A GENERALIZED LOSS RATIO METHOD DEALING WITH UNCERTAIN VOLUME MEASURES

BY

ULRICH RIEGEL

ABSTRACT

Unlike chain ladder, the loss ratio method requires volume measures. Typically, these volumes are assumed to be known. In practice, however, accurate volume measures are rarely available. We interpret the available volumes as estimators for the true volume measures and analyze the consequences for the loss ratio method. In particular, we calculate the mean squared error of prediction, including uncertainty of volume measures, and derive approximately optimal weights for the observed incremental loss ratios. We then introduce a generalization of the loss ratio method that is tailored to the situation of uncertain volume measures and calculate the prediction uncertainty of this generalized loss ratio method.

KEYWORDS

Loss ratio method, additive model, volume measures, IBNR method, mean squared error.

1. INTRODUCTION

The loss ratio method (or additive method) is one of the most popular reserving methods worldwide. According to the ASTIN report on non-life reserving (see ASTIN Working Party on Non-Life Reserving ASTIN Working Party on Non-Life Reserving (2016)) only chain ladder and Bornhuetter–Ferguson are more widely used.

The loss ratio method is based on the additive model that postulates the existence of *model volumes* $v_1, \dots, v_n$ and expected loss ratio increments $m_1, \dots, m_n$ such that

$$\mathrm{E}(S_{i,k}) = v_i \cdot m_k$$

holds for all incremental losses $S_{i,k}$ with $i, k \in \{1, \dots, n\}$. In literature, the model volumes v_i are usually assumed to be known (cf. Mack (2002), Hess *et al.* (2006) or Merz and Wüthrich (2009)). The loss ratio method then predicts the future

Astin Bulletin 48(2), 699-747. doi: 10.1017/asb.2018.5 © 2018 by Astin Bulletin. All rights reserved.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

increments $S_{i,k}$, $i + k > n + 1$, by

$$\widehat{S}_{i,k} := v_i \cdot \widehat{m}_k \quad \text{with} \quad \widehat{m}_k := \frac{\sum_{\nu=1}^{n-k+1} S_{\nu,k}}{\sum_{\nu=1}^{n-k+1} v_\nu} = \sum_{\nu=1}^{n-k+1} \frac{v_\nu}{\sum_{\mu=1}^{n-k+1} v_\mu} \cdot \frac{S_{\nu,k}}{v_\nu}. \tag{1}$$

In practice, however, it is difficult to obtain good volume measures. The actuary typically only has approximations such as trended premium (all lines of business), vehicle years (motor), payroll (worker's compensation), number of beds (medical malpractice) or tonnage (marine). The additive model and the loss ratio method do not address the fact that the available volume measures might deviate from the model volumes.

The problem of uncertain volume measures is not new and has been analyzed for other reserving methods in a number of papers. For instance, Gluck (1997) uses a random walk model for the "pure premium" in the Cape Cod method, Clark (2008) introduces a reserving method that allows the use of exposure information only for those years where it is reliable and Korn (2016) tackles the issue with Kalman filters for the loss ratios in the Cape Cod method. In the Bornhuetter–Ferguson method, the a priori estimates of the ultimate loss burdens can be interpreted as volume measures. The uncertainty of these a priori estimates is addressed in Saluz *et al.* (2011, 2014).

In the present paper, we denote the available volumes by $\widehat{v}_i$ and interpret them as estimators for the model volumes v_i, which are considered to be unknown parameters. In this setup, the predictions of the loss ratio method are not given by (1), but by

$$\widehat{S}_{i,k} := \widehat{v}_i \cdot \widehat{m}_k \quad \text{with} \quad \widehat{m}_k := \frac{\sum_{\nu=1}^{n-k+1} S_{\nu,k}}{\sum_{\nu=1}^{n-k+1} \widehat{v}_\nu}. \tag{2}$$

If the model volumes are known, i.e., $\widehat{v}_i = v_i$, then (1) and (2) are equal. We introduce an extension of the additive model that includes an assumption regarding the covariance structure of the volume estimates $\widehat{v}_i$. Similar models for the uncertainty of volume estimates have been used by Gluck (1997) and Riegel (2015) in different contexts. In addition, this *extended additive model* includes a lognormal assumption for the incremental claims of (at least) the first development year. This distributional assumption is very convenient for parameter estimation.

In this paper, we analyze the loss ratio method in the light of the extended additive model and introduce a *generalized loss ratio method* that is motivated by this stochastic model.

In the additive model, the variance minimizing weights for the observed loss ratios are inversely proportional to the model volumes (which are assumed to be known). In the extended additive model, however, the optimal weights are not inversely proportional to the volume estimates. We will see that higher weights should be used for more recent years.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Moreover, the mean squared error of prediction (MSEP) of the loss ratio method in the extended additive model is different from the MSEP in the additive model. In particular, if the variances of the volume estimates are large, the MSEP in the extended additive model is typically substantially higher, since the MSEP in the additive model neglects this source of uncertainty.

Parameter estimation is more difficult in the extended additive model than in the additive model. We provide a recursive approach that can be used to calculate estimators for the expected loss ratio increments $m_1, \ldots, m_n$ and for the variance parameters of the extended additive model. But that is not all. An important additional objective of the recursion is to provide *revised volume estimates* $\widehat{v}_i^{(\infty)}$ that are more accurate than the original volume estimates $\widehat{v}_i$.

This leads directly to the idea behind the generalized loss ratio method: We use the revised volume estimates $\widehat{v}_i^{(\infty)}$ instead of the volume estimates $\widehat{v}_i$ in the loss ratio method. In fact, the predictions of the generalized loss ratio method are defined as

$$\widehat{S}_{i,k}^{(\infty)} := \widehat{v}_i^{(\infty)} \cdot \widehat{m}_k^{(\infty)} \quad \text{with} \quad \widehat{m}_k^{(\infty)} := \sum_{\nu=1}^{n-k+1} \widehat{h}_{\nu,k}^{(\infty)} \frac{S_{\nu,k}}{\widehat{v}_\nu^{(\infty)}}$$

for $i + k > n + 1$. The weights $\widehat{h}_{\nu,k}^{(\infty)}$ are also calculated in the recursion and differ from the weights that are used for the increments $S_{\nu,k}/v_\nu$ in the loss ratio method, see (1). The generalized loss ratio method is really a generalization of the loss ratio method since the loss ratio method is the limiting case, as the variances of the volume estimates approach zero.

This paper is organized as follows: In Section 2, we introduce the extended additive model. In the following section, we define the loss ratio method in the context of the extended additive model and analyze some basic properties. Section 4 provides the recursive algorithm for parameter estimation in the extended additive model. We also describe a number of modifications of the algorithm that can be used to avoid problems in certain situations. For instance, it can make sense to use an educated guess instead of the calculated estimator for the variance parameter of the volume estimates if the triangle is small or if claims volatility is high. In Section 5, we calculate weights for the loss ratio method that are approximately optimal in the extended additive model. Section 6 is dedicated to the calculation of the MSEP of the loss ratio method in the extended additive model. The generalized loss ratio method is introduced in Section 7. In the last section, we calculate the prediction uncertainty of the generalized loss ratio method. The results are illustrated by a numerical example that is introduced in Section 3 and continued in the subsequent sections.

A reference implementation of the generalized loss ratio method is available in the `R` package `GLRM` that can be obtained from the author.

Throughout the paper, we assume that all random variables are defined on a fixed probability space $(\Omega, \mathcal{A}, P)$. To improve readability, several technical details have been moved to the appendix.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

2. Extended additive model

In this section, we introduce the *extended additive model*. The model is obtained from the additive model by adding an assumption for the uncertainty of volume estimates and a distributional assumption for the incremental claims. We will work in this stochastic model throughout the paper.

For $i, k \in \{1, \dots, n\}$, let $S_{i,k}$ denote the incremental claims of accident year i in development year k (payments or incurred amounts). We consider $S_{i,k}$ random variables of which we have observations for $i+k \leq n+1$, i.e., the observable claims data are described by the σ-algebra

$$\mathcal{D} := \sigma\{S_{i,k} \mid i + k \leq n + 1\}.$$

We assume that the claims data have been trended to the level of a fixed accident year, i.e., that inflation within each column has been removed. The conventional additive model consists of the following three assumptions (cf. Mack (2002)):

Model Assumption 1. *The increments $S_{i,k}$, $i, k \in \{1, \dots, n\}$, are independent.*

Model Assumption 2. *There are parameters $v_1, \dots, v_n > 0$ and parameters $m_1, \dots, m_n$ such that*

$$\mathrm{E}(S_{i,k}) = v_i \cdot m_k$$

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

Model Assumption 3. *There are parameters $s_1^2, \dots, s_n^2 > 0$ such that*

$$\mathrm{var}(S_{i,k}) = v_i \cdot s_k^2$$

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

In order to distinguish between the theoretical model parameters $v_1, \dots, v_n$ (which are deterministic) and the available volume measures (which are modeled as random variables below), we use the following definition:

Definition 2.1. *The model parameters $v_1, \dots, v_n$ are called model volumes.*

In literature, the model volumes $v_1, \dots, v_n$ are typically assumed to be known. In practice, however, the actuary uses volume measures such as trended premium, vehicle years, payroll or number of beds, which are rarely accurate. We assume that such volume measures $\widehat{v}_i$ are available for the accident years $i = 1, \dots, n$ and interpret them as observations of random variables that are used to estimate the model volumes v_i. In this paper, we refer to these random variables $\widehat{v}_1, \dots, \widehat{v}_n$ as *volume estimates*.

Model Assumption 4. *The volume estimate $\widehat{v}_n$ equals the model volume v_n. The estimators $\widehat{v}_1, \dots, \widehat{v}_{n-1}$ for the unknown model volumes $v_1, \dots, v_{n-1}$ satisfy*

$$\frac{1}{\widehat{v}_i} = \frac{1}{v_i} \cdot f_i \cdots f_{n-1},$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

with independent positive random factors f_i *with* $\mathrm{E}(f_i) = 1$ *and* $\mathrm{var}(f_i) = \varepsilon_i^2 > 0$. *Moreover, we assume that the* σ*-algebras* $\sigma\{f_1, \ldots, f_{n-1}\}$ *and* $\sigma\{S_{i,k} \mid i, k = 1, \ldots, n\}$ *are independent.*

Gluck (1997) and Riegel (2015) use similar assumptions to model uncertainty of volume estimates. Our Model Assumption 4 implies that $\widehat{v}_i^{-1}$ is an unbiased estimator for v_i^{-1} (see Proposition 2.3) and that

$$0 = \mathrm{var}(\widehat{v}_n) < \mathrm{var}(\widehat{v}_{n-1}) < \cdots < \mathrm{var}(\widehat{v}_1). \tag{3}$$

Note that this does not mean that $\widehat{v}_n$ is a better volume measure than $\widehat{v}_1$: The model parameters v_i and m_k are not uniquely determined, since for arbitrary $c > 0$, we can use the parameters cv_i and m_k/c instead. Therefore, we can choose the parameters in such a way that $v_n = \widehat{v}_n$. With this selection, (3) simply means that for small i we are less certain that $\widehat{v}_i$ is comparable to $\widehat{v}_n$ than for large i. For instance, if $\widehat{v}_1, \ldots, \widehat{v}_n$ denote trended premiums, then the selection $v_n = \widehat{v}_n$ corresponds to the case that all accident years are trended to the level of accident year n. Then, the smaller i, the larger the uncertainty in trending, where we attempt to bring the accident year i to the level of accident year n. For a more detailed discussion of the relationship between Model Assumption 4 and the uncertainty in premium trending see Appendix A.

In the Bornhuetter–Ferguson method, the a priori ultimates can be interpreted as volume estimates. However, the uncertainty of these estimates has a different origin. Often original pricings are used as a priori ultimates. The variances of these volume estimates then describe the uncertainty of the original pricings, whereas the motivation for our Model Assumption 4 is uncertainty caused by premium trending. Saluz *et al.* (2011, 2014) therefore use assumptions for the uncertainty of the a priori ultimates that differ substantially from our Model Assumption 4.

Eventually, we add a distributional assumption for the increments that is used below for parameter estimation.

Model Assumption 5. *For a given subset* $\mathcal{K} \subseteq \{1, \ldots, n-1\}$ *with* $1 \in \mathcal{K}$, *the increments* $S_{i,k}$ *are lognormally distributed for all* $i \in \{1, \ldots, n\}$ *and* $k \in \mathcal{K}$.

This assumption requires that at least the claims in the first development year are lognormally distributed. After removing large losses this is quite plausible in many cases: The central limit theorem for compound Poisson distributions (see Kaas *et al.* (2001)) suggests to use a normal distribution if the claims burden has a compound Poisson distribution with a sufficiently large expected number of claims. In this situation, the coefficient of variation is relatively small, i.e., lognormal and normal distribution have a very similar shape. For large coefficients of variation, the lognormal is far more plausible than the normal distribution. Methods for removing large losses are described in Gütschow *et al.* (2017) and Riegel (2014, 2016).

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

In practice, we recommend to choose $\mathcal{K} \subseteq \{1, \dots, n-1\}$ as large as possible, i.e., to include all development years where the observed increments are positive:

$$\mathcal{K} := \{k \in \{1, \dots, n-1\} \mid S_{i,k} > 0 \text{ for all } i \text{ with } i + k \leq n + 1\}.$$

If this set does not contain 1, i.e., if $S_{i,1} \leq 0$ for an $i \in \{1, \dots, n\}$, then the methods described in this paper are not applicable.

Definition 2.2. *The stochastic model determined by Model Assumptions 1–5 is called the extended additive model.*

As from now, we presume that the assumptions of the extended additive model are fulfilled. We use the abbreviation

$$\beta_{i,j} := \prod_{\nu=\max(i,j)}^{n-1} (1 + \varepsilon_\nu^2).$$

The following proposition is a consequence of Model Assumption 4. It will be used repeatedly in the following sections.

Proposition 2.3. *For $i, j \in \{1, \dots, n\}$, we have*

$$\mathrm{E}\left(\frac{1}{\widehat{v}_i}\right) = \frac{1}{v_i} \quad \textit{and} \quad \mathrm{E}\left(\frac{1}{\widehat{v}_i \widehat{v}_j}\right) = \frac{\beta_{i,j}}{v_i v_j}.$$

Therefore,

$$\mathrm{Cov}\left(\frac{1}{\widehat{v}_i}, \frac{1}{\widehat{v}_j}\right) = \frac{1}{v_i v_j} \left(\beta_{i,j} - 1\right).$$

For $i \leq j$, we have

$$\mathrm{E}\left(\frac{\widehat{v}_j}{\widehat{v}_i} \,\middle|\, \widehat{v}_j, \dots, \widehat{v}_n\right) = \mathrm{E}\left(\frac{\widehat{v}_j}{\widehat{v}_i}\right) = \frac{v_j}{v_i}.$$

Proof. We have

$$\mathrm{E}\left(\frac{1}{\widehat{v}_i}\right) = \mathrm{E}\left(\frac{1}{v_i} f_i \cdots f_{n-1}\right) = \frac{1}{v_i} \mathrm{E}(f_i) \cdots \mathrm{E}(f_{n-1}) = \frac{1}{v_i}$$

and for $i \leq j$

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{\widehat{v}_i \widehat{v}_j}\right) &= \mathrm{E}\left(\frac{1}{v_i v_j} f_i \cdots f_{j-1} f_j^2 \cdots f_{n-1}^2\right) \\
&= \frac{1}{v_i v_j} \mathrm{E}(f_i) \cdots \mathrm{E}(f_{j-1}) \, \mathrm{E}(f_j^2) \cdots \mathrm{E}(f_{n-1}^2) \\
&= \frac{1}{v_i v_j} \prod_{\nu=j}^{n-1} (1 + \varepsilon_\nu^2) = \frac{\beta_{i,j}}{v_i v_j}.
\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Moreover, for $i \le j$

$$\begin{aligned}\mathrm{E}\left(\frac{\widehat{v}_j}{\widehat{v}_i}\middle|\widehat{v}_j,\dots,\widehat{v}_n\right) &= \frac{v_j}{v_i}\,\mathrm{E}\left(\frac{f_i\cdots f_{n-1}}{f_j\cdots f_{n-1}}\middle|\widehat{v}_j,\dots,\widehat{v}_n\right)\\ &= \frac{v_j}{v_i}\,\mathrm{E}(f_i\cdots f_{j-1}\,|\,\widehat{v}_j,\dots,\widehat{v}_n)\\ &= \frac{v_j}{v_i}\,\mathrm{E}(f_i\cdots f_{j-1}) = \frac{v_j}{v_i}\,\mathrm{E}(f_i)\cdots\mathrm{E}(f_{j-1}) = \frac{v_j}{v_i}.\end{aligned}$$

■

In the following sections, we will repeatedly need the expectation of products and quotients of the lognormally distributed random variables $S_{i,k}$. It is convenient to use the notation

$$\gamma_{i,k} := 1 + \frac{v_i s_k^2}{v_i^2 m_k^2}$$

for $i \in \{1,\dots,n\}$ and $k \in \mathcal{K}$.

Proposition 2.4. *For $i \in \{1,\dots,n\}$, $k \in \mathcal{K}$ and $p \in \mathbb{Z}$, we have*

$$\mathrm{E}(S_{i,k}^p) = (v_i m_k)^p \cdot \gamma_{i,k}^{\frac{1}{2}p(p-1)}.$$

The proof uses standard properties of the lognormal distribution and can be obtained on request from the author.

3. Loss ratio method

The loss ratio method is associated with the additive model, which is determined by Model Assumptions 1–3 of the previous section. In this section, we discuss the loss ratio method in the light of the extended additive model of Section 2. In particular, we assume that the model volumes v_i are not known. Only the volume estimates $\widehat{v}_i$ are available.

Definition 3.1. *For $k = 1,\dots,n$, let $\mathbf{w}_k = (w_{1,k},\dots,w_{n-k+1,k})$ be weights such that*

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

for all k. The weights can be deterministic, but we also consider weights $\mathbf{w}_k$ that are (not necessarily measurable) functions of the observable random variables, i.e., of $\widehat{v}_i$ and $S_{i,k}$ with $i = 1,\dots,n$ and $k = 1,\dots,n-i+1$. Let $\mathbf{w} := (\mathbf{w}_1,\dots,\mathbf{w}_n)$ and let

$$\widehat{m}_k^{\mathbf{w}} := \sum_{i=1}^{n-k+1} w_{i,k}\frac{S_{i,k}}{\widehat{v}_i} \qquad (k = 1,\dots,n)$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

be the corresponding estimators for m_k. Then, for $i, k \in \{2, \dots, n\}$ with $i + k > n + 1$ the predictions

$$\widehat{S}^{\mathbf{w}}_{i,k} := \widehat{v}_i \cdot \widehat{m}^{\mathbf{w}}_k$$

for $S_{i,k}$ are called the predictions of the loss ratio method with weights $\mathbf{w}$*. We define the canonical weights* $\mathbf{w}^c$ *by*

$$w^c_{i,k} := \frac{\widehat{v}_i}{\sum_{j=1}^{n-k+1} \widehat{v}_j}.$$

If we use the canonical weights, then we drop the weights from the notation, i.e., we write $\widehat{m}_k := \widehat{m}^{\mathbf{w}^c}_k$ and say that $\widehat{S}_{i,k} := \widehat{S}^{\mathbf{w}^c}_{i,k}$ are the predictions of the loss ratio method.

The canonical weights are a natural selection for the weights, since they minimize the variances of the predictions in the additive model where the model volumes are assumed to be known, i.e., $\widehat{v}_i = v_i$ (cf. Mack (2002)). In practice, however, often actuarial judgement is used to select the weights for the loss ratio method.

The loss ratio method yields unbiased predictions for future claims development in the additive model. The following proposition shows that this remains true in the extended additive model in case of deterministic weights. Let

$$\mathcal{D}^k := \sigma\{S_{i,l} \mid i + l \le n + 1, l \ne k\} \subset \mathcal{D}.$$

Proposition 3.2. *The predictions of the loss ratio method with deterministic weights* $\mathbf{w}$ *are (conditionally) unbiased. More precisely, if* $\mathbf{w}$ *is deterministic, then we have*

$$\mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k} \mid \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n) = \mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k}) = \mathrm{E}(S_{i,k}) = \mathrm{E}(S_{i,k} \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n)$$

for all $i, k \in \{2, \dots, n\}$ with $i + k > n + 1$.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Proof. For $i+k > n+1$, we have

$$\begin{aligned}
\mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k} \mid \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n) &= \mathrm{E}(\widehat{v}_i \widehat{m}^{\mathbf{w}}_k \mid \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n) \\
&= \mathrm{E}\left(\widehat{v}_i \sum_{j=1}^{n-k+1} w_{j,k} \frac{S_{j,k}}{\widehat{v}_j} \,\middle|\, \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n \right) \\
&= \sum_{j=1}^{n-k+1} w_{j,k} \, \mathrm{E}\left(\frac{\widehat{v}_i}{\widehat{v}_j} S_{j,k} \,\middle|\, \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n \right) \\
&= \sum_{j=1}^{n-k+1} w_{j,k} \, \mathrm{E}\left(\frac{\widehat{v}_i}{\widehat{v}_j} \,\middle|\, \widehat{v}_i, \dots, \widehat{v}_n \right) \mathrm{E}(S_{j,k} \mid \mathcal{D}^k) \\
&= \sum_{j=1}^{n-k+1} w_{j,k} \frac{v_i}{v_j} \, \mathrm{E}(S_{j,k}) = \sum_{j=1}^{n-k+1} w_{j,k} \frac{v_i}{v_j} v_j m_k \\
&= v_i m_k = \mathrm{E}(S_{i,k}).
\end{aligned}$$

In particular, $\mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k} \mid \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n)$ is constant, which implies $\mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k} \mid \mathcal{D}^k, \widehat{v}_i, \dots, \widehat{v}_n) = \mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k})$. Eventually, we have $\mathrm{E}(S_{i,k}) = \mathrm{E}(S_{i,k} \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n)$ since $S_{i,k}$ and $\sigma\{\mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n\}$ are independent for $i+k > n+1$. ■

If the weights $\mathbf{w}$ are functions of the observable data, then the predictions $\widehat{S}^{\mathbf{w}}_{i,k}$ may not be measurable, i.e., $\mathrm{E}(\widehat{S}^{\mathbf{w}}_{i,k})$ may not exist. If the weights $\mathbf{w}$ are bounded and $\sigma(\mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n)$-measurable (which is in particular true for the canonical weights), then the predictions $\widehat{S}^{\mathbf{w}}_{i,k}$ are square integrable. However, the predictions are not unbiased in general.

