

Bootstrap consistency for the Mack bootstrap

Julia Steinmetz*, Carsten Jentsch

Department of Statistics, TU Dortmund University, D-44221 Dortmund, Germany

ARTICLE INFO

JEL classification:

C13
C18
C53
G22

Keywords:

Bootstrap consistency
Loss reserving
Mack's model
Mack bootstrap
Predictive inference

ABSTRACT

Mack's distribution-free chain ladder reserving model belongs to the most popular approaches in non-life insurance mathematics. Proposed to determine the first two moments of the reserve, it does not allow to identify the whole distribution of the reserve. For this purpose, Mack's model is usually equipped with a tailor-made bootstrap procedure. Although widely used in practice to estimate the reserve risk, no theoretical bootstrap consistency results exist that justify this approach.

To fill this gap in the literature, we adopt the framework proposed by Steinmetz and Jentsch (2022) to derive asymptotic theory in Mack's model. By splitting the reserve into two parts corresponding to process and estimation uncertainty, this enables - for the first time - a rigorous investigation also of the validity of the Mack bootstrap. We prove that the (conditional) distribution of the asymptotically dominating process uncertainty part is correctly mimicked by Mack's bootstrap if the parametric family of distributions of the individual development factors is correctly specified. Otherwise, this is not the case. In contrast, the (conditional) distribution of the estimation uncertainty part is generally not correctly captured by Mack's bootstrap. To tackle this, we propose an alternative Mack-type bootstrap, which is designed to capture also the distribution of the estimation uncertainty part.

We illustrate our findings by simulations and show that the newly proposed alternative Mack bootstrap performs superior to the Mack bootstrap.

1. Introduction

In a non-life insurance business an insurer needs to build up a reserve to be able to meet future obligations arising from incurred claims. The actual sizes of the claims are unknown at the time the reserves have to be built, since the claims are incurred, but either not been reported yet or they have been reported, but not settled yet. This process of forecasting of outstanding claims is called *reserving*. An accurate estimation of the outstanding claims is crucial for pricing future policies and for the assessment of the solvency of the insurer. A popular and widely used technique in practice to forecast future claims is the Chain Ladder Model (CLM), which provides an algorithm to predict future claims. In this respect, the most popular model is the recursive model proposed by Mack (1993), which extends the CLM by allowing also the calculation of the standard deviation of the reserve.

Alternatively, frameworks based on general linear models (GLMs) considered e.g. in Renshaw and Verrall (1998) make use of over-dispersed Poisson and Log-normal distributions for modeling mean and variance of the reserve. However, such parametric assumptions are often restrictive and the knowledge of the first two moments of the reserve is not satisfactory for actuaries to draw sufficient conclusions about the reserve risk and the solvency of the insurance company. The *reserve risk* is defined as the risk that the economic-valued reserve does not suffice to pay for all outstanding claims, which inevitably requires the knowledge distribution or at least of high quantiles of the reserve. For this purpose, England and Verrall (2006) proposed the Mack bootstrap which equips Mack's model with a tailor-made bootstrap procedure. Alternative bootstrap procedures for GLM-based setups have been addressed also in England and Verrall (1999; 2006), England (2002) and Pinheiro et al. (2003). Without providing any consistency results, Björkwall et al. (2009) review these bootstrap techniques and suggest alternative non-parametric and parametric bootstrap procedures. Similarly, Björkwall et al. (2010) suggest bootstrap techniques for the separation method, that takes calendar year effects into account. In recent years, bootstrap-based approaches have been favored by many actuaries, because such methods usually produce plausible distributions in

* Corresponding author.

E-mail addresses: steinmetz@statistik.tu-dortmund.de (J. Steinmetz), jentsch@statistik.tu-dortmund.de (C. Jentsch).

<https://doi.org/10.1016/j.insmathco.2024.01.001>

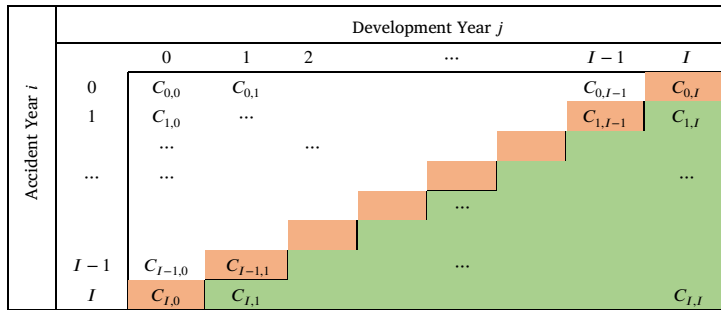
Received March 2023; Received in revised form December 2023; Accepted 3 January 2024

Available online 17 January 2024

0167-6687/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Table 1

Observed upper loss triangle \mathcal{D}_I (upper-left triangle; white and orange) with accident years (vertical axis), development years (horizontal axis), diagonal \mathcal{Q}_I (orange), and unobserved lower loss triangle \mathcal{D}_I^c (lower-right triangle; green). (For interpretation of the references to color please refer to the web version of this article.)



practice. However, as demonstrated by Gibson et al. (2007) and Bruce et al. (2008), Mack’s model and GLM-type models in combination with the bootstrap do not produce satisfactory results in certain situations. In this regard, refined approaches have been proposed to improve the finite sample performance. For example, Verdonck and Debruyne (2011) investigate the influence of outliers for the parameter estimation in the GLM framework and calculate its leverage on the CLM. Hartl (2010) propose to use deviance residuals instead of Pearson residuals for the GLM framework. Tee et al. (2017) provide an extensive case study for bootstrapping the GLM using a (over-dispersed) Poisson model, the Gamma model and the Log-normal model in combination with different residual types. Peremans et al. (2017) propose a robust bootstrap procedure in a GLM setting based on M-estimators using influence functions. Peters et al. (2010) compare the Mack bootstrap with a Bayesian bootstrap.

Nevertheless, already for the original Mack bootstrap method, the existing literature lacks a deeper and mathematically rigorous understanding. For this purpose, it is desirable to provide a suitable theoretical framework to be able to justify the application of the Mack bootstrap. Only recently, Steinmetz and Jentsch (2022) proposed a suitable theoretical (stochastic and asymptotic) framework, which allows the derivation of conditional and unconditional asymptotic theory for the reserve in Mack’s model. They split the reserve (centered around its best estimate) into two parts, that carry the *process uncertainty* and the *estimation uncertainty*, respectively. This allows to derive unconditional limiting distributions for both parts of the reserve, and when conditioning on the latest observed cumulative claims. As risk reserving is generally a prediction task, these *conditional* limiting distributions serve well as benchmarks for the corresponding Mack bootstrap distributions, when addressing the question of bootstrap consistency. While the conditional limiting distribution of the estimation uncertainty part turns out to be Gaussian under mild regularity conditions and when properly inflated, the conditional limiting distribution of the process uncertainty part will be generally non-Gaussian. Considering both parts jointly, the process uncertainty part dominates asymptotically, which leads to a non-Gaussian limiting distribution of the reserve in total.

In this paper, we adopt the theoretical framework introduced in Steinmetz and Jentsch (2022) to investigate the long-standing question of Mack bootstrap consistency. Our contributions are twofold. First, we derive bootstrap asymptotic theory for both parts of the (centered) Mack bootstrap reserve corresponding to process uncertainty and estimation uncertainty, respectively. We prove that the (conditional) bootstrap distribution of the asymptotically dominating process uncertainty part is correctly mimicked if the parametric family of distributions of the Mack bootstrap individual development factors is correctly specified. Otherwise, this will be generally not the case. In contrast, the (conditional) distribution of the estimation uncertainty part is generally not correctly captured. Second, inspired from our asymptotic findings, we propose an alternative Mack-type bootstrap, which is designed to capture also the distribution of the estimation uncertainty part.

The paper is organized as follows. Section 2 introduces the required notation and assumptions for the CLM, discusses parameter estimation in Mack’s model, and provides the asymptotic and stochastic framework of Steinmetz and Jentsch (2022). In Section 3, we discuss the Mack bootstrap approach as proposed by England and Verrall (2006). In Section 4, we summarize the (conditional) asymptotic results from Steinmetz and Jentsch (2022) for the process uncertainty and estimation uncertainty terms in Section 4.1, which will serve as benchmarks for the Mack bootstrap results. Then, in Section 4.2, we derive bootstrap asymptotic theory for both parts of the (centered) Mack bootstrap reserve corresponding to process uncertainty and estimation uncertainty, respectively. Based on these results, we propose an alternative Mack-type bootstrap in Section 5 and derive its asymptotic properties in Section 6. We illustrate our findings in simulations in Section 7 and show that the newly proposed alternative Mack-type bootstrap performs superior to the original Mack bootstrap in finite samples. Section 8 concludes. All proofs, auxiliary results and additional simulations are deferred to the appendix.

2. The chain ladder model

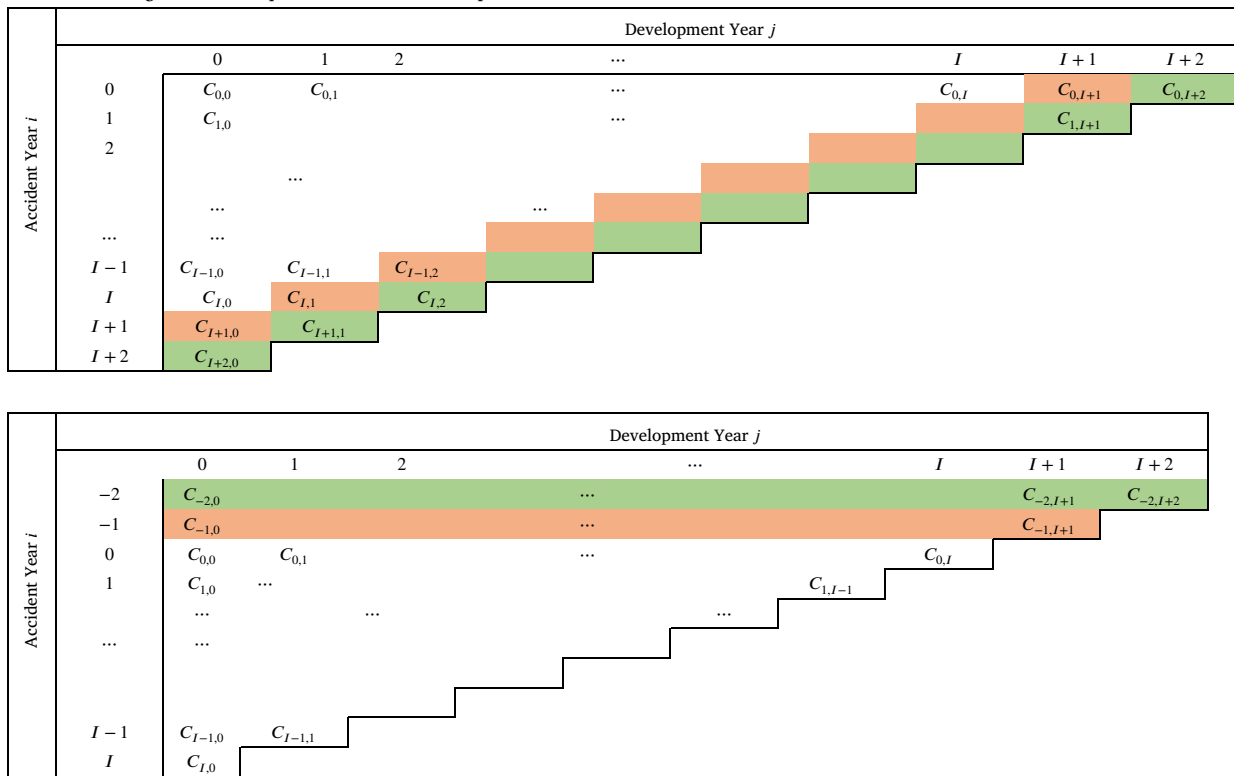
Reserves are the major part of the balance sheet for non-life insurance companies such that their accurate prediction is crucial. For this purpose, insurers summarize all observed claims of a business line in a loss triangle (upper-left triangle in Table 1). Its entries, the cumulative amount of claims $C_{i,j}$, are sorted by their years of accident i (vertical axis) and their years of occurrence j (horizontal axis), where $i, j = 0, \dots, I$ with $i + j \leq I$. Hence, the (observed) loss triangle contains all cumulative claims $C_{i,j}$ that have already been observed up to calendar year I . It constitutes the available data basis and is denoted by

$$D_I = \{C_{i,j} | i, j = 0, \dots, I, 0 \leq i + j \leq I\}. \tag{2.1}$$

The total aggregated amount of claims of the same calendar year k with $k = 0, \dots, I$ are lying on the same diagonal (from lower-left to upper-right corner) of the loss triangle. We denote these diagonals by $\mathcal{Q}_k = \{C_{k-i,i} | i = 0, \dots, k\}$. In this setup, I is the current calendar year corresponding to the most recent accident year and development period such that the diagonal \mathcal{Q}_I (orange diagonal in Table 1) summarizes the latest cumulative claim amounts collected in year I .

Table 2

Two asymptotic frameworks of growing loss triangles based on adding diagonals (upper panel) and by adding rows (lower panel). Both approaches lead to loss triangles that are equal in distributions (adapted from Steinmetz and Jentsch (2022)).



For the theoretical analysis of the prediction of the outstanding (unobserved) claims, it is useful to augment the (observed) *upper* loss triangle D_I by an unobserved *lower* triangle

$$D_I^c = \{C_{i,j} | i, j = 0, \dots, I, i + j > I\}$$

that contains all future claims that have *not* been observed (yet) up to time I (green triangle in Table 1). The resulting cumulative claim matrix is denoted by $C_I = (C_{i,j})_{i,j=0,\dots,I} = D_I \cup D_I^c$. For each accident year i , the main interest lies in the *reserves* at terminal time I , denoted by $R_{i,I}$, which is computed by taking the difference of the ultimate claim $C_{i,I}$ (last column), which is not observed (for $i > 0$) at time I , minus the latest observed claim $C_{i,I-i}$ (on the diagonal) at time I . Precisely, we define the reserve for accident year i by $R_{i,I} = C_{i,I} - C_{i,I-i}$ for $i = 0, \dots, I$ and the aggregated total amount of the reserve R_I by

$$R_I = \sum_{i=0}^I R_{i,I}, \tag{2.2}$$

noting that $R_{0,I} = C_{0,I} - C_{0,I} = 0$ by construction. Hence, for each accident year i and being in calendar year I , to get an estimate of $R_{i,I}$, we have to predict the unobserved ultimate claim $C_{i,I}$. Starting from $C_{i,I-i}$, this is done by predicting sequentially all future, yet (at time I) unobserved claims $\{C_{i,j} | j = I - i + 1, \dots, I\}$. By doing this for all $i = 0, \dots, I$, the whole unobserved lower loss triangle D_I^c has to be predicted, and by summing-up all predictions for $R_{i,I}$, we get a prediction also for R_I .

However, to make the CLM setup above accessible for the derivation of asymptotic theory for predictive inference, Steinmetz and Jentsch (2022) introduced a suitable stochastic and asymptotic framework for Mack’s model, which is adopted here as well and will be described in the following.

2.1. Asymptotic framework for reserve prediction

With the loss triangle D_I at hand, an asymptotic analysis conditional on the diagonal Q_I , which contains the most up-to-date information in the loss triangle, is of much interest for insurers. However, for this purpose, we will not rely on a seemingly “natural” asymptotic framework based on $I \rightarrow \infty$, where increasing I means adding new *diagonals* $Q_{I+h} = \{C_{I-i,i} | i = 0, \dots, I + h\}$, $h \geq 1$ to the loss triangle D_I (see Table 2, upper panel). Instead, as common in predictive inference (see e.g. Paparoditis and Shang (2021)), we employ a different asymptotic framework throughout this paper. That is, we keep the *latest* cumulative claims in D_I , that is, Q_I , fixed and let D_I grow by adding new *rows* of cumulative claims $\{C_{-h,i} | i = 0, \dots, I + h\}$, $h \geq 1$ (see Table 2, lower panel). Nevertheless, both versions of differently growing loss triangles displayed in Table 2 are equal in distribution. In what follows, all asymptotic results are derived under the framework that a sequence of (upper) loss triangles

$$D_{I,n} = \{C_{i,j} | i = -n, \dots, I, j = 0, \dots, I + n, -n \leq i + j \leq I\}, \quad n \in \mathbb{N}_0 = \{0, 1, 2, \dots\}, \tag{2.3}$$

is observed, where

$$Q_{I,n} = \{C_{I-i,i} | i = 0, \dots, I + n\}, \quad n \in \mathbb{N}_0, \tag{2.4}$$

denote the corresponding diagonals. Note that $D_{I,0} = D_I$, $Q_{I,0} = Q_I$ and that $D_{I,n}$ (and $Q_{I,n}$) is obtained by sequentially adding n rows of lengths $I + 2, I + 3, \dots, I + n + 1$, respectively, on top to D_I (see Table 2, lower panel). As before, for all $n \in \mathbb{N}_0$, we augment the (observed) upper loss triangle $D_{I,n}$ by an unobserved lower triangle $D_{I,n}^c = \{C_{i,j} | i = -n, \dots, I, j = 0, \dots, I + n, i + j > I\}$ that contains all future claims that have not been observed (yet) up to time I . Further, according to (2.2), the aggregated total amount of the reserve is denoted by

$$R_{I,n} = \sum_{i=-n}^I R_{i,I+n}, \quad n \in \mathbb{N}_0, \tag{2.5}$$

where $R_{i,I+n} = C_{i,I+n} - C_{i,I-i}$, $n \in \mathbb{N}_0$ and $R_{-n,I+n} = C_{-n,I+n} - C_{-n,I+n} = 0$ by construction.

While we keep I and n fixed in the expositions of the remainder of this section and of Section 3, we let $n \rightarrow \infty$ to derive the limiting distribution of the reserve in Section 4.

2.2. Mack’s distribution-free chain ladder reserving

By adopting the notion of the asymptotic framework described in Section 2.1, the conditions of Mack’s Model originally proposed in Mack (1993) can be summarized as follows.

Assumption 2.1 (Mack’s Model). For any $n \in \mathbb{N}_0$, let $C_{I,n} = (C_{i,j}, i = -n, \dots, I, j = 0, \dots, I + n)$ denote random variables on some probability space (Ω, \mathcal{A}, P) and suppose the following holds:

- (i) There exist so-called development factors f_0, \dots, f_{I+n-1} such that

$$E(C_{i,j+1} | C_{i,j}) = f_j C_{i,j}, \quad i = -n, \dots, I, j = 0, \dots, I + n - 1. \tag{2.6}$$

- (ii) There exist variance parameters $\sigma_0^2, \dots, \sigma_{I+n-1}^2$ such that

$$Var(C_{i,j+1} | C_{i,j}) = \sigma_j^2 C_{i,j}, \quad i = -n, \dots, I, j = 0, \dots, I + n - 1. \tag{2.7}$$

- (iii) The cumulative claims are stochastically independent over the accident years $i = -n, \dots, I$, that is, the cumulative claim matrix $C_{I,n}$ consists of independent rows $C_{i,\bullet} = (C_{i,0}, \dots, C_{i,I+n}), i = -n, \dots, I$.

For any $n \in \mathbb{N}_0$, based on the available data $D_{I,n}$, all development factors f_j and variance parameters σ_j^2 for $j = 0, \dots, I + n - 1$ are unknown and have to be estimated from $D_{I,n}$. The development factors f_0, \dots, f_{I+n-1} can be (consistently) estimated by $\hat{f}_{0,n}, \dots, \hat{f}_{I+n-1,n}$, where

$$\hat{f}_{j,n} = \frac{\sum_{i=-n}^{I-j-1} C_{i,j+1}}{\sum_{i=-n}^{I-j-1} C_{i,j}}, \quad j = 0, \dots, I + n - 1. \tag{2.8}$$

According to Mack (1993), these estimators are unbiased, i.e. $E(\hat{f}_{j,n}) = f_j$, and pairwise uncorrelated, i.e. $Cov(\hat{f}_{j,n}, \hat{f}_{k,n}) = 0$ for all $j \neq k$. By plugging-in the $\hat{f}_{j,n}$ ’s, the best estimate of the ultimate claim $\hat{C}_{i,I+n}$ (point predictor) of the ultimate claim $C_{i,I+n}$ is calculated by

$$\hat{C}_{i,I+n} = C_{i,I-i} \prod_{j=I-i}^{I+n-1} \hat{f}_{j,n}, \quad i = -n, \dots, I.$$

Consequently, given $C_{i,I-i}$, the best estimate $\hat{R}_{i,I+n}$ of the reserve $R_{i,I+n}$ is given by

$$\hat{R}_{i,I+n} = \hat{C}_{i,I+n} - C_{i,I-i} = C_{i,I-i} \left(\prod_{j=I-i}^{I+n-1} \hat{f}_{j,n} - 1 \right), \quad i = -n, \dots, I, \tag{2.9}$$

and the best estimate $\hat{R}_{I,n}$ of the total reserve $R_{I,n}$ defined in (2.5) computes to

$$\hat{R}_{I,n} = \sum_{i=-n}^I \hat{R}_{i,I+n} \tag{2.10}$$

noting that $\hat{R}_{-n,I+n} = 0$ due to $\prod_{j=I+n}^{I+n-1} \hat{f}_{j,n} := 1$. Furthermore, Mack (1993) proposed to estimate the variance parameters $\sigma_0^2, \dots, \sigma_{I+n-1}^2$ by

$$\hat{\sigma}_{j,n}^2 = \frac{1}{I + n - j - 1} \sum_{i=-n}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_{j,n} \right)^2, \quad j = 0, \dots, I + n - 2, \tag{2.11}$$

which are unbiased estimators, i.e. $E(\hat{\sigma}_{j,n}^2) = \sigma_j^2$, and by setting $\hat{\sigma}_{I+n-1,n}^2 = 0$.

Of particular interest is in the distribution of the difference of the stochastic (unobserved) reserve $R_{I,n}$ and its best estimate $\hat{R}_{I,n}$ (based on the observed data $D_{I,n}$), which is denoted as the *predictive root of the reserve* in the following. That is, by combining (2.5) and (2.10), it computes to

$$R_{I,n} - \widehat{R}_{I,n} = \sum_{i=-n}^I (R_{i,I+n} - \widehat{R}_{i,I+n}) = \sum_{i=-n}^I (C_{i,I+n} - \widehat{C}_{i,I+n}). \tag{2.12}$$

While a common approach to approximate an unknown (finite sample) distributions is the derivation of asymptotic theory, Mack’s conditions summarized in Assumption 2.1 are not (yet) sufficient to establish limiting distributions for the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$.

2.3. A fully-described stochastic framework of Mack’s model

Following Steinmetz and Jentsch (2022, Section 2.2), to establish a theoretical framework sufficient to derive asymptotic theory for parameter estimators $\widehat{f}_{j,n}$ and $\widehat{\sigma}_{j,n}^2$, which finally also enables the derivation of the limiting distributions of the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$, we introduce Assumptions 2.2, 2.3, and 2.5 on the stochastic mechanism that generates the cumulative claim matrix $C_{I,n}$, and Assumption 2.4 on the sequences of development factors and of variance parameters. They resemble the Assumptions 2.2, 2.3 and 3.3 as well as Assumption 4.1 in Steinmetz and Jentsch (2022), respectively. This framework will also allow to rigorously investigate consistency properties of the Mack bootstrap in Section 4.

The first assumption addresses the initial claims, i.e. the first column of $C_{I,n}$ (and of $D_{I,n}$).

Assumption 2.2 (Initial claims). Let the initial claims $(C_{I-n,0}, n \in \mathbb{N}_0)$ be independent and identically distributed (i.i.d.) random variables with support $[1, \infty)$, i.e. $C_{i,0} \geq 1$ for all i . Further, let $\mu_0 := E(C_{i,0}) \in [1, \infty)$ and $\tau_0^2 := Var(C_{i,0}) \in (0, \infty)$.

The independence of the initial claims is a direct consequence of Assumption 2.1 (iii). In addition, Assumption 2.2 imposes an identical distribution for the initial claims. The latter assumptions could be relaxed such that all $(C_{I-n,0}, n \in \mathbb{N}_0)$ belong to the same parametric family of distributions, but with possibly different parameters. In this case, additionally, we have to guarantee that both $\frac{1}{I+n+1} \sum_{i=-n}^I E(C_{i,0})$ and $\frac{1}{I+n+1} \sum_{i=-n}^I Var(C_{i,0})$ converge as $n \rightarrow \infty$. However, depending on the resampling scheme to be used in the upcoming Sections 3, 4 and 5, this relaxation may have to be taken into account for valid inference. In practice, the condition on the support $[1, \infty)$ of $C_{i,0}$ is not restrictive and can be relaxed to $C_{i,0}$ being bounded away from zero. While independence of the initial claims $(C_{I-n,0}, n \in \mathbb{N}_0)$, that is, of the first column of the cumulative claim matrix C_I , is usually a very plausible assumption in practice, allowing for dependency will also result in dependency in *all* other columns of C_I as well. Hence, all the theory developed in this paper would require different techniques of proof by relying on CLTs for dependent data, which would also alter the obtained (limiting) distributions.

In view of $E(C_{i,j+1}|C_{i,j})$ in (2.6), suppose that the cumulative claims $C_{i,j+1}$ are recursively defined by

$$C_{i,j+1} = C_{i,j} F_{i,j} = C_{i,0} \prod_{k=0}^j F_{i,k}, \quad i = -n, \dots, I, \quad j = 0, \dots, I + n - 1, \tag{2.13}$$

where the *individual development factors* $F_{i,j}$ are assumed to fulfill the following condition.

Assumption 2.3 (Conditional distribution of the individual development factors). Let the individual development factors $(F_{I-i,j}, i \in \mathbb{N}_0, j \in \mathbb{N}_0)$ be random variables with support (ϵ, ∞) for some $\epsilon \geq 0$ such that $F_{i,j}$ and $F_{k,l}$ are independent given $(C_{i,j}, C_{k,l})$ for all $(i, j) \neq (k, l)$ with

$$E(F_{i,j}|C_{i,j}) = f_j \quad \text{and} \quad Var(F_{i,j}|C_{i,j}) = \frac{\sigma_j^2}{C_{i,j}}. \tag{2.14}$$

Note that Mack’s original model setup in Assumption 2.1 is implied by Assumptions 2.2 and 2.3 together. Also note that the stochastic mechanism determined by (2.13) and Assumption 2.3 are assumed for the whole cumulative claim matrix $C_{I,n}$. However, recall that only those $C_{i,j}$ in $C_{I,n}$ are observed that are contained in the upper loss triangle $D_{I,n}$. Hence, by using the multiplicative relationship in (2.13), we have also *perfect* knowledge of $F_{i,j}$, $i = -n, \dots, I - 1, j = 0, \dots, I - i - 1$.

According to Lemma 2.4 in Steinmetz and Jentsch (2022), Assumptions 2.2 and 2.3 allow to derive formulas for the (unconditional) means and variances of $C_{i,j}$, $i = -n, \dots, I, j = 0, \dots, I + n$ leading to

$$E(C_{i,j}) = \mu_0 \prod_{k=0}^{j-1} f_k =: \mu_j \quad \text{and} \quad Var(C_{i,j}) = \tau_0^2 \prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1} \sigma_l^2 \prod_{n=l+1}^{j-1} f_n^2 \prod_{m=0}^{l-1} f_m =: \tau_j^2,$$

where μ_0 and τ_0^2 are defined in Assumption 2.2. Together with Assumption 2.4 below, according to Lemma 4.2 in Steinmetz and Jentsch (2022), both sequences $(\mu_j, j \in \mathbb{N}_0)$ and $(\tau_j^2, j \in \mathbb{N}_0)$ are non-negative, monotonically non-decreasing, and converging with $\mu_j \rightarrow \mu_\infty$ and $\tau_j^2 \rightarrow \tau_\infty^2$ as $j \rightarrow \infty$, where $\mu_\infty := \mu_0 \prod_{j=0}^\infty f_j$ and $\tau_\infty^2 := \tau_0^2 \prod_{k=0}^\infty f_k^2 + \mu_0 \sum_{l=0}^\infty \left(\prod_{m=0}^{l-1} f_m \right) \sigma_l^2 \left(\prod_{n=l+1}^\infty f_n^2 \right)$.

Assumption 2.4 (Development factors and variance parameters). Letting $n \rightarrow \infty$ in the setup of Assumptions 2.2 and 2.3 leads to

- (i) a sequence of development factors $(f_j, j \in \mathbb{N}_0)$ with $f_j \geq 1$ for all $j \in \mathbb{N}_0$ and $f_j \rightarrow 1$ as $j \rightarrow \infty$ such that $\prod_{j=0}^\infty f_j < \infty$, which is equivalent to $\sum_{j=0}^\infty (f_j - 1) < \infty$.
- (ii) a sequence of variance parameters $(\sigma_j^2, j \in \mathbb{N}_0)$ with $\sigma_0^2 > 0$ and $\sigma_j^2 \geq 0$ for all $j \in \mathbb{N}$ with $\sigma_j^2 \rightarrow 0$ as $j \rightarrow \infty$ such that $\sum_{j=0}^\infty (j + 1)^2 \sigma_j^2 < \infty$.

