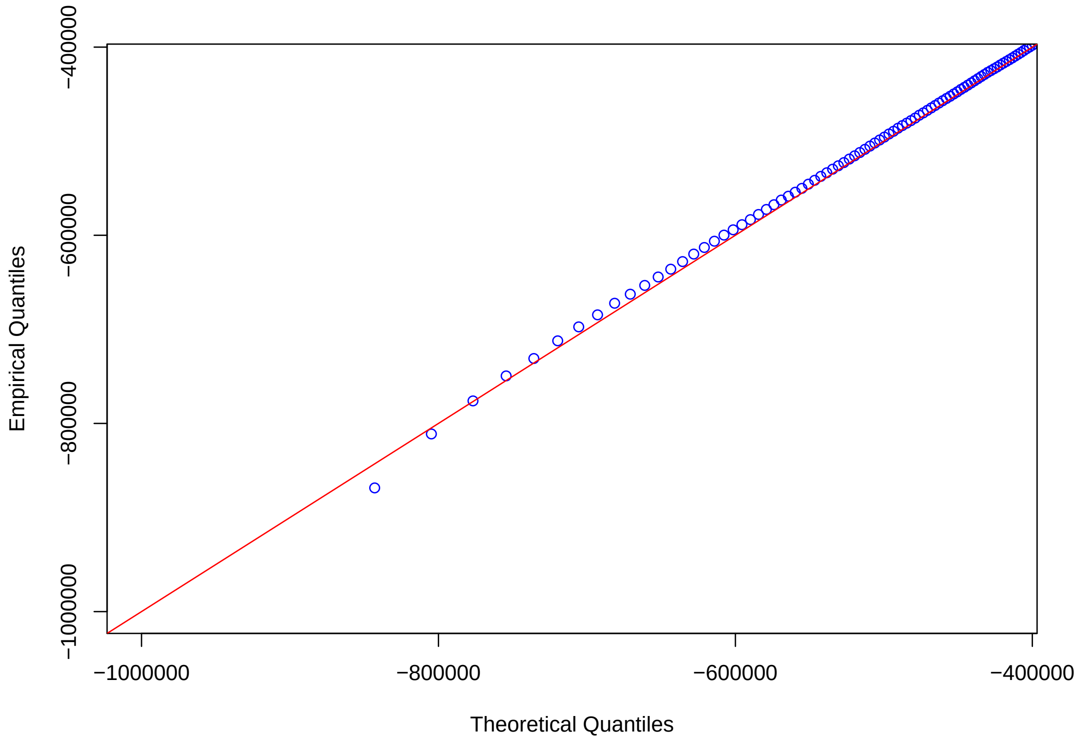

Figure 5.4: QQ-plot for lower quantiles $q \in (0, 0.1)$ to compare the left tail of the empirical density for the one-year CDR with the left tail of the fitted Gaussian density with mean 0 and standard deviation 292.879

the CDR in the one-year time horizon. Merz–Wüthrich [46] derived the MSEP formula for the ultimate claim uncertainty. In this chapter we did the same for the one-year CDR uncertainty. In contrast to the ECLR method by Dahms [16], where also MSEP formulas for the ultimate claim and the CDR uncertainty exist, the PIC method allows for the calculation of the full predictive distribution of the ultimate claim and the CDR via Monte-Carlo simulations. This implies that any other risk measure for example VaR or ES can be calculated for the ultimate claim uncertainty (long term risk) as well as for the CDR uncertainty (one-year risk).

Main results of the PIC reserving method:

For solvency considerations in Chapter 7 we summarize all quantities of interest derived in the PIC reserving method:

1. The predictor $\widehat{\mathcal{R}}^I$ for outstanding loss liabilities $\mathcal{R}^I$, see (2.5c) and (2.4c) given by

$$\widehat{\mathcal{R}}^I = \sum_{i=1}^{J} \left(\mathrm{E}[\, P_{i,J} | \, \mathcal{D}_J] - P_{i,J-i} \right), \tag{5.13}$$

see Corollary 5.7 and

$$\widehat{S}_{i,k}^{0|I,Bayes} := \mathrm{E}[\, P_{i,k} |\, \mathcal{D}_J] - \mathrm{E}[\, P_{i,k-1} |\, \mathcal{D}_J]$$

2. The estimator for the prediction uncertainty in terms of the (conditional) MSEP

$$\mathrm{msep}_{\mathcal{R}^I | \mathcal{D}_J} \left[\widehat{\mathcal{R}}^I \right], \tag{5.14}$$

given by Theorem 4.1 in Merz–Wüthrich [46].

3. The estimator for the CDR uncertainty in terms of the (conditional) MSEP

$$\mathrm{msep}_{\sum_{i=1}^{J} \mathrm{CDR}_i^{J+1} \middle| \mathcal{D}_J} [0] \,, \tag{5.15}$$

given by Theorem 5.12.

6 Paid-Incurred Chain Reserving Method with Dependence Modeling

As mentioned in the previous section, the classical PIC reserving method introduced in Merz–Wüthrich [46] is one of the first claims reserving methods which can cope with three sources of information: (i) claims payments for reported claims; (ii) incurred losses which correspond to the reported claim amounts; (iii) prior expert opinion which can be used to design the prior covariance matrix $\mathbf{V}$ and prior means. The initial version of the PIC reserving method assumes $\mathbf{V}$ to be diagonal and hence does not allow for dependence modeling between claims payments and incurred losses data. We revisit the problem of the classical PIC reserving method and generalize it to allow for appropriate dependence modeling. In this section we follow Happ–Wüthrich [31].

6.1 Notation and Model Assumptions

For the PIC model we consider three channels of information: (i) claims payments, which refer to the payments done for reported claims; (ii) incurred losses, which correspond to the reported claim amounts; (iii) prior expert opinion. As already mentioned in Chapter 5 the crucial observation is that the claims payments and incurred losses time series must reach the same ultimate value, because these two time series both converge to the total ultimate claim. By choosing appropriate model assumptions we force this property to hold true in our model. In the same way as in Chapter 5 we denote accident years by $i \in \{0, \dots, J\}$ and development years by $j \in \{0, \dots, J\}$. We assume that all claims are settled after the J-th development year. Cumulative claims payments in accident year i after j development periods are denoted by $P_{i,j}$ and the corresponding incurred losses by $I_{i,j}$. Moreover, for the ultimate claim we assume (force) $P_{i,J} = I_{i,J}$ with probability 1, which means that ultimately (at time J) they reach the same ultimate value. For an illustration we refer to Table 6.1. The PIC model with dependence is defined as follows:

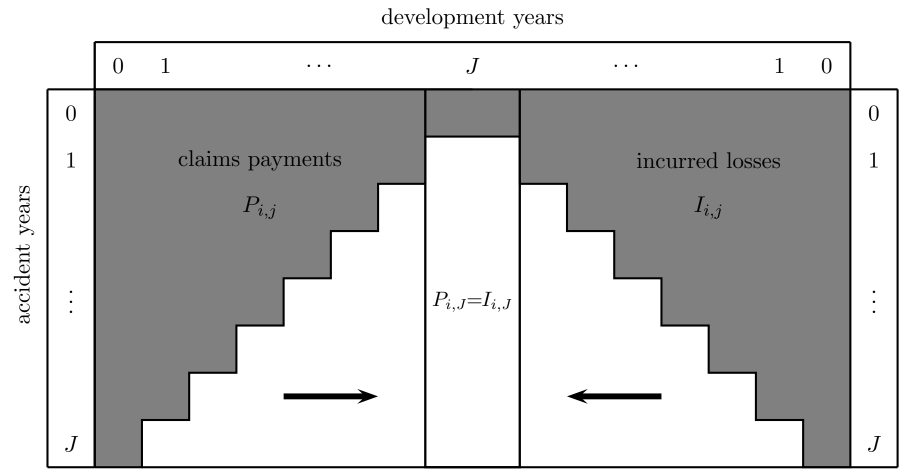

Table 6.1: Left-hand side: development triangle with cumulative claims payments $P_{i,j}$; right-hand side: development triangle with incurred losses $I_{i,j}$; both leading to the same ultimate claim $P_{i,J} = I_{i,J}$

Model Assumptions 6.1 (PIC model with dependence)

a) Conditionally, given the parameter vector $\mathbf{\Theta} := (\Psi_0; \Psi_1, \Phi_1, \Psi_2, \Phi_2, \dots, \Psi_J, \Phi_J)'$, *we assume:*

- *the random vectors* $\mathbf{\Xi}_i := (\zeta_{i,0}; \zeta_{i,1}, \xi_{i,1}, \zeta_{i,2}, \xi_{i,2}, \dots, \zeta_{i,J}, \xi_{i,J})' \in \mathbb{R}^{2J+1}$ *are i.i.d. with multivariate Gaussian distribution*

$$\mathbf{\Xi}_i \sim \mathcal{N}(\mathbf{\Theta}, \mathbf{V}) \qquad \text{for } i \in \{0, \dots, J\};$$

and positive definite covariance matrix $\mathbf{V} \in \mathbb{R}^{(2J+1)\times(2J+1)}$ *as well as individual development factors*

$$\zeta_{i,j} := \log \frac{I_{i,j}}{I_{i,j-1}} \qquad \text{and} \qquad \xi_{i,l} := \log \frac{P_{i,l}}{P_{i,l-1}}, \tag{6.1}$$

for $j \in \{0, \dots, J\}$ *and* $l \in \{1, \dots, J\}$, *where we have set* $I_{i,-1} := 1$;

- $P_{i,J} = I_{i,J}$, $\mathbb{P}$*-a.s., for all* $i \in \{0, \dots, J\}$.

b) The components of $\mathbf{\Theta}$ *are independent with prior distributions*

$$\Psi_j \sim \mathcal{N}\left(\psi_j, t_j^2\right) \text{ for } j \in \{0, \dots, J\} \qquad \text{and} \qquad \Phi_l \sim \mathcal{N}\left(\phi_l, s_l^2\right) \text{ for } l \in \{1, \dots, J\}$$

with prior parameters ψ_j, $\phi_l \in \mathbb{R}$ *and* $t_j^2 > 0$, $s_l^2 > 0$.

□

The only difference between Model Assumptions 5.1 of the PIC model and Model Assumptions 6.1 of the PIC model with dependence is that $I_{i,j}$ and $P_{i,j}$ have changed roles and $I_{i,j}$ are used as priors for $P_{i,j}$, see Remark 6.2 ii) for details.

Remarks 6.2 (PIC model with dependence)

i) *For* $\mathbf{V} = \mathrm{diag}(\tau_0^2; \tau_1^2, \sigma_1^2, \ldots, \tau_J^2, \sigma_J^2)$ *we obtain the PIC reserving model from* Merz–Wüthrich *[46]. In the following we allow for general covariance matrices* $\mathbf{V}$ *(as long as they are positive definite). In* (6.3) *below, we give an explicit choice that will be applied to a motor third party liability portfolio.*

ii) *The PIC model combines both cumulative payments and incurred losses data to get a unified predictor for the total ultimate claim that is based on both sources of information. Thereby, the model assumption* $P_{i,J} = I_{i,J}$ *guarantees that the total ultimate claim coincides for claims payments and incurred losses data. In particular, we obtain by* (6.1) *the identities*

$$I_{i,j} = I_{i,j-1} \ \exp\{\zeta_{i,j}\}, \qquad \textit{with initial value } I_{i,0} = \exp\{\zeta_{i,0}\},$$

and by backwards recursion

$$P_{i,j-1} = P_{i,j} \ \exp\{-\xi_{i,j}\}, \qquad \textit{with initial value } P_{i,J} = I_{i,J}. \tag{6.2}$$

Note that in comparison to Merz–Wüthrich *[46] we have exchanged the role of* $I_{i,j}$ *and* $P_{i,j}$*. In the original model of* Merz–Wüthrich *[46] the resulting claims reserves are completely symmetric in the exchange of* $I_{i,j}$ *and* $P_{i,j}$*. If we consider the model with dependence, as in Model Assumptions 6.1 above, it is more natural to use incurred losses* $I_{i,J}$ *as prior for claims payments* $P_{i,j}$*. This means that Hertig's log-normal model [32] for* $I_{i,j}$ *plays the role of the prior for Gogol's claims reserving model [28] for* $P_{i,j}$*, see also* Merz–Wüthrich *[46].*

iii) *If we have prior (expert) knowledge (as a third information channel) this can be used to design the prior distribution of* $\mathbf{\Theta}$*. If there is no prior knowledge we choose non-informative priors for* $\mathbf{\Theta}$*, that is we let* $t_j^2 \to \infty$ *and* $s_l^2 \to \infty$ *for* $j \in \{0, \ldots, J\}$ *and* $l \in \{1, \ldots, J\}$*.*

iv) *The assumption* $P_{i,J} = I_{i,J}$ *means that all claims are assumed to be settled after* J *development years and there is no so-called tail development factor. If there is a claims development beyond development year* J*, then one can extend the PIC model for the estimation of a tail development factor, see* Merz–Wüthrich *[42] for more details.*

v) *Under Model Assumptions 6.1 the distribution of the ultimate claims* $I_{i,J}$ *are a priori equal across accident years. However, given the observed data, we observe different posterior distributions for claims of different accident years. Therefore, the PIC reserving method allows for accident year variation (see Corollary 6.6). However, if knowledge of prior differences is available it should be incorporated in the prior means. This relaxation of the model assumption will still lead to closed form solutions. A similar effect can be achieved by considering (volume-) adjusted observations.*

vi) Conditional i.i.d. guarantees that we obtain a model of CL type, see (6.2), where CL factors do not depend on accident year i. Of course, this model assumption requires that the data considered need to be sufficiently regular. If this is not the case, one can introduce prior differences between accident years (see also last bullet point). These more general assumptions still lead to a closed form solution. The drawback is that the model might become over-parametrized and/or it requires extended expert knowledge.

vii) The covariance matrix $\mathbf{V}$ allows for modeling dependence within $\mathbf{\Xi}_i$. In particular, we will choose this covariance matrix such that the correlation between $\zeta_{i,j}$ and $\xi_{i,j}$ is positive because $P_{i,j}$ is contained in $I_{i,j}$ (and hence they are dependent).

This last bullet point is motivated by the following argument: a positive change (an increase) from $I_{i,j-1}$ to $I_{i,j}$ means that the claims adjusters increase their expectation in future claims payments. One part of this increased expectation is immediately paid in development period j (and hence contained in both $I_{i,j}$ and $P_{i,j}$) and the remaining increased expectation is paid with some settlement delay, which means that we also have higher expectations for $P_{i,l}$, $l > j$. This argument leads to the following possible explicit choice for the correlation matrix $\widetilde{\mathbf{V}}$ (note that we have to differentiate between the covariance matrix $\mathbf{V}$ and its associated correlation matrix $\widetilde{\mathbf{V}}$ of the random vector $\mathbf{\Xi}_i$)

$\widetilde{\mathbf{V}} \quad :=$ (6.3)

	$\zeta_{i,0}$	$\zeta_{i,1}$	$\xi_{i,1}$	$\zeta_{i,2}$	$\xi_{i,2}$	$\zeta_{i,3}$	$\xi_{i,3}$	$\zeta_{i,4}$	$\xi_{i,4}$	$\cdots$	$\zeta_{i,J}$	$\xi_{i,J}$
$\zeta_{i,0}$	1	0	ρ_1	0	ρ_2	0	0	0	0	$\cdots$	0	0
$\zeta_{i,1}$	0	1	ρ_0	0	ρ_1	0	ρ_2	0	0	$\cdots$	0	0
$\xi_{i,1}$	ρ_1	ρ_0	1	0	0	0	0	0	0	$\cdots$	0	0
$\zeta_{i,2}$	0	0	0	1	ρ_0	0	ρ_1	0	ρ_2	$\cdots$	0	0
$\xi_{i,2}$	ρ_2	ρ_1	0	ρ_0	1	0	0	0	0	$\cdots$	0	0
$\zeta_{i,3}$	0	0	0	0	0	1	ρ_0	0	ρ_1	$\cdots$	0	0
$\xi_{i,3}$	0	ρ_2	0	ρ_1	0	ρ_0	1	0	0	$\cdots$	0	0
$\zeta_{i,4}$	0	0	0	0	0	0	0	1	ρ_0	$\cdots$	0	0
$\xi_{i,4}$	0	0	0	ρ_2	0	ρ_1	0	ρ_0	1	$\cdots$	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\zeta_{i,J}$	0	0	0	0	0	0	0	0	0	$\cdots$	1	ρ_0
$\xi_{i,J}$	0	0	0	0	0	0	0	0	0	$\cdots$	ρ_0	1

The rational behind this correlation matrix is that the incurred losses increments $\zeta_{i,j}$ are (positively) correlated to the claims payments increments $\xi_{i,j}$, $\xi_{i,j+1}$ and $\xi_{i,j+2}$ with positive correlations ρ_0, ρ_1 and ρ_2, respectively. $\zeta_{i,0}$ plays the special role of the initial value for incurred losses $I_{i,0}$ (on the log scale), whereas the initial value for claims payments $P_{i,0}$ (on the log scale) can be defined by $\xi_{i,0} = \sum_{j=0}^{J} \zeta_{i,j} - \sum_{l=1}^{J} \xi_{i,l}$.

