Gauss–Markov Loss Prediction in a Linear Model

By Alexander Ludwig[1] and Klaus D. Schmidt[2,*]

1. zeb/rolfes.schierenbeck.associates, Hammer Strasse 165, D–48153 Münster, Germany
2. Lehrstuhl für Versicherungsmathematik, Technische Universität Dresden, D–01062 Dresden, Germany

June 10, 2010

Abstract

In a linear model for loss reserving, Gauss–Markov prediction is the natural principle of prediction: It minimizes the mean squared error of prediction over the class of all unbiased linear predictors, and it provides exact formulas for predictors and their mean squared error of prediction. Another advantage of Gauss–Markov prediction is in the fact that the Gauss–Markov predictor of a sum is just the sum of the Gauss–Markov predictors of the single terms of that sum such that essentially only the most elementary quantities have to be predicted.

The use of Gauss–Markov prediction in loss reserving is not new. For example, the additive (or incremental loss ratio) method and the Panning method are based on Gauss–Markov prediction in an appropriate linear model. Here we propose a systematic study of Gauss–Markov prediction in these and several related models. This leads to a variety of new methods of loss reserving, and for each of these models and methods we obtain straightforward estimators of the mean squared error of prediction.

To complete the discussion, we also explain certain limitations of the Gauss–Markov principle in connection with the chain–ladder method.

*Corresponding author. E–mail address: `klaus.d.schmidt@tu-dresden.de`

1 Introduction

For at least six decades, loss reserving was determined by a variety of heuristic methods, among which the most popular ones are the chain–ladder method described by Tarbell [1934] and the Bornhuetter–Ferguson method proposed by Bornhuetter and Ferguson [1972].

The first stochastic model for loss reserving is probably that of Hachemeister and Stanard [1975]. In their model, the incremental losses are independent and Poisson distributed with a multiplicative structure of the expectations, and it turns out that maximum–likelihood estimations leads to the chain–ladder predictors. Their model thus provides a first justification of the chain–ladder method, but because of the Poisson assumption it applies to claim numbers rather than claim amounts.[1]

About two decades later, a couple of papers appeared which considerably advanced the use of stochastic models in loss reserving. In one of these papers, Mack [1991] proposed a model in which the incremental losses are uncorrelated with a multiplicative structure of the expectations and variances and in which least squares estimation leads to the additive (or incremental loss ratio) method. Subsequently, Mack [1993] proposed another but similar model in which least squares estimation leads to the chain–ladder method.[2] In both of these papers, however, emphasis is on parameter estimation and not on prediction of future losses.

It is easy to see that the additive model of Mack [1991] is a linear model, and it follows from Schmidt and Schnaus [1996] that the chain–ladder model of Mack [1993] is a sequential linear model.[3] But this was certainly not the usual way of looking at these models at the time when they were published, and it is the merit of Halliwell [1996] of having pointed out that linear models are most useful in loss reserving since the Gauss–Markov principle provides not only estimators of parameters but also predictors of future losses.

About another decade later, linear models turned out to be a driving force for the development of new methods of loss reserving: Inspired by Braun [2004], Pröhl and Schmidt [2005] proposed a sequential linear model in which Gauss–Markov prediction leads to a multivariate version of the chain–ladder method[4] and Hess, Schmidt and Zocher [2006] proposed a linear model in which Gauss–Markov prediction leads to a multivariate version of the additive method. Both methods are of interest for

[1]Extensions of the model of Hachemeister and Stanard [1975], which allow for dependence within the accident years and in which maximum–likelihood estimation still produces the chain–ladder predictors of the ultimate cumulative losses were proposed by Schmidt and Wünsche [1998] and by Schmidt and Zocher [2005].

[2]It is remarkable that the assumptions of the model of Hachemeister and Standard [1972] and those of the model of Mack [1993] cannot be fulfilled simultaneously; see Hess and Schmidt [2002] for a comparison of a variety of models for the chain–ladder method.

[3]See Schmidt [2003] and Radtke and Schmidt [2004].

[4]Another paper which is in the spirit of Pröhl and Schmidt [2005] is that of Kremer [2005].

simultaneous prediction for dependent lines of business.[5] At the same time, Panning [2006] proposed a linear model which in a certain sense is intermediate between the linear model for the additive method and the sequential linear model for the chain–ladder method. More recently, Kloberdanz and Schmidt [2009] used a bivariate version of the additive model to approach the paid & incurred problem which was first studied by Halliwell [1997] and later by Quarg and Mack [2004, 2008].

At this point, it is useful to briefly review some basic aspects of linear and general linear models and of Gauss–Markov estimation and prediction in such models; a more precise discussion will be given in Section 3.

A *linear model* (or *regression model*) essentially consists in the assumption that the unknown expectations of certain random variables $X_1, \ldots, X_s$ can be expressed as linear functions of certain unknown parameters $\beta_1, \ldots, \beta_r$ with $r < s$. This means that, for every $i \in \{1, \ldots, s\}$, there exist known coefficients $a_{i,1}, \ldots, a_{i,r}$ such that

$$E[X_i] \;=\; \sum_{k=1}^{r} a_{i,k}\, \beta_k$$

The point is that in a linear model the s unknown expectations are explained by r unknown parameters such that the problem of estimating s expectations is reduced to that of estimating only $r < s$ parameters. A general principle for estimating the parameters in a linear model is *Gauss–Markov estimation* which consists in the computation of the *Gauss–Markov estimators* β_k^{GM} minimizing the *mean squared error of estimation*

$$E[(\widehat{\beta}_k - \beta_k)^2]\,.$$

over all estimators $\widehat{\beta}_k$ which are linear in $X_1, \ldots, X_s$ and unbiased for β_k. Thus, with respect to the mean squared error of estimation, the Gauss–Markov estimator β_k^{GM} is the *best linear unbiased estimator* of β_k.

In a *general linear model*, only the first $s_1 < s$ random variables are observable while the remaining $s_2 := s - s_1$ random variables are non–observable. In this case, Gauss–Markov estimation of the parameters is still possible by replacing s with s_1 in the previous identities, but the real problem is *Gauss–Markov prediction* of the non–observable random variables which consists in the computation of the *Gauss–Markov predictors* X_j^{GM} with $j \in \{s_1{+}1, \ldots, s_1{+}s_2\}$ minimizing the *mean squared error of prediction*

$$E[(\widehat{X}_j - X_j)^2]$$

over all predictors $\widehat{X}_j$ which are linear and unbiased for X_j in the sense that $E[\widehat{X}_j] = E[X_j]$. Thus, with respect to the mean squared error of prediction, the Gauss–Markov predictor X_j^{GM} is the *best linear unbiased predictor* of X_j.

Under mild conditions on the coefficients and the variances and covariances of the random variables, Gauss–Markov estimators and predictors exist and are unique. To

[5]See Schmidt [2006b] for a survey of the results of these papers.

determine Gauss–Markov estimators and predictors, the variances and covariances of the random variables must be known or have to be estimated but no further assumptions on their joint distribution have to be made.[6] Moreover, since Gauss–Markov estimators and predictors are linear and unbiased, it is evident that also the mean squared errors of estimation and prediction are determined by the variances and covariances.

Since loss reserving aims at the prediction of future losses from those observed in the past, every stochastic model for loss reserving typically has to consist of observable and non–observable random variables representing past and future losses. Therefore, general linear models provide a wide class of stochastic models which meet the basic requirement on every stochastic model for loss reserving.

Whenever it is judged to be appropriate, the use of general linear models in loss reserving is strongly recommendable since

- explicit formulas can be given for Gauss–Markov predictors of reserves and for their mean squared error of prediction, and
- estimators of the mean squared errors of prediction can be obtained by simply replacing unknown variances and covariances with appropriate estimators.

Of course, the choice of a particular stochastic model for loss reserving should not be determined by such technical advantages but rather by statistical analysis and actuarial judgement. In many cases, however, such considerations will not end up with a single model and the choice of a general linear model could be reasonable.

In the present paper we propose Gauss–Markov prediction in a general linear model as a common approach to the additive method, the Panning method and a new method which is a combination of both and could be extended further. We thus extend results of Ludwig, Schmeisser, and Thänert [2009].

This paper is organized as follows: We first present the typical data structure in loss reserving (Section 2) and discuss Gauss–Markov prediction in the general linear model (Section 3). We then apply the general results on Gauss–Markov prediction to the additive model (Section 4), the Panning model (Section 5), and the combined model (Section 6). For the sake of comparison, we also consider the Mack model for the chain–ladder method (Section 7), which because of its sequential structure presents certain difficulties with regard to the estimation of the mean squared errors of prediction for reserves.[7] Finally, we present a numerical example (Section 9) and we conclude with some remarks (Section 8).

[6]In particular, it is not necessary to assume that the random variables are jointly normally distributed. The popularity of the normal assumption is probably due to the fact that, if it holds, then the Gauss–Markov estimators agree with the maximum–likelihood estimators. While the normal assumption is inessential for Gauss–Markov estimation and prediction, it is of interest for the construction of confidence intervals or prediction intervals; these topics, however, will not be dealt with in the present paper.

[7]The use of plug–in estimators for estimating the mean squared errors of prediction is not possible in the Mack model; instead, certain approximations seem to be unavoidable in the construction of estimators or the mean squared errors of prediction and it appears to be difficult to quantify the approximation errors.

2 Data Structure

In the present paper, we consider a portfolio of risks and we assume that each claim of the portfolio is settled either in the accident year or in finitely many subsequent development years.

To model such a portfolio, we consider a family of square integrable random variables

$$\{Z_{i,k}\}_{i\in\{-m,\ldots,n\},k\in\{0,\ldots,n\}}$$

and we interpret the random variable $Z_{i,k}$ as the loss of *accident year* i which is settled with a delay of k years and hence in *development year* k and in *calendar year* $i+k$. We refer to $Z_{i,k}$ as the *incremental loss* of accident year i and development year k.

We assume that the incremental losses $Z_{i,k}$ are *observable* for *calendar years* $i+k \leq n$ and that they are *non–observable* for calendar years $i+k \geq n+1$. The observable incremental losses are represented by the following *run–off trapezoid*:

Accident	**Development Year**									
Year	0	1	$\ldots$	k	$\ldots$	$n-i$	$\ldots$	$n-1$	n
$-m$	$Z_{-m,0}$	$Z_{-m,1}$	$\ldots$	$Z_{-m,k}$	$\ldots$	$Z_{-m,n-i}$	$\ldots$	$Z_{-m,n-1}$	$Z_{-m,n}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	$\vdots$
0	$Z_{0,0}$	$Z_{0,1}$	$\ldots$	$Z_{0,k}$	$\ldots$	$Z_{0,n-i}$	$\ldots$	$Z_{0,n-1}$	$Z_{0,n}$
1	$Z_{1,0}$	$Z_{1,1}$	$\ldots$	$Z_{1,k}$	$\ldots$	$Z_{1,n-i}$	$\ldots$	$Z_{1,n-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$			
i	$Z_{i,0}$	$Z_{i,1}$	$\ldots$	$Z_{i,k}$	$\ldots$	$Z_{i,n-i}$			
$\vdots$	$\vdots$	$\vdots$		$\vdots$					
$n-k$	$Z_{n-k,0}$	$Z_{n-k,1}$	$\ldots$	$Z_{n-k,k}$					
$\vdots$	$\vdots$	$\vdots$							
$n-1$	$Z_{n-1,0}$	$Z_{n-1,1}$							
n	$Z_{n,0}$								

In the traditional case $m = 0$, the run–off trapezoid reduces to a *run–off triangle.* The case $m \geq 1$ is of interest, since it is always desirable to have more than one completely developed accident year and since this also turns out to be necessary for certain stochastic models which to some extent specify the joint distribution of the family of all incremental losses.

For the stochastic models to be considered in this paper, it is essential to linearize the run–off trapezoid of observable incremental losses and the triangle of non–observable incremental losses. Therefore, we define the random vectors

$$\mathbf{X}_1 := \begin{pmatrix} Z_{-m,0} \\ \vdots \\ Z_{n,0} \\ \hline \vdots \\ \hline Z_{-m,k} \\ \vdots \\ Z_{n-k,k} \\ \hline \vdots \\ \hline Z_{-m,n} \\ \vdots \\ Z_{0,n} \end{pmatrix} \qquad \text{and} \qquad \mathbf{X}_2 := \begin{pmatrix} Z_{n,1} \\ \hline \vdots \\ \hline Z_{n-k+1,k} \\ \vdots \\ Z_{n,k} \\ \hline \vdots \\ \hline Z_{1,n} \\ \vdots \\ Z_{n,n} \end{pmatrix}$$

such that $\mathbf{X}_1$ represents the run–off trapezoid of observable incremental losses and $\mathbf{X}_2$ represents the triangle of non–observable incremental losses.

The first problem is to *predict*

(1) the *accident year reserves*

$$R_i := \sum_{k=n-i+1}^{n} Z_{i,k}$$

for $i \in \{1, \ldots, n\}$,

(2) the *calendar year reserves*

$$R_{(c)} := \sum_{i=c-n}^{n} Z_{i,c-i}$$

for $c \in \{n+1, \ldots, 2n\}$, and

(3) the *total reserve*

$$R := \sum_{k=1}^{n} \sum_{i=n-k+1}^{n} Z_{i,k}.$$

In either case, the problem is to predict $\mathbf{d}'\mathbf{X}_2$ for a suitable vector $\mathbf{d}$.

The second problem is to estimate the mean squared error of prediction for the predictors of these reserves.

3 Gauss–Markov Prediction in the General Linear Model

The stochastic models of loss reserving to be studied in the present paper are special cases of the following general linear model for a random vector

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}$$

consisting of an observable part $\mathbf{X}_1$ and a non–observable part $\mathbf{X}_2$ with square integrable coordinates:

> **General Linear Model:** *There exist known matrices* $\mathbf{A}_1$ *and* $\mathbf{A}_2$ *and an unknown parameter vector* $\boldsymbol{\beta}$ *such that*
>
> $$E\left[\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}\right] = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{pmatrix} \boldsymbol{\beta}$$
>
> *Moreover,* $\mathbf{A}_1$ *has full column rank and* $\mathrm{var}[\mathbf{X}_1]$ *is invertible.*

The general linear model is more general than the traditional linear model since it involves the non–observable part $\mathbf{X}_2$. In particular, the problem is not only to estimate the parameter vector $\boldsymbol{\beta}$ but also to predict the non–observable random vector $\mathbf{X}_2$. The matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ are called the *design matrices* of the general linear model.

For the remainder of this section, we assume that the assumptions of the general linear model are fulfilled.

Following an idea of Hamer [1999], the best way to simultaneously estimate the parameter vector $\boldsymbol{\beta}$ and predict the non–observable random vector $\mathbf{X}_2$ is to predict a target quantity of the form

$$\mathbf{T} = \mathbf{C}_0\boldsymbol{\beta} + \mathbf{C}_1\mathbf{X}_1 + \mathbf{C}_2\mathbf{X}_2$$

with matrices $\mathbf{C}_0, \mathbf{C}_1, \mathbf{C}_2$ of suitable dimensions which also allows for the prediction of linear combinations of the coordinates of $\mathbf{X}_1$ and $\mathbf{X}_2$.

Since only $\mathbf{X}_1$ is observable, every random variable $\widehat{\mathbf{T}}$ which is a (measurable) transformation of $\mathbf{X}_1$ is said to be a *predictor* of $\mathbf{T}$.