Example 3.3. *Throughout this paper, we will apply the discussed approaches to the numerical example from Riegel (2014, 2016). In contrast to Riegel (2016), we do not interpret the (trended) premiums as model volumes v_i, but as volume estimates $\widehat{v}_i$ in the extended additive model. The example contains premiums and payments from a large German motor quota share. In this segment, claims volatility is relatively low and uncertainty in premium trending is relatively high.*

In Table 1, we show the volume estimates $\widehat{v}_i$ and the incremental payments $\widehat{S}_{i,k} := S_{i,k}$ for $i+k \le n+1$. The predictions $\widehat{S}_{i,k}$ with $i+k > n+1$ of the loss ratio method with canonical weights are shown in italic print. In addition, we provide the estimators $\widehat{m}_k$ from the loss ratio method.

4. Parameter estimation

In this section, we describe a recursive algorithm that can be used to estimate the parameters of the extended additive model. One objective, in particular, is

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 1

PREDICTIONS OF THE LOSS RATIO METHOD WITH CANONICAL WEIGHTS.

i	$\widehat{v}_i$	$\widehat{S}_{i,1}$	$\widehat{S}_{i,2}$	$\widehat{S}_{i,3}$	$\widehat{S}_{i,4}$	$\widehat{S}_{i,5}$	$\widehat{S}_{i,6}$	$\widehat{S}_{i,7}$	$\widehat{S}_{i,8}$
1	316,285	151,727	45,469	7,631	5,902	3,757	3,536	3,123	2,756
2	301,927	147,475	41,521	7,215	4,419	3,050	1,789	2,103	*2,631*
3	271,299	132,485	42,212	7,303	4,753	3,357	2,196	*2,293*	*2,364*
4	273,377	140,542	47,218	8,857	5,705	4,134	*2,311*	*2,311*	*2,382*
5	280,869	137,054	46,568	7,216	5,207	*3,453*	*2,375*	*2,374*	*2,447*
6	311,389	152,017	50,926	8,675	*5,605*	*3,829*	*2,633*	*2,632*	*2,713*
7	322,885	166,046	54,902	*8,627*	*5,812*	*3,970*	*2,730*	*2,729*	*2,814*
8	330,848	170,026	*52,352*	*8,840*	*5,955*	*4,068*	*2,797*	*2,797*	*2,883*
	$\widehat{m}_k$	49.71%	15.82%	2.67%	1.80%	1.23%	0.85%	0.85%	0.87%

to obtain revised volume estimates $\widehat{v}_i^{(\infty)}$ that are more accurate than the volume estimates $\widehat{v}_i$. The idea behind the generalized loss ratio method, which will be introduced in Section 7, is to replace the volume estimates $\widehat{v}_i$ by the revised volume estimates $\widehat{v}_i^{(\infty)}$ in the loss ratio method.

In order to reduce the number of parameters that have to be estimated, we assume from now on that $(\varepsilon_1, \dots, \varepsilon_{n-1})^t = \varepsilon \cdot (c_1, \dots, c_{n-1})^t$ with an *unknown* parameter $\varepsilon > 0$ and a *known* vector $\mathbf{c} = (c_1, \dots, c_{n-1})^t$ with $c_i > 0$. We recommend to use $\mathbf{c} := \mathbb{1}_{n-1} := (1, \dots, 1)^t \in \mathbb{R}^{n-1}$ unless there is a reason to select a different $\mathbf{c}$. This choice of $\mathbf{c}$ implies $\operatorname{var}(f_i) = \varepsilon^2$ for $i = 1, \dots, n-1$.

In Section 4.1, we describe the idea behind the recursion. In Section 4.2, we provide the theoretical background of the recursion. Section 4.3 contains the algorithm for computer programming. In Section 4.4, we describe several modifications to the algorithm that can be used to avoid issues in certain situations. Finally, we continue our numerical example in Section 4.5.

4.1. Outline of the approach

Let $\mathbf{X} := (\widehat{v}_1, \dots, \widehat{v}_n, (S_{i,k})_{i+k \le n+1})$ denote the random vector containing the observable data and let

$$\mathbf{\Theta} = (v_1^{-1}, \dots, v_n^{-1}, m_1, \dots, m_n, s_1^2, \dots, s_{n-1}^2, \varepsilon^2),$$

be the vector containing the parameters of the model (apart from s_n^2). Let

$$D_{\mathcal{K}} := \{(x_1, \dots, x_n) \mid x_k > 0 \text{ für } k \in \mathcal{K}\} \subset \mathbb{R}^n.$$

In Section 4.2, we define a map

$$\mathbf{\Psi} \colon (0, \infty)^n \times D_{\mathcal{K}} \times (0, \infty)^n \times \mathbf{X}(\Omega) \to (0, \infty)^n \times D_{\mathcal{K}} \times (0, \infty)^n,$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

with the components $\boldsymbol{\Psi} = (\Psi_1^v, \ldots, \Psi_n^v, \Psi_1^m, \ldots, \Psi_n^m, \Psi_1^s, \ldots, \Psi_{n-1}^s, \Psi^\varepsilon)$, which is determined by $\mathcal{K}$ and $\mathbf{c}$ such that

$$(\widehat{v}_i^*)^{-1} := \Psi_i^v(\boldsymbol{\Theta}, \mathbf{X}), \qquad \widehat{m}_k^* := \Psi_k^m(\boldsymbol{\Theta}, \mathbf{X}), \tag{4}$$

$$(\widehat{s}_k^*)^2 := \Psi_k^s(\boldsymbol{\Theta}, \mathbf{X}), \qquad (\widehat{\varepsilon}^*)^2 := \boldsymbol{\Psi}^\varepsilon(\boldsymbol{\Theta}, \mathbf{X}) \tag{5}$$

satisfy

$$\mathrm{E}\left((\widehat{v}_i^*)^{-1}\right) = v_i^{-1}, \qquad \mathrm{E}\left(\widehat{m}_k^*\right) = m_k, \tag{6}$$

$$\mathrm{E}\left((\widehat{s}_k^*)^2\right) = s_k^2, \qquad \mathrm{E}\left((\widehat{\varepsilon}^*)^2\right) = \varepsilon^2, \tag{7}$$

respectively, for all i, k and arbitrary model parameters $\boldsymbol{\Theta}$. With

$$\widehat{\boldsymbol{\Theta}}^* = ((\widehat{v}_1^*)^{-1}, \ldots, (\widehat{v}_n^*)^{-1}, \widehat{m}_1^*, \ldots, \widehat{m}_n^*, (s_1^*)^2, \ldots, (s_{n-1}^*)^2, (\varepsilon^*)^2),$$

equations (4)–(7) can be written concisely as

$$\widehat{\boldsymbol{\Theta}}^* := \boldsymbol{\Psi}(\boldsymbol{\Theta}, \mathbf{X}) \quad \text{and} \quad \mathrm{E}\left(\widehat{\boldsymbol{\Theta}}^*\right) = \boldsymbol{\Theta}. \tag{8}$$

The map $\boldsymbol{\Psi}$ is constructed in such a way that

$$\mathrm{var}\left((\widehat{v}_i^*)^{-1}\right) < \mathrm{var}\left(\widehat{v}_i^{-1}\right)$$

for $i = 1, \ldots, n-1$, i.e., $(\widehat{v}_i^*)^{-1}$ is a better estimator for v_i^{-1} than $\widehat{v}_i^{-1}$. The estimator $\widehat{m}_k^*$ is a linear combination of the estimators

$$\frac{S_{1,k}}{\widehat{v}_i^*}, \ldots, \frac{S_{n-k+1,k}}{\widehat{v}_i^*}.$$

This matches the idea of replacing the volume estimates $\widehat{v}_i$ in the loss ratio method with better volume estimates, when we introduce the generalized loss ratio method in Section 7.

The "estimators" $\widehat{v}_i^*$, $\widehat{m}_k^*$, $(\widehat{s}_k^*)^2$ and $(\widehat{\varepsilon}^*)^2$ cannot be used directly since they do not only depend on the observed data $\mathbf{X}$, but also on the unknown model parameters $\boldsymbol{\Theta}$. However, the map $\boldsymbol{\Psi}$ is used in Section 4.3 to estimate the parameters recursively. We start with initial estimators $\widehat{\boldsymbol{\Theta}}^{(0)} = ((\widehat{v}_1^{(0)})^{-1}, \ldots, (\widehat{v}_n^{(0)})^{-1}, \widehat{m}_1^{(0)}, \ldots, \widehat{m}_n^{(0)}, (s_1^{(0)})^2, \ldots, (s_{n-1}^{(0)})^2, (\varepsilon^{(0)})^2)$ and then calculate $\widehat{\boldsymbol{\Theta}}^{(\nu)} = ((\widehat{v}_1^{(\nu)})^{-1}, \ldots, (\widehat{v}_n^{(\nu)})^{-1}, \widehat{m}_1^{(\nu)}, \ldots, \widehat{m}_n^{(\nu)}, (s_1^{(\nu)})^2, \ldots, (s_{n-1}^{(\nu)})^2, (\varepsilon^{(\nu)})^2)$ for $\nu \geq 1$ inductively by

$$\widehat{\boldsymbol{\Theta}}^{(\nu+1)} := \boldsymbol{\Psi}(\widehat{\boldsymbol{\Theta}}^{(\nu)}, \mathbf{X}), \tag{9}$$

i.e., in each iteration, we replace the unknown parameters $\boldsymbol{\Theta}$ in (8) with their estimates $\widehat{\boldsymbol{\Theta}}^{(\nu)}$ from the previous iteration and hope that the estimators improve in each iteration. If the recursion converges, then we define $\widehat{v}_i^{(\infty)} := \lim_{\nu\to\infty} \widehat{v}_i^{(\nu)}$, $\widehat{m}_k^{(\infty)} := \lim_{\nu\to\infty} \widehat{m}_k^{(\nu)}$, $(\widehat{s}_k^{(\infty)})^2 := \lim_{\nu\to\infty} (\widehat{s}_k^{(\nu)})^2$ and $(\widehat{\varepsilon}^{(\infty)})^2 := \lim_{\nu\to\infty} (\widehat{\varepsilon}^{(\nu)})^2$.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Of course, the estimators $\widehat{v}_i^{(\infty)}$, $m_k^{(\infty)}$, $(\widehat{s}_k^{(\infty)})^2$ and $(\widehat{\varepsilon}^{(\infty)})^2$ will have a bias in general, but they are calculated in a consistent way and we may hope that $(\widehat{v}_i^{(\infty)})^{-1}$ is still a better estimate for v_i^{-1} than $\widehat{v}_i^{-1}$.

Alternatively, we can interpret the resulting estimators in terms of fixed points. According to (8), we have

$$\mathrm{E}\left(\boldsymbol{\Psi}(\boldsymbol{\Theta}, \mathbf{X})\right) = \mathrm{E}\left(\widehat{\boldsymbol{\Theta}}^*\right) = \boldsymbol{\Theta},$$

i.e., the parameter vector $\boldsymbol{\Theta}$ is a fixed point of the map

$$\mathrm{E}\left(\boldsymbol{\Psi}(\,\cdot\,, \mathbf{X})\right) : (0,\infty)^n \times D_{\mathcal{K}} \times (0,\infty)^n \to (0,\infty)^n \times D_{\mathcal{K}} \times (0,\infty)^n.$$

If the recursion converges, then

$$\widehat{\boldsymbol{\Theta}}^{(\infty)} = ((\widehat{v}_1^{(\infty)})^{-1}, \ldots, (\widehat{v}_n^{(\infty)})^{-1}, \widehat{m}_1^{(\infty)}, \ldots, \widehat{m}_n^{(\infty)}, (s_1^{(\infty)})^2, \ldots, (s_{n-1}^{(\infty)})^2, (\varepsilon^{(\infty)})^2)$$

is a fixed point of the map

$$\boldsymbol{\Psi}(\,\cdot\,, \mathbf{X}) \colon (0,\infty)^n \times D_{\mathcal{K}} \times (0,\infty)^n \to (0,\infty)^n \times D_{\mathcal{K}} \times (0,\infty)^n,$$

i.e., with our approach we hope that the *realized* map $\boldsymbol{\Psi}(\,\cdot\,, \mathbf{X})$ is similar to the *expected* map $\mathrm{E}\left(\boldsymbol{\Psi}(\,\cdot\,, \mathbf{X})\right)$, and therefore has similar fixed points.

4.2. Theoretical background of the recursion

As described in Section 4.1, the goal of this section is to calculate unbiased estimators $(\widehat{v}_i^*)^{-1}$, $\widehat{m}_k^*$, $(\widehat{s}_k^*)^2$ and $(\widehat{\varepsilon}^*)^2$ as functions of the model parameters v_i^{-1}, m_k, s_k^2 and ε^2 and the observable random variables $\mathbf{X}$.

Notation 4.1. *For $i, j \in \{1, \ldots, n\}$ and $k \in \mathcal{K}$ let*

$$\widehat{r}_{i,(0,0)} := \frac{1}{\widehat{v}_i} \quad \textit{and} \quad \widehat{r}_{i,(j,k)} := \frac{1}{\widehat{v}_j} \frac{S_{j,k}}{S_{i,k}} \gamma_{i,k}^{-1}.$$

Let $\mathcal{I}_n := \{(0,0)\}$ and for $i = 1, \ldots, n-1$

$$\mathcal{I}_i^* := \{(j,k) \,|\, k \in \mathcal{K},\ j \neq i,\ \max(i,j) + k \leq n+1\} \quad \textit{and} \quad \mathcal{I}_i := \mathcal{I}_i^* \cup \{(0,0)\}.$$

Remark 4.2. *As an immediate consequence of Proposition 2.4, we have*

$$\mathrm{E}(\widehat{r}_{i,(j,k)}) = \frac{1}{v_i}$$

for all $i = 1, \ldots, n$ and $(j,k) \in \mathcal{I}_i$, i.e., apart from $\widehat{v}_i^{-1}$, we have a number of additional unbiased estimators for v_i. The idea behind the estimators $\widehat{r}_{i,(j,k)}$ with $i > 0$ becomes obvious if we neglect the bias correction factor $\gamma_{i,k}^{-1}$ for a moment: Estimate v_i^{-1} in such a way that the incremental loss ratio $S_{i,k} \cdot \widehat{r}_{i,(j,k)}$ in development year i equals the corresponding incremental loss ratio $S_{j,k}/\widehat{v}_j$.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Notation 4.3. *Let $N_i := |\mathcal{I}_i|$. For each $i = 1, \dots, n$, we fix a bijection*

$$\varphi_i : \{1, \dots, N_i\} \to \mathcal{I}_i$$

with $\varphi_i(1) = (0, 0)$ and use the notation

$$\widehat{r}_{i,j} := \widehat{r}_{i,\varphi_i(j)} \quad \text{and} \quad \widehat{\mathbf{r}}_i := (\widehat{r}_{i,1}, \dots, \widehat{r}_{i,N_i})^t.$$

For $i < n$ the covariance matrix $\mathbf{R}_i := \mathrm{Cov}(\widehat{\mathbf{r}}_i)$ is calculated in Appendix D, where we also show that this covariance matrix is always positive definite. Let $\mathbf{g}_n := g_{n,1} := 1$,

$$\mathbf{g}_i := (g_{i,1}, \dots, g_{i,N_i}) := \frac{\mathbf{R}_i^{-1}\mathbb{1}_{N_i}}{\mathbb{1}_{N_i}^t \mathbf{R}_i^{-1}\mathbb{1}_{N_i}} \quad (i < n)$$

and

$$g_{i,(j,k)} := g_{i,\varphi_i^{-1}(j,k)}.$$

Then, we define $\widehat{v}_i^*$ by

$$\frac{1}{\widehat{v}_i^*} := \sum_{(j,k)\in\mathcal{I}_i} g_{i,(j,k)}\widehat{r}_{i,(j,k)}.$$

Proposition 4.4. *The unbiased estimator $(\widehat{v}_i^*)^{-1}$ for v_i^{-1} has minimal variance amongst all convex combinations of the estimators $\widehat{r}_{i,1}, \dots, \widehat{r}_{i,N_i}$.*

Proof. This is an immediate consequence of Proposition B.1. ∎

Lemma 4.5. *For $i, k \in \{1, \dots, n\}$ with $i + k \le n + 1$ let $\widehat{m}_{i,k}^* := S_{i,k}/\widehat{v}_i^*$. Then, $\widehat{m}_{i,k}^*$ is an estimator for m_k with bias*

$$b_{i,k} := m_k(\gamma_{i,k}^{-1} - 1) \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l=k}} g_{i,(j,k)}$$

(unbiased if $k \notin \mathcal{K}$).

Proof. We have

$$\begin{aligned}\widehat{m}_{i,k}^* = \frac{S_{i,k}}{\widehat{v}_i^*} &= g_{i,(0,0)}\frac{S_{i,k}}{\widehat{v}_i} + \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l\neq k}} g_{i,(j,l)}\frac{S_{i,k}}{\widehat{v}_j}\frac{S_{j,l}}{S_{i,l}}\gamma_{i,l}^{-1} + \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l=k}} g_{i,(j,k)}\frac{S_{i,k}}{\widehat{v}_j}\frac{S_{j,k}}{S_{i,k}}\gamma_{i,k}^{-1} \\ &= g_{i,(0,0)}\frac{S_{i,k}}{\widehat{v}_i} + \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l\neq k}} g_{i,(j,l)}\frac{S_{i,k}}{\widehat{v}_j}\frac{S_{j,l}}{S_{i,l}}\gamma_{i,l}^{-1} + \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l=k}} g_{i,(j,k)}\frac{S_{j,k}}{\widehat{v}_j}\gamma_{i,k}^{-1}\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

and therefore

$$\begin{aligned}\mathrm{E}(\widehat{m}^*_{i,k}) &= g_{i,(0,0)}m_k + \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l\neq k}} g_{i,(j,l)}m_k + \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l=k}} g_{i,(j,k)}m_k\gamma_{i,k}^{-1}\\ &= g_{i,(0,0)}m_k + \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l\neq k}} g_{i,(j,l)}m_k + \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l=k}} g_{i,(j,k)}m_k + \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l=k}} g_{i,(j,k)}m_k(\gamma_{i,k}^{-1}-1)\\ &= m_k + \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l=k}} g_{i,(j,k)}m_k(\gamma_{i,k}^{-1}-1) = m_k + b_{i,k}.\end{aligned}$$

■

Notation 4.6. *For $k = 1, \ldots, n$ let $\mathbf{b}_k := (b_{1,k}, \ldots, b_{n-k+1,k})^t$ and let $\mathbf{M}_k$ denote the covariance matrix of*

$$\widehat{\mathbf{m}}^*_k := (\widehat{m}^*_{1,k}, \widehat{m}^*_{2,k}, \ldots, \widehat{m}^*_{n-k+1,k})^t.$$

The covariance matrix $\mathbf{M}_k$ is calculated in Appendix E, where we also show that $\mathbf{M}_k$ is positive definite for $k > 1$. If $\mathbf{M}_k$ is positive definite, then we define

$$\mathbf{d}_k := (d_{1,k}, \ldots, d_{n-k+1,k})^t := \frac{(\mathbf{M}_k + \mathbf{b}_k\mathbf{b}_k^t)^{-1}\mathbb{1}_{n-k+1}}{\mathbb{1}^t_{n-k+1}(\mathbf{M}_k + \mathbf{b}_k\mathbf{b}_k^t)^{-1}\mathbb{1}_{n-k+1}}.$$

There might exist (rather exotic) cases where the matrix $\mathbf{M}_1$ is singular. In this case, we simply use $d_{1,1} = \cdots = d_{n,1} := 1/n$.

Proposition 4.7. *If $\mathbf{M}_k$ is positive definite, then the estimator*

$$\widetilde{m}^*_k := d_{1,k}\widehat{m}^*_{1,k} + \cdots + d_{n-k+1,k}\widehat{m}^*_{n-k+1,k}$$

for m_k minimizes the mean squared error amongst all convex combinations of the estimators $\widehat{m}^*_{1,k}, \ldots, \widehat{m}^*_{n-k+1,k}$.

Proof. This is an immediate consequence of Lemma 4.5 and Proposition B.1. ■

For $i + k \leq n + 1$, let

$$a_{i,k} := (\gamma_{i,k}^{-1} - 1) \sum_{\substack{(j,l)\in\mathcal{I}^*_i\\ l=k}} g_{i,(j,k)}$$

and $\mathbf{a}_k := (a_{1,k}, \ldots, a_{n-k+1,k})^t$. Then, we have $\mathbf{b}_k = m_k\mathbf{a}_k$ and Lemma 4.5 implies $\mathrm{E}(\widetilde{m}^*_k) = m_k(1 + \mathbf{d}_k^t\mathbf{a}_k)$. Therefore, we obtain the following.