Notational remark:

In comparison to Chapter 5 there will be many similarities in the notation and derivations in the following sections of this chapter. Most of the proofs in this chapter follow in a similar way as the corresponding proofs in Chapter 5 and we often refer to these proofs. However, there are some differences which are not obvious at first glance. Note that on the contrary to the Model Assumptions 5.1 in Chapter 5 the role of the paid and incurred ratios has changed in Model Assumptions 6.1. This implies that the matrix $\mathbf{B}$ in this chapter does not coincide with the matrix $\mathbf{B}$ in Chapter 5, although we choose the same symbols. Beside this, we use the basis vectors $\mathbf{e}_i$ in this chapter in a slightly different meaning in order to further simplify notation. Moreover, we focus in this chapter on the dependence structure of the paid and incurred ratios and hence do not derive the MSEP for the CDR (this can be done in the same way as in Chapter 5). This allows to leave out the time index $k \in \{J, J+1\}$ in this chapter simplifying the notation and calculations in comparison to Chapter 5 and making the derivations easier to understand.

6.2 Ultimate Claim Prediction for Known Parameters $\boldsymbol{\Theta}$

We can either work with the random vector $\boldsymbol{\Xi}_i \in \mathbb{R}^{2J+1}$ (see Model Assumptions 6.1) or with the logarithmized observations given by the random vector

$$\mathbf{X}_i := (\log I_{i,0}, \log P_{i,0}, \log I_{i,1}, \log P_{i,1}, \dots, \log I_{i,J-1}, \log P_{i,J-1}; \log I_{i,J})' \in \mathbb{R}^{2J+1}.$$

The consideration of $\boldsymbol{\Xi}_i$ was easier for the model definition and for the interpretation of the dependence structure; but often it is more straightforward if we directly work with $\mathbf{X}_i$ (under the explicit logarithmized cumulative observations). Similar to Section 5.3 there is a linear one-to-one correspondence $\mathbf{B}$ between $\boldsymbol{\Xi}_i$ and $\mathbf{X}_i$, such that $\mathbf{X}_i = \mathbf{B}\ \boldsymbol{\Xi}_i$. By this correspondence we obtain the following conditional multivariate Gaussian distribution for $\mathbf{X}_i$:

$$\mathbf{X}_i|_{\boldsymbol{\Theta}} \ = \ \mathbf{B}\ \boldsymbol{\Xi}_i|_{\boldsymbol{\Theta}} \ \sim \ \mathcal{N}\left(\boldsymbol{\mu} := \boldsymbol{\mu}(\boldsymbol{\Theta}) := \mathbf{B}\boldsymbol{\Theta}\ ,\ \boldsymbol{\Sigma} := \mathbf{B}\mathbf{V}\mathbf{B}'\right). \tag{6.4}$$

Conditionally, given the parameter vector $\boldsymbol{\Theta}$, the random vector $\mathbf{X}_i$ is multivariate Gaussian distributed. Our first aim is to study the conditional distribution of the ultimate claim $P_{i,J} = I_{i,J}$, conditionally given the parameter vector $\boldsymbol{\Theta}$ and the observations

$$\mathcal{D}_J = \{I_{i,j},\ P_{i,j}:\ i+j \le J,\ 0 \le i \le J,\ 0 \le j \le J\}$$

in the upper paid and incurred triangles, see Table 6.1.
For accident years $i \in \{1, \dots, J\}$, define $n := 2J+1$ and $q := q(i) := 2(J-i+1) \in \{2, \dots, 2J\}$. At time J we have for accident year i observations (given in the upper triangles $\mathcal{D}_J$)

$$\mathbf{X}_i^{(1)} := (\log I_{i,0}, \log P_{i,0}, \log I_{i,1}, \log P_{i,1}, \dots, \log I_{i,J-i}, \log P_{i,J-i})' \in \mathbb{R}^q,$$

and we would like to predict the lower triangles given by

$$\mathbf{X}_i^{(2)} := (\log I_{i,J-i+1}, \log P_{i,J-i+1}, \dots, \log I_{i,J-1}, \log P_{i,J-1}; \log I_{i,J})' \in \mathbb{R}^{n-q}. \tag{6.5}$$

This provides, for $i \in \{1, \ldots, J\}$, the following decomposition $\boldsymbol{\mu} = (\boldsymbol{\mu}_i^{(1)}, \boldsymbol{\mu}_i^{(2)}) = \mathbf{B}\boldsymbol{\Theta} \ \in \mathbb{R}^n$ of the conditional mean:

$$\mathrm{E}\left[\mathbf{X}_i^{(1)} \middle| \boldsymbol{\Theta}\right] = \boldsymbol{\mu}_i^{(1)} = \mathbf{B}_i^{(1)}\boldsymbol{\Theta} \in \mathbb{R}^q \qquad \text{and} \qquad \mathrm{E}\left[\mathbf{X}_i^{(2)} \middle| \boldsymbol{\Theta}\right] = \boldsymbol{\mu}_i^{(2)} = \mathbf{B}_i^{(2)}\boldsymbol{\Theta} \in \mathbb{R}^{n-q},$$

with partition of $\mathbf{B} \in \mathbb{R}^{n\times n}$ given by, for $i \in \{1, \ldots, J\}$,

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_i^{(1)} \\ \mathbf{B}_i^{(2)} \end{pmatrix} \qquad \text{with} \qquad \mathbf{B}_i^{(1)} \in \mathbb{R}^{q\times n} \ \text{ and } \ \mathbf{B}_i^{(2)} \in \mathbb{R}^{(n-q)\times n}. \tag{6.6}$$

In the same way we also decompose the covariance matrix which provides

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_i^{(11)} & \boldsymbol{\Sigma}_i^{(12)} \\ \boldsymbol{\Sigma}_i^{(21)} & \boldsymbol{\Sigma}_i^{(22)} \end{pmatrix} \qquad \text{with} \qquad \boldsymbol{\Sigma}_i^{(11)} \in \mathbb{R}^{q\times q}. \tag{6.7}$$

For $i = 0$ we set $q(0) := n$, $\mathbf{X}_0^{(1)} := \mathbf{X}_0 \in \mathbb{R}^n$, $\boldsymbol{\Sigma}_0^{(11)} := \boldsymbol{\Sigma}$ and $\mathbf{B}_0^{(1)} := \mathbf{B}$, but (6.6) and (6.7) do not hold in this case.

Having this notation, we provide the prediction of $\mathbf{X}_i^{(2)}$, conditionally given $\{\boldsymbol{\Theta}, \mathcal{D}_J\}$:

Lemma 6.3 (Conditional distribution) *Choose an accident year $i \in \{1, \ldots, J\}$. Under Model Assumptions 6.1 we have*

$$\mathbf{X}_i^{(2)}\Big|_{\{\boldsymbol{\Theta},\mathcal{D}_J\}} \ = \ \mathbf{X}_i^{(2)}\Big|_{\{\boldsymbol{\Theta},\mathbf{X}_i^{(1)}\}} \ \sim \ \mathcal{N}\left(\widetilde{\boldsymbol{\mu}}_i^{(2)} \ , \ \widetilde{\boldsymbol{\Sigma}}_i^{(22)} \right),$$

with the conditional mean and covariance matrix

$$\widetilde{\boldsymbol{\mu}}_i^{(2)} := \boldsymbol{\mu}_i^{(2)} + \boldsymbol{\Sigma}_i^{(21)}(\boldsymbol{\Sigma}_i^{(11)})^{-1}\left(\mathbf{X}_i^{(1)} - \boldsymbol{\mu}_i^{(1)}\right) \quad \textit{and} \quad \widetilde{\boldsymbol{\Sigma}}_i^{(22)} := \boldsymbol{\Sigma}_i^{(22)} - \boldsymbol{\Sigma}_i^{(21)}(\boldsymbol{\Sigma}_i^{(11)})^{-1}\boldsymbol{\Sigma}_i^{(12)}.$$

Proof: The proof follows in the same way as the proof of Lemma 5.3. □

An immediate consequence of Lemma 6.3 is the following corollary, which constitutes an analog of Corollary 5.5 for the PIC model with dependence.

Corollary 6.4 (Conditional distribution) *Under the assumptions and notation of Lemma 6.3 we obtain for the ultimate claim $I_{i,J} = P_{i,J}$, for $i \in \{1, \ldots, J\}$,*

$$\log I_{i,J}|_{\{\boldsymbol{\Theta},\mathcal{D}_J\}} \ \sim \ \mathcal{N}\left(\mathbf{e}_i' \ \widetilde{\boldsymbol{\mu}}_i^{(2)} \ , \ \mathbf{e}_i' \ \widetilde{\boldsymbol{\Sigma}}_i^{(22)} \ \mathbf{e}_i \ \right).$$

Proof: By Lemma 6.3 the random vector $\mathbf{X}_i^{(2)}$ has (conditional) distribution

$$\mathbf{X}_i^{(2)}\Big|_{\{\boldsymbol{\Theta},\mathcal{D}_J\}} \sim \ \mathcal{N}\left(\widetilde{\boldsymbol{\mu}}_i^{(2)}, \ \widetilde{\boldsymbol{\Sigma}}_i^{(22)} \ \right).$$

Since $\log I_{i,J}$ is the last entry of $\mathbf{X}_i^{(2)}$, see (6.5), we have that $\log I_{i,J} = \mathbf{e}_i' \ \mathbf{X}_i^{(2)}$. □

This corollary implies that, conditionally given the parameter vector $\boldsymbol{\Theta}$ and the observations $\mathcal{D}_J$, we get the ultimate claim predictor, for $i \in \{1, \ldots, J\}$,

$$\mathrm{E}[I_{i,J}|\,\boldsymbol{\Theta}, \mathcal{D}_J] = \exp\left\{\mathbf{e}_i' \ \widetilde{\boldsymbol{\mu}}_i^{(2)} + \mathbf{e}_i' \ \widetilde{\boldsymbol{\Sigma}}_i^{(22)} \ \mathbf{e}_i/2\right\}. \tag{6.8}$$

In the special case of a diagonal correlation matrix (6.3), i.e. $\rho_0 = \rho_1 = \rho_2 = 0$, this is exactly the predictor derived in Corollary 2.5 of MERZ–WÜTHRICH [46].

6.3 Estimation of Parameter $\boldsymbol{\Theta}$

The ultimate claim predictor (6.8) is still based on the unknown parameter vector $\boldsymbol{\Theta}$, namely

$$\begin{aligned} \mathbf{e}_i' \, \widetilde{\boldsymbol{\mu}}_i^{(2)} &= \mathbf{e}_i' \left(\boldsymbol{\mu}_i^{(2)} + \boldsymbol{\Sigma}_i^{(21)} (\boldsymbol{\Sigma}_i^{(11)})^{-1} \left(\mathbf{X}_i^{(1)} - \boldsymbol{\mu}_i^{(1)} \right) \right) \\ &= \mathbf{e}_i' \left(\mathbf{B}_i^{(2)} \, \boldsymbol{\Theta} + \boldsymbol{\Sigma}_i^{(21)} (\boldsymbol{\Sigma}_i^{(11)})^{-1} \left(\mathbf{X}_i^{(1)} - \mathbf{B}_i^{(1)} \, \boldsymbol{\Theta} \right) \right) \\ &= \boldsymbol{\Gamma}_i \, \boldsymbol{\Theta} + \mathbf{e}_i' \, \boldsymbol{\Sigma}_i^{(21)} (\boldsymbol{\Sigma}_i^{(11)})^{-1} \, \mathbf{X}_i^{(1)}, \end{aligned} \tag{6.9}$$

where we have defined

$$\boldsymbol{\Gamma}_i := \mathbf{e}_i' \left(\mathbf{B}_i^{(2)} - \boldsymbol{\Sigma}_i^{(21)} (\boldsymbol{\Sigma}_i^{(11)})^{-1} \mathbf{B}_i^{(1)} \right).$$

In particular, we see that $\mathbf{e}_i' \, \widetilde{\boldsymbol{\mu}}_i^{(2)}$ is an affine-linear function in $\boldsymbol{\Theta}$. We aim to calculate the posterior distribution of $\boldsymbol{\Theta}$, conditionally given the observations $\mathcal{D}_J$. The σ-field generated by $\mathcal{D}_J$ is the same as the one generated by $\widetilde{\mathcal{D}}_J = \{\mathbf{X}_0^{(1)}, \dots, \mathbf{X}_J^{(1)}\}$. Therefore, by a slight abuse of notation, we identify the observations $\widetilde{\mathcal{D}}_J$ with $\mathcal{D}_J$. The likelihood of the logarithmized observations, conditionally given $\boldsymbol{\Theta}$, is then written as, see also (6.4),

$$l_{\mathcal{D}_J}(\boldsymbol{\Theta}) \;\; \propto \;\; \prod_{i=0}^{J} \exp \left\{ -\frac{1}{2} \left(\mathbf{X}_i^{(1)} - \mathbf{B}_i^{(1)} \, \boldsymbol{\Theta} \right)' (\boldsymbol{\Sigma}_i^{(11)})^{-1} \left(\mathbf{X}_i^{(1)} - \mathbf{B}_i^{(1)} \, \boldsymbol{\Theta} \right) \right\}.$$

Under Model Assumptions 6.1 the posterior density of $\boldsymbol{\Theta}$, given $\mathcal{D}_J$, is given by

$$u\left(\boldsymbol{\Theta} \middle| \mathcal{D}_J\right) \;\propto\; l_{\mathcal{D}_J}(\boldsymbol{\Theta}) \; \exp \left\{ -\frac{1}{2} \left(\boldsymbol{\Theta} - \boldsymbol{\vartheta} \right)' \mathbf{T}^{-1} \left(\boldsymbol{\Theta} - \boldsymbol{\vartheta} \right) \right\}, \tag{6.10}$$

where the last term is the prior density of $\boldsymbol{\Theta}$ with prior mean given by

$$\boldsymbol{\vartheta} := (\psi_0; \psi_1, \phi_1, \psi_2, \phi_2, \dots, \psi_J, \phi_J)' \in \mathbb{R}^n,$$

and (diagonal) covariance matrix defined by

$$\mathbf{T} := \operatorname{diag}(t_0^2; t_1^2, s_1^2, t_2^2, s_2^2, \dots, t_J^2, s_J^2) \in \mathbb{R}^{n \times n}.$$

This immediately implies the following theorem:

Theorem 6.5 (Posterior distribution of $\boldsymbol{\Theta}$) *Under Model Assumptions 6.1, the posterior distribution of $\boldsymbol{\Theta}$, given $\mathcal{D}_J$, is a multivariate Gaussian distribution with posterior mean*

$$\boldsymbol{\vartheta}(\mathcal{D}_J) := \mathbf{T}(\mathcal{D}_J) \left[\mathbf{T}^{-1} \, \boldsymbol{\vartheta} + \sum_{i=0}^{J} (\mathbf{B}_i^{(1)})' (\boldsymbol{\Sigma}_i^{(11)})^{-1} \mathbf{X}_i^{(1)} \right],$$

and posterior covariance matrix

$$\mathbf{T}(\mathcal{D}_J) := \left(\mathbf{T}^{-1} + \sum_{i=0}^{J} (\mathbf{B}_i^{(1)})' (\boldsymbol{\Sigma}_i^{(11)})^{-1} \mathbf{B}_i^{(1)} \right)^{-1}.$$

Proof: The proof follows in the same way as the proof of Theorem 5.6. □

Remarks on credibility theory:

We define the matrix

$$\mathbf{S} := \left(\sum_{i=0}^{J} (\mathbf{B}_i^{(1)})' (\mathbf{\Sigma}_i^{(11)})^{-1} \mathbf{B}_i^{(1)} \right)^{-1}.$$

The existence of $\mathbf{S}$ follows by the fact that $\sum_{i=0}^{J} (\mathbf{B}_i^{(1)})' (\mathbf{\Sigma}_i^{(11)})^{-1} \mathbf{B}_i^{(1)}$ is as sum of *symmetric positive definite* (s.p.d) matrices also s.p.d. and hence invertible. Moreover, we define the credibility weights ($\mathbf{I}$ is the identity matrix)

$$\mathbf{A} := \left(\mathbf{T}^{-1} + \mathbf{S}^{-1}\right)^{-1} \mathbf{S}^{-1} \qquad \text{and} \qquad \mathbf{I} - \mathbf{A} = \left(\mathbf{T}^{-1} + \mathbf{S}^{-1}\right)^{-1} \mathbf{T}^{-1}, \tag{6.11}$$

see also formula (7.11) in Bühlmann–Gisler [14]. This implies for the posterior covariance matrix $\mathbf{T}(\mathcal{D}_J) = \mathbf{AS} = (\mathbf{I} - \mathbf{A})\mathbf{T}$ and we obtain the credibility formula for the posterior mean

$$\boldsymbol{\vartheta}(\mathcal{D}_J) = (\mathbf{I} - \mathbf{A})\, \boldsymbol{\vartheta} + \mathbf{A}\ \mathbf{Y}, \tag{6.12}$$

with compressed data

$$\mathbf{Y} := \mathbf{S} \left[\sum_{i=0}^{J} (\mathbf{B}_i^{(1)})' (\mathbf{\Sigma}_i^{(11)})^{-1} \mathbf{X}_i^{(1)} \right], \tag{6.13}$$

see Chapters 7 and 8 in Bühlmann–Gisler [14]. That is, the posterior mean $\boldsymbol{\vartheta}(\mathcal{D}_J)$ is a credibility weighted average between the prior mean $\boldsymbol{\vartheta}$ and the observations $\mathbf{Y}$ with credibility matrix $\mathbf{A}$. The crucial point why we obtain identical results using Theorem 6.5 and credibility theory is that the normal prior distribution and the normal conditional distribution in Model Assumptions 6.1 lead to the exact credibility case, see Bühlmann–Gisler [14] for details. Therefore, the compressed data vector (6.13) in the PIC method has a similar structure as the compressed data vector (4.14) used for the credibility predictors $\mathbf{F}_k^{I,Cred}$ in the (Bayesian) LSRM and (6.12) corresponds to the credibility predictor $\mathbf{F}_k^{I,Cred}$ for Bayesian LSRMs, see Subsection 4.2.3. Now we state an analog to Corollary 5.7 for the PIC model with dependence.