A predictor $\widehat{\mathbf{T}}$ is said to be an *admissible predictor* of $\mathbf{T}$ if there exists a matrix $\mathbf{Q}$ satisfying

$$\widehat{\mathbf{T}} = \mathbf{Q}\mathbf{X}_1$$

and

$$\mathbf{Q}\mathbf{A}_1 = \mathbf{C}_0 + \mathbf{C}_1\mathbf{A}_1 + \mathbf{C}_2\mathbf{A}_2$$

Because of the first identity, every admissible predictor $\widehat{\mathbf{T}}$ of $\mathbf{T}$ is *linear* (in $\mathbf{X}_0$), and because of the second identity it is also *unbiased* since

$$\begin{aligned} E[\widehat{\mathbf{T}}] &= E[\mathbf{Q}\mathbf{X}_1] \\ &= \mathbf{Q}E[\mathbf{X}_1] \\ &= \mathbf{Q}\mathbf{A}_1\boldsymbol{\beta} \\ &= (\mathbf{C}_0 + \mathbf{C}_1\mathbf{A}_1 + \mathbf{C}_2\mathbf{A}_2)\boldsymbol{\beta} \\ &= \mathbf{C}_0\boldsymbol{\beta} + \mathbf{C}_1\mathbf{A}_1\boldsymbol{\beta} + \mathbf{C}_2\mathbf{A}_2\boldsymbol{\beta} \\ &= \mathbf{C}_0\boldsymbol{\beta} + \mathbf{C}_1E[\mathbf{X}_1] + \mathbf{C}_2E[\mathbf{X}_2] \\ &= E[\mathbf{C}_0\boldsymbol{\beta} + \mathbf{C}_1\mathbf{X}_1 + \mathbf{C}_2\mathbf{X}_2] \\ &= E[\mathbf{T}] \end{aligned}$$

An admissible predictor $\widehat{\mathbf{T}}$ of $\mathbf{T}$ is said to be a *Gauss–Markov predictor* of $\mathbf{T}$ if it minimizes the *mean squared error of prediction*

$$E[(\widehat{\mathbf{T}} - \mathbf{T})'(\widehat{\mathbf{T}} - \mathbf{T})]$$

which is sometimes also called the *mean squared error of prediction* of $\widehat{\mathbf{T}}$ and is denoted by m.s.e.p.$[\widehat{\mathbf{T}}]$. Since every admissible predictor $\widehat{\mathbf{T}}$ of $\mathbf{T}$ is unbiased, we have $E[\widehat{\mathbf{T}} - \mathbf{T}] = \mathbf{0}$ and hence

$$\begin{aligned} E[(\widehat{\mathbf{T}} - \mathbf{T})'(\widehat{\mathbf{T}} - \mathbf{T})] &= E[\text{trace}((\widehat{\mathbf{T}} - \mathbf{T})(\widehat{\mathbf{T}} - \mathbf{T})')] \\ &= \text{trace}(E[(\widehat{\mathbf{T}} - \mathbf{T})(\widehat{\mathbf{T}} - \mathbf{T})')] \\ &= \text{trace}(\text{var}[\widehat{\mathbf{T}} - \mathbf{T}] + E[\widehat{\mathbf{T}} - \mathbf{T}]\, E[\widehat{\mathbf{T}} - \mathbf{T}]') \\ &= \text{trace}(\text{var}[\widehat{\mathbf{T}} - \mathbf{T}])\,. \end{aligned}$$

We have the following result:

3.1 Proposition (Gauss–Markov Theorem). *There exists a unique Gauss–Markov predictor* $\mathbf{T}^{\text{GM}}$ *of* $\mathbf{T}$ *and it satisfies*

$$\mathbf{T}^{\text{GM}} \;=\; \mathbf{C}\boldsymbol{\beta}^* + \mathbf{C}_1\mathbf{X}_1 + \mathbf{C}_2\mathbf{X}_2^*$$

with

$$\boldsymbol{\beta}^* \;:=\; (\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_1$$

and

$$\mathbf{X}_2^* \;:=\; \mathbf{A}_2\boldsymbol{\beta}^* + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\big(\mathbf{X}_1 - \mathbf{A}_1\boldsymbol{\beta}^*\big)$$

Moreover,

$$\text{var}[\mathbf{T}^{\text{GM}} - \mathbf{T}] \;=\; \mathbf{K}\text{var}[\boldsymbol{\beta}^*]\mathbf{K} + \mathbf{C}_2(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})\mathbf{C}_2'$$

with $\mathbf{K} := \mathbf{C} + \mathbf{C}_2\mathbf{A}_2 - \mathbf{C}_2\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\mathbf{C}_2'$ *and* $\text{var}[\boldsymbol{\beta}^*] \;=\; (\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}$.

Proposition 3.1 is well–known; see e. g. Rao and Toutenburg [1995], Radtke and Schmidt [2004], Schmidt [2004] and, in particular, Hamer [1999].

The previous result shows that Gauss–Markov prediction of the target quantity $\mathbf{T}$ is based on Gauss–Markov estimation of the parameter $\boldsymbol{\beta}$. Although the following result is a special case of Proposition 3.1, we state it because of its importance and for later reference:

3.2 Corollary. *The Gauss–Markov estimator $\boldsymbol{\beta}^{\mathrm{GM}}$ of $\boldsymbol{\beta}$ satisfies*

$$\boldsymbol{\beta}^{\mathrm{GM}} \;=\; (\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{X}_1$$

and

$$\mathrm{var}[\boldsymbol{\beta}^{\mathrm{GM}}] \;=\; (\mathbf{A}_1'\boldsymbol{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}.$$

In the models considered in this paper, we always have $\boldsymbol{\Sigma}_{12} = \mathbf{O}$. In this case, we obtain particularly simple formulas for the Gauss–Markov predictor of $\mathbf{X}_2$ and for the variance of the prediction error:

3.3 Corollary. *Assume that $\boldsymbol{\Sigma}_{12} = \mathbf{O}$. Then the Gauss–Markov predictor $\mathbf{X}_2^{\mathrm{GM}}$ of $\mathbf{X}_2$ satisfies*

$$\mathbf{X}_2^{\mathrm{GM}} \;=\; \mathbf{A}_2\boldsymbol{\beta}^{\mathrm{GM}}$$

and

$$\mathrm{var}[\mathbf{X}_2^{\mathrm{GM}} - \mathbf{X}_2] \;=\; \mathbf{A}_2\mathrm{var}[\boldsymbol{\beta}^{\mathrm{GM}}]\mathbf{A}_2' + \boldsymbol{\Sigma}_{22}\,.$$

Because of the previous result, the mean squared error of prediction

$$E[(\mathbf{X}_2^{\mathrm{GM}} - \mathbf{X}_2)'(\mathbf{X}_2^{\mathrm{GM}} - \mathbf{X}_2)] \;=\; \mathrm{trace}\Big(\mathrm{var}[\mathbf{X}_2^{\mathrm{GM}} - \mathbf{X}_2]\Big)$$

is the sum of the *estimation error* $\mathrm{trace}(\mathbf{A}_2\mathrm{var}[\boldsymbol{\beta}^{\mathrm{GM}}]\mathbf{A}_2')$ and the *random error* $\mathrm{trace}(\boldsymbol{\Sigma}_{22})$.

Finally, the Gauss–Markov predictor of a linear transformation $\mathbf{C}_2\mathbf{X}_2$ of $\mathbf{X}_2$ is easily obtained from the Gauss–Markov predictor of $\mathbf{X}_2$:

3.4 Corollary. *The Gauss–Markov predictor $(\mathbf{C}_2\mathbf{X}_2)^{\mathrm{GM}}$ of $\mathbf{C}_2\mathbf{X}_2$ satisfies*

$$(\mathbf{C}_2\mathbf{X}_2)^{\mathrm{GM}} \;=\; \mathbf{C}_2\mathbf{X}_2^{\mathrm{GM}}$$

and

$$\mathrm{var}[(\mathbf{C}_2\mathbf{X}_2)^{\mathrm{GM}} - \mathbf{C}_2\mathbf{X}_2] \;=\; \mathbf{C}_2\mathrm{var}[\mathbf{X}_2^{\mathrm{GM}} - \mathbf{X}_2]\mathbf{C}_2'\,.$$

Because of the previous result, Gauss–Markov prediction is linear in the sense that the Gauss–Markov predictor of a linear combination of non–observable random variables is the same linear combination of their Gauss–Markov predictors.

We shall also need a conditional version of the general linear model and of the Gauss–Markov Theorem. For a sub–σ–algebra $\mathcal{G} \subseteq \mathcal{F}$, the $\mathcal{G}$–conditional linear model is defined as follows:

$\mathcal{G}$–Conditional General Linear Model: *There exist observable $\mathcal{G}$–measurable random matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ and an unknown parameter vector $\boldsymbol{\beta}$ such that*

$$E^{\mathcal{G}}\left[\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix}\right] = \begin{pmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{pmatrix}\boldsymbol{\beta}\,.$$

Moreover, $\mathbf{A}_1$ *has full column rank and* $\mathrm{var}^{\mathcal{G}}[\mathbf{X}_1]$ *is invertible.*

Here and in the sequel, $E^{\mathcal{G}}[\mathbf{X}_i]$ and $\mathrm{var}^{\mathcal{G}}[\mathbf{X}_i]$ denote the $\mathcal{G}$–conditional expectation and the $\mathcal{G}$–conditional variance of $\mathbf{X}_i$, respectively; accordingly, $\mathrm{cov}^{\mathcal{G}}[\mathbf{X}_1, \mathbf{X}_2]$ denotes the $\mathcal{G}$–conditional covariance of $\mathbf{X}_1$ and $\mathbf{X}_2$.

The discussion of the $\mathcal{G}$–conditional general linear model is entirely analogous to that of the general linear model: Replace the admissible predictors by the $\mathcal{G}$–conditionally admissible predictors (which are obtained by replacing the matrix $\mathbf{Q}$ by a $\mathcal{G}$–measurable random matrix $\mathbf{Q}$ and which are linear and $\mathcal{G}$–conditionally unbiased in the sense that their $\mathcal{G}$–conditional expectation coincides with that of the target quantity) and replace the first and second order moments by their $\mathcal{G}$–conditional counterparts.

4 Gauss–Markov Loss Prediction in the Extended Additive Model

The extended additive model is defined as follows:

Extended Additive Model: *There exist known parameters* $v_i, w_i \in (0, \infty)$ *with* $i \in \{-m, \ldots, n\}$ *as well as unknown parameters* $\zeta_k \in \mathbb{R}$ *and* $\sigma_k^2 \in (0, \infty)$ *with* $k \in \{0, \ldots, n\}$ *such that the incremental losses satisfy*

$$\begin{aligned} E[Z_{i,k}] &= v_i\, \zeta_k \\ \operatorname{cov}[Z_{i,k}, Z_{j,l}] &= w_i\, \sigma_k^2\, \delta_{i,j}\, \delta_{k,l} \end{aligned}$$

for all $i, j \in \{-m, \ldots, n\}$ *and* $k, l \in \{0, \ldots, n\}$.

In the extended additive model, the accident year parameter v_i is usually referred to as a *volume measure* of accident year i; for example, the volume measure could be the total premium income or the number of contracts in the accident year. Since the first identity in the extended additive model can be written as

$$E[Z_{i,k}/v_i] = \zeta_k$$

the development year parameter ζ_k is the *expected incremental loss ratio* of development year k (with respect to the volume measures) and is assumed to be independent of the accident year such that the collection of these parameters forms a *development pattern*; see Schmidt and Zocher [2009].

The extended additive model extends the traditional additive model in which it is assumed that $m = 0$ and that $w_i = v_i$ holds for all $i \in \{-m, \ldots, n\}$; see Mack [1991], Radtke and Schmidt [2004], Hess, Schmidt and Zocher [2006], and Schmidt and Zocher [2009]. The reason for considering the extended additive model becomes evident from its comparison with the extended Panning model (Section 5) and with the combination of both models (Section 6).

Assume that the assumptions of the extended additive model are fulfilled. Then the expectation of the random vector $\mathbf{X}_1$ of all observable incremental losses satisfies

$$E\left[\left(\begin{array}{c} Z_{-m,0} \\ \vdots \\ Z_{n,0} \\ \hline \vdots \\ \hline Z_{-m,k} \\ \vdots \\ Z_{n-k,k} \\ \hline \vdots \\ \hline Z_{-m,n} \\ \vdots \\ Z_{0,n} \end{array}\right)\right] = \left(\begin{array}{ccccc} v_{-m} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ v_n & \cdots & 0 & \cdots & 0 \\ \hline \vdots & & \vdots & & \vdots \\ \hline 0 & \cdots & v_{-m} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & v_{n-k} & \cdots & 0 \\ \hline \vdots & & \vdots & & \vdots \\ \hline 0 & \cdots & 0 & \cdots & v_{-m} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & v_0 \end{array}\right) \left(\begin{array}{c} \zeta_0 \\ \vdots \\ \zeta_k \\ \vdots \\ \zeta_n \end{array}\right)$$

such that there exist a design matrix $\mathbf{A}_1$ having full column rank and a parameter vector $\boldsymbol{\beta}$ satisfying $E[\mathbf{X}_1] = \mathbf{A}_1\boldsymbol{\beta}$, and the expectation of the random vector $\mathbf{X}_2$ of all non–observable incremental losses satisfies

$$
E\left[\left(\begin{array}{c} Z_{n,1} \\ \hline \vdots \\ \hline Z_{n-k+1,k} \\ \vdots \\ Z_{n,k} \\ \hline \vdots \\ \hline Z_{1,n} \\ \vdots \\ Z_{n,n} \end{array}\right)\right]
=
\left(\begin{array}{cccccc}
0 & v_n & \cdots & 0 & \cdots & 0 \\ \hline
\vdots & \vdots & & \vdots & & \vdots \\ \hline
0 & 0 & \cdots & v_{n-k+1} & \cdots & 0 \\
\vdots & \vdots & & \vdots & & \vdots \\
0 & 0 & \cdots & v_{n-k} & \cdots & 0 \\ \hline
\vdots & \vdots & & \vdots & & \vdots \\ \hline
0 & 0 & \cdots & 0 & \cdots & v_1 \\
\vdots & \vdots & & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & \cdots & v_n
\end{array}\right)
\left(\begin{array}{c} \zeta_0 \\ \zeta_1 \\ \vdots \\ \zeta_k \\ \vdots \\ \zeta_n \end{array}\right).
$$

Moreover, the variance $\boldsymbol{\Sigma}_{11}$ of $\mathbf{X}_1$ satisfies

$$
\mathrm{var}\left[\left(\begin{array}{c} Z_{-m,0} \\ \vdots \\ Z_{n,0} \\ \hline \vdots \\ \hline Z_{-m,k} \\ \vdots \\ Z_{n-k,k} \\ \hline \vdots \\ \hline Z_{-m,n} \\ \vdots \\ Z_{0,n} \end{array}\right)\right]
=
\left(\begin{array}{ccc|c|ccc|c|ccc}
w_{-m}\sigma_0^2 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\
0 & \cdots & w_n\sigma_0^2 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \hline
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & 0 & 0 & \cdots & w_{-m}\sigma_k^2 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \cdots & w_{n-k}\sigma_k^2 & \cdots & 0 & \cdots & 0 \\ \hline
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & w_{-m}\sigma_n^2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & w_0\sigma_n^2
\end{array}\right)
$$

and is thus invertible, and the variance $\boldsymbol{\Sigma}_{22}$ of $\mathbf{X}_2$ satisfies

$$
\mathrm{var}\left[\left(\begin{array}{c} Z_{n,1} \\ \hline \vdots \\ \hline Z_{n-k+1,k} \\ \vdots \\ Z_{n,k} \\ \hline \vdots \\ \hline Z_{1,n} \\ \vdots \\ Z_{n,n} \end{array}\right)\right]
=
\left(\begin{array}{c|c|ccc|c|ccc}
w_n\sigma_1^2 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \hline
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & \cdots & w_{n-k+1}\sigma_k^2 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
\vdots & \cdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & w_n\sigma_k^2 & \cdots & 0 & \cdots & 0 \\ \hline
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \hline
0 & \cdots & 0 & \cdots & 0 & \cdots & w_1\sigma_n^2 & \cdots & 0 \\
\vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\
0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & w_n\sigma_n^2
\end{array}\right).
$$

Furthermore, we have $\boldsymbol{\Sigma}_{12} = \mathrm{cov}[\mathbf{X}_1, \mathbf{X}_2] = \mathbf{O}$. We thus obtain the following result:

4.1 Theorem. *The extended additive model is a linear model.*

In a first step, we compute the Gauss–Markov estimators of the coordinates of the parameter vector and their covariances:

4.2 Lemma (Gauss–Markov estimation of parameters). *In the extended additive model, the Gauss–Markov estimators of the coordinates of the parameter vector satisfy*

$$\zeta_k^{\mathrm{GM}} = \frac{\sum_{i=-m}^{n-k} v_i Z_{i,k}/w_i}{\sum_{i=-m}^{n-k} v_i^2/w_i}$$

and

$$\mathrm{cov}[\zeta_k^{\mathrm{GM}}, \beta_l^{\mathrm{GM}}] = \frac{1}{\sum_{i=-m}^{n-k} v_i^2/w_i}\, \sigma_k^2\, \delta_{k,l}$$

for all $k, l \in \{0, 1, \ldots, n\}$*.*

Proof. The coordinates of the random vector $\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{X}_1$ satisfy

$$\begin{pmatrix} v_{-m} & \cdots & v_{n-k} \end{pmatrix} \begin{pmatrix} w_{-m}\sigma_k^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{n-k}\sigma_k^2 \end{pmatrix}^{-1} \begin{pmatrix} Z_{-m,k} \\ \vdots \\ Z_{n-k,k} \end{pmatrix} = \left(\sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} \right) \frac{1}{\sigma_k^2} \,.$$

Moreover, the matrix $\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_1$ is diagonal and its diagonal elements satisfy

$$\begin{pmatrix} v_{-m} & \cdots & v_{n-k} \end{pmatrix} \begin{pmatrix} w_{-m}\sigma_k^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{n-k}\sigma_k^2 \end{pmatrix}^{-1} \begin{pmatrix} v_{-m} \\ \vdots \\ v_{n-k} \end{pmatrix} = \left(\sum_{i=-m}^{n-k} \frac{v_i^2}{w_i} \right) \frac{1}{\sigma_k^2} \,.$$

Because of Corollary 3.2, we have $\boldsymbol{\beta}^{\mathrm{GM}} = (\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{X}_1$ and hence

$$\zeta_k^{\mathrm{GM}} = \frac{\sum_{i=-m}^{n-k} v_i Z_{i,k}/w_i}{\sum_{i=-m}^{n-k} v_i^2/w_i}$$

which is the first identity, and we also have $\mathrm{var}[\boldsymbol{\beta}^{\mathrm{GM}}] = (\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}$ and hence

$$\mathrm{cov}[\zeta_k^{\mathrm{GM}}, \zeta_l^{\mathrm{GM}}] = \frac{1}{\sum_{i=-m}^{n-k} v_i^2/w_i}\, \sigma_k^2\, \delta_{k,l}$$

which is the second identity. □

In a second step, we compute the Gauss–Markov predictors of the non–observable incremental losses and the covariances of their prediction errors:

4.3 Lemma (Gauss–Markov prediction of incremental losses). *In the extended additive model, the Gauss–Markov predictors of the non–observable incremental losses satisfy*

$$Z_{i,k}^{\mathrm{GM}} = v_i\, \zeta_k^{\mathrm{GM}}$$

and

$$\mathrm{cov}[Z_{i,k}^{\mathrm{GM}} - Z_{i,k}, Z_{j,l}^{\mathrm{GM}} - Z_{j,l}] = \Big(v_i v_j\, \mathrm{var}[\zeta_k^{\mathrm{GM}}] + w_i\, \sigma_k^2\, \delta_{i,j} \Big)\, \delta_{k,l}$$

for all $i, j \in \{-m, \ldots, n\}$ *and* $k, l \in \{0, \ldots, n\}$ *such that* $\min\{i+k, j+l\} \geq n+1$*.*

Proof. Because of Corollary 3.3, we have $\mathbf{X}_2^{\mathrm{GM}} = \mathbf{A}_2 \boldsymbol{\beta}^{\mathrm{GM}}$. For all $i \in \{-m, \dots, n\}$ and $k \in \{0, \dots, n\}$ such that $i + k \geq n + 1$, this yields

$$Z_{i,k}^{\mathrm{GM}} = v_i \, \zeta_k^{\mathrm{GM}}$$

which is the first identity.
Corollary 3.3 also provides an identity for $\mathrm{var}[\mathbf{X}_2^{\mathrm{GM}} - \mathbf{X}_2]$, but in the present model the direct computation of the elements of this matrix seems to be more transparent. Consider $i, j \in \{-m, \dots, n\}$ and $k, l \in \{0, \dots, n\}$ such that $\min\{i+k, j+l\} \geq n+1$. Lemma 4.2 yields

$$\begin{aligned} \mathrm{cov}[Z_{i,k}^{\mathrm{GM}}, Z_{j,l}^{\mathrm{GM}}] &= \mathrm{cov}[v_i \zeta_k^{\mathrm{GM}}, v_j \zeta_l^{\mathrm{GM}}] \\ &= v_i v_j \, \mathrm{cov}[\zeta_k^{\mathrm{GM}}, \zeta_l^{\mathrm{GM}}] \\ &= v_i v_j \, \mathrm{var}[\zeta_k^{\mathrm{GM}}] \, \delta_{k,l} \end{aligned}$$

and we also have

$$\mathrm{cov}[Z_{i,k}, Z_{j,l}] = w_i \, \sigma_k^2 \, \delta_{i,j} \, \delta_{k,l} \; .$$

Since $Z_{i,k}^{\mathrm{GM}}$ and $Z_{j,l}^{\mathrm{GM}}$ are linear combinations of observable incremental losses whereas $Z_{i,k}$ and $Z_{j,l}$ are non–observable incremental losses, we have $\mathrm{cov}[Z_{i,k}^{\mathrm{GM}}, Z_{j,l}] = 0 = \mathrm{cov}[Z_{i,k}, Z_{j,l}^{\mathrm{GM}}]$ and hence

$$\begin{aligned} \mathrm{cov}[Z_{i,k}^{\mathrm{GM}} - Z_{i,k}, Z_{j,l}^{\mathrm{GM}} - Z_{j,l}] &= \mathrm{cov}[Z_{i,k}^{\mathrm{GM}}, Z_{j,l}^{\mathrm{GM}}] + \mathrm{cov}[Z_{i,k}, Z_{j,l}] \\ &= v_i v_j \, \mathrm{var}[\zeta_k^{\mathrm{GM}}] \, \delta_{k,l} + w_i \, \sigma_k^2 \, \delta_{i,j} \, \delta_{k,l}, \end{aligned}$$

which is the second identity. □

In a third step, we compute the Gauss–Markov predictors of reserves and their mean squared errors of prediction:

4.4 Theorem (Gauss–Markov prediction of reserves). *In the extended additive model,*
(1) *the Gauss–Markov predictors of the accident year reserves satisfy*

$$R_i^{\mathrm{GM}} = v_i \sum_{k=n-k+1}^{n} \zeta_k^{\mathrm{GM}}$$

and

$$\mathrm{cov}[R_i^{\mathrm{GM}} - R_i, R_j^{\mathrm{GM}} - R_j] = v_i v_j \sum_{k=n-i \wedge j+1}^{n} \mathrm{var}[\zeta_k^{\mathrm{GM}}] + w_i \left(\sum_{k=n-i+1}^{n} \sigma_k^2 \right) \delta_{i,j}$$

for all $i, j \in \{1, \dots, n\}$; in particular,

$$E[(R_i^{\mathrm{GM}} - R_i)^2] = v_i^2 \sum_{k=n-i+1}^{n} \mathrm{var}[\zeta_k^{\mathrm{GM}}] + w_i \sum_{k=n-i+1}^{n} \sigma_k^2$$

holds for all $i \in \{1, \dots, n\}$.

(2) *the Gauss–Markov predictors of the calendar year reserves satisfy*

$$R_{(c)}^{\mathrm{GM}} = \sum_{i=c-n}^{n} v_i\,\zeta_{c-i}^{\mathrm{GM}}$$

and

$$\mathrm{cov}[R_{(c)}^{\mathrm{GM}}-R_{(c)},R_{(d)}^{\mathrm{GM}}-R_{(d)}] = \sum_{i=c\vee d-n}^{n} v_i v_{i-|c-d|}\,\mathrm{var}[\zeta_{c\vee d-i}^{\mathrm{GM}}] + \left(\sum_{i=c-n}^{n} w_i\,\sigma_{c-i}^2\right)\delta_{c,d}$$

for all $c,d \in \{n+1,\ldots,2n\}$; *in particular,*

$$E[(R_{(c)}^{\mathrm{GM}}-R_{(c)})^2] = \sum_{i=c-n}^{n} v_i^2\,\mathrm{var}[\zeta_{c-i}^{\mathrm{GM}}] + \sum_{i=c-n}^{n} w_i\,\sigma_{c-i}^2$$

holds for all $c \in \{n+1,\ldots,2n\}$.

(3) *the Gauss–Markov predictor of the total reserve satisfies*

$$R^{\mathrm{GM}} = \sum_{k=1}^{n}\left(\sum_{i=n-k+1}^{n} v_i\right)\zeta_k^{\mathrm{GM}}$$

and

$$E[(R^{\mathrm{GM}}-R)^2] = \sum_{k=1}^{n}\left(\sum_{i=n-k+1}^{n} v_i\right)^2 \mathrm{var}[\zeta_k^{\mathrm{GM}}] + \sum_{k=1}^{n}\left(\sum_{i=n-k+1}^{n} w_i\right)\sigma_k^2\,.$$

Proof. Let us first consider the accident year reserves. We have

$$R_i = \sum_{k=n-i+1}^{n} Z_{i,k}$$

and, since Gauss–Markov prediction is linear, we obtain

$$R_i^{\mathrm{GM}} = \sum_{k=n-i+1}^{n} Z_{i,k}^{\mathrm{GM}}$$

which because of Lemma 4.3 gives the first identity. This yields

$$R_i^{\mathrm{GM}} - R_i = \sum_{k=n-i+1}^{n} (Z_{i,k}^{\mathrm{GM}}-Z_{i,k})$$

and because of Lemma 4.3 we obtain

$$\mathrm{cov}[R_i^{\mathrm{GM}}-R_i,R_j^{\mathrm{GM}}-R_j] = \mathrm{cov}\left[\sum_{k=n-i+1}^{n} (Z_{i,k}^{\mathrm{GM}}-Z_{i,k}),\ \sum_{l=n-j+1}^{n} (Z_{j,l}^{\mathrm{GM}}-Z_{j,l})\right]$$

$$
\begin{aligned}
&= \sum_{k=n-i+1}^{n} \sum_{l=n-j+1}^{n} \mathrm{cov}[Z^{\mathrm{GM}}_{i,k} - Z_{i,k}, Z^{\mathrm{GM}}_{j,l} - Z_{j,l}] \\
&= \sum_{k=n-i+1}^{n} \sum_{l=n-j+1}^{n} \Big(v_i v_j \,\mathrm{var}[\zeta^{\mathrm{GM}}_k] + w_i\, \sigma^2_k\, \delta_{i,j} \Big)\, \delta_{k,l} \\
&= \sum_{k=n-i\wedge j+1}^{n} \Big(v_i v_j \,\mathrm{var}[\zeta^{\mathrm{GM}}_k] + w_i\, \sigma^2_k\, \delta_{i,j} \Big) \\
&= v_i v_j \sum_{k=n-i\wedge j+1}^{n} \mathrm{var}[\zeta^{\mathrm{GM}}_k] + w_i \left(\sum_{k=n-i+1}^{n} \sigma^2_k \right) \delta_{i,j} \,,
\end{aligned}
$$

which is the second identity. Since Gauss–Markov predictors are unbiased, the third identity follows from the second.

Let us now consider the calendar year reserves. We have

$$
R_{(c)} = \sum_{i=c-n}^{n} Z_{i,c-i}
$$

and hence

$$
R^{\mathrm{GM}}_{(c)} = \sum_{i=c-n}^{n} Z^{\mathrm{GM}}_{i,c-i}
$$

which because of Lemma 4.3 gives the first identity. This yields

$$
R^{\mathrm{GM}}_{(c)} - R_{(c)} = \sum_{i=c-n}^{n} (Z^{\mathrm{GM}}_{i,c-i} - Z_{i,c-i})
$$

and because of Lemma 4.3 we obtain

$$
\begin{aligned}
\mathrm{cov}[R^{\mathrm{GM}}_{(c)} - R_{(c)}, R^{\mathrm{GM}}_{(d)} - R_{(d)}] &= \mathrm{cov}\left[\sum_{i=c-n}^{n} (Z^{\mathrm{GM}}_{i,c-i} - Z_{i,c-i}) \,, \sum_{j=d-n}^{n} (Z^{\mathrm{GM}}_{j,d-j} - Z_{j,d-j}) \right] \\
&= \sum_{i=c-n}^{n} \sum_{j=d-n}^{n} \mathrm{cov}[(Z^{\mathrm{GM}}_{i,c-i} - Z_{i,c-i}), (Z^{\mathrm{GM}}_{j,d-j} - Z_{j,d-j})] \Big] \\
&= \sum_{i=c-n}^{n} \sum_{j=d-n}^{n} \Big(v_i v_j \,\mathrm{var}[\zeta^{\mathrm{GM}}_{c-i}] + w_i\, \sigma^2_{c-i}\, \delta_{i,j} \Big)\, \delta_{c-i,d-j} \\
&= \sum_{i=c\vee d-n}^{n} v_i v_{i-|c-d|} \,\mathrm{var}[\zeta^{\mathrm{GM}}_{c\vee d-i}] + \left(\sum_{i=c-n}^{n} w_i\, \sigma^2_{c-i} \right) \delta_{c,d}
\end{aligned}
$$

which is the second identity.

Let us finally consider the total reserve. We have

$$
R = \sum_{k=1}^{n} \sum_{i=k+1}^{n} Z_{i,k}
$$

and, since Gauss–Markov prediction is linear, we obtain

$$R^{\mathrm{GM}} \;=\; \sum_{k=1}^{n} \sum_{i=k+1}^{n} Z_{i,k}^{\mathrm{GM}}$$

which because of Lemma 4.3 gives the first identity. This yields

$$R^{\mathrm{GM}} - R \;=\; \sum_{k=1}^{n} \sum_{i=k+1}^{n} (Z_{i,k}^{\mathrm{GM}} - Z_{i,k})$$

and because of Lemma 4.3 we obtain

$$\begin{aligned}
E[(R^{\mathrm{GM}}-R)^2] &= \mathrm{var}[R^{\mathrm{GM}}-R] \\
&= \mathrm{var}\left[\sum_{k=1}^{n} \sum_{i=n-k+1}^{n} (Z_{i,k}^{\mathrm{GM}} - Z_{i,k})\right] \\
&= \sum_{k=1}^{n} \mathrm{var}\left[\sum_{i=n-k+1}^{n} (Z_{i,k}^{\mathrm{GM}} - Z_{i,k})\right] \\
&= \sum_{k=1}^{n} \sum_{i=n-k+1}^{n} \sum_{j=n-k+1}^{n} \mathrm{cov}[Z_{i,k}^{\mathrm{GM}} - Z_{i,k}, Z_{j,k}^{\mathrm{GM}} - Z_{j,k}] \\
&= \sum_{k=1}^{n} \sum_{i=n-k+1}^{n} \sum_{j=n-k+1}^{n} \Big(v_i v_j \, \mathrm{var}[\zeta_k^{\mathrm{GM}}] + w_i \, \sigma_k^2 \, \delta_{i,j} \Big) \\
&= \sum_{k=1}^{n} \left(\sum_{i=n-k+1}^{n} v_i \right)^2 \mathrm{var}[\zeta_k^{\mathrm{GM}}] + \sum_{k=1}^{n} \left(\sum_{i=n-k+1}^{n} w_i \right) \sigma_k^2
\end{aligned}$$

which is the second identity. □

In the special case where $m = 0$ and $w_i = v_i$ holds for all $i \in \{-m, \ldots, n\}$, the Gauss–Markov predictors of incremental losses and reserves are identical with the predictors used in the traditional additive method of loss reserving. This means that Gauss–Markov prediction in the extended additive model provides simultaneously an extension of the additive method and its justification based on a general statistical principle.