The conditions imposed on the sequences of development factors $(f_j, j \in \mathbb{N}_0)$ and variance parameters $(\sigma_j^2, j \in \mathbb{N}_0)$ in Assumption 2.4 are rather mild. In practice, each claim has a finite, but possibly unknown horizon until it is finally settled, which varies by the insurance lines. Altogether, as done in Steinmetz and Jentsch (2022, Section 3.1), this setup allows to derive central limit theorems (CLTs) for (smooth functions of) the parameter estimators $\hat{f}_{j,n}$ for $n \rightarrow \infty$. For example, $f_j = 1 + \frac{1}{j}$ or $\sigma_j^2 = \frac{1}{j^s}$ with $s \leq 3$ are ruled out by Assumption 2.4 such that f_j and σ_j^2 have to converge fast enough to 1 and 0, respectively. Otherwise, $E(C_{i,j}) = \mu_j$ and $Var(C_{i,j}) = \tau_j^2$ would not converge for $j \rightarrow \infty$.

According to Steinmetz and Jentsch (2022, Section 3.2), the following additional assumption has to be imposed to derive a CLT also for $\hat{\sigma}_{j,n}^2$. Although the distributional properties of $\hat{\sigma}_{j,n}^2$ do not show asymptotically in the distribution of the reserve, $\sqrt{I+n}$ -consistency of $\hat{\sigma}_{j,n}^2$ as obtained in Steinmetz and Jentsch (2022, Theorem 3.5) is required for establishing the bootstrap asymptotic theory in Section 4.

Assumption 2.5 (Higher-order conditional moments of individual development factors). For all $i \in \mathbb{Z}, i \leq I, j \in \mathbb{N}_0$, suppose that conditional on $C_{i,j}$, the third and fourth (central) moments of the individual development factors $F_{i,j}$, that is, $E((F_{i,j} - f_j)^3 | C_{i,j})$ and $E((F_{i,j} - f_j)^4 | C_{i,j})$ exist such that both

$$\kappa_j^{(3)} = E(C_{i,j}^2 E((F_{i,j} - f_j)^3 | C_{i,j})) \quad \text{and} \quad \kappa_j^{(4)} = E(C_{i,j}^2 E((F_{i,j} - f_j)^4 | C_{i,j})) \tag{2.15}$$

exist and are finite, respectively.

For example, as shown in Steinmetz and Jentsch (2022, Example 3.4), Assumption 2.5 is fulfilled for certain distributions commonly used in practice, including (conditional) normal, gamma, and log-normal distribution.

Using (2.13), conditional on $Q_{I,n}$, the reserve $R_{I,n}$ can be written as

$$R_{I,n} = \sum_{i=-n}^I C_{i,I-i} \left(\prod_{j=I-i}^{I+n-1} F_{i,j} - 1 \right). \tag{2.16}$$

Hence, by plugging-in (2.9) and (2.16), the predictive root of the reserve from (2.12) becomes

$$R_{I,n} - \hat{R}_{I,n} = \sum_{i=-n}^I C_{i,I-i} \left(\prod_{j=I-i}^{I+n-1} F_{i,j} - \prod_{j=I-i}^{I+n-1} \hat{f}_{j,n} \right) = \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,i} - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right), \tag{2.17}$$

where we flipped the index i to $I - i$ in the last step.

3. Mack’s bootstrap scheme

The Mack bootstrap, introduced by England and Verrall (2006), equips Mack’s Model with a resampling procedure to estimate the whole distribution of the (predicted) reserve. It is very popular and widely used in practice as it describes a rather simple to implement algorithm to estimate the *reserve risk* by estimating high quantiles of the reserve distribution.

As proposed by England and Verrall (2006), to mimic the distribution of the predictive root of the reserve $R_{I,n} - \hat{R}_{I,n}$, the Mack bootstrap constructs a certain bootstrap version $R_{I,n}^* - \hat{R}_{I,n}^*$ of it. On the one hand, this bootstrap predictive root relies on the *same* best estimate of the reserve and centers $R_{I,n}^*$ also around $\hat{R}_{I,n}$. On the other hand, it constructs a certain *double*-bootstrap version of the reserve $R_{I,n}$, that is $R_{I,n}^*$, by combining *two* complementing (non-parametric and parametric) bootstrap approaches for resampling the individual development factors in the upper triangle and in the lower triangle:

- (i) First, a *non-parametric* residual-based bootstrap (see Step 4 below) is applied to construct bootstrap individual development factors $F_{i,j}^*, j = 0, \dots, I+n-1, i = -n, \dots, I-j-1$, that is, for the *upper triangle*, in order to get bootstrap development factor estimators $\hat{f}_{j,n}^*, j = 0, \dots, I+n-1$.
- (ii) Second, the bootstrap development factor estimators $\hat{f}_{j,n}^*$ from (i) together with a *parametric* bootstrap (see Step 5 below) are used to construct also bootstrap individual development factors $F_{i,j}^*, i = -n, \dots, I, j = 0, \dots, I+n-1$ and $i+j \geq I$, that is for the *lower triangle*. For this purpose, a parametric family of (conditional) bootstrap distributions has to be chosen.

Finally, as we are dealing with a prediction problem when estimating the reserve risk, the limiting properties of the predictive root of the reserve *conditional* on the latest observed cumulative claims are relevant and have to be mimicked by a suitable resampling procedure. For this purpose, the Mack bootstrap is employed to estimate the conditional distribution of $R_{I,n} - \hat{R}_{I,n}$ given $Q_{I,n}$ by the conditional bootstrap distribution of $R_{I,n}^* - \hat{R}_{I,n}^*$ given $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$.

3.1. Mack’s bootstrap algorithm

With the upper triangle $D_{I,n}$ at hand, Mack’s bootstrap algorithm is defined as follows:

- Step 1. Estimate the development factors f_j and the variance parameters σ_j^2 from $D_{I,n}$ by computing $\hat{f}_{j,n}$ and $\hat{\sigma}_{j,n}^2$ for $j = 0, \dots, I+n-1$ as defined in (2.8) and (2.11), respectively.
- Step 2. For all $j = 0, \dots, I+n-1$ with $\hat{\sigma}_{j,n}^2 > 0$, compute ‘residuals’

$$\hat{r}_{i,j} = \sqrt{C_{i,j}}(F_{i,j} - \hat{f}_{j,n}) / \hat{\sigma}_{j,n}, \quad i = -n, \dots, I-j-1, \tag{3.1}$$

and re-center and re-scale them to get $\tilde{r}_{i,j} = \frac{1}{s} (\hat{r}_{i,j} - \bar{r})$, where¹

$$\bar{r} = \frac{2}{(I+n+1)(I+n)-2} \sum_{k=0}^{I+n-2} \sum_{l=-n}^{I-k-1} \hat{r}_{l,k}, \quad s^2 = \frac{2}{(I+n+1)(I+n)-2} \sum_{k=0}^{I+n-2} \sum_{l=-n}^{I-k-1} (\hat{r}_{l,k} - \bar{r})^2.$$

Step 3. Draw randomly with replacement from the re-centered and re-scaled residuals $\tilde{r}_{i,j}$, $j = 0, \dots, I+n-2$, $i = -n, \dots, I-j-1$ to get 'bootstrap errors' $r_{i,j}^*$, $j = 0, \dots, I+n-1$, $i = -n, \dots, I-j-1$.

Step 4. Define the bootstrap individual development factors

$$F_{i,j}^* = \hat{f}_{j,n} + \frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} r_{i,j}^*, \quad j = 0, \dots, I+n-1, \quad i = -n, \dots, I-j-1, \tag{3.2}$$

let $\mathcal{F}_{I,n}^* = \{F_{i,j}^* | j = 0, \dots, I+n-1, i = -n, \dots, I-j-1\}$, and compute the Mack bootstrap development factor estimators

$$\hat{f}_{j,n}^* = \frac{\sum_{i=-n}^{I-j-1} C_{i,j} F_{i,j}^*}{\sum_{i=-n}^{I-j-1} C_{i,j}} = \hat{f}_{j,n} + \frac{\hat{\sigma}_{j,n} \sum_{i=-n}^{I-j-1} \sqrt{C_{i,j}} r_{i,j}^*}{\sum_{i=-n}^{I-j-1} C_{i,j}}, \quad j = 0, \dots, I+n-1. \tag{3.3}$$

Step 5. Choose a parametric family for the (conditional) bootstrap distributions of $F_{i,j}^*$ given $C_{i,j}^*$, $\mathcal{D}_{I,n}$ and $\mathcal{F}_{I,n}^*$ such that $F_{i,j}^* > 0$ a.s. with

$$E^*(F_{i,j}^* | C_{i,j}^*, \mathcal{F}_{I,n}^*) = \hat{f}_{j,n}^*, \quad Var^*(F_{i,j}^* | C_{i,j}^*, \mathcal{F}_{I,n}^*) = \frac{\hat{\sigma}_{j,n}^2}{C_{i,j}^*}, \quad i = -n, \dots, I, \quad j = I-i, \dots, I+n-1,$$

where $E^*(\cdot) := E^*(\cdot | \mathcal{D}_{I,n})$, $Var^*(\cdot) := Var^*(\cdot | \mathcal{D}_{I,n})$, etc. denote the Mack bootstrap mean, variance, etc., respectively, that is, conditional on the data $\mathcal{D}_{I,n}$. Then, given $Q_{I,n}^* = Q_{I,n}$, generate the bootstrap ultimate claims $C_{i,I+n}^*$ and the reserves $R_{i,I+n}^* = C_{i,I+n}^* - C_{i,I-i}^*$ for $i = -n, \dots, I$ using the recursion

$$C_{i,j+1}^* = C_{i,j}^* F_{i,j}^*, \quad j = I-i, \dots, I+n-1. \tag{3.4}$$

Step 6. Compute the bootstrap total reserve $R_{I,n}^* = \sum_{i=0}^{I+n} R_{I-i,I+n}^*$ and its bootstrap predictive root

$$R_{I,n}^* - \hat{R}_{I,n} = \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right). \tag{3.5}$$

Step 7. Repeat Steps 3 - 6 above B times, where B is large, to get bootstrap predictive roots $(R_{I,n}^* - \hat{R}_{I,n})^{(b)}$, $b = 1, \dots, B$, and denote by $q^*(\alpha)$ the α -quantile of their empirical distribution.

Step 8. Construct the $(1 - \alpha)$ equal-tailed prediction interval for $R_{I,n}$ as

$$\left[\hat{R}_{I,n} + q^*(\alpha/2), \hat{R}_{I,n} + q^*(1 - \alpha/2) \right].$$

Remark 3.1 (On Mack's bootstrap proposal).

(i) While the Mack bootstrap predictive root of the reserve $R_{I,n}^* - \hat{R}_{I,n}$ uses the same best estimate $\hat{R}_{I,n}$ for centering (as in $R_{I,n} - \hat{R}_{I,n}$), it relies on a certain type of double-bootstrap version $R_{I,n}^*$ of the total reserve $R_{I,n}$, which employs $\hat{f}_{j,n}^*$ instead of just $\hat{f}_{j,n}$, but uses $\hat{\sigma}_{j,n}^2$. However, although $E^*(F_{i,j}^* | C_{i,j}^*, \mathcal{D}_{I,n}, \mathcal{F}_{I,n}^*) = \hat{f}_{j,n}^*$ holds, we still have $E^*(F_{i,j}^* | C_{i,j}^*) = \hat{f}_{j,n}$ for the lower triangle individual development factors. In contrast, for the variances, we have $Var^*(F_{i,j}^* | C_{i,j}^*, \mathcal{F}_{I,n}^*) = \frac{\hat{\sigma}_{j,n}^2}{C_{i,j}^*}$, but

$$Var^*(F_{i,j}^* | C_{i,j}^*) = \frac{\hat{\sigma}_{j,n}^2}{C_{i,j}^*} + \frac{\hat{\sigma}_{j,n}^2}{\sum_{k=-n}^{I-j-1} C_{k,j}}.$$

(ii) Due to the fixed-design bootstrap in Step 4, which does not generate bootstrap cumulative claims $C_{i,j}^*$ (and consequently no bootstrap upper loss triangle $\mathcal{D}_{I,n}^*$), but only $F_{i,j}^*$'s, the bootstrap development factor estimators $\hat{f}_{j,n}^*$ and $\hat{f}_{k,n}^*$ defined in (3.3) are independent for $j \neq k$ conditional on $\mathcal{D}_{I,n}$. This is on contrast to the development factor estimators $\hat{f}_{j,n}$ and $\hat{f}_{k,n}$, which are asymptotically independent for $j \neq k$, but only uncorrelated in finite samples such that $E(\hat{f}_{j,n} \hat{f}_{k,n}^2) < 0$ for $j \neq k$.

(iii) The non-parametric bootstrap used to construct the $\hat{f}_{j,n}^*$'s in Step 4 uses residuals, but according to Assumption 2.2 and 2.3, there are no errors in Mack's model that are approximated by these residuals. In fact, each (possibly parametric) bootstrap proposal that successfully mimics the first and second conditional moments of $C_{i,j+1}$ given $C_{i,j}$ will correctly mimic the limiting distribution of the $\hat{f}_{j,n}$'s.

(iv) In view of the discussion above, a fully parametric implementation that uses the same parametric family from Step 5 also in Step 4 to get bootstrap development factors $F_{i,j}^*$'s can be used.

¹ Note that $\hat{\sigma}_{I+n-1,n}^2 = 0$ by construction such that $\hat{r}_{I+n-1,n}$ is excluded in (3.1) such that (at most) $(I+n+1)(I+n)/2 - 1 = ((I+n+1)(I+n) - 2)/2$ residuals can be computed. If $\hat{\sigma}_{j,n}^2 = 0$ holds also for other j , the corresponding residuals are excluded in (3.1) as well and the formulas for \bar{r} and s have to be adjusted accordingly. In the following, for notational convenience, we assume that only $\hat{\sigma}_{I+n-1,n}^2 = 0$ and $\hat{\sigma}_{j,n}^2 > 0$ holds for all $j = 0, \dots, I+n-2$ and all $n \in \mathbb{N}_0$.

- (v) A fully non-parametric approach that uses the non-parametric bootstrap from Step 4 also in Step 5 is thinkable, but would suffer from issues arising from potentially negative $F_{I,j}^*$'s leading to a reduced finite sample performance.

4. Asymptotic theory for the Mack bootstrap

Although the Mack bootstrap as proposed by England and Verrall (2006) and described in Section 3 is widely used in practice for reserve risk estimation, limiting results that confirm its consistency are still missing in the literature. In this section, based on the asymptotic and stochastic framework described in Section 2, we derive asymptotic theory for the Mack bootstrap, which enables a rigorous investigation of its consistency properties.

The Mack bootstrap is designed to mimic the distribution of the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$ conditional on $\mathcal{Q}_{I,n}$ based on the bootstrap distribution of the corresponding Mack bootstrap predictive root of the reserve $R_{I,n}^* - \widehat{R}_{I,n}^*$ conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$. Hence, a closer inspection of both expressions is advisable. Picking-up the representation of the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$ in (2.17), it can be decomposed into two additive parts that account for the prediction error and the estimation error, respectively. Precisely, by subtracting and adding $\sum_{i=0}^{I+n} C_{I-i,i} \prod_{j=i}^{I+n-1} f_j$, we get

$$\begin{aligned} R_{I,n} - \widehat{R}_{I,n} &= \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,j} - \prod_{j=i}^{I+n-1} f_j \right) + \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} f_j - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right) \\ &=: (R_{I,n} - \widehat{R}_{I,n})_1 + (R_{I,n} - \widehat{R}_{I,n})_2, \end{aligned} \tag{4.1}$$

where $(R_{I,n} - \widehat{R}_{I,n})_1$ represents the process uncertainty (that carries the process variance) and $(R_{I,n} - \widehat{R}_{I,n})_2$ the estimation uncertainty (that carries the estimation variance).

Similarly, for the Mack bootstrap predictive root of the reserve $R_{I,n}^* - \widehat{R}_{I,n}^*$ from (3.5), by subtracting and adding $\sum_{i=0}^{I+n} C_{I-i,i}^* \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^*$, we get

$$\begin{aligned} R_{I,n}^* - \widehat{R}_{I,n}^* &= \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^* - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^* \right) + \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^* - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right) \\ &=: (R_{I,n}^* - \widehat{R}_{I,n}^*)_1 + (R_{I,n}^* - \widehat{R}_{I,n}^*)_2, \end{aligned} \tag{4.2}$$

where $(R_{I,n}^* - \widehat{R}_{I,n}^*)_1$ and $(R_{I,n}^* - \widehat{R}_{I,n}^*)_2$ are the Mack bootstrap versions of $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$, respectively.

As main interest is in the distribution of the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$ conditional on $\mathcal{Q}_{I,n}$, in view of the decompositions (4.1) and (4.2), it is instructive to first consider separately the (limiting) distributions of $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$ conditional on $\mathcal{Q}_{I,n}$, respectively. They will serve as valuable benchmark distributions for the investigation of consistency properties of the Mack bootstrap in Section 4.2. Such asymptotic results have been established in Steinmetz and Jentsch (2022, Section 4). We will briefly summarize the relevant conditional limiting distributions below in Section 4.1.

4.1. Conditional asymptotics for the predictive root of the reserve

In the following, we review the conditional asymptotic results established in Steinmetz and Jentsch (2022, Theorems 4.3, 4.10, 4.12, and Corollary 4.13) separately for the process uncertainty term $(R_{I,n} - \widehat{R}_{I,n})_1$ in Section 4.1.1, for the estimation uncertainty term $(R_{I,n} - \widehat{R}_{I,n})_2$ in Section 4.1.2, as well as jointly for $R_{I,n} - \widehat{R}_{I,n}$ in Section 4.1.3, respectively.

4.1.1. Conditional asymptotics for reserve prediction: process uncertainty

Based on Theorem 4.3 from Steinmetz and Jentsch (2022), the following theorem provides the limiting distribution of the process uncertainty term $(R_{I,n} - \widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}$.

Theorem 4.1 (Asymptotics for $(R_{I,n} - \widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}$). *Suppose Assumptions 2.2, 2.3 and 2.4 hold. Then, as $n \rightarrow \infty$, conditionally on $\mathcal{Q}_{I,n}$, $(R_{I,n} - \widehat{R}_{I,n})_1$ converges in L_2 -sense to the non-degenerate random variable $(R_{I,\infty} - \widehat{R}_{I,\infty})_1$. That is, we have*

$$E \left(\left((R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,\infty} - \widehat{R}_{I,\infty})_1 \right)^2 \middle| \mathcal{Q}_{I,n} \right) \xrightarrow{P} 0, \tag{4.3}$$

where

$$(R_{I,\infty} - \widehat{R}_{I,\infty})_1 := \sum_{i=0}^{\infty} C_{I-i,i} \left(\prod_{j=i}^{\infty} F_{I-i,j} - \prod_{j=i}^{\infty} f_j \right) \sim \mathcal{G}_1. \tag{4.4}$$

Conditional on $\mathcal{Q}_{I,\infty} = (C_{I-i,i} | i \in \mathbb{N}_0)$, \mathcal{G}_1 has mean zero, $E((R_{I,\infty} - \widehat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty}) = 0$, and variance

$$Var \left((R_{I,\infty} - \widehat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty} \right) = \sum_{i=0}^{\infty} C_{I-i,i} \sum_{j=i}^{\infty} \left(\prod_{k=i}^{j-1} f_k \right) \sigma_j^2 \left(\prod_{l=j+1}^{\infty} f_l^2 \right) = O_P(1). \tag{4.5}$$

The (conditional) L_2 -convergence result in Theorem 4.1 immediately implies also (conditional) convergence in distribution. That is, for $n \rightarrow \infty$, we have

$$(R_{I,n} - \widehat{R}_{I,n})_1 | \mathcal{Q}_{I,n} \xrightarrow{d} (R_{I,\infty} - \widehat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty} \sim \mathcal{G}_1 | \mathcal{Q}_{I,\infty}. \tag{4.6}$$

Moreover, according to Theorem 4.1 (see also the discussion in (Steinmetz and Jentsch 2022, Remark 4.4)), the conditional limiting distribution $\mathcal{G}_1 | \mathcal{Q}_{I,\infty}$ will be typically *non-Gaussian* and depending on the (conditional) distribution of the individual development factors $F_{i,j} | C_{i,j}$.

4.1.2. *Conditional asymptotics for reserve prediction: estimation uncertainty*

In comparison to the conditional limiting result for $(R_{I,n} - \widehat{R}_{I,n})_1$ displayed in Theorem 4.1, the derivation of asymptotic results for $(R_{I,n} - \widehat{R}_{I,n})_2$ is rather different and also much more cumbersome. In particular, to obtain non-degenerate limiting distributions, we have to inflate $(R_{I,n} - \widehat{R}_{I,n})_2$ by $\sqrt{I+n+1}$ and the obtained (Gaussian) distribution relies on CLTs for (smooth functions of) development factor estimators $\widehat{f}_{j,n}$ established in (Steinmetz and Jentsch 2022, Section 3 and Appendix C). For the derivation of asymptotic theory, conditional on $\mathcal{Q}_{I,n}$, it is instructive to further decompose $(R_{I,n} - \widehat{R}_{I,n})_2$ to get

$$\begin{aligned} & (R_{I,n} - \widehat{R}_{I,n})_2 \\ &= \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} f_j - \prod_{j=i}^{I+n-1} f_{j,n}(\mathcal{Q}_{I,n}) \right) + \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} f_{j,n}(\mathcal{Q}_{I,n}) - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right) \\ &= (R_{I,n} - \widehat{R}_{I,n})_2^{(1)} + (R_{I,n} - \widehat{R}_{I,n})_2^{(2)}, \end{aligned} \tag{4.7}$$

where $(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ is measurable wrt $\mathcal{Q}_{I,n}$ and $f_{j,n}(\mathcal{Q}_{I,n}) := \mu_{j+1,n}^{(1)}(\mathcal{Q}_{I,n}) / \mu_{j,n}^{(2)}(\mathcal{Q}_{I,n})$ with $\mu_{j+1,n}^{(1)}(\mathcal{Q}_{I,n}) = E(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j+1} | \mathcal{Q}_{I,n})$ and $\mu_{j,n}^{(2)}(\mathcal{Q}_{I,n}) = E(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} | \mathcal{Q}_{I,n})$. The derivation of (conditional) asymptotic theory for $(R_{I,n} - \widehat{R}_{I,n})_2$ requires additional assumptions on the stochastic properties of the individual development factors $F_{i,j}$ summarized in Assumptions 4.2 and 4.3 below, which resemble Assumptions 4.6 and 4.8 in Steinmetz and Jentsch (2022).

Assumption 4.2 (*Support condition and variance parameters*). The individual development factors $(F_{i,j}, i \in \mathbb{Z}, i \leq I, j \in \mathbb{N}_0)$ are random variables with support (ϵ, ∞) for some $\epsilon > 0$ and the sequence of variance parameters $(\sigma_j^2, j \in \mathbb{N}_0)$ converges to 0 as $j \rightarrow \infty$ such that $\sum_{j=0}^{\infty} (j+1)^2 \frac{\sigma_j^2}{\epsilon^2} < \infty$.

While Assumption 4.2 is a sufficient condition for the limiting results below, it appears to be *not* a necessary condition. We conjecture that the theory derived in this Section 4.1.2 still holds (with technically more tedious proofs), when dropping the lower bound $\epsilon > 0$ and asking just for $\sum_{j=0}^{\infty} (j+1)^2 \sigma_j^2 < \infty$ as in Assumption 2.4. However, if Assumption 2.4 holds, the closeness to 0 of the lower bound of $F_{i,j}$ could violate the above Assumption 4.2. Nevertheless, for practical applications, $F_{i,j}$ is usually not close to 0 such that this condition appears to be not restrictive.

In addition to the condition on the support and the variance parameters in Assumption 4.2, a regularity condition for the *backward* conditional distribution of cumulative claim $C_{i,j}$ given $C_{i,j+1}$ is required.

Assumption 4.3 (*Backward conditional moments*). Assumptions 2.2, 2.3, 2.4 and 4.2 are fulfilled such that, for all $K \in \mathbb{N}_0, k \geq 0$ and $j, j_1, j_2 \in \{0, \dots, K\}, j_1 \leq j_2$, we have

$$|E(C_{i,j} | C_{i,j+k}) - E(C_{i,j} | C_{i,j+k+1})| \leq a_k X_i,$$

$$|Cov(C_{i,j_1}, C_{i,j_2} | C_{i,j_2+k}) - Cov(C_{i,j_1}, C_{i,j_2} | C_{i,j_2+k+1})| \leq b_k Y_i,$$

where $(X_i, i \in \mathbb{Z}, i \leq I), (Y_i, i \in \mathbb{Z}, i \leq I)$ are sequences of non-negative i.i.d. random variables with $E(X_i^{2+\delta}) < \infty$ for some $\delta > 0$ and $E(Y_i^2) < \infty$, and $(a_j, j \in \mathbb{N}_0)$ and $(b_j, j \in \mathbb{N}_0)$ are non-negative real-valued sequences with $\sum_{j=0}^{\infty} (j+1)^2 a_j < \infty$ and $\sum_{j=0}^{\infty} (j+1)^2 b_j < \infty$.

While Mack’s model is designed to generate loss triangles in a rather simple *forward* way according to the recursion (2.13), which allows to easily calculate *forward* conditional means $E(C_{i,j+1} | C_{i,j})$ and variances $Var(C_{i,j+1} | C_{i,j})$, it is *not* straightforward to calculate *backward* conditional means $E(C_{i,j} | C_{i,j+1})$ and variances $Var(C_{i,j} | C_{i,j+1})$. We refer to Example 4.9 in Steinmetz and Jentsch (2022) for an already very complicated analytical derivation of such backward moments in a simple toy example using uniform distributions.

Based on Theorem 4.10 in Steinmetz and Jentsch (2022), which relies on conditional CLTs for (smooth functions of) development factor estimators $\widehat{f}_{j,n}$ given $\mathcal{Q}_{I,n}$ stated in (Steinmetz and Jentsch 2022, Appendix C), the following theorem provides the limiting distribution of the estimation uncertainty term $(R_{I,n} - \widehat{R}_{I,n})_2$ conditional on $\mathcal{Q}_{I,n}$. While $(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ is measurable with respect to $\mathcal{Q}_{I,n}$, Assumptions 4.2 and 4.3 allow to prove asymptotic normality of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(2)}$ conditional on $\mathcal{Q}_{I,n}$.

Theorem 4.4 (*Asymptotics for $(R_{I,n} - \widehat{R}_{I,n})_2$ conditional on $\mathcal{Q}_{I,n}$*). Suppose Assumptions 2.2, 2.3, 2.4, 4.2 and 4.3 hold. Then, as $n \rightarrow \infty$, the following holds:

- (i) *Unconditionally, $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ converges in distribution to a non-degenerate limiting distribution $\mathcal{G}_2^{(1)}$. That is, we have*

$$\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(1)} \xrightarrow{d} \langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_{\infty}^{(1)} \rangle \sim \mathcal{G}_2^{(1)}, \tag{4.8}$$

where $\mathbf{Y}_{\infty}^{(1)} = (Y_i^{(1)}, i \in \mathbb{N}_0)$ denotes a centered Gaussian process with covariances

$$Cov(Y_{i_1}^{(1)}, Y_{i_2}^{(1)}) = \lim_{K \rightarrow \infty} \Sigma_{K, \prod f_j}^{(1)}(i_1, i_2), \quad i_1, i_2 \in \mathbb{N}_0,$$

where $\Sigma_{K, \prod f_j}^{(1)}(i_1, i_2)$ is defined in Corollary C.2 in Steinmetz and Jentsch (2022). Here, the two random sequences $\mathcal{Q}_{I, \infty}$ and $\mathbf{Y}_{\infty}^{(1)}$ are stochastically independent.

(ii) Conditionally on $\mathcal{Q}_{I, n}$, $\sqrt{I+n+1}(R_{I, n} - \hat{R}_{I, n})_2^{(2)}$ converges in distribution to a centered normal distribution. That is, we have

$$\sqrt{I+n+1}(R_{I, n} - \hat{R}_{I, n})_2^{(2)} | \mathcal{Q}_{I, n} \xrightarrow{d} \langle \mathcal{Q}_{I, \infty}, \mathbf{Y}_{\infty}^{(2)} \rangle | \mathcal{Q}_{I, \infty} \sim \mathcal{G}_2^{(2)} | \mathcal{Q}_{I, \infty}, \tag{4.9}$$

where $\mathcal{G}_2^{(2)} | \mathcal{Q}_{I, \infty} \sim \mathcal{N}(0, \Xi(\mathcal{Q}_{I, \infty})) | \mathcal{Q}_{I, \infty}$ is Gaussian with mean zero and variance

$$\Xi(\mathcal{Q}_{I, \infty}) = \lim_{K \rightarrow \infty} \mathcal{Q}_{I, K-I} \Sigma_{K, \prod f_j}^{(2)} \mathcal{Q}'_{I, K-I} = \lim_{K \rightarrow \infty} \mathcal{Q}_{I, K-I} (\Sigma_{K, \prod f_j} - \Sigma_{K, \prod f_j}^{(1)}) \mathcal{Q}'_{I, K-I}, \tag{4.10}$$

where $\Sigma_{K, \prod f_j}$ and $\Sigma_{K, \prod f_j}^{(1)}$ and $\Sigma_{K, \prod f_j}^{(2)}$ are defined in Corollary C.2 in Steinmetz and Jentsch (2022).