Corollary 6.6 (Ultimate claim predictor) *Under Model Assumptions 6.1 we obtain for $I_{i,J} = P_{i,J}$ the ultimate claim predictor, given observations $\mathcal{D}_J$,*

$$\mathrm{E}[I_{i,J} |\, \mathcal{D}_J] = \exp\left\{ \mathbf{\Gamma}_i\ \boldsymbol{\vartheta}(\mathcal{D}_J) + \mathbf{\Gamma}_i\ \mathbf{T}(\mathcal{D}_J)\ \mathbf{\Gamma}_i'/2 + \mathbf{e}_i'\ \mathbf{\Sigma}_i^{(21)} (\mathbf{\Sigma}_i^{(11)})^{-1}\ \mathbf{X}_i^{(1)} + \mathbf{e}_i'\ \widetilde{\mathbf{\Sigma}}_i^{(22)}\ \mathbf{e}_i/2 \right\}.$$

Proof:

The proof follows by Theorem 6.5 and (6.8) in the same way as the proof of Corollary 5.7. □

Corollary 6.6 gives the ultimate claim predictor that is now based on claims payments, incurred losses and prior expert information. In contrast to Merz–Wüthrich [46] we can now easily choose any meaningful covariance matrix $\mathbf{V}$ for $\mathbf{\Xi}_i$.

6.4 Prediction Uncertainty

In order to analyze the prediction uncertainty, we can now study the posterior predictive distribution of $\mathbf{I}_J = (I_{1,J}, \ldots, I_{J,J})$, which exactly corresponds to the column of unknown ultimate claims, given the observations $\mathcal{D}_J$. If $g(\cdot)$ is a sufficiently nice function we obtain

$$\mathrm{E}[g(\mathbf{I}_J)|\,\mathcal{D}_J] = \int_{\mathbf{x}\in\mathbb{R}^J} g(\mathbf{x})\ f\left(\mathbf{x}|\,\mathcal{D}_J\right)\ d\mathbf{x} = \int_{\mathbf{x}\in\mathbb{R}^J,\boldsymbol{\Theta}} g(\mathbf{x})\ f\left(\mathbf{x}|\,\boldsymbol{\Theta},\mathcal{D}_J\right)\ u\left(\boldsymbol{\Theta}|\,\mathcal{D}_J\right)\ d\mathbf{x}\ d\boldsymbol{\Theta},$$

where $u\left(\boldsymbol{\Theta}|\,\mathcal{D}_J\right)$ denotes the posterior density of $\boldsymbol{\Theta}$, given $\mathcal{D}_J$ (cf. (6.10)). Because the densities $f\left(\mathbf{x}|\,\boldsymbol{\Theta},\mathcal{D}_J\right)$ and $u\left(\boldsymbol{\Theta}|\,\mathcal{D}_J\right)$ are explicitly given by Corollary 6.4 and (6.10) and the conditional independence of accident years, given $\boldsymbol{\Theta}$, we can calculate the predictive values $\mathrm{E}[g(\mathbf{I}_J)|\,\mathcal{D}_J]$ numerically, for example using MC simulations. This allows for the analysis of any uncertainty and risk measure. If we consider the total ultimate claim $\sum_{i=1}^J I_{i,J}$ and the corresponding predictor $\sum_{i=1}^J \mathrm{E}[I_{i,J}|\,\mathcal{D}_J]$ the conditional MSEP is given by

$$\mathrm{msep}_{\sum_{i=1}^J I_{i,J}\big|\mathcal{D}_J}\left[\sum_{i=1}^J \mathrm{E}[I_{i,J}|\,\mathcal{D}_J]\right] = \mathrm{E}\left[\left(\sum_{i=1}^J I_{i,J} - \sum_{i=1}^J \mathrm{E}[I_{i,J}|\,\mathcal{D}_J]\right)^2\middle|\,\mathcal{D}_J\right] = \mathrm{Var}\left[\sum_{i=1}^J I_{i,J}\middle|\,\mathcal{D}_J\right].$$

Henceforth, we need to calculate this last conditional variance in order to obtain the conditional MSEP.

Theorem 6.7 *Under Model Assumptions 6.1 the conditional MSEP is given by*

$$\begin{aligned}\mathrm{msep}_{\sum_{i=1}^J I_{i,J}\big|\mathcal{D}_J}&\left[\sum_{i=1}^J \mathrm{E}[I_{i,J}|\,\mathcal{D}_J]\right]\\ &= \sum_{i,k=1}^J \mathrm{E}[I_{i,J}|\,\mathcal{D}_J]\ \ \mathrm{E}[I_{k,J}|\,\mathcal{D}_J]\ \ \left(\exp\left\{\boldsymbol{\Gamma}_i\ \mathbf{T}(\mathcal{D}_J)\ \boldsymbol{\Gamma}_k' + \mathbf{e}_i'\ \widetilde{\boldsymbol{\Sigma}}_i^{(22)}\ \mathbf{e}_i\ \mathbf{1}_{\{i=k\}}\right\} - 1\right),\end{aligned}$$

with $\mathrm{E}[I_{i,J}|\,\mathcal{D}_J]$ *given by Corollary 6.6.*

Proof: With the variance decoupling formula we obtain

$$\begin{aligned}\mathrm{Var}\left[\sum_{i=1}^J I_{i,J}\middle|\,\mathcal{D}_J\right] &= \sum_{i,k=1}^J \mathrm{Cov}[I_{i,J}, I_{k,J}|\,\mathcal{D}_J]\\ &= \sum_{i,k=1}^J \mathrm{Cov}[\mathrm{E}[I_{i,J}|\,\boldsymbol{\Theta},\mathcal{D}_J]\,,\mathrm{E}[I_{k,J}|\,\boldsymbol{\Theta},\mathcal{D}_J]|\,\mathcal{D}_J] + \sum_{i=1}^J \mathrm{E}[\mathrm{Var}[I_{i,J}|\,\boldsymbol{\Theta},\mathcal{D}_J]|\,\mathcal{D}_J]\,, \end{aligned}\tag{6.14}$$

where for the second term on the right hand side of (6.14) we have used the conditional independence of different accident years, given $\boldsymbol{\Theta}$. Thus, we need to calculate these last two terms. Using Corollary 6.4 and (6.9), we obtain

$$\mathrm{E}[I_{i,J}|\,\boldsymbol{\Theta},\mathcal{D}_J] = \exp\left\{\boldsymbol{\Gamma}_i\ \boldsymbol{\Theta} + \mathbf{e}_i'\ \boldsymbol{\Sigma}_i^{(21)}(\boldsymbol{\Sigma}_i^{(11)})^{-1}\ \mathbf{X}_i^{(1)} + \mathbf{e}_i'\ \widetilde{\boldsymbol{\Sigma}}_i^{(22)}\ \mathbf{e}_i/2\right\}, \tag{6.15}$$

and

$$\mathrm{Var}[\, I_{i,J} |\, \mathbf{\Theta}, \mathcal{D}_J] = \mathrm{E}[\, I_{i,J} |\, \mathbf{\Theta}, \mathcal{D}_J]^2 \left(\exp \left\{ \mathbf{e}_i' \; \widetilde{\mathbf{\Sigma}}_i^{(22)} \; \mathbf{e}_i \right\} - 1 \right). \tag{6.16}$$

We first treat the second term in (6.14). Using (6.16) leads to

$$\begin{aligned}
&\mathrm{E}[\mathrm{Var}[\, I_{i,J} |\, \mathbf{\Theta}, \mathcal{D}_J] |\, \mathcal{D}_J] \\
&= \mathrm{E}\Big[\exp \Big\{ 2\mathbf{\Gamma}_i \; \mathbf{\Theta} + 2\mathbf{e}_i' \; \mathbf{\Sigma}_i^{(21)} (\mathbf{\Sigma}_i^{(11)})^{-1} \; \mathbf{X}_i^{(1)} + \mathbf{e}_i' \; \widetilde{\mathbf{\Sigma}}_i^{(22)} \; \mathbf{e}_i \Big\} \Big|\, \mathcal{D}_J \Big] \left(\exp \left\{ \mathbf{e}_i' \; \widetilde{\mathbf{\Sigma}}_i^{(22)} \; \mathbf{e}_i \right\} - 1 \right) \\
&= \mathrm{E}[\, I_{i,J} |\, \mathcal{D}_J]^2 \exp \left\{ \mathbf{\Gamma}_i \; \mathbf{T}(\mathcal{D}_J) \; \mathbf{\Gamma}_i' \right\} \left(\exp \left\{ \mathbf{e}_i' \; \widetilde{\mathbf{\Sigma}}_i^{(22)} \; \mathbf{e}_i \right\} - 1 \right).
\end{aligned}$$

For the first term in (6.14) we need to consider

$$\begin{aligned}
&\mathrm{Cov}[\mathrm{E}[\, I_{i,J} |\, \mathbf{\Theta}, \mathcal{D}_J] \,, \mathrm{E}[\, I_{k,J} |\, \mathbf{\Theta}, \mathcal{D}_J] |\, \mathcal{D}_J] \\
&\quad = \exp \left\{ \mathbf{e}_i' \; \mathbf{\Sigma}_i^{(21)} (\mathbf{\Sigma}_i^{(11)})^{-1} \; \mathbf{X}_i^{(1)} + \mathbf{e}_i' \; \widetilde{\mathbf{\Sigma}}_i^{(22)} \; \mathbf{e}_i/2 + \mathbf{e}_k' \; \mathbf{\Sigma}_i^{(21)} (\mathbf{\Sigma}_i^{(11)})^{-1} \; \mathbf{X}_k^{(1)} + \mathbf{e}_k' \; \widetilde{\mathbf{\Sigma}}_i^{(22)} \; \mathbf{e}_k/2 \right\} \\
&\qquad \times \; \mathrm{Cov}[\exp \left\{ \mathbf{\Gamma}_i \; \mathbf{\Theta} \right\} , \exp \left\{ \mathbf{\Gamma}_k \; \mathbf{\Theta} \right\} |\, \mathcal{D}_J] \,,
\end{aligned}$$

where the last covariance term is given by

$$\begin{aligned}
&\mathrm{Cov}[\exp \left\{ \mathbf{\Gamma}_i \; \mathbf{\Theta} \right\} , \exp \left\{ \mathbf{\Gamma}_k \; \mathbf{\Theta} \right\} |\, \mathcal{D}_J] \\
&\qquad = \mathrm{E}[\exp \left\{ \mathbf{\Gamma}_i \; \mathbf{\Theta} \right\} |\, \mathcal{D}_J] \;\; \mathrm{E}[\exp \left\{ \mathbf{\Gamma}_k \; \mathbf{\Theta} \right\} |\, \mathcal{D}_J] \;\; \left(\exp \left\{ \mathbf{\Gamma}_i \; \mathbf{T}(\mathcal{D}_J) \; \mathbf{\Gamma}_k' \right\} - 1 \right).
\end{aligned}$$

Henceforth, using (6.16) we obtain for the first term in (6.14)

$$\mathrm{Cov}[\mathrm{E}[\, I_{i,J} |\, \mathbf{\Theta}, \mathcal{D}_J] \,, \mathrm{E}[\, I_{k,J} |\, \mathbf{\Theta}, \mathcal{D}_J] |\, \mathcal{D}_J] = \mathrm{E}[\, I_{i,J} |\, \mathcal{D}_J] \, \mathrm{E}[\, I_{k,J} |\, \mathcal{D}_J] \left(\exp \left\{ \mathbf{\Gamma}_i \; \mathbf{T}(\mathcal{D}_J) \; \mathbf{\Gamma}_k' \right\} - 1 \right).$$

Collecting all the terms completes the proof. □

Thus, we obtain a closed form solution for both, the ultimate claim predictors $\mathrm{E}[\, I_{i,J} |\, \mathcal{D}_J]$ and the corresponding prediction errors, measured by the conditional MSEP.

6.5 Example PIC Reserving Method with Dependence Modeling

We apply the PIC model with dependence to the motor third party liability data given in Table 7.8 and 7.9 below. In Model Assumptions 6.1 we work with logarithmized paid claims ratios $\zeta_{i,k}$ and logarithmized incurred losses ratios $\xi_{i,k}$, respectively (cf. (6.1)). That means that we have to transform the data in Table 7.8 and 7.9 into $\zeta_{i,j}$ and $\xi_{i,l}$. Due to the fact that there is no expert knowledge for the specific choice of the means in the prior distributions for Ψ_l and Φ_j we choose in Model Assumptions 6.1 non-informative priors, i.e. we let $t_j^2 \to \infty$ and $s_l^2 \to \infty$. This implies asymptotically for the credibility matrix $\mathbf{A} = \mathbf{I}$ in (6.11) and no prior knowledge is used in our calculations.

For Model Assumptions 6.1, it remains to choose a suitable covariance matrix $\mathbf{V}$. Here we present three different choices of correlation matrices $\widetilde{\mathbf{V}}$ of the type (6.3) , which are motivated by an ad-hoc estimate. The corresponding covariance matrix $\mathbf{V}$ is then given by

$$\mathbf{V} := \mathbf{Var}^{1/2}\ \widetilde{\mathbf{V}}\ \mathbf{Var}^{1/2}, \tag{6.17}$$

where

$$\mathbf{Var} := \operatorname{diag}\left(\sigma^2_{\zeta_0}; \sigma^2_{\zeta_1}, \sigma^2_{\xi_1}, \ldots, \sigma^2_{\zeta_J}, \sigma^2_{\xi_J}\right).$$

The estimator for the correlation matrix of the type (6.3) should be seen as an intuitive proposal for a correlation structure and not as an estimator being optimal in some mathematical sense.

Correlation Matrix Choice

The choice of a correlation matrix of type (6.3) reduces the number of parameters to be estimated in comparison to the estimation of a general correlation matrix. Note that we have decided for structure (6.3) by pure expert choice. For a correlation matrix of type (6.3), we have to choose ρ_l for $l \in \{0, 1, 2\}$ as

$$\rho_l = \operatorname{Cor}[\zeta_{i,k}, \xi_{i,k+l}] \qquad \text{for} \quad i = 1, \ldots, J \quad \text{and} \quad k = 1, \ldots, J-l,$$

and

$$\rho_l = \operatorname{Cor}[\zeta_{i,0}, \xi_{i,l}] \quad \text{for} \quad l \in \{1, 2\}.$$

We propose the following ad-hoc estimators for ρ_l.

1:

For the unknown means $\Psi_k = \mathrm{E}[\zeta_{i,k}]$ and $\Phi_k = \mathrm{E}[\xi_{i,k}]$ as well as variances $\sigma^2_{\zeta_k} = \operatorname{Var}[\zeta_{i,k}]$ and $\sigma^2_{\xi_k} = \operatorname{Var}[\xi_{i,k}]$ we use sample estimates

$$\hat{\Psi}_k := \frac{1}{J-k+1} \sum_{i=0}^{J-k} \zeta_{i,k} \text{ for } k = 0, \ldots, J \qquad \hat{\sigma}^2_{\zeta_k} := \frac{1}{J-k} \sum_{i=0}^{J-k} \left(\zeta_{i,k} - \hat{\Psi}_k\right)^2 \text{ for } k = 0, \ldots, J-1$$

and

$$\hat{\Phi}_k := \frac{1}{J-k+1} \sum_{i=0}^{J-k} \xi_{i,k} \text{ for } k = 1, \ldots, J \qquad \hat{\sigma}^2_{\xi_k} := \frac{1}{J-k} \sum_{i=0}^{J-k} \left(\xi_{i,k} - \hat{\Phi}_k\right)^2 \text{ for } k = 1, \ldots, J-1.$$

Since for the estimation of the last variance parameters $\hat{\sigma}^2_{\zeta_J}$ and $\hat{\sigma}^2_{\xi_J}$ there is only one observation in the observed triangle we use the well-known extrapolation formula

$$\hat{\sigma}^2_{\zeta_J} := \min\{\hat{\sigma}^2_{\zeta_{J-2}}, \hat{\sigma}^2_{\zeta_{J-1}}, \hat{\sigma}^4_{\zeta_{J-2}}/\hat{\sigma}^2_{\zeta_{J-1}}\} \text{ and } \hat{\sigma}^2_{\xi_J} := \min\{\hat{\sigma}^2_{\xi_{J-2}}, \hat{\sigma}^2_{\xi_{J-1}}, \hat{\sigma}^4_{\xi_{J-2}}/\hat{\sigma}^2_{\xi_{J-1}}\},$$

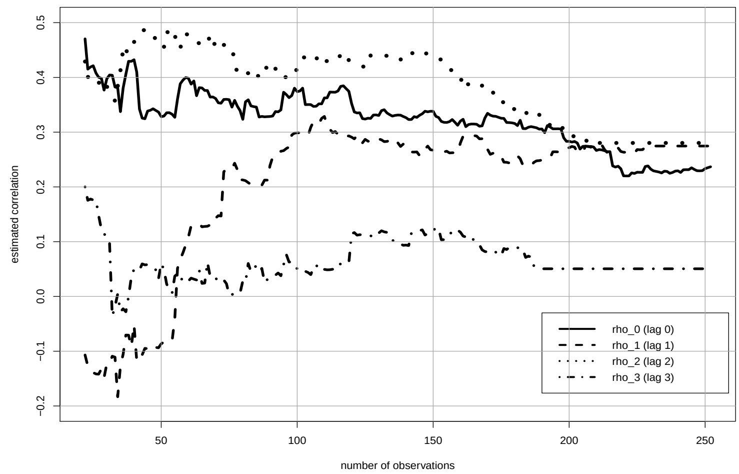

Figure 6.1: Correlation estimators $\hat{\rho}_l$ for ρ_l for $l \in \{0, 1, 2, 3\}$ as a function of the number of observations used for the estimation

see WÜTHRICH–MERZ [63].