Via Lemma 4.3 and Lemma 4.2, the mean squared errors of prediction depend on unknown variance parameters which may be estimated as follows:

4.5 Theorem (Estimation of variance parameters). *In the extended additive model with $m \geq 1$ and for every $k \in \{0, \ldots, n\}$, the random variable*

$$\widehat{\sigma}_k^2 \;:=\; \frac{1}{m+n-k} \sum_{i=-m}^{n-k} \frac{1}{w_i} \, (Z_{i,k} - v_i \zeta_k^{\mathrm{GM}})^2$$

is an unbiased estimator of σ_k^2.

Proof. Consider $i \in \{-m, \ldots, n-k\}$. By Lemma 4.2, we have $E[v_i \zeta_k^{\mathrm{GM}}] = E[Z_{i,k}]$ and thus

$$
\begin{aligned}
E[(Z_{i,k} - v_i \zeta_k^{\mathrm{GM}})^2] &= \mathrm{var}[Z_{i,k} - v_i \zeta_k^{\mathrm{GM}}] \\
&= \mathrm{var}[Z_{i,k}] - 2\, v_i \,\mathrm{cov}[Z_{i,k}, \zeta_k^{\mathrm{GM}}] + v_i^2 \,\mathrm{var}[\zeta_k^{\mathrm{GM}}] \;.
\end{aligned}
$$

Recall that

$$
\mathrm{var}[Z_{i,k}] = w_i\, \sigma_k^2 \;.
$$

Furthermore, using Lemma 4.2 and Lemma 4.3 we obtain

$$
\begin{aligned}
\mathrm{cov}[Z_{i,k}, \zeta_k^{\mathrm{GM}}] &= \mathrm{cov}\left[Z_{i,k}, \frac{\sum_{j=-m}^{n-k} v_j Z_{j,k}/w_j}{\sum_{j=-m}^{n-k} v_j^2/w_j} \right] \\
&= \sum_{j=-m}^{n-k} \frac{v_j/w_j}{\sum_{h=-m}^{n-k} v_h^2/w_h} \,\mathrm{cov}[Z_{i,k}, Z_{j,k}] \\
&= \sum_{j=-m}^{n-k} \frac{v_j/w_j}{\sum_{h=-m}^{n-k} v_h^2/w_h} \, w_i\, \sigma_k^2\, \delta_{i,j} \\
&= v_i \, \frac{1}{\sum_{h=-m}^{n-k} v_h^2/w_h} \, \sigma_k^2
\end{aligned}
$$

and Lemma 4.2 yields

$$
\mathrm{var}[\zeta_k^{\mathrm{GM}}] = \frac{1}{\sum_{h=-m}^{n-k} v_h^2/w_h} \, \sigma_k^2 \;.
$$

Therefore, we have

$$
\begin{aligned}
E[(Z_{i,k} - v_i \zeta_k^{\mathrm{GM}})^2] &= \mathrm{var}[Z_{i,k}] - 2\, v_i \,\mathrm{cov}[Z_{i,k}, \zeta_k^{\mathrm{GM}}] + v_i^2 \,\mathrm{var}[\zeta_k^{\mathrm{GM}}] \\
&= w_i\, \sigma_k^2 - 2\, v_i^2 \, \frac{1}{\sum_{h=-m}^{n-k} v_h^2/w_h} \, \sigma_k^2 + v_i^2 \, \frac{1}{\sum_{h=-m}^{n-k} v_h^2/w_h} \, \sigma_k^2 \\
&= \left(w_i - \frac{v_i^2}{\sum_{h=-m}^{n-k} v_h^2/w_h} \right) \sigma_k^2
\end{aligned}
$$

and hence

$$
\begin{aligned}
\sum_{i=-m}^{n-k} \frac{1}{w_i}\, E[(Z_{i,k} - v_i \zeta_k^{\mathrm{GM}})^2] &= \sum_{i=-m}^{n-k} \frac{1}{w_i} \left(w_i - \frac{v_i^2}{\sum_{h=-m}^{n-k} v_h^2/w_h} \right) \sigma_k^2 \\
&= \Big((m+1+n-k) - 1 \Big)\, \sigma_k^2 \\
&= (m+n-k)\, \sigma_k^2
\end{aligned}
$$

which proves the assertion. □

In the case $m = 0$, the assertion of Theorem 4.5 remains valid for $k \in \{0, \ldots, n-1\}$. To obtain an estimator of σ_n^2 also in this case, one may choose a parametric class $\{f_c \mid c \in C\}$ of real functions (e.g.,the class $\{f_{(a,b)} : \mathbb{R} \to \mathbb{R} \mid (a,b) \in (0,\infty)^2\}$ with $f_{(a,b)}(x) = a\, e^{-bx}$), determine $\widehat{c} \in C$ satisfying

$$\sum_{k=0}^{n-1} \Bigl(f_{\widehat{c}}(k) - \widehat{\sigma}_k^2 \Bigr)^2 \;=\; \inf_{c \in C} \sum_{k=0}^{n-1} \Bigl(f_c(k) - \widehat{\sigma}_k^2 \Bigr)^2$$

and define

$$\widehat{\sigma}_n^2 \;:=\; f_{\widehat{c}}(n)$$

If the sequence $\{\widehat{\sigma}_k^2\}_{k \in \{0,\ldots,n-1\}}$ is decreasing, one might alternatively define $\widehat{\sigma}_n^2 := \widehat{\sigma}_{n-1}^2$.

Now estimators of the mean squared errors of prediction can be obtained by replacing the variance parameters by their estimators in the formulas for the mean squared errors of prediction.

5 Gauss–Markov Loss Prediction in the Extended Panning Model

In the present section, we denote by $\mathcal{F}_0$ the σ–algebra generated by the family $\{Z_{i,0}\}_{i\in\{-m,\ldots,n\}}$ of the losses of development year 0. The extended Panning model is defined as follows:

> **Extended Panning Model:** *There exist known $\mathcal{F}_0$–measurable random parameters w_i with $w_i > 0$ and $i \in \{-m, \ldots, n\}$ as well as unknown parameters $\xi_k \in \mathbb{R}$ and $\sigma_k^2 \in (0, \infty)$ with $k \in \{0, \ldots, n\}$ such that the incremental losses satisfy*
>
> $$\begin{aligned} E^{\mathcal{F}_0}[Z_{i,k}] &= Z_{i,0}\,\xi_k \\ \mathrm{cov}^{\mathcal{F}_0}[Z_{i,k}, Z_{j,l}] &= w_i\,\sigma_k^2\,\delta_{i,j}\,\delta_{k,l} \end{aligned}$$
>
> *for all $i, j \in \{-m, \ldots, n\}$ and $k, l \in \{0, \ldots, n\}$. Moreover, $Z_{i,0} > 0$ holds for all $i \in \{-m, \ldots, n\}$.*

In the extended Panning model, the *initial losses* $Z_{i,0}$ replace the volume measures used in the extended additive model. Since the first identity in the extended Panning model implies

$$E[Z_{i,k}/Z_{i,0}] = \xi_k$$

the development year parameter ξ_k is assumed to be independent of the accident year such that the collection of these parameters forms a *development pattern*; see Schmidt and Zocher [2009].

The extended Panning model extends the traditional Panning model in which it is assumed that $m = 0$ and that $w_i = 1$ holds for all $i \in \{-m, \ldots, n\}$; see Panning [2006] and Schmidt and Zocher [2009]. The reason for considering the extended Panning model becomes evident from its comparison with the extended additive model (Section 4) and with the combination of both models (Section 6).

The following results are entirely analogous to those for the extended additive model and can be obtained by replacing the volume measures v_i used in the extended additive model by the initial losses $Z_{i,0}$ and by replacing the first and second order moments by their $\mathcal{F}_0$–conditional counterparts.

5.1 Theorem. *The extended Panning model is an $\mathcal{F}_0$–conditional linear model.*

5.2 Lemma (Gauss–Markov estimation of parameters). *In the extended Panning model, the $\mathcal{F}_0$–conditional Gauss–Markov estimators of the coordinates of the parameter vector satisfy*

$$\xi_k^{\mathrm{GM}} = \frac{\sum_{i=-m}^{n-k} Z_{i,0} Z_{i,k}/w_i}{\sum_{i=-m}^{n-k} Z_{i,0}^2/w_i}$$

and

$$\operatorname{cov}^{\mathcal{F}_0}[\xi_k^{\mathrm{GM}}, \xi_l^{\mathrm{GM}}] \;=\; \frac{1}{\sum_{i=-m}^{n-k} Z_{i,0}^2/w_i}\, \sigma_k^2\, \delta_{k,l}$$

for all $k, l \in \{0, 1, \dots, n\}$.

5.3 Lemma (Gauss–Markov prediction of incremental losses). *In the extended Panning model, the* $\mathcal{F}_0$*–conditional Gauss–Markov predictors of the non–observable incremental losses satisfy*

$$Z_{i,k}^{\mathrm{GM}} \;=\; Z_{i,0}\, \xi_k^{\mathrm{GM}}$$

and

$$\operatorname{cov}^{\mathcal{F}_0}[Z_{i,k}^{\mathrm{GM}} - Z_{i,k}, Z_{j,l}^{\mathrm{GM}} - Z_{j,l}] \;=\; \Big(Z_{i,0} Z_{j,0}\, \operatorname{var}^{\mathcal{F}_0}[\xi_k^{\mathrm{GM}}] + w_i\, \sigma_k^2\, \delta_{i,j}\Big)\, \delta_{k,l}$$

for all $i, j \in \{-m, \dots, n\}$ *and* $k, l \in \{0, \dots, n\}$ *such that* $\min\{i+k, j+l\} \geq n+1$.

5.4 Theorem (Gauss–Markov prediction of reserves). *In the extended Panning model,*

(1) *the* $\mathcal{F}_0$*–conditional Gauss–Markov predictors of the accident year reserves satisfy*

$$R_i^{\mathrm{GM}} \;=\; Z_{i,0} \sum_{k=n-k+1}^{n} \xi_k^{\mathrm{GM}}$$

and

$$\operatorname{cov}^{\mathcal{F}_0}[R_i^{\mathrm{GM}} - R_i, R_j^{\mathrm{GM}} - R_j] \;=\; Z_{i,0} Z_{j,0} \sum_{k=n-i\wedge j+1}^{n} \operatorname{var}^{\mathcal{F}_0}[\xi_k^{\mathrm{GM}}] + w_i \left(\sum_{k=n-i+1}^{n} \sigma_k^2 \right) \delta_{i,j}$$

for all $i, j \in \{1, \dots, n\}$; *in particular,*

$$E^{\mathcal{F}_0}[(R_i^{\mathrm{GM}} - R_i)^2] \;=\; Z_{i,0}^2 \sum_{k=n-i+1}^{n} \operatorname{var}^{\mathcal{F}_0}[\xi_k^{\mathrm{GM}}] + w_i \sum_{k=n-i+1}^{n} \sigma_k^2$$

holds for all $i \in \{1, \dots, n\}$.

(2) *the* $\mathcal{F}_0$*–conditional Gauss–Markov predictors of the calendar year reserves satisfy*

$$R_{(c)}^{\mathrm{GM}} \;=\; \sum_{i=c-n}^{n} Z_{i,0}\, \xi_{c-i}^{\mathrm{GM}}$$

and

$$\begin{aligned}
&\operatorname{cov}^{\mathcal{F}_0}[R_{(c)}^{\mathrm{GM}} - R_{(c)}, R_{(d)}^{\mathrm{GM}} - R_{(d)}] \\
&\quad= \sum_{i=c\vee d-n}^{n} Z_{i,0} Z_{i-|c-d|,0}\, \operatorname{var}^{\mathcal{F}_0}[\xi_{c\vee d-i}^{\mathrm{GM}}] + \left(\sum_{i=c-n}^{n} w_i\, \sigma_{c-i}^2 \right) \delta_{c,d}
\end{aligned}$$

for all $c, d \in \{n{+}1, \ldots, 2n\}$; *in particular,*

$$E^{\mathcal{F}_0}[(R^{\mathrm{GM}}_{(c)} - R_{(c)})^2] \;=\; \sum_{i=c-n}^{n} Z_{i,0}^2 \,\mathrm{var}^{\mathcal{F}_0}[\xi^{\mathrm{GM}}_{c-i}] + \sum_{i=c-n}^{n} w_i\, \sigma^2_{c-i}$$

holds for all $c \in \{n{+}1, \ldots, 2n\}$.

(3) *the* $\mathcal{F}_0$*–conditional Gauss–Markov predictor of the total reserve satisfies*

$$R^{\mathrm{GM}} \;=\; \sum_{k=1}^{n} \left(\sum_{i=n-k+1}^{n} Z_{i,0} \right) \xi^{\mathrm{GM}}_k$$

and

$$E^{\mathcal{F}_0}[(R^{\mathrm{GM}} - R)^2] \;=\; \sum_{k=1}^{n} \left(\sum_{i=n-k+1}^{n} Z_{i,0} \right)^2 \mathrm{var}^{\mathcal{F}_0}[\xi^{\mathrm{GM}}_k] + \sum_{k=1}^{n} \left(\sum_{i=n-k+1}^{n} w_i \right) \sigma^2_k \,.$$

In the special case where $m = 0$ and $w_i = 1$ holds for all $i \in \{-m, \ldots, n\}$, the Gauss–Markov predictors of incremental losses and reserves are identical with the predictors used in the traditional Panning method of loss reserving. This means that Gauss–Markov prediction in the extended Panning model provides simultaneously an extension of the Panning method and its justification based on a general statistical principle.

The unknown variance parameters may be estimated as follows:

5.5 Theorem (Estimation of variance parameters). *In the extended Panning model with* $m \geq 1$ *and for every* $k \in \{0, \ldots, n\}$, *the random variable*

$$\widehat{\sigma}^2_k \;:=\; \frac{1}{m+n-k} \sum_{i=-m}^{n-k} \frac{1}{w_i} (Z_{i,k} - Z_{i,0}\xi^{\mathrm{GM}}_k)^2$$

is an $\mathcal{F}_0$*–conditionally unbiased estimator of* σ^2_k.

The final remarks of Section 4 apply to the extended Panning model as well.

6 Gauss–Markov Loss Prediction in the Combined Model

Because of the similarity of the extended additive model and the extended Panning model, it is natural to consider convex combinations of these models. As in the previous section, we denote by $\mathcal{F}_0$ the σ–algebra generated by the family $\{Z_{i,0}\}_{i\in\{-m,\dots,n\}}$ of the losses of development year 0. The combined model is defined as follows:

Combined Model: *There exist known $\mathcal{F}_0$–measurable random parameters v_i, w_i with $v_i, w_i > 0$ and $i \in \{-m, \dots, n\}$ as well as unknown parameters $\zeta_k, \xi_k \in \mathbb{R}$ and $\sigma_k^2 \in (0,\infty)$ with $k \in \{0,\dots,n\}$ such that the incremental losses satisfy*

$$\begin{aligned} E^{\mathcal{F}_0}[Z_{i,k}] &= v_i\,\zeta_k + Z_{i,0}\,\xi_k \\ \mathrm{cov}^{\mathcal{F}_0}[Z_{i,k}, Z_{j,l}] &= w_i\,\sigma_k^2\,\delta_{i,j}\,\delta_{k,l} \end{aligned}$$

for all $i,j \in \{-m,\dots,n\}$ and $k,l \in \{0,\dots,n\}$. Moreover, $Z_{i,0} > 0$ holds for all $i \in \{-m,\dots,n\}$ and $v_i Z_{j,0} \neq v_j Z_{i,0}$ holds for some $i,j \in \{-m,\dots,0\}$ with $i \neq j$.

It is evident that the combined model combines the extended additive model and the extended Panning model: Formally, putting $\xi_k := 0$ yields the extended additive model and putting $\zeta_k := 0$ yields the extended Panning model. However, the analysis of the combined model turns out to be a bit more subtle than the analysis of the extended additive and Panning models.