Corollary 4.8. *The estimator*

$$\widehat{m}^*_k := \frac{\widetilde{m}^*_k}{1 + \mathbf{d}_k^t\mathbf{a}_k}$$

for m_k is unbiased.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

We define

$$\mathbf{h}_k := (h_{1,k}, \ldots, h_{n-k+1,k})^t := \frac{\mathbf{d}_k}{1 + \mathbf{d}_k^t \mathbf{a}_k}.$$

Then, we have

$$\widehat{m}_k^* = \sum_{i=1}^{n-k+1} h_{i,k} \widehat{m}_{i,k}^* = \sum_{i=1}^{n-k+1} h_{i,k} \frac{S_{i,k}}{\widehat{v}_i^*}.$$

Proposition 4.9. *Let* $\mathbf{e}_1^{(v)}, \ldots, \mathbf{e}_v^{(v)}$ *denote the canonical basis of* $\mathbb{R}^v$. *With*

$$\Gamma_{i,k} := \frac{s_k^2}{(\mathbf{e}_i^{(n-k+1)} - \mathbf{d}_k)^t (\mathbf{M}_k + \mathbf{b}_k \mathbf{b}_k^t)(\mathbf{e}_i^{(n-k+1)} - \mathbf{d}_k)},$$

we obtain the unbiased estimator

$$\widehat{(s_k^*)}^2 := \frac{1}{n-k+1} \sum_{i=1}^{n-k+1} \Gamma_{i,k} \cdot \left(\frac{S_{i,k}}{\widehat{v}_i^*} - \widetilde{m}_k^* \right)^2$$

for s_k^2 *with* $k < n$. *For* $k = n$, *we use*

$$\widehat{(s_n^*)}^2 := \min \left((\widehat{s}_{n-1}^*)^4 / (\widehat{s}_{n-2}^*)^2, (\widehat{s}_{n-2}^*)^2 \right).$$

Proof. The estimators $\widehat{(s_k^*)}^2$ are unbiased for $k < n$ since

$$\mathrm{E}\left((\widehat{m}_{i,k}^* - \widetilde{m}_k^*)^2 \right) = (\mathbf{e}_i^{(n-k+1)} - \mathbf{d}_k)^t (\mathbf{M}_k + \mathbf{b}_k \mathbf{b}_k^t)(\mathbf{e}_i^{(n-k+1)} - \mathbf{d}_k)$$

(cf. Proposition B.1). ∎

Notation 4.10. *For* $i = 1, \ldots, n-1$, *the covariance matrix* $\mathbf{Q}_i := \mathrm{Cov}(\widehat{\mathbf{r}}_i, \widehat{\mathbf{r}}_{i+1})$ *is calculated in Appendix D. Let*

$$\begin{aligned} \Lambda_i := {} & v_i^2 \cdot (\mathbf{e}_1^{(N_i)} - \mathbf{g}_i)^t \mathbf{R}_i (\mathbf{e}_1^{(N_i)} - \mathbf{g}_i) - 2 v_i v_{i+1} \cdot (\mathbf{e}_1^{(N_i)} - \mathbf{g}_i)^t \mathbf{Q}_i (\mathbf{e}_1^{(N_{i+1})} - \mathbf{g}_{i+1}) \\ & + v_{i+1}^2 \cdot (\mathbf{e}_1^{(N_{i+1})} - \mathbf{g}_{i+1})^t \mathbf{R}_{i+1} (\mathbf{e}_1^{(N_{i+1})} - \mathbf{g}_{i+1}) \end{aligned}$$

and

$$\Delta_i := \frac{\varepsilon^2}{\Lambda_i}.$$

Proposition 4.11. *The estimator*

$$\widehat{(\varepsilon^*)}^2 := \frac{1}{n-1} \sum_{i=1}^{n-1} \Delta_i \cdot \left[v_i \left(\frac{1}{\widehat{v}_i} - \frac{1}{\widehat{v}_i^*} \right) - v_{i+1} \left(\frac{1}{\widehat{v}_{i+1}} - \frac{1}{\widehat{v}_{i+1}^*} \right) \right]^2$$

for ε^2 *is unbiased.*

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Proof. This is a consequence of

$$\begin{aligned}\Lambda_i &= v_i^2 \cdot (\mathbf{e}_1^{(N_i)} - \mathbf{g}_i)^t \mathbf{R}_i (\mathbf{e}_1^{(N_i)} - \mathbf{g}_i) - 2v_i v_{i+1} \cdot (\mathbf{e}_1^{(N_i)} - \mathbf{g}_i)^t \mathbf{Q}_i (\mathbf{e}_1^{(N_{i+1})} - \mathbf{g}_{i+1}) \\ &\qquad + v_{i+1}^2 \cdot (\mathbf{e}_1^{(N_{i+1})} - \mathbf{g}_{i+1})^t \mathbf{R}_{i+1} (\mathbf{e}_1^{(N_{i+1})} - \mathbf{g}_{i+1}) \\ &= v_i^2 \, \mathrm{E}\left[\left(\frac{1}{\widehat{v}_i} - \frac{1}{\widehat{v}_i^*}\right)^2\right] - 2v_i v_{i+1} \, \mathrm{E}\left[\left(\frac{1}{\widehat{v}_i} - \frac{1}{\widehat{v}_i^*}\right)\left(\frac{1}{\widehat{v}_{i+1}} - \frac{1}{\widehat{v}_{i+1}^*}\right)\right] \\ &\qquad + v_{i+1}^2 \, \mathrm{E}\left[\left(\frac{1}{\widehat{v}_{i+1}} - \frac{1}{\widehat{v}_{i+1}^*}\right)^2\right] \\ &= \mathrm{E}\left(\left[v_i\left(\frac{1}{\widehat{v}_i} - \frac{1}{\widehat{v}_i^*}\right) - v_{i+1}\left(\frac{1}{\widehat{v}_{i+1}} - \frac{1}{\widehat{v}_{i+1}^*}\right)\right]^2\right).\end{aligned}$$

■

The unbiased estimators $\widehat{\boldsymbol{\Theta}}^* = ((\widehat{v}_1^*)^{-1}, \ldots, (\widehat{v}_n^*)^{-1}, \widehat{m}_1^*, \ldots, \widehat{m}_n^*, (s_1^*)^2, \ldots, (s_{n-1}^*)^2, (\varepsilon^*)^2)$ from Propositions 4.4, 4.9, 4.11 and from Corollary 4.8 can all be interpreted as functions of the parameter vector $\boldsymbol{\Theta}$ and the observable data $\mathbf{X}$. Using this interpretation, we have

$$\widehat{\boldsymbol{\Theta}}^* = \boldsymbol{\Psi}(\boldsymbol{\Theta}, \mathbf{X})$$

with a map $\boldsymbol{\Psi}$ that is determined by $\mathcal{K}$ and $\mathbf{c}$.

4.3. Algorithm for computer programming

We start the recursion with reasonable initial values $\widehat{v}_i^{(0)}$, $\widehat{m}_k^{(0)}$, $(\widehat{s}_k^{(0)})^2$ and $(\widehat{\varepsilon}^{(0)})^2$ and proceed as described in Section 4.1: We use the formulas for $(\widehat{v}_i^*)^{-1}$, $\widehat{m}_k^*$, $(\widehat{s}_k^*)^2$ and $(\widehat{\varepsilon}^*)^2$ from Section 4.2 and replace the unknown parameters v_i^{-1}, m_k, s_k^2 and ε^2 with the estimators $(\widehat{v}_i^{(\nu)})^{-1}$, $\widehat{m}_k^{(\nu)}$, $(\widehat{s}_k^{(\nu)})^2$ and $(\widehat{\varepsilon}^{(\nu)})^2$ to obtain $(\widehat{v}_i^{(\nu+1)})^{-1}$, $\widehat{m}_k^{(\nu+1)}$, $(\widehat{s}_k^{(\nu+1)})^2$ and $(\widehat{\varepsilon}^{(\nu+1)})^2$. In particular, this means that we have to replace $(\widehat{v}_i^*)^{-1} = \Psi_i^v(\boldsymbol{\Theta}, \mathbf{X})$ with $(\widehat{v}_i^{(\nu+1)})^{-1} = \Psi_i^v(\widehat{\boldsymbol{\Theta}}^{(\nu)}, \mathbf{X})$ in the formulas for $\widehat{m}_{i,k}^*$, $(\widehat{s}_k^*)^2$ and $(\widehat{\varepsilon}^*)^2$.

Start of the recursion:

We choose an initial estimate $(\widehat{\varepsilon}^{(0)})^2 > 0$. In practice, we have had good experiences with

$$(\widehat{\varepsilon}^{(0)})^2 := \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{1}{c_i^2} \left(\frac{\widehat{C}_{i,n}^{\mathrm{CL}}}{\widehat{C}_{i+1,n}^{\mathrm{CL}}} \cdot \frac{\widehat{v}_{i+1}}{\widehat{v}_i} - 1\right)^2,$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

where $\widehat{C}^{\mathrm{CL}}_{i,n}$ denotes the chain ladder prediction for $C_{i,n} := \sum_{k=1}^{n} S_{i,k}$. Moreover, let

$$
\begin{aligned}
(\widehat{\varepsilon}_i^{(0)})^2 &:= (\widehat{\varepsilon}^{(0)})^2 \cdot c_i^2, \\
\widehat{v}_i^{(0)} &:= \widehat{v}_i, \\
\widehat{m}_k^{(0)} &:= \frac{\sum_{i=1}^{n-k+1} S_{i,k}}{\sum_{i=1}^{n-k+1} \widehat{v}_i}, \\
(\widehat{s}_k^{(0)})^2 &:= \frac{1}{n-k} \sum_{i=1}^{n-k+1} \widehat{v}_i \left(\frac{S_{i,k}}{\widehat{v}_i} - \widehat{m}_k^{(0)} \right)^2 \quad (k = 1, \ldots, n-1), \\
(\widehat{s}_n^{(0)})^2 &:= \min\left((\widehat{s}_{n-1}^{(0)})^4 / (\widehat{s}_{n-2}^{(0)})^2, (\widehat{s}_{n-2}^{(0)})^2 \right).
\end{aligned}
$$

Step $\nu \to \nu + 1$:

We interpret $\gamma_{i,k} = \gamma_{i,k}(\boldsymbol{\Theta})$, $\mathbf{R}_i = \mathbf{R}_i(\boldsymbol{\Theta})$, $\mathbf{M}_i = \mathbf{M}_i(\boldsymbol{\Theta})$ and $\mathbf{Q}_i = \mathbf{Q}_i(\boldsymbol{\Theta})$ as functions of the unknown parameter vector $\boldsymbol{\Theta}$.

Define $\widehat{\gamma}_{i,k}^{(\nu+1)} := \gamma_{i,k}(\widehat{\boldsymbol{\Theta}}^{(\nu)})$, $\widehat{\mathbf{R}}_i^{(\nu+1)} := \mathbf{R}_i(\widehat{\boldsymbol{\Theta}}^{(\nu)})$, $\widehat{\mathbf{M}}_i^{(\nu+1)} := \mathbf{M}_i(\widehat{\boldsymbol{\Theta}}^{(\nu)})$ and $\widehat{\mathbf{Q}}_i^{(\nu+1)} := \mathbf{Q}_i(\widehat{\boldsymbol{\Theta}}^{(\nu)})$. For $i \in \{1, \ldots, n\}$ and $(j,k) \in \mathcal{I}_i^*$, let

$$
\widehat{r}_{i,(0,0)}^{(\nu+1)} := \frac{1}{\widehat{v}_i} \quad \text{and} \quad \widehat{r}_{i,(j,k)}^{(\nu+1)} := \frac{1}{\widehat{v}_j} \frac{S_{j,k}}{S_{i,k}} (\widehat{\gamma}_{i,k}^{(\nu+1)})^{-1}.
$$

Let $\widehat{\mathbf{g}}_n^{(\nu+1)} := \widehat{g}_{n,1}^{(\nu+1)} := 1$,

$$
\widehat{\mathbf{g}}_i^{(\nu+1)} := (\widehat{g}_{i,1}^{(\nu+1)}, \ldots, \widehat{g}_{i,N_i}^{(\nu+1)}) := \frac{(\widehat{\mathbf{R}}_i^{(\nu+1)})^{-1} \mathbb{1}_{N_i}}{\mathbb{1}_{N_i}^t (\widehat{\mathbf{R}}_i^{(\nu+1)})^{-1} \mathbb{1}_{N_i}} \quad (i < n)
$$

and let

$$
\widehat{g}_{i,(j,k)}^{(\nu+1)} := \widehat{g}_{i,\varphi_i^{-1}(j,k)}^{(\nu+1)}.
$$

Then, we obtain updated volume estimates $\widehat{v}_i^{(\nu+1)}$:

$$
\frac{1}{\widehat{v}_i^{(\nu+1)}} := \sum_{(j,k) \in \mathcal{I}_i} \widehat{g}_{i,(j,k)}^{(\nu+1)} \widehat{r}_{i,(j,k)}^{(\nu+1)}. \tag{10}
$$

We define

$$
\widehat{m}_{i,k}^{(\nu+1)} := \frac{S_{i,k}}{\widehat{v}_i^{(\nu+1)}}
$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

and

$$\widehat{b}_{i,k}^{(\nu+1)} := \widehat{m}_k^{(\nu)}((\widehat{\gamma}_{i,k}^{(\nu+1)})^{-1} - 1) \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l=k}} \widehat{g}_{i,(j,k)}^{(\nu+1)}.$$

Let $\widehat{\mathbf{b}}_k^{(\nu+1)} := (\widehat{b}_{1,k}^{(\nu+1)}, \dots, \widehat{b}_{n-k+1,k}^{(\nu+1)})^t$ and

$$\widehat{\mathbf{d}}_k^{(\nu+1)} := (\widehat{d}_{1,k}^{(\nu+1)}, \dots, \widehat{d}_{n-k+1,k}^{(\nu+1)})^t := \frac{(\widehat{\mathbf{M}}_k^{(\nu+1)} + \widehat{\mathbf{b}}_k^{(\nu+1)}(\widehat{\mathbf{b}}_k^{(\nu+1)})^t)^{-1}\mathbb{1}_{n-k+1}}{\mathbb{1}_{n-k+1}^t(\widehat{\mathbf{M}}_k^{(\nu+1)} + \widehat{\mathbf{b}}_k^{(\nu+1)}(\widehat{\mathbf{b}}_k^{(\nu+1)})^t)^{-1}\mathbb{1}_{n-k+1}}$$

and define

$$\widetilde{m}_k^{(\nu+1)} := \widehat{d}_{1,k}^{(\nu+1)}\widehat{m}_{1,k}^{(\nu+1)} + \cdots + \widehat{d}_{n-k+1,k}^{(\nu+1)}\widehat{m}_{n-k+1,k}^{(\nu+1)}.$$

Moreover, let

$$\widehat{a}_{i,k}^{(\nu+1)} := ((\widehat{\gamma}_{i,k}^{(\nu+1)})^{-1} - 1) \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l=k}} \widehat{g}_{i,(j,k)}^{(\nu+1)}$$

and with $\widehat{\mathbf{a}}_k^{(\nu+1)} := (\widehat{a}_{1,k}^{(\nu+1)}, \dots, \widehat{a}_{n-k+1,k}^{(\nu+1)})^t$

$$\widehat{\mathbf{h}}_k^{(\nu+1)} := (\widehat{h}_{1,k}^{(\nu+1)}, \dots, \widehat{h}_{n-k+1,k}^{(\nu+1)})^t := \frac{\widehat{\mathbf{d}}_k^{(\nu+1)}}{1 + (\widehat{\mathbf{d}}_k^{(\nu+1)})^t\,\widehat{\mathbf{a}}_k^{(\nu+1)}}$$

and

$$\widehat{m}_k^{(\nu+1)} := \widehat{h}_{1,k}^{(\nu+1)}\widehat{m}_{1,k}^{(\nu+1)} + \cdots + \widehat{h}_{n-k+1,k}^{(\nu+1)}\widehat{m}_{n-k+1,k}^{(\nu+1)}.$$

For $k < n$, we define

$$\Gamma_{i,k}^{(\nu+1)} := \frac{(s_k^{(\nu)})^2}{(\mathbf{e}_i^{(n-k+1)} - \widehat{\mathbf{d}}_k^{(\nu+1)})^t(\widehat{\mathbf{M}}_k^{(\nu+1)} + \widehat{\mathbf{b}}_k^{(\nu+1)}(\widehat{\mathbf{b}}_k^{(\nu+1)})^t)(\mathbf{e}_i^{(n-k+1)} - \widehat{\mathbf{d}}_k^{(\nu+1)})}$$

and obtain updated estimators

$$(\widehat{s}_k^{(\nu+1)})^2 := \frac{1}{n-k+1}\sum_{i=1}^{n-k+1}\Gamma_{i,k}^{(\nu+1)} \cdot \left(\frac{S_{i,k}}{\widehat{v}_i^{(\nu+1)}} - \widetilde{m}_k^{(\nu+1)}\right)^2, \tag{11}$$

for s_k^2. For $k = n$, we use

$$(\widehat{s}_n^{(\nu+1)})^2 := \min\left((\widehat{s}_{n-1}^{(\nu+1)})^4/(\widehat{s}_{n-2}^{(\nu+1)})^2, (\widehat{s}_{n-2}^{(\nu+1)})^2\right).$$

Let

$$\begin{aligned}\Lambda_i^{(\nu+1)} :=& (\widehat{v}_i^{(\nu)})^2 \cdot (\mathbf{e}_1^{(N_i)} - \widehat{\mathbf{g}}_i^{(\nu+1)})^t\mathbf{R}_i^{(\nu+1)}(\mathbf{e}_1^{(N_i)} - \widehat{\mathbf{g}}_i^{(\nu+1)}) \\ &- 2\widehat{v}_i^{(\nu)}\widehat{v}_{i+1}^{(\nu)} \cdot (\mathbf{e}_1^{(N_i)} - \widehat{\mathbf{g}}_i^{(\nu+1)})^t\mathbf{Q}_i^{(\nu+1)}(\mathbf{e}_1^{(N_{i+1})} - \widehat{\mathbf{g}}_{i+1}^{(\nu+1)}) \\ &+ (\widehat{v}_{i+1}^{(\nu)})^2 \cdot (\mathbf{e}_1^{(N_{i+1})} - \widehat{\mathbf{g}}_{i+1}^{(\nu+1)})^t\mathbf{R}_{i+1}^{(\nu+1)}(\mathbf{e}_1^{(N_{i+1})} - \widehat{\mathbf{g}}_{i+1}^{(\nu+1)}).\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

With

$$\Delta_i^{(\nu+1)} := \frac{(\widehat{\varepsilon}^{(\nu)})^2}{\Lambda_i^{(\nu+1)}}$$

we define

$$(\widehat{\varepsilon}^{(\nu+1)})^2 := \frac{1}{n-1}\sum_{i=1}^{n-1}\Delta_i^{(\nu+1)} \times \left[\widehat{v}_i^{(\nu)}\left(\frac{1}{\widehat{v}_i} - \frac{1}{\widehat{v}_i^{(\nu+1)}}\right) - \widehat{v}_{i+1}^{(\nu)}\left(\frac{1}{\widehat{v}_{i+1}} - \frac{1}{\widehat{v}_{i+1}^{(\nu+1)}}\right)\right]^2. \quad (12)$$

In case of convergence, we define

$$(\widehat{\varepsilon}^{(\infty)})^2 := \lim_{\nu\to\infty}(\widehat{\varepsilon}^{(\nu)})^2, \qquad (\widehat{\varepsilon}_i^{(\infty)})^2 := c_i^2\cdot(\widehat{\varepsilon}^{(\infty)})^2,$$
$$\widehat{\mathbf{g}}_i^{(\infty)} := \lim_{\nu\to\infty}\widehat{\mathbf{g}}_i^{(\nu)}, \qquad \widehat{v}_i^{(\infty)} := \lim_{\nu\to\infty}\widehat{v}_i^{(\nu)},$$
$$\widehat{\mathbf{h}}_k^{(\infty)} := \lim_{\nu\to\infty}\widehat{\mathbf{h}}_k^{(\nu)}, \qquad \widehat{m}_k^{(\infty)} := \lim_{\nu\to\infty}\widehat{m}_k^{(\nu)} \quad \text{and}$$
$$(\widehat{s}_k^{(\infty)})^2 := \lim_{\nu\to\infty}(\widehat{s}_k^{(\nu)})^2.$$

In the following sections, we use $\widehat{v}_i^{(\infty)}$, $\widehat{m}_k^{(\infty)}$, $(\widehat{\varepsilon}_i^{(\infty)})^2$ and $(\widehat{s}_k^{(\infty)})^2$ as estimators for v_i, m_k, ε_i^2 and s_k^2 and call $\widehat{v}_i^{(\infty)}$ the *revised volume estimates*.

4.4. Potential modifications to the algorithm

The following remarks describe modifications that can be used to tackle potential issues with the recursive parameter estimation in practice.

Remark 4.12. *The estimators resulting from the recursion can be unreliable for triangles with relatively few accident years. The algorithm also struggles with situations where volatility of loss ratios is mainly driven by claims volatility (and not uncertainty of volume estimates). Therefore, we suggest to factor in external information if available. In the following, we describe how to do this and provide some more possibilities for modifications.*

1. *If the recursion yields an implausible estimator $\widehat{\varepsilon}^{(\infty)}$, e.g., $\widehat{\varepsilon}^{(\infty)} = 0$, it makes sense to use an external estimator $\widehat{\varepsilon}_{ext}$ for ε. Such an external estimator can be included using*

$$(\widehat{\varepsilon}^{(\nu+1)})^2 := \widehat{\varepsilon}_{ext}^2 \qquad (\nu \geq 0)$$

instead of equation (12) in the algorithm in Section 4.3. Typically, a reliable external estimator $\widehat{\varepsilon}_{ext}$ is not available, of course. However, it is absolutely adequate to use an educated guess here. Even if the external estimator is not

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

accurate, it will often be more realistic than the assumption $\varepsilon = 0$*, which corresponds to the additive model.*

2. *If the actuary does not wish to pick an external estimator for* ε*, he might at least have an opinion on a lower bound* $\widehat{\varepsilon}_{min} > 0$ *and/or an upper bound* $\widehat{\varepsilon}_{max} > 0$*. For instance, a lower bound* $\widehat{\varepsilon}_{min}$ *can be considered by setting*

$$(\widetilde{\varepsilon}^{(\nu+1)})^2 := \frac{1}{n-1}\sum_{i=1}^{n-1} \Delta_i^{(\nu+1)} \cdot \left[\widehat{v}_i^{(\nu)} \left(\frac{1}{\widehat{v}_i} - \frac{1}{\widehat{v}_i^{(\nu+1)}} \right) - \widehat{v}_{i+1}^{(\nu)} \left(\frac{1}{\widehat{v}_{i+1}} - \frac{1}{\widehat{v}_{i+1}^{(\nu+1)}} \right) \right]^2$$

and

$$\widehat{\varepsilon}^{(\nu+1)} := \max\left[\widehat{\varepsilon}_{min}\,,\, \widetilde{\varepsilon}^{(\nu+1)} \right]$$

instead of equation (12) in the algorithm in Section 4.3.

3. *If the revised volume estimates* $\widehat{v}_i^{(\infty)}$ *deviate substantially from the volume estimates* $\widehat{v}_i$*, then the estimator* $\widehat{\varepsilon}^{(\infty)}$ *will be relatively large and the estimators* $(\widehat{s}_k^{(\infty)})^2$ *are substantially smaller than the estimators* $\widehat{s}_k^2 := (\widehat{s}_k^{(0)})^2$*, which are commonly used in the additive model. If a small external estimator* $\widehat{\varepsilon}_{ext}$ *is used in the recursion, it can happen that the resulting* $\widehat{v}_i^{(\infty)}$ *deviate more from the* $\widehat{v}_i$ *than we would expect, given* $\widehat{\varepsilon}_{ext}$*. We therefore recommend to limit the distance between* $\widehat{v}_i^{(\infty)}$ *and* $\widehat{v}_i$*, if an external estimator* $\widehat{\varepsilon}_{ext}$ *is used. This can be done in the following way: Define* $(\widetilde{\varepsilon}^{(\nu+1)})^2$ *instead of* $(\widehat{\varepsilon}^{(\nu+1)})^2$ *with formula (12) and set* $(\widehat{\varepsilon}^{(\nu+1)})^2 := \widehat{\varepsilon}_{ext}^2$*. After having incremented* ν *by 1, we define* $(\widehat{v}_i^{(\nu+1)})^{-1}$ *instead of* $(\widehat{v}_i^{(\nu+1)})^{-1}$ *by formula (10) and define*

$$\frac{1}{\widehat{v}_i^{(\nu+1)}} := \frac{1}{\widehat{v}_i} + \min(1, \widehat{\varepsilon}_{ext}/\widetilde{\varepsilon}^{(\nu)}) \cdot \left(\frac{1}{\widetilde{v}_i^{(\nu+1)}} - \frac{1}{\widehat{v}_i} \right).$$

This adjustment is used by default in the `GLRM` *package in case of an external estimator* $\widehat{\varepsilon}_{ext}$*.*

4. *If* $\widehat{\varepsilon}^{(\infty)}$ *is large (e.g., due to a large external estimator* $\widehat{\varepsilon}_{ext}$*), then the algorithm tends to reduce the estimator for* s_1^2 *at the cost of higher estimators for* $s_2^2, \ldots, s_n^2$*. It is possible to avoid this effect and to make the estimation of* $s_1^2, \ldots, s_n^2$ *more robust by fixing the ratios between the estimators.*

Let $\mathcal{S} \subseteq \{1, \ldots, n-1\}$ *and let* $\widehat{s}_k^2 := (\widehat{s}_k^{(0)})^2$ *denote the estimators from the loss ratio method with canonical weights. If we replace (11) in the algorithm in Section 4.3 for* $k \in \mathcal{S}$ *with*

$$(\widehat{s}^{(\nu+1)})^2 := \sum_{l \in \mathcal{S}} \frac{n-l}{\widehat{s}_l^2} \frac{1}{n-l+1} \sum_{i=1}^{n-l+1} \Gamma_{i,l}^{(\nu+1)} \left(\frac{S_{i,l}}{\widehat{v}_i^{(\nu+1)}} - \widetilde{m}_l^{(\nu+1)} \right)^2$$

and

$$(\widehat{s}_k^{(\nu+1)})^2 := \frac{\widehat{s}_k^2}{\sum_{l \in \mathcal{S}} n-l} \cdot (\widehat{s}^{(\nu+1)})^2,$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 2

RECURSION WITH VARIOUS START VALUES $\widehat{\varepsilon}^{(0)}$.

$\widehat{\varepsilon}^{(0)}$	$\widehat{\varepsilon}^{(1)}$	$\widehat{\varepsilon}^{(2)}$	$\widehat{\varepsilon}^{(3)}$	$\widehat{\varepsilon}^{(5)}$	$\widehat{\varepsilon}^{(10)}$	$\widehat{\varepsilon}^{(20)}$	$\widehat{\varepsilon}^{(100)}$
0.500%	0.959%	1.518%	1.925%	2.357%	2.614%	2.634%	2.634%
2.000%	2.146%	2.306%	2.433%	2.563%	2.629%	2.634%	2.634%
5.000%	3.306%	2.930%	2.796%	2.686%	2.637%	2.634%	2.634%
10.000%	3.918%	3.110%	2.884%	2.712%	2.639%	2.634%	2.634%

then we obtain estimators $(\widehat{s}_k^{(\infty)})^2$ *with* $(\widehat{s}_k^{(\infty)})^2/(\widehat{s}_l^{(\infty)})^2 = \widehat{s}_k^2/\widehat{s}_l^2$ *for all* $k, l \in \mathcal{S}$. *In the* `GLRM` *package, this modification is used by default with* $\mathcal{S} := \{1, \ldots, n-1\}$.

5. *Similar to* $\widehat{\varepsilon}_{ext}$, *it is possible to use external estimators (or lower and/or upper bounds) for the parameters* s_k^2. *For instance, if we assume a compound Poisson distribution for* $S_{i,1}$, *then an external estimator for* s_1^2 *can be calculated from individual claims data, see Riegel (2015).*

4.5. Example

Example 4.13. *We apply the recursion to the data from Example 3.3. All figures have been calculated using the* `GLRM` *package with default options, i.e., the largest possible* $\mathcal{K}$, $c_i = 1$ *for all* i *and the modifications of the recursion described in Remark 4.12, 3 and 4. Table 2 shows that the recursion converges to* $\widehat{\varepsilon}^{(\infty)} = 2.63\%$ *for a broad range of start values* $\widehat{\varepsilon}^{(0)}$. *Table 3 shows the volume estimates* $\widehat{v}_i$ *and the estimators*

$$\widehat{s}_k^2 := \frac{1}{n-k} \sum_{i=1}^{n-k+1} \widehat{v}_i \left(\frac{S_{i,k}}{\widehat{v}_i} - \widehat{m}_k \right)^2$$

that are normally used in the additive model (cf. Mack (2002)). Moreover, the table provides the estimators $\widehat{v}_i^{(\infty)}$ *and* $(\widehat{s}_k^{(\infty)})^2$ *resulting from the recursion with various external estimators* $\widehat{\varepsilon}_{ext}$ *for* ε *(cf. Remark 4.12, 1). The recursion, without fixing an external estimator for* ε, *leads to the same estimators for the parameters as the recursion with* $\widehat{\varepsilon}_{ext} = 2.63\%$.

Note that a higher $\widehat{\varepsilon}_{ext}$ *tends to result in lower estimators for the variance parameters* $s_1^2, \ldots, s_n^2$.

5. SELECTION OF WEIGHTS FOR THE LOSS RATIO METHOD

In literature, it is typically assumed that the model volumes are known, i.e., $\widehat{v}_i = v_i$ for $i = 1, \ldots, n$. If this is the case, then the canonical weights

$$w_{i,k}^c = \frac{\widehat{v}_i}{\sum_{j=1}^{n-k+1} \widehat{v}_j} = \frac{v_i}{\sum_{j=1}^{n-k+1} v_j}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 3

ESTIMATORS FOR v_μ AND s_μ^2 FROM THE LOSS RATIO METHOD AND FROM THE RECURSION WITH VARIOUS EXTERNAL ESTIMATORS $\widehat{\varepsilon}_{\text{EXT}}$.

	Loss Ratio Method		Recursion With External Estimator					
			$\widehat{\varepsilon}_{\text{ext}} = 1\%$		$\widehat{\varepsilon}_{\text{ext}} = 2.63\%$		$\widehat{\varepsilon}_{\text{ext}} = 5\%$	
μ	$\widehat{v}_\mu$	$\widehat{s}_\mu^2$	$\widehat{v}_\mu^{(\infty)}$	$(\widehat{s}_\mu^{(\infty)})^2$	$\widehat{v}_\mu^{(\infty)}$	$(\widehat{s}_\mu^{(\infty)})^2$	$\widehat{v}_\mu^{(\infty)}$	$(\widehat{s}_\mu^{(\infty)})^2$
1	316.285	6.50%	306.105	5.42%	295.360	4.28%	293.952	4.08%
2	301.927	5.42%	292.890	4.52%	282.017	3.57%	278.749	3.41%
3	271.299	0.28%	265.540	0.23%	260.276	0.19%	259.023	0.18%
4	273.377	0.15%	270.121	0.12%	270.826	0.10%	274.640	0.09%
5	280.869	0.12%	277.632	0.10%	272.822	0.08%	270.379	0.08%
6	311.389	0.22%	308.797	0.18%	302.922	0.14%	299.871	0.14%
7	322.885	0.13%	322.142	0.11%	321.462	0.09%	322.464	0.08%
8	330.848	0.08%	330.848	0.07%	330.848	0.05%	330.848	0.05%

minimize the variance of the estimator

$$\widehat{m}_k^{\mathbf{w}} = \sum_{i=1}^{n-k+1} w_{i,k} \frac{S_{i,k}}{\widehat{v}_i} = \sum_{i=1}^{n-k+1} w_{i,k} \frac{S_{i,k}}{v_i}$$

for m_k among all deterministic weights $w_{i,k}$ with $\sum_{i=1}^{n-k+1} w_{i,k} = 1$. In the extended additive model, we assume that the model volumes are unknown and that the $\widehat{v}_i$ are only estimators for the model volumes v_i. It is not surprising that this leads to different optimal weights. Let

$$\widehat{m}_{i,k} := \frac{S_{i,k}}{\widehat{v}_i}, \quad i + k \leq n + 1,$$

$\widehat{\mathbf{m}}_k := (\widehat{m}_{1,k}, \ldots, \widehat{m}_{n-k+1,k})$ and $\mathbf{\Sigma}_k := \text{Cov}(\widehat{\mathbf{m}}_k)$.

Proposition 5.1. *The matrix $\mathbf{\Sigma}_k$ is positive definite and its entries can be calculated as follows:*

$$\text{Cov}(\widehat{m}_{i,k}\,,\, \widehat{m}_{j,k}) = m_k^2(\beta_{i,j} - 1) + \delta_{i,j} \frac{s_k^2 \beta_{i,i}}{v_i},$$

where $\delta_{k,l}$ denotes the Kronecker symbol, i.e.,

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Proof. We apply Proposition C.1 to $(S_{i,k})_{i+k\le n+1}$ and $(\widehat{v}_1^{-1},\dots,\widehat{v}_{n-1}^{-1})$ to see that $\mathbf{\Sigma}_k$ is positive definite. For the entries of $\mathbf{\Sigma}_k$, we have

$$\begin{aligned}\operatorname{var}(\widehat{m}_{i,k}) &= \mathrm{E}\left(\frac{S_{i,k}^2}{\widehat{v}_i^2}\right) - \mathrm{E}\left(\frac{S_{i,k}}{\widehat{v}_i}\right)^2 = \frac{v_i s_k^2 + v_i^2 m_k^2}{v_i^2}\beta_{i,i} - \frac{v_i^2 m_k^2}{v_i^2} \\ &= m_k^2(\beta_{i,i}-1) + \frac{s_k^2\beta_{i,i}}{v_i}\end{aligned}$$

and for $i \neq j$

$$\begin{aligned}\operatorname{Cov}(\widehat{m}_{i,k},\widehat{m}_{j,k}) &= \mathrm{E}\left(\frac{S_{i,k}S_{j,k}}{\widehat{v}_i\widehat{v}_j}\right) - \mathrm{E}\left(\frac{S_{i,k}}{\widehat{v}_i}\right)\mathrm{E}\left(\frac{S_{j,k}}{\widehat{v}_j}\right) \\ &= \frac{\mathrm{E}(S_{i,k})\,\mathrm{E}(S_{j,k})}{v_i v_j}\beta_{i,j} - \frac{\mathrm{E}(S_{i,k})\,\mathrm{E}(S_{j,k})}{v_i v_j} = m_k^2\left(\beta_{i,j}-1\right).\end{aligned}$$

■

Applying Proposition B.1, we obtain optimal weights

$$\mathbf{w}_k := \frac{\mathbf{\Sigma}_k^{-1}\mathbb{1}_{n-k+1}}{\mathbb{1}_{n-k+1}^t\mathbf{\Sigma}_k^{-1}\mathbb{1}_{n-k+1}}$$

for the estimation of m_k. If the recursion from Section 4 converges, let $\widehat{\mathbf{\Sigma}}_k^{(\infty)}$ be the matrix obtained from $\mathbf{\Sigma}_k$ by replacing ε_i^2 with $(\widehat{\varepsilon}_i^{(\infty)})^2$, s_k^2 with $(\widehat{s}_k^{(\infty)})^2$, m_k with $\widehat{m}_k^{(\infty)}$ and v_i with $\widehat{v}_i^{(\infty)}$. Then, the weights

$$\widehat{\mathbf{w}}_k^{(\infty)} := \frac{(\widehat{\mathbf{\Sigma}}_k^{(\infty)})^{-1}\mathbb{1}_{n-k+1}}{\mathbb{1}_{n-k+1}^t(\widehat{\mathbf{\Sigma}}_k^{(\infty)})^{-1}\mathbb{1}_{n-k+1}}$$

are approximately optimal for the loss ratio method in the extended additive model. If the recursion from Section 4 does not converge, we simply use the canonical weights, i.e., $\widehat{\mathbf{w}}_k^{(\infty)} := \mathbf{w}^c$ in order to get weights for all possible realizations of the observable random variables $\widehat{v}_1,\dots,\widehat{v}_n$ and $S_{i,k}$ with $i+k \le n+1$.

Example 5.2. *We continue with Example 3.3: Based on the recursion without using an external estimator for ε, i.e., with $\widehat{\varepsilon}^{(\infty)} = 2.63\%$, the approximately optimal weights for the loss ratio method in the extended additive model are displayed in Table 4. For the purpose of comparison, Table 5 contains the canonical weights for the loss ratio method.*

In the extended additive model, we give far more weight to the more recent years. This makes sense since in each column the volume estimates of the most recent years are most comparable to the volume estimates of the accident years we wish to predict.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 4

APPROXIMATELY OPTIMAL WEIGHTS FOR THE LOSS RATIO METHOD ($\widehat{\varepsilon}^{(\infty)} = 2.63\%$).

i	$\widehat{w}_{i,1}^{(\infty)}$	$\widehat{w}_{i,2}^{(\infty)}$	$\widehat{w}_{i,3}^{(\infty)}$	$\widehat{w}_{i,4}^{(\infty)}$	$\widehat{w}_{i,5}^{(\infty)}$	$\widehat{w}_{i,6}^{(\infty)}$	$\widehat{w}_{i,7}^{(\infty)}$	$\widehat{w}_{i,8}^{(\infty)}$
1	0.1%	5.4%	11.3%	16.4%	24.2%	34.7%	50.7%	100.0%
2	0.1%	6.0%	11.7%	16.8%	24.1%	33.6%	49.3%	
3	0.3%	7.1%	12.5%	17.7%	24.0%	31.7%		
4	0.8%	10.0%	15.9%	21.9%	27.8%			
5	2.2%	14.2%	20.1%	27.1%				
6	7.0%	22.5%	28.6%					
7	21.7%	34.8%						
8	67.8%							

TABLE 5

CANONICAL WEIGHTS FOR THE LOSS RATIO METHOD.

i	$w_{i,1}^{c}$	$w_{i,2}^{c}$	$w_{i,3}^{c}$	$w_{i,4}^{c}$	$w_{i,5}^{c}$	$w_{i,6}^{c}$	$w_{i,7}^{c}$	$w_{i,8}^{c}$
1	13.1%	15.2%	18.0%	21.9%	27.2%	35.6%	51.2%	100.0%
2	12.5%	14.5%	17.2%	20.9%	26.0%	33.9%	48.8%	
3	11.3%	13.1%	15.5%	18.8%	23.3%	30.5%		
4	11.3%	13.2%	15.6%	18.9%	23.5%			
5	11.7%	13.5%	16.0%	19.5%				
6	12.9%	15.0%	17.7%					
7	13.4%	15.5%						
8	13.7%							

6. PREDICTION UNCERTAINTY OF THE LOSS RATIO METHOD

A formula for the MSEP of the loss ratio method is well-known for the additive model (cf. Mack (2002)). In this section, we estimate the MSEP using the extended additive model from Section 2, which includes uncertainty of volume estimates.

Let $\mathcal{L} := \{(i,k) \mid i,k \in \{2,\dots,n\}, i+k > n+1\}$. For $(i,k) \in \mathcal{L}$, let $\alpha_{i,k} \in [0,1]$ and $\boldsymbol{\alpha} := (\alpha_{i,k})_{(i,k)\in\mathcal{L}}$. Let

$$\widehat{R}_{\boldsymbol{\alpha}}^{\mathbf{w}} := \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} \widehat{S}_{i,k}^{\mathbf{w}} \quad \text{and} \quad R_{\boldsymbol{\alpha}} := \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} S_{i,k}.$$

If $\alpha_{i,k} = 1$ for all $(i,k) \in \mathcal{L}$, then $R_{\boldsymbol{\alpha}}$ is the total reserve for all accident years. If

$$\alpha_{i,k} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

then R_α is the reserve for the accident year j. In this case, we write $\widehat{R}^{\mathbf{w}}_j := \widehat{R}^{\mathbf{w}}_\alpha$ and $R_j := R_\alpha$.

Assume now that the weights $\mathbf{w}$ are deterministic. Actually, we are interested in the conditional MSEP

$$\begin{aligned}\operatorname{mse}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n) &:= \mathrm{E}\left((R_\alpha - \widehat{R}^{\mathbf{w}}_\alpha)^2 \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n\right) \\ &= \operatorname{var}(R_\alpha) + (\widehat{R}^{\mathbf{w}}_\alpha - \mathrm{E}(R_\alpha))^2 \\ &= \operatorname{var}(R_\alpha) + (\widehat{R}^{\mathbf{w}}_\alpha - \mathrm{E}(\widehat{R}^{\mathbf{w}}_\alpha))^2.\end{aligned}$$

Since it is not possible to calculate $(\widehat{R}^{\mathbf{w}}_\alpha - \mathrm{E}(\widehat{R}^{\mathbf{w}}_\alpha))^2$, it is common to use the approximation

$$(\widehat{R}^{\mathbf{w}}_\alpha - \mathrm{E}(\widehat{R}^{\mathbf{w}}_\alpha))^2 \approx \operatorname{var}(\widehat{R}^{\mathbf{w}}_\alpha),$$

which leads to the unconditional MSEP

$$\operatorname{mse}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha) := \mathrm{E}\left((R_\alpha - \widehat{R}^{\mathbf{w}}_\alpha)^2\right) = \operatorname{var}(R_\alpha) + \operatorname{var}(\widehat{R}^{\mathbf{w}}_\alpha)$$

with the *process variance* $\operatorname{var}(R_\alpha)$ and the *squared parameter estimation error* $\operatorname{var}(\widehat{R}^{\mathbf{w}}_\alpha)$.

Theorem 6.1. *If the weights* $\mathbf{w}$ *are deterministic, then we have*

$$\operatorname{mse}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha) = \operatorname{var}(R_\alpha) + \operatorname{var}(\widehat{R}^{\mathbf{w}}_\alpha)$$

with

$$\operatorname{var}(R_\alpha) = \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}^2 v_i s_k^2$$

and

$$\begin{aligned}\operatorname{var}(\widehat{R}^{\mathbf{w}}_\alpha) = &\sum_{(i,k)\in\mathcal{L}} \sum_{(j,l)\in\mathcal{L}} \alpha_{i,k}\alpha_{j,l} v_i v_j \\ &\times \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} w_{\nu,k} m_k \cdot w_{\kappa,l} m_l \cdot \left[\left(\prod_{\xi=\max(\nu,\kappa)}^{\min(i,j)-1} (1+\varepsilon_\xi^2)\right) - 1\right] \\ &+ \sum_{(i,k)\in\mathcal{L}} \sum_{(j,l)\in\mathcal{L}} \delta_{k,l} \cdot \alpha_{i,k}\alpha_{j,l} v_i v_j \sum_{\nu=1}^{n-k+1} w_{\nu,k}^2 \frac{s_k^2}{v_\nu} \left(\prod_{\xi=\nu}^{\min(i,j)-1} (1+\varepsilon_\xi^2)\right),\end{aligned}$$

where $\delta_{k,l}$ *denotes the Kronecker symbol.*

A proof of this theorem is found in Appendix F.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Remark 6.2. *For $\varepsilon_1^2, \dots, \varepsilon_{n-1}^2 = 0$, we have $\widehat{v}_i = v_i$ for all i and the formula for the squared parameter estimation error yields*

$$\operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}}) = \sum_{k=2}^{n} \sum_{i=n-k+2}^{n} \sum_{j=n-k+2}^{n} \alpha_{i,k}\alpha_{j,k} v_i v_j \sum_{\nu=1}^{n-k+1} w_{\nu,k}^2 \frac{s_k^2}{v_\nu}.$$

In particular, if we use the canonical weights

$$w_{i,k} = w_{i,k}^c = \frac{\widehat{v}_i}{\sum_{j=1}^{n-k+1} \widehat{v}_j} = \frac{v_i}{\sum_{j=1}^{n-k+1} v_j},$$

then we obtain the well-known formula

$$\operatorname{var}(\widehat{R}_\alpha) = \sum_{k=2}^{n} \sum_{i=n-k+2}^{n} \sum_{j=n-k+2}^{n} \alpha_{i,k}\alpha_{j,k} \frac{v_i v_j}{\sum_{\nu=1}^{n-k+1} v_\nu} s_k^2$$

from the additive model (cf. Mack (2002)).

If the weights $\mathbf{w}$ are not deterministic but bounded and $\sigma(\mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n)$-measurable, then the predictions $\widehat{R}_\alpha^{\mathbf{w}}$ are square integrable and a similar calculation as above leads to

$$\operatorname{mse}_{R_\alpha}(\widehat{R}_\alpha^{\mathbf{w}}) := \mathrm{E}\left((R_\alpha - \widehat{R}_\alpha^{\mathbf{w}})^2\right) = \operatorname{var}(R_\alpha) + \operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}}) + (\mathrm{E}(\widehat{R}_\alpha^{\mathbf{w}}) - \mathrm{E}(R_\alpha^{\mathbf{w}}))^2.$$

The bias $\mathrm{E}(\widehat{R}_\alpha^{\mathbf{w}}) - \mathrm{E}(R_\alpha^{\mathbf{w}})$ and $\operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}})$ are difficult to calculate in general. For special cases like the canonical weights, linear approximation can be used to obtain results, but the formulas become very lengthy. If the weights $\mathbf{w}$ are not $\sigma(\mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n)$-measurable, then $\operatorname{mse}_{R_\alpha}(\widehat{R}_\alpha^{\mathbf{w}})$ may not exist since $\widehat{R}_\alpha^{\mathbf{w}}$ may not be measurable.

To avoid these problems with non-deterministic weights, we interpret the formula for $\operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}})$ from Theorem 6.1 as function $\varphi(\mathbf{w}) := \operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}})$ of the deterministic weights $\mathbf{w}$. For arbitrary weights $\mathbf{w}$ (which are functions of the observable data), we then define

$$\operatorname{var}^\bullet(\widehat{R}_\alpha^{\mathbf{w}}) := \varphi \circ \mathbf{w}$$

and

$$\operatorname{mse}^\bullet_{R_\alpha}(\widehat{R}_\alpha^{\mathbf{w}}) := \operatorname{var}(R_\alpha) + \operatorname{var}^\bullet(\widehat{R}_\alpha^{\mathbf{w}}).$$

This means that we fix the weights during the calculation of the MSEP and treat them as if they had already been determined before observing the random variables $\widehat{v}_1, \dots, \widehat{v}_n$ and $S_{i,k}$ with $i + k \le n + 1$. With this approach, we neglect the variability of the weights in the calculation of the mean squared error.

Estimator 6.3. *If we replace the unknown parameters v_i, m_k, s_k^2 and ε_i^2 in the formulas of Theorem 6.1 with the estimators $\widehat{v}_i^{(\infty)}$, $\widehat{m}_k^{(\infty)}$, $(\widehat{s}_k^{(\infty)})^2$ and $(\widehat{\varepsilon}_i^{(\infty)})^2$ from*

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Section 4, then we obtain estimators

$$\widehat{\mathrm{mse}}^{\bullet}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha) = \widehat{\mathrm{var}}(R_\alpha) + \widehat{\mathrm{var}}^{\bullet}(\widehat{R}^{\mathbf{w}}_\alpha)$$

for mean squared error, process variance and squared parameter estimation error.

It is not clear whether the approximately optimal weights $\widehat{\mathbf{w}}^{(\infty)}$ from Section 5 are always measurable with respect to $\sigma(\mathcal{D}, \widehat{v}_1, \ldots, \widehat{v}_n)$. This means that the MSEP $\mathrm{mse}_{R_\alpha}(\widehat{R}^{\mathbf{w}^{(\infty)}}_\alpha)$ may not exist. Therefore, we recommend to use a simulation approach to assess the prediction uncertainty including the variability of the weights. The following simulation approach, which considers the variability of non-deterministic weights, is implemented in the `R` package `GLRM`.

Simulation Approach 6.4. *Let* $\mathbf{w}$ *be weights that may be functions of the observable random variables* $\widehat{v}_i$ *and* $S_{i,k}$ *with* $i = 1, \ldots, n$ *and* $i + k \le n + 1$. *Let* M *be the number of simulations. In the* ν*th simulation, the estimated parameters* $\widehat{v}_i^{(\infty)}$, $\widehat{\varepsilon}_i^{(\infty)}$, $\widehat{m}_k^{(\infty)}$ *and* $(\widehat{s}_k^{(\infty)})^2$ *are used to draw a random sample* $\widehat{v}_i[\nu]$ *for each* $\widehat{v}_i$ *and a random sample* $S_{i,k}[\nu]$ *for each* $S_{i,k}$ $(i, k \in \{1, \ldots, n\})$. *At this, we use lognormal distributions for the* $S_{i,k}$ *with* $k \in \mathcal{K}$ *and for the factors* f_i *from Model Assumption 4. For* $k \notin \mathcal{K}$, *we use normal distributions for* $S_{i,k}$. *Let* $\mathbf{w}[\nu]$ *denote the weights calculated from* $\widehat{v}_i[\nu]$ *and* $S_{i,k}[\nu]$ *and let* $\widehat{S}^{\mathbf{w}[\nu]}_{i,k}[\nu]$, $i + k > n + 1$, *be the predictions of the loss ratio method with weights* $\mathbf{w}[\nu]$, *based on the simulated data* $\widehat{v}_i[\nu]$ *and* $S_{i,k}[\nu]$. *With*

$$\widehat{R}^{\mathbf{w}[\nu]}_\alpha[\nu] := \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} \widehat{S}^{\mathbf{w}[\nu]}_{i,k}[\nu] \quad \text{and} \quad R_\alpha[\nu] := \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} S_{i,k}[\nu],$$

we then define

$$simMSE_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha) := \frac{1}{M}\sum_{\nu=1}^{M} (\widehat{R}^{\mathbf{w}[\nu]}_\alpha[\nu] - R_\alpha[\nu])^2.$$

If $\widehat{R}^{\mathbf{w}}_\alpha$ *is square integrable and if the number of simulations* M *is large, then* $simMSE_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha)$ *is a good approximation for* $\mathrm{mse}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha)$.

If we use the weights $\mathbf{w}^{\bullet}$ *obtained from the actually observed data instead of* $\mathbf{w}[\nu]$ *in every simulation and define* $\widehat{R}^{\mathbf{w}^{\bullet}}_\alpha[\nu]$ *in analogy to* $\widehat{R}^{\mathbf{w}[\nu]}_\alpha[\nu]$, *then*

$$simMSE^{\bullet}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha) := \frac{1}{M}\sum_{\nu=1}^{M} (\widehat{R}^{\mathbf{w}^{\bullet}}_\alpha[\nu] - R_\alpha[\nu])^2$$

is a good approximation for $\mathrm{mse}^{\bullet}_{R_\alpha}(\widehat{R}^{\mathbf{w}}_\alpha)$ *if* M *is large.*

Example 6.5. *We continue our numerical example: In Table 6, we show the reserves* $\widehat{R}_i$ *and* $\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}}$ *calculated with canonical weights* $\mathbf{w}^c$ *and the approximately optimal weights* $\widehat{\mathbf{w}}^{(\infty)}$ *from Table 4 (i.e., with* $\widehat{\varepsilon}^{(\infty)} = 2.63\%$*), respectively. In the columns headed "Additive model", we show the standard errors* $\widehat{\mathrm{s.e.}}_{R_i}(\widehat{R}_i) :=$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 6

PREDICTIONS OF THE LOSS RATIO METHOD WITH CANONICAL AND APPROXIMATELY OPTIMAL WEIGHTS AND THEIR STANDARD ERRORS IN THE ADDITIVE AND THE EXTENDED ADDITIVE MODEL.

	Predicted Reserves		Additive Model		Extended Additive Model With $\widehat{\varepsilon}^{(\infty)} = 2.63\%$	
i	$\widehat{R}_i$	$\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}}$	$\widehat{\text{s.e.}}_{R_i}(\widehat{R}_i)$	$\widehat{\text{s.e.}}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})$	$\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i)$	$\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})$
1	0.00	0.00	0.00	0.00	0.00	0.00
2	2.63	2.63	0.68	0.68	0.54	0.54
3	4.66	4.65	0.95	0.95	0.78	0.78
4	7.00	7.00	1.30	1.30	1.10	1.10
5	10.65	10.69	1.48	1.48	1.28	1.28
6	17.41	17.53	1.75	1.75	1.64	1.63
7	26.68	26.94	2.07	2.08	2.19	2.15
8	79.69	81.81	5.02	5.56	5.72	5.37
Total	148.73	151.25	7.71	8.09	8.98	8.42

$\widehat{\text{mse}}_{R_i}(\widehat{R}_i)^{1/2}$ and $\widehat{\text{s.e.}}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}}) := \widehat{\text{mse}}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})^{1/2}$, calculated with the formulas from Remark 6.2. For the approximately optimal weights, we have applied the formulas from Remark 6.2 with the weights $\widehat{\mathbf{w}}^{(\infty)}$. The last two columns contain the estimators $\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i) := \widehat{\text{mse}}^{\bullet}_{R_i}(\widehat{R}_i)^{1/2}$ and $\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}}) := \widehat{\text{mse}}^{\bullet}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})^{1/2}$ for the standard errors in the extended additive model, where we have used the estimators $\widehat{v}_i^{(\infty)}$, $\widehat{\varepsilon}_i^{(\infty)}$, $\widehat{m}_k^{(\infty)}$ and $(\widehat{s}_k^{(\infty)})^2$ resulting from the recursion of Section 4 without fixing an external estimator for ε.

We see that the standard errors derived from the extended additive model are substantially higher than those in the additive model. The reason is that the additive model does not take the uncertainty of the volume estimates into account and therefore underestimates the prediction uncertainty. Moreover, as expected, we obtain a lower standard error for the canonical weights in the additive model, whereas in the extended additive model, the standard error with the weights $\widehat{\mathbf{w}}^{(\infty)}$ is lower. Table 7 contains simulated prediction errors

$$simS.E._{R_i}(\widehat{R}_i^{\mathbf{w}}) := simMSE_{R_i}(\widehat{R}_i^{\mathbf{w}})^{1/2} \quad and \quad simS.E.^{\bullet}_{R_i}(\widehat{R}_i^{\mathbf{w}}) := simMSE^{\bullet}_{R_i}(\widehat{R}_i^{\mathbf{w}})^{1/2}$$

for $\mathbf{w} = \mathbf{w}^c$ and $\mathbf{w} = \mathbf{w}^{(\infty)}$, with and without considering the volatility of the weights, respectively.

We see that the simulated prediction errors with fixed weights $simS.E.^{\bullet}_{R_i}(\widehat{R}_i)$ and $simS.E.^{\bullet}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})$ from Table 7 are very similar to the estimated standard errors $\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i)$ and $\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})$ in Table 6. The simulated prediction errors $simS.E._{R_i}(\widehat{R}_i)$ and $simS.E._{R_i}(\widehat{R}_i^{\widehat{\mathbf{w}}^{(\infty)}})$, which include the volatility of the weights, are surprisingly even slightly lower than the prediction errors with fixed weights.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 7

SIMULATED PREDICTION ERRORS INCLUDING AND EXCLUDING THE VARIABILITY OF THE WEIGHTS (100.000 SIMULATIONS).

i	simS.E.$_{R_i}(\widehat{R}_i)$	simS.E.$_{R_i}(\widehat{R}_i^{\mathfrak{m}^{(\infty)}})$	simS.E.$^{\bullet}_{R_i}(\widehat{R}_i)$	simS.E.$^{\bullet}_{R_i}(\widehat{R}_i^{\mathfrak{m}^{(\infty)}})$
1	0.00	0.00	0.00	0.00
2	0.54	0.54	0.54	0.54
3	0.78	0.78	0.78	0.78
4	1.10	1.09	1.10	1.09
5	1.28	1.28	1.28	1.28
6	1.63	1.62	1.63	1.63
7	2.18	2.13	2.18	2.15
8	5.69	5.31	5.70	5.36
Total	8.93	8.36	8.96	8.43

7. A GENERALIZED LOSS RATIO METHOD

Up to now, we have used the estimators $(\widehat{\varepsilon}_i^{(\infty)})^2$, $(\widehat{s}_k^{(\infty)})^2$, $\widehat{m}_k^{(\infty)}$ and $\widehat{v}_i^{(\infty)}$ from Section 4 in the context of the loss ratio method: in Section 5 for the estimation of approximately optimal weights and in Section 6 for the calculation of the mean squared error (cf. Estimator 6.3).

Now, we go a step further: we use the estimators $\widehat{m}_k^{(\infty)}$ and $\widehat{v}_i^{(\infty)}$ from Section 4 to define a generalization of the loss ratio method.

Definition 7.1. ***(Generalized loss ratio method)*** *If the recursion of Section 4 converges, then for $i+k > n+1$ the predictions*

$$\widehat{S}_{i,k}^{(\infty)} := \widehat{v}_i^{(\infty)} \cdot \widehat{m}_k^{(\infty)} = \widehat{v}_i^{(\infty)} \cdot \sum_{\nu=1}^{n-k+1} \widehat{h}_{\nu,k}^{(\infty)} \frac{S_{\nu,k}}{\widehat{v}_\nu^{(\infty)}}$$

for $S_{i,k}$ are called the predictions of the generalized loss ratio method.

Essentially, the generalized loss ratio method is a loss ratio method, where we replace the volume estimates $\widehat{v}_i$ with the revised volume estimates $\widehat{v}_i^{(\infty)}$. However, the weights $\widehat{h}_{1,k}^{(\infty)}, \ldots, \widehat{h}_{n-k+1,k}^{(\infty)}$ do not always add up to 1. In order to have the full analogy to Definition 3.1, where we require $\sum_{i=1}^{n-k+1} w_{i,k} = 1$, we could alternatively use the weights $\widehat{d}_{i,k}^{(\infty)} := \lim_{\nu\to\infty} \widehat{d}_{i,k}^{(\nu)}$ instead of $\widehat{h}_{i,k}^{(\infty)}$. These weights always add up to 1. In practice, the difference seems to be small. For all examples we have looked at, the results are similar.

Definition 7.1 really is a generalization of the loss ratio method: If we use an external estimator $\widehat{\varepsilon}_{\text{ext}}$ for ε (see Remark 4.12), then the loss ratio method (with canonical weights) is the limiting case for $\widehat{\varepsilon}_{\text{ext}} \to 0$.

In order to obtain predictions $\widehat{S}_{i,k}^{(\infty)}$, $i+k > n+1$, for all possible realizations of the observable random variables $\widehat{v}_1, \ldots, \widehat{v}_n$ and $S_{i,k}$ with $i+k \le n+1$, we

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 8

PREDICTED RESERVES OF LOSS RATIO, GENERALIZED LOSS RATIO AND CHAIN LADDER.

	Loss Ratio Method	Generalized Loss Ratio Method			ChainLadder
		$\widehat{\varepsilon}_{\text{ext}} = 1\%$	$\widehat{\varepsilon}_{\text{ext}} = 2.63\%$	$\widehat{\varepsilon}_{\text{ext}} = 5\%$	
i	$\widehat{R}_i$	$\widehat{R}_i^{(\infty)}$	$\widehat{R}_i^{(\infty)}$	$\widehat{R}_i^{(\infty)}$	$\widehat{R}_i^{\text{CL}}$
1	0.00	0.00	0.00	0.00	0.00
2	2.63	2.64	2.63	2.61	2.59
3	4.66	4.71	4.78	4.79	4.80
4	7.00	7.14	7.41	7.57	7.77
5	10.65	10.84	10.99	10.94	11.01
6	17.41	17.73	17.90	17.80	17.87
7	26.68	27.28	27.94	28.16	28.56
8	79.69	81.26	83.01	83.32	83.43
Total	148.73	151.59	154.66	155.20	156.02

use the predictions of the loss ratio method, i.e., $\widehat{S}_{i,k}^{(\infty)} := \widehat{S}_{i,k} = \widehat{v}_i \widehat{m}_k$, if the recursion from Section 4 does not converge.

Example 7.2. *We now apply this method to our numerical example. In Table 8, we show the predictions $\widehat{R}_i^{(\infty)}$ of the generalized loss ratio method for a number of external estimators $\widehat{\varepsilon}_{ext}$, in particular also for $\widehat{\varepsilon}_{ext} = 2.63\%$, which is the limit $\widehat{\varepsilon}^{(\infty)}$ if we do not use an external estimator. For the purpose of comparison, we additionally show the predicted reserves $\widehat{R}_i$ for the loss ratio method and $\widehat{R}_i^{CL}$ for the chain ladder method.*

For $\widehat{\varepsilon}_{ext} = 2.63\%$ and $\widehat{\varepsilon}_{ext} = 5\%$, the revised volume estimates $\widehat{v}_i^{(\infty)}$ deviate substantially from the volume estimates $\widehat{v}_i$ (cf. Table 3) and we obtain predictions of the generalized loss ratio method that are closer to the chain ladder predictions than to the predictions of the loss ratio method.

8. PREDICTION UNCERTAINTY OF THE GENERALIZED LOSS RATIO METHOD

Like in Section 6, let $\mathcal{L} := \{(i,k) \mid i,k \in \{2,\dots,n\}, i+k > n+1\}$. For $(i,k) \in \mathcal{L}$, let $\alpha_{i,k} \in [0,1]$ and $\boldsymbol{\alpha} := (\alpha_{i,k})_{(i,k)\in\mathcal{L}}$. We consider the prediction

$$\widehat{R}_{\boldsymbol{\alpha}}^{(\infty)} := \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} \widehat{S}_{i,k}^{(\infty)}$$

for the reserve $R_{\boldsymbol{\alpha}} = \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} S_{i,k}$. We are interested in the accuracy of the prediction $\widehat{R}_{\boldsymbol{\alpha}}^{(\infty)}$ given the observed data. Typically, the prediction uncertainty is measured by the conditional MSEP

$$\operatorname{mse}_{R_{\boldsymbol{\alpha}}}(\widehat{R}_{\boldsymbol{\alpha}}^{(\infty)} \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n) := \mathrm{E}\left((R_{\boldsymbol{\alpha}} - \widehat{R}_{\boldsymbol{\alpha}}^{(\infty)})^2 \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n\right).$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

However, we have a similar issue like in Section 6, where we discussed the prediction uncertainty of the loss ratio method with non-deterministic weights: This expression is only defined if $\widehat{R}_\alpha^{(\infty)}$ is a random variable. However, we cannot guarantee that $\widehat{R}_\alpha^{(\infty)}$ is measurable.

Let us assume for a moment that $\widehat{R}_\alpha^{(\infty)}$ is a random variable. Then, we have

$$\begin{aligned}\operatorname{mse}_{R_\alpha}(\widehat{R}_\alpha^{(\infty)} \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n) &= \operatorname{var}(R_\alpha \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n) + (\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2 \\ &= \operatorname{var}(R_\alpha) + (\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2.\end{aligned}$$

This leads us to the alternative definition

$$\operatorname{mse}'_{R_\alpha}(\widehat{R}_\alpha^{(\infty)} \mid \mathcal{D}, \widehat{v}_1, \dots, \widehat{v}_n) := \operatorname{var}(R_\alpha) + (\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2$$

of the conditional MSEP, which is defined even if $\widehat{R}_\alpha^{(\infty)}$ is not measurable. The question is: How to approximate $(\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2$? Obviously, the approach

$$(\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2 \approx \mathrm{E}((\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2)$$

is not possible. In analogy to Section 6, we use the approximation

$$(\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2 \approx \mathrm{E}^\bullet((\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2)$$

where $\mathrm{E}^\bullet$ means that we treat the weights $\widehat{\mathbf{g}}_i^{(\infty)}$ and $\widehat{\mathbf{h}}_i^{(\infty)}$ as constants when we calculate the expectation. A formal definition of $\mathrm{E}^\bullet$ can be obtained from the author. Note that $\mathrm{E}^\bullet((\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2)$ depends on the observed data (just like $(\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2$). Overall, this means that we measure the accuracy of the prediction with the MSEP

$$\begin{aligned}\operatorname{mse}^\bullet_{R_\alpha}(\widehat{R}_\alpha^{(\infty)}) &:= \operatorname{var}(R_\alpha) + \mathrm{E}^\bullet((\widehat{R}_\alpha^{(\infty)} - \mathrm{E}(R_\alpha))^2) \\ &= \operatorname{var}(R_\alpha) + (\mathrm{E}^\bullet(\widehat{R}_\alpha^{(\infty)}) - \mathrm{E}(R_\alpha))^2 + \operatorname{var}^\bullet(\widehat{R}_\alpha^{(\infty)}).\end{aligned}$$

With this approach, we neglect the fact that the weights $\widehat{\mathbf{g}}_i^{(\infty)}$ and $\widehat{\mathbf{h}}_k^{(\infty)}$ depend on the observed data, i.e., we calculate the prediction error as if these weights had been selected before observing the data $\widehat{v}_1, \dots, \widehat{v}_n$ and $S_{i,k}$ with $i + k \le n + 1$.

The process variance $\operatorname{var}(R_\alpha)$ has been calculated in Section 6 and Appendix F. As in Section 6, we use the estimator

$$\widehat{\operatorname{var}}(R_\alpha) = \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}^2 \widehat{v}_i^{(\infty)} (\widehat{s}_k^{(\infty)})^2.$$

In order to calculate the remaining components of the MSEP, we use the Taylor polynomial of degree $d \ge 1$ to approximate $\widehat{v}_i^{(\infty)}$:

$$\widehat{v}_i^{(\infty)} = v_i \cdot \frac{1}{\frac{v_i}{\widehat{v}_i^{(\infty)}}} \approx v_i \cdot \sum_{\nu=0}^{d} (-1)^\nu \left(\frac{v_i}{\widehat{v}_i^{(\infty)}} - 1\right)^\nu.$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

In practice, it is acceptable to use $d = 1$, i.e., the linear approximation

$$\widehat{v}_i^{(\infty)} \approx v_i \cdot \left(1 - \left(\frac{v_i}{\widehat{v}_i^{(\infty)}} - 1\right)\right) = v_i \cdot \left(2 - \frac{v_i}{\widehat{v}_i^{(\infty)}}\right).$$

The relative error of this linear approximation is below 4% for $\widehat{v}_i^{(\infty)}/v_i \in [0.8, 1.2]$ and below 10% for $\widehat{v}_i^{(\infty)}/v_i \in [0.7, 1.3]$. Of course, the approximation improves with increasing d, in fact

$$\lim_{d\to\infty} v_i \cdot \sum_{\nu=0}^{d} (-1)^\nu \left(\frac{v_i}{\widehat{v}_i^{(\infty)}} - 1\right)^\nu = \widehat{v}_i^{(\infty)}$$

for $\widehat{v}_i^{(\infty)} \in (0, 2v_i)$. However, even for $d = 1$, the following calculations are computationally intensive.

For whole numbers κ with $0 \le \kappa \le d$, let $\omega_{\kappa,d} := \sum_{\nu=\kappa}^{d} (-1)^\kappa \binom{\nu}{\kappa}$. Using the Taylor polynomial of degree $d \ge 1$, we obtain the following.

$$\begin{aligned}
\widehat{S}_{i,k}^{(\infty)} &= \widehat{v}_i^{(\infty)} \cdot \widehat{m}_k^{(\infty)} \\
&\approx v_i \cdot \sum_{\nu=0}^{d} (-1)^\nu \left(\frac{v_i}{\widehat{v}_i^{(\infty)}} - 1\right)^\nu \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \frac{S_{\mu,k}}{\widehat{v}_\mu^{(\infty)}} \\
&= v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \frac{S_{\mu,k}}{\widehat{v}_\mu^{(\infty)}} \cdot \sum_{\nu=0}^{d} (-1)^\nu \sum_{\kappa=0}^{\nu} \binom{\nu}{\kappa} \left(\frac{v_i}{\widehat{v}_i^{(\infty)}}\right)^\kappa (-1)^{\nu-\kappa} \\
&= v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \frac{S_{\mu,k}}{\widehat{v}_\mu^{(\infty)}} \cdot \sum_{\kappa=0}^{d} (-1)^\kappa \left(\frac{v_i}{\widehat{v}_i^{(\infty)}}\right)^\kappa \sum_{\nu=\kappa}^{d} \binom{\nu}{\kappa} \\
&= v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \frac{S_{\mu,k}}{\widehat{v}_\mu^{(\infty)}} \cdot \sum_{\kappa=0}^{d} \omega_{\kappa,d} \left(\frac{v_i}{\widehat{v}_i^{(\infty)}}\right)^\kappa \\
&= v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \cdot \sum_{\kappa=0}^{d} \omega_{\kappa,d} \frac{v_i^\kappa S_{\mu,k}}{\widehat{v}_\mu^{(\infty)} \left(\widehat{v}_i^{(\infty)}\right)^\kappa}.
\end{aligned}$$

Let $\mathcal{J}_i := \{i\} \times \mathcal{I}_i$ and $\mathcal{J} := \bigcup_{i=1}^{n} \mathcal{J}_i$. For $\mathrm{j}_1, \ldots, \mathrm{j}_\kappa \in \mathcal{J}$, we use the abbreviations

$$\widehat{g}_{\mathrm{j}_1,\ldots,\mathrm{j}_\kappa}^{(\infty)} := \widehat{g}_{\mathrm{j}_1}^{(\infty)} \cdots \widehat{g}_{\mathrm{j}_\kappa}^{(\infty)} \quad \text{and} \quad \widehat{r}_{\mathrm{j}_1,\ldots,\mathrm{j}_\kappa} := \widehat{r}_{\mathrm{j}_1} \cdots \widehat{r}_{\mathrm{j}_\kappa}.$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Then, we have

$$\begin{aligned}
\mathrm{E}^\bullet(\widehat{R}_\alpha^{(\infty)}) &= \mathrm{E}^\bullet\left(\sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}\widehat{S}_{i,k}^{(\infty)}\right) = \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}\,\mathrm{E}^\bullet(\widehat{S}_{i,k}^{(\infty)}) \\
&\approx \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}\,\mathrm{E}^\bullet\left(v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \cdot \sum_{\kappa=0}^{d} \omega_{\kappa,d} \frac{v_i^\kappa\, S_{\mu,k}}{\widehat{v}_\mu^{(\infty)}\big(\widehat{v}_i^{(\infty)}\big)^\kappa}\right) \\
&= \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \cdot \sum_{\kappa=0}^{d} \omega_{\kappa,d} v_i^\kappa \cdot \mathrm{E}^\bullet\left(\frac{S_{\mu,k}}{\widehat{v}_\mu^{(\infty)}\big(\widehat{v}_i^{(\infty)}\big)^\kappa}\right) \\
&= \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} v_i \cdot \sum_{\mu=1}^{n-k+1} \widehat{h}_{\mu,k}^{(\infty)} \cdot \sum_{\kappa=0}^{d} \omega_{\kappa,d} v_i^\kappa \cdot \\
&\quad \times \sum_{(\mathrm{j}_1,\ldots,\mathrm{j}_\kappa)\in\mathcal{J}_i^\kappa} \sum_{\mathrm{j}_{\kappa+1}\in\mathcal{J}_\mu} \widehat{g}^{(\infty)}_{\mathrm{j}_1,\ldots,\mathrm{j}_{\kappa+1}}\,\mathrm{E}(S_{i,k}\widehat{r}_{\mathrm{j}_1,\ldots,\mathrm{j}_{\kappa+1}}).
\end{aligned}$$

The expectations $\mathrm{E}(S_{i,k}\widehat{r}_{\mathrm{j}_1,\ldots,\mathrm{j}_{\kappa+1}})$ are calculated in Appendix G, using the additional assumption that the factors f_i from Model Assumption 4 are lognormally distributed. This expectation depends on the unknown parameters ε_i^2, v_i, m_k and s_k^2. If we replace them with their estimators $(\widehat{\varepsilon}_i^{(\infty)})^2$, $\widehat{v}_i^{(\infty)}$, $\widehat{m}_k^{(\infty)}$ and $(\widehat{s}_k^{(\infty)})^2$, we obtain an estimator $\widehat{\mathrm{E}}^\bullet(\widehat{R}_\alpha^{(\infty)})$ for $\mathrm{E}^\bullet(\widehat{R}_\alpha^{(\infty)})$. For

$$\mathrm{E}(R_\alpha) = \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} v_i m_k,$$

we use the estimator

$$\widehat{\mathrm{E}}(R_\alpha) := \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}\widehat{v}_i^{(\infty)}\widehat{m}_k^{(\infty)} = \widehat{R}_\alpha^{(\infty)},$$

which leads to the estimator

$$\widehat{\mathrm{Bias}}^\bullet(\widehat{R}_\alpha^{(\infty)}) := \widehat{\mathrm{E}}^\bullet(\widehat{R}_\alpha^{(\infty)}) - \widehat{R}_\alpha^{(\infty)}$$

for the bias $\mathrm{E}^\bullet(\widehat{R}_\alpha^{(\infty)}) - \mathrm{E}(R_\alpha)$.

For the variance of the prediction, we have

$$\mathrm{var}^\bullet(\widehat{R}_\alpha^{(\infty)}) = \mathrm{var}\left(\sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}\widehat{S}_{i,k}^{(\infty)}\right) = \sum_{(i,k)\in\mathcal{L}} \sum_{(j,l)\in\mathcal{L}} \alpha_{i,k}\alpha_{j,l}\,\mathrm{Cov}^\bullet(\widehat{S}_{i,k}^{(\infty)}, \widehat{S}_{j,l}^{(\infty)})$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

with

$$\begin{aligned}\mathrm{Cov}^{\bullet}(\widehat{S}_{i,k}^{(\infty)}, \widehat{S}_{j,l}^{(\infty)}) &\approx \\ &\approx v_i v_j \cdot \sum_{\mu=1}^{n-k+1} \sum_{\nu=1}^{n-l+1} \widehat{h}_{\mu,k}^{(\infty)} \widehat{h}_{\nu,l}^{(\infty)} \sum_{\kappa=0}^{d} \sum_{\lambda=0}^{d} \omega_{\kappa,d}\omega_{\lambda,d} v_i^{\kappa} v_j^{\lambda} \\ &\times \mathrm{Cov}^{\bullet}\left(\frac{S_{\mu,k}}{\widehat{v}_{\mu}^{(\infty)} \big(\widehat{v}_i^{(\infty)}\big)^{\kappa}}, \frac{S_{\nu,l}}{\widehat{v}_{\nu}^{(\infty)} \big(\widehat{v}_j^{(\infty)}\big)^{\lambda}} \right)\end{aligned}$$

and

$$\begin{aligned}&\mathrm{Cov}^{\bullet}\left(\frac{S_{\mu,k}}{\widehat{v}_{\mu}^{(\infty)} \big(\widehat{v}_i^{(\infty)}\big)^{\kappa}}, \frac{S_{\nu,l}}{\widehat{v}_{\nu}^{(\infty)} \big(\widehat{v}_j^{(\infty)}\big)^{\lambda}} \right) \\ &= \sum_{(\mathfrak{j}_1,\ldots,\mathfrak{j}_\kappa)\in\mathcal{J}_i^{\kappa}} \sum_{\mathfrak{j}_{\kappa+1}\in\mathcal{J}_{\mu}} \sum_{(\mathfrak{k}_1,\ldots,\mathfrak{k}_\kappa)\in\mathcal{J}_j^{\lambda}} \sum_{\mathfrak{k}_{\lambda+1}\in\mathcal{J}_{\nu}} \widehat{g}^{(\infty)}_{\mathfrak{j}_1,\ldots,\mathfrak{j}_{\kappa+1},\mathfrak{k}_1,\ldots,\mathfrak{k}_{\lambda+1}} \\ &\times \mathrm{Cov}(S_{\mu,k}\widehat{r}_{\mathfrak{j}_1,\ldots,\mathfrak{j}_{\kappa+1}}, S_{\nu,l}\widehat{r}_{\mathfrak{k}_1,\ldots,\mathfrak{k}_{\lambda+1}}).\end{aligned}$$

The covariances $\mathrm{Cov}(S_{i,k}\widehat{r}_{\mathfrak{j}_1,\ldots,\mathfrak{j}_{\kappa+1}}, S_{j,l}\widehat{r}_{\mathfrak{k}_1,\ldots,\mathfrak{k}_{\lambda+1}})$ are calculated in Appendix G, using the additional assumption that the factors f_i from Model Assumption 4 are lognormally distributed.

The variance of the prediction $\mathrm{var}^{\bullet}(\widehat{R}_{\alpha}^{(\infty)})$ depends on the unknown parameters ε_i^2, v_i, m_k and s_k^2. If we replace these parameters with their estimators $(\widehat{\varepsilon}_i^{(\infty)})^2$, $\widehat{v}_i^{(\infty)}$, $\widehat{m}_k^{(\infty)}$ and $(\widehat{s}_k^{(\infty)})^2$, we obtain an estimator $\widehat{\mathrm{var}}^{\bullet}(\widehat{R}_{\alpha}^{(\infty)})$. Eventually, we obtain the estimator

$$\widehat{\mathrm{mse}}^{\bullet}_{R_\alpha}(\widehat{R}_{\alpha}^{(\infty)}) := \widehat{\mathrm{var}}(R_\alpha) + \big[\widehat{\mathrm{Bias}}^{\bullet}(\widehat{R}_{\alpha}^{(\infty)})\big]^2 + \widehat{\mathrm{var}}^{\bullet}(\widehat{R}_{\alpha}^{(\infty)})$$

for $\mathrm{mse}^{\bullet}_{R_\alpha}(\widehat{R}_{\alpha}^{(\infty)})$.

Simulation Approach 8.1. *As mentioned above, the calculation of* $\widehat{\mathrm{mse}}^{\bullet}_{R_\alpha}(\widehat{R}_{\alpha}^{(\infty)})$ *is computationally intensive. In practice, we therefore recommend to use the simulation approach for the prediction uncertainty that is implemented in the* `R` *package* `GLRM`*. An additional benefit of this approach is that it is also possible to simulate the prediction uncertainty including the variability of the weights* $\widehat{\mathbf{g}}_i^{(\infty)}$ *and* $\widehat{\mathbf{h}}_i^{(\infty)}$*. The approach works entirely analogously to the Simulation Approach 6.4: If we apply the generalized loss ratio method (including the recursion for parameter estimation and the calculation of the weights* $\widehat{\mathbf{g}}_i^{(\infty)}$ *and* $\widehat{\mathbf{h}}_i^{(\infty)}$*) in every simulation, then we obtain a simulated value*

$$simMSE_{R_\alpha}(\widehat{R}_{\alpha}^{(\infty)})$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

TABLE 9

PREDICTION UNCERTAINTY OF THE GENERALIZED LOSS RATIO METHOD.

	Generalized Loss Ratio Method		
i	$\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})$	$\text{simS.E.}_{R_i}(\widehat{R}_i^{(\infty)})$	$\text{simS.E.}^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})$
1	0.00	0.00	0.00
2	0.54	0.54	0.54
3	0.77	0.77	0.77
4	1.07	1.07	1.07
5	1.21	1.22	1.21
6	1.44	1.48	1.45
7	1.76	1.83	1.76
8	4.39	4.62	4.39
Total	6.63	6.99	6.63

that can be used as an approximation for $\text{mse}'_{R_\alpha}(\widehat{R}_\alpha^{(\infty)} \mid \mathcal{D}, \widehat{v}_1, \ldots, \widehat{v}_n)$. *If we use the same weights* $\widehat{\mathbf{g}}_i^{(\infty)}$ *and* $\widehat{\mathbf{h}}_i^{(\infty)}$, *that have been calculated from the actually observed data, in every simulation, then we obtain a simulated value*

$$simMSE^{\bullet}_{R_\alpha}(\widehat{R}_\alpha^{(\infty)})$$

which is an approximation for $\text{mse}^{\bullet}_{R_\alpha}(\widehat{R}_\alpha^{(\infty)})$.

Example 8.2. *We continue with our numerical example: In Table 9, we show the standard errors* $\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)}) := \widehat{\text{mse}}^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})^{1/2}$ *of the generalized loss ratio method with* $\widehat{\varepsilon}^{(\infty)} = 2.63\%$ *and* $d = 1$.

Moreover, Table 9 contains simulated prediction errors

$$simS.E._{R_i}(\widehat{R}_i^{(\infty)}) := simMSE_{R_i}(\widehat{R}_i^{(\infty)})^{1/2} \quad \text{and}$$
$$simS.E.^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)}) := simMSE^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})^{1/2}$$

with and without considering the volatility of the weights $\widehat{\mathbf{g}}_i^{(\infty)}$ *and* $\widehat{\mathbf{h}}_i^{(\infty)}$, *respectively.*

We see that the simulated prediction errors $simS.E.^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})$ *are very similar to the estimated standard error* $\widehat{\text{s.e.}}^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})$. *The prediction errors* $simS.E._{R_i}(\widehat{R}_i^{(\infty)})$ *are higher than* $simS.E.^{\bullet}_{R_i}(\widehat{R}_i^{(\infty)})$ *but they are still substantially lower than the prediction errors of the loss ratio method in Table 7.*

ACKNOWLEDGEMENTS

The author thanks the two anonymous referees for their valuable comments, which helped to improve the paper substantially.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

References

ASTIN Working Party on Non-Life Reserving (2016) Non-life reserving practices report, http://www.actuaries.org.
CLARK, D.R. (2008) *Reserving with incomplete exposure information*. CAS E-Forum, Fall 2008.
GLUCK, S.M. (1997) *Balancing development and trend in loss reserve analysis*. Conference Paper, 1997 Annual Meeting of the Casualty Actuarial Society.
GÜTSCHOW, T., HESS, K.T. and SCHMIDT, K.D. (2017) Separation of small and large claims on the basis of collective models. *Scandinavian Actuarial Journal*, pp. 1–19. doi: 10.1080/03461238.2017.1394364
HESS, K.T., SCHMIDT, K.D. and ZOCHER, M. (2006) Multivariate loss prediction in the multivariate additive model. *Insurance Mathematics and Economics*, **39**, 185–191.
HÄRDLE, W.K. and HLÁVKA, Z. (2007) *Multivariate Statistics: Exercises and Solutions.* New York: Springer.
JONES, B.D. (2002) *An introduction to premium trend.* CAS Study Note.
KAAS, R., GOOVAERTS, M., DHAENE, J. and DENUIT, M. (2001) *Modern Actuarial Risk Theory.* Boston: Kluwer Academic Publishers.
KORN, U. (2016) *An extension to the Cape Cod method with credibility weighted smoothing*. CAS E-Forum, Summer 2016.
MACK, TH. (2002) *Schadenversicherungsmathematik.* Karlsruhe: VVW.
MARKOWITZ, H. (1952) Portfolio selection. *The Journal of Finance*, **7**(1), 77–91.
MERZ, M. and WÜTHRICH M. V. (2009) Prediction error of the multivariate additive loss reserving method of dependent lines of business. *Variance*, **3**, 131–151.
RIEGEL, U. (2014) A bifurcation approach for attritional and large losses in chain ladder calculations. *ASTIN Bulletin*, **44**(1), 127–172.
RIEGEL, U. (2015) A quantitative study of chain ladder based pricing approaches for long-tail quota shares. *ASTIN Bulletin*, **45**(2), 267–307.
RIEGEL, U. (2016) Bifurcation of attritional and large losses in an additive IBNR environment. *Scandinavian Actuarial Journal*, **2016**(7), 604–623.
SALUZ, A., BÜHLMANN, H., GISLER, A. and MORICONI, F. (2014) Bornhuetter-Ferguson Reserving Method with Repricing. Available at SSRN: https://ssrn.com/abstract=2697167
SALUZ, A., GISLER, A. and WÜTHRICH, M.V. (2011) Development pattern and prediction error for the stochastic Bornhuetter–Ferguson claims reserving method. *ASTIN Bulletin*, **41**(2), 279–313.

ULRICH RIEGEL
Munich Reinsurance Company
Königinstrasse 107
80802 Munich, Germany
E-Mail: uriegel@munichre.com

APPENDIX A. COMMENTS ON TRENDING

Without adjustments to the data, it is often not plausible to assume that a claims triangle and the corresponding premiums fulfil the assumptions of the additive model. Therefore, losses and premiums are typically trended in order to bring all accident years to the same monetary level and to make them comparable. Premium trending is a very complex task (see Jones (2002)) and is never accurate in practice. In this appendix, we will show how Model Assumption 4 in Section 2 is related to uncertainty in premium trending.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Let $p_1^*, \dots, p_n^*$ be the (untrended) premiums and let $S_{i,k}^*$, $i, k \in \{1, \dots, n\}$ denote the incremental claims before trending. For our purpose, it is convenient to consider *accident year trending* that is one of the most popular trending methods in practice. For this trending method, we assume that there exists a claims index $(I_1^S, \dots, I_n^S) \in (0, \infty)^n$ such that the indexed claims

$$S_{i,k}^\bullet := \frac{I_n^S}{I_i^S} \cdot S_{i,k}^*$$

satisfy the assumptions of the additive model. In particular, there exist model volumes $v_1, \dots, v_n$ and parameters $m_1, \dots, m_n$ such that

$$\mathrm{E}(S_{i,k}^\bullet) = v_i \cdot m_k$$

for all i, k. The parameters v_i and m_k are not uniquely determined: For arbitrary $c > 0$, we can use the parameters cv_i and m_k/c instead. Therefore, we can choose the parameters in such a way that $v_n = p_n^*$. Let

$$p_i^\bullet := \frac{I_n^S}{I_i^S} \cdot p_i^*$$

and define the *premium quality index* $(I_1^Q, \dots, I_n^Q)$ by

$$I_i^Q := \frac{p_i^\bullet}{v_i}.$$

Then, we have

$$\frac{I_n^Q}{I_i^Q} \cdot \frac{I_n^S}{I_i^S} \cdot p_i^* = \frac{I_n^Q}{I_i^Q} \cdot p_i^\bullet = v_i$$

for all i. In reality, the indices I_i^S and I_i^Q are not known, of course. The actuary uses estimates $\widehat{I}_i^S$ and $\widehat{I}_i^Q$ instead, i.e., the trended incremental losses and the trended premiums are calculated as

$$S_{i,k} := \frac{\widehat{I}_n^S}{\widehat{I}_i^S} \cdot S_{i,k}^* \quad \text{and} \quad \widehat{v}_i := \frac{\widehat{I}_n^Q}{\widehat{I}_i^Q} \frac{\widehat{I}_n^S}{\widehat{I}_i^S} p_i^*.$$

We now assume that the loss index is known, i.e., $\widehat{I}_i^S = I_i^S$ for all i. This approximation is acceptable since the loss ratios $S_{i,k}/\widehat{v}_i$ do not depend on the estimators $\widehat{I}_i^S$ for the loss index. Then, we have trended losses $S_{i,k} = S_{i,k}^\bullet$ and the trended premiums

$$\widehat{v}_i = \frac{\widehat{I}_n^Q}{\widehat{I}_i^Q} \cdot p_i^\bullet$$

which are used as estimators for the model volumes v_i. Assume that the annual changes of the premium quality index can only be estimated up to independent random errors e_i with $\mathrm{E}(e_i) = 0$ and $\mathrm{var}(e_i) = \varepsilon_i^2$, i.e.,

$$\frac{\widehat{I}_i^Q}{\widehat{I}_{i+1}^Q} = \frac{I_i^Q}{I_{i+1}^Q}(1 + e_i).$$

Let $f_i := 1 + e_i$. Then, we have $\widehat{v}_n = p_n^\bullet = p_n^* = v_n$ and

$$\frac{1}{\widehat{v}_i} = \frac{\widehat{I}_i^Q}{\widehat{I}_n^Q} \cdot \frac{1}{p_i^\bullet} = \frac{I_i^Q}{I_n^Q} \cdot f_i \cdots f_{n-1} \cdot \frac{1}{p_i^\bullet} = \frac{1}{v_i} \cdot f_i \cdots f_{n-1},$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

which is the formula from Model Assumption 4. In particular, we have

$$0 = \operatorname{var}(\widehat{v}_n) < \operatorname{var}(\widehat{v}_{n-1}) < \cdots < \operatorname{var}(\widehat{v}_1).$$

This does not mean that $\widehat{v}_n$ is a better volume measure than $\widehat{v}_1$. In our trending approach, we try to bring all accident years to the level of accident year n. The smaller i, the more years lie between the accident year i and the accident year n, which implies higher uncertainty when we attempt to make the accident year i comparable to the accident year n.

APPENDIX B. MINIMUM MEAN SQUARED ERROR WEIGHTING OF ESTIMATORS

Minimum variance weighting is a fairly common concept that goes back at least to Markowitz's portfolio theory (see Markowitz (1952)). Given a set of unbiased estimators, the variance minimizing weights can be calculated just like the *variance efficient portfolio weights* (see Härdle and Hlávka (2007), Theorem 17.1). The following proposition provides a slightly more general statement for biased estimators:

Proposition B.1. *For $i = 1, \ldots, n$, let T_i be an estimator for a parameter t with bias b_i. Let $\mathbf{T} := (T_1, \ldots, T_n)^t$, $\mathbf{b} := (b_1, \ldots, b_n)^t$ and assume that the covariance matrix*

$$\mathbf{S} := \operatorname{Cov}(\mathbf{T})$$

is positive definite. Then, $\mathbf{S} + \mathbf{b}\mathbf{b}^t$ is positive definite and a linear combination $g_1 T_1 + \cdots + g_n T_n$ with $g_1 + \cdots + g_n = 1$ has minimal mean squared error $\mathrm{E}((\mathbf{g}^t\mathbf{T} - t)^2)$, amongst all such linear combinations, if and only if $\mathbf{g} = (g_1, \ldots, g_n)^t$ satisfies

$$\mathbf{g} = \frac{(\mathbf{S} + \mathbf{b}\mathbf{b}^t)^{-1}\mathbb{1}_n}{\mathbb{1}_n^t(\mathbf{S} + \mathbf{b}\mathbf{b}^t)^{-1}\mathbb{1}_n}.$$

If $\mathbf{e}_i$ denotes the ith unit vector in $\mathbb{R}^n$, then

$$\mathrm{E}((T_i - \mathbf{g}^t\mathbf{T})^2) = (\mathbf{e}_i - \mathbf{g})^t(\mathbf{S} + \mathbf{b}\mathbf{b}^t)(\mathbf{e}_i - \mathbf{g}).$$

The proof of Proposition B.1 can be obtained on request from the author.

Remark B.2. *If T_1 is unbiased and $\operatorname{Cov}(T_1, T_i) = 0$ for $i = 2, \ldots, n$, then*

$$(\mathbf{S} + \mathbf{b}\mathbf{b}^t)^{-1} = \left(\begin{array}{c|c} \operatorname{var}(T_1)^{-1} & 0 \ldots 0 \\ \hline 0 & \\ \vdots & * \\ 0 & \end{array}\right),$$

and we obtain a positive optimal weight

$$g_1 = \frac{\operatorname{var}(T_1)^{-1}}{\mathbb{1}_n^t(\mathbf{S} + \mathbf{b}\mathbf{b}^t)^{-1}\mathbb{1}_n} > 0$$

for T_1.

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

APPENDIX C. CRITERIA FOR POSITIVE DEFINITE COVARIANCE MATRICES

The following two propositions provide criteria that ensure that a covariance matrix is positive definite. These criteria are used in the following appendices and in Section 5.

Proposition C.1. *Let* $\mathbf{X} = (X_1, \dots, X_k)^t$ *and* $\mathbf{Y} = (Y_1, \dots, Y_l)^t$ *be independent random vectors with positive definite covariance matrices. Then, the covariance matrix of*

$$(X_1, \dots, X_k, X_1 Y_1, \dots, X_k Y_1, \dots, X_1 Y_l, \dots X_k Y_l)^t$$

is positive definite.

Proof. We consider a linear combination

$$Z := \sum_{\mu=1}^{k} \lambda_\mu X_\mu + \sum_{\mu=1}^{k} \sum_{\nu=1}^{l} \lambda_{\mu,\nu} X_\mu Y_\nu$$

with $\operatorname{var}(Z) = 0$. Then, we have

$$0 = \operatorname{var}(Z \mid Y_1, \dots, Y_l) = \operatorname{var}\left(\sum_{\mu=1}^{k} X_\mu \cdot \left(\lambda_\mu + \sum_{\nu=1}^{l} \lambda_{\mu,\nu} Y_\nu \right) \,\middle|\, Y_1, \dots, Y_l \right)$$

almost surely. Since the covariance matrix of $\mathbf{X}$ is positive definite and since $\mathbf{X}$ and $\mathbf{Y}$ are independent this implies

$$\lambda_\mu + \sum_{\nu=1}^{l} \lambda_{\mu,\nu} Y_\nu = 0$$

almost surely for $\mu = 1, \dots, k$. This then implies

$$\operatorname{var}\left(\sum_{\nu=1}^{l} \lambda_{\mu,\nu} Y_\nu \right) = 0$$

for $\mu = 1, \dots, k$ and since $\operatorname{Cov}(\mathbf{Y})$ is positive definite, we obtain

$$\lambda_{\mu,\nu} = 0$$

for all $\mu = 1, \dots, k$ and $\nu = 1, \dots, l$. Consequently,

$$Z = \sum_{\mu=1}^{k} \lambda_\mu X_\mu.$$

As $\operatorname{Cov}(\mathbf{X})$ is positive definite, we obtain $\lambda_\mu = 0$ for $\mu = 1, \dots, k$. ■

Proposition C.2. *Let* $X_1, \dots, X_k$ *be independent with positive variances and let* $\mathcal{P} \subseteq \{1, \dots, k\}$ *be a non-empty subset such that* X_μ *is lognormally distributed for all* $\mu \in \mathcal{P}$. *Let*

$$\mathcal{Q} := \{(\mu, \nu) \in \{1, \dots, k\} \times \mathcal{P} \mid \mu \neq \nu\}.$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

Then, the vector

$$\left(\frac{X_\mu}{X_\nu}\right)_{(\mu,\nu)\in\mathcal{Q}}$$

has a positive definite covariance matrix.

Remark C.3. *In Proposition C.2, we require that X_μ is lognormally distributed for $\mu \in \mathcal{P}$ because we will use the proposition for this situation. In the proof, we only use*

$$P(X_\mu = 0) = 0, \quad \operatorname{var}(X_\mu^{-1}) < \infty \quad \textit{and} \quad |\operatorname{Cov}(X_\mu, X_\mu^{-1})| < \sqrt{\operatorname{var}(X_\mu)\operatorname{var}(X_\mu^{-1})}$$

for $\mu \in \mathcal{P}$, which is a weaker requirement.

Proof. Let

$$Z = \sum_{(\mu,\nu)\in\mathcal{Q}} \lambda_{\mu,\nu}\frac{X_\mu}{X_\nu}$$

be a linear combination with $\operatorname{var}(Z) = 0$. Let $\mathcal{S}_i$ be the σ-algebra generated by $\{X_\mu \mid \mu \neq i\}$. Then, we have for $i \in \{1,\dots,k\} \setminus \mathcal{P}$

$$0 = \operatorname{var}(Z \mid \mathcal{S}_i) = \operatorname{var}\left(X_i \sum_{\nu\in\mathcal{P}} \lambda_{i,\nu}\frac{1}{X_\nu} \,\middle|\, \mathcal{S}_i\right) = \left(\sum_{\nu\in\mathcal{P}} \lambda_{i,\nu}\frac{1}{X_\nu}\right)^2 \cdot \operatorname{var}(X_i)$$

almost surely. Since X_ν, $\nu \in \mathcal{P}$, are independent, this implies $\lambda_{i,\nu} = 0$. Let $\mathcal{P}_i := \mathcal{P} \setminus \{i\}$ and $\mathcal{R}_i := \{1,\dots,k\} \setminus \{i\}$. Then, we have for $i \in \mathcal{P}$

$$\begin{aligned}
0 = \operatorname{var}(Z \mid \mathcal{S}_i) &= \operatorname{var}\left(X_i \sum_{\nu\in\mathcal{P}_i} \lambda_{i,\nu}\frac{1}{X_\nu} + \frac{1}{X_i}\sum_{\mu\in\mathcal{R}_i} \lambda_{\mu,i} X_\mu \,\middle|\, \mathcal{S}_i\right) \\
&\geq \left(\sum_{\nu\in\mathcal{P}_i} \lambda_{i,\nu}\frac{1}{X_\nu}\right)^2 \cdot \operatorname{var}(X_i) + \left(\sum_{\nu\in\mathcal{R}_i} \lambda_{\mu,i} X_\mu\right)^2 \cdot \operatorname{var}\left(\frac{1}{X_i}\right) \\
&\quad - 2 \cdot \left|\left(\sum_{\nu\in\mathcal{P}_i} \lambda_{i,\nu}\frac{1}{X_\nu}\right)\left(\sum_{\nu\in\mathcal{R}_i} \lambda_{\mu,i} X_\mu\right)\right| \cdot \left|\operatorname{Cov}\left(X_i, \frac{1}{X_i}\right)\right| \\
&\geq \left(\sum_{\nu\in\mathcal{P}_i} \lambda_{i,\nu}\frac{1}{X_\nu}\right)^2 \cdot \operatorname{var}(X_i) + \left(\sum_{\nu\in\mathcal{R}_i} \lambda_{\mu,i} X_\mu\right)^2 \cdot \operatorname{var}\left(\frac{1}{X_i}\right) \\
&\quad - 2 \cdot \left|\left(\sum_{\nu\in\mathcal{P}_i} \lambda_{i,\nu}\frac{1}{X_\nu}\right)\left(\sum_{\nu\in\mathcal{R}_i} \lambda_{\mu,i} X_\mu\right)\right| \cdot \sqrt{\operatorname{var}(X_i)\operatorname{var}\left(\frac{1}{X_i}\right)} \\
&= \left[\left|\sum_{\nu\in\mathcal{P}_i} \lambda_{i,\nu}\frac{1}{X_\nu}\right| \cdot \sqrt{\operatorname{var}(X_i)} - \left|\sum_{\nu\in\mathcal{R}_i} \lambda_{\mu,i} X_\mu\right| \cdot \sqrt{\operatorname{var}\left(\frac{1}{X_i}\right)}\right]^2 \geq 0
\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

almost surely. We can replace "$\geq$" by "$=$" and obtain

$$\left|\left(\sum_{\nu\in\mathcal{P}_i}\lambda_{i,\nu}\frac{1}{X_\nu}\right)\left(\sum_{\nu\in\mathcal{R}_i}\lambda_{\mu,i}X_\mu\right)\right|\cdot\left|\operatorname{Cov}\left(X_i,\frac{1}{X_i}\right)\right|$$
$$=\left|\left(\sum_{\nu\in\mathcal{P}_i}\lambda_{i,\nu}\frac{1}{X_\nu}\right)\left(\sum_{\nu\in\mathcal{R}_i}\lambda_{\mu,i}X_\mu\right)\right|\cdot\sqrt{\operatorname{var}(X_i)\operatorname{var}\left(\frac{1}{X_i}\right)}$$

almost surely. Since X_i is lognormally distributed, we have $|\operatorname{Cov}(X_i, X_i^{-1})| < \sqrt{\operatorname{var}(X_i)\operatorname{var}(X_i^{-1})}$ and therefore

$$\left(\sum_{\nu\in\mathcal{P}_i}\lambda_{i,\nu}\frac{1}{X_\nu}\right)\left(\sum_{\mu\in\mathcal{R}_i}\lambda_{\mu,i}X_\mu\right)=0$$

almost surely. Consequently,

$$\left(\sum_{\nu\in\mathcal{P}_i}\lambda_{i,\nu}\frac{1}{X_\nu}\right)^2\cdot\operatorname{var}(X_i)+\left(\sum_{\mu\in\mathcal{R}_i}\lambda_{\mu,i}X_\mu\right)^2\cdot\operatorname{var}\left(\frac{1}{X_i}\right)=0$$

and thus

$$\sum_{\nu\in\mathcal{P}_i}\lambda_{i,\nu}\frac{1}{X_\nu}=0\quad\text{and}\quad\sum_{\mu\in\mathcal{R}_i}\lambda_{\mu,i}X_\mu=0$$

almost surely. Since $X_1,\dots,X_k$ are independent, this implies $\lambda_{i,\nu}=0$ and $\lambda_{\mu,i}=0$ for all $\nu\in\mathcal{P}_i$ and $\mu\in\mathcal{R}_i$. ■

APPENDIX D. CALCULATION OF THE COVARIANCE MATRICES $\mathbf{R}_i$ AND $\mathbf{Q}_i$

Let $\mathcal{U}:=\{(i,k)\mid i,k\in\{1,\dots,n\}, i+k\leq n+1\}$, $\mathcal{J}_i:=\{i\}\times\mathcal{I}_i$ and $\mathcal{J}:=\bigcup_{i=1}^n\mathcal{J}_i$. Let $\mathbb{Z}^{\mathcal{U}}$ denote the set of maps $\mathcal{U}\to\mathbb{Z}$. We define

$$\chi^{(1)}\colon\mathcal{J}\to\mathbb{Z}^{\mathcal{U}},\quad \mathfrak{j}\mapsto\chi_{\mathfrak{j}}^{(1)}\quad\text{and}\quad\xi^{(1)}\colon\mathcal{J}\to\mathbb{Z}^{\mathcal{U}},\quad\mathfrak{j}\mapsto\xi_{\mathfrak{j}}^{(1)}$$

by

$$\chi_{i,(j,k)}^{(1)}(\mu,\nu):=\begin{cases}1 & \text{if } (\mu,\nu)=(j,k)\\ -1 & \text{if } (\mu,\nu)=(i,k)\\ 0 & \text{else}\end{cases}$$

and

$$\xi_{i,(j,k)}^{(1)}(\mu,\nu):=\begin{cases}-1 & \text{if } (\mu,\nu)=(i,k)\\ 0 & \text{else.}\end{cases}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

For $\mathrm{j}_1, \dots, \mathrm{j}_k \in \mathcal{J}$, let

$$\chi_{\mathrm{j}_1,\dots,\mathrm{j}_k}(\mu,\nu) := \chi^{(1)}_{\mathrm{j}_1}(\mu,\nu) + \cdots + \chi^{(1)}_{\mathrm{j}_k}(\mu,\nu),$$
$$\xi_{\mathrm{j}_1,\dots,\mathrm{j}_k}(\mu,\nu) := \xi^{(1)}_{\mathrm{j}_1}(\mu,\nu) + \cdots + \xi^{(1)}_{\mathrm{j}_k}(\mu,\nu) \quad \text{and}$$
$$\rho_{\mathrm{j}_1,\dots,\mathrm{j}_k}(\mu,\nu) := \frac{1}{2}\chi_{\mathrm{j}_1,\dots,\mathrm{j}_k}(\mu,\nu)\cdot\Big(\chi_{\mathrm{j}_1,\dots,\mathrm{j}_k}(\mu,\nu)-1\Big) + \xi_{\mathrm{j}_1,\dots,\mathrm{j}_k}(\mu,\nu).$$

Moreover, we consider the maps

$$\pi : \mathcal{J} \to \mathbb{Z}, \quad (i,(j,k)) \mapsto \begin{cases} i & \text{if } (j,k) = (0,0) \\ j & \text{else} \end{cases}$$

and

$$\tau : \mathcal{J} \to \mathbb{Z}, \quad (i,(j,k)) \mapsto i.$$

Note that $\pi(\mathrm{j}) = j$ means that $\widehat{r}_{\mathrm{j}}$ contains the factor $\widehat{v}_j^{-1}$, and $\tau(\mathrm{j}) = i$ means that $\widehat{r}_{\mathrm{j}}$ is an estimator for v_i^{-1}. Subsequently, we use the conventions $0^0 := 1$ and $\gamma^0_{\mu,\nu} := 1$ for $\nu \notin \mathcal{K}$.

Proposition D.1. *For $i = 1, \dots, n-1$, the covariance matrix $\mathbf{R}_i = \mathrm{Cov}(\widehat{\mathbf{r}}_i)$ is positive definite. The matrix $\mathbf{R}_i$ has the entries $\mathrm{Cov}(\widehat{r}_{\mathrm{j}_1}, \widehat{r}_{\mathrm{j}_2})$ with $\mathrm{j}_1, \mathrm{j}_2 \in \mathcal{J}_i$. The matrix $\mathbf{Q}_i$ has the entries $\mathrm{Cov}(\widehat{r}_{\mathrm{j}_1}, \widehat{r}_{\mathrm{j}_2})$ with $\mathrm{j}_1 \in \mathcal{J}_i$, $\mathrm{j}_2 \in \mathcal{J}_{i+1}$. For arbitrary $\mathrm{j}_1, \mathrm{j}_2 \in \mathcal{J}$, we have*

$$\mathrm{Cov}(\widehat{r}_{\mathrm{j}_1}, \widehat{r}_{\mathrm{j}_2}) = \frac{\beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)}}{v_{\tau(\mathrm{j}_1)} v_{\tau(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)} - \frac{1}{v_{\tau(\mathrm{j}_1)} v_{\tau(\mathrm{j}_2)}}.$$

Proof. It follows from Proposition C.2 that the covariance matrix of the random variables

$$\frac{S_{j,k}}{S_{i,k}}, \quad (j,k) \in \mathcal{I}_i^*$$

is positive definite. Applying Proposition C.1 to the random variables $\widehat{v}_1^{-1}, \dots \widehat{v}_{n-1}^{-1}$ and $S_{j,k}/S_{i,k}$, $(j,k) \in \mathcal{I}_i^*$ we see that $\mathbf{R}_i$ is positive definite.

We have

$$\widehat{r}_{\mathrm{j}} = \frac{1}{\widehat{v}_{\pi(\mathrm{j})}} \prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi_{\mathrm{j}}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\xi_{\mathrm{j}}(\mu,\nu)}$$

and therefore

$$\widehat{r}_{\mathrm{j}_1} \cdot \widehat{r}_{\mathrm{j}_2} = \frac{1}{\widehat{v}_{\pi(\mathrm{j}_1)} \widehat{v}_{\pi(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\xi_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)}.$$

With Propositions 2.3 and 2.4, we obtain

$$\begin{aligned}\mathrm{E}(\widehat{r}_{\mathrm{j}_1}\widehat{r}_{\mathrm{j}_2}) &= \frac{\beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)}}{v_{\pi(\mathrm{j}_1)} v_{\pi(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} (v_\mu m_\nu)^{\chi_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)} \\ &= \frac{\beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)}}{v_{\tau(\mathrm{j}_1)} v_{\tau(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)},\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

i.e.,

$$\begin{aligned}\operatorname{Cov}(\widehat{r}_{\mathrm{j}_1}, \widehat{r}_{\mathrm{j}_2}) &= \mathrm{E}(\widehat{r}_{\mathrm{j}_1}\widehat{r}_{\mathrm{j}_2}) - \mathrm{E}(\widehat{r}_{\mathrm{j}_1})\,\mathrm{E}(\widehat{r}_{\mathrm{j}_2}) \\ &= \frac{\beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)}}{v_{\tau(\mathrm{j}_1)}v_{\tau(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)} - \frac{1}{v_{\tau(\mathrm{j}_1)}v_{\tau(\mathrm{j}_2)}}.\end{aligned}$$

■

APPENDIX E. CALCULATION OF THE COVARIANCE MATRIX $\mathbf{M}_i$

We use the notation from Appendix D. For $\mathrm{j}_1, \ldots, \mathrm{j}_k \in \mathcal{J}$ and $(i,k), (j,l) \in \mathcal{U}$ let

$$\chi^{(i,k)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) := \begin{cases} \chi_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) + 1 & \text{if } (\mu,\nu) = (i,k) \\ \chi_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) & \text{else} \end{cases}$$

and

$$\chi^{(i,k),(j,l)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) := \begin{cases} \chi^{(i,k)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) + 1 & \text{if } (\mu,\nu) = (j,l) \\ \chi^{(i,k)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) & \text{else.} \end{cases}$$

Moreover, let

$$\rho^{(i,k)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) := \frac{1}{2}\chi^{(i,k)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu)\cdot\left(\chi^{(i,k)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) - 1\right) + \xi_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu),$$

$$\rho^{(i,k),(j,l)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) := \frac{1}{2}\chi^{(i,k),(j,l)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu)\cdot\left(\chi^{(i,k),(j,l)}_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu) - 1\right) + \xi_{\mathrm{j}_1,\ldots,\mathrm{j}_k}(\mu,\nu).$$

We will now use the notation

$$\gamma_{i,k} := 1 + \frac{v_i s_k^2}{v_i^2 m_k^2}$$

also for $k \notin \mathcal{K}$ with $m_k \neq 0$. If $m_k = 0$, then we read

$$\gamma^0_{i,k} := 1 \quad \text{and} \quad m_k^2\gamma_{i,k} := \lim_{m\to 0} m^2\left(1 + \frac{v_i s_k^2}{v_i^2 m^2}\right) = \frac{s_k^2}{v_i}.$$

We continue to use the convention $0^0 := 1$.

Proposition E.1. *The covariance matrix $\mathbf{M}_k$ is positive definite for $k > 1$. For the entries $\operatorname{Cov}(\widehat{m}^*_{i_1,k}, \widehat{m}^*_{i_2,k})$ of $\mathbf{M}_k$, we have*

$$\operatorname{Cov}(\widehat{m}^*_{i_1,k}, \widehat{m}^*_{i_2,k}) = \sum_{\mathrm{j}_1\in\mathcal{J}_{i_1}}\sum_{\mathrm{j}_2\in\mathcal{J}_{i_2}} g_{\mathrm{j}_1}g_{\mathrm{j}_2}\operatorname{Cov}(S_{i_1,k}\widehat{r}_{\mathrm{j}_1}, S_{i_2,k}\widehat{r}_{\mathrm{j}_2}).$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

For $\mathrm{j}_1 \in \mathcal{J}_{i_1}$ *and* $\mathrm{j}_2 \in \mathcal{J}_{i_2}$, *we have*

$$\operatorname{Cov}(S_{i_1,k}\widehat{r}_{\mathrm{j}_1},\ S_{i_2,k}\widehat{r}_{\mathrm{j}_2}) = \beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)} m_k^2 \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} - m_k^2 \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1}^{(i_1,k)}(\mu,\nu)+\rho_{\mathrm{j}_2}^{(i_2,k)}(\mu,\nu)}.$$

Proof. We first show that that matrix $\mathbf{M}_k$ is positive definite for $k > 1$. Let $k \in \{2, \ldots, n\}$ and let $\lambda_1, \ldots, \lambda_{n-k+1}$ such that

$$\operatorname{var}\left(\sum_{i=1}^{n-k+1} \lambda_i \widehat{m}_{i,k}^*\right) = 0.$$

We have

$$\sum_{i=1}^{n-k+1} \lambda_i \widehat{m}_{i,k}^* = \sum_{i=1}^{n-k+1} \lambda_i \left(g_{i,(0,0)} \frac{S_{i,k}}{\widehat{v}_i} + \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l\neq k}} g_{i,(j,l)} \frac{S_{i,k}}{\widehat{v}_j} \frac{S_{j,l}}{S_{i,l}} \gamma_{i,l}^{-1} + \sum_{\substack{(j,l)\in\mathcal{I}_i^* \\ l=k}} g_{i,(j,k)} \frac{S_{j,k}}{\widehat{v}_j} \gamma_{i,k}^{-1} \right).$$

Applying Propositions C.1 and C.2, we see that the covariance matrix of the random variables

$$\left\{ \frac{S_{i,k}}{\widehat{v}_i},\ \frac{S_{i,k}}{\widehat{v}_j}\frac{S_{j,l}}{S_{i,l}} \,\middle|\, i = 1, \ldots, n-k+1 \text{ and } (j,l) \in \mathcal{I}_i^* \text{ with } l \neq k \right\}$$

is positive definite. In particular, we obtain

$$\lambda_i g_{i,(j,l)} \gamma_{i,l}^{-1} = 0$$

for all $i = 1, \ldots, n-k+1$ and $(j,l) \in \mathcal{I}_i^*$ with $l \neq k$. Therefore, it is sufficient to show that for each $i \in \{1, \ldots, n-k+1\}$ there exists $(j,l) \in \mathcal{I}_i^*$ with $l \neq k$ such that $g_{i,(j,l)} \neq 0$. Let $i \in \{1, \ldots, n-k+1\}$ be arbitrary but fixed. Assume that $g_{i,(j,l)} = 0$ for all $(j,l) \in \mathcal{I}_i^*$ with $l \neq k$. In particular, we then have $g_{i,(j,1)} = 0$ for $j \in \{1, \ldots, n\} \setminus \{i\}$. Consequently, the weights $g_{i,(j,l)}$ are optimal for the estimators

$$\{\widehat{r}_{i,(j,l)} \mid (j,l) \in \mathcal{I}_i^* \text{ with } l \neq 1\} \cup \{\widehat{r}_{i,(n,1)}\}.$$

Note that $\widehat{r}_{i,(n,1)}$ and $\sigma\{\widehat{r}_{i,(j,l)} \mid (j,l) \in \mathcal{I}_i^* \text{ with } l \neq 1\}$ are independent. Therefore, Remark B.2 yields $g_{i,(n,1)} > 0$ that contradicts $g_{i,(j,1)} = 0$ for all $j \in \{1, \ldots, n\} \setminus \{i\}$. Consequently, we have shown $\lambda_i = 0$ for all i, which implies that $\mathbf{M}_k$ is positive definite for $k > 1$.

We have

$$S_{i,k}\widehat{r}_{\mathrm{j}} = \frac{1}{\widehat{v}_{\pi(\mathrm{j})}} \prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi_{\mathrm{j}}^{(i,k)}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\xi_{\mathrm{j}}(\mu,\nu)}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

and therefore

$$S_{i_1,k}\widehat{r}_{\mathrm{j}_1} \cdot S_{i_2,k}\widehat{r}_{\mathrm{j}_2} = \frac{1}{\widehat{v}_{\pi(\mathrm{j}_1)}\widehat{v}_{\pi(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\xi_{\mathrm{j}_1,\mathrm{j}_2}(\mu,\nu)}.$$

With Proposition 2.4, we obtain

$$\mathrm{E}(S_{i,k}\widehat{r}_{\mathrm{j}}) = \mathrm{E}\left(\frac{1}{\widehat{v}_{\pi(\mathrm{j})}}\right) \prod_{(\mu,\nu)\in\mathcal{U}} (v_\mu m_\nu)^{\chi_{\mathrm{j}}^{(i,k)}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\rho_{\mathrm{j}}^{(i,k)}(\mu,\nu)} \quad \text{and}$$

$$\mathrm{E}(S_{i_1,k}\widehat{r}_{\mathrm{j}_1} \cdot S_{i_2,k}\widehat{r}_{\mathrm{j}_2}) = \mathrm{E}\left(\frac{1}{\widehat{v}_{\pi(\mathrm{j}_1)}\widehat{v}_{\pi(\mathrm{j}_2)}}\right) \prod_{(\mu,\nu)\in\mathcal{U}} (v_\mu m_\nu)^{\chi_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)},$$

i.e.,

$$\begin{aligned}
\mathrm{Cov}(S_{i_1,k}\widehat{r}_{\mathrm{j}_1},\ S_{i_2,k}\widehat{r}_{\mathrm{j}_2}) &= \mathrm{E}(S_{i_1,k}\widehat{r}_{\mathrm{j}_1} \cdot S_{i_2,k}\widehat{r}_{\mathrm{j}_2}) - \mathrm{E}(S_{i_1,k}\widehat{r}_{\mathrm{j}_1})\,\mathrm{E}(S_{i_2,k}\widehat{r}_{\mathrm{j}_2}) \\
&= \frac{\beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)}}{v_{\pi(\mathrm{j}_1)}v_{\pi(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} (v_\mu m_\nu)^{\chi_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} \\
&\quad - \frac{1}{v_{\pi(\mathrm{j}_1)}v_{\pi(\mathrm{j}_2)}} \prod_{(\mu,\nu)\in\mathcal{U}} (v_\mu m_\nu)^{\chi_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} \cdot \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1}^{(i_1,k)}(\mu,\nu)+\rho_{\mathrm{j}_2}^{(i_2,k)}(\mu,\nu)} \\
&= \frac{v_{i_1}v_{i_2}}{v_{\tau(\mathrm{j}_1)}v_{\tau(\mathrm{j}_2)}}\beta_{\pi(\mathrm{j}_1),\pi(\mathrm{j}_2)}m_k^2 \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1,\mathrm{j}_2}^{(i_1,k),(i_2,k)}(\mu,\nu)} \\
&\quad - \frac{v_{i_1}v_{i_2}}{v_{\tau(\mathrm{j}_1)}v_{\tau(\mathrm{j}_2)}}m_k^2 \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho_{\mathrm{j}_1}^{(i_1,k)}(\mu,\nu)+\rho_{\mathrm{j}_2}^{(i_2,k)}(\mu,\nu)}.
\end{aligned}$$

Since $\mathrm{j}_1 \in \mathcal{J}_{i_1}$ and $\mathrm{j}_2 \in \mathcal{J}_{i_2}$, we have $\tau(\mathrm{j}_1) = i_1$ and $\tau(\mathrm{j}_2) = i_2$, i.e.,

$$\frac{v_{i_1}v_{i_2}}{v_{\tau(\mathrm{j}_1)}v_{\tau(\mathrm{j}_2)}} = 1.$$

■

APPENDIX F. PROOF OF THEOREM 6.1

Since the increments $S_{i,k}$ are independent, we have

$$\begin{aligned}
\mathrm{var}(R_{\boldsymbol{\alpha}}) &= \mathrm{var}\left(\sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}S_{i,k}\right) = \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}^2\,\mathrm{var}(S_{i,k}) \\
&= \sum_{(i,k)\in\mathcal{L}} \alpha_{i,k}^2 v_i s_k^2.
\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

In the calculation of $\operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}})$, we use

$$\widehat{S}_{i,k}^{\mathbf{w}} = \widehat{v}_i \widehat{m}_k^{\mathbf{w}} = \frac{v_i}{f_i \cdots f_{n-1}} \sum_{\nu=1}^{n-k+1} w_{\nu,k} \frac{S_{\nu,k}}{\widehat{v}_\nu}$$

$$= \frac{v_i}{f_i \cdots f_{n-1}} \sum_{\nu=1}^{n-k+1} w_{\nu,k} \frac{S_{\nu,k}}{v_\nu} f_\nu \cdots f_{n-1} = v_i \sum_{\nu=1}^{n-k+1} w_{\nu,k} \frac{S_{\nu,k}}{v_\nu} f_\nu \cdots f_{i-1}$$

for $(i,k) \in \mathcal{L}$. For $\operatorname{Cov}(\widehat{S}_{i,k}^{\mathbf{w}}, \widehat{S}_{j,l}^{\mathbf{w}})$, we have to distinguish between the two cases $k = l$ and $k \neq l$.

Case $k \neq l$:

$$\operatorname{Cov}(\widehat{S}_{i,k}^{\mathbf{w}}, \widehat{S}_{j,l}^{\mathbf{w}}) = v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} \frac{w_{\nu,k} w_{\kappa,l}}{v_\nu v_\kappa} \operatorname{Cov}(S_{\nu,k} f_\nu \cdots f_{i-1}, S_{\kappa,l} f_\kappa \cdots f_{j-1})$$

$$= v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} w_{\nu,k} w_{\kappa,l} \frac{\operatorname{E}(S_{\nu,k}) \operatorname{E}(S_{\kappa,l})}{v_\nu v_\kappa}$$

$$\times \operatorname{var}(f_{\max(\nu,\kappa)} \cdots f_{\min(i,j)-1})$$

$$= v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} w_{\nu,k} w_{\kappa,l} m_k m_l \operatorname{var}(f_{\max(\nu,\kappa)} \cdots f_{\min(i,j)-1})$$

$$= v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} w_{\nu,k} m_k \cdot w_{\kappa,l} m_l \cdot \left[\left(\prod_{\xi=\max(\nu,\kappa)}^{\min(i,j)-1} (1+\varepsilon_\xi^2) \right) - 1 \right].$$

Case $k = l$:

$$\operatorname{Cov}(\widehat{S}_{i,k}^{\mathbf{w}}, \widehat{S}_{j,l}^{\mathbf{w}}) = v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} \frac{w_{\nu,k} w_{\kappa,l}}{v_\nu v_\kappa} \operatorname{Cov}(S_{\nu,k} f_\nu \cdots f_{i-1}, S_{\kappa,l} f_\kappa \cdots f_{j-1})$$

$$= v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\substack{\kappa=1 \\ \kappa \neq \nu}}^{n-l+1} \frac{w_{\nu,k} w_{\kappa,l}}{v_\nu v_\kappa} \operatorname{Cov}(S_{\nu,k} f_\nu \cdots f_{i-1}, S_{\kappa,l} f_\kappa \cdots f_{j-1})$$

$$+ v_i v_j \sum_{\nu=1}^{n-k+1} w_{\nu,k}^2 \left(\frac{\operatorname{E}(S_{\nu,k}^2)}{v_\nu^2} \operatorname{var}(f_\nu \cdots f_{\min(i,j)-1}) + \frac{\operatorname{var}(S_{\nu,k})}{v_\nu^2} \right)$$

$$= v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\substack{\kappa=1 \\ \kappa \neq \nu}}^{n-l+1} w_{\nu,k} w_{\kappa,l} m_k m_l \operatorname{var}(f_{\max(\nu,\kappa)} \cdots f_{\min(i,j)-1})$$

$$+ v_i v_j \sum_{\nu=1}^{n-k+1} w_{\nu,k}^2 \left(\left(m_k^2 + \frac{s_k^2}{v_\nu} \right) \operatorname{var}(f_\nu \cdots f_{\min(i,j)-1}) + \frac{s_k^2}{v_\nu} \right)$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

$$= v_i v_j \sum_{\nu=1}^{n-k+1} \sum_{\kappa=1}^{n-l+1} w_{\nu,k} m_k \cdot w_{\kappa,l} m_l \cdot \left[\left(\prod_{\xi=\max(\nu,\kappa)}^{\min(i,j)-1} (1+\varepsilon_\xi^2) \right) - 1 \right]$$
$$+ v_i v_j \sum_{\nu=1}^{n-k+1} w_{\nu,k}^2 \frac{s_k^2}{v_\nu} \left(\prod_{\xi=\nu}^{\min(i,j)-1} (1+\varepsilon_\xi^2) \right).$$

Theorem 6.1 now follows from

$$\operatorname{var}(\widehat{R}_\alpha^{\mathbf{w}}) = \operatorname{var}\left(\sum_{(i,k)\in\mathcal{L}} \alpha_{i,k} \widehat{S}_{i,k}^{\mathbf{w}} \right) = \sum_{(i,k)\in\mathcal{L}} \sum_{(j,l)\in\mathcal{L}} \alpha_{i,k}\alpha_{j,l} \operatorname{Cov}(\widehat{S}_{i,k}^{\mathbf{w}}, \widehat{S}_{j,l}^{\mathbf{w}}).$$

APPENDIX G. CALCULATION OF COVARIANCES FOR THE MSEP OF THE GENERALIZED LOSS RATIO METHOD

For $i_1, \dots, i_k \in \{1, \dots, n\}$, we define

$$\beta_{i_1,i_2,\dots,i_k} := \prod_{\kappa=2}^{k} \prod_{\nu=i_{(\kappa)}}^{i_{(\kappa+1)}-1} (1+\varepsilon_\nu^2)^{\frac{1}{2}(\kappa^2-\kappa)},$$

where $i_{(1)} \le \cdots \le i_{(k)} \le i_{(k+1)} := n$ with $\{i_{(1)}, \dots, i_{(k)}\} = \{i_1, \dots, i_k\}$. Empty products are set to 1. Note that for $k = 2$, this definition is consistent with the definition of $\beta_{i,j}$ in Section 2.

Proposition G.1. *Let the factors $f_1, \dots, f_{n-1}$ from Model Assumption 4 be lognormally distributed. Then, we have*

$$\mathrm{E}\left(\frac{1}{\widehat{v}_{i_1}} \cdots \frac{1}{\widehat{v}_{i_k}} \right) = \frac{\beta_{i_1,i_2,\dots,i_k}}{v_{i_1} \cdots v_{i_k}}$$

for $i_1, \dots, i_k \in \{1, \dots, n\}$.

Proof. Similar to Proposition 2.4, it is easy to show that

$$\mathrm{E}(f_i^k) = (1+\varepsilon_i^2)^{\frac{1}{2}(k^2-k)}$$

for all $k \in \mathbb{Z}$. W.l.o.g. let $i_1 \le i_2 \le \cdots \le i_k$. Then,

$$\begin{aligned}
\mathrm{E}\left(\frac{1}{\widehat{v}_{i_1}} \cdots \frac{1}{\widehat{v}_{i_k}} \right) &= \mathrm{E}\left(\frac{1}{v_{i_1}} f_{i_1} \cdots f_{i_2-1} \frac{1}{v_{i_2}} f_{i_2}^2 \cdots f_{i_3-1}^2 \cdots \frac{1}{v_{i_k}} f_{i_k}^k \cdots f_{n-1}^k \right) \\
&= \frac{1}{v_{i_1}} \mathrm{E}(f_{i_1}) \cdots \mathrm{E}(f_{i_2-1}) \frac{1}{v_{i_2}} \mathrm{E}(f_{i_2}^2) \cdots \mathrm{E}(f_{i_3-1}^2) \cdots \\
&\quad \times \frac{1}{v_{i_k}} \mathrm{E}(f_{i_k}^k) \cdots \mathrm{E}(f_{n-1}^k) \\
&= \beta_{i_1,i_2,\dots,i_k} \cdot \frac{1}{v_{i_1}} \cdots \frac{1}{v_{i_k}}.
\end{aligned}$$

■

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

In the following proposition, we use the notation introduced in Appendices D and E.

Proposition G.2. *Let $f_1, \dots, f_{n-1}$ be lognormally distributed and let $\mathfrak{j}_1, \dots, \mathfrak{j}_\kappa, \mathfrak{k}_1, \dots, \mathfrak{k}_\lambda \in \mathcal{J}$. Then,*

$$\mathrm{E}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa}) = \frac{\beta_{\pi(\mathfrak{j}_1),\dots,\pi(\mathfrak{j}_\kappa)}}{v_{\tau(\mathfrak{j}_1)}\cdots v_{\tau(\mathfrak{j}_\kappa)}} v_i m_k \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)}$$

and

$$\begin{aligned}\mathrm{Cov}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa},\ S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_1}) &= \\ = \frac{\beta_{\pi(\mathfrak{j}_1),\dots,\pi(\mathfrak{j}_\kappa),\pi(\mathfrak{k}_1),\dots,\pi(\mathfrak{k}_\lambda)}}{v_{\tau(\mathfrak{j}_1)}\cdots v_{\tau(\mathfrak{j}_\kappa)}v_{\tau(\mathfrak{k}_1)}\cdots v_{\tau(\mathfrak{k}_\lambda)}} v_i m_k v_j m_l \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho^{(i,k),(j,l)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa,\mathfrak{k}_1,\dots,\mathfrak{k}_\lambda}(\mu,\nu)} & \\ - \frac{\beta_{\pi(\mathfrak{j}_1),\dots,\pi(\mathfrak{j}_\kappa)}\beta_{\pi(\mathfrak{k}_1),\dots,\pi(\mathfrak{k}_\lambda)}}{v_{\tau(\mathfrak{j}_1)}\cdots v_{\tau(\mathfrak{j}_\kappa)}v_{\tau(\mathfrak{k}_1)}\dots v_{\tau(\mathfrak{k}_\lambda)}} v_i m_k v_j m_l \prod_{(\mu,\nu)\in\mathcal{U}} \gamma_{\mu,\nu}^{\rho^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)+\rho^{(j,l)}_{\mathfrak{k}_1,\dots,\mathfrak{k}_\lambda}(\mu,\nu)}&.\end{aligned}$$

Proof. We have

$$\begin{aligned}S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa} &= \frac{S_{i,k}}{\widehat{v}_{\pi(\mathfrak{j}_1)}\cdots\widehat{v}_{\pi(\mathfrak{j}_\kappa)}} \prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)}\cdot\gamma_{\mu,\nu}^{\xi_{\mathfrak{j}_1\dots,\mathfrak{j}_\kappa}(\mu,\nu)} \\ &= \frac{1}{\widehat{v}_{\pi(\mathfrak{j}_1)}\cdots\widehat{v}_{\pi(\mathfrak{j}_\kappa)}} \prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)}\cdot\gamma_{\mu,\nu}^{\xi_{\mathfrak{j}_1\dots,\mathfrak{j}_\kappa}(\mu,\nu)}\end{aligned}$$

and therefore

$$\begin{aligned}\mathrm{E}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa}) &= \mathrm{E}\left(\frac{1}{\widehat{v}_{\pi(\mathfrak{j}_1)}\cdots\widehat{v}_{\pi(\mathfrak{j}_\kappa)}}\right)\prod_{(\mu,\nu)\in\mathcal{U}}(v_\mu m_\nu)^{\chi^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)}\cdot\gamma_{\mu,\nu}^{\rho^{(i,k)}_{\mathfrak{j}_1\dots,\mathfrak{j}_\kappa}(\mu,\nu)} \\ &= \frac{\beta_{\pi(\mathfrak{j}_1),\dots,\pi(\mathfrak{j}_\kappa)}}{v_{\pi(\mathfrak{j}_1)}\cdots v_{\pi(\mathfrak{j}_\kappa)}}\prod_{(\mu,\nu)\in\mathcal{U}}(v_\mu m_\nu)^{\chi^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)}\cdot\gamma_{\mu,\nu}^{\rho^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)} \\ &= \frac{\beta_{\pi(\mathfrak{j}_1),\dots,\pi(\mathfrak{j}_\kappa)}}{v_{\tau(\mathfrak{j}_1)}\cdots v_{\tau(\mathfrak{j}_\kappa)}} v_i m_k \prod_{(\mu,\nu)\in\mathcal{U}}\gamma_{\mu,\nu}^{\rho^{(i,k)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa}(\mu,\nu)}.\end{aligned}$$

Similarly, we have

$$\begin{aligned}S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa}\cdot S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_\lambda} &= \\ = \frac{S_{i,k}S_{j,l}}{\widehat{v}_{\pi(\mathfrak{j}_1)}\cdots\widehat{v}_{\pi(\mathfrak{j}_\kappa)}\widehat{v}_{\pi(\mathfrak{k}_1)}\cdots\widehat{v}_{\pi(\mathfrak{k}_\lambda)}} &\prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa,\mathfrak{k}_1,\dots,\mathfrak{k}_\lambda}(\mu,\nu)}\cdot\gamma_{\mu,\nu}^{\xi_{\mathfrak{j}_1\dots,\mathfrak{j}_\kappa,\mathfrak{k}_1\dots,\mathfrak{k}_\lambda}(\mu,\nu)} \\ = \frac{1}{\widehat{v}_{\pi(\mathfrak{j}_1)}\cdots\widehat{v}_{\pi(\mathfrak{j}_\kappa)}\widehat{v}_{\pi(\mathfrak{k}_1)}\cdots\widehat{v}_{\pi(\mathfrak{k}_\lambda)}} &\prod_{(\mu,\nu)\in\mathcal{U}} S_{\mu,\nu}^{\chi^{(i,k),(j,l)}_{\mathfrak{j}_1,\dots,\mathfrak{j}_\kappa,\mathfrak{k}_1,\dots,\mathfrak{k}_\lambda}(\mu,\nu)}\cdot\gamma_{\mu,\nu}^{\xi_{\mathfrak{j}_1\dots,\mathfrak{j}_\kappa,\mathfrak{k}_1\dots,\mathfrak{k}_\lambda}(\mu,\nu)}\end{aligned}$$

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press

and therefore

$$\begin{aligned}
&\mathrm{E}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa}\cdot S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_\lambda}) = \\
&\quad= \frac{\beta_{\pi(\mathfrak{j}_1),\ldots,\pi(\mathfrak{j}_\kappa),\pi(\mathfrak{k}_1),\ldots,\pi(\mathfrak{k}_\lambda)}}{\widehat{v}_{\pi(\mathfrak{j}_1)}\cdots\widehat{v}_{\pi(\mathfrak{j}_\kappa)}\widehat{v}_{\pi(\mathfrak{k}_1)}\cdots\widehat{v}_{\pi(\mathfrak{k}_\lambda)}}\cdot\prod_{(\mu,\nu)\in\mathcal{U}}(v_\mu m_\nu)^{\chi^{(i,k),(j,l)}_{\mathfrak{j}_1,\ldots,\mathfrak{j}_\kappa,\mathfrak{k}_1,\ldots,\mathfrak{k}_\lambda}(\mu,\nu)} \\
&\qquad\cdot\gamma_{\mu,\nu}^{\rho^{(i,k),(j,l)}_{\mathfrak{j}_1,\ldots,\mathfrak{j}_\kappa,\mathfrak{k}_1,\ldots,\mathfrak{k}_\lambda}(\mu,\nu)} \\
&\quad= \frac{\beta_{\pi(\mathfrak{j}_1),\ldots,\pi(\mathfrak{j}_\kappa),\pi(\mathfrak{k}_1),\ldots,\pi(\mathfrak{k}_\lambda)}}{v_{\tau(\mathfrak{j}_1)}\cdots v_{\tau(\mathfrak{j}_\kappa)}v_{\tau(\mathfrak{k}_1)}\cdots v_{\tau(\mathfrak{k}_\lambda)}}v_i m_k v_j m_l\prod_{(\mu,\nu)\in\mathcal{U}}\gamma_{\mu,\nu}^{\rho^{(i,k),(j,l)}_{\mathfrak{j}_1,\ldots,\mathfrak{j}_\kappa,\mathfrak{k}_1,\ldots,\mathfrak{k}_\lambda}(\mu,\nu)}.
\end{aligned}$$

The formula for $\mathrm{Cov}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa},\ S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_1})$ follows with

$$\begin{aligned}
&\mathrm{Cov}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa},\ S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_1}) = \\
&\qquad= \mathrm{E}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa}\cdot S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_1}) - \mathrm{E}(S_{i,k}\widehat{r}_{\mathfrak{j}_1}\cdots\widehat{r}_{\mathfrak{j}_\kappa})\,\mathrm{E}(S_{j,l}\widehat{r}_{\mathfrak{k}_1}\cdots\widehat{r}_{\mathfrak{k}_1}).
\end{aligned}$$

■

https://doi.org/10.1017/asb.2018.5 Published online by Cambridge University Press