2:

We consider for accident year $i \in \{0, \dots, J\}$ the standardized logarithmized ratios

$$\tilde{\zeta}_{i,k} := \frac{\zeta_{i,k} - \hat{\Psi}_k}{\hat{\sigma}_{\zeta_k}} \quad \text{for } k = 0, \dots J - i \quad \text{and} \quad \tilde{\xi}_{i,k} := \frac{\xi_{i,k} - \hat{\Phi}_k}{\hat{\sigma}_{\xi_k}} \text{ for } k = 1, \dots J - i.$$

3:

We use the correlation estimator for ρ_l given by

$$\hat{\rho}_l := \sum_{k=1}^{J-l} \sum_{i=1}^{J-k-l} \tilde{\zeta}_{i,k} \tilde{\xi}_{i,k+l} \text{ for } l \in \{0, 1, 2, 3\}. \tag{6.18}$$

According to the correlation estimators (6.18) we obtain for $\hat{\rho}_l$ with $l = 0, 1, 2, 3$ as a function of the number of observations the values given in Figure 6.1. We see in Figure 6.1 that the assumption of positive correlatedness between $\zeta_{i,k}$ and $\xi_{i,k+l}$ for $l \in \{0, 1, 2\}$ is evident. For $l = 3$ or higher time lags the correlation estimation is comparably small (about 5%) and will therefore be neglected in our following considerations. For the sample estimators (6.18) we obtain:

$\hat{\rho}_0$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
23%	27%	28%	5%

(6.19)

In order to study correlation sensitivities, we make three explicit correlation choices, see Table 6.2, based on the correlation estimates in (6.19) and compare it to the uncorrelated case (case 0) treated in MERZ–WÜTHRICH [46]. For these cases we have to check whether the corresponding covariance matrix $\mathbf{V}$ is positive definite (see Model Assumptions 6.1). To calculate $\mathbf{V}$, we use the identity (6.17) and obtain

$$\mathbf{V} := \widehat{\mathbf{Var}}^{1/2} \widetilde{\mathbf{V}}_l \widehat{\mathbf{Var}}^{1/2},$$

case 0	case 1	case 2	case 3
$\rho_0 = \rho_1 = \rho_2 = 0\%$	$\rho_0 = 30\%, \rho_1 = 25\%, \rho_2 = 40\%$	$\rho_0 = 30\%, \rho_1 = 25\%, \rho_2 = 30\%$	$\rho_0 = 25\%, \rho_1 = 25\%, \rho_2 = 30\%$

Table 6.2: Uncorrelated case and three explicit choices for correlations

where $\widehat{\mathbf{Var}}$ denotes the variance estimates

$$\widehat{\mathbf{Var}} := \operatorname{diag}\left(\hat{\sigma}^2_{\zeta_0}; \hat{\sigma}^2_{\zeta_1}, \hat{\sigma}^2_{\xi_1}, \ldots, \hat{\sigma}^2_{\zeta_J}, \hat{\sigma}^2_{\xi_J}\right).$$

Since $\widehat{\mathbf{Var}}^{1/2}$ is diagonal we only have to check, whether the matrix $\widetilde{\mathbf{V}}_l$ is positive definite. The eigenvalues of the estimated correlation matrix $\widetilde{\mathbf{V}}_l$ are for the four cases strictly positive, the smallest being $9.2 \cdot 10^{-12}$ and the largest being $3.8 \cdot 10^{-03}$ and hence the corresponding covariance matrix $\mathbf{V}$ fulfills Model Assumptions 6.1.
Based on these choices for the covariance matrix $\mathbf{V}$, we calculate the ultimate claim reserves and the conditional MSEP.

Remarks 6.8 (Example PIC reserving method with dependence modeling)

i) *We choose by means of explorative data analysis explicit covariance matrices* $\mathbf{V}$*. This was partly done by intuitive expert knowledge.*

ii) *The derivation of an optimal estimator* $\widetilde{\mathbf{V}}$ *for the covariance matrix* $\mathbf{V}$ *with good statistical properties is not trivial and should be subject to more statistical research. Therefore, we present an ad-hoc estimator for the correlation matrix and use the resulting estimates as an orientation for different explicit choices for the correlation structure (case 1-3).*

iii) *Model Assumptions 6.1 allow for arbitrary covariance matrices as long as they are positive definite. If sufficient data for a robust estimation of their* $n(n+1)/2$ *entries is available, there is no need to reduce to correlations up to lag 2. However, we believe (due to over-parametrization) that an arbitrary correlation structure is not a feasible alternative and expert opinion always needs to specify additional structure.*

iv) *Positive definiteness of* $\mathbf{V}$ *should always be checked because most intuitive choices do not provide a positive definite covariance matrix.*

Claims Reserves and Prediction Uncertainty

1) Claims reserves at time J:
We consider the expected outstanding loss liabilities (claims reserves)

$$\hat{R}(\mathcal{D}_J) := E[I_{i,J}|\mathcal{D}_J] - P_{i,J-i}$$

acc. year	$\hat{R}(\mathcal{D}_J)$ case 0	$\hat{R}(\mathcal{D}_J)$ case 1	$D(\hat{R}(\mathcal{D}_J))$ case 1	$\hat{R}(\mathcal{D}_J)$ case 2	$D(\hat{R}(\mathcal{D}_J))$ case 2	$\hat{R}(\mathcal{D}_J)$ case 3	$D(\hat{R}(\mathcal{D}_J))$ case 3
1	7.726	7.729	0,0%	7.729	0,0%	7.728	0,0%
2	12.084	12.090	0,0%	12.089	0,0%	12.087	0,0%
3	15.196	15.537	2,2%	15.423	1,5%	15.397	1,3%
4	9.916	8.291	16,4%	8.664	12,6%	8.718	12,1%
5	20.746	21.310	2,7%	21.169	2,0%	21.096	1,7%
6	23.675	24.111	1,8%	24.102	1,8%	24.047	1,6%
7	33.328	33.410	0,2%	33.749	1,3%	33.683	1,1%
8	35.740	37.369	4,6%	37.327	4,4%	37.146	3,9%
9	40.144	38.695	3,6%	39.669	1,2%	39.767	0,9%
10	53.888	48.764	9,5%	51.602	4,2%	51.788	3,9%
11	62.825	59.284	5,6%	61.134	2,7%	61.233	2,5%
12	79.164	77.724	1,8%	78.716	0,6%	78.352	1,0%
13	89.437	81.510	8,9%	85.614	4,3%	85.572	4,3%
14	88.300	79.565	9,9%	82.942	6,1%	83.358	5,6%
15	122.534	107.575	12,2%	115.540	5,7%	116.508	4,9%
16	126.151	108.955	13,6%	117.667	6,7%	118.831	5,8%
17	126.202	119.794	5,1%	122.695	2,8%	122.682	2,8%
18	127.522	124.947	2,0%	126.287	1,0%	125.897	1,3%
19	152.078	143.847	5,4%	147.725	2,9%	148.060	2,6%
20	185.586	170.054	8,4%	175.798	5,3%	177.062	4,6%
21	251.803	246.960	1,9%	248.818	1,2%	248.554	1,3%
total	1.664.045	1.567.522	5,8%	1.614.459	3,0%	1.617.568	2,8%

Table 6.3: Claims reserves in the classical PIC model and PIC model with dependence

at time J. The percental difference between claims reserves with and without dependence is denoted by $D(\hat{R}(\mathcal{D}_J))$. We observe in Table 6.3 that in the first case the claims reserves are about 6% lower than the claims reserves without dependence. In the other two cases the difference is still about 3%. This shows that the specific choice of correlation structure has a crucial impact on the size of claims reserves.

2) Prediction uncertainty at time J:

In Table 6.4 we provide the MSEP for our four explicit correlation structures, see Table 6.2. We observe that the prediction uncertainty measured in terms of the conditional MSEP for the PIC

	$\text{msep}^{1/2}$ case 0	$\text{msep}^{1/2}$ case 1	$\text{msep}^{1/2}$ case 2	$\text{msep}^{1/2}$ case 3
total	40.606	48.010	49.145	48.922
in % of claims reserves	2,44%	3,06%	3,04%	3,02%

Table 6.4: Prediction uncertainty $\text{msep}^{1/2}$ for the classical PIC model and the PIC model with dependence

model with dependence is higher than in the classical PIC model. The reason is that positive correlations of the type (6.3) between paid and incurred ratios in Model Assumptions 6.1 increase the correlations in the ultimate outcomes, and hence the uncertainties. This means that not considering dependence in the PIC model clearly underestimates the total uncertainty.

6.6 Conclusions

In the original PIC model of Merz–Wüthrich [46] a unified predictor of the ultimate claim based on incurred losses and claims payments as well as the corresponding prediction uncertainty in terms of the conditional MSEP can be derived analytically. The main criticism is that the original PIC model does not allow for modeling dependencies between claims payments and incurred losses as it is observed in the data.

In this paper, we generalize the original PIC model so that it allows for modeling dependence between claims payments and incurred losses data. This is motivated by the fact that on the one hand, claims payments are contained in incurred losses data and on the other hand, incurred losses contain additional information, which influences future claims payments data. The data in our example (see Tables 7.8 and 7.9) confirms this hypothesis of dependence between claims payments and incurred losses data (see Table 6.1).

We have seen in the sensitivity analysis that dependence modeling in the PIC method has a crucial impact on the claims reserves and the corresponding (conditional) MSEP (see Tables 6.3 and 6.4). Therefore, the classical PIC model of Merz–Wüthrich [46] underestimates the prediction uncertainty, see Table 6.4, due to the missing dependence structure within accident years. For a better understanding of the influence of prior choices on the reserves and its uncertainty it might be useful to provide a sensitivity analysis of the method to the choice of priors, which should be subject to an extended case study in future work.

Summarizing, the benefits of the PIC method with dependence modeling are that

- two different channels of information are combined to get a unified ultimate loss predictor;
- dependence structures between paid and incurred data can be modeled appropriately;
- prior expert knowledge can be used to design the prior distributions of the parameter vector $\mathbf{\Theta}$, otherwise we can choose non-informative priors for $\mathbf{\Theta}$. Prior expert opinion should also be used for the design of appropriate correlation structures;
- we can calculate the ultimate claim and the conditional MSEP analytically;
- the CDR prediction uncertainty can be calculated, see Happ–Wüthrich [30] or Chapter 5;
- the full predictive distribution can be derived via MC simulations. This allows for the

calculation of any risk measure like VaR or ES.

Main results of the PIC reserving method with dependence:

For solvency considerations in Chapter 7 we summarize all quantities of interest derived in the PIC reserving method:

1. The predictor $\widehat{\mathcal{R}}^I$ for outstanding loss liabilities $\mathcal{R}^I$, see (2.5c) and (2.4c) given by

$$\widehat{\mathcal{R}}^I = \sum_{i=1}^{J} \left(\mathrm{E}[\, I_{i,J} |\, \mathcal{D}_J] - P_{i,J-i} \right), \tag{6.20}$$

   see Corollary 6.6 and the predictor for incremental claim payments

$$\widehat{S}_{i,k+1}^{0|J,Bayes} := \mathrm{E}[\, P_{i,k+1} |\, \mathcal{D}_J] - \mathrm{E}[\, P_{i,k} |\, \mathcal{D}_J] \,. \tag{6.21}$$

2. The estimator for the prediction uncertainty in terms of the (conditional) MSEP

$$\mathrm{msep}_{\mathcal{R}^I|\mathcal{D}_J} \left[\widehat{\mathcal{R}}^I \right] \tag{6.22}$$

   given by Theorem 6.7.

3. The estimator for the CDR uncertainty in terms of the (conditional) MSEP

$$\mathrm{msep}_{\sum_{i=1}^{J} \mathrm{CDR}_i^{J+1} \big| \mathcal{D}_J} [0] \tag{6.23}$$

   can be derived in the same way as in Theorem 5.12.

7 Solvency

Recently, regulatory authorities have established new solvency requirements in order to maintain the long-term function of insurance companies and to protect policyholders from losses. As already mentioned in the introduction the European supervision authority EIOPA will regulate insurance companies in Europe through the Solvency II framework. Solvency II will presumably be obligatory for European insurance companies from January 2016. In Switzerland the regulatory solvency framework *Swiss Solvency Test* (SST) is already implemented since 2006 and obligatory for all insurance companies in Switzerland. In this chapter we consider the issue which conditions are required by the regulator w.r.t. appropriate reserves in solvency frameworks like Solvency II and SST. Moreover, we point out how these requirements can be fulfilled in claims reserving frameworks.

A central task in insurance companies is to build up "appropriate" reserves to meet future loss liabilities and to provide solvency. In the regulatory' point of view reserves should be such that they provide protection against almost all possible adverse events. The International Association of Insurance Supervisors [34] states:

> "*Solvency – 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,...*"

Following this regulatory statement we consider the issue which regulatory conditions are required for appropriate reserves.

In the previous chapters we introduced distribution-free as well as distribution-based claims reserving methods, namely LSRMs in Chapter 4 and the PIC reserving method (with dependence) in Chapters 5 and 6. All these methods provide predictors $\widehat{\mathcal{R}}^I$ for the outstanding loss liabilities $\mathcal{R}^I$ at time I, see (4.77) for LSRMs, (5.13) and (6.20) for the PIC reserving method (with dependence). Of course, depending on the method used for claims reserving (i.e. for the prediction of $\mathcal{R}^I$), there may result rather different reserves. However, the regulatory authorities do not explicitly require the usage of certain claims reserving methods, see European

COMMISSION [23], FOPI [24] and FOPI [25]. As already mentioned, for example, the regulatory authority EIOPA and FINMA only require the reserves to be *best-estimate valuation of liabilities* (BEL).

> *"Mathematisch ausgedrückt sind die versicherungstechnischen Bedarfsrückstellungen eine bedingt erwartungstreue Schätzung des bedingten Erwartungswertes der zukünftigen Zahlungsflüsse aufgrund der zum Zeitpunkt der Schätzung vorliegenden Information. Sie gelten damit als Best-Estimate, sind also weder auf der vorsichtigen noch auf der unvorsichtigen Seite und enthalten insbesondere keine bewussten Verstärkungen."*
> (FINMA Rundschreiben 2008/42)

Remarks 7.1 (BEL)

i) Note that neither in Solvency II nor in the SST exact mathematical definitions are given for BEL. Of course, the conditional expectation $\mathrm{E}\big[\mathcal{R}^I \big| \mathcal{D}^I\big]$ *is the best* $\mathcal{D}^I$*-measurable predictor for* $\mathcal{R}^I$ *when using the conditional MSEP as risk measure. However, many widely used claims reserving methods do not allow for the exact derivation of* $\mathrm{E}\big[\mathcal{R}^I \big| \mathcal{D}^I\big]$*, for example the CL method, the CLR method and the class of (Bayesian) LSRMs, see Chapters 3 and 4.*

ii) The analytical derivation of the conditional expectation $\mathrm{E}\big[\mathcal{R}^I \big| \mathcal{D}^I\big]$ *of the outstanding claims payments at time* I *is possible for some models such as the PIC model (with dependence) in* MERZ–WÜTHRICH *[46] and* HAPP–WÜTHRICH *[31], Hertig's [32] log-normal model and the gamma-gamma model in* WÜTHRICH *[60]. For more models which allow for the analytical derivation of the conditional expectation* $\mathrm{E}\big[\mathcal{R}^I \big| \mathcal{D}^I\big]$ *of the outstanding claims payments at time* I*, see* WÜTHRICH–MERZ *[63].*

Following claims reserving tradition we have used the term *reserves* for the value of the predictor $\widehat{\mathcal{R}}^I$ of outstanding loss liabilities at time I. From now on we only consider predictors $\widehat{\mathcal{R}}^I$ which are BEL. In the following we will use the term *reserves* for the amount (current value of all assets hold by the insurance company) available to cover all insurance liabilities. This amount at time I will be denoted by RES^I. The acceptability of reserves of an insurance company w.r.t regulatory solvency requirements is specified in Definitions 7.2 and 7.6.

7.1 Regulatory Requirements on Reserves

In order to build up reserves for outstanding loss liabilities an insurance company is obliged to predict very carefully all future loss liability cash flows based on the data available. So far we have discussed the concept of BEL. From the regulatory point of view, it is not acceptable to take the amount of BEL^I as reserves at time I. BEL^I only reflects the average outcome of loss liabilities and does not provide any protection against shortfalls in the CDR. At first a consensus has to be found and formulated which functions and requirements reserves for outstanding loss liabilities should fulfill. There is a general agreement that reserves should support the short- and longterm function and solvency of an insurance company. In the short run reserves are required to provide financial strength and solvency in the current accounting year I, i.e. reserves should be higher than the "current value" of liabilities. The Solvency II guideline [23] states:

> *"Liabilities should be valued at the amount for which they could be transferred, or settled, between knowledgeable willing parties in an arm's length transaction."*

This requirement will be later referred to as *accounting condition*, see Wüthrich–Merz [62], and ensures that there are sufficient reserves to transfer the run-off to a knowledgeable third party. The BEL does not fulfill this requirement, because no risk-averse market participant would bear the run-off risk for the price of BEL.