Assume that the assumptions of the combined model are fulfilled. Then the $\mathcal{F}_0$–conditional expectation of the random vector $\mathbf{X}_1$ of all observable incremental losses satisfies

$$E^{\mathcal{F}_0}\left[\begin{pmatrix} Z_{-m,0} \\ \vdots \\ Z_{n,0} \\ \hline \vdots \\ \hline Z_{-m,k} \\ \vdots \\ Z_{n-k,k} \\ \hline \vdots \\ \hline Z_{-m,n} \\ \vdots \\ Z_{0,n} \end{pmatrix}\right] = \left(\begin{array}{ccccc|ccccc} v_{-m} & \cdots & 0 & \cdots & 0 & Z_{-m,0} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ v_n & \cdots & 0 & \cdots & 0 & Z_{n,0} & \cdots & 0 & \cdots & 0 \\ \hline \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \hline 0 & \cdots & v_{-m} & \cdots & 0 & 0 & \cdots & Z_{-m,0} & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & v_{n-k} & \cdots & 0 & 0 & \cdots & Z_{n-k,0} & \cdots & 0 \\ \hline \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \hline 0 & \cdots & 0 & \cdots & v_{-m} & 0 & \cdots & 0 & \cdots & Z_{-m,0} \\ \vdots & & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & v_0 & 0 & \cdots & 0 & \cdots & Z_{0,0} \end{array}\right) \begin{pmatrix} \zeta_0 \\ \vdots \\ \zeta_k \\ \vdots \\ \zeta_n \\ \hline \xi_0 \\ \vdots \\ \xi_k \\ \vdots \\ \xi_n \end{pmatrix}$$

such that there exist an $\mathcal{F}_0$–measurable random design matrix $\mathbf{A}_1$ having full column rank and a parameter vector $\boldsymbol{\beta}$ satisfying $E^{\mathcal{F}_0}[\mathbf{X}_1] = \mathbf{A}_1\boldsymbol{\beta}$. A similar identity holds for the $\mathcal{F}_0$–conditional expectation of the random vector $\mathbf{X}_2$ of all non–observable incremental losses. Moreover, the $\mathcal{F}_0$–conditional variance of the random vector $\mathbf{X}$ is the same as in the extended additive model and the extended Panning model.

6.1 Theorem. *The combined model is an $\mathcal{F}_0$–conditional linear model.*

For a concise and transparent presentation of the result for the combined model, we now introduce some auxiliary random variables. By assumption, we have

$$\sum_{i=-m}^{0} \sum_{j=-m}^{0} (v_i Z_{j,0} - v_j Z_{i,0})^2 \;>\; 0$$

We may thus define, for $k \in \{0, \ldots, n\}$ and $r, s \in \{0, 1, 2\}$,

$$Y_k^{(r,s)} \;:=\; 2\, \frac{\sum_{i=-m}^{n-k} v_i^r Z_{i,0}^s / w_i}{\sum_{i=-m}^{n-k} \sum_{j=-m}^{n-k} (v_i Z_{j,0} - v_j Z_{i,0})^2 / w_i w_j}\,.$$

and straightforward calculation shows that

$$Y_k^{(r,s)} \;=\; \frac{\sum_{i=-m}^{n-k} v_i^r Z_{i,0}^s / w_i}{(\sum_{i=-m}^{n-k} v_i^2 / w_i)(\sum_{i=-m}^{n-k} Z_{i,0}^2 / w_i) - (\sum_{i=-m}^{n-k} v_i Z_{i,0} / w_i)^2}\,.$$

Note that these random variables are $\mathcal{F}_0$–measurable.

6.2 Lemma (Gauss–Markov estimation of parameters). *In the combined model, the $\mathcal{F}_0$–conditional Gauss–Markov estimators of the coordinates of the parameter vector satisfy*

$$\begin{aligned} \zeta_k^{\mathrm{GM}} &= Y_k^{(0,2)} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} - Y_k^{(1,1)} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} \\ \xi_k^{\mathrm{GM}} &= Y_k^{(2,0)} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} - Y_k^{(1,1)} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} \end{aligned}$$

as well as

$$\begin{aligned} \mathrm{cov}^{\mathcal{F}_0}[\zeta_k^{\mathrm{GM}}, \zeta_l^{\mathrm{GM}}] &= Y_k^{(0,2)}\, \sigma_k^2\, \delta_{k,l} \\ \mathrm{cov}^{\mathcal{F}_0}[\zeta_k^{\mathrm{GM}}, \xi_l^{\mathrm{GM}}] &= -\,Y_k^{(1,1)}\, \sigma_k^2\, \delta_{k,l} \\ \mathrm{cov}^{\mathcal{F}_0}[\xi_k^{\mathrm{GM}}, \xi_l^{\mathrm{GM}}] &= Y_k^{(2,0)}\, \sigma_k^2\, \delta_{k,l} \end{aligned}$$

and, in particular,

$$\mathrm{cov}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix}, \begin{pmatrix} \zeta_l^{\mathrm{GM}} \\ \xi_l^{\mathrm{GM}} \end{pmatrix}\right] = \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix}\right] \delta_{k,l}$$

for all $k, l \in \{0, 1, \ldots, n\}$.

Proof. We have

$$\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{X}_1 = \begin{pmatrix} \left(\frac{1}{\sigma_k^2} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} \right)_{k \in \{0,\dots,n\}} \\ \left(\frac{1}{\sigma_k^2} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} \right)_{k \in \{0,\dots,n\}} \end{pmatrix}$$

and

$$\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_1 = \begin{pmatrix} \operatorname{diag}\left(\frac{1}{\sigma_k^2} \sum_{i=-m}^{n-k} \frac{v_i^2}{w_i} \right)_{k \in \{0,\dots,n\}} & \operatorname{diag}\left(\frac{1}{\sigma_k^2} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,0}}{w_i} \right)_{k \in \{0,\dots,n\}} \\ \operatorname{diag}\left(\frac{1}{\sigma_k^2} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,0}}{w_i} \right)_{k \in \{0,\dots,n\}} & \operatorname{diag}\left(\frac{1}{\sigma_k^2} \sum_{i=-m}^{n-k} \frac{Z_{i,0}^2}{w_i} \right)_{k \in \{0,\dots,n\}} \end{pmatrix}.$$

Therefore, we have

$$\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_1 = \begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V} & \mathbf{W} \end{pmatrix}$$

with suitable diagonal matrices $\mathbf{U}, \mathbf{V}, \mathbf{W}$, and we also have

$$\begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V} & \mathbf{W} \end{pmatrix}^{-1} = \begin{pmatrix} (\mathbf{U} - \mathbf{V}\mathbf{W}^{-1}\mathbf{V})^{-1} & -\mathbf{U}^{-1}\mathbf{V}(\mathbf{W} - \mathbf{V}\mathbf{U}^{-1}\mathbf{V})^{-1} \\ -\mathbf{W}^{-1}\mathbf{V}(\mathbf{U} - \mathbf{V}\mathbf{W}^{-1}\mathbf{V})^{-1} & (\mathbf{W} - \mathbf{V}\mathbf{U}^{-1}\mathbf{V})^{-1} \end{pmatrix}.$$

Therefore, straightforward calculation yields

$$(\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_1)^{-1} = \begin{pmatrix} \operatorname{diag}(Y_k^{(0,2)} \sigma_k^2)_{k \in \{0,\dots,n\}} & -\operatorname{diag}(Y_k^{(1,1)} \sigma_k^2)_{k \in \{0,\dots,n\}} \\ -\operatorname{diag}(Y_k^{(1,1)} \sigma_k^2)_{k \in \{0,\dots,n\}} & \operatorname{diag}(Y_k^{(2,0)} \sigma_k^2)_{k \in \{0,\dots,n\}} \end{pmatrix}.$$

We thus obtain

$$(\mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{A}_1)^{-1} \mathbf{A}_1' \boldsymbol{\Sigma}_{11}^{-1} \mathbf{X}_1 = \begin{pmatrix} \left(Y_k^{(0,2)} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} - Y_k^{(1,1)} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} \right)_{k \in \{0,\dots,n\}} \\ \left(Y_k^{(2,0)} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} - Y_k^{(1,1)} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} \right)_{k \in \{0,\dots,n\}} \end{pmatrix}$$

and hence

$$\begin{aligned} \zeta_k^{\text{GM}} &= Y_k^{(0,2)} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} - Y_k^{(1,1)} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} \\ \xi_k^{\text{GM}} &= Y_k^{(2,0)} \sum_{i=-m}^{n-k} \frac{Z_{i,0} Z_{i,k}}{w_i} - Y_k^{(1,1)} \sum_{i=-m}^{n-k} \frac{v_i Z_{i,k}}{w_i} \,. \end{aligned}$$

The above identity for $(\mathbf{A}_1'\mathbf{\Sigma}_{11}^{-1}\mathbf{A}_1)^{-1}$ also yields

$$\begin{aligned}
\mathrm{cov}^{\mathcal{F}_0}[\zeta_k^{\mathrm{GM}}, \zeta_l^{\mathrm{GM}}] &= \;\; Y_k^{(0,2)}\, \sigma_k^2\, \delta_{k,l} \\
\mathrm{cov}^{\mathcal{F}_0}[\zeta_k^{\mathrm{GM}}, \xi_l^{\mathrm{GM}}] &= -Y_k^{(1,1)}\, \sigma_k^2\, \delta_{k,l} \\
\mathrm{cov}^{\mathcal{F}_0}[\xi_k^{\mathrm{GM}}, \xi_l^{\mathrm{GM}}] &= \;\; Y_k^{(2,0)}\, \sigma_k^2\, \delta_{k,l}
\end{aligned}$$

which completes the proof. □

We can now compute the Gauss–Markov predictors of the non–observable incremental losses and the covariances of their prediction errors:

6.3 Lemma (Gauss–Markov prediction of incremental losses). *In the combined model, the $\mathcal{F}_0$–conditional Gauss–Markov predictors of the non–observable incremental losses satisfy*

$$Z_{i,k}^{\mathrm{GM}} = v_i\, \zeta_k^{\mathrm{GM}} + Z_{i,0}\, \xi_k^{\mathrm{GM}}$$

and

$$\begin{aligned}
&\mathrm{cov}^{\mathcal{F}_0}[Z_{i,k}^{\mathrm{GM}} - Z_{i,k}, Z_{j,l}^{\mathrm{GM}} - Z_{j,l}] \\
&= \left(\begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix} \right] \begin{pmatrix} v_j \\ Z_{j,0} \end{pmatrix} + w_i \sigma_k^2 \delta_{i,j} \right) \delta_{k,l}
\end{aligned}$$

for all $i,j \in \{-m,\dots,n\}$ and $k,l \in \{0,\dots,n\}$ such that $\min\{i+k, j+l\} \geq n+1$; in particular,

$$E^{\mathcal{F}_0}[(Z_{i,k}^{\mathrm{GM}} - Z_{i,k})^2] = \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix} \right] \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix} + w_i\, \sigma_k^2$$

holds for all $i \in \{-m,\dots,n\}$ and $k \in \{0,\dots,n\}$ such that $i+k \geq n+1$.

Proof. The first identity is evident. Furthermore, Lemma 6.2 yields

$$\begin{aligned}
\mathrm{cov}^{\mathcal{F}_0}[Z_{i,k}^{\mathrm{GM}}, Z_{j,l}^{\mathrm{GM}}] &= \mathrm{cov}^{\mathcal{F}_0}[v_i\, \zeta_k^{\mathrm{GM}} + Z_{i,0}\xi_k^{\mathrm{GM}}, v_j\, \zeta_l^{\mathrm{GM}} + Z_{j,0}\, \xi_l^{\mathrm{GM}}] \\
&= \mathrm{cov}^{\mathcal{F}_0}\left[\begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix}, \begin{pmatrix} v_j \\ Z_{j,0} \end{pmatrix}' \begin{pmatrix} \zeta_l^{\mathrm{GM}} \\ \xi_l^{\mathrm{GM}} \end{pmatrix} \right] \\
&= \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{cov}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix}, \begin{pmatrix} \zeta_l^{\mathrm{GM}} \\ \xi_l^{\mathrm{GM}} \end{pmatrix} \right] \begin{pmatrix} v_j \\ Z_{j,0} \end{pmatrix} \\
&= \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix} \right] \begin{pmatrix} v_j \\ Z_{j,0} \end{pmatrix} \delta_{k,l} \; .
\end{aligned}$$

Since $\mathrm{cov}^{\mathcal{F}_0}[Z_{i,k}^{\mathrm{GM}}, Z_{j,l}] = 0 = \mathrm{cov}^{\mathcal{F}_0}[Z_{i,k}, Z_{j,l}^{\mathrm{GM}}]$ and

$$\mathrm{cov}^{\mathcal{F}_0}[Z_{i,k} Z_{j,l}] = w_i\, \sigma_k^2\, \delta_{i,j}\delta_{k,l}$$

we obtain

$$\begin{aligned}
&\mathrm{cov}^{\mathcal{F}_0}[Z^{\mathrm{GM}}_{i,k} - Z_{i,k}, Z^{\mathrm{GM}}_{j,l} - Z_{j,l}] \\
&= \mathrm{cov}^{\mathcal{F}_0}[Z^{\mathrm{GM}}_{i,k}, Z^{\mathrm{GM}}_{j,l}] + \mathrm{cov}^{\mathcal{F}_0}[Z_{i,k} Z_{j,l}] \\
&= \left(\begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta^{\mathrm{GM}}_k \\ \xi^{\mathrm{GM}}_k \end{pmatrix} \right] \begin{pmatrix} v_j \\ Z_{j,0} \end{pmatrix} \delta_{k,l} + w_i \sigma^2_k \delta_{i,j} \right) \delta_{k,l}
\end{aligned}$$

which is the second identity. □

The following result on the Gauss–Markov predictors of reserves and their expected squared prediction errors is formally identical with the results for the extended additive model and the extended Panning model:

6.4 Theorem (Gauss–Markov prediction of reserves). *In the combined model,*

(1) *the $\mathcal{F}_0$–conditional Gauss–Markov predictors of the accident year reserves satisfy*

$$R^{\mathrm{GM}}_i = v_i \sum_{k=n-k+1}^{n} \zeta^{\mathrm{GM}}_k + Z_{i,0} \sum_{k=n-k+1}^{n} \xi^{\mathrm{GM}}_k$$

and

$$\begin{aligned}
&\mathrm{cov}^{\mathcal{F}_0}[R^{\mathrm{GM}}_i - R_i, R^{\mathrm{GM}}_j - R_j] \\
&= \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \left(\sum_{k=n-i\wedge j+1}^{n} \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta^{\mathrm{GM}}_k \\ \xi^{\mathrm{GM}}_k \end{pmatrix} \right] \right) \begin{pmatrix} v_j \\ Z_{j,0} \end{pmatrix} + w_i \left(\sum_{k=n-i+1} \sigma^2_k \right) \delta_{i,j}
\end{aligned}$$

for all $i, j \in \{1, \dots, n\}$; in particular,

$$\begin{aligned}
&E^{\mathcal{F}_0}[(R^{\mathrm{GM}}_i - R_i)^2] \\
&= \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \left(\sum_{k=n-i+1}^{n} \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta^{\mathrm{GM}}_k \\ \xi^{\mathrm{GM}}_k \end{pmatrix} \right] \right) \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix} + w_i \sum_{k=n-i+1} \sigma^2_k
\end{aligned}$$

holds for all $i \in \{1, \dots, n\}$.

(2) *the $\mathcal{F}_0$–conditional Gauss–Markov predictors of the calendar year reserves satisfy*

$$R^{\mathrm{GM}}_{(c)} = \sum_{i=c-n}^{n} \left(v_i \, \zeta^{\mathrm{GM}}_{c-i} + Z_{i,0} \, \xi^{\mathrm{GM}}_{c-i} \right)$$

and

$$\begin{aligned}
&\mathrm{cov}^{\mathcal{F}_0}[R^{\mathrm{GM}}_{(c)} - R_{(c)}, R^{\mathrm{GM}}_{(d)} - R_{(d)}] \\
&= \sum_{i=c\vee d-n}^{n} \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta^{\mathrm{GM}}_{c\vee d-i} \\ \xi^{\mathrm{GM}}_{c\vee d-i} \end{pmatrix} \right] \begin{pmatrix} v_{i-|c-d|} \\ Z_{i-|c-d|,0} \end{pmatrix} + \left(\sum_{i=c\vee d-n}^{n} w_i \sigma^2_{c-i} \right) \delta_{c,d}
\end{aligned}$$

for all $c, d \in \{n{+}1, \ldots, 2n\}$; *in particular,*

$$E^{\mathcal{F}_0}[(R_{(c)}^{\mathrm{GM}} - R_{(c)})^2] = \sum_{i=c-n}^{n} \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_{c-i}^{\mathrm{GM}} \\ \xi_{c-i}^{\mathrm{GM}} \end{pmatrix}\right] \begin{pmatrix} v_i \\ Z_{i,0} \end{pmatrix} + \sum_{i=c-n}^{n} w_i \sigma_{c-i}^2$$

holds for all $c \in \{n{+}1, \ldots, 2n\}$.

(3) *the* $\mathcal{F}_0$*–conditional Gauss–Markov predictor of the total reserve satisfies*

$$R^{\mathrm{GM}} = \sum_{k=1}^{n} \sum_{i=n-k+1}^{n} \Big(v_i\, \zeta_k^{\mathrm{GM}} + Z_{i,0}\, \xi_k^{\mathrm{GM}} \Big)$$

and

$$E^{\mathcal{F}_0}[(R_i^{\mathrm{GM}} - R_i)^2] = \sum_{k=1}^{n} \begin{pmatrix} \sum_{i=n-k+1}^{n} v_i \\ \sum_{i=n-k+1}^{n} Z_{i,0} \end{pmatrix}' \mathrm{var}^{\mathcal{F}_0}\left[\begin{pmatrix} \zeta_k^{\mathrm{GM}} \\ \xi_k^{\mathrm{GM}} \end{pmatrix}\right] \begin{pmatrix} \sum_{i=n-k+1}^{n} v_i \\ \sum_{i=n-k+1}^{n} Z_{i,0} \end{pmatrix} + \sum_{k=1}^{n} \Bigg(\sum_{i=n-k+1}^{n} w_i \Bigg) \sigma_k^2 \,.$$

The proof of Theorem 6.4 is analogous to that of Theorem 4.4 (using Lemma 6.3 instead of Lemma 4.3).

Finally, the unknown variance parameters may be estimated as follows:

6.5 Theorem (Estimation of variance parameters). *In the combined model with* $m \geq 2$ *and for every* $k \in \{0, \ldots, n\}$, *the random variable*

$$\widehat{\sigma}_k^2 := \frac{1}{m+n-k-1} \sum_{i=-m}^{n-k} \frac{1}{w_i} \Big(Z_{i,k} - (v_i \zeta_k^{\mathrm{GM}} + Z_{i,0} \xi_k^{\mathrm{GM}}) \Big)^2$$

is an $\mathcal{F}_0$*–conditionally unbiased estimator of* σ_k^2.

The proof of Theorem 6.5 is analogous to that of Theorem 4.5 (using Lemmas 6.2 and 6.3 instead of Lemmas 4.2 and 4.3).

In the case $m = 1$, the assertion of Theorem 6.5 remains valid for $k \in \{0, \ldots, n{-}1\}$, and in the case $m = 0$ it remains valid for $k \in \{0, \ldots, n{-}2\}$. Thus, the final remarks of Section 4 apply *mutatis mutandis* to the combined model as well.

7 Loss Prediction in the Mack Model

For the sake of comparison, the present section provides a brief discussion of the famous Mack model for the chain–ladder method. In a sense to be made precise below, the Mack model is related to linear models but it is not a linear model as such.

For $k \in \{0, \ldots, n\}$, we denote by $\mathcal{F}_k$ the σ–algebra generated by the family

$$\{S_{j,l}\}_{l\in\{0,\ldots,k\},\, j\in\{-m,\ldots,n-l\}}$$

of all observable cumulative losses up to development year k and, for $i \in \{-m, \ldots, n\}$ and $k \in \{0, \ldots, n\}$, we denote by $\mathcal{F}_{i,k}$ the σ–algebra generated by the family

$$\{S_{i,l}\}_{l\in\{0,\ldots,k\}}$$

of all cumulative losses of accident year i up to development year k; note that the definition of $\mathcal{F}_0$ is in accordance with that used in Sections 5 and 6. The Mack model is defined as follows:

Mack Model: *The accident years are independent* (*in the sense that the family of σ–algebras $\{\mathcal{F}_{i,n}\}_{i\in\{-m,\ldots,n\}}$ is independent*) *and, for every development year $k \in \{1, \ldots, n\}$, there exist unknown parameters $\varphi_k \in \mathbb{R}$ and $\sigma_k^2 \in (0, \infty)$ such that the cumulative losses satisfy*

$$\begin{aligned} E^{\mathcal{F}_{i,k-1}}[S_{i,k}] &= S_{i,k-1}\,\varphi_k \\ \mathrm{var}^{\mathcal{F}_{i,k-1}}[S_{i,k}] &= S_{i,k-1}\,\sigma_k^2 \end{aligned}$$

for all $i \in \{-m, \ldots, n\}$. Moreover, $S_{i,k} > 0$ holds for all $i \in \{-m, \ldots, n\}$ and $k \in \{0, \ldots, n-1\}$.

In the Mack model, the cumulative losses $S_{i,k}$ replace the incremental losses used in the models considered before, the cumulative losses $S_{i,k-1}$ replace the volume measures used in the extended additive model and the initial losses used in the extended Panning model, and they also replace the accident year parameters w_i used in each of these models. Since the first identity in the Mack model implies

$$E[S_{i,k}/S_{i,k-1}] = \varphi_k$$

the development year parameter φ_k is assumed to be independent of the accident year and the collection of these parameters forms a *development pattern*; see Schmidt and Zocher [2009].

The Mack model is due to Mack [1993] who assumed that $m = 0$.

Assume that the assumptions of the Mack model are fulfilled. Then we have, for every $k \in \{1, \ldots, n\}$,

$$E^{\mathcal{F}_{k-1}}\left[\begin{pmatrix} S_{-m,k} \\ \vdots \\ S_{n-k,k} \end{pmatrix}\right] = \begin{pmatrix} S_{-m,k-1} \\ \vdots \\ S_{n-k,k-1} \end{pmatrix} \varphi_k$$

and

$$E^{\mathcal{F}_{k-1}}[S_{n-k+1,k}] \;=\; S_{n-k+1,k-1}\,\varphi_k$$

as well as

$$\mathrm{var}^{\mathcal{F}_{k-1}}\left[\begin{pmatrix} S_{-m,k} \\ \vdots \\ S_{n-k,k} \end{pmatrix}\right] \;=\; \begin{pmatrix} S_{-m,k-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S_{n-k,k-1} \end{pmatrix} \sigma_k^2$$

and

$$\mathrm{var}^{\mathcal{F}_{k-1}}[S_{n-k+1,k}] \;=\; S_{n-k+1,k-1}\,\sigma_k^2$$

and we also have

$$\mathrm{cov}^{\mathcal{F}_{k-1}}[S_{i,k}, S_{j,k}] \;=\; 0$$

for all $i, j \in \{-m, \ldots, n-k+1\}$ such that $i \neq j$; see Schmidt and Schnaus [1996]. We thus obtain the following result:

7.1 Theorem. *For every development year $k \in \{1, \ldots, n\}$, the Mack model provides an $\mathcal{F}_{k-1}$–conditional linear model for the family $\{S_{i,k}\}_{i\in\{-m,\ldots,n-k+1\}}$.*

Because of Theorem 7.1, the Mack model may be called a *sequential linear model*.

Let us first consider Gauss–Markov estimation of the parameter in the conditional linear models provided by the Mack model:

7.2 Lemma (Gauss–Markov estimation of parameters). *In the Mack model and for every development year $k \in \{1, \ldots, n\}$, the $\mathcal{F}_{k-1}$–conditional Gauss–Markov estimator of the parameter φ_k satisfies*

$$\varphi_k^{\mathrm{GM}} \;=\; \frac{\sum_{i=-m}^{n-k} S_{i,k}}{\sum_{i=-m}^{n-k} S_{i,k-1}}$$

and

$$\mathrm{var}^{\mathcal{F}_{k-1}}[\varphi_k^{\mathrm{GM}}] \;=\; \frac{1}{\sum_{i=-m}^{n-k} S_{i,k-1}}\,\sigma_k^2 \;.$$

The linear models for the families $\{S_{i,k}\}_{i\in\{-m,\ldots,n-k+1\}}$ cannot be extended to the families $\{S_{i,k}\}_{i\in\{-m,\ldots,n\}}$ since the cumulative losses $S_{i,k-1}$ with $i \in \{n-k+2, \ldots, n\}$ are non–observable and hence cannot be part of the design matrix of a conditional linear model (in which the design matrix is assumed to be observable); therefore, Gauss–Markov prediction is possible only for the non–observable cumulative losses $S_{n-k+1,k}$ of the first non–observable calendar year $n+1$:

7.3 Lemma (Gauss–Markov prediction of cumulative losses). *In the Mack model and for every development year $k \in \{1, \ldots, n\}$, the $\mathcal{F}_{k-1}$–conditional Gauss–Markov predictor of the cumulative loss $S_{n-k+1,k}$ satisfies*

$$S^{\mathrm{GM}}_{n-k+1,k} \;=\; S_{n-k+1,k-1}\varphi^{\mathrm{GM}}.$$

At this point, let us recall that, for every $k \in \{1, \ldots, n\}$, the *chain–ladder factor* φ^{CL}_k is defined as

$$\varphi^{\mathrm{CL}}_k \;:=\; \frac{\sum_{i=-m}^{n-k} S_{i,k}}{\sum_{i=-m}^{n-k} S_{i,k-1}}$$

and that, for all $i, k \in \{0, \ldots, n\}$ such that $i + k \geq n$, the *chain–ladder predictor* of the cumulative loss $S_{i,k}$ (which is non–observable for $i+k \geq n+1$) is defined as

$$S^{\mathrm{CL}}_{i,k} \;:=\; S_{i,n-i} \prod_{l=n-i+1}^{k} \varphi^{\mathrm{CL}}_l$$

(such that $S^{\mathrm{CL}}_{i,n-i} = S_{i,n-i}$). Thus, Lemmas 7.2 and 7.3 assert that

$$\varphi^{\mathrm{GM}}_k \;=\; \varphi^{\mathrm{CL}}_k$$

and

$$S^{\mathrm{GM}}_{n-k+1,k} \;=\; S^{\mathrm{CL}}_{n-k+1,k}$$

holds for all $k \in \{1, \ldots, n\}$. Since Gauss–Markov predictors are unbiased, the previous identity yields

$$E[S^{\mathrm{CL}}_{i,k} - S_{i,k}] \;=\; 0$$

and hence

$$E[(S^{\mathrm{CL}}_{i,k} - S_{i,k})^2] \;=\; \mathrm{var}[S^{\mathrm{CL}}_{i,k} - S_{i,k}]$$

for all $i, k \in \{1, \ldots, n\}$ such that $i + k = n + 1$, and it can be shown that these identities are also true for all $i, k \in \{1, \ldots, n\}$ such that $i + k \geq n + 2$.

Following Mack [1993], however, one should consider the $\mathcal{F}_n$–conditional mean squared error of prediction

$$E^{\mathcal{F}_n}[(S^{\mathrm{CL}}_{i,k} - S_{i,k})^2] \;=\; \mathrm{var}^{\mathcal{F}_n}[S^{\mathrm{CL}}_{i,k} - S_{i,k}] + \Big(E^{\mathcal{F}_n}[S^{\mathrm{CL}}_{i,k} - S_{i,k}]\Big)^2$$

instead of the unconditional mean squared error of prediction $E[(S^{\mathrm{CL}}_{i,k} - S_{i,k})^2]$. Since

$$E^{\mathcal{F}_n}[S^{\mathrm{CL}}_{i,k}] \;=\; S_{i,n-i} \prod_{l=n-i+1}^{k} \varphi^{\mathrm{CL}}_l$$

$$E^{\mathcal{F}_n}[S_{i,k}] \;=\; S_{i,n-i} \prod_{l=n-i+1}^{k} \varphi_l$$

we have

$$E^{\mathcal{F}_n}[S_{i,k}^{\text{CL}} - S_{i,k}] \;=\; S_{i,n-i}\left(\prod_{l=n-i+1}^{k} \varphi_l^{\text{CL}} - \prod_{l=n-i+1}^{k} \varphi_l\right)$$

which shows that the chain–ladder predictors fail to be $\mathcal{F}_n$–conditionally unbiased. Thus, the bias does not vanish in the identity for the $\mathcal{F}_n$–conditional mean squared error of prediction, which is most unfortunate since obviously plug–in estimators cannot be used to estimate the bias. By contrast, Mack [1993] has shown that the $\mathcal{F}_n$–conditional variance of the prediction error satisfies

$$\text{var}^{\mathcal{F}_n}[S_{i,k}^{\text{CL}} - S_{i,k}] \;=\; S_{i,n-i} \sum_{l=n-i+1}^{k} \left(\prod_{h=n-i+1}^{l-1} \varphi_h\right) \sigma_l^2 \left(\prod_{h=l+1}^{k} \varphi_h^2\right)$$

(which provides the identity

$$\text{var}^{\mathcal{F}_n}[S_{i,k}] \;=\; S_{i,n-i} \sum_{l=n-i+1}^{k} \left(\prod_{h=n-i+1}^{l-1} \varphi_h\right) \sigma_l^2 \left(\prod_{h=l+1}^{k} \varphi_h^2\right)$$

needed in Theorem 7.4 below). In conclusion, estimation of the bias causes a serious difficulty in the estimation of the $\mathcal{F}_n$–conditional mean squared error of prediction of the chain–ladder predictor of a non–observable cumulative loss.

These observations also apply to the chain–ladder predictors of non–observable incremental losses which are defined as

$$Z_{i,k}^{\text{CL}} \;:=\; S_{i,k}^{\text{CL}} - S_{i,k-1}^{\text{CL}}$$

and, in particular, to the chain–ladder predictors of reserves which are defined as

$$\begin{aligned} R_i^{\text{CL}} &:= \sum_{k=n-i+1}^{n} Z_{i,k}^{\text{CL}} \\ R_{(c)}^{\text{CL}} &:= \sum_{i=c-n}^{n} Z_{i,c-i}^{\text{CL}} \\ R^{\text{CL}} &:= \sum_{k=1}^{n} \sum_{i=n-k+1}^{n} Z_{i,k}^{\text{CL}} \;. \end{aligned}$$

This can be seen from the following result:

7.4 Theorem (Chain–ladder prediction of reserves). *In the Mack model,*
(1) *the chain–ladder predictors of the accident year reserves satisfy*

$$E^{\mathcal{F}_n}[R_i^{\text{CL}} - R_i] \;=\; S_{i,n-i}\left(\prod_{k=n-i+1}^{n} \varphi_k^{\text{CL}} - \prod_{k=n-i+1}^{n} \varphi_k\right)$$

and

$$E^{\mathcal{F}_n}[(R_i^{\mathrm{CL}} - R_i)^2] \;=\; \left(S_{i,n-i}\left(\prod_{k=n-i+1}^{n} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n} \varphi_k\right)\right)^2 + \mathrm{var}^{\mathcal{F}_n}[S_{i,n}]$$

as well as

$$\mathrm{cov}^{\mathcal{F}_n}[R_i^{\mathrm{CL}} - R_i, R_j^{\mathrm{CL}} - R_j] \;=\; \mathrm{var}^{\mathcal{F}_n}[S_{i,n}]\,\delta_{i,j}$$

(2) *the chain–ladder predictors of the calendar year reserves satisfy*

$$\begin{aligned} &E^{\mathcal{F}_n}[R_{(c)}^{\mathrm{CL}} - R_{(c)}] \\ &= \sum_{i=c-n}^{n} S_{i,n-i}\left(\left(\prod_{k=n-i+1}^{c-i-1} \varphi_k^{\mathrm{CL}}\right)(\varphi_{c-i}^{\mathrm{CL}} - 1) - \left(\prod_{k=n-i+1}^{c-i-1} \varphi_k\right)(\varphi_{c-i} - 1)\right) \end{aligned}$$

and

$$\begin{aligned} &E^{\mathcal{F}_n}[(R_{(c)}^{\mathrm{CL}} - R_{(c)})^2] \\ &= \left(\sum_{i=c-n}^{n} S_{i,n-i}\left(\left(\prod_{k=n-i+1}^{c-i-1} \varphi_k^{\mathrm{CL}}\right)(\varphi_{c-i}^{\mathrm{CL}} - 1) - \left(\prod_{k=n-i+1}^{c-i-1} \varphi_k\right)(\varphi_{c-i} - 1)\right)\right)^2 \\ &\quad + \sum_{i=c-n}^{n}\left(\mathrm{var}^{\mathcal{F}_n}[S_{i,c-i-1}]\,(\varphi_{c-i} - 1)^2 + S_{i,n-i}\left(\prod_{k=n-i+1}^{c-i-1} \varphi_k\right)\sigma_{c-i}^2\right) \end{aligned}$$

(3) *the chain–ladder predictor of the total reserve satisfies*

$$E^{\mathcal{F}_n}[R^{\mathrm{CL}} - R] \;=\; \sum_{i=1}^{n} S_{i,n-i}\left(\prod_{k=n-i+1}^{n} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n} \varphi_k\right)$$

and

$$E^{\mathcal{F}_n}[(R^{\mathrm{CL}} - R)^2] \;=\; \left(\sum_{i=1}^{n} S_{i,n-i}\left(\prod_{k=n-i+1}^{n} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n} \varphi_k\right)\right)^2 + \sum_{i=1}^{n} \mathrm{var}^{\mathcal{F}_n}[S_{i,n}]$$

A proof of Theorem 7.4 will be given in the Appendix.

Theorem 7.4 provides explicit formulas for the $\mathcal{F}_n$–conditional mean squared errors of prediction, but the use of plug–in estimators in these formulas is not recommendable since it would result in wiping out a part of the $\mathcal{F}_n$–conditional mean squared errors of prediction.

7.5 Theorem (Estimation of variance parameters). *In the Mack model with* $m \geq 1$ *and for every* $k \in \{1, \ldots, n\}$, *the random variable*

$$\widehat{\sigma}_k^2 \;:=\; \frac{1}{m+n-k} \sum_{i=-m}^{n-k} \frac{1}{S_{i,k-1}}\,(S_{i,k} - S_{i,k-1}\varphi_k^{\mathrm{CL}})^2$$

is an $\mathcal{F}_{k-1}$–conditionally unbiased estimator of σ_k^2.

As noted before, the use of plug–in estimators for the parameters of the development pattern in the formulas provided by Theorem 7.4 is not recommendable. Mack [1993] proposed the estimators

$$\begin{aligned}
&\widehat{E^{\mathcal{F}_n}}[(R_i^{\mathrm{CL}}-R_i)^2] \\
&:= (S_{i,n}^{\mathrm{CL}})^2 \sum_{k=n-i+1}^{n} \left(\frac{1}{\sum_{h=-m}^{n-k} S_{h,k}} + \frac{1}{S_{i,k}^{\mathrm{CL}}} \right) \frac{\widehat{\sigma}_k^2}{\varphi_k^{\mathrm{CL}}} \\
&= (S_{i,n}^{\mathrm{CL}})^2 \sum_{k=n-i+1}^{n} \frac{1}{\sum_{h=-m}^{n-k} S_{h,k}} \frac{\widehat{\sigma}_k^2}{\varphi_k^{\mathrm{CL}}} + (S_{i,n}^{\mathrm{CL}})^2 \sum_{k=n-i+1}^{n} \frac{1}{S_{i,k}^{\mathrm{CL}}} \frac{\widehat{\sigma}_k^2}{\varphi_k^{\mathrm{CL}}}
\end{aligned}$$

for the $\mathcal{F}_n$–conditional mean squared errors of prediction of the accident year reserves and

$$\begin{aligned}
&\widehat{E^{\mathcal{F}_n}}[(R^{\mathrm{CL}}-R)^2] \\
&:= \sum_{i=1}^{n} \sum_{j=1}^{n} S_{i,n}^{\mathrm{CL}} S_{j,n}^{\mathrm{CL}} \sum_{k=n-i\wedge j+1}^{n} \frac{1}{\sum_{h=-m}^{n-k} S_{h,k}} \frac{\widehat{\sigma}_k^2}{\varphi_k^{\mathrm{CL}}} + \sum_{i=1}^{n} (S_{i,n}^{\mathrm{CL}})^2 \sum_{k=n-i+1}^{n} \frac{1}{S_{i,k}^{\mathrm{CL}}} \frac{\widehat{\sigma}_k^2}{\varphi_k^{\mathrm{CL}}}
\end{aligned}$$

for the $\mathcal{F}_n$–conditional mean squared error of prediction of the total reserve. The construction of each of these estimators involves certain approximations.

Apparently, no estimators have been proposed in the literature for the conditional mean squared errors of prediction of the calendar year reserves.

8 Remarks

In the Panning model and in the combined model, it would be sufficient to assume $Z_{i,0} > 0$ only for $i \in \{1, \ldots, n\}$, but then the formulas for predictors and mean squared errors of prediction have would have to be modified as to avoid divisions by zero.

The accident year parameters w_i may e. g.,be chosen as follows:
- In the extended additive model, one may choose $w_i := 1$ (corresponding to the traditional Panning model) or $w_i := v_i$ (traditional additive model) or, more generally, $w_i := \alpha + \beta v_i$ with $\alpha, \beta \in [0, 1]$ and $\alpha + \beta = 1$.
- In the extended Panning model, one may choose $w_i := 1$ (traditional Panning model) or $w_i := v_i$ (corresponding to the traditional additive model) or $w_i := Z_{i,0}$ (in analogy with the traditional additive model) or, more generally, $w_i := \alpha + \beta v_i + \gamma Z_{i,0}$ with $\alpha, \beta, \gamma \in [0, 1]$ and $\alpha + \beta + \gamma = 1$.
- In the combined model, one may choose $w_i := 1$ (corresponding to the traditional Panning model) or $w_i := v_i$ (corresponding to the traditional additive model) or $w_i := Z_{i,0}$ or, more generally, $w_i := \alpha + \beta v_i + \gamma Z_{i,0}$ with $\alpha, \beta, \gamma \in [0, 1]$ and $\alpha + \beta + \gamma = 1$.

The combined model uses volume measures and initial losses as regressors and thus provides an example for a broad class of general linear models combining different sources of information on the accident years. As there are several possible choices for the volume measure, like the number of contracts, the premium income, market statistics or even information on a similar portfolio of risks, one might want to use some of them simultaneously; also, as for example in excess–of–loss reinsurance, one might want to use several volume measures but avoid initial losses. In both cases, it is straightforward to construct appropriate modifications of the combined model and the analysis of the resulting models would follow the lines of Section 6.

For the additive method and the Panning method, the principle of Gauss–Markov prediction in an appropriate linear model shows that, under certain assumptions on the first and second order moments of the incremental losses,
- the predictors used in these methods are unbiased and minimize the mean squared error of prediction, and
- the mean squared errors of prediction can be estimated by the simple use of plug–in estimators for the unknown variance parameters.

In addition, the systematic use of Gauss–Markov prediction in a linear model leads to variations and combinations of these methods; see Section 9 below for nine such methods using the available information in a slightly different way. The analysis of results from different but similar methods may be useful to study the sensitivity of result with respect to model variations and to analyze the impact of loss development data and volume measures; see also Schmidt and Zocher [2009] for a similar discussion of another family of models and methods.

Unfortunately, the situation is not that comfortable for the chain–ladder method. While the Mack model was certainly a breakthrough in stochastic modelling for the chain–ladder method and provides a partial justification of that method, it seems that in this model

- the question of whether or not the chain–ladder predictors minimize the mean squared error of prediction cannot be settled and that
- the construction of estimators of the mean squared errors of prediction presents a serious problem and seems to require certain delicate approximations.

This is due to the sequential character of the Mack model, which provides a linear model for every development year but not for the entire loss development.

9 A Numerical Example

In the present section we illustrate the results of this paper by a numerical example. In the example, we consider a portfolio of auto liability and use the incremental losses provided by Braun [2004], truncated at development year 9, and the volume measures proposed by Merz and Wüthrich [2009]. These data are presented in Table 1.

For each of the additive model, the Panning model and the combined model we consider the three cases in which the accident year parameters of the variances are chosen as $w_i = 1$, $w_i = v_i$, and $w_i = Z_{i,0}$, respectively, and we also consider the Mack model in which the corresponding parameters are the cumulative losses $S_{i,k-1}$. The Gauss–Markov estimators of the parameters ζ_k (additive model and combined model), ξ_k (Panning model and combined model), and φ_k (Mack model) are displayed in Tables 2–5.

In the combined model, the signs of the Gauss–Markov estimators given Table 4 show that the volume measures and the initial losses have an opposite effect on the Gauss–Markov predictors of reserves; see Theorem 6.4. The Gauss–Markov predictors of the reserves of accident years 1–9, the total reserves and the reserves of calendar years 10–18 are displayed in Table 6.

The *standard error of prediction* is defined as the square root of the mean squared error of prediction and measures uncertainty in the monetary unit. The estimated standard errors of prediction are displayed in Table 7.

As an alternative measure of uncertainty, one could also consider the *coefficient of variation* which is defined as the ratio between the standard error of prediction and the predictor and is dimension–free. The coefficients of variation are displayed in Table 8.

Of course, the choice of a stochastic model should not be driven by the numerical results which it produces. Nevertheless, model selection should perhaps proceed in steps, starting with the choice of a plausible *class* of models (like the class of general linear models) and subsequently shrinking this class to only a few models or even a single one. In this process a comparative analysis of a family of similar models could help to obtain some insight into some of the characteristics of these models.

For the example considered here, we make the following observations:

- The choice of regressors (volume measures in the additive model, initial losses in the Panning model, and both of them in the combined model) may affect the predictors and the standard errors of prediction. For example, for the Panning model, the predictors of the total reserves are smaller and the standard errors of calendar year 10 are larger than for the additive model and the combined model.

- The choice of the accident year parameters may affect the predictors and the standard errors of prediction. For example, for $w_i = 1$, the total reserves are larger and the standard errors are smaller than for $w_i = v_i$ and $w_i = Z_{i,0}$.
- For the Mack model, the predictors are in the range of those obtained for the other models but the standard errors are larger.

Such considerations combined with actuarial judgement could help to determine estimates of reserves and estimates of standard errors of prediction for the portfolio under consideration.

Nevertheless, such an analysis for a particular portfolio cannot justify a general preference for a particular stochastic model.

Accident Year	Volume Measure	Development Year									
		0	1	2	3	4	5	6	7	8	9
-4	413213	114423	133538	65021	31358	27139	-377	9889	4477	-316	7108
-3	537988	152296	152879	71438	41686	22009	25315	7961	4843	-113	1593
-2	589145	144325	162919	106365	50432	55224	7951	8234	1409	2061	669
-1	523419	145904	161732	79458	46642	29384	15811	3598	5527	-2484	462
0	501498	170333	171168	92601	36227	11872	18760	3180	3538	948	-875
1	598345	189643	171480	85734	61226	18479	13556	7523	1964	88	
2	608376	179022	217202	101080	56183	28362	29791	11244	12568		
3	698993	205908	210139	104397	45277	34888	30193	17563			
4	704129	210951	215478	98618	62846	52435	22824				
5	903557	213426	295796	140211	82259	59209					
6	947326	249508	330502	142126	122023						
7	1134129	258425	427587								
8	1538916	368762	540304								
9	1487234	394997									

Table 1: Incremental Losses

Gauss–Markov Estimators	Development Year									
	0	1	2	3	4	5	6	7	8	9
$w_i = 1$	0.2605	0.3368	0.1642	0.0934	0.0570	0.0326	0.0158	0.0091	0.0001	0.0030
$w_i = v_i$	0.2680	0.3290	0.1613	0.0905	0.0558	0.0317	0.0155	0.0091	0.0001	0.0035
$w_i = Z_{i,0}$	0.2648	0.3307	0.1626	0.0911	0.0573	0.0311	0.0156	0.0090	0.0001	0.0036

Table 2: Gauss–Markov Estimators in the Additive Model

Gauss–Markov Estimators	Development Year								
	1	2	3	4	5	6	7	8	9
$w_i = 1$	1.2747	0.6003	0.3308	0.1955	0.1121	0.0535	0.0313	0.0004	0.0100
$w_i = v_i$	1.2021	0.5769	0.3167	0.1890	0.1091	0.0522	0.0312	0.0002	0.0116
$w_i = Z_{i,0}$	1.2258	0.5891	0.3220	0.1964	0.1083	0.0531	0.0313	0.0002	0.0123

Table 3: Gauss–Markov Estimators in the Panning Model

Gauss–Markov Estimators	Development Year 1	2	3	4	5	6	7	8	9
$w_i = 1$	0.4795	0.2686	0.1731	0.1731	-0.0300	0.0305	0.0033	0.0024	0.0148
	-0.5505	-0.2023	-0.4139	-0.4139	0.2140	-0.0504	0.0199	-0.0077	-0.0419
$w_i = v_i$	0.4444	0.2403	0.1421	0.1896	-0.0340	0.0335	0.0047	0.0011	0.0177
	-0.4302	-0.2886	-0.1832	-0.4714	0.2246	-0.0618	0.0150	-0.0035	-0.0502
$w_i = Z_{i,0}$	0.4545	0.2542	0.1393	0.1861	-0.0414	0.0292	0.0003	0.0032	0.0146
	-0.4679	-0.3392	-0.1735	-0.4589	0.2499	-0.0471	0.0303	-0.0108	-0.0392

Table 4: Gauss–Markov Estimators in the Combined Model

Gauss–Markov Estimators	Development Year 1	2	3	4	5	6	7	8	9
	2.2258	1.2694	1.1204	1.0668	1.0354	1.0168	1.0097	1.0001	1.0037

Table 5: Gauss–Markov Estimators in the Mack Model

Year	Additive Model			Panning Model			Combined Model			Mack Model
	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	
1	1792	2089	2165	1890	2195	2336	919	1086	1304	2054
2	1912	2160	2258	1859	2100	2241	1581	1822	1874	2415
3	8567	8842	8896	8756	8833	9026	8232	8485	8588	8762
4	19763	19804	19937	20073	20068	20459	19024	18942	19200	20232
5	54806	54017	53717	44226	43588	43812	47548	47272	44396	52994
6	111440	109465	110578	100490	98103	100217	114045	114790	113047	116698
7	239298	233738	235656	189570	183455	187008	265053	265185	259631	251872
8	577322	565374	569989	491876	474513	484091	619938	613617	610210	562574
9	1058893	1035648	1042712	1030382	983097	1002726	1061093	1052418	1050462	1028283
Total	2073790	2031136	2045907	1888941	1815952	1851916	2137432	2123616	2108712	2045884
10	962268	940978	947253	902762	859493	876786	979515	965487	966517	943140
11	505930	495009	499106	457482	440535	449395	539568	536922	534841	498805
12	288908	281751	284390	253895	245074	250111	302808	302239	298209	285564
13	163703	160341	161950	142379	138618	141695	158496	158669	155306	163089
14	85982	84427	83876	74307	72919	73147	81916	81565	78020	85532
15	40543	40394	40590	35258	35053	35668	42187	42770	41316	40861
16	17173	17583	17656	15079	15360	15628	19610	20461	19767	18026
17	4829	5460	5706	3841	4330	4621	7846	8973	8498	5568
18	4454	5193	5380	3938	4570	4865	5486	6530	6239	5300

Table 6: Gauss–Markov Predictors of Reserves

Year	Additive Model			Panning Model			Combined Model			Mack Model
	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	
1	3672	4260	4458	3823	4428	4619	4627	5255	5661	5507
2	4046	4645	4730	4125	4738	4821	4598	5241	5439	6523
3	5816	6616	6722	5879	6702	6782	6504	7381	7575	7940
4	7213	8122	8252	7456	8420	8505	8012	8963	9244	9688
5	12257	15329	14299	11468	14555	13423	18370	20789	19572	17168
6	18424	22991	22327	20582	25468	24802	18160	21689	20921	28047
7	24595	30909	28394	27585	34861	31879	33790	38016	35710	38512
8	33753	44489	42401	43931	57054	54984	41550	50472	47660	55852
9	43298	56745	56753	72712	90441	91254	40463	54210	53349	129110
Total	86154	101944	100194	109448	130181	129282	113638	126545	121534	176968
10	41519	52118	51402	71084	86744	86557	41168	52158	50498	
11	31861	39778	38650	43229	53290	52786	33925	41131	39429	
12	25884	34347	32733	29612	39777	38020	30784	36890	35322	
13	20602	28982	27921	22623	31995	31032	26262	31509	30190	
14	13984	19671	19057	12946	18666	17902	20091	24483	23357	
15	8860	11802	11264	8567	11801	11172	13289	15601	14868	
16	7334	9780	9340	6859	9445	8940	11274	13729	13186	
17	5899	8354	7987	5439	8027	7642	9457	12073	11676	
18	5318	7602	7437	5183	7540	7375	6467	8997	8806	

Table 7: Standard Errors of Prediction

Year	Additive Model			Panning Model			Combined Model			Mack Model
	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	$w_i = 1$	$w_i = v_i$	$w_i = Z_{i,0}$	
1	204.97%	203.94%	205.95%	202.25%	201.76%	197.77%	503.36%	484.10%	434.04%	268.09%
2	211.56%	215.06%	209.53%	221.84%	225.59%	215.17%	290.77%	287.65%	290.26%	270.14%
3	67.91%	74.82%	75.56%	68.55%	75.88%	75.13%	79.00%	86.99%	88.20%	90.62%
4	36.50%	41.01%	41.39%	37.15%	41.96%	41.57%	42.11%	47.32%	48.14%	47.89%
5	22.36%	28.38%	26.62%	25.93%	33.39%	30.64%	38.63%	43.98%	44.09%	32.40%
6	16.53%	21.00%	20.19%	20.48%	25.96%	24.75%	15.92%	18.89%	18.51%	24.03%
7	10.28%	13.22%	12.05%	14.55%	19.00%	17.05%	12.75%	14.34%	13.75%	15.29%
8	5.85%	7.87%	7.44%	8.93%	12.02%	11.36%	6.70%	8.23%	7.81%	9.93%
9	4.09%	5.48%	5.44%	7.06%	9.20%	9.10%	3.81%	5.15%	5.08%	12.56%
Total	4.15%	5.02%	4.90%	5.79%	7.17%	6.98%	5.32%	5.96%	5.76%	8.65%
10	4.31%	5.54%	5.43%	7.87%	10.09%	9.87%	4.20%	5.40%	5.22%	
11	6.30%	8.04%	7.74%	9.45%	12.10%	11.75%	6.29%	7.66%	7.37%	
12	8.96%	12.19%	11.51%	11.66%	16.23%	15.20%	10.17%	12.21%	11.84%	
13	12.59%	18.07%	17.24%	15.89%	23.08%	21.90%	16.57%	19.86%	19.44%	
14	16.26%	23.30%	22.72%	17.42%	25.60%	24.47%	24.53%	30.02%	29.94%	
15	21.85%	29.22%	27.75%	24.30%	33.67%	31.32%	31.50%	36.48%	35.99%	
16	42.70%	55.62%	52.90%	45.49%	61.49%	57.20%	57.49%	67.10%	66.71%	
17	122.14%	153.00%	139.97%	141.61%	185.37%	165.39%	120.52%	134.54%	137.40%	
18	119.40%	146.41%	138.23%	131.62%	164.95%	151.61%	117.87%	137.78%	141.15%	

Table 8: Coefficients of Variation

Appendix

Here we present a proof of Theorem 7.4:

Proof. We have $R_i^{\mathrm{CL}} - R_i = S_{i,n}^{\mathrm{CL}} - S_{i,n}$ and hence

$$\begin{aligned} E^{\mathcal{F}_n}[R_i^{\mathrm{CL}} - R_i] &= E^{\mathcal{F}_n}[S_{i,n}^{\mathrm{CL}} - S_{i,n}] \\ &= S_{i,n-i}\left(\prod_{k=n-i+1}^{n} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n} \varphi_k\right). \end{aligned}$$

Since the accident years are independent, we also have

$$\begin{aligned} \mathrm{cov}^{\mathcal{F}_n}[R_i^{\mathrm{CL}} - R_i, R_j^{\mathrm{CL}} - R_j] &= \mathrm{cov}^{\mathcal{F}_n}[S_{i,n}^{\mathrm{CL}} - S_{i,n}, S_{j,n}^{\mathrm{CL}} - S_{j,n}] \\ &= \mathrm{cov}^{\mathcal{F}_n}[S_{i,n}, S_{j,n}] \\ &= \mathrm{var}^{\mathcal{F}_n}[S_{i,n}]\, \delta_{i,j} \end{aligned}$$

In particular, we have

$$\begin{aligned} E^{\mathcal{F}_n}[(R_i^{\mathrm{CL}} - R_i)^2] &= \mathrm{var}^{\mathcal{F}_n}[R_i^{\mathrm{CL}} - R_i] + \left(E^{\mathcal{F}_n}[R_i^{\mathrm{CL}} - R_i]\right)^2 \\ &= \mathrm{var}^{\mathcal{F}_n}[S_{i,n}] + S_{i,n-i}^2\left(\prod_{k=n-i+1}^{n} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n} \varphi_k\right)^2. \end{aligned}$$

This proves (1).
We have

$$\begin{aligned} R_{(c)}^{\mathrm{CL}} - R_{(c)} &= \sum_{i=c-n}^{n} ((S_{i,c-i}^{\mathrm{CL}} - S_{i,c-i-1}^{\mathrm{CL}}) - (S_{i,c-i} - S_{i,c-i-1})) \\ &= \sum_{i=c-n}^{n} (S_{i,c-i}^{\mathrm{CL}} - S_{i,c-i}) - \sum_{i=c-n}^{n} (S_{i,c-i-1}^{\mathrm{CL}} - S_{i,c-i-1}) \end{aligned}$$

and hence

$$\begin{aligned} &E^{\mathcal{F}_n}[R_{(c)}^{\mathrm{CL}} - R_{(c)}] \\ &= \sum_{i=c-n}^{n} E^{\mathcal{F}_n}[(S_{i,c-i}^{\mathrm{CL}} - S_{i,c-i})] - \sum_{i=c-n}^{n} E^{\mathcal{F}_n}[(S_{i,c-i-1}^{\mathrm{CL}} - S_{i,c-i-1})] \\ &= \sum_{i=c-n}^{n} S_{i,n-i}\left(\prod_{k=n-i+1}^{c-i} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{c-i} \varphi_k\right) - \sum_{i=c-n}^{n} S_{i,n-i-1}\left(\prod_{k=n-i+1}^{c-i-1} \varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{c-i-1} \varphi_k\right) \\ &= \sum_{i=c-n}^{n} S_{i,n-i}\left(\left(\prod_{k=n-i+1}^{c-i-1} \varphi_k^{\mathrm{CL}}\right)(\varphi_{c-i}^{\mathrm{CL}} - 1) - \left(\prod_{k=n-i+1}^{c-i-1} \varphi_k\right)(\varphi_{c-i} - 1)\right). \end{aligned}$$

Since the accident years are independent, we also have

$$\mathrm{var}^{\mathcal{F}_n}[R_{(c)}^{\mathrm{CL}} - R_{(c)}] = \mathrm{var}^{\mathcal{F}_n}\left[\sum_{i=c-n}^{n}\left(\left(S_{i,c-i}^{\mathrm{CL}} - S_{i,c-i-1}^{\mathrm{CL}}\right) - \left(S_{i,c-i} - S_{i,c-i-1}\right)\right)\right]$$

$$\begin{aligned} &= \operatorname{var}^{\mathcal{F}_n}\left[\sum_{i=c-n}^{n}\Bigl(S_{i,c-i}-S_{i,c-i-1}\Bigr)\right] \\ &= \sum_{i=c-n}^{n}\operatorname{var}^{\mathcal{F}_n}[S_{i,c-i}-S_{i,c-i-1}] \end{aligned}$$

as well as

$$\operatorname{var}^{\mathcal{F}_n}[S_{i,c-i}-S_{i,c-i-1}] = \operatorname{var}^{\mathcal{F}_n}[S_{i,c-i-1}]\,(\varphi_{c-i}-1)^2 + S_{i,n-i}\left(\prod_{k=n-i+1}^{c-i-1}\varphi_k\right)\sigma_{c-i}^2$$

and hence

$$\operatorname{var}^{\mathcal{F}_n}[R_{(c)}^{\mathrm{CL}}-R_{(c)}] = \sum_{i=c-n}^{n}\left(\operatorname{var}^{\mathcal{F}_n}[S_{i,c-i-1}]\,(\varphi_{c-i}-1)^2 + S_{i,n-i}\left(\prod_{k=n-i+1}^{c-i-1}\varphi_k\right)\sigma_{c-i}^2\right).$$

In particular, we have

$$\begin{aligned} &E^{\mathcal{F}_n}[(R_{(c)}^{\mathrm{CL}}-R_{(c)})^2] \\ &= \operatorname{var}^{\mathcal{F}_n}[R_{(c)}^{\mathrm{CL}}-R_{(c)}] + \Bigl(E^{\mathcal{F}_n}[R_{(c)}^{\mathrm{CL}}-R_{(c)}]\Bigr)^2 \\ &= \sum_{i=c-n}^{n}\left(\operatorname{var}^{\mathcal{F}_n}[S_{i,c-i-1}]\,(\varphi_{c-i}-1)^2 + S_{i,n-i}\left(\prod_{k=n-i+1}^{c-i-1}\varphi_k\right)\sigma_{c-i}^2\right) \\ &\quad + \left(\sum_{i=c-n}^{n}S_{i,n-i}\left(\left(\prod_{k=n-i+1}^{c-i-1}\varphi_k^{\mathrm{CL}}\right)(\varphi_{c-i}^{\mathrm{CL}}-1) - \left(\prod_{k=n-i+1}^{c-i-1}\varphi_k\right)(\varphi_{c-i}-1)\right)\right)^2. \end{aligned}$$

This proves (2).
We have $R^{\mathrm{CL}} - R = \sum_{i=1}^{n}(R_i^{\mathrm{CL}}-R_i)$ and hence

$$\begin{aligned} E^{\mathcal{F}_n}[R^{\mathrm{CL}}-R] &= \sum_{i=1}^{n}E^{\mathcal{F}_n}[R_i^{\mathrm{CL}}-R_i] \\ &= \sum_{i=1}^{n}S_{i,n-i}\left(\prod_{k=n-i+1}^{n}\varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n}\varphi_k\right) \end{aligned}$$

From (1) we obtain

$$\begin{aligned} \operatorname{var}^{\mathcal{F}_n}[R^{\mathrm{CL}}-R] &= \sum_{i=1}^{n}\operatorname{var}^{\mathcal{F}_n}[R_i^{\mathrm{CL}}-R_i] \\ &= \sum_{i=1}^{n}\operatorname{var}^{\mathcal{F}_n}[S_{i,n}]\,. \end{aligned}$$

In particular, we have

$$\begin{aligned} E^{\mathcal{F}_n}[(R^{\mathrm{CL}}-R)^2] &= \operatorname{var}^{\mathcal{F}_n}[R^{\mathrm{CL}}-R] + \Bigl(E^{\mathcal{F}_n}[R^{\mathrm{CL}}-R]\Bigr)^2 \\ &= \sum_{i=1}^{n}\operatorname{var}^{\mathcal{F}_n}[S_{i,n}] + \left(\sum_{i=1}^{n}S_{i,n-i}\left(\prod_{k=n-i+1}^{n}\varphi_k^{\mathrm{CL}} - \prod_{k=n-i+1}^{n}\varphi_k\right)\right)^2. \end{aligned}$$

This proves (3). □

References

Bornhuetter, R.L., and R.E. Ferguson, "The Actuary and IBNR," *Proceedings of the Casualty Actuarial Society* 59, 1972, pp. 181–195.

Braun, C., "The prediction error of the chain–ladder method applied to correlated run–off triangles," ASTIN Bulletin 34, 2004, 399–423.

Hachemeister, C.A., and J.N. Stanard, "IBNR claims count estimation with static lag functions," 1975, unpublished.

Halliwell, L.J., "Loss prediction by generalized least squares," *Proceedings of the Casualty Actuarial Society* 83, 1996, pp. 436–489.

Halliwell, L.J., "Conjoint prediction of paid and incurred losses," Casualty Actuarial Society *Forum*, Summer 1997, pp. 241–379.

Hamer, M.D., "Loss prediction by generalized least squares – Discussion of Halliwell (1996)," *Proceedings of the Casualty Actuarial Society* 86, 1999, pp. 748–763.

Hess, K.T., and K.D. Schmidt, "A comparison of models for the chain–ladder method," *Insurance Mathematics and Economics* 31, 2002, pp. 351–364.

Hess, K.T., K.D. Schmidt, and M. Zocher, "Multivariate loss prediction in the multivariate additive model," *Insurance Mathematics and Economics* 39, 2006, pp. 185–191.

Kloberdanz, K., and K.D. Schmidt, "Loss prediction in a linear model under a linear constraint," *AStA Advances in Statistical Analysis* 93, 2009, pp. 205–220.

Kremer, E., "The correlated chain–ladder method for reserving in case of correlated claims developments," *Blätter DGVFM* 27, 2005, pp. 315–322.

Ludwig, A., C. Schmeisser, and K. Thänert, *Linear Models in Loss Reserving*, Research Report, Technische Universität Dresden, May 2009.

Mack, T., "A simple parametric model for rating automobile insurance or estimating IBNR claims reserves," *ASTIN Bulletin* 21, 1991, pp. 93–103.

Mack, T., "Distribution–free alculation of the standard error of chain–ladder reserve estimates," *ASTIN Bulletin* 23, 1993, pp. 213–225.

Merz, M., and M.V. Wüthrich, "Prediction error of the multivariate additive loss reserving method for dependent lines of business," *Variance* 3, 2009, pp. 131–151.

Panning, W., "Measuring loss reserve uncertainties," Casualty Actuarial Society *Forum*, Fall 2006, pp. 237–267.

Pröhl, C., and K.D. Schmidt, "Multivariate chain–ladder," *Dresdner Schriften zur Versicherungsmathematik* 3/2005.

Quarg, G., and T. Mack, Munich chain–ladder," *Blätter DGVFM* 26, 2004, 597–630.

Quarg, G., and T. Mack, "Munich chain–ladder: A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses," *Variance* 2, 2008, pp. 266–299.

Radtke, M., and K.D. Schmidt K.D. (eds.), *Handbuch zur Schadenreservierung*, Karlsruhe: Verlag Versicherungswirtschaft, 2004.

Schmidt, K.D., "Modèles et Méthodes de Réservation," Lectures held at the University of Strasbourg, May 2003.

Schmidt, K.D., "Prediction," *Encyclopedia of Actuarial Science*, vol. 3, 2004, pp. 1317–1321. Chichester: Wiley.

Schmidt, K.D., "Methods and models of loss reserving based on run–off triangles: A unifying survey," Casualty Actuarial Society *Forum*, Fall 2006a, pp. 269-317.

Schmidt, K.D., "Optimal and additive loss reserving for dependent lines of business," Casualty Actuarial Society *Forum*, Fall 2006b, pp. 319-351.

Schmidt, K.D., and A. Schnaus, "An extension of Mack's model for the chain–ladder method," *ASTIN Bulletin* 26, 1996, pp. 247–262.

Schmidt, K.D., and A. Wünsche, "Chain–ladder, marginal–sum and maximum–likelihood estimation," *Blätter DGVM* 23, 1998, pp. 267–277.

Schmidt, K.D., and M. Zocher, "Loss Reserving and Hofmann distributions," *Mitteilungen SAV*, 2005, pp. 127–162.

Schmidt, K.D., and M. Zocher, "The Bornhuetter–Ferguson principle," *Variance* 2, 2008, pp. 85–110.

Tarbell, T. F., "Incurred but not reported claims reserves," *Proceedings of the Casualty Actuarial Society* 20, 1934, pp. 275–280.