According to Theorem 4.4(ii), in contrast to $\mathcal{G}_1 | \mathcal{Q}_{I, \infty}$ in Theorem 4.1, the limiting distribution $\mathcal{G}_2^{(2)} | \mathcal{Q}_{I, \infty}$ is Gaussian. Together with Theorem 4.4(i), conditional on $\mathcal{Q}_{I, n}$, the estimation uncertain term $\sqrt{I+n+1}(R_{I, n} - \hat{R}_{I, n})_2$ is Gaussian with mean $\langle \mathcal{Q}_{I, \infty}, \mathbf{Y}_{\infty}^{(2)} \rangle$ and variance $\Xi(\mathcal{Q}_{I, \infty})$.

4.1.3. Conditional asymptotics for the whole predictive root of the reserve

By combining the results derived for $(R_{I, n} - \hat{R}_{I, n})_1$ and $(R_{I, n} - \hat{R}_{I, n})_2$ in Theorems 4.1 and 4.4, respectively, joint asymptotic results for $R_{I, n} - \hat{R}_{I, n}$ conditional on $\mathcal{Q}_{I, n}$ can also be established.

Theorem 4.5 (Asymptotics for $R_{I, n} - \hat{R}_{I, n}$ conditional on $\mathcal{Q}_{I, n}$). Suppose the assumptions of Theorems 4.1 and 4.4 hold. Then, conditional on $\mathcal{Q}_{I, n}$, $(R_{I, n} - \hat{R}_{I, n})_1$ and $(R_{I, n} - \hat{R}_{I, n})_2$ are stochastically independent, and $R_{I, n} - \hat{R}_{I, n} | \mathcal{Q}_{I, n}$ converges in distribution to $\mathcal{G}_1 | \mathcal{Q}_{I, \infty}$. That is, we have

$$R_{I, n} - \hat{R}_{I, n} | \mathcal{Q}_{I, n} = (R_{I, n} - \hat{R}_{I, n})_1 + (R_{I, n} - \hat{R}_{I, n})_2 | \mathcal{Q}_{I, n} \xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I, \infty}. \tag{4.11}$$

According to Theorem 4.4, $(R_{I, n} - \hat{R}_{I, n})_2$ requires an inflation factor $\sqrt{I+n+1}$ to get convergence to a non-degenerate limiting distribution. As this is not the case for $(R_{I, n} - \hat{R}_{I, n})_1$ in Theorem 4.1, the process uncertainty term $(R_{I, n} - \hat{R}_{I, n})_1$ asymptotically dominates the predictive root of the reserve $R_{I, n} - \hat{R}_{I, n}$.

Hence, we can conclude that asymptotic normality of the (predictive root of the) reserve does generally *not* hold, which casts the common practice to use a normal approximation for the reserve in Mack’s model into doubt. Moreover, the shape of $\mathcal{G}_1 | \mathcal{Q}_{I, \infty}$ does depend on the true (conditional) distribution family of the individual development factors $F_{i, j} | C_{i, j}$.

4.2. Conditional bootstrap asymptotics for the Mack bootstrap predictive root of the reserve

In view of the decomposition $R_{I, n} - \hat{R}_{I, n} = (R_{I, n} - \hat{R}_{I, n})_1 + (R_{I, n} - \hat{R}_{I, n})_2$ in (4.1) and the conditional limiting distributions of $(R_{I, n} - \hat{R}_{I, n})_1$ and $(R_{I, n} - \hat{R}_{I, n})_2$ gathered in Section 4.1, it is instructive to consider the corresponding Mack bootstrap quantities $(R_{I, n}^* - \hat{R}_{I, n})_1$ and $(R_{I, n}^* - \hat{R}_{I, n})_2$ from (4.2) and check whether they are correctly mimicking such limiting distributions. While $(R_{I, n} - \hat{R}_{I, n})_1$ and $(R_{I, n} - \hat{R}_{I, n})_2$ are analyzed conditional on $\mathcal{Q}_{I, n}$, the bootstrap quantities $(R_{I, n}^* - \hat{R}_{I, n})_1$ and $(R_{I, n}^* - \hat{R}_{I, n})_2$ have to be considered conditional on $\mathcal{Q}_{I, n}^* = \mathcal{Q}_{I, n}$, but also on $\mathcal{D}_{I, n}$.

4.2.1. Conditional bootstrap asymptotics for reserve prediction: process uncertainty

For the derivation of bootstrap asymptotics, we have to impose additional smoothness properties of the parametric family of (conditional) distributions of the individual development factors to assure that consistent estimation of development factors and variance parameters implies also consistent estimation of the whole distribution.

Assumption 4.6 (Parametric family of (conditional) distributions of $F_{i, j}$). The (conditional) distribution $F_{i, j} | C_{i, j}$, $i \in \mathbb{Z}$, $i \leq I$, $j \in \mathbb{N}_0$, belongs to a parametric family of distributions \mathbf{H} such that:

- (i) A distribution $\mathcal{H} \in \mathbf{H}$ is uniquely specified by its first two (conditional) moments. That is, for all $i \in \mathbb{Z}$, $i \leq I$, $j \in \mathbb{N}_0$ and all $c \in (0, \infty)$ the conditional distribution of $F_{i, j} | C_{i, j} = c$ is uniquely determined by $E(F_{i, j} | C_{i, j} = c) = f_j$ and $Var(F_{i, j} | C_{i, j} = c) = \frac{\sigma_j^2}{c}$ according to (2.14).
- (ii) The distributions $\mathcal{H} \in \mathbf{H}$ are continuous in f_j and σ_j^2 . That is, for all $i \in \mathbb{Z}$, $i \leq I$, $j \in \mathbb{N}_0$ and for all $c \in (0, \infty)$, the conditional distribution of $F_{i, j} | C_{i, j} = c$ is continuous in a neighborhood of (f_j, σ_j^2) .

As the limiting distribution derived in Theorem 4.1 is generally non-Gaussian and depends on the (conditional) distribution (family) of the individual development factors, we require also that the bootstrap individual development factors $F_{i, j}^*$, $i = -n, \dots, I$, $j = I - i, \dots, I + n - 1$, that is, for the lower triangle, follow the true parametric family of (conditional) distributions as the $F_{i, j}$ ’s according to Assumption 4.6.

Assumption 4.7 ((Conditional) distributions of $F_{i, j}^*$ in lower triangle). For any $n \in \mathbb{N}_0$, the (conditional) distribution of $F_{i, j}^* | C_{i, j}^*, \mathcal{D}_{I, n}, \mathcal{F}_{I, n}^*$, $j = 0, \dots, I + n - 1$ and $i = -n, \dots, I - j - 1$ in Step 5 of the Mack Bootstrap in Section 3.1 belongs to the true parametric family of (conditional) distributions \mathbf{H} used to generate $F_{i, j} | C_{i, j}$ according to Assumption 4.6. That is, we have

$$F_{i, j}^* | (C_{i, j}^* = x, \hat{f}_{j, n}^* = y, \hat{\sigma}_{j, n}^2 = z) \stackrel{d}{=} F_{i, j} | (C_{i, j} = x, f_j = y, \sigma_j^2 = z) \quad \text{for all } (x, y, z)' \in (0, \infty).$$

Together with the setup of Theorem 4.1, the Assumptions 4.6 and 4.7 allow to prove the following result.

Theorem 4.8 (Bootstrap asymptotics for $(R_{I,n}^* - \hat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$). Suppose Assumptions 2.2, 2.3, 2.4, 2.5, 4.6, and 4.7 hold. Then, as $n \rightarrow \infty$, conditionally on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, $(R_{I,n}^* - \hat{R}_{I,n})_1$ converges in distribution to $\mathcal{G}_1 | \mathcal{Q}_{I,\infty}$ in probability, which is the (limiting) distribution of $(R_{I,\infty} - \hat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty}$ according to (4.6) described in Theorem 4.1. Moreover, for all $n \in \mathbb{N}_0$, it holds $E^*((R_{I,n}^* - \hat{R}_{I,n})_1 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) = 0$ and, for $n \rightarrow \infty$, we have

$$Var^*((R_{I,n}^* - \hat{R}_{I,n})_1 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) \rightarrow Var((R_{I,\infty} - \hat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty}) \tag{4.12}$$

in probability, where $Var((R_{I,\infty} - \hat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty}) = O_p(1)$ as given in (4.5). Consequently, as $n \rightarrow \infty$, we have

$$d_2\left(\mathcal{L}\left((R_{I,n} - \hat{R}_{I,n})_1 | \mathcal{Q}_{I,n}\right), \mathcal{L}^*\left((R_{I,n}^* - \hat{R}_{I,n})_1 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}\right)\right) \rightarrow 0$$

in probability, where $\mathcal{L}^*(\cdot)$ denotes a bootstrap distribution conditional on $\mathcal{D}_{I,n}$, and d_2 is the Mallows metric, that is defined for two distributions G and H as $d_2(G, H) = \inf(E\|X - Y\|^2)^{\frac{1}{2}}$, where the infimum is taken over all joint distributions of (X, Y) with marginals $X \sim G$ and $Y \sim H$.

4.2.2. Conditional bootstrap asymptotics for reserve prediction: estimation uncertainty

In view of the decomposition $(R_{I,n} - \hat{R}_{I,n})_2 = (R_{I,n} - \hat{R}_{I,n})_2^{(1)} + (R_{I,n} - \hat{R}_{I,n})_2^{(2)}$ in (4.7), for the derivation of corresponding bootstrap asymptotic theory, it is seemingly instructive to further decompose also its bootstrap counterpart $(R_{I,n}^* - \hat{R}_{I,n})_2$ in the same way conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$. That is, by taking into account the specific definition of $\hat{f}_{j,n}^*$ in (3.3), we get

$$\begin{aligned} & (R_{I,n}^* - \hat{R}_{I,n})_2 \\ &= \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* - \prod_{j=i}^{I+n-1} f_{j,n}^*(\mathcal{Q}_{I,n}) \right) + \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} f_{j,n}^*(\mathcal{Q}_{I,n}) - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right) \\ &= (R_{I,n}^* - \hat{R}_{I,n})_2^{(1)} + (R_{I,n}^* - \hat{R}_{I,n})_2^{(2)}, \end{aligned} \tag{4.13}$$

where $f_{j,n}^*(\mathcal{Q}_{I,n}) := \mu_{j+1,n}^{*(1)}(\mathcal{Q}_{I,n}) / \mu_{j,n}^{*(2)}(\mathcal{Q}_{I,n})$ with $\mu_{j+1,n}^{*(1)}(\mathcal{Q}_{I,n}) := E^*\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} F_{i,j}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}\right)$ and $\mu_{j,n}^{*(2)}(\mathcal{Q}_{I,n}) := E^*\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}\right)$. Now, for $\mu_{j+1,n}^{*(1)}(\mathcal{Q}_{I,n})$, we get

$$\mu_{j+1,n}^{*(1)}(\mathcal{Q}_{I,n}) = \frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} E^*(F_{i,j}^* | C_{i,I-i}^* = C_{i,I-i}) = \frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} E^*(F_{i,j}^*)$$

where we used that $C_{i,j}$ is measurable with respect to $\mathcal{D}_{I,n}$ and that $F_{i,j}^*$ is stochastically independent of the condition $C_{i,I-i}^* = C_{i,I-i}$ given $\mathcal{D}_{I,n}$. This is because the Mack bootstrap relies on a fixed-design approach based on the $C_{i,j}$'s instead of recursively generating $C_{i,j}^*$ to get a whole bootstrap loss triangle $D_{I,n}^*$. Altogether, using $E^*(F_{i,j}^*) = \hat{f}_{j,n}$, we get

$$f_{j,n}^*(\mathcal{Q}_{I,n}) = \frac{\hat{f}_{j,n} \left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} \right)}{\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j}} = \hat{f}_{j,n}$$

leading to $(R_{I,n}^* - \hat{R}_{I,n})_2^{(2)} = 0$ such that $(R_{I,n}^* - \hat{R}_{I,n})_2 = (R_{I,n}^* - \hat{R}_{I,n})_2^{(1)}$. Hence, in comparison to $(R_{I,n} - \hat{R}_{I,n})_2$, which was decomposed into two parts $(R_{I,n} - \hat{R}_{I,n})_2^{(1)}$ and $(R_{I,n} - \hat{R}_{I,n})_2^{(2)}$, such an analogous decomposition of $(R_{I,n}^* - \hat{R}_{I,n})_2$ does not exist. However, for $n \rightarrow \infty$, it remains to check the limiting properties of $(R_{I,n}^* - \hat{R}_{I,n})_2$ in the following. In contrast to the derivation of the conditional limiting result obtained in Theorem 4.4(ii), which relies on conditional CLTs for the development factor estimators $\hat{f}_{j,n}$ as stated in (Steinmetz and Jentsch 2022, Appendix C), the derivation of the limiting properties of $(R_{I,n}^* - \hat{R}_{I,n})_2$ rely on unconditional bootstrap CLTs for the Mack bootstrap development factor estimators $\hat{f}_{j,n}^*$, that is, without conditioning on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$. For this purpose, to prove asymptotic normality for the $\hat{f}_{j,n}^*$'s by justifying a Lyapunov condition, we have to impose additional regularity conditions on the estimators for the development factors and variance parameters.

Assumption 4.9 (Uniform boundedness condition). Suppose that the development factor estimators $\hat{f}_{j,n}$, $j = 0, \dots, I+n-1$ and the variance parameter estimators $\hat{\sigma}_{j,n}^2$, $j = 0, \dots, I+n-2$ fulfill

$$\sup_{j=0, \dots, I+n-1} \frac{\hat{f}_{j,n}}{f_j} = O_p(1) \quad \text{and} \quad \sup_{j=0, \dots, I+n-2} \frac{\sigma_j^2}{\hat{\sigma}_{j,n}^2} = O_p(1)$$

for $n \rightarrow \infty$. Moreover, for $\kappa_j^{(4)}$ defined in (2.15), suppose that $((\kappa_j^{(4)} / \sigma_j^4), j \in \mathbb{N}_0)$ is a bounded sequence.

This allows for the following asymptotic result for the Mack bootstrap estimation uncertainty part.

Theorem 4.10 (Bootstrap asymptotics for $(R_{I,n}^* - \hat{R}_{I,n})_2$ conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$). Suppose Assumptions 2.2, 2.3, 2.5, 2.4, 4.2 and 4.9 hold. Then, as $n \rightarrow \infty$, conditionally on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$, $\sqrt{I+n+1}(R_{I,n}^* - \hat{R}_{I,n})_2$ converges in distribution to $\tilde{G}_2|Q_{I,\infty}$ in probability, where $\tilde{G}_2|Q_{I,\infty} \sim \mathcal{N}(0, \tilde{\Xi}(Q_{I,\infty})|Q_{I,\infty})$ is a conditional Gaussian distribution with conditional mean zero and conditional variance

$$\tilde{\Xi}(Q_{I,\infty}) = \lim_{K \rightarrow \infty} Q_{I,K-I} \Sigma_{K,\Pi f_j} Q'_{I,K-I}, \tag{4.14}$$

where $\Sigma_{K,\Pi f_j} = \Sigma_{K,\Pi f_j}^{(1)} + \Sigma_{K,\Pi f_j}^{(2)}$ is defined in Corollary C.2 in Steinmetz and Jentsch (2022).

$$\tilde{\Xi}(Q_{I,\infty}) = \sum_{i_1, i_2=0}^{\infty} C_{I-i_1, i_1} C_{I-i_2, i_2} \sum_{j=\max(i_1, i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{l=\max(i_1, i_2), l \neq j}^{\infty} f_l^2 \prod_{m=\min(i_1, i_2)}^{\max(i_1, i_2)-1} f_m = O_P(1).$$

Consequently, as $n \rightarrow \infty$, we have

$$d_2 \left(\mathcal{L} \left(\sqrt{I+n+1}(R_{I,n} - \hat{R}_{I,n})_2 | Q_{I,n} \right), \mathcal{L}^* \left(\sqrt{I+n+1}(R_{I,n}^* - \hat{R}_{I,n})_2 | Q_{I,n}^* = Q_{I,n} \right) \right) \rightarrow 0 \text{ in prob.,}$$

because the limiting normal distribution of $\sqrt{I+n+1}(R_{I,n}^* - \hat{R}_{I,n})_2$ conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$ deviates in its (zero) mean and variance $\tilde{\Xi}(Q_{I,\infty})$ from that of $\sqrt{I+n+1}(R_{I,n} - \hat{R}_{I,n})_2$ conditional on $Q_{I,n}$, which has mean $\langle Q_{I,\infty}, Y_{\infty}^{(2)} \rangle$ and variance $\Xi(Q_{I,\infty})$ according to Theorem 4.4.

4.2.3. Conditional bootstrap asymptotics for the whole predictive root of the reserve

As in Section 4.1.3, combining the results for $(R_{I,n}^* - \hat{R}_{I,n})_1$ and $(R_{I,n}^* - \hat{R}_{I,n})_2$ from Theorems 4.8 and 4.10, respectively, joint asymptotics for $R_{I,n}^* - \hat{R}_{I,n}$ conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$ can be obtained.

Theorem 4.11 (Bootstrap asymptotics for $R_{I,n}^* - \hat{R}_{I,n}$ conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$). Suppose the assumptions of Theorems 4.8 and 4.10 hold. Then, conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$, $(R_{I,n}^* - \hat{R}_{I,n})_1$ and $(R_{I,n}^* - \hat{R}_{I,n})_2$ are uncorrelated, and $R_{I,n}^* - \hat{R}_{I,n} | (Q_{I,n}^* = Q_{I,n}, D_{I,n})$ converges in distribution to $\mathcal{G}_1 | Q_{I,\infty}$. That is, we have

$$R_{I,n}^* - \hat{R}_{I,n} | \left(Q_{I,n}^* = Q_{I,n}, D_{I,n} \right) = (R_{I,n}^* - \hat{R}_{I,n})_1 + (R_{I,n}^* - \hat{R}_{I,n})_2 | \left(Q_{I,n}^* = Q_{I,n}, D_{I,n} \right) \xrightarrow{d} \mathcal{G}_1 | Q_{I,\infty}$$

in probability.

As already observed in Theorem 4.4 for the estimation uncertainty term $(R_{I,n} - \hat{R}_{I,n})_2$, its Mack bootstrap version requires also an inflation factor $\sqrt{I+n+1}$ to establish convergence towards a non-degenerate limiting distribution. As this is not the case for the process uncertainty term $(R_{I,n} - \hat{R}_{I,n})_1$ in Theorem 4.1 and its bootstrap version in Theorem 4.8, the process uncertainty terms will asymptotically dominate the predictive roots $R_{I,n} - \hat{R}_{I,n}$ and $R_{I,n}^* - \hat{R}_{I,n}$. Hence, although the limiting bootstrap distribution of $\sqrt{I+n+1}(R_{I,n}^* - \hat{R}_{I,n})_2$ conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$ in Theorem 4.10 does not correctly mimic the corresponding limiting behavior of $\sqrt{I+n+1}(R_{I,n} - \hat{R}_{I,n})_2$ conditional on $Q_{I,n}$ in Theorem 4.4, the whole bootstrap predictive root $R_{I,n}^* - \hat{R}_{I,n}$ still mimics the limiting distribution of the predictive root $R_{I,n} - \hat{R}_{I,n}$ correctly.

Hence, in view of the concepts of asymptotic validity and asymptotic pertinence of a bootstrap prediction approach discussed in Pan and Politis (2016), the Mack bootstrap can be regarded as asymptotically valid, but not as asymptotically pertinent under the stated conditions.

Remark 4.12 (On the asymptotic results for the Mack bootstrap).

- (i) A closer inspection of the decompositions in (4.1) and (4.2) reveals some inconsistencies:
 - While a term based on products of f_j 's is added to and subtracted from $R_{I,n} - \hat{R}_{I,n}$ to get $(R_{I,n} - \hat{R}_{I,n})_1$ and $(R_{I,n} - \hat{R}_{I,n})_2$, a term using products of $\hat{f}_{j,n}^*$'s instead of \hat{f}_j 's, which would be the natural choice, is added to and subtracted from $R_{I,n}^* - \hat{R}_{I,n}$ to get $(R_{I,n}^* - \hat{R}_{I,n})_1$ and $(R_{I,n}^* - \hat{R}_{I,n})_2$.
 - Consequently, while $(R_{I,n} - \hat{R}_{I,n})_1$ relies on products of $F_{i,j}$'s centered around products of f_j 's, its Mack bootstrap version $(R_{I,n}^* - \hat{R}_{I,n})_1$ relies on products of $F_{i,j}^*$'s, which are not (naturally) centered around products of $\hat{f}_{j,n}^*$'s, but around products of $\hat{f}_{j,n}$'s.
 - Moreover, while $(R_{I,n} - \hat{R}_{I,n})_2$ relies on differences between products of development parameters f_j and products of their estimators $\hat{f}_{j,n}$, its Mack bootstrap version $(R_{I,n}^* - \hat{R}_{I,n})_2$ relies on differences between products of bootstrap development factor estimators $\hat{f}_{j,n}^*$ and products of estimators $\hat{f}_{j,n}$. Hence, the sign of $(R_{I,n}^* - \hat{R}_{I,n})_2$ is flipped in comparison to $(R_{I,n} - \hat{R}_{I,n})_2$. This may have a negative effect in finite samples, but as the limiting conditional distribution is Gaussian and hence symmetric, this will not be an issue asymptotically.
 - According to the latter observation, also the terms $(R_{I,n}^* - \hat{R}_{I,n})_2^{(1)}$ and $(R_{I,n}^* - \hat{R}_{I,n})_2^{(2)}$ in the seemingly natural decomposition of the bootstrap estimation uncertainty term in (4.13) are switched in comparison to $(R_{I,n} - \hat{R}_{I,n})_2^{(1)}$ and $(R_{I,n} - \hat{R}_{I,n})_2^{(2)}$.
- (ii) The bootstrap consistency result for the Mack bootstrap process uncertainty part conditional on $Q_{I,n}$ in Theorem 4.8 requires the correct choice of the true family of (conditional) distributions of the $F_{i,j}$'s also for the $F_{i,j}^*$'s in Step 5 of Section 3.1. Otherwise, only the first and second moments of the conditional distribution will be correctly mimicked asymptotically, but not necessarily the whole distribution.
- (iii) The uniform boundedness conditions in Assumption 4.9 are required to establish a Lyapunov condition for bootstrap CLTs for (smooth functions of) $\hat{f}_{j,n}^*$ because the Mack bootstrap draws bootstrap errors $r_{i,j}^*$ from residuals computed from all columns in $D_{I,n}$.
- (iv) The bootstrap inconsistency result for the Mack bootstrap estimation uncertainty part conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$ in Theorem 4.10 is because the bootstrap approach in Step 4 is not taking the condition $Q_{I,n}^* = Q_{I,n}$ into account. Hence, the (always larger!) variance-covariance

matrix $\Sigma_{K, \Pi f_j}$ shows in the conditional limiting distribution instead of $\Sigma_{K, \Pi f_j}^{(2)}$ obtained in Theorem 4.4. Moreover, a decomposition of $(R_{I,n}^* - \hat{R}_{I,n})_2$ resembling the decomposition of $(R_{I,n} - \hat{R}_{I,n})_2$ in (4.7) does not exist.

- (v) While a bootstrap procedure is called asymptotically valid if it mimics correctly the (conditional) predictive distribution in the limit, it should not only mimic the asymptotically dominating part, which captures the prediction (i.e. process) uncertainty. Ideally, it should also mimic as much as possible of the asymptotically negligible parts to achieve a better approximation to the finite sample distribution, which is of actual interest in applications. In the bootstrap literature, such higher-order consistency results are achieved using so-called Edgeworth expansions as discussed in detail in Hall (1992). Hence, in the context of the Mack bootstrap, mimicking also the asymptotically negligible part, that captures the uncertainty due to model parameter estimation, is closely related to the concept coined asymptotic pertinence in Pan and Politis (2016) for time series prediction, which is also discussed by Beutner et al. (2021) from a slightly different perspective.

The discussion above motivates an alternative notion of a Mack-type bootstrap to be introduced in the following section that is designed to eliminate the raised issues. In particular, it should respect the *conditioning on* $Q_{I,n}^* = Q_{I,n}$ and it should generate a whole *bootstrap loss triangle* $D_{I,n}^*$ in a *backward* manner starting from the diagonal $Q_{I,n}^* = Q_{I,n}$. See e.g. Paparoditis and Shang (2021) for bootstrap predictive inference in a functional time series setup.

5. An alternative Mack-type bootstrap scheme

According to the findings and the discussion in Section 4, the original Mack bootstrap proposal is not capable of mimicking the conditional distribution of the estimation uncertainty part correctly. Although it is asymptotically dominated by the process uncertainty part, it is generally desirable to construct a Mack-type bootstrap that addresses this issue to enable a better finite sample performance.

For this purpose, we propose an alternative Mack-type bootstrap in this section to mimic the distribution of the predictive root of the reserve $R_{I,n} - \hat{R}_{I,n}$ using an *alternative* bootstrap predictive root of the reserve $R_{I,n}^+ - \hat{R}_{I,n}^+$ to be defined below. To distinguish it from the original Mack bootstrap proposal in Section 3, we denote all related bootstrap quantities and operations with a “+” instead of a “*”. This novel approach deviates from the original Mack bootstrap scheme from Section 3 in several ways:

- (i) First, given the loss triangle $D_{I,n}$ and conditional on $Q_{I,n}^+ = Q_{I,n}$, where $Q_{I,n}^+ = \{C_{I-i,i}^+ | i = 0, \dots, I+n\}$, a recursive *backward* bootstrap approach is employed to generate a whole bootstrap *upper triangle*

$$D_{I,n}^+ = \left\{ C_{i,j}^+ | i = -n, \dots, I, j = 0, \dots, I+n, -n \leq i+j \leq I \right\}.$$

- (ii) Second, the development factor estimators $\hat{f}_{j,n}$ computed from $D_{I,n}$ are used for a *parametric* bootstrap to construct bootstrap individual development factors $F_{i,j}^+, i = -n, \dots, I, j = 0, \dots, I+n-1$ and $i+j \geq I$, that is for the *lower triangle*, which also allows to construct $R_{I,n}^+$.
- (iii) Third, for the construction of the bootstrap predictive root of the reserve $R_{I,n}^+ - \hat{R}_{I,n}^+$, the bootstrap reserve $R_{I,n}^+$ is *not* centered around its best estimate $\hat{R}_{I,n}$, but around a suitable bootstrap version $\hat{R}_{I,n}^+$.

Analogous to the original Mack bootstrap, the alternative Mack bootstrap is employed to estimate the conditional distribution of $R_{I,n} - \hat{R}_{I,n}$ given $Q_{I,n}$ by the conditional bootstrap distribution of $R_{I,n}^+ - \hat{R}_{I,n}^+$ given $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$.

5.1. An alternative Mack-type bootstrap algorithm

With the upper triangle $D_{I,n}$ at hand, the alternative Mack-type bootstrap algorithm is defined as follows:

- Step 1. Estimate the development factors f_j and the variance parameters σ_j^2 from the data by computing $\hat{f}_{j,n}$ and $\hat{\sigma}_{j,n}^2$ for $j = 0, \dots, I+n-1$ as defined in (2.8) and (2.11), respectively.
- Step 2. Choose a parametric family for the (conditional) bootstrap distributions of the *backward* individual development factors $G_{i,j}^+$ given $C_{i,j+1}^+$ and $D_{I,n}$ such that $G_{i,j}^+ > 0$ a.s. with

$$E^+(G_{i,j}^+ | C_{i,j+1}^+) = \hat{f}_{j,n}^{-1}, \quad Var^+(G_{i,j}^+ | C_{i,j+1}^+) = \frac{\hat{\sigma}_{j,n}^2}{C_{i,j+1}^+}, \quad j = 0, \dots, I+n-1, i = -n, \dots, I-j-1.$$

Then, given $Q_{I,n}^+ = Q_{I,n}$, generate *backwards* a bootstrap loss triangle $D_{I,n}^+$ using the recursion

$$C_{i,j}^+ = C_{i,j+1}^+ G_{i,j}^+, \quad j = 0, \dots, I+n-1, i = -n, \dots, I-j-1.$$

- Step 3. Compute bootstrap development factor estimators by $\hat{f}_{j,n}^+$ for $j = 0, \dots, I+n-1$, which are defined as $\hat{f}_{j,n}$ in (2.8), but are calculated from the bootstrap loss triangle $D_{I,n}^+$. That is, we compute

$$\hat{f}_{j,n}^+ = \frac{\sum_{i=-n}^{I-j-1} C_{i,j+1}^+}{\sum_{i=-n}^{I-j-1} C_{i,j}^+} = \frac{\sum_{i=-n}^{I-j-1} C_{i,j+1}^+}{\sum_{i=-n}^{I-j-1} C_{i,j+1}^+ G_{i,j}^+}, \quad j = 0, \dots, I+n-1. \tag{5.1}$$

- Step 4. Choose a parametric family for the (conditional) bootstrap distributions of $F_{i,j}^+$ given $C_{i,j}^+$ and $D_{I,n}$ such that $F_{i,j}^+ > 0$ a.s. with

$$E^+(F_{i,j}^+ | C_{i,j}^+) = \hat{f}_{j,n}, \quad Var^+(F_{i,j}^+ | C_{i,j}^+) = \frac{\hat{\sigma}_{j,n}^2}{C_{i,j}^+}, \quad i = -n, \dots, I, j = I-i, \dots, I+n-1.$$

Then, given $Q_{I,n}^+ = Q_{I,n}$, generate the bootstrap ultimate claims $C_{i,I+n}^+$ and the reserves $R_{i,I+n}^+ = C_{i,I+n}^+ - C_{i,I-i}^+$ for $i = -n, \dots, I$ using the recursion

$$C_{i,j+1}^+ = C_{i,j}^+ F_{i,j}^+, \quad j = I - i, \dots, I + n - 1. \tag{5.2}$$

Step 5. Compute the bootstrap total reserve $R_{I,n}^+ = \sum_{i=-n}^I R_{i,I+n}^+$ and the alternative Mack bootstrap predictive root of the reserve

$$R_{I,n}^+ - \widehat{R}_{I,n}^+ = \sum_{i=0}^{I+n} C_{I-i,i}^+ \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^+ - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+ \right),$$

where the centering term $\widehat{R}_{I,n}^+$ is a bootstrap version of the best estimate $\widehat{R}_{I,n}$, that is defined by

$$\widehat{R}_{I,n}^+ = \sum_{i=-n}^I C_{i,I-i}^+ \prod_{j=I-i}^{I+n-1} \widehat{f}_{j,n}^+ = \sum_{i=0}^{I+n} C_{I-i,i}^+ \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+. \tag{5.3}$$

Step 6. Repeat Steps 2 - 5 above B times, where B is large, to get $(R_{I,n}^+ - \widehat{R}_{I,n}^+)^{(b)}$, $b = 1, \dots, B$ bootstrap predictive roots, and denote by $q^+(\alpha)$ the α -quantile of their empirical distribution.