7.1.1 Market-Value Margin

A risk-averse market participant is only willing to take over the run-off of the insurance portfolio and to transfer the associated liabilities on his balance sheet, if he is paid the amount of BEL plus a *market-value margin* (MVM) for bearing the run-off risk. Therefore, the MVM accounts for the risk aversion of market participants and is a compensation payment for bearing the run-off risk. The MVM is also often called *risk-margin* (RM), *safety margin* (SF) or *cost-of-capital* (CoC) margin. The regulatory authority FOPI states for the MVM:

> *"The Market Value Margin (MVM) is the additional amount on top of the best estimate which is required by a willing buyer in an arms-length transaction to assume the liabilities the loss reserves are held to meet..."*

In accordance to the last quotation the *fair value of liabilities* (FVL) in accounting year I is defined by

$$\mathrm{FVL}^I := \mathrm{BEL}^I + \mathrm{MVM}^I.$$

In this context the FVL^I is the (pseudo-market) price to transfer the run-off of the insurance portfolio to a third party. The ability of an insurance company to transfer the insurance run-off portfolio to a third party at time I is required by regulatory authorities, see FOPI [24] and FOPI [25].

Definition 7.2 (Accounting Condition) *In accounting year I reserves RES^I fulfill the accounting condition, if the insurance company is able to transfer the insurance run-off portfolio* *to a willing knowledgeable party, i.e.*

$$\mathrm{RES}^I \geq \mathrm{FVL}^I = \mathrm{BEL}^I + \mathrm{MVM}^I.$$

□

By definition of the MVM^I, the FVL^I is the (market) price for which the run-off portfolio can be transferred to a third party in accounting year I. At this point arises the problem that there is no market for trading insurance contract liabilities of insurance run-off portfolios and we can not take the MVM as the difference between market price FVL^I of the run-off portfolio and its BEL^I. Thus, marked-to-model approaches have to be used to determine MVMs. The *International Actuarial Association* (IAA) [33] divides all approaches to the problem of determining (marked-to-model) MVMs into the following four classes:

- Quantile based methods using risk measures like the VaR, the *Conditional Tail Expectation* (CTE) and the ES also called *Conditional VaR* (TVaR).
- Cost-of-Capital (CoC) approach. CoC is defined as the cost of financing the *Solvency Capital Requirements* (SCR) (protection against adverse events, see the definition below) in all future accounting years up to the complete settlement of the insurance portfolio.
- Discount related methods: The MVM is defined as the difference between i) the discounted cash flows using risk-free interest rates and ii) the discounted cash flows using the risk-free interest rate plus a risk-adjustment. Often probability distortions, see WÜTHRICH–MERZ [62], are used to incorporate risk aversion leading to MVMs.
- Simple methods using a fixed percentage of the BEL as MVM.

The IAA prefers among these four categories the CoC approach, since this approach is most risk sensitive and used in life as well as in non-life insurance. However, this approach is the most sophisticated one and is feasible only under restrictive model assumptions or simplifications, see ROBERT [51] and SALZMANN–WÜTHRICH [53]. In recent publications probability distortions and deflators (to model risk-aversion) have been applied to calculate MVM, see WÜTHRICH ET AL. [61] or WÜTHRICH-MERZ [62] in a market-consistent full balance sheet approach.
In the following we present the CoC approach for the calculations of MVMs.

Market-Value Margin in a Cost-of-Capital Approach

In Section 2.6 we saw that the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ constitutes the amount by which BEL^I is to be adjusted at time $I+1$ to have best-estimate valuation of liabilities BEL^{I+1} also in accounting year $I+1$. In order to protect against shortfalls in the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ the insurance company needs to hold sufficient reserves. These required reserves are determined by a conditional (given $\mathcal{D}^I$) risk measure ρ_I which evaluates the shortfall risk of the claims development result $\mathrm{CDR}^{\mathcal{M},I+1}$ and provides protection against shortfalls. For the definition of conditional risk measures, let $(\Omega, \mathcal{D}, \mathbb{D}, \mathbb{P})$ be a filtered probability space with σ-field $\mathcal{D}$ and filtration $\mathbb{D} := (\mathcal{D}^n)_{0 \le n \le I+J}$ with $\mathcal{D}^n \subseteq \mathcal{D}$. Let

$$\mathcal{L}^0(\Omega, \mathcal{D}^n, \mathbb{P}) := \{\, X \mid X \text{ is } \mathcal{D}^n\text{-measurable and } X < \infty \ \mathbb{P}\text{-a.s.}\} \,.$$

be the space of almost sure finite $\mathcal{D}^n$-measurable random variables. For all times $n \in \{I, \dots, I+J-1\}$ from now we define the following conditional risk measures which determine the amount required at time n for protection against shortfalls in the claims development result $\mathrm{CDR}^{\mathcal{M},n+1}$. Let $\mathcal{A}_n \subset \mathcal{L}^0(\Omega, \mathcal{D}, \mathbb{P})$ be such that i) $c_n \in \mathcal{A}_n$ for all $\mathcal{D}^n$-measurable $c_n \in \mathcal{L}^0(\Omega, \mathcal{D}^n, \mathbb{P})$, ii) $X + Y \in \mathcal{A}_n$ for $X, Y \in \mathcal{A}_n$ and iii) $\lambda_n X \in \mathcal{A}_n$ for $X \in \mathcal{A}_n$ and $\mathcal{D}^n$-measurable λ_n.

Definition 7.3 (Conditional risk measure) *A mapping*

$$\rho_n : \mathcal{A}_n \longmapsto \mathcal{L}^0(\Omega, \mathcal{D}^n, \mathbb{P}) : X \longmapsto \rho_n(X)$$

is called conditional risk measure, if it is finite a.s. on $\mathcal{D}^n$-measurable random variables.

□

In a general CoC approach at each time $n \in \{I, \dots, I+J-1\}$ from today (I) to the last year before the final run-off $(I+J-1)$ a conditional risk measure determines the amount ρ_n which the insurance company needs to hold at time n for protection against shortfalls in the claims development result $\mathrm{CDR}^{\mathcal{M},n+1}$ at time $n+1$. However, the insurance company does not hold the amount ρ_n in its books, but the price for providing this capital. The idea behind that is to buy at time n a reinsurance contract for the price (premium) of $c\,\rho_n$ where $c > 0$ is the cost-of-capital rate. According to this reinsurance contract the seller of the contract (another insurance company or an investor) provides the insurance company with the required amount ρ_n at time n. This money will be used to compensate losses in the case of a shortfall in the $\mathrm{CDR}^{\mathcal{M},n+1}$. Alternatively, one can think of the issue of an insurance bond with the nominal value ρ_n with rate of return c at time n. Viewed from time I, this strategy generates CoC cash flows up to the final settlement of the run-off portfolio

$$c\,\rho_I, \ \dots \ , c\,\rho_{I+J-1}. \tag{7.1}$$

Note that in the general CoC setting the CoC loading ρ_n in (7.1) at time $n \in \{I, \dots, I+J-1\}$ is $\mathcal{D}^n$-measurable (i.e. observable only at time n) but must be evaluated at time $I \le n$. This means that the insurance company has to build up reserves $\widehat{\mathrm{CoC}}^I$ at time I for the CoC cash flows in (7.1). By setting $\mathrm{MVM}^{I,CoC} := \widehat{\mathrm{CoC}}^I$ the fair value of liabilities (risk-adjusted reserves) FVL^I at time I is then given by

$$\mathrm{FVL}^I = \mathrm{BEL}^I + \mathrm{MVM}^{I,CoC}. \tag{7.2}$$

Since there are many ways to define the risk measures ρ_n (from very simple to very sophisticated risk measures) in the CoC approach, there are various different ways to calculate risk-adjusted reserves in a CoC setting. For risk measures which reflect the whole run-off risk it is generally far from straightforward to calculate CoC market-value margins, see SALZMANN–WÜTHRICH [53], ROBERT [51] and WÜTHRICH–MERZ [62]. Therefore, in practice the following regulatory solvency approach is used for the calculation of CoC market-value margins. This approach meets the simplicity requirements in practice and is proposed in the SST [25] and discussed in SALZMANN–WÜTHRICH [53].

Definition 7.4 (Regulatory Cost-of-Capital approach) *Let $\phi > 0$ be a constant, $c > 0$ the cost-of-capital rate and*

$$\widehat{\mathcal{R}}^{n|I} := \sum_{m \in \mathcal{M}} \sum_{i=0}^{I} \sum_{k=n-i}^{J-1} \widehat{S}_{i,k+1}^{m|I}$$

be the predicted at time I outstanding loss liabilities after time $n \in \{I, \ldots, I+J-1\}$ and let $\mathbf{S}^{n+1} \in \mathbb{L}^{n+1}$. The regulatory risk measure is then defined by

$$\rho_n := \rho_n(\mathbf{S}^{n+1}) := \phi \, \frac{|\widehat{\mathcal{R}}^{n|I}|}{|\widehat{\mathcal{R}}^{I}|} \, \mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]^{1/2}. \tag{7.3}$$

The regulatory CoC market-value margin is then defined by

$$\mathrm{MVM}^{I,CoC} := \sum_{n=I}^{I+J-1} c\, \rho_n = c\, \phi \, \frac{\sum_{n=I}^{I+J-1} |\widehat{\mathcal{R}}^{n|I}|}{|\widehat{\mathcal{R}}^{I}|} \, \mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]^{1/2}. \tag{7.4}$$

□

Remarks 7.5 (Regulatory Cost-of-Capital approach)

i) $\rho_n = \rho_n(\mathbf{S}^{n+1})$ *with* $n \in \{I \ldots, I+J-1\}$ *is the amount required for protection against CDR shortfalls at time $n+1$. In general ρ_n is $\mathcal{D}^n$-measurable. However, in this regulatory CoC approach ρ_n are $\mathcal{D}^I$-measurable and hence observable at time I for all $n \in \{I, \ldots, I+J-1\}$.*

ii) *ρ_n comprises only risk of the one-year claims development result* $\mathrm{CDR}^{\mathcal{M},I+1}$. *Run-off risks for accounting years $I+1, \ldots, I+J-1$ are not captured.*

iii) *The risk in the claims development result* $\mathrm{CDR}^{\mathcal{M},I+1}$ *viewed from time I is measured by the conditional MSEP, see (7.3), and not by the risk measures VaR or ES proposed in Solvency II and SST, see* EUROPEAN COMMISSION *[23] and* FOPI *[24]. The constant ϕ can be used to compensate for the difference between MSEP and VaR or ES, respectively. If, for example, the claims development result* $\mathrm{CDR}^{\mathcal{M},I+1}$ *is assumed to be conditionally normal distributed*

$$\mathrm{CDR}^{\mathcal{M},I+1}|_{\mathcal{D}^I} \sim \mathcal{N}(\mu, \sigma^2) \tag{7.5}$$

with mean $\mu := 0$ *and variance* $\sigma^2 := \text{msep}_{\text{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]$ *we choose the parameter*

$$\phi := \frac{\Psi^{-1}_{\mu,\sigma^2}(q)}{\text{msep}_{\text{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]^{1/2}} \tag{7.6}$$

for any security level $q \in (0,1)$*, where* Ψ_{μ,σ^2} *denotes the distribution function of a normal distribution with mean* μ *and variance* σ^2*. The choice* (7.6) *replaces in* (7.4) *the MSEP by the risk measure VaR. In the same way, the MSEP in* (7.4) *can be replaced by any other risk measure like ES as required in the SST.*

iv) Equation (7.3) *shows that risk-bearing capital* $\rho_n(\mathbf{S}^{n+1})$ *at time* $n \in \{I, \ldots, I+J-1\}$ *is proportional to the conditional MSEP of the claims development result* $\text{CDR}^{\mathcal{M},I+1}$*. The proportionality factor depends only on the outstanding loss liability predictors* $\widehat{\mathcal{R}}^I$ *given by* (2.5c)*. This makes the regulatory CoC approach widely applicable in insurance practice.*

vi) The simplifications in the regulatory CoC approach (only the claims development result $\text{CDR}^{\mathcal{M},I+1}$ *is considered and the conditional MSEP is used as risk measure) are necessary to make the approach accessible to most claims reserving methods. In the case of a general CoC approach with VaR or ES as risk measure the calculations of MVMs become infeasible for almost all claims reserving methods. For two exceptions under restrictive model assumptions, see* Robert *[51] and Section 10.3.2 in* Wüthrich–Merz *[62].*

For a detailed discussion of more sophisticated CoC approaches we refer to Salzmann–Wüthrich [53] and Wüthrich–Merz [62].

7.1.2 Solvency Capital Requirements

The MVM in the accounting condition (cf. Definition 7.2) guarantuees transferability of the run-off portfolio to a third party in accounting year I. We presented in Definition 7.4 the regulatory CoC approach which is widely used to provide such MVMs and is applicable to all methods (except for MCL method) in this thesis. Additionally, the insurance company has to fulfill the accounting condition also in the next accounting year $I+1$ with "high probability". This condition will be referred to as *insurance contract condition*, see Wüthrich–Merz [62], and maintains that the insurance company has the financial strength to fulfill the accounting condition with high probability also in accounting year $I+1$. The requirement "with high probability" is determined by a conditional risk measure ρ_I based on the information $\mathcal{D}^I$ at time I. Let ρ_I be a conditional risk measure as in Definition 7.3 which is additionally translation invariant, i.e.

$$\rho_I(X+c) = \rho_I(X) + c \text{ for all } X \in \mathcal{A}_{I+1} \text{ and } \mathcal{D}^I\text{-measurable } c, \tag{7.7}$$

see McNeil et al. [41]. Based on this risk measure ρ_I the *solvency capital requirement* (SCR), see European Commission [23] and FOPI [24], for accounting year I is implicitly defined by

$$\mathrm{FVL}^I + \mathrm{SCR}^I := \rho_I\left(S_{I+1}^{\mathcal{M}} + \mathrm{BEL}^{I+1} + \mathrm{MVM}^{I+1}\right). \tag{7.8}$$

By definition the solvency capital requirement SCR^I is the minimum amount needed at time I in addition to the fair value of liabilities FVL^I to pay out in accounting year $I+1$ the loss liabilities $S_{I+1}^{\mathcal{M}}$ and to fulfill the accounting condition with "high probability" measured by ρ_I.

Definition 7.6 (Insurance Contract Condition) *Let ρ_I be a translation invariant (conditional) risk measure. In accounting year I the insurance contract condition is fulfilled, if*

$$\mathrm{RES}^I \geq \mathrm{FVL}^I + \mathrm{SCR}^I = \rho_I\left(S_{I+1}^{\mathcal{M}} + \mathrm{BEL}^{I+1} + \mathrm{MVM}^{I+1}\right),$$

where RES^I denotes the value at time I of all assets hold by the insurance company to meet outstanding loss liabilities.

□

Adding the solvency capital requirements SCR^I to the fair value of liabilities FVL^I creates an additional (beside the MVM) protection against adverse events and maintains that the insurance company holds with high probability (measured by ρ_I) at least the amount of $\mathrm{FVL}^{I+1} = \mathrm{BEL}^{I+1} + \mathrm{MVM}^{I+1}$ at time $I+1$. That means that if the insurance contract condition is fulfilled at time I the insurance company fulfills the accounting condition, see Definition 7.2, at time $I+1$ (in the next accounting year) with high probability. In this way the regulator can establish a condition at time I that provides with high probability that an insurance company has the financial strength to transfer its run-off portfolio to a third knowledgeable party at time $I+1$.

7.1.3 Final Regulatory Reserves

The amount of the solvency capital requirements SCR^I in (7.8) mainly depends on the choice of the risk measure ρ_I. Regulatory authorities propose the following two risk measures:

Value-at-Risk

The Value-at-Risk (VaR) is a quantile-based risk measure and widely used in the banking sector. It also became a central component of new regulatory requirements in Solvency II.

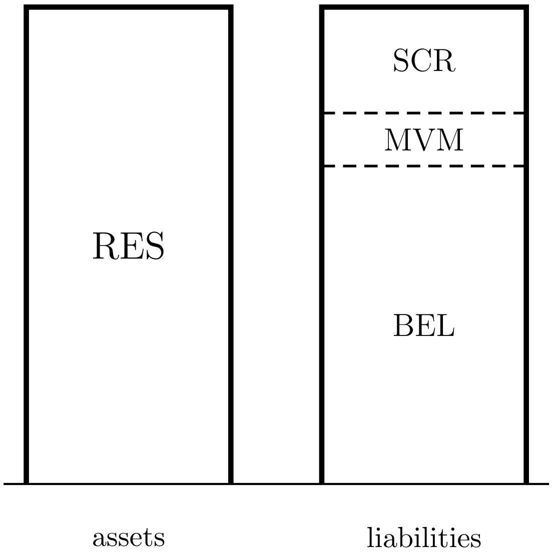

Figure 7.1: Reserves consist of BEL, MVM (together satisfying accounting condition) and SCR (satisfying the insurance contract condition)

Definition 7.7 ((Conditional) Value-at-Risk (VaR)) *Let X be a random variable on a probability space $(\Omega, \mathcal{D}, \mathbb{P})$ and $\mathcal{G} \subseteq \mathcal{D}$ a σ-field. For the distribution of $X|_{\mathcal{G}}$ with distribution function $F_{\mathcal{G}}$ (in the case of their existence) the conditional VaR for security level $\alpha \in (0,1)$ is defined by*

$$\mathrm{VaR}_\alpha[X|\,\mathcal{G}] := \inf\{x \in \mathbb{R} | F_{\mathcal{G}}(x) \geq \alpha\}.$$

□

The VaR with $\alpha = 99.5\%$ will be required in Solvency II calculations for determining appropriate reserves, see European Commission [23]. Often the $\mathrm{VaR}_{99.5\%}$ is associated with an one-in-200-year event. For a detailed discussion on VaR, see Artzner et al. [7].

Expected Shortfall (ES)

The VaR will be required in Solvency II whereas the ES is the risk measure incorporated in the SST framework.

Definition 7.8 ((Conditional) Expected Shortfall (ES)) *Under the assumptions of Definition 7.7 the ES is defined by*

$$\mathrm{ES}_\alpha[X|\,\mathcal{G}] := \frac{1}{1-\alpha} \int_\alpha^1 \mathrm{VaR}_x[X|\,\mathcal{G}]\, dx.$$

□

The ES for security level $\alpha \in (0,1)$ is a normalized average over all $\mathrm{VaR}_x[X|\mathcal{G}]$ with $x \in (\alpha, 1)$. If X describes losses the ES can be interpreted as the average loss in the worst $(1-\alpha)\cdot 100$ % of the cases. The ES is also called *conditional Value-at-Risk* (CVaR). Since the VaR is monotone increasing in α, we obtain $\mathrm{VaR}_\alpha[X|\mathcal{G}] \leq \mathrm{ES}_\alpha[X]$. This implies that for the same security level α the ES is a more conservative risk measure than the VaR. Moreover, the ES has the advantage that it better reflects the tail behaviour of a distribution than the VaR. In contrary to the value-at-risk VaR_α with security level α which is by definition the $(1-\alpha)$-quantile of the distribution $F_\mathcal{G}$ the ES_α also incorporates the shape of the distribution function behind the α-quantile, i.e. $F_\mathcal{G}(x)$ for $x \in (\alpha, 1)$. For details and a further discussion on VaR and ES, see ACERBI-TASCHE [1].

The insurance contract condition in Definition 7.6 is now formulated with the explicit risk measures required under Solvency II and SST:

- Solvency II: $\mathrm{VaR}_{99.5\%}$ is required as risk measure. This means that

$$\mathrm{RES}^I \geq \mathrm{VaR}_{99.5\%}\left[\left(S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1} + \mathrm{MVM}^{I+1}\right)\middle|\,\mathcal{D}^I\right] \tag{7.9}$$

  At time I only a one-in-200-year event leads to a financial situation where accounting condition in Definition 7.2 can not be fulfilled at time $I+1$.

- SST: The risk measure $\mathrm{ES}_{99\%}$ is required, i.e.

$$\mathrm{RES}^I \geq \mathrm{ES}_{99\%}\left[\left(S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1} + \mathrm{MVM}^{I+1}\right)\middle|\,\mathcal{D}^I\right]. \tag{7.10}$$

  Only an event worse than the average loss in the worst 1 % of the cases leads to a financial situation where accounting condition in Definition 7.2 can not be fulfilled at time $I+1$.

However, the exact calculation of the right hand side in (7.9) and (7.10) is infeasible in most claims reserving methods, because many claims reserving methods allow only for the calculation of the first two moments (or appropriate estimates) of $\mathcal{R}^I$ and $\mathrm{CDR}^{\mathcal{M},I+1}$. The knowledge of the distribution of $S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1} + \mathrm{MVM}^{I+1}$ would allow for the calculation of the right hand side of (7.9) and (7.10). Unfortunately, this distribution is unknown in most claims reserving methods.

This shows that the regulatory requirements (accounting condition in Definition 7.2 and insurance contract condition in Definition 7.6) for reserves exceed the possibilities of most current claims reserving methods. Therefore, an agreement has to be found which simplifications are permitted to make these requirements accessible to current claims reserving methods.

7.1.4 Simplifications for Regulatory Solvency Requirements

We consider proposals for simplifications given in the SST, see FOPI [24] and FOPI [25], to make current claims reserving methods accessible for the accounting and insurance contract conditions in Definitions 7.2 and 7.6. As already mentioned in the beginning of this chapter all methods presented in this thesis, namely the class of (Bayesian) LSRMs in Chapter 4 and the PIC reserving method (with dependence) in Chapters 5 and 6 allow for the derivation of the following risk characteristics and the MVM, according to the regulatory CoC approach in

	predictor	prediction uncertainty
$\mathcal{R}^I$	$\widehat{\mathcal{R}}^I$	$\text{msep}_{\mathcal{R}^I \mid \mathcal{D}^I}\left[\widehat{\mathcal{R}}^I\right]$
$\text{CDR}^{\mathcal{M},I+1}$	0	$\text{msep}_{\text{CDR}^{\mathcal{M},I+1} \mid \mathcal{D}^I}\left[0\right]$

(7.11)

Definition 7.4.

Simplification I

Viewed from time I, it is difficult to derive the distribution of the market-value margin MVM^{I+1} at time I, if the method used for calculating the MVM is sophisticated (for example a general CoC approach). Therefore, we approximate MVM^{I+1} by omitting in (7.4) the first summand, i.e.

$$\text{MVM}^{I+1} \simeq \widehat{\text{MVM}}^{I+1} := c\,\phi\, \frac{\sum_{n=I+1}^{I+J-1} |\widehat{\mathcal{R}}^{n|I}|}{|\widehat{\mathcal{R}}^I|}\, \text{msep}_{\text{CDR}^{\mathcal{M},I+1} \mid \mathcal{D}^I}\left[0\right]^{1/2}. \tag{7.12}$$

Since $\widehat{\text{MVM}}^{I+1}$ is $\mathcal{D}^I$-measurable we obtain with (7.12) for conditional risk measures, which are translation invariant, e.g. VaR and ES, for the right hand side of (7.8)

$$\rho_I\left(S_{I+1}^{\mathcal{M}} + \text{BEL}^{I+1} + \text{MVM}^{I+1}\right) \simeq \rho_I\left(S_{I+1}^{\mathcal{M}} + \text{BEL}^{I+1}\right) + \widehat{\text{MVM}}^{I+1}. \tag{7.13}$$

Simplification II

Applying Simplification I it remains to calculate

$$\rho_I\left(S_{I+1}^{\mathcal{M}} + \text{BEL}^{I+1}\right) \tag{7.14}$$

for $\rho_I = \text{VaR}_{99.5\%}$ and $\rho_I = \text{ES}_{99\%}$. The simplification strategy for the calculation of (7.14) is as follows:
Viewed from time I most claims reserving methods allow for the derivation of the first two moments of the quantity $S_{I+1}^{\mathcal{M}} + \text{BEL}^{I+1}$. The unknown distribution of the quantity $S_{I+1}^{\mathcal{M}} + \text{BEL}^{I+1}$ is approximated by a log-normal distribution with the same first two moments using the methods of moments for calibration. This is illustrated in Figure 7.2. The risk measures $\rho_I = \text{VaR}_{99.5\%}$ and $\rho_I = \text{ES}_{99\%}$ of the log-normal distribution can then be calculated easily. This

approximation method simplifies the calculation of the right hand side of (7.14) to the task of calculating (or estimating) the first two (conditional) moments of the quantity $S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1}$: Let us assume that the model used by the reserving actuary allows for the calculation of the conditional expectation of outstanding loss liabilities

$$\mathrm{BEL}^n = \widehat{\mathcal{R}}^n := \mathrm{E}[\mathcal{R}^n | \mathcal{D}^n] \text{ for } n \in \{I, \ldots, I+J-1\}, \tag{7.15}$$

i.e. the outstanding loss liability predictors are best predictors. The tower property of conditional expectations (cf. Williams [59]) then implies

$$\begin{aligned} \mathrm{E}\left[S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1} \middle| \mathcal{D}^I \right] &= \widehat{\mathcal{R}}^I = \mathrm{BEL}^I \\ \mathrm{Var}\left[S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1} \middle| \mathcal{D}^I \right] &= \mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]\,. \end{aligned} \tag{7.16}$$

For the PIC reserving method (with dependence) the condition in (7.15) is fulfilled, see (5.13) in Chapter 5 and (6.20) in Chapter 6 and hence we obtain the first two (conditional) moments of the quantity $S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1}$ by (7.16) (for the detailed structure, see (5.15) and (6.23)). For the class of Bayesian LSRMs we approximate the right hand side of (7.16) by corresponding estimates given in (4.77) and (4.79). Then we approximate (7.14) by

$$\rho_I\left(S^{\mathcal{M}}_{I+1} + \mathrm{BEL}^{I+1}\right) \simeq \rho_I(X) \quad \text{with} \quad X \sim \mathcal{LN}\left(\mathrm{BEL}^I, \widehat{\mathrm{msep}}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}[0]\right),$$

where $\mathcal{LN}$ denotes a log-normal distribution.

7.2 Example for Regulatory Reserves

We revisit the example given for the Bayesian LSRMs in Chapter 4. In this example we consider an insurance company with three insurance portfolios 1, 2 and 3, i.e. $m \in \{0, 1, 2\}$ in the LSRM terminology, where each portfolio corresponds to an individual LoB of the insurance company, see Tables 7.5, 7.6 and 7.7. In this section we calculate reserves which meet central regulatory solvency requirements in the SST. The best-estimate valuation of liabilities resulting from the Bayesian LSRM are given in Table 7.1. The overall best-estimate valuation of liabilities are then given by aggregation over $m \in \{0, 1, 2\}$. For the calculation of the market-value margin MVM^I we apply the regulatory CoC approach given in Definition 7.4 and as risk measure for the solvency capital requirements SCR^I we use the expected shortfall $\mathrm{ES}_{99\%}$ under Approximations I–II as proposed in the SST. In the run-off trapezoids in Tables 7.5–7.7 there are 21 accident years, i.e. $I = 20$, and we consider 11 development years, i.e. $J = 10$.

Best-Estimate Valuation of Liabilities

For the calculation of the best-estimate valuation of liabilities BEL^{20} at time 20 we apply the specific Bayesian LSRM used in Chapter 4. At time 20 the best-estimate valuation of liabilities

LoB 1	0	1	2	3	4	5	6	7	8	9	10	BEL_0^I
0	118	369	745	34	0	131	0	0	95	0	0	0
1	124	533	206	27	24	2	25	0	0	-76	0	0
2	556	1648	1290	-496	-15	35	-560	0	12	0	0	0
3	1646	705	141	15	105	0	-4	-853	0	0	0	0
4	317	569	4	0	60	40	0	0	0	0	0	0
5	242	677	299	6	5	20	0	0	0	0	0	0
6	203	409	10	17	28	-20	0	0	0	0	0	0
7	492	913	280	-17	85	-11	62	0	0	0	0	0
8	321	828	579	135	14	0	0	0	0	0	0	0
9	609	500	174	11	-41	2	0	0	0	0	0	0
10	492	1135	-5	50	0	0	0	0	-51	0	0	0
11	397	396	75	21	75	0	0	0	0	0	0	0
12	523	575	377	14	0	0	0	0	0	-8	0	-8
13	1786	1165	419	-341	182	78	36	0	8	-12	0	-4
14	241	224	71	60	56	0	0	-30	3	-5	0	-32
15	327	295	-45	6	0	0	-29	-46	5	-7	0	-77
16	275	245	9	0	12	13	-21	-33	4	-5	0	-43
17	89	238	51	4	19	9	-15	-23	2	-4	0	-11
18	295	6	95	-11	16	8	-12	-20	2	-3	0	-20
19	151	255	104	-11	16	8	-12	-20	2	-3	0	84
20	315	287	110	-11	17	8	-13	-21	2	-3	0	376
Total												267

LoB 2	0	1	2	3	4	5	6	7	8	9	10	BEL_1^I
0	92	350	99	0	-13	0	0	0	0	0	0	0
1	451	626	8	93	34	5	0	0	0	0	0	0
2	404	313	117	15	0	1	0	0	0	0	0	0
3	203	369	241	62	3	32	2	184	-7	0	-103	0
4	352	482	214	24	16	0	0	0	0	0	0	0
5	504	742	26	81	-68	0	0	0	0	0	0	0
6	509	499	53	0	0	10	0	0	0	0	0	0
7	229	351	50	40	2	0	0	0	0	0	0	0
8	324	491	56	-12	8	-90	0	0	0	0	0	0
9	508	297	101	63	2	0	0	0	0	0	0	0
10	354	287	192	9	0	0	0	0	0	0	0	0
11	431	416	7	61	3	0	0	0	0	0	-8	-8
12	205	625	148	56	14	0	0	0	0	0	-11	-11
13	522	612	-70	138	0	8	0	0	-1	0	-18	-19
14	567	358	-10	42	-4	0	0	7	0	0	-7	-1
15	1238	686	110	-137	0	0	0	10	-1	0	-10	-1
16	355	648	134	27	32	-2	0	7	0	0	-7	-2
17	312	368	2	4	1	-1	0	5	0	0	-5	-1
18	246	106	66	15	1	-1	0	4	0	0	-4	14
19	91	327	42	15	1	-1	0	4	0	0	-4	56
20	130	263	44	16	1	-1	0	4	0	0	-5	322
Total												349

LoB 3	0	1	2	3	4	5	6	7	8	9	10	BEL_2^I
0	268	188	29	-2	0	0	0	0	0	0	0	0
1	268	252	57	2	0	0	0	0	0	0	0	0
2	385	583	49	2	0	0	0	0	0	0	0	0
3	251	491	53	136	0	0	0	0	0	0	0	0
4	456	449	257	2	0	0	0	0	27	0	0	0
5	477	809	90	0	-3	0	0	0	0	0	0	0
6	405	594	173	24	0	14	0	0	0	0	0	0
7	443	489	20	13	19	8	20	0	0	0	0	0
8	477	569	290	26	13	0	0	0	0	0	0	0
9	581	565	170	46	29		0	0	0	0	0	0
10	401	596	232	19	33	3	-20	0	0	0	0	0
11	474	304	161	382	45	26	0	0	0	0	0	0
12	649	771	287	2	0	0	0	-71	0	0	0	0
13	911	1024	369	3	2	0	0	53	4	0	0	4
14	508	546	47	-30	0	0	0	-1	2	0	0	1
15	389	401	78	41	660	0	0	-1	2	0	0	1
16	373	625	93	64	46	2	0	-1	2	0	0	3
17	276	577	79	16	26	2	0	-1	1	0	0	28
18	465	355	39	19	22	1	0	-1	1	0	0	43
19	343	279	69	19	22	1	0	-1	1	0	0	111
20	254	285	73	20	23	2	0	-1	1	0	0	403
Total												594

Table 7.1: Predicted incremental claim information for LoB 1, 2 and 3

BEL for each LoB is given in Table 7.1. The overall BEL then result as

$$\begin{aligned} \mathrm{BEL}^{20} &= \mathrm{BEL}_0^{20} + \mathrm{BEL}_1^{20} + \mathrm{BEL}_2^{20} \\ &= 267 + 349 + 594 \\ &= 1210, \end{aligned} \tag{7.17}$$

where BEL_0^{20}, BEL_1^{20} and BEL_2^{20} denote the best-estimate valuation of liabilities for LoB 1, 2 and 3, respectively. In order to fulfill the accounting condition in Definition 7.2, which guarantees transferability of the run-off portfolio to a third party, we add to the BEL^{20} in (7.17) a market-value margin MVM^{20}.

Market-Value Margin

As proposed in the SST guidelines we use the regulatory CoC approach given in Definition 7.4 with a cost-of-capital rate $c = 6\%$ to calculate the MVM. We choose the parameter $\phi = 3$, since this choice provides $\mathrm{VaR}_{99.5}\left[\mathrm{CDR}^{\mathcal{M},I+1}\right] \simeq 3 \cdot \mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}\left[0\right]^{1/2}$ in the case that $\mathrm{CDR}^{\mathcal{M},I+1}$ is normally distributed with zero mean and variance $\sigma^2 = \mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}\left[0\right]$. For the calculation of the MVM according to Definition 7.4 we apply the expected pattern of BEL given by Table 7.2. The identity $\mathrm{msep}_{\mathrm{CDR}^{\mathcal{M},I+1}|\mathcal{D}^I}\left[0\right]^{1/2} = 646$ is given in (7.18). We obtain for

n	20	21	22	23	24	25	26	27	28	29
$\widehat{\mathcal{R}}^{n\|I}$	1210	136	-86	-100	-108	-77	-39	-17	-12	-5

Table 7.2: Expected pattern of BEL for calendar years $n = 20, \ldots, 29$

the MVM at time $I = 20$ calculated by the regulatory CoC approach

$$\begin{aligned} \mathrm{MVM}^{20,CoC} &= \sum_{n=20}^{29} 0.06 \cdot 3 \cdot \rho_n \\ &= 0.06 \cdot 3 \cdot \frac{(1210 + 136 + 86 + 100 + 108 + 77 + 39 + 17 + 12 + 5)}{1210} \cdot 646 \\ &= 172. \end{aligned}$$

This implies for the fair value of liabilities FVL^{20} at time $I = 20$

$$\begin{aligned} \mathrm{FVL}^{20} &= \mathrm{BEL}^{20} + \mathrm{MVM}^{20,CoC} \\ &= 1210 + 172 \\ &= 1382. \end{aligned}$$

This fair value of the run-off liabilities (in a marked-to-model view) is the price the run-off portfolio can be transferred to a third willing and knowledgeable party. Hereby it is assumed that both the insurance company and the third party (investor) agree to use the same claims reserving method for calculating the best-estimate valuation of liabilities BEL^{20} and the regulatory CoC approach for the calculation of the market-value margin MVM^{20}.