Step 7. Construct the $(1 - \alpha)$ equal-tailed prediction interval for $R_{I,n}$ as

$$\left[\widehat{R}_{I,n} + q^+(\alpha/2), \widehat{R}_{I,n} + q^+(1 - \alpha/2) \right].$$

Remark 5.1 (On the alternative Mack-type bootstrap).

- (i) In comparison to the Mack bootstrap from Section 3.1, the bootstrap reserve is not a double-bootstrap quantity anymore, the centering is based on a bootstrap version of the best estimate, and the bootstrap for the upper loss triangle is backwards starting in the diagonal.
- (ii) The conditional distribution for the $G_{i,j}^+ | C_{i,j+1}^+$ can be chosen in different ways. For instance, this can be done non-parametrically similar to Steps 2 - 4 in Section 3.1 or using the parametric family of distributions used in Step 5 in Section 3.1. However, it is crucial to mimic sufficiently well the first and second backward conditional moments, that is, $E(C_{ij}^+ | C_{i,j+1}^+)$ and $Var(C_{ij}^+ | C_{i,j+1}^+)$, respectively.

6. Asymptotic theory for the alternative Mack bootstrap

By adopting the general strategy of Section 4 to investigate the consistency properties of the original Mack bootstrap, the alternative Mack predictive root of the reserve $R_{I,n}^+ - \widehat{R}_{I,n}^+$ can be decomposed also into a prediction error part and an estimation error part, respectively. That is, by adding and subtracting $\sum_{i=0}^{I+n} C_{I-i,i}^+ \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+$, we get

$$\begin{aligned} R_{I,n}^+ - \widehat{R}_{I,n}^+ &= \sum_{i=0}^{I+n} C_{I-i,i}^+ \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^+ - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+ \right) + \sum_{i=0}^{I+n} C_{I-i,i}^+ \left(\prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+ - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+ \right) \\ &=: (R_{I,n}^+ - \widehat{R}_{I,n}^+)_1 + (R_{I,n}^+ - \widehat{R}_{I,n}^+)_2, \end{aligned} \tag{6.1}$$

where $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ and $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2$ are the alternative Mack bootstrap versions of $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$, respectively.

6.1. Conditional bootstrap asymptotics for the alternative Mack bootstrap predictive root

As in Section 4.2 for the Mack bootstrap, we have to check whether, conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$, the alternative Mack bootstrap quantities $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ and $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2$ in (4.2) are correctly mimicking the limiting distributions of $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$ given $Q_{I,n}$, respectively.

6.1.1. Conditional bootstrap asymptotics for reserve prediction: process uncertainty

The process uncertainty part $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ of the alternative Mack bootstrap differs from the $(R_{I,n}^* - \widehat{R}_{I,n}^*)_1$ as the $F_{i,j}^+$'s in $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ use $\widehat{f}_{j,n}$ instead of $\widehat{f}_{j,n}^*$ and as $\prod_{j=i}^{I+n-1} F_{I-i,j}^+$ is centered around $\prod_{j=i}^{I+n-1} \widehat{f}_{j,n}$ instead of $\prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^*$ accordingly. However, by using very similar arguments, we get the same limiting result also for the process uncertainty part $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ of the alternative Mack bootstrap.

Theorem 6.1 (Bootstrap asymptotics for $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$). Suppose Assumptions 2.2, 2.3, 2.4, 2.5, 4.6 and 4.7 (for $F_{i,j}^+$ instead of $F_{i,j}^*$) hold. Then, as $n \rightarrow \infty$, conditionally on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$, $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_1$ converges in distribution to $\mathcal{C}_1 | Q_{I,\infty}$ in probability, which is the (limiting) distribution of $(R_{I,\infty} - \widehat{R}_{I,\infty})_1 | Q_{I,\infty}$ according to (4.6) described in Theorem 4.1. Moreover, for all $n \in \mathbb{N}_0$, it holds $E^+((R_{I,n}^+ - \widehat{R}_{I,n}^+)_1 | Q_{I,n}^+ = Q_{I,n}) = 0$ and, for $n \rightarrow \infty$, we have

$$Var^+ \left((R_{I,n}^+ - \widehat{R}_{I,n}^+)_1 | Q_{I,n}^+ = Q_{I,n} \right) \longrightarrow Var \left((R_{I,\infty} - \widehat{R}_{I,\infty})_1 | Q_{I,\infty} \right) \quad \text{in prob.}, \tag{6.2}$$

where $Var((R_{I,\infty} - \widehat{R}_{I,\infty})_1 | Q_{I,\infty}) = O_p(1)$ as given in (4.5). Consequently, as $n \rightarrow \infty$, we have

$$d_2 \left(\mathcal{L} \left((R_{I,n} - \widehat{R}_{I,n})_1 | Q_{I,n} \right), \mathcal{L}^+ \left((R_{I,n}^+ - \widehat{R}_{I,n}^+)_1 | Q_{I,n}^+ = Q_{I,n} \right) \right) \longrightarrow 0 \quad \text{in prob.}$$

6.1.2. Conditional bootstrap asymptotics for reserve prediction: estimation uncertainty

In view of the decomposition of $(R_{I,n} - \widehat{R}_{I,n})_2$ in (4.7), conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$, its alternative Mack bootstrap counterpart $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2$ can be also decomposed further. That is, we have

$$\begin{aligned} & (R_{I,n}^+ - \widehat{R}_{I,n}^+)_2 \\ &= \sum_{i=0}^{I+n} C_{I-i,i}^+ \left(\prod_{j=i}^{I+n-1} \widehat{f}_{j,n} - \prod_{j=i}^{I+n-1} f_{j,n}^+(Q_{I,n}) \right) + \sum_{i=0}^{I+n} C_{I-i,i}^+ \left(\prod_{j=i}^{I+n-1} f_{j,n}^+(Q_{I,n}) - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^+ \right) \\ &= (R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(1)} + (R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(2)}, \end{aligned} \tag{6.3}$$

where $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(1)}$ is measurable with respect to $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$ and $f_{j,n}^+(Q_{I,n}) := \mu_{j+1,n}^{+(1)}(Q_{I,n})/\mu_{j,n}^{+(2)}(Q_{I,n})$ with $\mu_{j+1,n}^{+(1)}(Q_{I,n}) := E^+(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j+1}^+ | Q_{I,n}^+ = Q_{I,n})$ as well as $\mu_{j,n}^{+(2)}(Q_{I,n}) := E^+(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j}^+ | Q_{I,n}^+ = Q_{I,n})$.

In comparison to Theorem 4.10, the uniform boundedness condition in Assumption 4.9 can be dropped, but the derivation of (conditional) bootstrap asymptotic theory and consistency results for $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2$ requires additional assumptions on the backward individual development factors $G_{i,j}^+$ from Step 3 in Section 5.1. Precisely, it has to be guaranteed that the backward conditional mean $E(C_{i,j} | C_{i,j+1})$ and the backward conditional variance $Var(C_{i,j} | C_{i,j+1})$ are consistently mimicked by their alternative Mack bootstrap counterparts $E^+(C_{i,j}^+ | C_{i,j+1}^+)$ and $Var^+(C_{i,j}^+ | C_{i,j+1}^+)$, respectively, such that the corresponding limiting distributions obtained in Steinmetz and Jentsch (2022, Theorem C.1) are correctly mimicked.

Assumption 6.2 (Consistent estimation of backward moments). For $n \rightarrow \infty$, suppose that the (conditional) bootstrap distributions of the backward individual development factors $G_{i,j}^+$, $j = 0, \dots, I + n - 1$ and $i = -n, \dots, I - j - 1$ given $C_{i,j+1}^+$ and $D_{I,n}$ are chosen in Step 2 in Section 5.1 such that:

- (i) For each fixed $K \in \mathbb{N}_0$, let $\underline{f}_{K,n}^+(Q_{I,n}) = (f_{0,n}^+(Q_{I,n}), f_{1,n}^+(Q_{I,n}), \dots, f_{K,n}^+(Q_{I,n}))'$ and define $\widehat{\underline{f}}_{K,n} = (\widehat{f}_{0,n}, \widehat{f}_{1,n}, \dots, \widehat{f}_{K,n})'$. Then, conditional on $Q_{I,n}^+ = Q_{I,n}$, we have

$$J_n^{1/2} \left(\underline{f}_{K,n}^+(Q_{I,n}) - \widehat{\underline{f}}_{K,n} \right) | \left(Q_{I,n}^+ = Q_{I,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K,f}^{(1)} \right),$$

where $J_n^{1/2} = \text{diag} \left(\sqrt{I+n-j}, j = 0, \dots, K \right)$ is a diagonal $(K+1) \times (K+1)$ matrix of inflation factors and the variance-covariance matrix $\Sigma_{K,f}^{(1)}$ is defined in Theorem C.1 in Steinmetz and Jentsch (2022).

- (ii) For each fixed $K \in \mathbb{N}_0$, let $\widehat{\underline{f}}_{K,n}^+ = (\widehat{f}_{0,n}^+, \widehat{f}_{1,n}^+, \dots, \widehat{f}_{K,n}^+)$. Then, conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$, we have

$$J_n^{1/2} \left(\widehat{\underline{f}}_{K,n}^+ - \underline{f}_{K,n}^+(Q_{I,n}) \right) | \left(Q_{I,n}^+ = Q_{I,n}, D_{I,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K,f}^{(2)} \right) \quad \text{in prob.}$$

where the variance-covariance matrix $\Sigma_{K,f}^{(2)}$ is defined in Theorem C.1 in Steinmetz and Jentsch (2022).

In concordance to the derivation of the conditional limiting result obtained in Theorem 4.4(ii), which relies on conditional CLTs for the development factor estimators $\widehat{f}_{j,n}$ given in (Steinmetz and Jentsch 2022, Appendix C), the conditional bootstrap CLTs in Assumption 6.2 allow to state the following theorem, which provides the limiting distribution of the alternative Mack bootstrap estimation uncertainty term $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2$ conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$. Precisely, while $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(1)}$ is measurable with respect to $D_{I,n}$, Assumption 6.2 allows to establish asymptotic normality of $\sqrt{I+n+1}(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(2)}$ conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$.

Theorem 6.3 (Bootstrap asymptotics for $(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2$ conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$). Suppose Assumptions 2.2, 2.3, 2.5, 2.4, 4.2 and 6.2 hold. Then, as $n \rightarrow \infty$, the following holds:

- (i) Conditional on $Q_{I,n}^+ = Q_{I,n}$, $\sqrt{I+n+1}(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(1)}$ converges in distribution to the non-degenerate limiting distribution $\mathcal{G}_2^{(1)}$. That is, we have

$$\sqrt{I+n+1}(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(1)} | \left(Q_{I,n}^+ = Q_{I,n} \right) \xrightarrow{d} \langle Q_{I,\infty}, \mathbf{Y}_\infty^{(1)} \rangle \sim \mathcal{G}_2^{(1)}, \tag{6.4}$$

where $\mathbf{Y}_\infty^{(1)} = (Y_i^{(1)}, i \in \mathbb{N}_0)$ denotes a centered Gaussian process with covariances

$$\text{Cov}(Y_{i_1}^{(1)}, Y_{i_2}^{(1)}) = \lim_{K \rightarrow \infty} \Sigma_{K, \prod f_j}^{(1)}(i_1, i_2), \quad i_1, i_2 \in \mathbb{N}_0,$$

where $\Sigma_{K, \prod f_j}^{(1)}(i_1, i_2)$ is defined in Corollary C.2 in Steinmetz and Jentsch (2022). Here, the sequences $Q_{I,\infty}$ and $\mathbf{Y}_\infty^{(1)}$ are independent.

- (ii) Conditionally on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$, $\sqrt{I+n+1}(R_{I,n}^+ - \widehat{R}_{I,n}^+)_2^{(2)}$ converges in distribution to $\mathcal{G}_2 | Q_{I,\infty}$ in probability, where $\mathcal{G}_2 | Q_{I,\infty} \sim \mathcal{N}(0, \Xi(Q_{I,\infty})) | Q_{I,\infty}$ is the (conditional) limiting distribution obtained in Theorem 4.4(ii).

Consequently, as $n \rightarrow \infty$, we have

$$d_K \left(\mathcal{L} \left((R_{I,n} - \widehat{R}_{I,n})_2 | Q_{I,n} \right), \mathcal{L}^+ \left((R_{I,n}^+ - \widehat{R}_{I,n}^+)_2 | Q_{I,n}^+ = Q_{I,n} \right) \right) \rightarrow 0 \quad \text{in prob.}$$

6.1.3. Conditional bootstrap asymptotics for the whole predictive root of the reserve

As in Sections 4.1.3 and 4.2.3, combining the results for $(R_{I,n}^+ - \hat{R}_{I,n}^*)_1$ and $(R_{I,n}^+ - \hat{R}_{I,n}^*)_2$ from Theorems 6.1 and 6.3, we get joint asymptotics for $R_{I,n}^+ - \hat{R}_{I,n}^+$ conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$.

Theorem 6.4 (Bootstrap asymptotics for $R_{I,n}^+ - \hat{R}_{I,n}^+$ conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$). Suppose the assumptions of Theorems 6.1 and 6.3 hold. Then, conditional on $Q_{I,n}^+ = Q_{I,n}$ and $D_{I,n}$, $(R_{I,n}^+ - \hat{R}_{I,n}^*)_1$ and $(R_{I,n}^+ - \hat{R}_{I,n}^*)_2$ are stochastically independent, and $R_{I,n}^+ - \hat{R}_{I,n}^+$ converges in distribution to $G_1 | Q_{I,\infty}$ in probability. That is, we have

$$R_{I,n}^+ - \hat{R}_{I,n}^+ | (Q_{I,n}^+ = Q_{I,n}, D_{I,n}) = (R_{I,n}^+ - \hat{R}_{I,n}^*)_1 + (R_{I,n}^+ - \hat{R}_{I,n}^*)_2 | (Q_{I,n}^+ = Q_{I,n}, D_{I,n}) \xrightarrow{d} G_1 | Q_{I,\infty}$$

in probability.

According to the discussion below Theorem 4.11 and in view of the concepts of asymptotic validity and asymptotic pertinence of bootstrap predictive inference in Pan and Politis (2016), the alternative Mack bootstrap can be regarded as asymptotically valid and asymptotically pertinent under the stated conditions.

Remark 6.5 (Backward vs. forward bootstrapping). While a backward bootstrap appears to be natural in time series setups addressed in Pan and Politis (2016), they also propagate a simpler forward bootstrap to capture the estimation uncertainty in bootstrap prediction. Asymptotically, in their setup, both approaches are indeed equivalent due to the intrinsic stationarity assumption. However, in Mack’s Model setup considered here, this is not the case and the (fixed-design) forward bootstrap of England and Verrall (2006) does not correctly capture the conditional limiting distribution of the estimation uncertainty part.

7. Simulation study

In this section, we compare the original Mack bootstrap from Section 3 and the alternative Mack bootstrap from Section 5 to illustrate our theoretical findings from Sections 4 and 6 by means of simulations of several parameter scenarios. Additionally, we simulate a Mack-type bootstrap, which uses a forward bootstrap approach in Step 2 of Section 5.1, but coincides otherwise with the alternative Mack bootstrap.

7.1. Simulation setup

To assure comparability, according to the setup introduced in Section 2, we pick up the simulation setup employed in Steinmetz and Jentsch (2022). Although the theory derived in this paper provides asymptotic results only for the (more general) conditional distribution of the predictive root of the reserve, in the following, we show simulation results in terms of actual coverage rates of unconditional and conditional bootstrap-based prediction intervals of nominal 95% level. This is because the simulation setups for the unconditional case are straightforward, while those for the conditional case are not. Next, we discuss both setups in detail.

7.1.1. Simulation setup for the unconditional predictive distribution

In the notion of the asymptotic framework introduced in Section 2.1, let $I = 10$ and choose $n \in \{0, 10, 20, 30, 40\}$ leading to effective number of accident years $I + n + 1 \in \{11, 21, 31, 41, 51\}$. For each n and for different parameter scenarios specified below, we generate $M = 500$ cumulative claim matrices $C_{I,n}^{(m)} = \{C_{i,j}^{(m)} | i = -n, \dots, I, j = 0, \dots, I + n\}$, $m = 1, \dots, 500$ consisting of (upper) loss triangles $D_{I,n}^{(m)} = \{C_{i,j}^{(m)} | i = -n, \dots, I, j = 0, \dots, I - i\}$ (having diagonals $Q_{I,n}^{(m)}$), that will serve as the data basis, and their corresponding unobserved (lower) triangles $D_{I,n}^{c(m)} = \{C_{i,j}^{(m)} | i = -n, \dots, I, j = I - i + 1, \dots, I + n\}$. For this purpose, we generate the entries in their first columns $C_{\cdot,0}^{(m)}$ (independently) from a uniform distribution and for the conditional distribution of the individual developments factors $F_{i,j}$ given $C_{i,j}$, that is, for

$$F_{i,j} | C_{i,j} \sim \left(f_j, \frac{\sigma_j^2}{C_{i,j}} \right), \quad i = -n, \dots, I, j = 0, \dots, I + n, \tag{7.1}$$

we use one of the following parametric families of distributions:

- (DGP1) gamma distribution,
- (DGP2) log-normal distribution,
- (DGP3) left-tail truncated normal distribution (truncated at 0.1).

In all scenarios, the development factors and the variance parameters fulfill $f_j > 1$ and $\sigma_j^2 > 0$ with f_j and σ_j^2 decreasing (sufficiently fast) to 1 and 0, respectively. Precisely, we use exponentially decreasing sequences $(f_j)_{j \in \mathbb{N}_0}$ and $(\sigma_j^2)_{j \in \mathbb{N}_0}$ with $f_j = 1 + e^{-1-0.2j}$ and $\sigma_j^2 = 509,518 \cdot e^{-1-0.7j}$. Moreover, we distinguish between two Setups a) and b), where the parameters are exactly the same in both cases, but the first column $C_{\cdot,0} = (C_{-n,0}, \dots, C_{I,0})'$ of the (upper) loss triangle is uniformly distributed on $[120 \times 10^6, 350 \times 10^6]$ in Setup a) and on $[120 \times 10^4, 350 \times 10^4]$ in Setup b). As the results for both setups are similar, we show only those for Setup a) here and report the results for Setup b) in the appendix.

Finally, to evaluate their finite sample performances in terms of coverage rates, for each setup (DGP1)-(DGP3) above and for each loss triangle $D_{I,n}^{(m)}$, $m = 1, \dots, 500$, we perform three different Mack-type bootstraps based on 10,000 bootstrap replications each to construct unconditional bootstrap prediction intervals for the corresponding (true) reserve $R_{I,n}^{(m)}$ computed from $D_{I,n}^{(m)}$ and $D_{I,n}^{c(m)}$. In each case, we collect the coverage rates and the average lengths of the prediction intervals, which are both reported in Table 3 for Setup a). Concerning the different Mack-type bootstraps, we apply the following three approaches:

Table 3

Coverage rates and average interval lengths (both in thousand) for *unconditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) for different (true and bootstrap) parametric families, for $I = 10$ and different n in Setup a).

		oMB		iMB		aMB	
		true parametric family: gamma					
bootstrap parametric family	n	coverage rate	mean length	coverage rate	mean length	coverage rate	mean length
gamma	10	93%	219,466	93%	223,971	95%	218,834
	20	94%	259,173	94%	261,455	95%	257,277
	30	94%	258,372	94%	257,452	95%	255,836
	40	95%	256,274	96%	256,659	95%	255,228
	50	96%	255,725	96%	254,992	95%	253,689
log-normal	10	96%	218,538	96%	218,652	92%	214,489
	20	97%	258,830	97%	259,848	94%	257,321
	30	97%	253,692	96%	253,977	95%	252,820
	40	96%	252,651	96%	252,810	94%	250,371
	50	97%	252,652	96%	251,524	95%	250,789
truncated normal	10	96%	240,822	97%	241,649	95%	236,048
	20	98%	261,582	98%	262,596	94%	260,702
	30	97%	257,797	97%	258,689	94%	257,514
	40	97%	254,185	97%	254,511	95%	253,715
	50	98%	251,655	97%	251,793	95%	250,660
true parametric family: log-normal							
gamma	10	95%	218,660	95%	223,069	95%	218,457
	20	96%	260,833	97%	263,322	95%	259,962
	30	96%	256,580	96%	257,444	95%	255,694
	40	96%	255,190	96%	255,038	95%	252,441
	50	95%	254,618	96%	253,993	95%	253,270
log-normal	10	96%	219,631	95%	220,068	95%	217,652
	20	97%	259,538	94%	261,318	94%	258,688
	30	96%	261,415	96%	257,238	92%	256,082
	40	96%	255,460	97%	255,830	95%	252,970
	50	95%	256,653	95%	254,952	95%	254,117
truncated normal	10	90%	220,822	90%	223,233	92%	218,041
	20	91%	262,582	91%	264,185	90%	261,458
	30	91%	259,797	92%	261,309	91%	258,845
	40	92%	254,185	90%	254,155	92%	253,369
	50	91%	251,655	91%	251,048	93%	250,056
true parametric family: truncated normal							
gamma	10	96%	219,413	96%	225,086	94%	219,368
	20	97%	264,707	97%	265,887	93%	263,748
	30	97%	256,867	95%	257,283	94%	255,583
	40	96%	253,522	96%	254,215	95%	253,929
	50	94%	253,747	96%	252,074	95%	252,374
log-normal	10	94%	216,148	94%	219,752	93%	214,597
	20	96%	258,819	96%	262,787	95%	258,786
	30	97%	255,696	96%	255,762	95%	254,451
	40	97%	253,746	96%	253,772	95%	253,218
	50	95%	252,812	97%	251,879	95%	251,152
truncated normal	10	98%	218,204	99%	219,216	94%	218,126
	20	98%	264,476	98%	266,306	93%	263,599
	30	97%	259,273	97%	259,449	94%	258,405
	40	97%	258,576	97%	258,777	95%	256,373
	50	96%	257,198	97%	257,155	95%	255,792

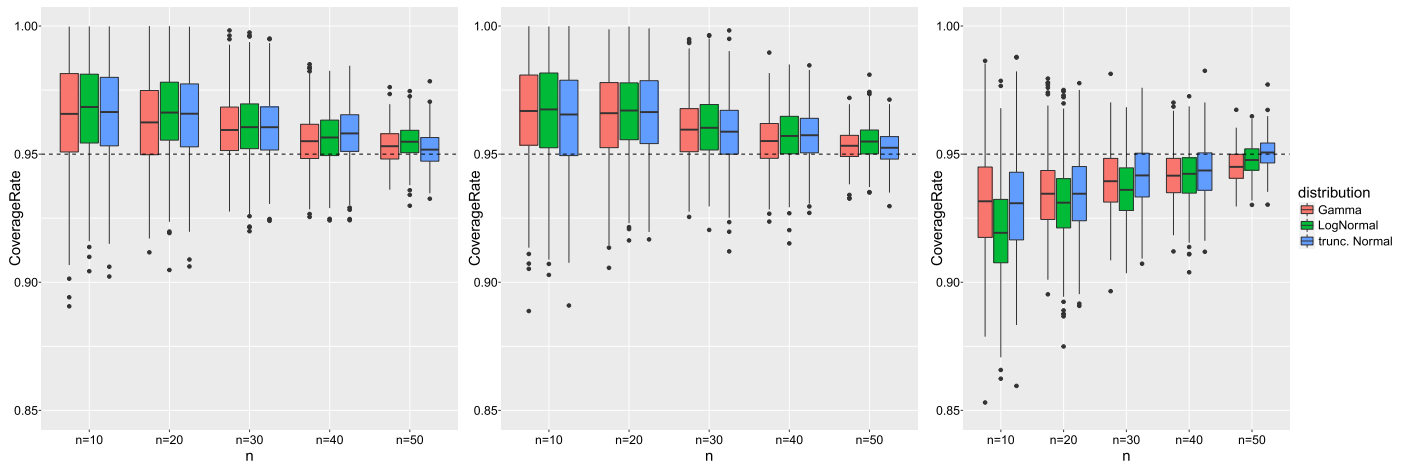


Fig. 1. Box plots of coverage rates for conditional 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) with gamma (red), log-normal (green) and truncated normal (blue), respectively. The true parametric family is gamma for $I = 10$ and different n in Setup a). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

- (oMB) original Mack bootstrap (from Section 3),
- (aMB) alternative Mack-type bootstrap (from Section 5),
- (iMB) intermediate Mack-type bootstrap (using a forward bootstrap in Step 2 of Section 5).

The intermediate Mack-type bootstrap is included to distinguish between the effects caused by the backward resampling proposed in Step 2 and by the different centering used in Step 5 of Section 5.1 on the finite sample performance. For this purpose, we introduce a novel centering term $\hat{R}_{I,n}^{++}$ defined by

$$\hat{R}_{I,n}^{++} = \sum_{i=0}^{I+n} C_{I-i,i}^+ \prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* \tag{7.2}$$

which deviates from $\hat{R}_{I,n}^+$ in (5.3) as it relies on $\hat{f}_{j,n}^*$ in (3.3), but is based on (parametrically generated)

$$F_{i,j}^* | C_{i,j}^*, D_{I,n} \sim \left(\hat{f}_{j,n}^*, \frac{\hat{\sigma}_{j,n}^2}{C_{i,j}^*} \right), \tag{7.3}$$

instead of $\hat{f}_{j,n}^+$ defined in (5.1). This choice of the centering term still resembles the decomposition in (6.1), that shares the (sign) properties of (4.1), which is not the case for (4.2). For all bootstraps, whenever a parametric distribution is used to generate the upper bootstrap loss triangle, we choose the same parametric distribution family used already for the lower triangle (to generate $R_{I,n}^+$ and $R_{I,n}^*$). However, as we do not know the correct parametric family of distributions of the $F_{i,j}$'s, we make use of all three parametric families in (DGP1)-(DGP3) for all three bootstrap approaches, respectively, to also investigate the effect of a misspecified parametric family of distributions to generate $R_{I,n}^*$ and $R_{I,n}^+$.

7.1.2. Simulation setup for the conditional predictive distribution

In contrast to the evaluation of the unconditional performance in terms of coverage rates of prediction intervals in Section 7.1.1, the simulation setup to evaluate the conditional performance given the diagonal $Q_{I,n}^{(m)}$, is not straightforward. This is because, for this purpose, we would have to generate (upper) loss triangles $D_{I,n}^{(m)}$ according to Mack's model from Section 2.2 that all have the same diagonal $Q_{I,n}^{(m)}$. To achieve this, starting with $Q_{I,n}^{(m)}$, we would need the true backward conditional distribution $C_{i,j}$ given $C_{i,j+1}$, but this is usually not possible to get analytically and closed-form expressions are already difficult to obtain in simple toy examples; see Example 4.9 in Steinmetz and Jentsch (2022).

Hence, as a workaround, we come up with an approximate solution to that problem that makes use of the idea of the alternative Mack-type bootstrap from Section 5. For each diagonal $Q_{I,n}^{(m)}$, $m = 1, \dots, 500$ obtained in Section 7.1.1, we generate backwards new (upper) loss triangles $\tilde{D}_{I,n}^{(m,s)}$, $s = 1, \dots, 1000$. For this purpose, we use the same parametric family (that was used for $F_{i,j}$ given $C_{i,j}$ in (7.1) to generate $D_{I,n}^{(m)}$) to simulate backward individual development factors $G_{i,j}$ given $C_{i,j+1}$ (to define $C_{i,j} = G_{i,j} C_{i,j+1}$ for the upper loss triangle $\tilde{D}_{I,n}^{(m,s)}$) having the conditional distribution

$$G_{i,j} | C_{i,j+1} \sim \left(\frac{1}{f_j}, \frac{\sigma_j^2}{C_{i,j+1}} \right), \quad j = 0, \dots, I+n-1, \quad i = -n, \dots, I-j-1 \tag{7.4}$$

making use of the true parameters f_j and σ_j^2 specified above. Again using the same parametric family, for each of these upper loss triangles $\tilde{D}_{I,n}^{(m,s)}$, $s = 1, \dots, 1000$, we generate an unobserved (lower) triangle $\tilde{D}_{I,n}^{c(m,s)}$ using individual development factors $F_{i,j}$ according to (7.1) (to define $C_{i,j+1} = F_{i,j} C_{i,j}$ for the lower loss triangle $\tilde{D}_{I,n}^{c(m,s)}$). Altogether, this results in cumulative claim matrices $\tilde{C}_{I,n}^{(m,s)} = \tilde{D}_{I,n}^{(m,s)} \cup \tilde{D}_{I,n}^{c(m,s)}$, $m = 1, \dots, 500$, $s = 1, \dots, 1000$ such that, for each fixed s , the diagonals of all (upper) loss triangles $\tilde{D}_{I,n}^{(m,s)}$, $m = 1, \dots, 500$ coincide and are equal to $Q_{I,n}^{(m)}$. Hence, this enables the evaluation of the finite sample performance of bootstrap-based conditional prediction intervals in terms of coverage rates and average lengths by applying the approach from Section 7.1.1 for all $s = 1, \dots, 1000$ to $\tilde{C}_{I,n}^{(m,s)}$, $m = 1, \dots, 500$. Results for Setup a) are reported in Figs. 1, 2 and

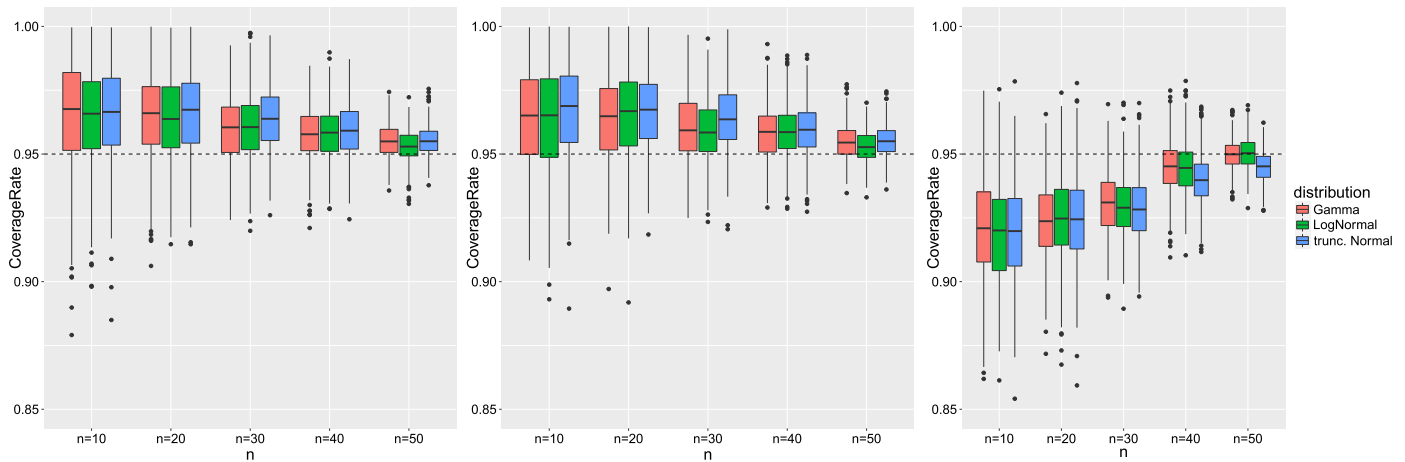


Fig. 2. Box plots of coverage rates for *conditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) with gamma (red), log-normal (green) and truncated normal (blue), respectively. The true parametric family is log-normal for $I = 10$ and different n in Setup a).

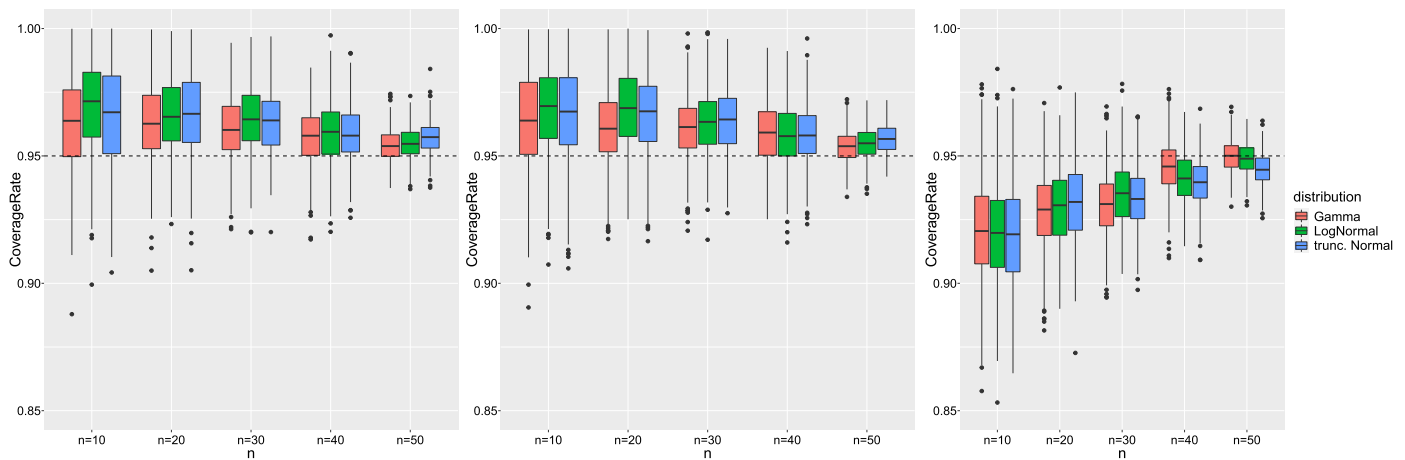


Fig. 3. Box plots of coverage rates for *conditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) with gamma (red), log-normal (green) and truncated normal (blue), respectively. The true parametric family is truncated normal for $I = 10$ and different n in Setup a).

3 for true distributions gamma, log-normal and truncated normal, respectively. Results for Setup b) can be found in Figs. 4, 5, 6 and Table 4 in the appendix.

7.2. Simulation results

First, we consider the simulation results for the *unconditional* bootstrap prediction intervals reported in Table 3. We find that, in most cases, the original Mack bootstrap as well as the intermediate Mack bootstrap lead to (comparably large) coverage rates that are larger than the envisaged 95%, while the alternative Mack bootstrap comes generally close to 95%. This is the case irrespective of the size of the observed (upper) loss triangle and irrespective of using the correctly specified parametric family of distributions or not. The only exception is the case, where the true (conditional) distribution of the individual development factors is a log-normal and a truncated normal is used for bootstrapping and all three methods lead to undercoverage. Moreover, in almost all scenarios, the average lengths of the unconditional prediction intervals of the alternative Mack bootstrap are smaller than those of the original and the intermediate Mack bootstraps. With the average lengths of the latter two bootstraps of comparable size, the difference to the alternative Mack bootstrap vanishes with increasing n , that is, with increasing loss triangle size $I + n + 1$.

These results perfectly agree to the findings of Theorem 4.10, where the (conditional) variance $\Xi(Q_{I,\infty})$, which is mimicked by the original Mack bootstrap and by the intermediate Mack-type bootstrap, is generally larger than the variance $\Xi(Q_{I,\infty})$ found in Theorem 4.4, which is mimicked by the alternative Mack bootstrap correctly according to Theorem 6.3. The finding that the average interval lengths of the original/intermediate Mack bootstrap and the alternative Mack bootstrap get closer for increasing n is well explained by the fact that their asymptotically dominating process uncertainty parts coincide, while the impact of the asymptotically negligible estimation uncertainty parts vanishes.

Next, we consider the simulation results for the *conditional* bootstrap prediction intervals reported in Figs. 1, 2 and 3 for DGP1-3. According to the simulation setup described in Section 7.1.2, we obtain coverage rates for *each* diagonal $Q_{I,n}^{(m)}$, $m = 1, \dots, 500$ as generated in Section 7.1.1. These 500 coverage rates are presented in Fig. 1 using box plots. Similar to the results obtained in the unconditional case, the original Mack bootstrap and the intermediate Mack bootstrap perform comparably and tend to overcoverage, while the alternative Mack bootstrap tends to undercoverage. For all three bootstrap approaches, the tendency to overcoverage or undercoverage vanishes for increasing loss triangle size $I + n + 1$, that is, all coverage rates get closer to the 95% level with interquartile ranges decreasing as well. Not reported in Fig. 1, in all scenarios and for all diagonals

$Q_{I,n}^{(m)}$, $m = 1, \dots, 500$, the average lengths of the conditional bootstrap prediction intervals are smaller for the alternative Mack bootstrap than for the original/intermediate Mack bootstrap. Again, this fits well to our theory derived in Sections 4 and 6.

With the original Mack bootstrap and intermediate Mack bootstrap showing a similar performance in terms of coverage rates and average interval lengths in most cases, correctly mimicking the (asymptotically vanishing) estimation uncertainty part (achieved by the alternative Mack bootstrap) appears to be more crucial in finite samples than using the natural centering (used for intermediate/alternative Mack bootstrap).

8. Conclusion

In this paper, we adopt the stochastic and asymptotic framework that was proposed by Steinmetz and Jentsch (2022) to derive asymptotic theory in Mack’s model, also for investigating the consistency properties of the Mack bootstrap proposal.

For this purpose, the (conditional) asymptotic theory derived in Steinmetz and Jentsch (2022) serves well as benchmark results for the Mack bootstrap approximations. By splitting the predictive root of the reserve into two additive parts corresponding to process and estimation uncertainty, our approach enables - for the first time - a rigorous investigation of the validity of the Mack bootstrap. We prove that the (conditional) distribution of the asymptotically dominating process uncertainty part is correctly mimicked by the Mack bootstrap if the parametric family of distributions of the individual development factors is correctly specified in Mack’s bootstrap. Otherwise, this will be generally not the case. In contrast, the corresponding (conditional) distribution of the estimation uncertainty part is generally not correctly captured by the bootstrap. Altogether, as the process uncertainty part dominates asymptotically, this proves asymptotic validity of the Mack bootstrap for the whole predictive root of the reserve. However, it also proves that asymptotic pertinence in the sense of Pan and Politis (2016) does not hold.

To remedy this, we propose a more natural alternative Mack-type bootstrap, that uses a different centering and that is designed to capture correctly also the (conditional) distribution of the estimation uncertainty part by using a backward resampling approach. Under suitable assumptions, we demonstrate that the newly proposed alternative Mack-type bootstrap can be indeed asymptotically valid and pertinent.

Our findings are illustrated by simulations, which show that the alternative Mack-type bootstrap tends to perform superior to the original Mack bootstrap in finite samples. An intermediate Mack-type bootstrap provides evidence that the backward resampling is mainly responsible for this improvement.

CRediT authorship contribution statement

Julia Steinmetz: Writing – original draft. Carsten Jentsch: Writing – review & editing.

Declaration of competing interest

None declared.

Data availability

No data was used for the research described in the article.

Appendix A. Auxiliary results for Section 4

A.1. Mack bootstrap asymptotics for parameter estimators

The following theorem is the Mack bootstrap version of the (unconditional!) Theorem 3.1 in Steinmetz and Jentsch (2022) adapted to the asymptotic framework of Section 2.1.

Theorem A.1 (Asymptotic normality of $\hat{f}_{j,n}^*$ conditional on $D_{I,n}$). Suppose Assumptions 2.2, 2.3, 2.5 and 4.9 are satisfied and let $\hat{f}_{j,n}^*$, $j = 0, \dots, I + n - 1$ be defined as in (3.5) according to the Mack bootstrap scheme of Section 3.1. Then, as $n \rightarrow \infty$, the following holds:

(i) For each fixed $j \in \mathbb{N}_0 = \{0, 1, 2, \dots\}$, we have

$$\sqrt{I + n - j} \left(\hat{f}_{j,n}^* - \hat{f}_{j,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \frac{\sigma_j^2}{\mu_j} \right) \text{ in probability,}$$

where “ \xrightarrow{d} ” denotes convergence in distribution.

(ii) For each fixed $K \in \mathbb{N}_0$, let $\underline{\hat{f}}_{K,n}^* = (\hat{f}_{0,n}^*, \hat{f}_{1,n}^*, \dots, \hat{f}_{K,n}^*)'$ be the $(K + 1)$ -dimensional Mack bootstrap version of $\underline{\hat{f}}_{K,n} = (\hat{f}_{0,n}, \hat{f}_{1,n}, \dots, \hat{f}_{K,n})'$. Then, we have

$$J^{1/2} \left(\underline{\hat{f}}_{K,n}^* - \underline{\hat{f}}_{K,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \underline{\Sigma}_{K,\underline{f}} \right) \text{ in probability,}$$

where $J^{1/2} = \text{diag} \left(\sqrt{I + n + 1 - j}, j = 0, \dots, K \right)$ is a diagonal $(K + 1) \times (K + 1)$ matrix of inflation factors and $\underline{\Sigma}_{K,\underline{f}} = J_g(\underline{\mu}_K) \underline{\Sigma}_{K,\underline{C}} J_g(\underline{\mu}_K)' =$

$\text{diag} \left(\frac{\sigma_0^2}{\mu_0}, \frac{\sigma_1^2}{\mu_1}, \dots, \frac{\sigma_K^2}{\mu_K} \right)$ is a diagonal $(K + 1) \times (K + 1)$ covariance matrix, where

$$\underline{\Sigma}_{K,\underline{C}} = \text{Cov}(\underline{C}_{i,K}) = \left(\begin{matrix} \left(\prod_{k=\min(j_1,j_2)}^{\max(j_1,j_2)-1} f_k \right) \tau_{\min(j_1,j_2)}^2 \\ j_1, j_2 = 0, \dots, K + 1 \end{matrix} \right) \tag{A.1}$$

is a $(K + 1) \times (K + 1)$ matrix,

$$J_g(\underline{x}) = \begin{pmatrix} -\frac{x_1}{x_0^2} & \frac{1}{x_0} & 0 & \dots & 0 \\ 0 & -\frac{x_2}{x_1^2} & \frac{1}{x_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\frac{x_{K+1}}{x_K^2} & \frac{1}{x_K} \end{pmatrix} \tag{A.2}$$

is a $(K + 1) \times (K + 2)$ matrix, and $\underline{\mu}_{-K} = (\mu_0, \dots, \mu_{K+1})'$ as derived in the proof of Theorem 3.1 in Steinmetz and Jentsch (2022).

As the *unconditional* limiting distributions obtained in Theorem A.1 above and in Theorem 3.1 in Steinmetz and Jentsch (2022) coincide, the Mack bootstrap is *unconditionally*, that is without conditioning on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$, consistent for an arbitrary, but fixed number of estimators of development factors. That is, for each fixed $K \in \mathbb{N}_0$, we have

$$d_K \left(\mathcal{L}^* \left(J^{1/2} \left(\widehat{\underline{f}}_{-K,n}^* - \widehat{\underline{f}}_{-K,n} \right) \right), \mathcal{L} \left(J^{1/2} \left(\widehat{\underline{f}}_{-K,n} - \underline{f}_{-K} \right) \right) \right) = o_P(1),$$

where $\underline{f}_{-K} = (f_0, f_1, \dots, f_K)'$ and d_K denotes the Kolmogorov distance between two probability distributions.

The following direct corollary is the Mack bootstrap version of Corollary 3.2 in Steinmetz and Jentsch (2022) adapted to the asymptotic framework of Section 2.1.

Corollary A.2 (Asymptotic normality for products of $\widehat{f}_{j,n}^*$'s conditional on $D_{I,n}$). Suppose the assumptions of Theorem A.1 hold. Then, as $n \rightarrow \infty$, the following holds:

(i) For each fixed $K \in \mathbb{N}_0$ and $i = 0, \dots, K$, we have

$$\sqrt{I+n+1} \left(\prod_{j=i}^K \widehat{f}_{j,n}^* - \prod_{j=i}^K \widehat{f}_{j,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \sum_{j=i}^K \frac{\sigma_j^2}{\mu_j} \prod_{l=i, l \neq j}^K f_l^2 \right) \text{ in probability.}$$

(ii) For each fixed $K \in \mathbb{N}_0$, we have also joint convergence, that is,

$$\sqrt{I} \left(\prod_{i=0, \dots, K}^K \widehat{f}_{j,n}^* - \prod_{i=0, \dots, K}^K \widehat{f}_{j,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K, \Pi f_j} \right) \text{ in probability,}$$

where $\Sigma_{K, \Pi f_j} = J_h(\underline{f}_{-K}) \Sigma_{K, \underline{f}} J_h(\underline{f}_{-K})' = (\Sigma_{K, \Pi f_j}(i_1, i_2))_{i_1, i_2=0, \dots, K}$ is a $(K + 1) \times (K + 1)$ covariance matrix with entries

$$\Sigma_{K, \Pi f_j}(i_1, i_2) = \sum_{j=\max(i_1, i_2)}^K \frac{\sigma_j^2}{\mu_j} \prod_{l=\max(i_1, i_2), l \neq j}^K f_l^2 \prod_{m=\min(i_1, i_2)}^{\max(i_1, i_2)-1} f_m,$$

for $i_1, i_2 = 0, \dots, K$. Here, $\Sigma_{K, \underline{f}}$ is defined in Theorem A.1(ii) and

$$J_h(\underline{x}) = \begin{pmatrix} \prod_{l=0, l \neq 0}^K x_l & \prod_{l=0, l \neq 1}^K x_l & \dots & \prod_{l=0, l \neq K}^K x_l \\ 0 & \prod_{l=1, l \neq 1}^K x_l & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \prod_{l=K, l \neq K}^K x_l \end{pmatrix} \tag{A.3}$$

as derived in the proof of Corollary 3.2 in Steinmetz and Jentsch (2022).

A.2. Proof of Theorem A.1

By construction of the Mack bootstrap estimators $\widehat{f}_{j,n}^*$, $j = 0, \dots, I + n - 1$ according to (3.5, for each fixed $K \in \mathbb{N}_0$, the $K + 1$ estimators $\widehat{f}_{0,n}^*, \widehat{f}_{1,n}^*, \dots, \widehat{f}_{K,n}^*$ are independent conditional on $D_{I,n}$. Hence, it is actually sufficient to prove part (i). For any fixed j and from (2.10) and (3.5), using $C_{i,j+1} = C_{i,j} F_{i,j}$, we get immediately

$$\begin{aligned} \sqrt{I+n-j} \left(\widehat{f}_{j,n}^* - \widehat{f}_{j,n} \right) &= \sqrt{I+n-j} \left(\frac{\sum_{i=-n}^{I-j-1} C_{i,j} F_{i,j}^*}{\sum_{k=-n}^{I-j-1} C_{k,j}} - \frac{\sum_{i=-n}^{I-j-1} C_{i,j}}{\sum_{k=-n}^{I-j-1} C_{k,j}} \widehat{f}_{j,n} \right) \\ &= \sum_{i=-n}^{I-j-1} \frac{C_{i,j} \left(F_{i,j}^* - \widehat{f}_{j,n} \right)}{\frac{1}{\sqrt{I+n-j}} \sum_{k=-n}^{I-j-1} C_{k,j}} \\ &=: \sum_{i=-n}^{I-j-1} Z_{i,n}. \end{aligned}$$

Noting that, for all j , $(Z_{i,n}, i = -n, \dots, I - j - 1, n \in \mathbb{N}_0)$ forms a triangular array of random variables that are independent conditional on $D_{I,n}$, we can make use of a (conditional) Lyapunov CLT to prove asymptotic normality. First, for the bootstrap mean, using measurability of all $C_{i,j}$'s and of $\widehat{f}_{j,n}$ in $Z_{i,n}$ with respect to $D_{I,n}$, we get

$$E^*(Z_{i,n}^*) = E^* \left(\frac{C_{i,j} (F_{i,j}^* - \hat{f}_{j,n})}{\frac{1}{\sqrt{I+n-j}} \sum_{k=-n}^{I-j-1} C_{k,j}} \right) = \frac{C_{i,j}}{\frac{1}{\sqrt{I+n-j}} \sum_{k=-n}^{I-j-1} C_{k,j}} (E^*(F_{i,j}^*) - \hat{f}_{j,n}).$$

Further, by the construction of Mack’s bootstrap, for any fixed j and $i = -n, \dots, I - j - 1$, we have $E^*(r_{i,j}^*) = 0$ such that

$$E^*(F_{i,j}^*) = E^* \left(\hat{f}_{j,n} + \frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} r_{i,j}^* \right) = \hat{f}_{j,n} + \frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} E^*(r_{i,j}^*) = \hat{f}_{j,n}$$

leading to $E^*(Z_{i,n}^*) = 0$. Second, for the bootstrap variance, we get

$$Var^*(Z_{i,n}^*) = \frac{C_{i,j}^2}{\left(\frac{1}{\sqrt{I+n-j}} \sum_{k=-n}^{I-j-1} C_{k,j} \right)^2} Var^*(F_{i,j}^*)$$

and, from the particular construction of Mack’s bootstrap leading to $E^*(r_{ij}^*) = 0$ and $E^*(r_{ij}^{*2}) = 1$, we obtain

$$\begin{aligned} Var^*(F_{i,j}^*) &= E^* \left(\left(\hat{f}_{j,n} + \frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} r_{i,j}^* \right)^2 \right) - \hat{f}_{j,n}^2 \\ &= \hat{f}_{j,n}^2 + 2\hat{f}_{j,n} \frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} E^*(r_{i,j}^*) + \left(\frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} \right)^2 E^*(r_{i,j}^{*2}) - \hat{f}_{j,n}^2 \\ &= \frac{\hat{\sigma}_{j,n}^2}{C_{i,j}} \end{aligned}$$

such that

$$Var^*(Z_{i,n}^*) = \frac{C_{i,j} \hat{\sigma}_{j,n}^2}{\left(\frac{1}{\sqrt{I+n-j}} \sum_{k=-n}^{I-j-1} C_{k,j} \right)^2}$$

and, altogether,

$$\begin{aligned} Var^* \left(\sum_{i=-n}^{I-j-1} Z_{i,n}^* \right) &= \sum_{i=-n}^{I-j-1} \frac{C_{i,j} \hat{\sigma}_{j,n}^2}{\left(\frac{1}{\sqrt{I+n-j}} \sum_{k=-n}^{I-j-1} C_{k,j} \right)^2} = \frac{\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} \right) \hat{\sigma}_{j,n}^2}{\left(\frac{1}{I+n-j} \sum_{k=-n}^{I-j-1} C_{k,j} \right)^2} \\ &= \frac{\hat{\sigma}_{j,n}^2}{\frac{1}{I+n-j} \sum_{k=-n}^{I-j-1} C_{k,j}}. \end{aligned}$$

Letting $n \rightarrow \infty$, making use of Assumption 2.5, we get $\hat{\sigma}_{j,n}^2 \rightarrow \sigma_j^2$ by Theorem 3.5 in Steinmetz and Jentsch (2022), as well as

$$\frac{1}{I+n-j} \sum_{k=-n}^{I-j-1} C_{k,j} \xrightarrow{p} \mu_j$$

by a WLLN using that, for all j , $(C_{k,j}, k \in \mathbb{Z}, k \leq I - j - 1)$ are iid by Assumption 2.1(iii) with (finite) mean μ_j and variance τ_j^2 according to Assumptions 2.2 and 2.3, respectively.

Finally, it remains to prove a Lyapunov condition to complete the proof. Choosing $\delta = 2$ for the Lyapunov condition, for any j , it is sufficient to show that

$$\sum_{i=-n}^{I-K-1} E^* \left(\left(\frac{C_{i,j} (F_{i,j}^* - \hat{f}_{j,n})}{\frac{1}{\sqrt{I+n-K}} \sum_{k=-n}^{I-K-1} C_{k,j}} \right)^4 \right) \xrightarrow{p} 0.$$

Due to measurability of all $C_{i,j}$ ’s with respect to $D_{I,n}$, we get

$$\begin{aligned} &\sum_{i=-n}^{I-K-1} E^* \left(\left(\frac{C_{i,j} (F_{i,j}^* - \hat{f}_{j,n})}{\frac{1}{\sqrt{I+n-K}} \sum_{k=-n}^{I-K-1} C_{k,j}} \right)^4 \right) \\ &= \frac{1}{\left(\frac{1}{I+n-K} \sum_{k=-n}^{I-K-1} C_{k,j} \right)^4} \frac{1}{(I+n-K)^2} \sum_{i=-n}^{I-K-1} C_{i,j}^4 E^* \left((F_{i,j}^* - \hat{f}_{j,n})^4 \right). \end{aligned}$$

Further, as $\frac{1}{I+n-K} \sum_{k=-n}^{I-K-1} C_{k,j} = O_P(1)$, it is sufficient to show that

$$\frac{1}{I+n-K} \sum_{i=-n}^{I-K-1} C_{i,j}^4 E^* \left(\left(F_{i,j}^* - \hat{f}_{j,n} \right)^4 \right) = O_P(1).$$

For this purpose, we have to compute $E^*((F_{i,j}^* - \hat{f}_{j,n})^4)$ next. By plugging-in for $F_{i,j}^*$, we get

$$E^* \left(\left(F_{i,j}^* - \hat{f}_{j,n} \right)^4 \right) = E^* \left(\left(\frac{\hat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} r_{i,j}^* \right)^4 \right) = \frac{\hat{\sigma}_{j,n}^4}{C_{i,j}^2} E^* \left(r_{i,j}^{*4} \right)$$

leading to

$$\frac{1}{I+n-K} \sum_{i=-n}^{I-K-1} C_{i,j}^4 E^* \left(\left(F_{i,j}^* - \hat{f}_{j,n} \right)^4 \right) = \frac{\hat{\sigma}_{j,n}^4}{I+n-K} \sum_{i=-n}^{I-K-1} C_{i,j}^2 E^* \left(r_{i,j}^{*4} \right).$$

Further, as $\frac{\hat{\sigma}_{j,n}^4}{I+n-K} \sum_{i=-n}^{I-K-1} C_{i,j}^2 = O_P(1)$, it remains to show that $E^* \left(r_{i,j}^{*4} \right) = O_P(1)$ holds as well. By construction, we have²

$$E^* \left(r_{i,j}^{*4} \right) = \frac{2}{(I+n+1)(I+n)-2} \sum_{s=0}^{I+n-2} \sum_{t=-n}^{I-s-1} \tilde{r}_{t,s}^4.$$

In the following, suppose for convenience that $\tilde{r}_{t,s} = \hat{r}_{t,s}$. However, the arguments for $\tilde{r}_{t,s}$ including re-centering (and re-scaling) are essentially the same, but tedious and lengthy. In this case, by plugging-in for $\tilde{r}_{t,s}$, we get

$$\begin{aligned} E^* \left(r_{i,j}^{*4} \right) &= \frac{2}{(I+n+1)(I+n)-2} \sum_{s=0}^{I+n-2} \sum_{t=-n}^{I-s-1} \left(\frac{\sqrt{C_{t,s}}(F_{t,s} - \hat{f}_{s,n})}{\hat{\sigma}_{s,n}} \right)^4 \\ &= \frac{2}{(I+n+1)(I+n)-2} \sum_{s=0}^{I+n-2} \sum_{t=-n}^{I-s-1} \frac{C_{t,s}^2}{\hat{\sigma}_{s,n}^4} \left(F_{t,s}^4 - 4F_{t,s}^3 \hat{f}_{s,n} + 6F_{t,s}^2 \hat{f}_{s,n}^2 - 4F_{t,s} \hat{f}_{s,n}^3 + \hat{f}_{s,n}^4 \right) \end{aligned}$$

By Assumption 4.9, for $n \rightarrow \infty$, we have

$$\sup_{j=0, \dots, I+n-1} \frac{\hat{f}_{j,n}}{f_j} = O_P(1) \quad \text{and} \quad \sup_{j=0, \dots, I+n-2} \frac{\sigma_j^2}{\hat{\sigma}_{j,n}^2} = O_P(1).$$

Hence, we can bound $E^*(r_{i,j}^{*4})$ above by

$$\begin{aligned} &O_P(1) \left(\frac{2}{(I+n+1)(I+n)-2} \sum_{s=0}^{I+n-2} \sum_{t=-n}^{I-s-1} \frac{C_{t,s}^2}{\sigma_s^4} \left(F_{t,s}^4 - 4F_{t,s}^3 f_s + 6F_{t,s}^2 f_s^2 - 4F_{t,s} f_s^3 + f_s^4 \right) \right) \\ &= O_P(1) \left(\frac{2}{(I+n+1)(I+n)-2} \sum_{s=0}^{I+n-2} \sum_{t=-n}^{I-s-1} \frac{C_{t,s}^2}{\sigma_s^4} (F_{t,s} - f_s)^4 \right). \end{aligned}$$

Finally, the term in brackets on the last right-hand side is a sum consisting of non-negative summands, which is also $O_P(1)$ as its expectation is bounded because the $\kappa_j^{(4)}$'s defined in (2.21) are assumed to form a bounded sequence $((\kappa_j^{(4)}/\sigma_j^4), j \in \mathbb{N}_0)$ again according to Assumption 4.9. \square

A.3. Proof of Corollary A.2

The proof follows from an application of the delta method and Theorem A.1 and is completely analogous to the proof of Corollary 3.2 in Steinmetz and Jentsch (2022). \square

Appendix B. Proofs of Section 4

B.1. Proof of Theorem 4.8

As the (conditional) L_2 -convergence result in Theorem 4.1 implies the (conditional) convergence in distribution in (4.6), for $n \rightarrow \infty$, it remains to show

$$(R_{I,n}^* - \hat{R}_{I,n})_1 \left(\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n} \right) \xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I,\infty} \tag{B.1}$$

² Note that we implicitly assume that only $\hat{\sigma}_{I+n-1,n}^2$ is estimated as zero; see Section 3.1.

with $E^*((R_{I,n}^* - \hat{R}_{I,n})_1 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) \rightarrow 0$ and

$$Var^* \left((R_{I,n}^* - \hat{R}_{I,n})_1 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \rightarrow Var \left((R_{I,\infty} - \hat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty} \right) \tag{B.2}$$

in probability, respectively.

Nevertheless, the asymptotic theory for $(R_{I,n}^* - \hat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$ is not straightforward as it is composed of sums and products consisting asymptotically of infinitely many summands and factors. Hence, we decompose $(R_{I,n}^* - \hat{R}_{I,n})_1$ by truncating these sums and products to be able to apply Proposition 6.3.9 in Brockwell and Davis (1991). For this purpose, let $K \in \mathbb{N}_0$ be fixed and suppose $I, n \in \mathbb{N}_0$ are large enough such that $K < I + n - 1$. Then, we have

$$\begin{aligned} (R_{I,n}^* - \hat{R}_{I,n})_1 &= \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* \right) \\ &= \sum_{i=0}^K C_{I-i,i}^* \left(\prod_{j=i}^K F_{I-i,j}^* - \prod_{j=i}^K \hat{f}_{j,n}^* \right) \\ &\quad + \sum_{i=0}^K C_{I-i,i}^* \left(\prod_{j=i}^K F_{I-i,j}^* \left(\prod_{l=K+1}^{I+n-1} F_{I-i,l}^* - 1 \right) - \prod_{j=i}^K \hat{f}_{j,n}^* \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* - 1 \right) \right) \\ &\quad + \sum_{i=K+1}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* \right) \\ &=: A_{1,K,I,n}^* + A_{2,K,I,n}^* + A_{3,K,I,n}^* \end{aligned}$$

Hence, to derive the claimed conditional limiting distribution, it suffices to show that, a) for all $K \in \mathbb{N}_0$, $A_{1,K,I,n}^* | (\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}) \xrightarrow{d} \mathcal{G}_{1,K} | \mathcal{Q}_{I,\infty}$ in probability as $n \rightarrow \infty$ for some (conditional) distribution $\mathcal{G}_{1,K} | \mathcal{Q}_{I,\infty}$, b) $\mathcal{G}_{1,K} | \mathcal{Q}_{I,\infty} \xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I,\infty}$ as $K \rightarrow \infty$, and c) that, for all $\epsilon > 0$, we have

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P^* \left(|A_{2,K,I,n}^*| > \epsilon | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) = 0 \quad \text{and} \quad \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P^* \left(|A_{3,K,I,n}^*| > \epsilon | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) = 0. \tag{B.3}$$

We begin with showing part a). The parametric family of (conditional) distributions used to generate the $F_{i,j} | C_{i,j}$ and $F_{i,j}^* | C_{i,j}^*$ is continuous with respect to $C_{i,j}$, f_j , σ_j^2 and $C_{i,j}^*$, $\hat{f}_{j,n}^*$, $\hat{\sigma}_{j,n}^2$, respectively, by Assumption 4.6. Hence, as $\hat{f}_{j,n} - f_j = O_P((I+n-1)^{-1/2})$, $\hat{f}_{j,n}^* - \hat{f}_{j,n} = O_P((I+n-1)^{-1/2})$ and $\hat{\sigma}_{j,n}^2 - \sigma_j^2 = O_P((I+n-1)^{-1/2})$ holds for all fixed $j \in \mathbb{N}_0$, we can conclude that, for all fixed $K \in \mathbb{N}_0$ and as $n \rightarrow \infty$, that

$$\sum_{i=0}^K C_{I-i,i}^* \left(\prod_{j=i}^K F_{I-i,j}^* - \prod_{j=i}^K \hat{f}_{j,n}^* \right) | (\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}) \xrightarrow{d} \sum_{i=0}^K C_{I-i,i} \left(\prod_{j=i}^K F_{I-i,j} - \prod_{j=i}^K f_j \right) | \mathcal{Q}_{I,\infty} \tag{B.4}$$

in probability, which proves $A_{1,K,I,n}^* | (\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}) \xrightarrow{d} \mathcal{G}_{1,K} | \mathcal{Q}_{I,\infty}$. For part b), by letting also $K \rightarrow \infty$, we get immediately

$$\sum_{i=0}^K C_{I-i,i} \left(\prod_{j=i}^K F_{I-i,j} - \prod_{j=i}^K f_j \right) | \mathcal{Q}_{I,\infty} \xrightarrow{d} \sum_{i=0}^{\infty} C_{I-i,i} \left(\prod_{j=i}^{\infty} F_{I-i,j} - \prod_{j=i}^{\infty} f_j \right) | \mathcal{Q}_{I,\infty} \sim \mathcal{G}_1 | \mathcal{Q}_{I,\infty}, \tag{B.5}$$

which proves $\mathcal{G}_{1,K} | \mathcal{Q}_{I,\infty} \xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I,\infty}$. Before we prove part c), let us also consider mean and variance of $A_{1,K,I,n}^*$ (conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$). For the mean, using measurability of $C_{I-i,i}$ with respect to $\mathcal{D}_{I,n}$ and the law of iterated expectations, we have

$$\begin{aligned} E^* (A_{1,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) &= E^* (E^* (A_{1,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^*) | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) \\ &= \sum_{i=0}^K C_{I-i,i} E^* \left(E^* \left(\prod_{j=i}^K F_{I-i,j}^* - \prod_{j=i}^K \hat{f}_{j,n}^* \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \\ &= 0 \end{aligned}$$

due to

$$\begin{aligned} E^* \left(\prod_{j=i}^K F_{I-i,j}^* - \prod_{j=i}^K \hat{f}_{j,n}^* \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) &= E^* \left(\prod_{j=i}^K F_{I-i,j}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) - \prod_{j=i}^K \hat{f}_{j,n}^* \\ &= \prod_{j=i}^K \hat{f}_{j,n}^* - \prod_{j=i}^K \hat{f}_{j,n}^* = 0 \end{aligned} \tag{B.6}$$

using similar arguments as used to show $E(\prod_{j=i}^K F_{I-i,j}) = \prod_{j=i}^K f_j$. Similarly, using the law of total variance and (B.6), we get for the variance

$$\begin{aligned}
 Var^*(A_{1,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) &= E^* \left(Var^*(A_{1,K,I,n}^* | Q_{I,n}^* = Q_{I,n}, F_{I,n}^*) | Q_{I,n}^* = Q_{I,n} \right) \\
 &\quad + Var^* \left(E^*(A_{1,K,I,n}^* | Q_{I,n}^* = Q_{I,n}, F_{I,n}^*) | Q_{I,n}^* = Q_{I,n} \right) \\
 &= E^* \left(\sum_{i=0}^K C_{I-i,i} \sum_{j=i}^K \left(\prod_{k=i}^{j-1} \hat{f}_{k,n}^* \right) \hat{\sigma}_{j,n}^2 \left(\prod_{l=j+1}^K \hat{f}_{l,n}^{*2} \right) | Q_{I,n}^* = Q_{I,n} \right) \\
 &= \sum_{i=0}^K C_{I-i,i} \sum_{j=i}^K \hat{\sigma}_{j,n}^2 E^* \left(\left(\prod_{k=i}^{j-1} \hat{f}_{k,n}^* \right) \left(\prod_{l=j+1}^K \hat{f}_{l,n}^{*2} \right) \right)
 \end{aligned} \tag{B.7}$$

due to the fact that $\hat{f}_{k,n}^*$'s are independent of the condition $Q_{I,n}^* = Q_{I,n}$ and because of

$$E^*(A_{1,K,I,n}^* | Q_{I,n}^* = Q_{I,n}, F_{I,n}^*) = \sum_{i=0}^K C_{I-i,i} \sum_{j=i}^K \left(\prod_{k=i}^{j-1} \hat{f}_{k,n}^* \right) \hat{\sigma}_{j,n}^2 \left(\prod_{l=j+1}^K \hat{f}_{l,n}^{*2} \right)$$

obtained by similar arguments as used in the proof of (Steinmetz and Jentsch 2022, Theorem 4.3) and using the measurability of $C_{I-i,i}$ and $\hat{\sigma}_{j,n}^2$ with respect to $D_{I,n}$. Now, using similar arguments as in (Steinmetz and Jentsch 2022, Theorem 4.7) and exploiting the fact that the $\hat{f}_{k,n}^*$'s are stochastically independent conditional on $D_{I,n}$, for the expectation in (B.7), we get

$$\begin{aligned}
 E^* \left(\left(\prod_{k=i}^{j-1} \hat{f}_{k,n}^* \right) \left(\prod_{l=j+1}^K \hat{f}_{l,n}^{*2} \right) \right) &= \left(\prod_{k=i}^{j-1} E^* \left(\hat{f}_{k,n}^* \right) \right) \left(\prod_{l=j+1}^K E^* \left(\hat{f}_{l,n}^{*2} \right) \right) \\
 &= \left(\prod_{k=i}^{j-1} \hat{f}_{k,n} \right) \left(\prod_{l=j+1}^K \left(\frac{\hat{\sigma}_{l,n}^2}{\sum_{k=-n}^{I-l-1} C_{k,l}} + \hat{f}_{l,n}^2 \right) \right) \\
 &= \left(\prod_{k=i}^{j-1} \hat{f}_{k,n} \right) \left(\prod_{l=j+1}^K \hat{f}_{l,n}^2 \right) + O_P \left(\frac{1}{I+n} \right)
 \end{aligned}$$

due to, for all $c \in \{0, \dots, K\}$, we have

$$E^* \left(\hat{f}_{c,n}^{*2} | B_{I,n}(c) \right) = \frac{\hat{\sigma}_{c,n}^2}{\sum_{k=-n}^{I-c-1} C_{k,c}} + \hat{f}_{c,n}^2, \tag{B.8}$$

where $B_{I,n}(k) = \{C_{i,j} | i = -n, \dots, I, j = 0, \dots, k, i + j \leq I + n\}$ denotes all elements of $D_{I,n}$ up to its k th column, and because of $\hat{\sigma}_{l,n}^2 \rightarrow \sigma_l^2$ in probability for all $l \in \{0, \dots, K\}$ and

$$\frac{1}{\sum_{k=-n}^{I-l-1} C_{k,l}} \leq \frac{1}{(I+n-l)\epsilon^l} \leq \frac{1}{(I+n-K)\epsilon^K} = O \left(\frac{1}{I+n} \right)$$

as K is fixed. This leads to

$$Var^*(A_{1,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) = \sum_{i=0}^K C_{I-i,i} \sum_{j=i}^K \left(\prod_{k=i}^{j-1} \hat{f}_{k,n} \right) \hat{\sigma}_{j,n}^2 \left(\prod_{l=j+1}^K \hat{f}_{l,n}^2 \right) + O_P \left(\frac{1}{I+n} \right) \tag{B.9}$$

$$\xrightarrow{p} \sum_{i=0}^K C_{I-i,i} \sum_{j=i}^K \left(\prod_{k=i}^{j-1} f_k \right) \sigma_j^2 \left(\prod_{l=j+1}^K f_l^2 \right) \tag{B.10}$$

as $n \rightarrow \infty$ for all K fixed, because $\hat{f}_{j,n} - f_j = O_P((I+n-1)^{-1/2})$ and $\hat{\sigma}_{j,n}^2 - \sigma_j^2 = O_P((I+n-1)^{-1/2})$ for all $j \in \mathbb{N}_0$. Finally, letting $K \rightarrow \infty$, we get

$$\sum_{i=0}^K C_{I-i,i} \sum_{j=i}^K \left(\prod_{k=i}^{j-1} f_k \right) \sigma_j^2 \left(\prod_{l=j+1}^K f_l^2 \right) \rightarrow \sum_{i=0}^{\infty} C_{I-i,i} \sum_{j=i}^{\infty} \left(\prod_{k=i}^{j-1} f_k \right) \sigma_j^2 \left(\prod_{l=j+1}^{\infty} f_l^2 \right), \tag{B.11}$$

which equals $Var((R_{I,\infty} - \hat{R}_{I,\infty})_1 | Q_{I,\infty})$. Hence, it remains to show part c) to complete the proof. We begin with showing part c) for $A_{2,K,I,n}^*$. By similar arguments used above, for the mean, we have $E^*(A_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) = 0$ due to $E^*(A_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}, F_{I,n}^*) = 0$ and, for the variance, we have $Var^*(A_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) = E^*((A_{2,K,I,n}^*)^2 | Q_{I,n}^* = Q_{I,n}) = E^*(E^*((A_{2,K,I,n}^*)^2 | Q_{I,n}^* = Q_{I,n}, F_{I,n}^*) | Q_{I,n}^* = Q_{I,n})$. For the inner expectation, using stochastic independence over accident years leading to stochastic independent summands of $A_{2,K,I,n}^*$ (conditional on $Q_{I,n}^* = Q_{I,n}$, $D_{I,n}$ and $F_{I,n}^*$), we get

$$\begin{aligned}
 &E^*((A_{2,K,I,n}^*)^2 | Q_{I,n}^* = Q_{I,n}, F_{I,n}^*) \\
 &= \sum_{i=0}^K C_{I-i,i}^2 E^* \left(\left(\prod_{j=i}^K F_{I-i,j}^* \left(\prod_{l=K+1}^{I+n-1} F_{I-i,l}^* - 1 \right) - \prod_{j=i}^K \hat{f}_{j,n}^* \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* - 1 \right) \right)^2 \middle| Q_{I,n}^* = Q_{I,n}, F_{I,n}^* \right) \\
 &= \sum_{i=0}^K C_{I-i,i}^2 \left[E^* \left(\left(\prod_{j=i}^K F_{I-i,j}^* \left(\prod_{l=K+1}^{I+n-1} F_{I-i,l}^* - 1 \right) \right)^2 \middle| Q_{I,n}^* = Q_{I,n}, F_{I,n}^* \right) - \left(\prod_{j=i}^K \hat{f}_{j,n}^* \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* - 1 \right) \right)^2 \right].
 \end{aligned}$$

For the term corresponding to the first term in brackets on the last right-hand side, we get

$$\begin{aligned} & \sum_{i=0}^K C_{I-i,i}^2 E^* \left(\left(\prod_{j=i}^K F_{I-i,j}^* \left(\prod_{l=K+1}^{I+n-1} F_{I-i,l}^* - 1 \right) \right)^2 \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \\ &= \sum_{i=0}^K C_{I-i,i}^2 E^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^{*2} - 2 \left(\prod_{j=i}^K F_{I-i,j}^{*2} \right) \left(\prod_{l=K+1}^{I+n-1} F_{I-i,l}^* \right) + \prod_{j=i}^K F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right). \end{aligned} \tag{B.12}$$

Using linearity of expectations, for the first expectation on the last right-hand side of (B.12), due to $F_{i,j}^* = \frac{C_{i,j+1}^*}{C_{i,j}^*}$ and $C_{I-i,i}^* = C_{I-i,i}$, we get

$$\begin{aligned} & E^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \\ &= E^* \left(\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^{*2} \right) E^* \left(F_{I-i,I+n-1}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, C_{I-i,i}^*, \dots, C_{I-i,I+n-1}^*, \mathcal{F}_{I,n}^* \right) \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \\ &= E^* \left(\prod_{j=i}^{I+n-2} F_{I-i,j}^{*2} \left(\frac{\hat{\sigma}_{I+n-1,n}^2}{C_{I-i,I+n-1}^*} + \hat{f}_{I+n-1,n}^{*2} \right) \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \\ &= E^* \left(\prod_{j=i}^{I+n-2} F_{I-i,j}^{*2} \frac{1}{C_{I-i,I+n-1}^*} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \hat{\sigma}_{I+n-1,n}^2 + E^* \left(\prod_{j=i}^{I+n-2} F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \hat{f}_{I+n-1,n}^{*2} \\ &= E^* \left(\prod_{j=i}^{I+n-2} F_{I-i,j}^* \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \frac{\hat{\sigma}_{I+n-1,n}^2}{C_{I-i,i}} + E^* \left(\prod_{j=i}^{I+n-2} F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \hat{f}_{I+n-1,n}^{*2} \\ &= \frac{\prod_{j=i}^{I+n-2} \hat{f}_{j,n}^* \hat{\sigma}_{I+n-1,n}^2}{C_{I-i,i}} + E^* \left(\prod_{j=i}^{I+n-2} F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \hat{f}_{I+n-1,n}^{*2}. \end{aligned}$$

By recursively plugging-in, we get

$$E^* \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) = \frac{1}{C_{I-i,i}} \sum_{k=i}^{I+n-1} \left(\prod_{j=i}^{k-1} \hat{f}_{j,n}^* \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^{I+n-1} \hat{f}_{h,n}^{*2} \right) + \prod_{j=i}^{I+n-1} \hat{f}_{j,n}^{*2}.$$

Similarly, for the second expectation in (B.12), we get

$$\begin{aligned} & E^* \left(-2 \left(\prod_{j=i}^K F_{I-i,j}^{*2} \right) \left(\prod_{l=K+1}^{I+n-1} F_{I-i,l}^* \right) \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \\ &= -2 \left(\frac{1}{C_{I-i,i}} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n}^* \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \hat{f}_{h,n}^{*2} \right) + \prod_{j=i}^K \hat{f}_{j,n}^{*2} \right) \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* \right) \end{aligned}$$

and for the third one, we have

$$E^* \left(\prod_{j=i}^K F_{I-i,j}^{*2} \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) = \frac{1}{C_{I-i,i}} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n}^* \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \hat{f}_{h,n}^{*2} \right) + \prod_{j=i}^K \hat{f}_{j,n}^{*2}.$$

Altogether, for all $K < I + n - 1$, this leads to

$$\begin{aligned} & E^* \left((A_{2,K,I,n}^*)^2 \middle| \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^* \right) \\ &= \sum_{i=0}^K C_{I-i,i} \left[\sum_{k=i}^{I+n-1} \left(\prod_{j=i}^{k-1} \hat{f}_{j,n}^* \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^{I+n-1} \hat{f}_{h,n}^{*2} \right) - 2 \left(\sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n}^* \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \hat{f}_{h,n}^{*2} \right) \right) \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* \right) \right. \\ & \quad \left. + \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n}^* \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \hat{f}_{h,n}^{*2} \right) \right]. \end{aligned}$$

Plugging-in and making use of the fact that the $\hat{f}_{j,n}^*$'s are stochastically independent conditional on $\mathcal{D}_{I,n}$ and $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$, this leads to

$$\begin{aligned}
 & Var^*(A_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) \\
 &= \sum_{i=0}^K C_{I-i,i} \left[\sum_{k=i}^{I+n-1} \hat{\sigma}_{k,n}^2 \left(\prod_{j=i}^{k-1} E^*(\hat{f}_{j,n}^* | Q_{I,n}^* = Q_{I,n}) \right) \left(\prod_{h=k+1}^{I+n-1} E^*(\hat{f}_{h,n}^{*2} | Q_{I,n}^* = Q_{I,n}) \right) \right. \\
 &\quad \left. - 2 \sum_{k=i}^K \hat{\sigma}_{k,n}^2 \left(\prod_{j=i}^{k-1} E^*(\hat{f}_{j,n}^* | Q_{I,n}^* = Q_{I,n}) \right) \left(\prod_{h=k+1}^K E^*(\hat{f}_{h,n}^{*2} | Q_{I,n}^* = Q_{I,n}) \right) \left(\prod_{l=K+1}^{I+n-1} E^*(\hat{f}_{l,n}^* | Q_{I,n}^* = Q_{I,n}) \right) \right. \\
 &\quad \left. + \sum_{k=i}^K \hat{\sigma}_{k,n}^2 \left(\prod_{j=i}^{k-1} E^*(\hat{f}_{j,n}^*) \right) \left(\prod_{h=k+1}^K E^*(\hat{f}_{h,n}^{*2}) \right) \right] \\
 &= \sum_{i=0}^K C_{I-i,i} \left[\sum_{k=i}^{I+n-1} \hat{\sigma}_{k,n}^2 \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \left(\prod_{h=k+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right. \\
 &\quad \left. - 2 \sum_{k=i}^K \hat{\sigma}_{k,n}^2 \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right. \\
 &\quad \left. + \sum_{k=i}^K \hat{\sigma}_{k,n}^2 \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right] \\
 &= \sum_{i=0}^K C_{I-i,i} \left[\sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right. \\
 &\quad \times \left(\left(\prod_{h=K+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) - \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right) \\
 &\quad + \sum_{k=K+1}^{I+n-1} \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \\
 &\quad \left. + \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \left(1 - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right]
 \end{aligned}$$

obtained by re-arranging terms and due to $E^*(\hat{f}_{j,n}^* | Q_{I,n}^* = Q_{I,n}) = E^*(\hat{f}_{j,n}^*) = \hat{f}_{j,n}$ and

$$E^*(\hat{f}_{j,n}^{*2} | Q_{I,n}^* = Q_{I,n}) = E^*(\hat{f}_{j,n}^{*2}) = \hat{f}_{j,n}^2 + \frac{\hat{\sigma}_{j,n}^2}{\sum_{p=-n}^{I-j-1} C_{p,j}}$$

for all j . Next, to argue that $Var^*(A_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) \geq 0$ vanishes in probability for $K \rightarrow \infty$ and $n \rightarrow \infty$ afterwards, it suffices to show that its unconditional expectation is bounded for $K \rightarrow \infty$ and that its bound converges to zero as $n \rightarrow \infty$. We get

$$\begin{aligned}
 & E \left(Var^*(A_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) \right) \\
 &= \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right. \\
 &\quad \times \left. \left(\prod_{h=K+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right] \\
 &\quad + \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=K+1}^{I+n-1} \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right] \\
 &\quad + \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \left(1 - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right] \\
 &= \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right. \\
 &\quad \times \left. \left(\prod_{h=K+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right]
 \end{aligned}$$

$$\begin{aligned}
 & + \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=K+1}^{I+n-1} \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right] \\
 & + \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \left(1 - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right].
 \end{aligned}$$

Now, let us consider the three terms on the last right-hand side separately. Using $\sum_{p=-n}^{I-h-1} C_{p,h} \geq (I+n-h)e^h \geq e^h$, the first one can be bounded by

$$\begin{aligned}
 & \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) \right) \left(\prod_{h=K+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right] \\
 & = \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^{I+n-1} \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) \right) \right] \\
 & - \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) \right) \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right].
 \end{aligned}$$

Next, using the law of iterated expectations and

$$E \left[\hat{f}_{c,n}^2 + \frac{\hat{\sigma}_{c,n}^2}{e^c} \mid \mathcal{B}_{I,n}(c) \right] = \frac{\sigma_c^2}{\sum_{k=-n}^{I-c-1} C_{k,c}} + f_c^2 + \frac{\sigma_c^2}{e^c} \leq f_c^2 + \frac{2\sigma_c^2}{e^c}$$

for all $c \in \{0, \dots, I+n-1\}$, where $\mathcal{B}_{I,n}(k) = \{C_{i,j} \mid i = -n, \dots, I, j = 0, \dots, k, i+j \leq I\}$, the first term on the right-hand side above becomes

$$\sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) \right) \left(\prod_{l=K+1}^{I+n-1} \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) \right) \right]$$

and, similarly, for the second term, we obtain

$$- \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) \right) \left(\prod_{l=K+1}^{I+n-1} f_l \right) \right].$$

Together, this term becomes

$$\sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{e^h} \right) \right) \left(\prod_{l=K+1}^{I+n-1} \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) - \prod_{l=K+1}^{I+n-1} f_l \right) \right],$$

which, using similar arguments as above, can be bounded by

$$\left[\sum_{i=0}^K \mu_i \sum_{k=i}^K \left(\prod_{j=i}^{k-1} f_j \right) \sigma_k^2 \left(\prod_{h=k+1}^K \left(f_h^2 + \frac{2\sigma_h^2}{e^h} \right) \right) \right] \left(\prod_{l=K+1}^{I+n-1} \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) - \prod_{l=K+1}^{I+n-1} f_l \right).$$

Now, letting $n \rightarrow \infty$, we get the following upper bound

$$\mu_\infty^2 \left(\prod_{h=0}^\infty \left(f_h^2 + \frac{2\sigma_h^2}{e^h} \right) \right) \sum_{i=0}^K \sum_{k=i}^K \sigma_k^2 \left(\prod_{l=K+1}^\infty \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) - \prod_{l=K+1}^\infty f_l \right) < \infty,$$

which is finite using $\prod_{j=0}^\infty x_j < \infty$ if and only if $\sum_{j=0}^\infty (x_j - 1) < \infty$ for $x_j \geq 1$ for all j , and as we have

$$\begin{aligned}
 \sum_{h=0}^\infty \left(f_h^2 + \frac{2\sigma_h^2}{e^h} - 1 \right) & = \sum_{h=0}^\infty (f_h^2 - 1) + \sum_{h=0}^\infty \frac{2\sigma_h^2}{e^h} = \sum_{h=0}^\infty (f_h - 1)(f_h + 1) + 2 \sum_{h=0}^\infty \frac{\sigma_h^2}{e^h} \\
 & \leq \sup_{h \in \mathbb{N}_0} (f_h + 1) \sum_{h=0}^\infty (f_h - 1) + 2 \sum_{h=0}^\infty \frac{\sigma_h^2}{e^h} < \infty
 \end{aligned}$$

by Assumptions 2.4 and 4.2. Now, letting also $K \rightarrow \infty$, the term $\sum_{i=0}^K \sum_{k=i}^K \sigma_k^2$ also remains bounded due to

$$\sum_{i=0}^K \sum_{k=i}^K \sigma_k^2 = \sum_{j=0}^K \sigma_j^2 \sum_{i=0}^j 1 = \sum_{j=0}^K (j+1) \sigma_j^2 \leq \sum_{j=0}^\infty (j+1) \sigma_j^2 < \infty.$$

Finally, as $f_l \rightarrow 1$ and $\sigma_l^2/e^l \rightarrow 0$ for $l \rightarrow \infty$, we get $\prod_{l=K+1}^\infty (f_l^2 + \frac{2\sigma_l^2}{e^l}) \rightarrow 1$ and $\prod_{l=K+1}^\infty f_l \rightarrow 1$ for $K \rightarrow \infty$ leading to

$$\prod_{l=K+1}^{\infty} \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) - \prod_{l=K+1}^{\infty} f_l \rightarrow 0.$$

Similarly, using the same arguments, the second term in the representation of $E(Var^*(A_{2,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}))$ above can be bounded by

$$\sum_{i=0}^K \mu_i \sum_{k=K+1}^{I+n-1} \left(\prod_{j=i}^{k-1} f_j \right) \sigma_k^2 \left(\prod_{h=k+1}^{I+n-1} \left(f_h^2 + \frac{2\sigma_h^2}{e^h} \right) \right),$$

which, for $n \rightarrow \infty$, can be bounded by

$$\mu_{\infty}^2 \left(\prod_{h=0}^{\infty} \left(f_h^2 + \frac{2\sigma_h^2}{e^h} \right) \right) \sum_{i=0}^K \sum_{k=K+1}^{\infty} \sigma_k^2 \leq \mu_{\infty}^2 \left(\prod_{h=0}^{\infty} \left(f_h^2 + \frac{2\sigma_h^2}{e^h} \right) \right) \sum_{k=K+1}^{\infty} (k+1)\sigma_k^2 < \infty,$$

which vanishes for $K \rightarrow \infty$.

Finally, for the third term in the representation of $E(Var^*(A_{2,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}))$, we get

$$\begin{aligned} & \sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \left(1 - \prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} \right) \right] \\ &= \left(\sum_{i=0}^K E \left[C_{I-i,i} \sum_{k=i}^K \left(\prod_{j=i}^{k-1} \hat{f}_{j,n} \right) \hat{\sigma}_{k,n}^2 \left(\prod_{h=k+1}^K \left(\hat{f}_{h,n}^2 + \frac{\hat{\sigma}_{h,n}^2}{\sum_{p=-n}^{I-h-1} C_{p,h}} \right) \right) \right] \right) \left(1 - \prod_{l=K+1}^{I+n-1} f_l \right) \end{aligned}$$

While, for $n \rightarrow \infty$, the second factor $1 - \prod_{l=K+1}^{I+n-1} f_l$ can be bounded by $1 - \prod_{l=K+1}^{\infty} f_l$, which converges to 0 due to $\prod_{l=K+1}^{\infty} f_l \rightarrow 1$ for $K \rightarrow \infty$, the first factor above can be bounded by

$$\sum_{i=0}^K \mu_i \sum_{k=i}^K \left(\prod_{j=i}^{k-1} f_j \right) \sigma_k^2 \left(\prod_{h=k+1}^K \left(f_h^2 + \frac{\sigma_h^2}{e^h} \right) \right),$$

which, for $n \rightarrow \infty$, can be bounded further by

$$\mu_{\infty}^2 \left(\prod_{h=0}^{\infty} \left(f_h^2 + \frac{\sigma_h^2}{e^h} \right) \right) \sum_{j=0}^K (j+1)\sigma_j^2,$$

which is bounded as $\sum_{j=0}^K (j+1)\sigma_j^2 \rightarrow \sum_{j=0}^{\infty} (j+1)\sigma_j^2 < \infty$ for $K \rightarrow \infty$. This completes the first part of c) for term $A_{2,K,I,n}^*$. Continuing with $A_{3,K,I,n}^*$ to prove also the second part of c), we have $E^*(A_{3,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) = 0$ using the law of iterated expectations. By using similar calculations as for $A_{2,K,I,n}^*$, we get

$$E^*((A_{3,K,I,n}^*)^2 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^*) = \sum_{i=K+1}^{I+n} C_{I-i,i} \sum_{j=i}^{I+n-1} \left(\prod_{h=i}^{j-1} \hat{f}_{h,n} \right) \hat{\sigma}_{j,n}^2 \left(\prod_{l=j+1}^{I+n-1} \hat{f}_{l,n}^2 \right)$$

and

$$\begin{aligned} Var^*(A_{3,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) &= E^* \left(E^*((A_{3,K,I,n}^*)^2 | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{F}_{I,n}^*) | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \\ &= \sum_{i=K+1}^{I+n} C_{I-i,i} \sum_{j=i}^{I+n-1} \left(\prod_{h=i}^{j-1} \hat{f}_{h,n} \right) \hat{\sigma}_{j,n}^2 \left(\prod_{l=j+1}^{I+n-1} \left(\hat{f}_{l,n}^2 + \frac{\hat{\sigma}_{l,n}^2}{\sum_{k=-n}^{I-j+1} C_{k,l}} \right) \right). \end{aligned}$$

Hence, to show that $Var^*(A_{3,K,I,n}^*) \geq 0$ vanishes in probability, we prove that $E(Var^*(A_{3,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}))$ is bounded for $n \rightarrow \infty$ and that its bound converges to zero as $K \rightarrow \infty$. By plugging-in and using similar arguments as above for $A_{2,K,I,n}^*$, we get

$$\begin{aligned} E(Var^*(A_{3,K,I,n}^* | \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n})) &\leq \sum_{i=K+1}^{I+n} \mu_i \sum_{j=i}^{I+n-1} \left(\prod_{h=i}^{j-1} f_h \right) \sigma_j^2 \left(\prod_{l=j+1}^{I+n-1} \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) \right) \\ &\leq \mu_{\infty}^2 \left(\prod_{l=0}^{\infty} \left(f_l^2 + \frac{2\sigma_l^2}{e^l} \right) \right) \sum_{i=K+1}^{I+n} \sum_{j=i}^{I+n-1} \sigma_j^2, \end{aligned}$$

which is bounded for $n \rightarrow \infty$ due to

$$\sum_{i=K+1}^{I+n} \sum_{j=i}^{I+n-1} \sigma_j^2 \leq \sum_{j=K+1}^{I+n-1} (j+1)\sigma_j^2 \rightarrow \sum_{j=K+1}^{\infty} (j+1)\sigma_j^2 < \infty$$

and this bound vanishes for $K \rightarrow \infty$. \square

B.2. Proof of Theorem 4.10

Similar to the proof of Theorem 4.8 for the conditional limiting behavior of $(R_{I,n}^* - \hat{R}_{I,n})_1$ and to the proof of Theorem 4.7 in Steinmetz and Jentsch (2022) for the (unconditional!) limiting behavior of $(R_{I,n} - \hat{R}_{I,n})_2$, we decompose $(R_{I,n}^* - \hat{R}_{I,n})_2$ by truncating the sums and products to be

able to apply Proposition 6.3.9 in Brockwell and Davis (1991). For this purpose, let $K \in \mathbb{N}_0$ be fixed and suppose $I, n \in \mathbb{N}_0$ are large enough such that $K < I + n - 1$. Then, after inflating $(R_{I,n}^* - \hat{R}_{I,n})_2$ with $\sqrt{I + n + 1}$, we get

$$\begin{aligned} & \sqrt{I + n + 1} \left(R_{I,n}^* - \hat{R}_{I,n} \right)_2 \\ &= \sqrt{I + n + 1} \sum_{i=0}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right) \\ &= \sqrt{I + n + 1} \sum_{i=0}^K C_{I-i,i}^* \left(\prod_{j=i}^K \hat{f}_{j,n}^* - \prod_{j=i}^K \hat{f}_{j,n} \right) \\ & \quad + \sqrt{I + n + 1} \sum_{i=0}^K C_{I-i,i}^* \left(\prod_{j=i}^K \hat{f}_{j,n}^* \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* - 1 \right) - \prod_{j=i}^K \hat{f}_{j,n} \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} - 1 \right) \right) \\ & \quad + \sqrt{I + n + 1} \sum_{i=K+1}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right) \\ &=: B_{1,K,I,n}^* + B_{2,K,I,n}^* + B_{3,K,I,n}^* \end{aligned}$$

Hence, to derive the claimed conditional limiting distribution, it suffices to show that, a) for all $K \in \mathbb{N}_0$, $B_{1,K,I,n}^* | (Q_{I,n}^* = Q_{I,n}, D_{I,n}) \xrightarrow{d} \tilde{\mathcal{G}}_{2,K} | Q_{I,\infty}$ in probability as $n \rightarrow \infty$ for some (conditional) distribution $\tilde{\mathcal{G}}_{2,K} | Q_{I,\infty}$, b) $\tilde{\mathcal{G}}_{2,K} | Q_{I,\infty} \xrightarrow{d} \mathcal{G}_2 | Q_{I,\infty}$ as $K \rightarrow \infty$, and c) that, for all $\epsilon > 0$, we have

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P^* \left(|B_{2,K,I,n}^*| > \epsilon | Q_{I,n}^* = Q_{I,n} \right) = 0 \quad \text{and} \quad \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P^* \left(|A_{3,K,I,n}^*| > \epsilon | Q_{I,n}^* = Q_{I,n} \right) = 0.$$

We begin with part a). That is, for each fixed $K \in \mathbb{N}_0$, we consider

$$B_{1,K,I,n}^* = \sqrt{I + n + 1} \sum_{i=0}^K C_{I-i,i}^* \left(\prod_{j=i}^K \hat{f}_{j,n}^* - \prod_{j=i}^K \hat{f}_{j,n} \right), \tag{B.13}$$

where

$$\hat{f}_{j,n} = \frac{\sum_{k=-n}^{I-j-1} C_{k,j+1}}{\sum_{k=-n}^{I-j-1} C_{k,j}} \quad \text{and} \quad \hat{f}_{j,n}^* = \frac{\sum_{k=-n}^{I-j-1} C_{k,j} F_{k,j}^*}{\sum_{k=-n}^{I-j-1} C_{k,j}}.$$

In contrast to the situation in the proof of Theorem 4.7 in Steinmetz and Jentsch (2022), where all $\hat{f}_{j,n}$'s are indeed affected by conditioning on $Q_{I,n}$, here, conditional on $D_{I,n}$, all $\hat{f}_{j,n}^*$'s are independent of the condition $Q_{I,n}^* = Q_{I,n}$. Hence, for $n \rightarrow \infty$, the (unconditional!) asymptotic bootstrap theory derived in Theorem A.1 and Corollary A.2 lead to

$$B_{1,K,I,n}^* | (Q_{I,n}^* = Q_{I,n}, D_{I,n}) \xrightarrow{d} \langle Q_{I,K-I}, \mathbf{Y}_K \rangle | Q_{I,\infty}$$

in probability, where $Q_{I,K-I} = \{C_{I-i,i} | i = 0, \dots, I + (K - I) = K\}$, $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in \mathbb{R}^{K+1} , and $\mathbf{Y}_K = (Y_i, i = 0, \dots, K)$ is a $(K + 1)$ -dimensional multivariate normally distributed random variable with $\mathbf{Y}_K \sim \mathcal{N} \left(0, \Sigma_{K, \prod f_j} \right)$ with $\Sigma_{K, \prod f_j}$ defined in Corollary A.2.

Further, letting $K \rightarrow \infty$, we get $\langle Q_{I,K-I}, \mathbf{Y}_K \rangle | Q_{I,\infty} \xrightarrow{d} \langle Q_{I,\infty}, \mathbf{Y}_\infty \rangle | Q_{I,\infty}$, where $Q_{I,\infty} = \{C_{I-i,i} | i \in \mathbb{N}_0\}$, and $\mathbf{Y}_\infty = (Y_i, i \in \mathbb{N}_0)$ denotes a centered Gaussian process with covariance

$$\text{Cov}(Y_{i_1}, Y_{i_2}) = \lim_{K \rightarrow \infty} \Sigma_{K, \prod f_j}(i_1, i_2) = \sum_{j=\max(i_1, i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{l=\max(i_1, i_2), l \neq j}^{\infty} f_l^2 \prod_{m=\min(i_1, i_2)}^{\max(i_1, i_2)-1} f_m \tag{B.14}$$

for $i_1, i_2 \in \mathbb{N}_0$. Moreover, as $Q_{I,\infty}$ and \mathbf{Y}_∞ are stochastically independent, conditional on $Q_{I,\infty}$, the variance of $\langle Q_{I,\infty}, \mathbf{Y}_\infty \rangle$ computes to

$$\begin{aligned} \text{Var} \left(\langle Q_{I,\infty}, \mathbf{Y}_\infty \rangle | Q_{I,\infty} \right) &= \sum_{i=0}^{\infty} \text{Var}(C_{I-i,i} Y_i | Q_{I,\infty}) + \sum_{\substack{i_1, i_2=0 \\ i_1 \neq i_2}}^{\infty} \text{Cov}(C_{I-i_1, i_1} Y_{i_1}, C_{I-i_2, i_2} Y_{i_2} | Q_{I,\infty}) \\ &= \sum_{i=0}^{\infty} C_{I-i,i}^2 \text{Var}(Y_i) + \sum_{\substack{i_1, i_2=0 \\ i_1 \neq i_2}}^{\infty} C_{I-i_1, i_1} C_{I-i_2, i_2} \text{Cov}(Y_{i_1}, Y_{i_2}) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{j=0}^{\infty} C_{I-i_1}^2 \sum_{j=1}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{\substack{l=1 \\ l \neq j}}^{\infty} f_l^2 + \sum_{\substack{i_1, i_2=0 \\ i_1 \neq i_2}}^{\infty} C_{I-i_1, i_1} C_{I-i_2, i_2} \sum_{j=\max(i_1, i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{\substack{l=\max(i_1, i_2) \\ l \neq j}}^{\infty} f_l^2 \prod_{m=\min(i_1, i_2)}^{\max(i_1, i_2)-1} f_m \\
 &= \sum_{j=0}^{\infty} \frac{\sigma_j^2}{\mu_j f_j^2} \left(\sum_{i=0}^j C_{I-i, i}^2 \prod_{k=i}^j f_k^2 \right) + 2 \sum_{j=1}^{\infty} \frac{\sigma_j^2}{\mu_j} \left(\sum_{i_1=1}^j C_{I-i_1, i_1} \sum_{i_2=0}^{i_1-1} C_{I-i_2, i_2} \prod_{l=i_2}^{i_1-1} f_l \right) \prod_{k=j}^{\infty} f_k^2 \\
 &= \sum_{j=0}^{\infty} \frac{\sigma_j^2}{\mu_j f_j^2} \left(\frac{j+1}{j+1} \sum_{i=0}^j C_{I-i, i}^2 \prod_{k=i}^j f_k^2 \right) + 2 \sum_{j=1}^{\infty} \frac{\sigma_j^2}{\mu_j} \left(\frac{j}{j} \sum_{i_1=1}^j C_{I-i_1, i_1} \frac{1}{i_1} \sum_{i_2=0}^{i_1-1} C_{I-i_2, i_2} \prod_{l=i_2}^{i_1-1} f_l \right) \prod_{k=j}^{\infty} f_k^2 \\
 &\leq \mu_{\infty}^2 \sum_{j=0}^{\infty} \frac{\sigma_j^2}{\mu_j f_j^2} (j+1) \left(\frac{1}{j+1} \sum_{i=0}^j C_{I-i, i}^2 \right) + 2 \mu_{\infty}^2 \sum_{j=1}^{\infty} \frac{\sigma_j^2}{\mu_j} (j+1)^2 \left(\frac{1}{j} \sum_{i_1=1}^j C_{I-i_1, i_1} \frac{1}{i_1} \sum_{i_2=0}^{i_1-1} C_{I-i_2, i_2} \right) \\
 &= O_P(1)
 \end{aligned}$$

due to $\sum_{j=0}^{\infty} (j+1)^2 \sigma_j^2 < \infty$ by Assumption 2.4.

We continue with showing part c) for $B_{2,K,I,n}^*$. Using similar arguments as above, we have to consider

$$\begin{aligned}
 &B_{2,K,I,n}^*(Q_{I,n}^* = Q_{I,n}, D_{I,n}) \\
 &= \sqrt{I+n+1} \sum_{i=0}^K C_{I-i, i}^* \left(\prod_{j=i}^K \hat{f}_{j,n}^* \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n}^* - 1 \right) - \prod_{j=i}^K \hat{f}_{j,n} \left(\prod_{l=K+1}^{I+n-1} \hat{f}_{l,n} - 1 \right) \right) | (Q_{I,n}^* = Q_{I,n}, D_{I,n}).
 \end{aligned}$$

Using the unbiasedness of $\hat{f}_{j,n}^*$ conditional on $Q_{I,n}^* = Q_{I,n}, D_{I,n}$ for $\hat{f}_{j,n}$, that is,

$$E^*(\hat{f}_{j,n}^* | Q_{I,n}^* = Q_{I,n}) = E^*(\hat{f}_{j,n}^*) = \hat{f}_{j,n}$$

for all j , and the independence of the $\hat{f}_{j,n}^*$'s conditional on $Q_{I,n}^* = Q_{I,n}$ and $D_{I,n}$, we have $E^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) = 0$ by construction. Hence, it remains to show that $Var^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n})$ is bounded in probability for $n \rightarrow \infty$ and its bound vanishes for $K \rightarrow \infty$ afterwards. Now, to compute the bootstrap variance $Var^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n})$, for any fixed $K \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$ large enough such that $K < I+n-1$, we get

$$\begin{aligned}
 &Var^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) \\
 &= (I+n+1) \sum_{i_1, i_2=0}^K C_{I-i_1, i_1} C_{I-i_2, i_2} Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^* - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^* | Q_{I,n}^* = Q_{I,n} \right) \\
 &= (I+n+1) \sum_{i_1, i_2=0}^K C_{I-i_1, i_1} C_{I-i_2, i_2} Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^* - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^* \right) \\
 &\leq 2(I+n+1) \sum_{i_2=0}^K \sum_{i_1=0}^{i_2} C_{I-i_1, i_1} C_{I-i_2, i_2} Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^* - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^* \right),
 \end{aligned}$$

using that, conditional on $D_{I,n}$, the $\hat{f}_{j,n}^*$'s are independent of the condition $Q_{I,n}^* = Q_{I,n}$.

To calculate the covariance on the last right-hand side, for $i_1 \leq i_2$, first, we consider the mixed moment

$$\begin{aligned}
 &E^* \left(\left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^* \right) \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^* - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^* \right) \right) \\
 &= E^* \left(\left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n}^* \right) \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^{*2} \right) \right) - 2E^* \left(\left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n}^* \right) \left(\prod_{j_2=i_2}^K \hat{f}_{j_2,n}^{*2} \right) \left(\prod_{j_3=K+1}^{I+n-1} \hat{f}_{j_3,n}^* \right) \right) \\
 &\quad + E^* \left(\left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n}^* \right) \left(\prod_{j_2=i_2}^K \hat{f}_{j_2,n}^{*2} \right) \right) \\
 &= \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left(\prod_{j_2=i_2}^{I+n-1} \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \\
 &\quad - 2 \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left(\prod_{j_2=i_2}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \left(\prod_{j_3=K+1}^{I+n-1} \hat{f}_{j_3,n} \right) \\
 &\quad + \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left(\prod_{j_2=i_2}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right),
 \end{aligned}$$

since $\hat{f}_{j,n}^*$ and $\hat{f}_{k,n}^*$ are independent for $j \neq k$ and $j, k \in \{0, \dots, I + n - 1\}$ conditional on $D_{I,n}$. Similarly, we have

$$E^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^* \right) = \prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1} - \prod_{j_1=i_1}^K \hat{f}_{j_1}$$

leading to

$$\begin{aligned} & Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^* - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^* \right) \\ &= \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left(\prod_{j_2=i_2}^{I+n-1} \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^2 \right) \\ &\quad - 2 \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left(\prod_{j_2=i_2}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^2 \right) \left(\prod_{j_3=K+1}^{I+n-1} \hat{f}_{j_3,n} \right) \\ &\quad + \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left(\prod_{j_2=i_2}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^2 \right) \\ &= \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1,n} \right) \left[\sum_{j_3=i_2}^{I+n-1} \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \right. \\ &\quad \left. - 2 \sum_{j_3=i_2}^K \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left(\prod_{j_2=j_3+1}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} \hat{f}_{j_4,n} \right) \right. \\ &\quad \left. + \sum_{j_3=i_2}^K \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left(\prod_{j_2=j_3+1}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \right]. \end{aligned}$$

By rearranging the terms in brackets on the last right-hand side above, it becomes

$$\begin{aligned} & \sum_{j_3=K+1}^{I+n-1} \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \\ &+ \sum_{j_3=i_2}^K \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \left(1 - \prod_{j_4=K+1}^{I+n-1} \hat{f}_{j_4,n} \right) \\ &+ \sum_{j_3=i_2}^K \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left[\prod_{j_2=j_3+1}^{I+n-1} \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) - \prod_{j_2=j_3+1}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right] \\ &\quad \times \left(\prod_{j_4=K+1}^{I+n-1} \hat{f}_{j_4,n} \right) \\ &+ \sum_{j_3=i_2}^K \frac{\hat{\sigma}_{j_3,n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k,j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \hat{f}_{j_4,n}^2 \right) \left(\prod_{j_2=j_3+1}^K \left(\hat{f}_{j_2,n}^2 + \frac{\hat{\sigma}_{j_2,n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k,j_2}} \right) \right) \left(1 - \prod_{j_4=K+1}^{I+n-1} \hat{f}_{j_4,n} \right). \end{aligned}$$

Now, following the same steps as in the proof of Theorem 4.7 in Steinmetz and Jentsch (2022), we can compute the unconditional expectation of the above. Using $C_{i,j} > e^j$, $E(\hat{f}_{c,n}|B_{I,n}(c)) = f_c$, $E(\hat{\sigma}_{c,n}^2|B_{I,n}(c)) = \sigma_c^2$ as well as

$$E(\hat{f}_{c,n}^2|B_{I,n}(c)) = \frac{\sigma_c^2}{\sum_{k=-n}^{I-c-1} C_{k,c}} + f_c^2 \leq \frac{\sigma_c^2}{(I+n-c)e^c} + f_c^2 \leq \frac{\sigma_c^2}{e^c} + f_c^2 \tag{B.15}$$

for all $c \in \{0, \dots, I + n - 1\}$, where $B_{I,n}(k) = \{C_{i,j} | i = -n, \dots, I, j = 0, \dots, k, i + j \leq I + n\}$, we can argue that $Var^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) \geq 0$ vanishes in probability for $n \rightarrow \infty$ and $K \rightarrow \infty$ afterwards, by showing that its unconditional expectation is bounded for $n \rightarrow \infty$ and that its bound converges to zero as $K \rightarrow \infty$.

Using that $E(Cov^*(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1,n}^* - \prod_{j_1=i_1}^K \hat{f}_{j_1,n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2,n}^* - \prod_{j_2=i_2}^K \hat{f}_{j_2,n}^*))$ can be bounded by

$$\begin{aligned} & \left(\prod_{j_1=i_1}^{i_2-1} f_{j_1} \right) \left\{ \sum_{j_3=K+1}^{I+n-1} \frac{\sigma_{j_3}^2}{(I+n-j_3)e^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{e^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{e^{j_2}} \right) \right) \right. \\ & \left. + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{(I+n-j_3)e^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{e^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{e^{j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} - 1 \right) \right\} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{(I+n-j_3)\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \\
 & \quad \times \left[\prod_{j_2=j_3+1}^K \left(f_{j_2}^2 + \frac{2\sigma_{j_2,n}^2}{\epsilon^{j_2}} \right) \left(\prod_{j_2=K+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) - 1 \right) \right] \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} \right) \\
 & + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{(I+n-j_3)\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^K \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} - 1 \right) \Big\},
 \end{aligned}$$

we can bound also $E(\text{Var}^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n}))$ from above. Precisely, putting everything together, we get

$$\begin{aligned}
 & E(\text{Var}^*(B_{2,K,I,n}^* | Q_{I,n}^* = Q_{I,n})) \\
 & \leq 2 \sum_{i_2=0}^K \sum_{i_1=0}^{i_2} \mu_{i_1} \mu_{i_2} \left(\prod_{j_1=i_1}^{i_2-1} f_{j_1} \right) \\
 & \quad \times \left\{ \sum_{j_3=K+1}^{I+n-1} \frac{(I+n+1)\sigma_{j_3}^2}{(I+n-j_3)\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \right. \\
 & \quad + \sum_{j_3=i_2}^K \frac{(I+n+1)\sigma_{j_3}^2}{(I+n-j_3)\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2,n}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} - 1 \right) \\
 & \quad + \sum_{j_3=i_2}^K \frac{(I+n+1)\sigma_{j_3}^2}{(I+n-j_3)\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \\
 & \quad \quad \times \left[\prod_{j_2=j_3+1}^K \left(f_{j_2}^2 + \frac{2\sigma_{j_2,n}^2}{\epsilon^{j_2}} \right) \left(\prod_{j_2=K+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) - 1 \right) \right] \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} \right) \\
 & \quad \left. + \sum_{j_3=i_2}^K \frac{(I+n+1)\sigma_{j_3}^2}{(I+n-j_3)\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^K \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} - 1 \right) \right\}
 \end{aligned}$$

and the leading term of the last right-hand side becomes

$$\begin{aligned}
 & 2 \sum_{i_2=0}^K \sum_{i_1=0}^{i_2} \mu_{i_1} \mu_{i_2} \left(\prod_{j_1=i_1}^{i_2-1} f_{j_1} \right) \\
 & \quad \times \left\{ \sum_{j_3=K+1}^{I+n-1} \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \right. \\
 & \quad + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2,n}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} - 1 \right) \\
 & \quad + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left[\prod_{j_2=j_3+1}^K \left(f_{j_2}^2 + \frac{2\sigma_{j_2,n}^2}{\epsilon^{j_2}} \right) \left(\prod_{j_2=K+1}^{I+n-1} \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) - 1 \right) \right] \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} \right) \\
 & \quad \left. + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left(\prod_{j_4=i_2}^{j_3-1} \left(f_{j_4}^2 + \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \left(\prod_{j_2=j_3+1}^K \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_4=K+1}^{I+n-1} f_{j_4} - 1 \right) \right\},
 \end{aligned}$$

which can be bounded further by

$$\begin{aligned}
 & 2\mu_\infty \left(\prod_{j_4=0}^\infty \left(f_{j_4}^2 + \frac{2\sigma_{j_4}^2}{\epsilon^{j_4}} \right) \right) \sum_{i_2=0}^K \sum_{i_1=0}^{i_2} \mu_{i_1} \mu_{i_2} \left\{ \sum_{j_3=K+1}^\infty \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left(\prod_{j_4=K+1}^\infty f_{j_4} - 1 \right) \right. \\
 & \quad \left. + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left[\left(\prod_{j_2=K+1}^\infty \left(f_{j_2}^2 + \frac{2\sigma_{j_2}^2}{\epsilon^{j_2}} \right) - 1 \right) \left(\prod_{j_4=K+1}^\infty f_{j_4} \right) + \sum_{j_3=i_2}^K \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \left(\prod_{j_4=K+1}^\infty f_{j_4} - 1 \right) \right] \right\}.
 \end{aligned}$$

Now, considering the four terms in brackets separately, for the first one, we can argue that it vanishes asymptotically due to

$$\sum_{i_2=0}^K \sum_{i_1=0}^{i_2} \mu_{i_1} \mu_{i_2} \sum_{j_3=K+1}^\infty \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \leq \left(\frac{1}{K+1} \sum_{i_2=0}^K \mu_{i_2} \right) \left(\frac{1}{i_2+1} \sum_{i_1=0}^{i_2} \mu_{i_1} \right) \sum_{j_3=K+1}^\infty (j_3+1)^2 \frac{\sigma_{j_3}^2}{\epsilon^{j_3}} \rightarrow 0$$

for $K \rightarrow \infty$, because the sequence $(\frac{1}{j+1} \sum_{i=0}^j \mu_i, j \in \mathbb{N}_0)$ is converging and, consequently, also bounded, and due to $\sum_{j=0}^\infty (j+1)^2 \frac{\sigma_j^2}{\epsilon^j} < \infty$ by Assumption

4.2. Similarly, using that $\prod_{j=K+1}^\infty f_j \rightarrow 1$ and $\prod_{j=K+1}^\infty \left(f_j^2 + \frac{2\sigma_j^2}{\epsilon^j} \right) \rightarrow 1$ for $K \rightarrow \infty$, we can also show that the other three terms vanish asymptotically. This completes the first part of c) for $B_{2,K,I,n}^*$.

Similarly, for showing part c) for $B_{3,K,I,n}^*$, we have to consider

$$B_{3,K,I,n}^* | (Q_{I,n}^* = Q_{I,n}, D_{I,n}) = \sqrt{I+n+1} \sum_{i=K+1}^{I+n} C_{I-i,i}^* \left(\prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right) | (Q_{I,n}^* = Q_{I,n}, D_{I,n}).$$

By the same arguments as used above for $B_{2,K,I,n}^*$, we get $E^*(B_{3,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) = 0$ and for any fixed $K \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$ large enough such that $K < I+n-1$, we have

$$\begin{aligned} & Var^*(B_{3,K,I,n}^* | Q_{I,n}^* = Q_{I,n}) \\ &= (I+n+1) \sum_{i_1, i_2=K+1}^{I+n-1} C_{I-i_1, i_1} C_{I-i_2, i_2} Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^* | Q_{I,n}^* = Q_{I,n} \right) \\ &\leq 2(I+n+1) \sum_{i_2=K+1}^{I+n-1} \sum_{i_1=K+1}^{i_2} C_{I-i_1, i_1} C_{I-i_2, i_2} Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^* \right). \end{aligned}$$

To calculate the covariance on the last right-hand side, for $i_1 \leq i_2$, we consider the mixed moment

$$E^* \left(\left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1, n}^* \right) \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^{*2} \right) \right),$$

which is just the first term of the mixed moment of the covariance calculated for $B_{2,K,I,n}^*$. By using similar calculations to get $E^*(\hat{f}_{c,n}^{*2})$ (for $B_{2,K,I,n}^*$), we obtain

$$\begin{aligned} & Cov^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n}^*, \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^* \right) \\ &= \left[E^* \left(\left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1, n}^* \right) \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^{*2} \right) \right) - E^* \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n}^* \right) E^* \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^* \right) \right] \\ &= \left[\left(\prod_{j_1=i_1}^{i_2-1} E^* \left(\hat{f}_{j_1, n}^* \right) \right) \left(\prod_{j_2=i_2}^{I+n-1} E^* \left(\hat{f}_{j_2, n}^{*2} \right) \right) - \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n} \right) \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n} \right) \right] \\ &= \left[\left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1, n} \right) \left(\prod_{j_2=i_2}^{I+n-1} \left(\hat{f}_{j_2, n}^2 + \frac{\hat{\sigma}_{j_2, n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k, j_2}} \right) \right) - \left(\prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n} \right) \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n} \right) \right] \\ &= \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1, n} \right) \left[\prod_{j_2=i_2}^{I+n-1} \left(\hat{f}_{j_2, n}^2 + \frac{\hat{\sigma}_{j_2, n}^2}{\sum_{k=-n}^{I-j_2-1} C_{k, j_2}} \right) - \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n} \right] \\ &= \sum_{j_4=i_2}^{I+n-1} \frac{\hat{\sigma}_{j_4, n}^2}{\sum_{k=-n}^{I-j_4-1} C_{k, j_4}} \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1, n} \right) \left(\prod_{j_2=i_2}^{j_4-1} \hat{f}_{j_2, n}^2 \right) \left(\prod_{j_3=j_4+1}^{I+n-1} \left(\hat{f}_{j_3, n}^2 + \frac{\hat{\sigma}_{j_3, n}^2}{\sum_{k=-n}^{I-j_3-1} C_{k, j_3}} \right) \right) \\ &\leq \sum_{j_4=i_2}^{I+n-1} \frac{(I+n+1)\hat{\sigma}_{j_4, n}^2}{(I+n-j_4)\epsilon^{j_4}} \left(\prod_{j_1=i_1}^{i_2-1} \hat{f}_{j_1, n} \right) \left(\prod_{j_2=i_2}^{j_4-1} \hat{f}_{j_2, n}^2 \right) \left(\prod_{j_3=j_4+1}^{I+n-1} \left(\hat{f}_{j_3, n}^2 + \frac{\hat{\sigma}_{j_3, n}^2}{\epsilon^{j_3}} \right) \right). \end{aligned}$$

Noting that all involved summands and factors are non-negative, taking expectations of the last right-hand side and using the law of iterative expectations and $C_{i,j} > \epsilon^j$, $E(\hat{f}_{c,n} | \mathcal{B}_{I,n}(c)) = f_c$, $E(\hat{\sigma}_{c,n}^2 | \mathcal{B}_{I,n}(c)) = \sigma_c^2$ as well as (B.15), we get

$$\sum_{j_4=i_2}^{I+n-1} \frac{\sigma_{j_4}^2}{(I+n-j_4)\epsilon^{j_4}} \left(\prod_{j_1=i_1}^{i_2-1} f_{j_1} \right) \left(\prod_{j_2=i_2}^{j_4-1} \left(f_{j_2}^2 + \frac{\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_3=j_4+1}^{I+n-1} \left(f_{j_3}^2 + \frac{2\sigma_{j_3}^2}{\epsilon^{j_3}} \right) \right)$$

such that the leading term of $E(Var^*(B_{3,K,I,n}^* | Q_{I,n}^* = Q_{I,n}))$ becomes

$$\begin{aligned} & 2 \sum_{i_2=K+1}^{I+n-1} \sum_{i_1=K+1}^{i_2} \mu_{i_1} \mu_{i_2} \sum_{j_4=i_2}^{I+n-1} \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \left(\prod_{j_1=i_1}^{i_2-1} f_{j_1} \right) \left(\prod_{j_2=i_2}^{j_4-1} \left(f_{j_2}^2 + \frac{\sigma_{j_2}^2}{\epsilon^{j_2}} \right) \right) \left(\prod_{j_3=j_4+1}^{I+n-1} \left(f_{j_3}^2 + \frac{2\sigma_{j_3}^2}{\epsilon^{j_3}} \right) \right) \\ &\leq 2\mu_\infty \left(\prod_{j_3=0}^\infty \left(f_{j_3}^2 + \frac{2\sigma_{j_3}^2}{\epsilon^{j_3}} \right) \right) \sum_{i_2=K+1}^{I+n-1} \sum_{i_1=K+1}^{i_2} \mu_{i_1} \mu_{i_2} \sum_{j_4=i_2}^{I+n-1} \frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \\ &\leq 2\mu_\infty \left(\prod_{j_3=0}^\infty \left(f_{j_3}^2 + \frac{2\sigma_{j_3}^2}{\epsilon^{j_3}} \right) \right) \sum_{i_2=K+1}^{I+n-1} \sum_{i_1=K+1}^{i_2} \mu_{i_1} \mu_{i_2} \sum_{j_4=i_2}^{I+n-1} \frac{\sigma_{j_4}^2}{\epsilon^{j_4}}. \end{aligned}$$

For the triple sum on the last right-hand side, we get

$$\begin{aligned} \sum_{i_2=K+1}^{I+n-1} \sum_{i_1=K+1}^{i_2} \mu_{i_1} \mu_{i_2} \sum_{j_4=i_2}^{I+n-1} \frac{\sigma_{j_4}^2}{e^{j_4}} &= \sum_{i_2=1}^{I+n-K-1} \sum_{i_1=K+1}^{i_2+K} \mu_{i_1} \mu_{i_2+K} \sum_{j_4=i_2+K}^{I+n-1} \frac{\sigma_{j_4}^2}{e^{j_4}} \\ &= \sum_{i_2=1}^{I+n-K-1} i_2 \left(\frac{1}{i_2} \sum_{i_1=K+1}^{i_2+K} \mu_{i_1} \right) \mu_{i_2+K} \sum_{j_4=i_2+K}^{I+n-1} \frac{\sigma_{j_4}^2}{e^{j_4}} \\ &\leq \text{const.} \sum_{i_2=1}^{I+n-K-1} i_2 \mu_{i_2+K} \sum_{j_4=i_2+K}^{I+n-1} \frac{\sigma_{j_4}^2}{e^{j_4}} \\ &= \text{const.} \sum_{j=1+K}^{I+n-1} \frac{\sigma_j^2}{e^j} \sum_{l=1}^{j-K} l \mu_{l+K}. \end{aligned}$$

Further, the sequence $(\mu_i, i \in \mathbb{N}_0)$ shares the properties of $(C_{I-i,i}, i \in \mathbb{N}_0)$ in a deterministic sense such that $\sum_{l=1}^{j-K} l \mu_{l+K} \leq \text{const.} j^2$. Consequently, we have

$$\sum_{j=1+K}^{I+n-1} \frac{\sigma_j^2}{e^j} \sum_{l=1}^{j-K} l \mu_{l+K} \leq \text{const.} \sum_{j=1+K}^{\infty} j^2 \frac{\sigma_j^2}{e^j} \rightarrow 0$$

as $K \rightarrow \infty$ by Assumption 4.2. \square

B.3. Proof of Theorem 4.11

The proof is analogous to the proof of Theorem 4.12 and Corollary 4.13 in Steinmetz and Jentsch (2022). The claimed uncorrelatedness of $(R_{I,n}^* - \hat{R}_{I,n})_1$ and $(R_{I,n}^* - \hat{R}_{I,n})_2$ conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, follows from

$$\begin{aligned} &Cov^* \left((R_{I,n}^* - \hat{R}_{I,n})_1, \sqrt{I+n+1} (R_{I,n}^* - \hat{R}_{I,n})_2 \mid \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \\ &= Cov^* \left(\sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,j}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* \right), \sqrt{I+n+1} \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} \hat{f}_{j,n}^* - \prod_{j=i}^{I+n-1} \hat{f}_{j,n} \right) \mid \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \\ &= \sqrt{I+n+1} \sum_{i_1, i_2=0}^{I+n} E^* \left(C_{I-i_1, i_1} \left(\prod_{j_1=i_1}^{I+n-1} F_{I-i_1, j_1}^* - \prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n}^* \right) C_{I-i_2, i_2} \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2}^* - \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n} \right) \mid \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \\ &= 0 \end{aligned}$$

since for all $i_1, i_2 = 0, \dots, I+n$, we have

$$\begin{aligned} &E^* \left(C_{I-i_1, i_1} C_{I-i_2, i_2} \left(\prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n}^* - \prod_{j_2=i_2}^{I+n-1} \hat{f}_{j_2, n} \right) E^* \left(\prod_{j_1=i_1}^{I+n-1} F_{I-i_1, j_1}^* - \prod_{j_1=i_1}^{I+n-1} \hat{f}_{j_1, n}^* \mid \mathcal{F}_{I,n}^*, \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \mid \mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n} \right) \\ &= 0, \end{aligned}$$

because the inner conditional expectation on the last right-hand side is zero. \square

Appendix C. Proofs of Section 6

C.1. Proof of Theorem 6.1

Following the technique of proof in Theorem 4.8 and using $\hat{f}_{j,n} - f_j = O_P((I+n-1)^{-1/2})$, $\hat{f}_{j,n}^* - \hat{f}_{j,n} = O_{P^*}((I+n-1)^{-1/2})$ and $\hat{\sigma}_{j,n}^2 - \sigma_j^2 = O_P((I+n-1)^{-1/2})$ leads to the same limiting result also for the process uncertainty part $(R_{I,n}^+ - \hat{R}_{I,n}^+)_1$ of the alternative Mack bootstrap. \square

C.2. Proof of Theorem 6.3

Following the technique of proof in Theorem 4.10 and exploiting the limiting properties from Assumption 6.2, we get the claimed asymptotic results. \square

C.3. Proof of Theorem 6.4

Based on the results established in Theorems 6.1 and 6.3, the arguments are completely analogous to those used in the proof of Theorem 4.11. \square

Appendix D. Conditional versions of the CLTs from Steinmetz and Jentsch (2022)

For the sake of completeness, in Theorem D.1 and Corollary D.2 below, we summarize the results from Theorem C.1(ii,iv) and Corollary C.2(ii,iv) in Steinmetz and Jentsch (2022).

Theorem D.1 (Asymptotic normality of \hat{f}_j conditionally on $\mathcal{Q}_{I,n}$; Theorem C.1(ii,iv) in Steinmetz and Jentsch (2022)). Suppose Assumptions 2.2, 2.3, 2.4, 4.2 and 4.3 are satisfied. Then, as $n \rightarrow \infty$, the following holds:

(i) For each fixed $K \in \mathbb{N}_0$, let $\underline{f}_{\underline{K}} = (f_0, f_1, \dots, f_K)'$ and define

$$\underline{f}_{\underline{K},n}(\mathcal{Q}_{I,n}) = (f_{0,n}(\mathcal{Q}_{I,n}), f_{1,n}(\mathcal{Q}_{I,n}), \dots, f_{K,n}(\mathcal{Q}_{I,n}))'.$$

Then, unconditionally, we have

$$J_n^{1/2} \left(\underline{f}_{\underline{K},n}(\mathcal{Q}_{I,n}) - \underline{f}_{\underline{K}} \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K,\underline{f}}^{(1)} \right),$$

where $J_n^{1/2} = \text{diag} \left(\sqrt{I+n+1-j}, j=0, \dots, K \right)$ is a diagonal $(K+1) \times (K+1)$ matrix of inflation factors and the variance-covariance matrix

$$\Sigma_{K,\underline{f}}^{(1)} = J_g \left(\underline{\mu}_{\underline{K}} \right) \Sigma_{K,\underline{C}}^{(1)} J_g \left(\underline{\mu}_{\underline{K}} \right)',$$

where $\Sigma_{K,\underline{C}}^{(1)}$ is defined in (D.1), has entries

$$\begin{aligned} & \Sigma_{K,\underline{f}}^{(1)}(j_1, j_2) \\ &= \frac{f_{j_1} f_{j_2} E \left(E(C_{i,j_1} | C_{i,\infty}) E(C_{i,j_2} | C_{i,\infty}) \right) + E \left(E(C_{i,j_1+1} | C_{i,\infty}) E(C_{i,j_2+1} | C_{i,\infty}) \right)}{\mu_{j_1} \mu_{j_2}} \\ &+ \frac{-f_{j_2} E \left(E(C_{i,j_1+1} | C_{i,\infty}) E(C_{i,j_2} | C_{i,\infty}) \right) - f_{j_1} E \left(E(C_{i,j_1} | C_{i,\infty}) E(C_{i,j_2+1} | C_{i,\infty}) \right)}{\mu_{j_1} \mu_{j_2}} \end{aligned}$$

for $j_1, j_2 = 0, \dots, K$.

(ii) For each fixed $K \in \mathbb{N}_0$, let $\hat{\underline{f}}_{\underline{K},n} = (\hat{f}_{0,n}, \hat{f}_{1,n}, \dots, \hat{f}_{K,n})'$. Then, conditionally on $\mathcal{Q}_{I,n}$, we have

$$J_n^{1/2} \left(\hat{\underline{f}}_{\underline{K},n} - \underline{f}_{\underline{K},n}(\mathcal{Q}_{I,n}) \right) | \mathcal{Q}_{I,n} \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K,\underline{f}}^{(2)} \right),$$

where the variance-covariance matrix

$$\Sigma_{K,\underline{f}}^{(2)} = J_g \left(\underline{\mu}_{\underline{K}} \right) \Sigma_{K,\underline{C}}^{(2)} J_g \left(\underline{\mu}_{\underline{K}} \right)',$$

where $\Sigma_{K,\underline{C}}^{(2)}$ is defined in (D.2), has entries $\Sigma_{K,\underline{C}}^{(2)}(j, j) = \sigma_{f_j,2}^2 = \frac{\sigma_j^2}{\mu_j} - \sigma_{f_j,1}^2$ for $j = 0, \dots, K$ and $\Sigma_{K,\underline{C}}^{(2)}(j_1, j_2) = -\Sigma_{K,\underline{C}}^{(1)}(j_1, j_2)$ for $j_1, j_2 = 0, \dots, K, j_1 \neq j_2$.

We obtain

$$\Sigma_{K,\underline{f}}^{(1)} = J_g \left(\underline{\mu}_{\underline{K}} \right) \Sigma_{K,\underline{C}}^{(1)} J_g \left(\underline{\mu}_{\underline{K}} \right)' \tag{D.1}$$

and

$$\Sigma_{K,\underline{C}}^{(2)} = E(\text{Var}(\underline{C}_{i,K} | C_{i,\infty})), \tag{D.2}$$

where $C_{i,\infty} = C_{i,0} \prod_{k=0}^{\infty} F_{i,k}$. Note that, due to the law of total variance, we have

$$\Sigma_{K,\underline{C}} = \Sigma_{K,\underline{C}}^{(1)} + \Sigma_{K,\underline{C}}^{(2)} \quad \text{and} \quad \Sigma_{K,\underline{f}} = \Sigma_{K,\underline{f}}^{(1)} + \Sigma_{K,\underline{f}}^{(2)}, \tag{D.3}$$

where $\Sigma_{K,\underline{f}} = \text{diag} \left(\frac{\sigma_0^2}{\mu_0}, \frac{\sigma_1^2}{\mu_1}, \dots, \frac{\sigma_K^2}{\mu_K} \right)$.

Corollary D.2 (Asymptotic normality for products of $\hat{f}_{j,n}$'s conditionally on $\mathcal{Q}_{I,n}$; Corollary C.2(ii,iv) in Steinmetz and Jentsch (2022)). Suppose the assumptions of Theorem D.1 hold. Then, as $n \rightarrow \infty$, the following holds:

(i) For each fixed $K \in \mathbb{N}_0$, unconditionally, we have also joint convergence, that is,

$$\sqrt{I+n+1} \left(\prod_{j=0}^K f_{j,n}(\mathcal{Q}_{I,n}) - \prod_{j=0}^K f_{j,n} \right) \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K,\prod f_j}^{(1)} \right)$$

where $\Sigma_{K,\prod f_j}^{(1)} = J_h(\underline{f}_{\underline{K}}) \Sigma_{K,\underline{f}}^{(1)} J_h(\underline{f}_{\underline{K}})'$ with $J_h(\cdot)$ as defined in (A.3).

(ii) For each fixed $K \in \mathbb{N}_0$, conditionally on $\mathcal{Q}_{I,n}$, we have also joint convergence, that is,

$$\sqrt{I+n+1} \left(\prod_{j=0}^K \hat{f}_{j,n} - \prod_{j=0}^K f_{j,n}(\mathcal{Q}_{I,n}) \right) | \mathcal{Q}_{I,n} \xrightarrow{d} \mathcal{N} \left(0, \Sigma_{K,\prod f_j}^{(2)} \right)$$

where $\Sigma_{K,\prod f_j}^{(2)} = \Sigma_{K,\prod f_j} - \Sigma_{K,\prod f_j}^{(1)}$, where $\Sigma_{K,\prod f_j}$ is defined in Corollary 3.2(ii) in Steinmetz and Jentsch (2022).

Appendix E. Additional simulation results

Table 4

Coverage rates and average interval lengths (in thousand) for *unconditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) for different (true and bootstrap) parametric families, for $I = 10$ and different n in Setup b).

bootstrap parametric family	n	oMB		iMB		aMB	
		true distribution: gamma					
		coverage rate	mean length	coverage rate	mean length	coverage rate	mean length
gamma	10	91%	21,168	91%	21,123	90%	20,913
	20	90%	25,569	90%	25,771	90%	25,409
	30	93%	25,154	93%	25,265	92%	24,907
	40	96%	25,056	95%	24,828	93%	24,637
	50	95%	25,183	96%	25,050	94%	24,983
log-normal	10	91%	21,405	91%	21,702	92%	21,234
	20	92%	25,264	91%	25,195	93%	24,898
	30	94%	24,960	93%	24,929	94%	24,849
	40	96%	24,386	95%	24,565	94%	24,114
	50	97%	24,587	96%	24,579	95%	24,419
truncated normal	10	90%	21,247	90%	22,106	91%	21,047
	20	92%	26,650	93%	25,753	92%	25,100
	30	93%	26,161	95%	25,232	93%	24,695
	40	96%	24,394	97%	24,347	94%	24,053
	50	95%	24,603	96%	24,600	95%	24,248
true distribution: log-normal							
gamma	10	92%	21,618	93%	21,716	93%	20,913
	20	94%	25,529	92%	25,771	94%	25,406
	30	94%	25,101	93%	25,275	94%	24,924
	40	95%	25,017	96%	25,832	95%	24,643
	50	96%	25,139	96%	25,054	95%	24,983
log-normal	10	89%	21,361	89%	21,702	90%	21,226
	20	93%	25,218	93%	25,194	94%	24,891
	30	94%	26,099	92%	26,325	93%	24,950
	40	96%	24,335	96%	24,360	93%	24,106
	50	97%	24,542	95%	24,580	95%	24,414
truncated normal	10	89%	21,216	89%	22,097	89%	21,045
	20	90%	26,588	90%	26,756	90%	25,103
	30	92%	26,099	91%	26,228	92%	24,682
	40	92%	24,357	92%	24,338	93%	24,051
	50	93%	24,558	92%	24,600	93%	24,245
true distribution: truncated normal							
gamma	10	97%	22,134	97%	22,047	95%	21,433
	20	98%	26,369	98%	26,337	96%	26,246
	30	98%	25,723	97%	25,798	96%	25,543
	40	98%	25,279	98%	25,220	96%	25,096
	50	98%	25,375	97%	25,375	96%	25,180
log-normal	10	97%	21,602	97%	21,919	96%	21,536
	20	98%	26,399	98%	26,454	97%	26,202
	30	96%	25,930	97%	25,972	96%	25,887
	40	97%	25,414	96%	25,467	96%	25,128
	50	98%	25,418	97%	25,471	96%	25,324
truncated normal	10	90%	21,927	92%	22,114	93%	21,666
	20	92%	25,987	91%	25,862	92%	25,710
	30	93%	25,251	93%	25,321	94%	25,128
	40	96%	25,617	97%	25,611	95%	25,074
	50	96%	24,819	96%	25,000	95%	24,748

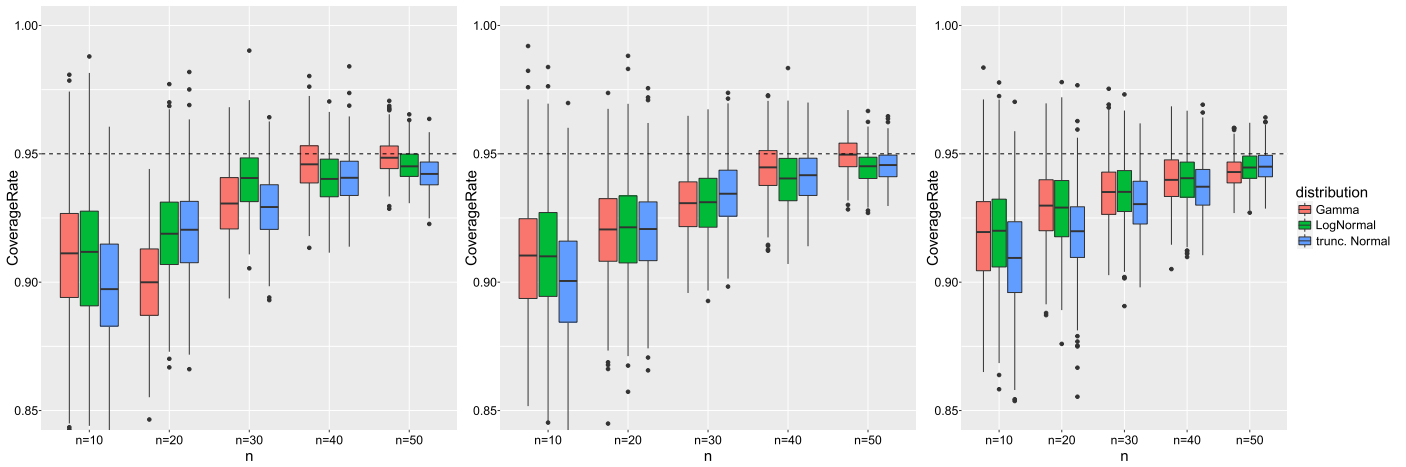


Fig. 4. Box plots of coverage rates for *conditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) with gamma (red), log-normal (green) and truncated normal (blue), respectively. The true parametric family is gamma for $I = 10$ and different n in Setup b).

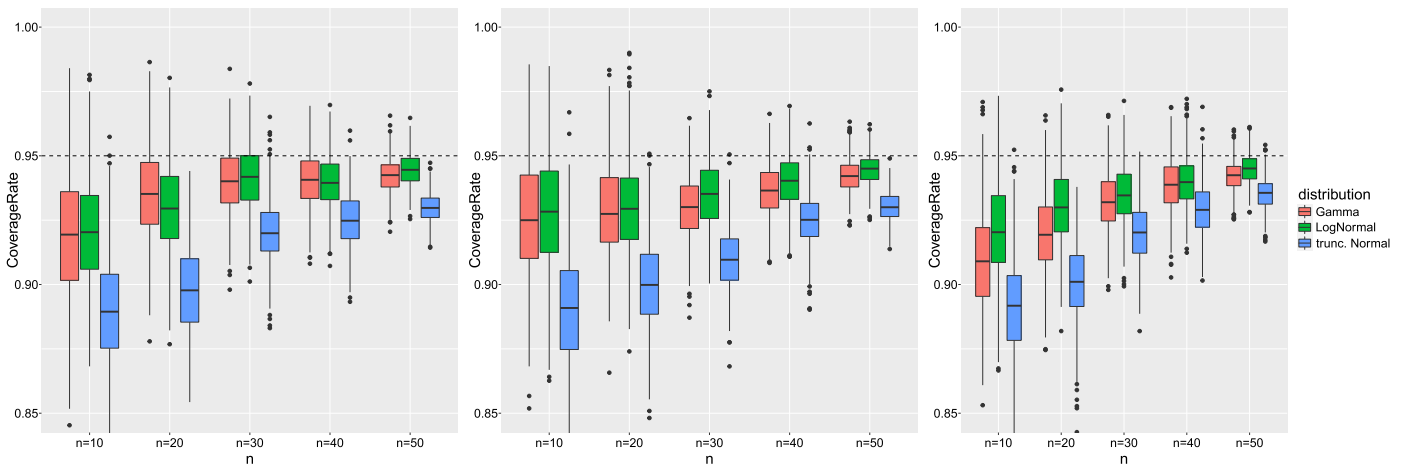


Fig. 5. Box plots of coverage rates for *conditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) with gamma (red), log-normal (green) and truncated normal (blue), respectively. The true parametric family is log-normal for $I = 10$ and different n in Setup b).

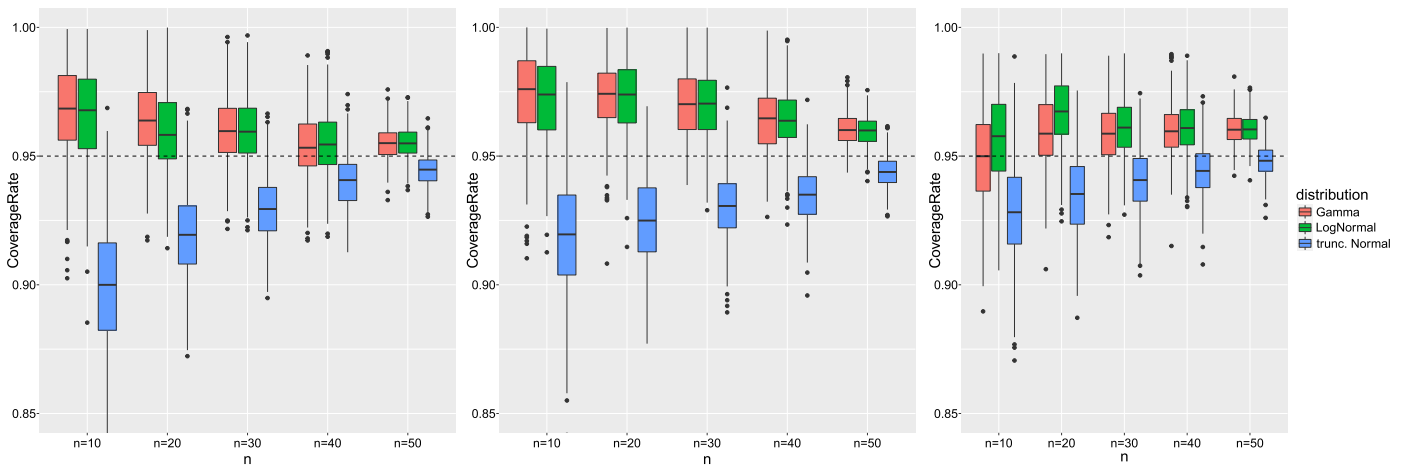


Fig. 6. Box plots of coverage rates for *conditional* 95% bootstrap prediction intervals using the original Mack bootstrap (oMB), the intermediate Mack bootstrap (iMB) and the alternative Mack bootstrap (aMB) with gamma (red), log-normal (green) and truncated normal (blue), respectively. The true parametric family is truncated normal for $I = 10$ and different n in Setup b).

References

- Beutner, E., Heinemann, A., Smeekes, S., 2021. A justification of conditional confidence intervals. *Electronic Journal of Statistics* 15 (1), 2517–2565.
- Björkwall, S., Hössjer, O., Ohlsson, E., 2009. Non-parametric and parametric bootstrap techniques for age-to-age development factor methods in stochastic claims reserving. *Scandinavian Actuarial Journal* 2009 (4), 306–331.
- Björkwall, S., Hössjer, O., Ohlsson, E., 2010. Bootstrapping the separation method in claims reserving. *ASTIN Bulletin: The Journal of the IAA* 40 (2), 845–869.
- Brockwell, P.J., Davis, R.A., 1991. *Time Series: Theory and Methods*. Springer, New York.
- Bruce, N., Chen, C., Dunne, G., Hinder, I., McMurrrough, T., Meyers, G., White, A., Wright, T., 2008. Best estimates and reserving uncertainty. Actuarial profession general insurance (GIRO) convention. Available at <https://www.actuaries.org.uk/system/files/documents/pdf/bhprizegibson.pdf>.
- England, P., 2002. Addendum to “analytic and bootstrap estimates of prediction errors in claims reserving”. *Insurance: Mathematics and Economics* 31 (3), 461–466.
- England, P., Verrall, R., 1999. Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics* 25 (3), 281–293.
- England, P.D., Verrall, R.J., 2006. Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science* 1 (2), 221.
- Gibson, L., Archer-Lock, P., Bruce, N., Collins, A., Dunne, G., Felisky, K., Hamilton, A., Jewell, M., Lo, J., Locke, J., Marshall, D., Nicholson, E., Thomas, L., Wilcox, S., Winer, J., Wright, T., 2007. Best estimates and reserving uncertainty. Actuarial profession general insurance (GIRO) convention. Available at <https://www.actuaries.org.uk/system/files/documents/pdf/bhprizegibson.pdf>.
- Hall, P., 1992. *The Bootstrap and Edgeworth Expansion*. Springer, New York.
- Hartl, T., 2010. Bootstrapping generalized linear models for development triangles using deviance residuals. In: *CAS E-Forum Fall*.
- Mack, T., 1993. Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin: The Journal of the IAA* 23 (2), 213–225.
- Pan, L., Politis, D.N., 2016. Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *Journal of Statistical Planning and Inference* 177, 1–27.
- Papadoditis, E., Shang, H.L., 2021. Bootstrap prediction bands for functional time series. *Journal of the American Statistical Association* 118 (542), 972–986.
- Peremans, K., Segaert, P., Van Aelst, S., Verdonck, T., 2017. Robust bootstrap procedures for the chain-ladder method. *Scandinavian Actuarial Journal* 2017 (10), 870–897.
- Peters, G.W., Wüthrich, M.V., Shevchenko, P.V., 2010. Chain ladder method: Bayesian bootstrap versus classical bootstrap. *Insurance: Mathematics and Economics* 47 (1), 36–51.
- Pinheiro, P.J., Andrade e Silva, J.M., de Lourdes Centeno, M., 2003. Bootstrap methodology in claim reserving. *The Journal of Risk and Insurance* 70 (4), 701–714.
- Renshaw, A.E., Verrall, R.J., 1998. A stochastic model underlying the chain-ladder technique. *British Actuarial Journal* 4 (4), 903–923.
- Steinmetz, J., Jentsch, C., 2022. Asymptotic theory for Mack’s model. *Insurance: Mathematics and Economics* 107, 223–268.
- Tee, L., Käärrik, M., Viin, R., 2017. On comparison of stochastic reserving methods with bootstrapping. *Risks* 5 (1), 2.
- Verdonck, T., Debruyne, M., 2011. The influence of individual claims on the chain-ladder estimates: analysis and diagnostic tool. *Insurance: Mathematics and Economics* 48 (1), 85–98.