Solvency Capital Requirements

We follow the approximation outline given in the SST to derive the SCR. In the specific Bayesian LSRM under consideration we obtain

$$\begin{aligned} \mathrm{BEL}^{20} &= 1210, \\ \widehat{\mathrm{msep}}_{\mathrm{CDR}^{\mathcal{M},21}|\mathcal{D}^{20}}\left[0\right]^{1/2} &= \sqrt{480^2+207^2+240^2+294^2} = 646, \end{aligned} \tag{7.18}$$

where 480, 207 and 240 are the square roots of $\widehat{\mathrm{msep}}_{\mathrm{CDR}^{0,21}|\mathcal{D}^{20}}[0]$, $\widehat{\mathrm{msep}}_{\mathrm{CDR}^{1,21}|\mathcal{D}^{20}}[0]$ and $\widehat{\mathrm{msep}}_{\mathrm{CDR}^{2,21}|\mathcal{D}^{20}}[0]$, respectively and 294 corresponds to the covariance terms between the different LoBs (cf. Estimator 4.25).

Now we fit by the method of moments a log-normally distributed random variable X to these empirical moments. This leads to

$$\begin{aligned} \mathrm{E}[X] &= \exp\left\{\mu + \frac{\sigma^2}{2}\right\} = 1210 \\ \mathrm{Var}[X] &= \exp\{2\mu+\sigma^2\}\left(\exp\{\sigma^2\}-1\right) = 646^2 \end{aligned}$$

which imply $\widehat{\mu} = 6.973024$ and $\widehat{\sigma} = 0.500702$. The corresponding density $f(x)$ of X is then given by

$$f(x) = \begin{cases} \frac{1}{0.500702\sqrt{2\pi}x}\exp\left\{\frac{-(\log(x)-6.973024)^2}{2\cdot 0.500702^2}\right\} & \text{for} \quad x>0 \\ 0 & \text{for} \quad x\le 0, \end{cases}$$

see Figure 7.2 for an illustration. Then the ES can be calculated by the well-known formula

$$\mathrm{ES}_\alpha[X] = \frac{1}{1-\alpha} e^{\mu+\sigma^2/2}\Phi(\sigma-\Phi^{-1}(\alpha)),$$

where $\alpha \in (0,1)$ denotes a security level and Φ the distribution function of the standard normal distribution. Using the security level $\alpha = 99\%$ and the parameter estimates $\widehat{\mu} = 6.973024$ and $\widehat{\sigma} = 0.500702$ resulting from the method of moments we obtain

$$\mathrm{ES}_{99\%}[X] = 4108.2.$$

By Definition 7.6 we obtain for the solvency capital requirements

$$\begin{aligned} \mathrm{SCR}^{20} &= \rho_{20}\left(S_{21}^{\mathcal{M}} + \mathrm{BEL}^{21} + \mathrm{MVM}^{21}\right) - \mathrm{FVL}^{20} \\ &\simeq \rho_{20}\left(S_{21}^{\mathcal{M}} + \mathrm{BEL}^{21}\right) + \widehat{\mathrm{MVM}}^{21} - \mathrm{FVL}^{20} \\ &= \mathrm{ES}_{99\%}\left[S_{21}^{\mathcal{M}} + \mathrm{BEL}^{21}\middle|\,\mathcal{D}^{20}\right] + \widehat{\mathrm{MVM}}^{21} - \mathrm{FVL}^{20} \\ &\simeq 4108.2 + 55.7 - 1382 \\ &= 2781.9. \end{aligned} \tag{7.19}$$

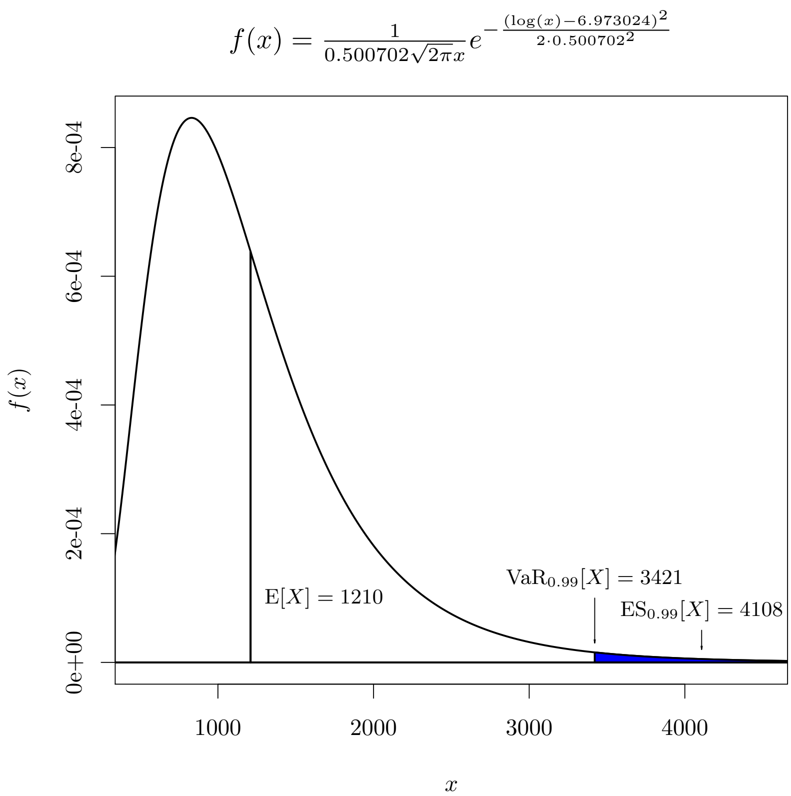

Figure 7.2: The calibrated log-normal distribution with $\widehat{\mu} = 6.973024$ and $\widehat{\sigma} = 0.500702$ used as an approximation for the distribution of the quantity $S_{21}^{\mathcal{M}} + \text{BEL}^{21}$ and corresponding expected value, VaR and ES for the security level $\alpha = 0.99$

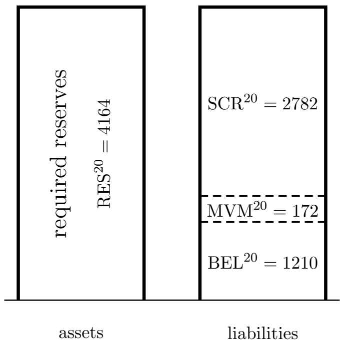

Figure 7.3: Best-estimate valuation of liabilities BEL^{20}, market-value margin MVM^{20} (together satisfying accounting condition) and solvency capital requirements SCR^{20} (satisfying the insurance contract condition) leading to the overall reserves

Reserves

Summing up the results for best-estimate valuation of liabilities BEL^{20}, the market-value margin MVM^{20} and the solvency capital requirements SCR^{20} we obtain

$$\begin{aligned} \mathrm{RES}^{20} &= \mathrm{BEL}^{20} + \mathrm{MVM}^{20} + \mathrm{SCR}^{20} \\ &= 1210 + 172 + 2781.9 \\ &= 4163.9. \end{aligned}$$

This means that the overall reserves RES^{20} must be at least 4137.1 to fulfill the regulatory solvency requirements (accounting condition, see Definition 7.2, and insurance contract condition, see Definition 7.6). For an illustration, see Figures 7.2 and 7.3. At this stage it becomes obvious that the Simplifications I–II are crucial for the calculation of MVM and SCR in all distribution-free claims reserving methods.

At a first glance the required amount of $\mathrm{RES}^{20} = 4163.9$ seems very high and we will analyze this fact in more detail. The biggest loading of the reserves RES^{20} is contributed by the solvency capital requirements $\mathrm{SCR}^{20} = 2781.9$ and the expected shortfall $\mathrm{ES}_{99\%}\left[S_{21}^{\mathcal{M}} + \mathrm{BEL}^{21} \middle| \mathcal{D}^{20} \right]$ of the log-normal distribution is the main risk driver of this quantity, see (7.19). The crucial point is that the ES of the heavy-tailed log-normal distribution mainly depends on the variance of the log-normal distribution. In (7.16) the variance is replaced by the conditional MSEP of the Bayesian LSRM. In our specific Bayesian LSRM in Chapter 4 the square root of the conditional MSEP is about 50% of the predicted outstanding loss liabilities, see (7.18), what is much higher than in many other claims reserving methods. This is due to the fact that the Bayesian LSRM is chosen in such a way that all three LoBs are highly correlated, for details see Section 4.3. Hence the associated MSEP of the CDR is much higher then in most other claims reserving methods leading to a high SCR. This explains the high value for the reserves RES^{20}.

Conclusions and Outlook

In this thesis we considered the problem of claims reserving as one of the main actuarial tasks in non-life insurance practice. In Chapter 2 we started with the introduction of typical loss liability cash flows (claims payments) generated by classical non-life insurance run-off portfolios. For a complete risk assessment of these run-off portfolios all claims payments at any time in the future have to be predicted based on all relevant information available at time of prediction (prediction step). In Chapter 3 we gave a brief introduction in classical widely-used distribution-free claims reserving methods. We saw that most of these classical methods are very limited w.r.t. the information sources they can incorporate. To solve this problem we presented three claims reserving methods which can cope with various different sources of information.

Model I: At first we considered the class of LSRMs in Chapter 4. This model class was recently presented by Dahms [17] and generalizes almost all distribution-free claims reserving methods, given in Chapter 3. However, expert knowledge w.r.t. the claims development pattern can not be included in a mathematically consistent way. We considered LSRMs in a Bayesian model setup and approximated the Bayes predictors by their corresponding credibility based predictors. This led to the class of Bayesian LSRMs. This model class additionally allows for including expert knowledge/external data w.r.t. the development pattern. Such credibility based methods are often applied for pricing problems, if there is only a scarce data base available and extern prior knowledge is to be included.

Models II and III: Beside the class of distribution-free claims reserving methods there are various approaches to claims reserving based on distributional assumptions. An important representative among them is the PIC reserving method introduced in Merz-Wüthrich [46] which allows to combine paid and incurred data simultaneously. These data sources are often available in insurance practice and hence should be utilized for prediction purposes. We recapitulated this model and showed how the one-year CDR prediction uncertainty can be quantified. The classical PIC reserving method assumes the paid and incurred ratios to be independent. Therefore, in a second step, we generalized the classical PIC reserving method in that way that it respects dependence structures often observed in practice in paid and incurred data. This led to the PIC reserving method with dependence modeling.

Concluding, we considered in this thesis three methods: The Bayesian LSRMs, PIC reserving method and PIC reserving method with dependence modeling. All three models allow for the incorporation of data sources often available in insurance practice and hence take account for the requirement that all predictions should be based on all data available. Moreover, all three methods allow for the derivation of best-estimate valuations of liabilities BEL and estimates for the ultimate claim as well as the CDR prediction uncertainty. On top of that, the PIC reserving method (with dependence modeling) provides the whole predictive distribution of the ultimate claim and CDR. As we pointed out in Chapter 7 BEL, MVM and SCR can be calculated in each model based on the BEL and the CDR prediction uncertainty.

We state two questions to be answered in future research:

1. The (Bayesian) LSRMs is a large class of distribution-free claims reserving methods. Therefore, it would be helpful to have a criterion for model selection. Having such a criterion the specific LSRM is chosen which provides the best fit to the data w.r.t. this model selection criterion. A similar problem remains in the PIC reserving method with dependence. For the explicit choice of the covariance matrix $\mathbf{V}$ an estimator has to be found which is optimal w.r.t. to some optimality criterion.

2. All calculations in the three methods are provided on a nominal scale, i.e. time value of money (stochastic discounting) is not considered. Of course, this in-line with the status quo of classical claims reserving literature but does not accommodate recent developments of market-consistent valuation techniques, see WÜTHRICH–MERZ [62]. Therefore, we put the question: To what extend these methods can be generalized to a full market-consistent valuation approach ?

   All three methods allow for the calculation of BEL. Based on these BEL the valuation portfolio, see WÜTHRICH–MERZ [62], can be calculated for each method leading to a market-consistent BEL. The assessment of the prediction uncertainty in terms of the MSEP for the ultimate claims and for the CDR in the market-consistent valuation setup turns out to be more sophisticated.

   For the wide model class of (Bayesian) LSRMs it seems possible to extend this model class w.r.t. market-consistent valuation. However, to the best of our knowledge no scientific contribution considering this issue exists so far. Making LSRMs accessible to market-consistent valuation would allow to consider most distribution-free claims reserving methods (CL, BF, CLR methods), widely used in practice, in a market-consistent valuation approach.

   Moreover, for the PIC reserving method (with dependence modeling) and for most so far existing distributional claims reserving methods the question how these methods can be embedded in a valuation approach should be subject to further research.

Accompanying Publications

Parts of this thesis have already been published:

I:

Credibility for the Linear Stochastic Reserving Methods

René Dahms   Sebastian Happ

Submitted

II:

Claims Developement Result in the Paid-Incurred Chain Reserving Method

Sebastian Happ   Michael Merz   Mario V. Wüthrich

Insurance: Mathematics and Economics, Volume 51, Issue 1, July 2012, pp. 66–72

III:

Paid–Incurred Chain Reserving Method with Dependence Modeling

Sebastian Happ   Mario V. Wüthrich

Astin Bulletin, Volume 43, Issue 1, January 2013, pp. 1–20

Authors information

René Dahms: rene.dahms@math.ethz.ch

Sebastian Happ: sebastian.happ@wiso.uni-hamburg.de

Michael Merz: michael.merz@wiso.uni-hamburg.de

Mario V. Wüthrich: mario.wuethrich@math.ethz.ch

Bibliography

[1] Acerbi, C. & Tasche, D. (2002): Expected shortfall: a natural coherent alternative to Value-at-Risk. *Economic notes*, 31(2): 379–388.

[2] AISAM-ACME (2007): AISAM-ACME study on non-life long tail liabilities. Reserve risk and risk margin assessment under Solvency II, 17 October 2007.

[3] Alai, D. H., Merz, M. & Wüthrich, M. V. (2009): Mean square error of prediction in the Bornhuetter-Ferguson claims reserving method. *Annals of Actuarial Science*, 4(1): 7–31.

[4] Alai, D. H., Merz, M. & Wüthrich, M. V. (2011): Prediction uncertainty in the Bornhuetter-Ferguson claims reserving method: revisited. *Annals of Actuarial Science*, 5(1): 7–17.

[5] Antonio, K. & Beirlant, J. (2007): Actuarial statistics with generalized linear mixed models. *Insurance: Mathematics and Economics*, 40(1): 58–76.

[6] Arbenz, P. & Salzmann, R. (2011): On a combination of multiplicative and additive stochastic loss reserving methods, Pre–print, ETH Zürich.

[7] Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999): Coherent measures of risk. *Mathematical finance*, 9(3): 203–228.

[8] Bornhuetter, R. & Ferguson, R. (1972): The actuary and IBNR. *Proc. CAS*, LIX: 181–195.

[9] Brockwell, P. J. & Davis, R. A. (2009): *Time Series: Theory and Methods*. Berlin: Springer.

[10] Buchwalder, M., Bühlmann, H., Merz, M. & Wüthrich, M. (2006): The mean square error of prediction in the chain ladder reserving method. *ASTIN Bulletin*, 36(2): 521–542.

[11] Bühlmann, H. (1983): Estimation of IBNR reserves by the methods of chain ladder, cape cod and complementary loss ratio, International Summer School.

[12] Bühlmann, H. (1983): Chain ladder, cape cod and complementary loss ratio, International Summer School, unpublished.

[13] Bühlmann, H., De Felice, M., Gisler, A., Moriconi, F. & Wüthrich, M. V. (2009): Recursive credibility formula for chain ladder factors and the claims development result. *ASTIN Bulletin*, 39(1): 275–306.

[14] Bühlmann, H. & Gisler, A. (2005): *A Course in Credibility Theory and its Applications.* Berlin: Springer.

[15] Dahms, R. & Happ, S. (2013): Credibility for the linear stochastic reserving methods, Submitted.

[16] Dahms, R. (2008): A loss reserving method for incomplete claim data. *Bulletin Swiss Association of Actuaries*, 127–148.

[17] Dahms, R. (2012): Linear stochastic reserving methods. *ASTIN Bulletin*, 42(1): 1–34.

[18] Dahms, R., Merz, M. & Wüthrich, M. V. (2009): Claims development result for combined claims incurred and claims paid data. *Bulletin Francais d'Actuariat*, 9(18): 5–39.

[19] De Jong, P. (2012): Modeling dependence between loss triangles. *North American Actuarial Journal*, 16(1): 74–86.

[20] Diers, D. & Linde, M. (2013): The multi-year non-life insurance risk in the additive loss reserving model. *Insurance: Mathematics and Economics*, 52(3): 590–598.

[21] England, P. D. & Verrall, R. J. (2002): Stochastic claims reserving in general insurance. *British Actuarial Journal*, 8(3): 443–518.

[22] England, P. & Verrall, R. (1999): Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics*, 25(3): 281–293.

[23] European Commission (2010): *QIS5 technical specifications, annex to call for advice from CEIOPS on QIS5.*

[24] Federal Office of Private Insurance (2006): *Technical document on the Swiss Solvency Test.*

[25] Federal Office of Private Insurance (2006): *The Swiss experience with market consistent technical provisions - the Cost of Capital approach.*

[26] Gerber, H. (1997): *Life Insurance Mathematics.* Berlin: Springer.

[27] Gisler, A. & Wüthrich, M. V. (2008): Credibility for the chain ladder reserving method. *ASTIN Bulletin*, 38(2): 565–600.

[28] Gogol, D. (1993): Using expected loss ratios in reserving. *Insurance: Mathematics and Economics*, 12(3): 297–299.

[29] Haberman, S. & Renshaw, A. E. (1996): Generalized linear models and actuarial science. *The Statistician*, 407–436.

[30] Happ, S., Merz, M. & Wüthrich, M. V. (2012): Claims development result in the paid-incurred chain reserving method. *Insurance: Mathematics and Economics*, 51(1): 66–72.

[31] Happ, S. & Wüthrich, M. V. (2013): Paid-incurred chain reserving method with dependence modeling. *ASTIN Bulletin*, 43(1): 1–20.

[32] Hertig, J. (1985): A statistical approach to the IBNR-reserves in marine insurance. *ASTIN Bulletin*, 15(2): 171–183.

[33] International Actuarial Association IAA (2009): *Measurement of liabilities for insurance contracts: current estimates and risk margins.*

[34] International Association of Insurance Supervisors (2005): *Glossary of Terms.*

[35] Koller, M. (2012): *Stochastic Models in Life Insurance*. Berlin: Springer.

[36] Kolmogorov, A. N., Fomin, S. V. & Fomin, S. (1999): *Elements of the Theory of Functions and Functional Analysis*. Dover: Dover Publications.

[37] Lang, S. (2002): *Algebra*. New York: Springer.

[38] Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, 23(2): 213–225.

[39] Mack, T. (2008): The prediction error of Bornhuetter/Ferguson. *ASTIN Bulletin*, 38(1): 87–103.

[40] Mack, T. (1999): The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor. *ASTIN Bulletin*, 29(2): 361–366.

[41] McNeil, A. J., Frey, R. & Embrechts, P. (2010): *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton: Princeton university press.

[42] Merz, M. & Wüthrich, M. V. (2013): Estimation of tail development factors in the paid-incurred chain reserving method. *Variance*, 7(1): 61–73.

[43] Merz, M. & Wüthrich, M. V. (2006): A credibility approach to the Munich chain-ladder method. *Blätter der DGVFM*, 27(4): 619–628.

[44] Merz, M. & Wüthrich, M. V. (2012): Full and 1-year runoff risk in the credibility-based additive loss reserving method. *Applied Stochastic Models in Business and Industry*, 28(4): 362–380.

[45] Merz, M. & Wüthrich, M. V. (2008): Modelling the claims development result for solvency purposes. *Casualty Actuarial Society E-Forum*, Fall: 542–568.

[46] Merz, M. & Wüthrich, M. V. (2010): Paid-incurred chain claims reserving method. *Insurance: Mathematics and Economics*, 46(3): 568–579.

[47] Merz, M. & Wüthrich, M. V. (2007): Prediction error of the expected claims development result in the chain ladder method. *Bulletin of Swiss Association of Actuaries*, 1: 117–137.

[48] Murphy, D. (1994): Unbiased loss development factors. *PCAS*, 81: 154–207.

[49] Ohlsson, E. & Lauzeningks, J. (2009): The one-year non-life insurance risk. *Insurance: Mathematics and Economics*, 45(2): 203–208.

[50] Quarg, G. & Mack, T. (2004): Munich chain ladder. *Blätter der DGVFM*, 26(4): 597–630.

[51] Robert, C. Y. (2013): Market value margin calculations under the cost of capital approach within a Bayesian chain ladder framework. *Insurance: Mathematics and Economics*, 53(1): 216–229.

[52] Saluz, A., Gisler, A. & Wüthrich, M. V. (2011): Development pattern and prediction error for the stochastic Bornhuetter-Ferguson claims reserving method. *ASTIN Bulletin*, 41(2): 279–313.

[53] Salzmann, R. & Wüthrich, M. V. (2010): Cost-of-capital margin for a general insurance liability runoff. *ASTIN Bulletin*, 40(2): 415.

[54] Salzmann, R. & Wüthrich, M. V. (2012): Modeling accounting year dependence in runoff triangles. *European Actuarial Journal*, 2(2): 227–242.

[55] Scollnik, D. P. (2001): Actuarial modeling with MCMC and BUGS. *North American Actuarial Journal*, 5(2): 96–124.

[56] Shi, P. & Frees, E. (2011): Dependent loss reserving using copulas. *ASTIN Bulletin*, 41(2): 449–486.

[57] Taylor, G. C. (2000): *Loss Reserving : An Actuarial Perspective*. London: Kluwer Academic Boston.

[58] Taylor, G., McGuire, G. & Greenfield, A. (2003): *Loss Reserving: Past, Present and Future*. Centre for Actuarial Studies, Department of Economics, University of Melbourne.

[59] Williams, D. (1991): *Probability with Martingales*. Cambridge: Cambridge Univ Press.

[60] Wüthrich, M. V. (2010): Accounting year effects modeling in the stochastic chain ladder reserving method. *North American Actuarial Journal*, 14(2): 235–255.

[61] Wüthrich, M. V., Embrechts, P. & Tsanakas, A. (2011): Risk margin for a non-life insurance run-off. *Statistics and Risk Modeling*, 28(4): 299–317.

[62] Wüthrich, M. V. & Merz, M. (2013): *Financial Modeling, Actuarial Valuation and Solvency in Insurance*. Heidelberg: Springer.

[63] Wüthrich, M. V. & Merz, M. (2008): *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley.

[64] Wüthrich, M. V., Merz, M. & Lysenko, N. (2009): Uncertainty of the claims development result in the chain ladder method. *Scandinavian Actuarial Journal*, 2009(1): 63–84.

Data Sets

Data sets used in the contributions of this thesis:

i\ j	0	1	2	3	4	5	6	7	8	9
0	1'216'632	1'347'072	1'786'877	2'281'606	2'656'224	2'909'307	3'283'388	3'587'549	3'754'403	3'921'258
1	798'924	1'051'912	1'215'785	1'349'939	1'655'312	1'926'210	2'132'833	2'287'311	2'567'056	
2	1'115'636	1'387'387	1'930'867	2'177'002	2'513'171	2'931'930	3'047'368	3'182'511		
3	1'052'161	1'321'206	1'700'132	1'971'303	2'298'349	2'645'113	3'003'425			
4	808'864	1'029'523	1'229'626	1'590'338	1'842'662	2'150'351				
5	1'016'862	1'251'420	1'698'052	2'105'143	2'385'339					
6	948'312	1'108'791	1'315'524	1'487'577						
7	917'530	1'082'426	1'484'405							
8	1'001'238	1'376'124								
9	841'930									

Table 7.3: Cumulative claims payments

i\j	0	1	2	3	4	5	6	7	8	9
0	3'362'115	5'217'243	4'754'900	4'381'677	4'136'883	4'094'140	4'018'736	3'971'591	3'941'391	3'921'258
1	2'640'443	4'643'860	3'869'954	3'248'558	3'102'002	3'019'980	2'976'064	2'946'941	2'919'955	
2	2'879'697	4'785'531	4'045'448	3'467'822	3'377'540	3'341'934	3'283'928	3'257'827		
3	2'933'345	5'299'146	4'451'963	3'700'809	3'553'391	3'469'505	3'413'921			
4	2'768'181	4'658'933	3'936'455	3'512'735	3'385'129	3'298'998				
5	3'228'439	5'271'304	4'484'946	3'798'384	3'702'427					
6	2'927'033	5'067'768	4'066'526	3'704'113						
7	3'083'429	4'790'944	4'408'097							
8	2'761'163	4'132'757								
9	3'045'376									

Table 7.4: Incurred losses

i \ j	0	1	2	3	4	5	6	7	8	9	10
0	118	369	745	34	0	131	0	0	95	0	0
1	124	533	206	27	24	2	25	0	0	-76	0
2	556	1648	1290	-496	-15	35	-560	0	12	0	0
3	1646	705	141	15	105	0	-4	-853	0	0	0
4	317	569	4	0	60	40	0	0	0	0	0
5	242	677	299	6	5	20	0	0	0	0	0
6	203	409	10	17	28	-20	0	0	0	0	0
7	492	913	280	-17	85	-11	62	0	0	0	0
8	321	828	579	135	14	0	0	0	0	0	0
9	609	500	174	11	-41	2	0	0	0	0	0
10	492	1135	-5	50	0	0	0	0	-51	0	0
11	397	396	75	21	75	0	0	0	0	0	
12	523	575	377	14	0	0	0	0	0		
13	1786	1165	419	-341	182	78	36	0			
14	241	224	71	60	56	0	0				
15	327	295	-45	6	0	0					
16	275	245	9	0	12						
17	89	238	51	4							
18	295	6	95								
19	151	255									
20	315										

Table 7.5: Business unit 1

i \ j	0	1	2	3	4	5	6	7	8	9	10
0	92	350	99	0	-13	0	0	0	0	0	0
1	451	626	8	93	34	5	0	0	0	0	0
2	404	313	117	15	0	1	0	0	0	0	0
3	203	369	241	62	3	32	2	184	-7	0	-103
4	352	482	214	24	16	0	0	0	0	0	0
5	504	742	26	81	-68	0	0	0	0	0	0
6	509	499	53	0	0	10	0	0	0	0	0
7	229	351	50	40	2	0	0	0	0	0	0
8	324	491	56	-12	8	-90	0	0	0	0	0
9	508	297	101	63	2	0	0	0	0	0	0
10	354	287	192	9	0	0	0	0	0	0	0
11	431	416	7	61	3	0	0	0	0	0	
12	205	625	148	56	14	0	0	0	0		
13	522	612	-70	138	0	8	0	0			
14	567	358	-10	42	-4	0	0				
15	1238	686	110	-137	0	0					
16	355	648	134	27	32						
17	312	368	2	4							
18	246	106	66								
19	91	327									
20	130										

Table 7.6: Business unit 2

i \ j	0	1	2	3	4	5	6	7	8	9	10
0	268	188	29	-2	0	0	0	0	0	0	0
1	268	252	57	2	0	0	0	0	0	0	0
2	385	583	49	2	0	0	0	0	0	0	0
3	251	491	53	136	0	0	0	0	0	0	0
4	456	449	257	2	0	0	0	0	27	0	0
5	477	809	90	0	-3	0	0	0	0	0	0
6	405	594	173	24	0	14	0	0	0	0	0
7	443	489	20	13	19	8	20	0	0	0	0
8	477	569	290	26	13	0	0	0	0	0	0
9	581	565	170	46	29	0	0	0	0	0	0
10	401	596	232	19	33	3	-20	0	0	0	0
11	474	304	161	382	45	26	0	0	0	0	
12	649	771	287	2	0	0	0	-71	0		
13	911	1024	369	3	2	0	0	53			
14	508	546	47	-30	0	0	0				
15	389	401	78	41	660	0					
16	373	625	93	64	46						
17	276	577	79	16							
18	465	355	39								
19	343	279									
20	254										

Table 7.7: Business unit 3

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	136.367	195.757	213.788	227.545	240.136	249.647	260.260	271.207	285.847	290.814	300.035	305.988	312.930	316.524	319.988	323.268	325.798	326.231	327.722	328.616	332.724	337.137
1	143.135	212.658	233.989	252.155	267.259	280.669	295.863	310.584	322.938	332.021	341.670	349.924	356.908	361.126	367.795	370.548	373.051	376.942	378.314	379.716	381.388	
2	146.469	219.759	241.983	257.063	272.848	285.437	313.398	329.903	340.223	353.075	362.015	371.042	375.094	379.430	382.385	389.057	391.468	395.824	397.531	399.724		
3	158.518	232.128	256.752	276.593	292.807	310.757	322.837	339.751	352.613	366.707	378.735	385.394	394.505	402.618	409.044	412.422	415.624	421.409	424.117			
4	158.633	224.457	249.797	267.676	285.455	303.548	320.282	340.976	352.487	361.300	374.500	388.449	397.848	402.989	408.151	414.016	416.098	419.528				
5	153.215	225.074	249.688	267.753	285.294	307.116	324.791	341.238	353.420	369.549	382.016	390.301	395.206	403.634	406.302	407.819	411.082					
6	153.185	215.699	235.609	255.384	272.749	290.988	304.081	319.717	334.457	352.992	372.879	383.645	394.634	401.194	407.377	410.387						
7	150.974	217.545	242.400	260.473	279.436	299.797	317.991	336.679	352.929	373.339	397.542	407.145	416.136	429.445	435.980							
8	141.432	205.018	225.339	241.315	260.098	277.061	296.286	312.645	330.538	338.629	349.021	357.775	366.468	372.513								
9	141.554	207.510	230.597	250.393	272.538	294.008	321.253	346.836	366.865	381.705	391.678	404.292	411.770									
10	141.899	206.157	229.510	246.710	262.735	280.171	303.956	324.354	343.041	356.874	368.163	380.622										
11	145.037	215.127	240.970	260.457	280.524	304.118	322.331	345.629	357.081	370.673	384.000											
12	135.739	203.999	232.176	250.014	277.500	298.976	323.555	339.853	352.098	364.883												
13	135.350	209.545	236.220	256.710	276.576	293.467	305.436	320.329	336.143													
14	132.847	203.592	227.902	249.914	270.477	286.129	301.347	317.801														
15	135.951	205.450	229.862	250.624	266.371	280.202	300.874															
16	131.151	193.635	215.365	234.202	247.325	262.034																
17	130.188	190.262	213.586	226.115	242.768																	
18	118.505	174.622	192.852	206.808																		
19	118.842	177.671	199.872																			
20	121.011	185.856																				
21	132.116																					

Table 7.8: Cumulative claims payments $P_{i,j}$, $i + j \leq 21$, from a motor third party liability

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
0	370.661	369.491	358.572	344.401	346.665	342.370	344.807	344.529	346.520	344.774	345.346	344.592	339.747	340.214	338.824	338.687	338.119	337.285	336.834	336.271	336.647	337.137
1	419.289	416.218	393.048	384.661	383.508	376.001	382.602	383.968	384.388	383.803	387.951	389.023	386.576	389.510	388.477	388.833	389.687	390.600	389.190	388.147	388.548	
2	437.568	427.975	415.222	407.791	406.460	418.175	411.091	406.724	410.959	413.780	414.044	410.973	410.142	409.120	410.870	414.860	412.483	410.468	410.964	410.768		
3	450.610	448.418	431.294	424.361	432.825	427.773	436.947	434.592	437.282	437.315	438.579	439.037	440.049	439.942	439.561	437.789	437.218	439.781	438.979			
4	468.979	450.942	432.591	435.162	423.583	420.876	424.588	424.632	429.180	431.788	429.673	433.284	430.848	432.002	431.379	429.732	429.539	428.875				
5	443.189	434.315	433.325	425.332	424.921	422.432	421.233	422.904	424.484	426.661	431.462	431.676	434.368	433.599	431.820	432.994	433.037					
6	435.307	422.149	412.660	405.190	402.367	402.597	411.096	412.873	420.973	432.325	437.067	433.610	435.824	435.656	434.796	436.011						
7	475.948	438.817	432.413	436.710	436.904	443.156	448.347	463.171	468.412	476.070	474.593	473.324	477.058	472.283	473.524							
8	447.021	422.678	405.919	399.462	400.047	398.297	406.939	404.834	409.056	411.421	412.002	410.739	409.744	413.587								
9	457.229	444.054	436.390	436.853	442.292	453.494	456.363	460.272	459.591	456.975	455.336	454.500	453.068									
10	462.989	464.776	447.833	432.893	432.124	442.743	451.994	451.534	450.528	450.845	448.398	442.810										
11	484.915	468.800	454.958	447.601	461.106	470.358	465.346	468.879	461.537	456.753	453.919											
12	462.028	429.610	438.929	454.797	468.116	468.721	469.907	463.823	459.524	452.385												
13	450.908	456.030	476.259	483.129	476.952	464.941	453.391	445.089	434.103													
14	426.385	428.504	456.796	449.886	445.397	432.021	412.353	402.565														
15	461.078	477.458	480.960	471.869	462.978	444.670	437.203															
16	444.123	430.684	433.664	419.422	403.126	396.903																
17	433.830	407.931	393.723	371.800	361.853																	
18	418.202	374.855	338.598	324.790																		
19	426.853	373.282	351.590																			
20	410.810	394.477																				
21	405.597																					

Table 7.9: Incurred losses $I_{i,j}$, $i + j \leq 21$, from a motor third party liability

Eidesstattliche Versicherung

Hiermit erkläre ich, Sebastian Happ, an Eides statt, dass ich die vorliegende Dissertation mit dem Titel

"**Stochastic Claims Reserving under Consideration of Various Different Sources of Information**"

selbständig und ohne fremde Hilfe verfasst habe.
Andere als die von mir angegebenen Quellen und Hilfsmittel habe ich nicht benutzt.

Ort und Datum: Hamburg, den 02.07.2014.
Unterschrift: