Statistical Inference for the Reserve Risk

Dissertation

by

Julia Steinmetz

born in Schweinfurt

in partial fulfillment of
the requirements for the degree of
Doktor der Naturwissenschaften (Dr. rer. nat.)

Submitted:	Dortmund, February 2023
Primary referee:	Prof. Dr. Carsten Jentsch
Secondary referee:	Prof. Dr. Henryk Zähle
Third referee:	Prof. Dr. Markus Pauly

Day of oral examination:	15. June. 2023

I would like to express my gratitude to Prof. Dr. Carsten Jentsch for supporting my research and for giving me this opportunity. In particular, I would like to thank him for the great confidence he placed all the time in me and my doing, whether research or teaching.
Further, I want to thank Prof. Dr. Henryk Zähle and Prof. Dr. Markus Pauly for agreeing to be my second and third supervisor, respectively, and to invest the related time and effort.
I would like to thank my colleagues at the chair of Business and Social Statistics for the helpful conversations regarding my research or life which always gave me new energy.

I thank my parents for all the efforts you have ever made for me, for all the opportunities you have given me, and for the life path you have shown me.
Also, I would like to thank my fiancé Oliver Kuchler for always supporting me.

Abstract

The major part of the liability of an insurance company's balance belongs to the reserves. Reserves are built to pay for all future, known or unknown, claims that happened so far. Also, they ensure the solvability of the insurance company. Hence an accurate prediction of the outstanding claims to determine the reserve is important. For non-life insurance companies, Mack (1993) proposed a distribution-free approach to calculate the first two moments of the reserve. This approach is still popular and a widely used technique in practice. Using a normal approximation together with the derived first two moments, it is widely used to conduct statistical inference including the estimation of large quantiles of the reserve and the determination of the reserve risk. However, Mack's model lacks a rigorous justification in the literature for such a normal approximation for the reserve.
Alternatively, to derive the whole distribution of the reserve and the large quantiles of the reserve to determine the reserve risk, Mack's model was equipped with a tailor-made bootstrap by England and Verrall (2006) which combines a residual based bootstrap non-parametric step with a parametric bootstrap. So far surprisingly, there are no theoretical bootstrap consistency results that justify this approach, although it is widely used in practice.

In this cumulative dissertation, we derive first asymptotic theory for the unconditional and conditional limit distribution of the reserve risk. Therefore, we enhance the assumptions from Mack's model and derive a fully stochastic framework.
The distribution of the reserve risk can be split up into two additive random parts covering the process and parameter uncertainty. The process uncertainty part dominates asymptotically and is in general non-Gaussian distributed unconditional and conditional on the whole observed loss triangle or the last observed diagonal of the loss triangle.
In contrast, the parameter uncertainty part is measurable with respect to the whole observed upper loss triangle. Properly inflated, the parameter uncertainty part is Gaussian distributed conditional on the last observed diagonal of the loss triangle, and unconditional, it leads to a non-Gaussian distribution. Hence, the parameter uncertainty part is asymptotically negligible.
In total, the reserve risk has asymptotically the same distribution as the process uncertainty part since this part dominates asymptotically leading to a non-Gaussian distribution

conditional and unconditional.
These results question the common practice of using a normal approximation for the reserve risk within Mack's model. We illustrate our results through simulations and show that our setup covers cases where the limiting distributions of the reserve risk can deviate significantly from a Gaussian distribution.

Using the theoretical asymptotic distribution results regarding the distribution of the reserve risk, we can now establish bootstrap consistency results, where the derived distribution of the reserve risk serves as a benchmark. Splitting the reserve risk into two additive parts enables a rigorous investigation of the validity of the Mack bootstrap. If the parametric family of distributions of the individual development factors is correctly specified, we prove that the (conditional) distribution of the asymptotically dominating process uncertainty part is correctly mimicked by the proposed Mack bootstrap approach. If not, this is in general not the case. On the contrary, the corresponding (conditional) distribution of the estimation uncertainty part is generally not correctly captured by the Mack bootstrap.
To address this issue, we propose an alternative Mack bootstrap, which uses a different centering and is designed to capture also the distribution of the estimation uncertainty part correctly.
We illustrate our findings in simulations and show that our newly alternatively proposed Mack bootstrap performs superior to the original Mack bootstrap in finite samples.

Zusammenfassung

Der größte Teil der Bilanz einer Versicherungsgesellschaft entfällt auf die Rückstellungen (Reserven). Rückstellungen werden gebildet, um für alle zukünftigen, bekannten oder unbekannten, aber bisher eingetretenen Schäden aufzukommen. Zudem gewährleisten diese die Solvabilität des Versicherungsunternehmens. Daher ist eine präzise Vorhersage der ausstehenden Schäden zur Bestimmung der Rückstellungen wichtig. Für Nichtlebensversicherer hat Mack (1993) einen verteilungsfreien Ansatz zur Berechnung der ersten beiden Momente der Rückstellung vorgeschlagen. Dieser Ansatz ist nach wie vor sehr weit verbreitet und wird in der Praxis sehr häufig verwendet. Unter Verwendung einer Normalapproximation zusammen mit den ermittelten ersten beiden Momenten wird dieser Ansatz häufig verwendet, um statistische Inferenz zu betreiben, einschließlich der Schätzung hoher Quantile der Reserve zur Bestimmung des Reserverisikos. Allerdings fehlt in der Literatur eine stichhaltige Rechtfertigung für eine solche Normalapproximation der Reserve für das Modell von Mack.
Als Alternative zur Herleitung der kompletten Verteilung der Reserve und hoher Quantile der Reserve zur Bestimmung des Reserverisikos wurde das Modell von Mack von England and Verrall (2006) mit einem auf das Modell von Mack zugeschnittenen Bootstrap versehen, der einen nicht-parametrischen Bootstrap-Schritt auf der Basis von Residuen mit einem parametrischen Bootstrap kombiniert. Bislang gibt es keine theoretischen Bootstrap-Konsistenzergebnisse, die diesen Ansatz rechtfertigen, obwohl dieser in der Praxis vielfach verwendet wird.

In dieser kumulativen Dissertation leiten wir zunächst asymptotische Theorie für die unbedingte und bedingte Grenzverteilung des Reserverisikos her. Dazu erweitern wir die Annahmen des Modells von Mack und leiten einen vollständigen stochastischen Modellrahmen ab.
Die Verteilung des Reserverisikos kann in zwei additive Teile aufgeteilt werden, die die Prozess- und die Parameterunsicherheit abdecken. Der Teil, der die Prozessunsicherheit abbildet, dominiert asymptotisch und ist im Allgemeinen nicht-gaußförmig verteilt, und zwar unbedingt und bedingt auf das beobachtete obere Schadendreieck oder die letzte beobachtete Diagonale des Schadendreiecks.
Im Gegensatz dazu ist der Teil, der die Parameterunsicherheit abbildet, messbar in Bezug auf das beobachtete obere Schadendreieck. Mit dem korrekte Faktor multipliziert ist der Parameterunsicherheitsteil bedingt auf die letzte beobachtete Diagonale des Schadendreiecks gaußverteilt bzw. unbedingt nicht gaußverteilt. Daher ist der Parameterunsicherheitsteil asymptotisch vernachlässigbar.
Zusammengenommen hat das Reserverisiko asymptotisch dieselbe Verteilung wie der

Prozessunsicherheitsteil, da dieser Teil asymptotisch dominiert und somit ist das Reserverisiko nicht-gauß verteilt im Allgemeinen, bedingt und unbedingt.
Diese Ergebnisse stellen die gängige Praxis der Verwendung einer Normalapproximation für das Reserverisiko für das Modell von Mack in Frage. Durch Simulationen werden die Ergebnisse veranschaulicht und zeigen, dass das Simulationssetup auch Situationen abdeckt, in denen die Grenzverteilungen des Reserverisikos erheblich von einer Gauß-Verteilung abweicht.

Unter Verwendung der theoretischen, asymptotischen Ergebnisse bezüglich der Verteilung des Reserverisikos werden Bootstrap-Konsistenzergebnisse ermittelt, wobei die abgeleitete Verteilung des Reserverisikos als Vergleichsgröße dient. Die Aufspaltung des Reserverisikos in zwei additive Teile ermöglicht eine präzise Untersuchung der "Validity" des Bootstraps von Macks. Unter der Annhame, dass die parametrische Familie der Verteilungen der einzelnen Abwicklungsfaktoren korrekt bestimmt ist, kann man beweisen, dass die (bedingte) Verteilung des asymptotisch dominierenden Prozessunsicherheitsteils durch den vorgeschlagenen Mack Bootstrap korrekt simuliert wird. Falls nicht, ist dies im Allgemeinen nicht der Fall. Im Gegensatz dazu wird die entsprechende (bedingte) Verteilung des Schätzungsunsicherheitsteils im Allgemeinen nicht korrekt durch den Mack Bootstrap wiedergegeben.
Um dieses Problem zu beheben, wird ein neuer Mack Bootstrap Ansatz vorgeschlagen, der so konzipiert ist, dass die Verteilung des Schätzunsicherheitsteils zudem richtig berücksichtigt wird.
Die Ergebnisse werden in Simulationen veranschaulicht und es wird gezeigt, dass der neu vorgeschlagene Mack Bootstrap dem ursprünglichen Mack Bootstrap in endlichen Stichproben vorzuziehen ist.

Contents

List of Papers

The following paper and preprint are part of this cumulative thesis:

Article 1 Steinmetz, J. and Jentsch, C. (2022). Asymptotic Theory for Mack's Model. Insurance Mathematics and Economics. **107**, 223-268.
DOI: `10.1016/j. insmatheco.2022.08.007` (Copyright Elsevier).
The reuse of this article is granted under the terms of copyright of Elsevier. (`https://www.elsevier.com/about/policies/copyright/permissions`).

Contribution of the author:
The author of this thesis conducted the mathematical proofs and implemented the simulations under Prof. Jentsch's supervision. She also prepared and structured the manuscript with input from Prof. Jentsch.

Article 2 Steinmetz, J. and Jentsch, C. (2022). Bootstrap Consistency for the Mack Bootstrap.

The paper is submitted for publication.

Contribution of the author:
The author of this thesis conducted the mathematical proofs and implemented the simulations under Prof. Jentsch's supervision. She also prepared and structured the manuscript with input from Prof. Jentsch.

Notation

$C_{i,j}$	cumulative claim for accident year i and development year j
$C^*_{i,j}$	Mack bootstrap cumulative claim for accident year i and development year j
$C^+_{i,j}$	alternative Mack bootstrap cumulative claim for accident year i and development year j
$\mathcal{C}_I$	cumulative claim matrix at time I
$\mathcal{C}_{I,n}$	cumulative claim matrix at time I (asymptotic view)
d_2	Mallows' metric between two probability distributions and for $r = 2$
d_K	Kolmogorov distance between two probability distributions
$\mathcal{D}_I$	upper loss triangle at time I
$\mathcal{D}_{I,n}$	upper loss triangle at time I (asymptotic framework)
$\mathcal{D}_{I,\infty}$	upper loss triangle at time I for $n \to \infty$
$\mathcal{D}^+_{I,n}$	upper loss triangle generated by the backward bootstrap
$\mathcal{D}^c_I$	lower loss triangle at time I
f_j	development factor for development year j
$\widehat{f}_j$	estimator of the development factor for development year j
$\widehat{f}_{j,n}$	estimator of the development factor for development year j (asymptotic framework)
$F_{i,j}$	individual development factor for accident year i and development year j
$F^*_{i,j}$	forward Mack bootstrap individual development factor for accident year i and development year j
$F^+_{i,j}$	forward alternative Mack bootstrap individual development factor for accident year i and development year j
$\widehat{f}^*_j$	Mack bootstrap estimator of the development factor for development year j
$\widehat{f}^*_{j,n}$	Mack bootstrap estimator of the development factor for development year j (asymptotic framework)
$\widehat{f}^+_{j,n}$	alternative Mack bootstrap estimator of the development factor for development year j (asymptotic framework)

$\mathcal{F}^*_{I,n}$	set of $F^*_{i,j}$ for $j = 0, \ldots, I+n-1$ and $i = -n, \ldots, I-j-1$
$F_{i,j} \vert C_{i,j}$	$F_{i,j}$ conditional on $C_{i,j}$
$G^+_{i,j}$	backward alternative Mack bootstrap individual development factor for accident year i and development year j
L_2	convergence in the 2nd mean
$\mathcal{L}(\cdot)$	distribution
$\mathcal{L}^*(\cdot)$	Mack bootstrap distribution conditional on $\mathcal{D}_{I,n}$
$\mathcal{L}^+(\cdot)$	alternative Mack bootstrap distribution conditional on $\mathcal{D}_{I,n}$
μ_0	expectation of $C_{i,0}$
μ_j	expectation of $C_{i,j}$
μ_∞	asymptotic expectation of $C_{i,j}$ for $j \to \infty$
$\mathbb{N}_0$	natural numbers including 0
$o_P(1)$	convergence in probability to 0
$O_P(1)$	stochastic boundness by a constant
$o(1)$	convergence to 0
$O(1)$	boundness by a constant
$\mathcal{Q}_k$	diagonal at time k, $k = 0, \ldots .I$
$\mathcal{Q}_I$	last diagonal at time I
$\mathcal{Q}_{I,n}$	last diagonal at time I (asymptotic framework)
$\mathcal{Q}_{I,\infty}$	last diagonal at time I for $n \to \infty$
$\mathcal{Q}^*_{I,n}$	Mack bootstrap diagonal which is equal to $\mathcal{Q}_{I,n}$
$\mathcal{Q}^+_{I,n}$	alternative Mack bootstrap diagonal which is equal to $\mathcal{Q}_{I,n}$
$\widehat{r}_{i,j}$	estimated 'residual' for accident year i and development year j
$\widetilde{r}_{i,j}$	centered and re-scaled residual for accident year i and development year j
$r^*_{i,j}$	bootstrap 'error' for accident year i and development year j
R_i	individual reserve for accident year i
$R_{i,n}$	individual reserve for accident year i (asymptotic framework)
R_I	total reserve
$R_{I,n}$	total reserve for accident year i (asymptotic framework)
$R_{I,n} - \widehat{R}_{I,n}$	predictive root of the reserve
$(R_{I,n} - \widehat{R}_{I,n})_1$	first part of the predictive root of the reserve
$(R_{I,n} - \widehat{R}_{I,n})_2$	second part of the predictive root of the reserve
$R^*_{I,n} - \widehat{R}_{I,n}$	Mack bootstrap predictive root of the reserve
$(R^*_{I,n} - \widehat{R}_{I,n})_1$	first part of the Mack bootstrap predictive root of the reserve
$(R^*_{I,n} - \widehat{R}_{I,n})_2$	second part of the Mack bootstrap predictive root of the reserve
$R^+_{I,n} - \widehat{R}^+_{I,n}$	alternative Mack bootstrap predictive root of the reserve
$(R^+_{I,n} - \widehat{R}^+_{I,n})_1$	first part of the alternative Mack bootstrap predictive root of the reserve

$(R^+_{I,n} - \widehat{R}^+_{I,n})_2$	second part of the alternative Mack bootstrap predictive root of the reserve
σ_j^2	variance parameter for development year j
$\widehat{\sigma}_j^2$	estimator of the variance parameter for development year j
$\widehat{\sigma}_{j,n}^2$	estimator of the variance parameter for development year j (asymptotic framework)
τ_0^2	variance of $C_{i,0}$
τ_j^2	variance of $C_{i,j}$
τ_∞^2	asymptotic variance of $C_{i,j}$ for $j \to \infty$
$\xrightarrow{d}$	convergence in distribution
$\xrightarrow{p}$	convergence in probability
$X \sim (\mu, \sigma^2)$	X belongs to a certain distribution with mean μ and variance σ^2
$X \sim \mathcal{G}$	X belongs to the distribution $\mathcal{G}$
P^*, E^*, Var^*, Cov^*	bootstrap probability, expectation, variance, and covariance conditional on $\mathcal{D}_I/\mathcal{D}_{I,n}$
E^+, Var^+	alternative Mack bootstrap expectation, variance conditional on $\mathcal{D}_{I,n}$
$E_\mathcal{Q}, Var_\mathcal{Q}, Cov_\mathcal{Q}$	expectation, variance, and covariance conditional on $\mathcal{Q}_I/\mathcal{Q}_{I,n}$

List of Abbreviations

aMB	alternative Mack bootstrap
CDR	Claims Developemnt Result
cdf	cumulative distribution function
CLM	Chain Ladder Model
CLT	Central Limit Theorem
GLM	Generalized Linear Model
IBNR	incurred but not reported yet
i.i.d.	identical, independently distributed
iMB	intermediate Mack bootstrap
MSEP	mean squared error of prediction
oMB	orginal Mack bootstrap
pdf	probability density function
RBNS	reported but not settled yet
RMMSE	root of the overall mean of the mean squared error
WLLN	Weak law of Large Numbers
wrt	with respect to

List of Figures

List of Tables

1 Introduction

Due to the law, insurance companies are committed to setting up capital to meet their future liabilities to policyholders related to their earlier and current insurance contracts. This capital is referred to as provisions for outstanding claims or reserves. For non-life insurance companies, reserves are often the major part of the liability side of the balance sheet. Non-life insurance contracts are usually one-year contracts and have to renew every year. The fundamental issue is that the claims' actual sizes are unknown when the reserves have to be set. In general, we distinguish between two major claim types:

- Claims that are incurred but not reported yet. We call them *IBNR*-claims.
- Claims that are reported but not finally settled, i.e., the exact amount of the claims are not ultimately determined yet. We call them *RBNS*-claims.

In this thesis, we do not separately model *IBNR*- and *RBNS*- claims. We use a model to predict the outstanding claim amount in total for one business line at a certain point in time.

The typical loss adjustment process for a non-life insurance company is described in Figure 1. The accident date, the reporting date, and the payout date of the amount of the claims does not need to coincide (cf. Wüthrich and Merz (2008) for more details). The duration of the process of a loss adjustment has several reasons.

- There is always a certain period between the occurrence of an insured event and the reporting of this event. The time difference between the occurrence of a claim and the notification of the claim can take several years.
- After an insured event has been reported to the relevant insurance company, it can sometimes take several years before the full extent of the claim is known. Hence, it can take several years until a claim is closed.
- After an insurance claim has been closed, the insurance claim may have to be reopened. Further liabilities may be imposed on the insurance company, which may have to be settled before the insurance case can be closed again.

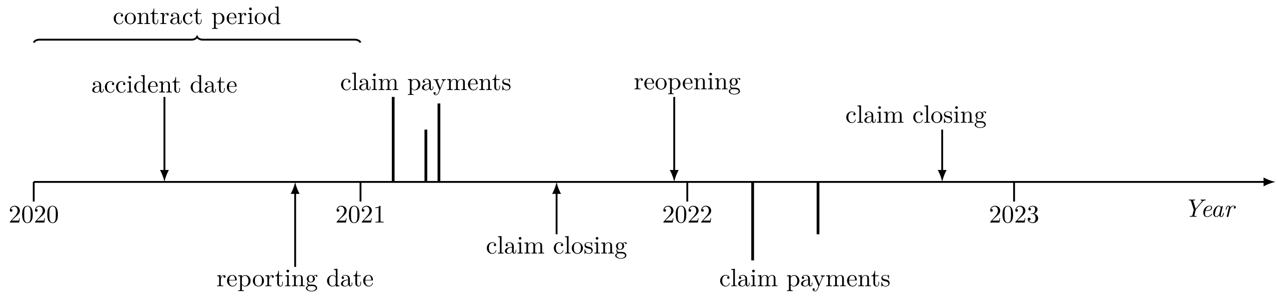

Figure 1: Typical process of loss adjustment for a non-life insurance.

Further, we distinguish between two types of insurance lines, short and long term, depending on how long it takes until a claim is finally closed. Claims that belong to short term business lines are paid out around the first 5 years after the contract period. For example, fire and water insurance contracts belong to this type. In contrast, for long term business lines, it can take decades until a claim is finally closed, e.g., liability insurance contracts. For modeling, we do not distinguish between the types of lines of business.

The process of prediction of outstanding claims is called reserving. Actuaries need to set adequate reserves based on historical data to ensure the insurance company's solvency. The careful calculation of the reserves is important. If the reserves are underestimated, the insurance company may not be solvent for its undertakings. In contrast, if the reserves are overestimated, the insurance company unnecessarily holds additional capital instead of using this capital for other purposes, e.g., investments with higher risks and, hence, potentially higher returns.

In this regard, the reserving actuaries are more interested in a reasonable reserve range than the best estimate of the reserve. The so-called reserve risk covers the uncertainty of the prediction of the claims such that it covers the risk that the best estimate of the reserve is not sufficient to pay for all outstanding claims.
Traditional deterministic algorithms are often sufficient for the best estimate of outstanding claims and the overall variability around the best estimate of the reserve but not sufficient in deriving the (conditional) limit distribution of the reserve and especially estimating large quantiles of the reserve. Therefore, over the past three decades, stochastic claims reserving methods have been developed extensively. Often traditional methods have been equipped with a tailor-made bootstrap algorithm to simulate the distribution of the reserve and to get an estimate for large quantiles.

Section 1.1 discusses the concept and definition of the reserve risk. Section 1.2 introduces the notation and Mack's model. This model determines the best estimate of the reserve and its standard deviation. Mack's model is used in Chapters 2 and 3 as a benchmark model. Section 1.3 deals briefly with the thematic environments of the main contents of this dissertation. Section 1.4 contains the extended summaries of the two articles. Afterward, in Chapters 2 and 3, the articles are attached. Chapter 4 concludes and gives an outlook for further potential research.

For a more detailed introduction to claims reserving in non-life insurance companies, we refer to the monographs Wüthrich and Merz (2008) and Hindley (2017).

1.1 Reserve Risk

The reserve risk and the premium risk are the major risks of a non-life insurance company. The premium risk is the risk that losses in an insurance period will exceed the premiums collected for that period. The premium risk is not further considered in this dissertation. In contrast, the reserve risk is defined as the risk that the best estimate of the reserve for known and unknown claims, that happened in the past, is not sufficient to pay for all outstanding claims. Here, the term best estimate represents the expected value of the range of potential outcomes for future claims. Hence, the reserve risk covers the fluctuation around the expected value. Therefore, we need first to set suitable assumptions, define a model for the reserve, find a method to obtain the best estimate of the reserves, and assess the uncertainty around the best estimate. The risk is often calculated by the *Mean Squared Error of Prediction* (MSEP). For a single future claim payment $S_{i,j}$ (random variable) for the accident year i and development year j, the MSEP is the expected squared difference between the future claim $S_{i,j}$ and its point prediction $\widehat{S}_{i,j}$. If we apply the two assumptions that the $S_{i,j}$ and $\widehat{S}_{i,j}$ are unbiased, i.e. $E(\widehat{S}_{i,j}) = E(S_{i,j})$ and uncorrelated, i.e. $E(S_{i,j}\widehat{S}_{i,j}) = E(\widehat{S}_{i,j})E(\widehat{S}_{i,j})$ then the prediction variance of the claim can be decomposed into two additive parts, covering the process and estimating uncertainty.

$$
\begin{aligned}
E((S_{i,j} - \widehat{S}_{i,j})^2) &= E((S_{i,j} - E(S_{i,j}))^2) + E((\widehat{S}_{i,j} - E(\widehat{S}_{i,j}))^2) \\
&= \text{ process variance } \quad + \text{estimation variance.}
\end{aligned}
$$

Consequently, the prediction variance can be estimated separately determining the process variance and the estimation variance. The process variance is the uncertainty that arises from the actual claims development process which is described by the chosen stochastic model. The estimation variance, on the other hand, describes the uncertainty in the parameter estimation.

The goal of many reserving methods (e.g. Mack (1993), Buchwalder et al. (2006), England and Verrall (1999)) is to reasonably estimate the MSEP of the reserve. In order to draw conclusions about the whole distribution of the reserve risk and its large quantiles the derived estimated MSEP and the best estimate are equipped with some distributional assumptions,

Accident Year i	Development Year j					
	0	1	2	$\cdots$	$I-1$	I
0	$C_{0,0}$	$C_{0,1}$			$C_{0,I-1}$	$C_{0,I}$
1	$C_{1,0}$	$\cdots$			$C_{1,I-1}$	$C_{1,I}$
			$\cdots$			
$\cdots$	$\cdots$			$\cdots$		$\cdots$
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$		$\cdots$		$C_{I-1,I}$
I	$C_{I,0}$	$C_{I,1}$		$\cdots$		$C_{I,I}$

Table 1: Upper (observed) loss triangle $\mathcal{D}_I$ and (unobserved) lower loss triangle $\mathcal{D}_I^c$ (in light gray).

e.g. normal distribution. The literature lacks a justification for the applied distribution. Also, a parametric assumption imposed on the distribution of the reserve might be too restrictive and can result in a misleading estimate of the tail of the distribution of the reserve.

1.2 Notation and Mack's Model

To estimate the outstanding claims and their reserve, insurance companies summarize all observed individual claims data of a business line in an upper loss triangle organized as shown in Table 1. Its entries, the cumulative amount of claims $C_{i,j}$, are sorted by their accident year i (vertical axis) and their development year j (horizontal axis) for $i = 0, \ldots, I$ and $j = 0, \ldots, J$ with $i + j \leq I$. In the following, we have $I = J$. Hence, the (observed) loss triangle contains the cumulative claims $C_{i,j}$ that have already been observed up to year I. It constitutes the available data basis and is denoted by

$$\mathcal{D}_I = \{C_{i,j} | i, j = 0, \ldots, I, \; 0 \leq \; i + j \leq I\} . \tag{1.1}$$

To predict the outstanding claims, we need to estimate claims which are unobserved at time I. Therefore, we augment the (observed) *upper* loss triangle by an unobserved *lower* triangle $\mathcal{D}_I^c = \{C_{i,j} | i, j = 0, \ldots, I, \; i + j > I\}$ (cf. Table 1 light gray lower triangle) that contains all future claims that have not been observed (yet) up to time I.
For each accident year i, the main interest lies in the reserves (loss liabilities) $R_{i,I}$ at terminal time I, which is the difference of the so-called ultimate claim $C_{i,I}$ and the last observed diagonal element $C_{i,I-i}$ of the loss triangle, where we assume that after $I + 1$

years the claims are finally settled. Precisely, we define

$$R_{i,I} = C_{i,I} - C_{i,I-i}, \quad i = 0, \dots, I, \tag{1.2}$$

and denote the aggregated total amount of the reserve by

$$R_I = \sum_{i=0}^{I} R_{i,I}, \tag{1.3}$$

noting that $R_{0,I} = C_{0,I} - C_{0,I} = 0$ by construction.

In practice, a widely used distribution-free model to determine the mean of the reserve for each accident year i and in total is the so-called *Chain Ladder Model* (CLM). The estimates of the reserve by the CLM are often referred to as the best estimates of the reserve. Mack (1993) enhanced the CLM by the assumption regarding the variance and stated a formula for the MSEP of the reserve. This model is often denoted as *Mack's model* and relies on three fundamental model assumptions without any distributional assumptions summarized as follows.

Assumption 1.1 (Mack's Model)

(i) There exist so-called development factors $f_0, \dots, f_{I-1} > 0$ *such that*

$$E(C_{i,j+1}|C_{i,j}, \dots, C_{i,0}) = f_j C_{i,j} \tag{1.4}$$

for all $i = 0, \dots, I$ *and* $j = 0, \dots, I-1$.

(ii) There exist variance parameters $\sigma_0^2, \dots, \sigma_{I-1}^2 > 0$ *such that*

$$Var(C_{i,j+1}|C_{i,j}, \dots, C_{i,0}) = \sigma_j^2 C_{i,j} \tag{1.5}$$

for all $i = 0, \dots, I$ *and* $j = 0, \dots, I-1$.

(iii) The cumulative payments are stochastically independent over the accident years $i = 0, \dots, I$, *that is* $(C_{i,0}, \dots, C_{i,I})$ *and* $(C_{k,0}, \dots, C_{k,I})$ *are independent for* $i \neq k$.

Using Assumption 1.1 (i) and (ii), we get recursively the mean and the variance of the ultimate claim $C_{i,I}$ for $i = 0, \dots, I$ conditional on the observed loss triangle:

$$E(C_{i,I}|\mathcal{D}_I) = E(C_{i,I}|C_{i,I-i}, \dots, C_{i,0}) = C_{i,I-i} \prod_{j=I-i}^{I-1} f_j, \tag{1.6}$$

$$Var(C_{i,I}|\mathcal{D}_I) = Var(C_{i,I}|C_{i,I-i}, \dots, C_{i,0}) = C_{i,I-i} \sum_{j=I-i}^{I-1} f_{I-i} \dots f_{j-1} \sigma_j^2 f_{j+1}^2 \dots f_{I-1}^2, \tag{1.7}$$

where $\prod_{j=I}^{I-1} f_j = 1$ and $\sum_{j=I}^{I-1} f_{I-i} \dots f_{j-1}\sigma_j^2 f_{j+1}^2 \dots f_{I-1}^2 = 0$ by construction. As all development factors f_j and variance parameters σ_j^2 are generally unknown, they have to be estimated from the available data at time I, i.e. the observed upper loss triangle $\mathcal{D}_I$. By the CLM the development factors $f_0, \dots, f_{I-1}$ can be estimated by $\widehat{f}_0, \dots, \widehat{f}_{I-1}$, where

$$\widehat{f}_j = \frac{\sum\limits_{i=0}^{I-j-1} C_{i,j+1}}{\sum\limits_{i=0}^{I-j-1} C_{i,j}} \tag{1.8}$$

for $j = 0, \dots, I-1$. As shown by Mack (1993), $\widehat{f}_j$ are unbiased estimators for f_j, i.e. $E(\widehat{f}_j) = f_j$ for $j = 0, \dots, I-1$, and $\widehat{f}_0, \dots, \widehat{f}_{I-1}$ are pairwise uncorrelated, i.e. $Cov(\widehat{f}_j, \widehat{f}_k) = 0$ for all $j \neq k$.

By plugging-in the $\widehat{f}_j$'s in (1.6), conditional on $C_{i,I-i}, \dots, C_{i,0}$, the best estimate $\widehat{C}_{i,I}$ (point predictor) of the ultimate claim $C_{i,I}$ becomes

$$\widehat{C}_{i,I} = C_{i,I-i} \prod_{j=I-i}^{I-1} \widehat{f}_j, \quad i = 0, \dots, I. \tag{1.9}$$

Consequently, given $C_{i,I-i}$, the best estimate $\widehat{R}_{i,I}$ of the reserve $R_{i,I}$ of accident year i defined in (1.2) is given by

$$\widehat{R}_{i,I} = \widehat{C}_{i,I} - C_{i,I-i} = C_{i,I-i}\left(\prod_{j=I-i}^{I-1} \widehat{f}_j - 1\right), \quad i = 0, \dots, I, \tag{1.10}$$

and the best estimate $\widehat{R}_I$ of the total reserve R_I defined in (1.3) computes to

$$\widehat{R}_I = \sum_{i=0}^{I} \widehat{R}_{i,I} = \sum_{i=0}^{I} C_{i,I-i}\left(\prod_{j=I-i}^{I-1} \widehat{f}_j - 1\right) \tag{1.11}$$

noting that $\widehat{R}_{0,I} = 0$ due to $\prod_{j=I}^{I-1} \widehat{f}_j = 1$.

Furthermore, Mack (1993) proposed to estimate the variance parameters $\sigma_0^2, \dots, \sigma_{I-2}^2$ by $\widehat{\sigma}_0^2, \dots, \widehat{\sigma}_{I-2}^2$, where

$$\widehat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2, \quad j = 0, \dots, I-2, \tag{1.12}$$

which are unbiased estimators for σ_j^2, i.e. $E(\widehat{\sigma}_j^2) = \sigma_j^2$. Note that (1.12) does not cover the estimation of σ_{I-1}^2. If $\widehat{f}_{I-1} = 1$ and if the development of the claim is believed to be finished after $I-1$ years, we can set $\widehat{\sigma}_{I-1}^2 = 0$. Alternatively, Mack (1993) proposed to extrapolate $\widehat{\sigma}_{I-1}^2$ by using a log-linear regression or by setting $\widehat{\sigma}_{I-1}^2 = min(\widehat{\sigma}_{I-2}^4/\widehat{\sigma}_{I-3}^2, min(\widehat{\sigma}_{I-3}^2, \widehat{\sigma}_{I-2}^2))$.

Of particular interest is the difference between the stochastic (unobserved) reserve R_I and its best estimate $\widehat{R}_I$ (based on the observed data $\mathcal{D}_I$). We call the difference in the following the *predictive root of the reserve* and it computes to

$$R_I - \widehat{R}_I = \sum_{i=0}^{I} \left(C_{i,I} - C_{i,I-i} - \left(\widehat{C}_{i,I} - C_{i,I-i} \right) \right) = \sum_{i=0}^{I} \left(C_{i,I} - \widehat{C}_{i,I} \right). \qquad (1.13)$$

The mean squared error of prediction of $\widehat{R}_I$ given $\mathcal{D}_I$ is defined by

$$MSEP\left(\widehat{R}_I|\mathcal{D}_I\right) = E\left((R_I - \widehat{R}_I)^2|\mathcal{D}_I\right), \qquad (1.14)$$

and Mack (1993) derived the following formula for the MSEP:

$$MSEP\left(\widehat{R}_I|\mathcal{D}_I\right) = \sum_{i=0}^{I} \left(\widehat{C}_{i,I}^2 \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\widehat{C}_{i,j}} + \widehat{C}_{i,I}^2 \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}} \right) \qquad (1.15)$$

$$+ 2 \sum_{\substack{i,l=0 \\ i<l}}^{I} \left(\widehat{C}_{i,I} \widehat{C}_{l,I} \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}} \right), \qquad (1.16)$$

where $\widehat{C}_{i,j} = C_{i,I-i} \widehat{f}_{I-i} \cdots \widehat{f}_{j-1}$ for $j > I - i$ are the estimated values of the future claims $C_{i,j}$ and $\widehat{C}_{i,I-i} = C_{i,I-i}$. In the above formula, the summands in (1.15) consist of two terms corresponding to the *process variance* and *estimation variance* (of parameter estimates) of $R_{i,I}$, respectively. The expression in (1.16) reflects the linear dependence between $\widehat{R}_{i,I}$ and $\widehat{R}_{l,I}$, $i \neq l$. It contains their covariances and belongs also to the *estimation variance*. The MSEP can be rewritten as

$$MSEP\left(\widehat{R}_I|\mathcal{D}_I\right) = \sum_{i=0}^{I} \left(C_{i,I-i} \sum_{j=I-i}^{I-1} \widehat{\sigma}_j^2 \prod_{k=I-i}^{j-1} \widehat{f}_k \prod_{l=j+1}^{I-1} \widehat{f}_l^2 + C_{i,I-i}^2 \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}} \prod_{\substack{l=I-i \\ l \neq j}}^{I-1} \widehat{f}_l^2 \right)$$

$$+ 2 \sum_{\substack{i,l=0 \\ i<l}}^{I} \left(C_{i,I-i} C_{l,I-l} \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}} \prod_{n=I-l}^{I-i-1} \widehat{f}_n \prod_{\substack{m=I-i \\ m \neq j}}^{I-1} \widehat{f}_m^2 \right).$$

The term which belongs to the process variance follows straightforwardly by plugging $\widehat{f}_j$'s and $\widehat{\sigma}_j^2$'s in (1.7). The calculation of the estimation uncertainty is not straightforward as for the process uncertainty since $\widehat{R}_I$ is measurable with respect to $\mathcal{D}_I$. Therefore, Mack approximated $(f_k - \widehat{f}_k)^2$ for $k = 0, \ldots I - 1$ by $E((f_k - \widehat{f}_k)^2|\mathcal{B}_k)$, where $\mathcal{B}_k := \{C_{i,j}|j = 0, \ldots, k, i + j \leq I\}$, which means the denominator of $\widehat{f}_j$ is kept fixed, to derive a formula for the estimation variance. Buchwalder et al. (2006) proposed a slightly different formula for the estimation variance.

However, the knowledge of the first two moments of the reserve is not sufficient to determine the whole distribution of the predictive root of the reserve, i.e. $R_I - \widehat{R}_I$. In practice, it is important to estimate large quantiles of the (conditional) distribution of the predictive root of the reserve to approximate the *reserve risk*. For this purpose, a common approach is to assume a certain parametric family of distributions for either the reserve of a single accident year or of the total reserve and to estimate their distributions by estimating their parameters. In this regard, it is more common to model the distribution of the *total* reserve to take diversification effects between the single accident years into account. However, a parametric assumption imposed on the reserve might be too restrictive and result in a misleading estimate of the tail of the distribution of the predictive root of the reserve.
Hence, a non-parametric analysis to derive the limiting distribution of the reserve without restricting considerations to any parametric assumptions might be more beneficial in practice. However, given Mack's distribution-free reserving framework, this is not possible. Therefore, England and Verrall (2002) suggest a tailor-made bootstrap algorithm for Mack's Model to simulate the distribution of the reserve. Their approach combines a residual based bootstrap non-parametric step with a parametric bootstrap. Although it is widely used in applications to estimate the reserve risk, no theoretical bootstrap consistency results exist that justify this approach.
So far, it was not possible to derive bootstrap consistency results for Mack's Model since the true asymptotic (conditional) distribution of the reserve was unknown.

In Chapter 2 (Paper 1) we define a general stochastic framework that allows deriving asymptotic theory for Mack's Model. We obtain separately the limit (conditional) distribution of the process and the estimation uncertainty and then jointly to get the limit distribution of the predictive root of the reserve.
In Chapter 3 (Paper 2) we use the derived distribution as a benchmark to show bootstrap consistency results.

1.3 Predictive Distribution and its Bootstrap Application

We want to predict the outstanding claims based on the observed claim data to be able to get an estimate for the reserve.
Therefore, we introduce briefly in the following the main statistical ideas which are used in Chapter 2 (Paper 1) to derive the limit (conditional) distribution of the predictive root of the reserve and in Chapter 3 (Paper 2) to derive bootstrap consistency results.

Predictive Distribution

We often think of a prediction as a single scalar (or vector) value that the model predicts will be the output from a given input. The prediction from a model, e.g., regression or time series, is a probability distribution of the values that could be the output. The single prediction we are used to seeing is often the mean of that distribution.

In the following, we discuss a simple example of a predictive distribution.

We denote X as a random variable, that we want to predict based on the observed data. We assume that the distribution of X is completely specified except for the unknown parameters μ and σ^2, where $E(X) = \mu$ and $Var(X) = \sigma^2$. Since the parameters μ and σ^2 of the distribution of X are unknown, they must be estimated from the observations. We estimate μ by the sample mean of the observations and denote it with $\widehat{\mu}$ and σ^2 by the sample variance and denote it with $\widehat{\sigma}^2$. We denote $\widehat{X}$ as the point estimator for X and we get here $\widehat{X} = \widehat{\mu}$. Of particular interest is the difference between the stochastic (unobserved) X and its best estimate $\widehat{X}$ (based on the observed data), which is denoted as the predictive root. By adding and subtracting μ we get

$$X - \widehat{X} = X - \widehat{\mu} = (X - \mu) - (\widehat{\mu} - \mu). \tag{1.17}$$

We call in the following the (conditional) distribution of $X - \mu$ process uncertainty, which depends only on the distributional class of X, and the (conditional) distribution of $\widehat{\mu} - \mu$ estimation uncertainty. Because the number of observations is finite, the estimate $\widehat{\mu}$ is also subject to random fluctuations. Often, although both distributions depend on the observations, the (conditional) limit distributions can be derived separately for each part and then jointly to derive the limit distribution of the predictive root.
In addition, if we get that $\widehat{\mu}$ has a rate of convergence a_n, i.e., $a_n(\widehat{\mu} - \mu)$ has a well-defined, non-trivial asymptotic distribution where $a_n \to \infty$ as $n \to \infty$ and n denotes the number of observations, the estimation uncertainty part becomes asymptotically negligible. The process uncertainty part will asymptotically dominate since the estimation part has to be properly inflated. In total, the limit distribution of the predictive root of the reserve will depend asymptotically on the limit distribution of the process uncertainty part.

Prediction models should consider the process uncertainty and the uncertainty involved in the parameter estimation, which is especially beneficial in the finite sample application, although ladder one is asymptotically negligible.

Mack's formula for the MSEP reflects the process and estimation uncertainty, as seen in (1.14). The prediction variance of the reserve is split into two separate additive parts capturing the process and estimation uncertainty.

So far, the predictive distribution for the reserve based on Mack's model is not known. The assumptions in Mack's model are not sufficient to derive the full distribution of the individual reserve $R_{i,I}$ for each accident year i and the total reserve R_I given the observed data. In Mack's model, there are no assumptions made regarding the distribution of the claims just about the first two conditional moments. The independence assumption allows that the limit distribution of the reserve for each accident year can be derived separately and the limit distributions can be just added up to get the limit distribution of the total reserve.
In addition, the asymptotic distribution of the estimators of the parameters in Mack's model - development factors and variance parameters - are unknown. Pešta and Hudecová (2012) showed only that $\widehat{f}_j$ is a consistent estimator for f_j for any fixed j, i.e. $\widehat{f}_j$ converges in probability to f_j, if $\sum_{i=0}^{I} C_{i,0} \to \infty$ as $I \to \infty$. For the variance parameters, no such consistency results exist in the literature.

Additional assumptions are necessary to derive

- a fully stochastic model for the claims,
- the asymptotic limit distributions for the estimators of the parameters, and
- the (conditional) asymptotic limit distribution of the process and estimation uncertainty, and then jointly for the predictive root of the reserve.

Also, suitable assumptions regarding the asymptotic behavior of the parameters in Mack's model are required since consistency results of the parameters can only be obtained for a fixed number of development years j and an increasing number of accident years I.

In the last decades, so-called bootstrap approaches have become popular among actuaries to approximate the MSEP and to approximate the distribution and high quantiles of the predictive root of the reserve by simulations, since the limit distribution of the reserve is unknown.

Bootstrap Approach

The most popular one is the bootstrap scheme first introduced by Efron (1979) for independent and identically distributed (i.i.d.) observations. For i.i.d. random variables $X_1, \ldots, X_n$, the bootstrap idea is briefly described as follows. To approximate the distribution of some statistic $T_n = T_n(X_1, \ldots, X_n)$, we resample randomly with replacement from the original observations $X_1, \ldots, X_n$ n times to get a bootstrap data set $X_1^*, \ldots, X_n^*$. We use the generated data set to compute the corresponding statistic $T_n^* = T_n(X_1^*, \ldots, X_n^*)$. We repeat this procedure B- times, where B is large, and the empirical distribution of these B values of T_n^* is used to approximate the desired distribution. The proceeding described above can only be applied with the i.i.d. assumption.

Since the claims $C_{i,j}$ for accident year i and development year j for $i + j \leq I$ are usually not independent and not identically distributed, Efron's bootstrap can not be applied here. Therefore, we need another bootstrap type.
The Mack bootstrap proposed by England and Verrall (2006) combines a residual-based non-parametric resampling step together with a parametric bootstrap.

Parametric Bootstrap

A parametric assumption is used to generate the resampling sample X_i^* for all $i = 1, \ldots, n$ through

$$X_i^* \sim F \quad \text{for } i = 1, \ldots, n,$$

where F is a suitable distribution that depends on the parameters of the observed data like the sample mean $\widehat{\mu} = \overline{X} = \frac{1}{n} \sum_{i=1}^n X_i$ and variance $\widehat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X})^2$. As above, we use the generated data set to compute the corresponding bootstrap statistic $T_n^* = T_n(X_1^*, \ldots, X_n^*)$. We repeat this procedure B- times, where B is large, and the empirical distribution of these B values of T_n^* is used to approximate the desired distribution.

To apply a parametric bootstrap to Mack's model, we have to enhance the assumptions from Mack's model by additional assumptions such that the claims $C_{i,j}$ for all i, j can be described by a fully stochastic model together with distributional assumptions.

Residual Bootstrap

In general, residual bootstrap methods have in common that some parametric (e.g., linear regression, generalized linear regression) model is fitted to the data at first, and the classical i.i.d. bootstrap is applied to the estimated residuals afterward, which are assumed to be i.i.d. random variables at least approximately. Also, this bootstrap type is referred to as a non-parametric bootstrap approach.
England and Verrall (1999) proposed a residual-based bootstrap algorithm for a generalized linear model (GLM) framework for claims to estimate the MSEP. Later, they suggested using their approach to estimate large quantiles of the reserve (cf. England (2002) and England and Verrall (2006)). England and Verrall (2006) also proposed a tailor-made bootstrap for Mack's model, whereas a first step they propagate a residual bootstrap, although Mack's model is not defined by residuals. A residual bootstrap approach for loss triangles allows resampling the residuals across the whole loss triangle, i.e. resampling is done independent of the accident year i and development year j. However, the same estimators of the parameters in Mack's model are obtained for a weighted regression model with standard normal distributed errors. Nevertheless, the normal assumption may lead to negative claims $C_{i,j}$, which violates the recursive model structure of Mack's model.

Besides the bootstrap type, an appropriate bootstrap approach for a predictive model should also mimic the process and the estimation uncertainty correctly.

Bootstrap Consistency

We use the asymptotic (conditional) limit distribution of the process and estimation uncertainty as a benchmark distribution to derive bootstrap consistency results.
We call a bootstrap asymptotically valid if the asymptotic dominating limit distribution of the process uncertainty part is mimicked correctly by the bootstrap approach. If in addition, the bootstrap is also able to mimic the asymptotic limit distribution of the parameter uncertainty part, we call it pertinent.

The requirement of a bootstrap procedure to not only mimic the asymptotically dominating part of the (conditional) predictive distribution that captures the process uncertainty (i.e. asymptotic validity) but also the asymptotically negligible part capturing the uncertainty due to model parameter estimation is closely related to the concept coined asymptotic pertinence in Pan and Politis (2016a,b) for prediction in a time series context. The concept is also discussed in Beutner et al. (2021) from a different perspective. They argue that asymptotic validity of prediction intervals is a fundamental property, but it does not tell the whole story. Prediction intervals are particularly useful if they can also capture the uncertainty involved in model estimation, since it is beneficial in finite samples, although the estimation uncertainty is asymptotically negligible.

Since the asymptotic (conditional) limit distribution of the reserve is unknown so far, Chapter 2 (Paper 1) defines a fully stochastic framework based on Mack's model and derives the asymptotic results, capturing the uncertainty of the process and the estimation parts at first separately and then jointly.

Based on the results obtained in Chapter 2 (Paper 1), we show bootstrap consistency results in Chapter 3 (Paper 2). We show the proposed tailor-made bootstrap by England and Verrall (2002) based on Mack's model is only valid, since it can only capture the process uncertainty part under mild assumptions, but not the estimation uncertainty part. Hence, we propose an alternative Mack bootstrap, which is designed to capture also the estimation uncertainty part and is therefore called pertinent.

1.4 Extended Summary of the Articles

Extended Summary of Paper 1

In the first paper, we focus on deriving the asymptotic, unconditional and conditional distribution of the reserve risk based on Mack's model. Mack (1993) proposed a model to calculate the first two moments of the reserve. A normal approximation, together with the calculated moments, is often applied to conduct statistical inference and to estimate large quantiles. The literature lacks a rigorous justification for such a normal distribution. Paper 1 fills this gap and shows that, in general, the distribution of the reserve is non-Gaussian.

We enhance the assumptions from Mack's model to derive a general stochastic framework with rather mild assumptions.
We prove that the estimated parameters in Mack's model are consistent for an increasing number of accident years and a fixed number of development years. Since the claims $C_{i,j}$ for $i = 0, \ldots, I - j$, and fixed j are independent and identically distributed with finite first and second moment, the requirements of a Central Limit Theorem (CLT) are fulfilled. Using the delta method, we show properly inflated (smooth functions of) estimators of the development factors follow a normal distribution for a fixed number of development years and an increasing number of accident years. Imposing additional assumptions regarding the claims' higher (conditional) moments, we also prove a CLT for the estimators of the variance parameters for a fixed number of development years and an increasing number of accident years. We show that the asymptotic variance derived by the CLT of the variance parameters depends on the (conditional) third and fourth moments of the individual development factors. The CLT for the development factors does not depend on the (conditional) distribution of the development factors since the first two moments of the individual development factors are already determined by Mack's model.

Afterward, we define an appropriate asymptotic view of a loss triangle for prediction, where the observed loss triangle is growing by accident years (row-wise) instead of calendar years (diagonal-wise). Both approaches lead to loss triangles that are equal in distributions. Additionally, we set mild assumptions about the limit behavior of the development factors and the variance parameters for an increasing number of development years.

Using the defined stochastic model and the asymptotic view of a loss triangle, we define the so-called predictive root of the reserve. The predictive root of the reserve is defined as the difference between the stochastic reserve and its best estimate. This definition follows directly by the definition of the reserve risk that the best estimate of the reserve is not sufficient to pay for all outstanding claims.
The predictive root of the reserve can be split up into two additive parts covering the

process and parameter uncertainty. The process uncertainty covers the uncertainty involved in the development of the claims, whereas the estimation uncertainty covers the uncertainty involved in the estimation of the parameters for the development factors. We derive the (un-)conditional limit distribution for each part at first separately and then jointly.
The process uncertainty part dominates asymptotically and is, in general, non-Gaussian distributed unconditional and conditional on the whole observed loss triangle or the last observed diagonal of the loss triangle.
In contrast, the parameter uncertainty part is measurable with respect to the whole observed loss triangle. Properly inflated the estimation uncertainty part converges unconditionally to a non-Gaussian distribution and conditional on the last observed diagonal of the loss triangle to a Gaussian distribution using the derived asymptotic results for the estimators of the development factors.
Together, the predictive root of the reserve converges to the limit distribution of the process uncertainty conditional and unconditional, which is generally non-Gaussian. These findings cast the common practice of using a normal approximation together with the estimated moments by Mack's model for the distribution of the reserve into doubt.

Also, the findings are illustrated by a simulation study and show that the setup covers cases where the limiting distributions of the reserve risk might deviate substantially from a Gaussian distribution.

Extended Summary of Paper 2

The second paper focuses on the tailor-made Mack bootstrap algorithm proposed by England and Verrall (2006) based on Mack's model. The goal is to simulate the whole distribution of the reserve and to get estimates for large quantiles of the predictive root of the reserve to quantify the reserve risk. Although this bootstrap algorithm is widely used in applications to estimate the reserve risk, so far no theoretical bootstrap consistency results exist in the literature that justifies this approach. Paper 2 shows - for the first time - that the proposed Mack bootstrap is only valid, and therefore we suggest a new alternative Mack bootstrap algorithm.

In Paper 2, we use the stochastic and asymptotic framework from Paper 1.

The predictive root of the reserve is defined as the difference between the stochastic reserve and its best estimate. The definition of the predictive root of the reserve is motivated by the reserve risk such that the best estimate is not sufficient to cover all unknown and known claims in the future. Motivated by the definition of the reserve risk, England and Verrall (2006) keep the best estimate of the reserve fixed for their bootstrap algorithm. To mimic the stochastic reserve by the bootstrap, England and Verrall (2006) suggest a non-parametric re-sampling of residuals to derive bootstrap estimators for the development factors. Then a parametric bootstrap step is applied using the bootstrap estimators to simulate the lower loss triangle, i.e. the future claims, and to derive the stochastic reserve. Thus, the Mack bootstrap predictive root computes the difference between the stochastic bootstrap reserve constructed by the double bootstrap (non-parametric step with an additional parametric step) and the best estimate estimated on the observed data.

We discuss in detail the particularities and (asymptotic) properties of the Mack bootstrap predictive root and compare them with the (asymptotic) properties of the predictive root of the reserve derived in Paper 1. We split the predictive root of the reserve and its bootstrap version into two additive parts corresponding to process and estimation uncertainty. The conditional asymptotic distributions of the two parts of the predictive root of the reserve are used as benchmark distributions. Comparing the derived conditional asymptotic limit distributions from Paper 1 with the bootstrap distributions - separately for each part - allows a rigorous investigation of the validity of the Mack bootstrap.
We prove that the distribution conditional on the observed claims data of the 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 proposal. Otherwise, this is generally not the case.
In contrast, the corresponding (conditional) distribution of the estimation uncertainty part is generally not correctly captured by the Mack bootstrap. Since the asymptotic variance

of the (conditional) bootstrap distribution of the estimation uncertainty is bigger compared to the asymptotic variance of the (conditional) distribution of the estimation uncertainty. Hence, the limit distribution of the estimation uncertainty part is not correctly mimicked by Mack's bootstrap.
Together, the (conditional) bootstrap distribution of the predictive root of the reserve will be correctly mimicked if the parametric family of distributions of the individual development factors is correctly specified in Mack's bootstrap since the process uncertainty part is the asymptotically dominating process. We call the bootstrap predictive root only valid and not pertinent since the parameter uncertainty will not be mimicked correctly. Especially for finite samples it is beneficial to capture also the parameter uncertainty correctly.

To tackle this, we propose a more naturally alternative Mack-type bootstrap that uses a different centering and is designed to capture also the distribution of the estimation uncertainty part correctly.
We propose to generate recursively backward starting at the diagonal $\mathcal{Q}^{+}_{I,n} = \mathcal{Q}_{I,n}$ new upper loss triangles and to calculate the bootstrap estimators for the development factors based on these newly generated triangles. We show that the variance of the backward bootstrap estimation part is asymptotically the same as the variance of the estimation part conditional on the last observed diagonal.
We discuss the properties of this alternative Mack bootstrap extensively and compare them with the properties of the original Mack bootstrap. The alternative bootstrap is called pertinent since it mimics the process and parameter uncertainty part correctly if the parametric family of distributions of the individual development factors is correctly specified. Otherwise, this will be, in general, not the case.

By simulations, we illustrate our findings and show that the newly proposed Mack-type bootstrap performs superior to the original Mack bootstrap in a finite sample. Also, an intermediate Mack bootstrap provides evidence that the backward resampling appears to be critical and responsible for this improvement

2 Chapter

Asymptotic Theory for Mack's Model

This chapter was published in Insurance: Mathematics and Economics, 107, Julia Steinmetz and Carsten Jentsch, Asymptotic Theory for Mack's Model, 223-268, Copyright Elsevier (2022).

Abstract
The distribution-free chain ladder reserving model by Mack (1993) belongs to the most popular approaches in non-life insurance mathematics. It was originally proposed to determine the first two moments of the reserve distribution. Together with a normal approximation, it is commonly applied to conduct statistical inference including the estimation of large quantiles of the reserve and determination of the reserve risk. However, for Mack's model, the literature lacks a rigorous justification of such a normal approximation for the reserve.
In this paper, we propose a general stochastic framework which allows to derive asymptotic theory for Mack's model. For increasing number of accident years, we establish central limit theorems for the parameter estimators in Mack's model. In particular, these results enable us to derive also unconditional and conditional limiting distributions for the reserve. For this purpose, the reserve risk is split into two random parts that carry the process uncertainty and the estimation uncertainty, respectively. Unconditionally, but also when conditioning on either the whole observed loss triangle or on its diagonal, we show that the limiting distribution of the first part that corresponds to the process uncertainty will be usually non-Gaussian. When properly inflated, the second part corresponding to the estimation uncertainty is measurable with respect to the loss triangle and, unconditionally, turns out to be asymptotically non-Gaussian as well. By contrast, when conditioning only on the diagonal, this results in a Gaussian limit. As the process uncertainty part dominates asymptotically, this leads overall to a non-Gaussian limiting distribution for the reserve in both cases.
These findings cast the common practice to use a normal approximation for the reserve in Mack's model into doubt. We illustrate our findings by simulations and show that our setup covers cases, where the limiting distributions of the reserve risk might deviate substantially from a Gaussian distribution.

Keywords: Chain ladder model; claims reserving; conditional limiting theory; limiting distribution; Mack's model; reserve risk

JEL: C13, C18, C53, G22

2.1 Introduction

Insurance companies are committed to set capital reserves to be able to meet their future liabilities related to their earlier and current insurance contracts. The claim settlement of a non-life insurance contract can take several years until a claim is finally closed. The typical process of loss adjustment for a non-life insurance company is described in Figure 2. The fundamental issue is that the actual sizes of the claims are unknown at the time the reserves have to be set since these outstanding claims have either not been reported yet or they have been reported, but not settled yet. This process of prediction of outstanding claims is called *reserving*. For non-life insurance companies reserves are often the major part of the liability side of the balance sheet. Hence, an accurate estimation 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 determine recursively the best estimate of the future claims by using a set of development factors. First of all, the CLM does not impose any distributional assumptions on the (future) claims. In this respect, the most popular extension of the CLM is the recursive model proposed by Mack (1993), who equipped the CLM with formulas to calculate the standard deviation of the reserve.

There exist various other models to estimate future claims (see e.g. the monographs by Wüthrich and Merz (2008) and Hindley (2017)). For example, frameworks based on general linear models (GLMs) by Renshaw and Verrall (1998) make use e.g. of over-dispersed Poisson and log-normal distributions. These models often have age-cohort assumptions. For incremental claims data, Harnau and Nielsen (2018) and Kuang and Nielsen (2020) developed an asymptotic theory for an over-dispersed Poisson and a log-normal framework, respectively. They assume that the incremental claims have infinitely divisible distributions and derive central limit theorems (CLTs) for the estimators and proposed a $t-$distribution for the centered and standardized reserve.

However, Mack's model is still one of the most popular and frequently used approaches in practice. It is often favoured by actuaries due to its simple and straightforward application, see e.g. Gisler (2019). Mack's distribution-free algorithm allows for best estimation of outstanding liabilities, but the knowledge of solely the first two moments is usually not sufficient to calculate the reserve risk and to make conclusions about the solvency of the insurance company. The *reserve risk* is defined as the risk that the economic value of the reserve (best estimate) does not suffice to pay for all outstanding claims. To get such insights, the knowledge of high quantiles or the whole distribution of the reserve is usually inevitable. Hence, actuaries are more interested in the construction of a well-founded range of the reserve rather than the best estimate alone. Together with a normal approximation,

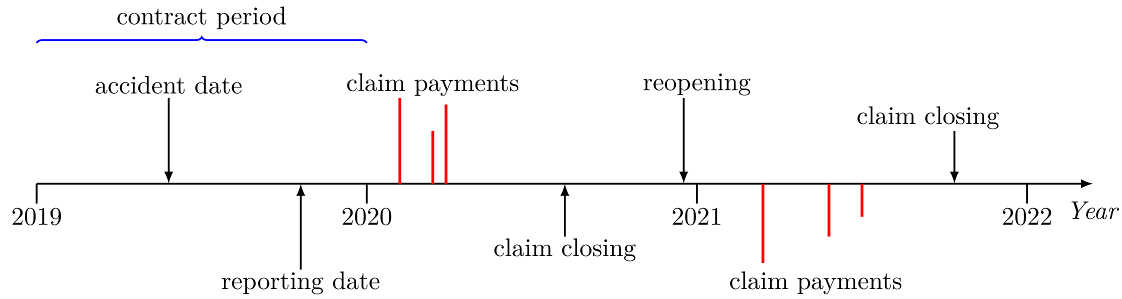

Figure 2: Typical process of loss adjustment for a non-life insurance.

Mack's model is commonly applied to conduct statistical inference, which includes the estimation of high quantiles of the reserve and the determination of the reserve risk.

In view of its popularity, it is surprising that there is no rigorous justification in form of meaningful asymptotic distribution theory for Mack's model and the justification of a normal approximation in this context. Nevertheless, a deeper understanding is required to understand in which situations Mack's model can be consistently applied in real data applications.

In this paper, we establish a rigorous model framework that allows to derive asymptotic theory for the parameter estimators in Mack's model for increasing number of accident years. The assumptions imposed in our model framework are mild such that real data sets can easily meet these assumptions. In particular, we do not assume any parametric distribution class that generates the claims. Our results provide further insights about the unconditional and conditional limiting behaviour of the reserve risk and allow to better assess the uncertainty of the best estimate of the reserve. The stochastic properties of Mack's model are usually not completely determined such that there exist various stochastic models which agree with the imposed assumptions (see e.g. Mack (1994), Renshaw and Verrall (1998)). For this purpose, we propose a fully-described stochastic framework that agrees with Mack's model setup and allows for the derivation of asymptotic normality of parameter estimators. The established theory goes beyond estimation consistency results that were achieved by Pešta and Hudecová (2012) for the development factors. More precisely, we will have a closer look at the reserve centered around its best estimate. Throughout the paper, borrowing the notation from time series analysis, we call this difference the *predictive root of the reserve*. We split this predictive root into two (random) parts which carry the process uncertainty and the estimation uncertainty, respectively. This approach coincides with Mack (1993), who already distinguishes between *process variance* and *estimation variance*. For these two parts, we aim to derive (conditional) asymptotic distributions for increasing number of accident years. With this goal in mind, we first derive CLTs for the parameter estimators in Mack's model. In particular, the limiting distributions of the estimated development factors enable us to derive the limiting distributions of the second part of the reserve risk, which corresponds to the estimation uncertainty. Unconditionally, but also when conditioning on the whole observed loss triangle or on its diagonal, we show that the limiting distribution of the first part that corresponds to the process uncertainty will be usually non-Gaussian. When properly inflated, the second part corresponding to the estimation uncertainty is measurable with respect to the loss triangle and, unconditionally, turns out to asymptotically non-Gaussian as well. By contrast, when conditioning only on the diagonal, this results in a Gaussian limit. As the process uncertainty part dominates asymptotically, this leads overall to a non-Gaussian limiting distribution for the reserve in both cases. These findings cast the common practice to use a normal approximation for the reserve in Mack's model into doubt.

The paper is organized as follows. In Section 2.2, we consider the CLM setup and introduce Mack's model and the required notation in Section 2.2.1. The complementing stochastic framework suitable to derive asymptotic theory in Mack's model and some implications are discussed in Section 2.2.2. In Section 2.3, we establish several versions of CLTs for the parameter estimators in Mack's model. These results are employed in Section 2.4 to derive unconditional and conditional limiting distributions separately for both random parts of the predictive root of the reserve, which correspond to process and estimation uncertainty, respectively, as well as jointly. In particular, in contrast to common belief (and common practice), the limiting distribution of the reserve is in general non-Gaussian such that asymptotic standard inference based on a normal approximation is not justified. Section 2.5 contains a simulation study to illustrate our findings. In particular, we demonstrate that the distribution of the reserve risk might deviate substantially from a Gaussian distribution. Instead, generally, it will depend on the specific distribution of the individual development factors also in the limit. Section 2.6 concludes. All proofs and additional simulation results are deferred to an appendix.

Accident Year i	Development Year j					
	0	1	2	$\cdots$	$I-1$	I
0	$C_{0,0}$	$C_{0,1}$				$C_{0,I}$
1	$C_{1,0}$	$\cdots$			$C_{1,I-1}$	
	$\cdots$		$\cdots$			
$\cdots$	$\cdots$					
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$				
I	$C_{I,0}$					

Table 2: Loss triangle $\mathcal{D}_I$ with accident years (vertical axis), development years (horizontal axis) and diagonal $\mathcal{Q}_I$ (orange).

2.2 The Chain Ladder Model

Accurate prediction of outstanding claims is one of the most crucial tasks for insurance companies. For this purpose, insurers summarize all observed claims of a business line in a loss triangle organized as shown in Table 2. Its entries, the cumulative amount of claims $C_{i,j}$, are sorted by their accident year i (vertical axis) and their development year j (horizontal axis) for $i, j = 0, \dots, I$ with $i + j \leq I$. Hence, the (observed) loss triangle contains the cumulative claims $C_{i,j}$ that have already been observed up to year I. It constitutes the available data basis and is denoted by

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

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

$$\mathcal{Q}_k = \{C_{k-i,i} | i = 0, \dots, k\}\,. \tag{2.2}$$

In this setup, I is the current calendar year corresponding to the most recent accident year and development period such that the diagonal $\mathcal{Q}_I$ (see Table 2) summarizes the latest cumulative claim amounts collected in year I.

With the goal to theoretically analyze the prediction of outstanding claims in the chain ladder model, it is useful to augment the (observed) *upper* loss triangle by an unobserved *lower* triangle $\{C_{i,j} | i = 0, \dots, I,\ j = 0, \dots, I,\ i + j > I\}$ that contains all future claims that have not been observed (yet) up to time I. The resulting cumulative claim *matrix* is denoted by $\mathcal{C}_I = (C_{i,j})_{i,j=0,\dots,I}$. For each accident year i, the main interest lies in the *reserves* (loss liabilities) $R_{i,I}$ at *terminal* time I, which is computed by taking the difference of the so-called ultimate claim $C_{i,I}$, which is not observed (for $i > 0$) at time I, minus the last claim $C_{i,I-i}$ observed at time I. Precisely, we define

$$R_{i,I} = C_{i,I} - C_{i,I-i}, \quad i = 0, \ldots, I, \tag{2.3}$$

and denote the aggregated total amount of the reserve by

$$R_I = \sum_{i=0}^{I} R_{i,I}, \tag{2.4}$$

noting that $R_{0,I} = C_{0,I} - C_{0,I} = 0$ by construction. Hence, in calendar year I, to get an estimate of the reserve $R_{i,I}$, we have to predict the unobserved ultimate claim $C_{i,I}$ based on the observed upper loss triangle $\mathcal{D}_I$. Starting from the latest observed claim $C_{i,I-i}$ for some accident year i, this is done by predicting recursively all future, yet (at time I) unobserved claims $\{C_{i,j}|j = I-i+1, \ldots, I\}$. By doing this for all accident years $i = 0, \ldots, I$, this allows us also to predict R_I by aggregating all predictions for $R_{i,I}$.

2.2.1 Distribution-free chain ladder reserving

A widely used model in practice to determine the mean and the variance of the reserve for each accident year i is the distribution-free Chain Ladder Model (CLM) proposed by Mack (1993). Often denoted as *Mack's model*, it relies on three fundamental (stochastic) model assumptions summarized as follows.

Assumption 2.1 (Mack's model) *Let $C_{i,j}$, $i, j = 0, \ldots, I$ denote random variables on some probability space $(\Omega, \mathcal{A}, P)$ and suppose the following holds:*

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

$$E(C_{i,j+1}|C_{i,j}) = f_j C_{i,j} \tag{2.5}$$

for all $i = 0, \ldots, I$ and $j = 0, \ldots, I-1$.

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

$$Var(C_{i,j+1}|C_{i,j}) = \sigma_j^2 C_{i,j} \tag{2.6}$$

for all $i = 0, \ldots, I$ and $j = 0, \ldots, I-1$.

(iii) The cumulative payments are stochastically independent over the accident years $i = 0, \ldots, I$, that is, the cumulative claim matrix $\mathcal{C}_I = (C_{i,j})_{i,j=0,\ldots,I}$ consists of independent rows $C_{i,\bullet} = (C_{i,0}, \ldots, C_{i,I})$, $i = 0, \ldots, I$.

Alternatively, as in the original paper by Mack (1993), the conditional mean and variance in (2.5) and (2.6) are phrased to be conditional on *all* previously observed cumulative claim amounts $C_{i,0}, \ldots, C_{i,j}$. Either way, the stochastic model assumptions given in Assumption 2.1 allow for simple and easy to interpret formulas for the conditional mean and conditional variance of an ultimate claim $C_{i,I}$ given the last observed claim $C_{i,I-i}$. Precisely, we get

$$E(C_{i,I}|\mathcal{Q}_I) = E(C_{i,I}|C_{i,I-i}) = C_{i,I-i} \prod_{j=I-i}^{I-1} f_j, \tag{2.7}$$

$$Var(C_{i,I}|\mathcal{Q}_I) = Var(C_{i,I}|C_{i,I-i}) = C_{i,I-i} \sum_{j=I-i}^{I-1} f_{I-i} \ldots f_{j-1} \sigma_j^2 f_{j+1}^2 \ldots f_{I-1}^2 \tag{2.8}$$

for all $i = 0, \ldots, I$, respectively. As all development factors f_j and variance parameters σ_j^2 in (2.7) and (2.8) are generally unknown, they have to be estimated from the data available in $\mathcal{D}_I$. By exploiting the CLM fulfilling Assumption 2.1, the development factors $f_0, \ldots, f_{I-1}$ can be estimated by $\widehat{f}_0, \ldots, \widehat{f}_{I-1}$, where

$$\widehat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \tag{2.9}$$

for $j = 0, \ldots, I-1$. As shown by Mack (1993), $\widehat{f}_j$ is an unbiased estimator for f_j, i.e. $E(\widehat{f}_j) = f_j$, $j = 0, \ldots, I-1$, and $\widehat{f}_0, \ldots, \widehat{f}_{I-1}$ are pairwise uncorrelated, i.e. $Cov(\widehat{f}_j, \widehat{f}_k) = 0$ for all $j \neq k$.

By plugging-in the $\widehat{f}_j$'s in (2.7), conditional on $C_{i,I-i}$, the best estimate $\widehat{C}_{i,I}$ (point predictor) of the ultimate claim $C_{i,I}$ becomes

$$\widehat{C}_{i,I} = C_{i,I-i} \prod_{j=I-i}^{I-1} \widehat{f}_j \tag{2.10}$$

for $i = 0, \ldots, I$. Consequently, given $C_{i,I-i}$, the best estimate $\widehat{R}_{i,I}$ of the reserve $R_{i,I}$ of accident year i defined in (2.3) is given by

$$\widehat{R}_{i,I} = \widehat{C}_{i,I} - C_{i,I-i} = C_{i,I-i} \left(\prod_{j=I-i}^{I-1} \widehat{f}_j - 1 \right), \quad i = 0, \ldots, I, \tag{2.11}$$

and the best estimate $\widehat{R}_I$ of the total reserve R_I defined in (2.4) computes to

$$\widehat{R}_I = \sum_{i=0}^{I} \widehat{R}_{i,I} = \sum_{i=0}^{I} C_{i,I-i} \left(\prod_{j=I-i}^{I-1} \widehat{f}_j - 1 \right) \tag{2.12}$$

noting that $\widehat{R}_{0,I} = 0$ due to $\prod_{j=I}^{I-1} \widehat{f}_j = 1$. Furthermore, Mack (1993) proposed to estimate the variance parameters $\sigma_0^2, \ldots, \sigma_{I-2}^2$ by $\widehat{\sigma}_0^2, \ldots, \widehat{\sigma}_{I-2}^2$, where

$$\widehat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j \right)^2, \quad j = 0, \ldots, I-2, \tag{2.13}$$

which is an unbiased estimator for σ_j^2, i.e. $E(\widehat{\sigma}_j^2) = \sigma_j^2$. Note that (2.13) does not cover the estimation of σ_{I-1}^2. However, if $\frac{1}{I-j-1}$ is replaced by $\frac{1}{I-j}$, setting $j = I-1$ in (2.13) naturally leads to $\widehat{\sigma}_{I-1}^2 = 0$, because $\widehat{f}_{I-1}$ is estimated by only one observed pair of claims $(C_{0,I-1}, C_{0,I})$, i.e. $\widehat{f}_{I-1} = \frac{C_{0,I}}{C_{0,I-1}}$. Hence, it is reasonable to set $\widehat{\sigma}_{I-1}^2 = 0$. Alternatively, Mack (1993) proposed to extrapolate $\widehat{\sigma}_{I-1}^2$ by using a log-linear regression or by setting $\widehat{\sigma}_{I-1}^2 = min(\widehat{\sigma}_{I-2}^4/\widehat{\sigma}_{I-3}^2, min(\widehat{\sigma}_{I-3}^2, \widehat{\sigma}_{I-2}^2))$.

Particular interest is in the difference of the stochastic (unobserved) reserve R_I and its best estimate $\widehat{R}_I$. Therefore, we define the mean squared error of prediction (MSEP) of $\widehat{R}_I$ given $\mathcal{D}_I$ by

$$MSEP\left(\widehat{R}_I|\mathcal{D}_I\right) = E\left((R_I - \widehat{R}_I)^2|\mathcal{D}_I\right). \tag{2.14}$$

Mack (1993) derived a formula for the MSEP, that is

$$MSEP\left(\widehat{R}_I|\mathcal{D}_I\right) = \sum_{i=0}^{I} \left(\widehat{C}_{i,I}^2 \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\widehat{C}_{i,j}} + \widehat{C}_{i,I}^2 \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}} \right) \tag{2.15}$$

$$+ 2 \sum_{\substack{i,l=0 \\ i<l}}^{I} \left(\widehat{C}_{i,I} \widehat{C}_{l,I} \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}} \right), \tag{2.16}$$

where the summands of the first sum on the right-hand side above in (2.15) consist of two terms corresponding to the *process variance* and *estimation variance* (of parameter estimates) of $R_{i,I}$, respectively. The second expression in (2.16) reflects the linear dependence between $\widehat{R}_{i,I}$ and $\widehat{R}_{l,I}$, $i \neq l$ and contains their covariances. Alternatively, the MSEP of $\widehat{R}_I$ can be rewritten as

$$MSEP\left(\widehat{R}_I|\mathcal{D}_I\right) = \sum_{i=0}^{I} \left(C_{i,I-i} \sum_{j=I-i}^{I-1} \widehat{\sigma}_j^2 \prod_{k=I-i}^{j-1} \widehat{f}_k \prod_{l=j+1}^{I-1} \widehat{f}_l^2 + C_{i,I-i}^2 \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}} \prod_{\substack{l=I-i \\ l\neq j}}^{I-1} \widehat{f}_l^2 \right)$$

$$+ 2 \sum_{\substack{i,l=0 \\ i<l}}^{I} \left(C_{i,I-i} C_{l,I-l} \sum_{j=I-i}^{I-1} \frac{\widehat{\sigma}_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}} \prod_{n=I-l}^{I-i-1} \widehat{f}_n \prod_{\substack{m=I-i \\ m\neq j}}^{I-1} \widehat{f}_m^2 \right).$$

As presented above, the model setup proposed by Mack (1993) allows to estimate the mean and the variance of the reserve R_I conditional on $\mathcal{D}_I$ as well as its MSEP. Buchwalder et al. (2006) derive another formula for the MSEP of the reserve, where they model $C_{i,j}$ using an $AR(1)$ time series. There exists a broad discussion which formula for the MSEP should be preferred (see e.g. Mack et al. (2006) and Gisler (2021)). Recently, Siegenthaler (2021) points out that the MSEP formulas by Mack (1993) and by Buchwalder et al. (2006) lead to very similar results whose differences are not material for applications in actuarial practice. In a GLM framework, Lindholm et al. (2020) derive the MSEP of the reserve and show that the MSEP coincides with Mack's MSEP in a special case. However, the knowledge of the first two moments of the reserve will be not sufficient to determine its whole distribution. But in practice, it is important to be able to estimate also high quantiles of the (conditional) distribution of $\widehat{R}_I$ to approximate the *reserve risk* as e.g. the value-at-risk. For this purpose, a common approach is to assume a certain parametric family of distributions for either the reserve of a single accident year $\widehat{R}_i$ or of the total reserve $\widehat{R}_I$, and to estimate their distributions by estimating their parameters. In this regard, it is more common to model the distribution of the *total* reserve to take diversification effects between the single accidents years into account. However, a parametric assumption imposed on the reserve might be too restrictive and may result in misleading conclusions drawn from the prediction $\widehat{R}_I$ of R_I. Hence, a non-parametric analysis to derive the limiting distribution of the reserve without restricting considerations to any parametric assumptions might be more beneficial in practice. However, in view of Mack's distribution-free reserving framework discussed in this section, this is not yet possible. In the following section, to complement this framework, we will state additional conditions on the random mechanism that generates the cumulative claim matrix $\mathcal{C}_I = (C_{i,j})_{i,j=0,\ldots,I}$ to be able to derive asymptotic theory for parameter estimators in Mack's model in Sections 2.3 and for the predictive root of the reserve in Section 2.4.

2.2.2 A fully-described stochastic framework of Mack's model

As discussed in Section 2.2.1, the classical set of assumptions for Mack's model summarized in Assumption 2.1 alone is not yet sufficient to estimate the whole distribution of the reserve. In this section, we will complement Assumption 2.1 with some mild conditions on the stochastic properties of the cumulative claims. By doing this, we establish a fully-described stochastic framework that allows for the derivation of a general asymptotic theory for increasing number of accident years. In the subsequent Section 2.3, we will establish CLTs for (smooth functions of) parameter estimators $\widehat{f}_j$ and $\widehat{\sigma}_j^2$. These results will enable us also to investigate the unconditional limiting distributions of the reserve $\widehat{R}_I$, which will be addressed in Section 2.4. Conditional versions of the CLTs for the parameter estimators, which allow for the derivation also of the conditional limiting distributions of the reserve, are given in the appendix.

Precisely, to establish a framework sufficient to be able to derive asymptotic theory, we introduce three additional assumptions on the stochastic mechanism that generates the cumulative claim matrix $\mathcal{C}_I = (C_{i,j})_{i,j=0,\ldots,I}$. The first one addresses the initial claims, i.e. the first column $C_{\bullet,0} = (C_{0,0}, \ldots, C_{I,0})'$ of $\mathcal{C}_I$.

Assumption 2.2 (Initial claims) *Suppose that the initial claims $C_{\bullet,0} = (C_{0,0}, \ldots, C_{I,0})'$ are 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 $\tau_0^2 := Var(C_{i,0}) \in (0, \infty)$.*

Note that the independence between the initial claims is a common assumption that is in particular a direct consequence of Assumption 2.1 (iii). In addition, Assumption 2.2 also imposes an identical distribution for the initial claims. The condition on the support $[1, \infty)$ of $C_{i,0}$ can be relaxed and a condition that $C_{i,0}$ is bounded away from zero will be also sufficient as well to derive asymptotic theory. However, in practice, it will be not restrictive to assume a support $[1, \infty)$.

In view of the multiplicative structure of $E(C_{i,I}|C_{i,I-i})$ in (2.7), let the (random) cumulative claims $C_{i,j+1}$, $i = 0, \ldots, I$ and $j = 0, \ldots, I-1$, be recursively defined by

$$C_{i,j+1} = C_{i,j} F_{i,j} = C_{i,0} \prod_{k=0}^{j} F_{i,k}, \tag{2.17}$$

where the *individual development factors* $F_{i,j}$, which satisfy $F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$ by construction, are assumed to fulfill the following condition.

Assumption 2.3 (Conditional distribution of individual development factors) *Let the individual development factors $F_{i,j}$, $i = 0, \ldots, I$, $j = 0, \ldots, I-1$ 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 conditional mean and conditional variance*

$$E(F_{i,j}|C_{i,j}) = f_j \quad \textit{and} \quad Var(F_{i,j}|C_{i,j}) = \frac{\sigma_j^2}{C_{i,j}}. \tag{2.18}$$

It is important to note that both Assumptions 2.2 and 2.3 together imply Mack's original setup of Assumption 2.1. In Sections 2.3 and 2.4, we will see that the stochastic framework described by Assumptions 2.2 and 2.3 is appropriate to derive asymptotic theory in Mack's model framework. Also note that the stochastic mechanism described by (2.17) and Assumption 2.3 is assumed for the whole cumulative claim matrix $\mathcal{C}_I$. However, recall that only those $C_{i,j}$ in $\mathcal{C}_I$ are observed up to year I that are contained in the upper loss triangle $\mathcal{D}_I$ defined in (2.1). Hence, by using the multiplicative relationship in the first identity of (2.17), knowledge of $\mathcal{D}_I$ implies that we have also perfect knowledge of the individual development factors $F_{i,j}$, $i = 0, \ldots, I-1$, $j = 0, \ldots, I-i-1$.

In addition to the formulas for the conditional mean and variance of $C_{i,j+1}$ given $C_{i,j}$ obtained from Assumption 2.1 (or, equivalently, from (2.17) and Assumption 2.3), i.e.

$$E(C_{i,j+1}|C_{i,j}) = f_j C_{i,j} \quad \text{and} \quad Var(C_{i,j+1}|C_{i,j}) = \sigma_j^2 C_{i,j} \tag{2.19}$$

for all $i = 0, \ldots, I$ and $j = 0, \ldots, I-1$, the following lemma provides some useful formulas also for the unconditional mean, variance and covariances of the cumulative claims $C_{i,j}$, which turn out to be useful in the sequel.

Lemma 2.4 (Mean, variance and covariances of $C_{i,j}$) *Suppose Assumption 2.2 and 2.3 hold. Then, for each $i, i_1, i_2, j, j_1, j_2 = 0, \ldots, I$, we have*

$$E(C_{i,j}) = \mu_0 \prod_{k=0}^{j-1} f_k =: \mu_j, \tag{2.20}$$

$$Var(C_{i,j}) = \tau_0^2 \prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1} \left(\prod_{m=0}^{l-1} f_m \right) \sigma_l^2 \left(\prod_{n=l+1}^{j-1} f_n^2 \right) =: \tau_j^2, \tag{2.21}$$

$$Cov(C_{i,j_1}, C_{i,j_2}) = \left(\prod_{k=\min(j_1,j_2)}^{\max(j_1,j_2)-1} f_k \right) \tau^2_{\min(j_1,j_2)} \tag{2.22}$$

and $Cov(C_{i_1,j_1}, C_{i_2,j_2}) = 0$, whenever $i_1 \neq i_2$.

So far, based on the stochastic framework for Mack's model established by Assumptions 2.2 and 2.3, we have summarized some results for *fixed* number of development years. In the next section, we make use of this stochastic framework and let the number of accident years I go to infinity to establish asymptotic theory in form of CLTs for (smooth functions of) *finitely* many parameter estimators $\widehat{f}_j$ and $\widehat{\sigma}_j^2$ in Mack's model. In Section 2.4, we make use of a slightly adjusted asymptotic notion suitable particularly for the purpose of *prediction*, which is equivalent in distribution. Hence, these CLTs as well as their conditional versions in the appendix, will be useful to derive closed-form expressions to describe the limiting distributions for the best estimate of the reserve $\widehat{R}_I$.

2.3 Asymptotic Theory for Mack's model: Parameter Estimation

In this section, for asymptotic considerations, we wish to increase the number of accident years I and let $I \to \infty$. Recall that for any fixed I, we are dealing with an $(I+1) \times (I+1)$

cumulative claim matrix $\mathcal{C}_I$, where we observe the upper loss triangle $\mathcal{D}_I$. Hence, letting $I \to \infty$ means, that the cumulative claim matrix $\mathcal{C}_I$ grows in *both* dimensions. However, in this section, we are dealing only with the estimation of a fixed *finite* number of development factors f_j and variance parameters σ_j^2. Hence, $I \to \infty$ describes foremost an increasing number of rows (accident years) in the loss triangle throughout this section.

2.3.1 Asymptotic normality for $\widehat{f}_j$

To derive asymptotic theory for estimated development factors $\widehat{f}_j$, recall that $\widehat{f}_j$ is defined as the ratio of two sums over (neighbouring) columns $C_{\bullet,j+1}$ and $C_{\bullet,j}$ of equal length $I-j$ as defined in (2.9). Further note that for each fixed j, based on Assumptions 2.2 and 2.3, the cumulative claims in one column $C_{\bullet,j}$, i.e. $C_{i,j}$, $i = 0, \ldots, I$, are i.i.d. with mean μ_j and variance τ_j^2 as defined in Lemma 2.4. Moreover, for any fixed K, the vectors $(C_{i,0}, \ldots, C_{i,K+1})'$, $i = 0, \ldots, I$ are i.i.d. with mean vector $(\mu_0, \ldots, \mu_{K+1})'$ and variance-covariance matrix

$$Cov((C_{i,0}, \ldots, C_{i,K+1})') = \begin{pmatrix} \left(\prod_{k=\min(j_1,j_2)}^{\max(j_1,j_2)-1} f_k \right) \tau^2_{\min(j_1,j_2)} \\ j_1, j_2 = 0, \ldots, K+1 \end{pmatrix}. \tag{2.23}$$

This i.i.d.-ness meets the typical requirements of a law of large numbers (LLN) and of a CLT for averages of the form $\frac{1}{I-j} \sum_{i=0}^{I-j-1} C_{i,j}$. In particular, letting $I \to \infty$, Assumption 2.2 implies $\frac{1}{I} \sum_{i=0}^{I-1} C_{i,0} \xrightarrow{p} \mu_0$ and, consequently, as $\mu_0 \in [1, \infty)$, also $\sum_{i=0}^{I} C_{i,0} \xrightarrow{p} \infty$. As shown by Pešta and Hudecová (2012), the latter condition is sufficient for estimation consistency $\widehat{f}_j \xrightarrow{p} f_j$ for any fixed j, where "$\xrightarrow{p}$" denotes convergence in probability.

Altogether, the stochastic framework introduced above allows us to prove a CLT for the estimators of the development factors $\widehat{f}_j$ for any fixed j or finitely many $j \in \{0, 1, \ldots, K\}$ for all fixed $K \in \mathbb{N}_0$.

Theorem 2.5 (Asymptotic normality of $\widehat{f}_j$) *Suppose Assumptions 2.2 and 2.3 are satisfied. Then, as $I \to \infty$, the following holds:*

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

$$\sqrt{I-j}\left(\widehat{f}_j - f_j\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma_j^2}{\mu_j}\right),$$

where "$\xrightarrow{d}$" denotes convergence in distribution.

(ii) For each fixed $K \in \mathbb{N}_0$, let $\widehat{\underline{f}}_K = (\widehat{f}_0, \widehat{f}_1, \ldots, \widehat{f}_K)'$ be the $(K+1)$-dimensional estimator of $\underline{f}_K = (f_0, f_1, \ldots, f_K)'$. Then, we have

$$J^{1/2}\left(\widehat{\underline{f}}_K - \underline{f}_K\right) \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\underline{f}}\right),$$

where $J^{1/2} = diag\left(\sqrt{I-j}, j = 0, \ldots, K\right)$ is a diagonal $(K+1)\times(K+1)$ matrix of inflation factors and $\mathbf{\Sigma}_{K,\underline{f}} = J_g(\underline{\mu}_K)\mathbf{\Sigma}_{K,\underline{C}}J_g(\underline{\mu}_K)' = diag\left(\frac{\sigma_0^2}{\mu_0}, \frac{\sigma_1^2}{\mu_1}, \ldots, \frac{\sigma_K^2}{\mu_K}\right)$ is a diagonal $(K+1)\times(K+1)$ covariance matrix. $\mathbf{\Sigma}_{K,\underline{C}}$ and $J_g(\underline{\mu}_K)$ are defined in (2.68) and (2.72).

Recall that for each fixed $K \in \mathbb{N}_0$, the estimators $\widehat{f}_0, \ldots, \widehat{f}_K$ are known to be pairwise uncorrelated already for finite I just based on Assumption 2.1. The asymptotic normality achieved in Theorem 2.5 complements this result and due to the diagonal limiting covariance matrix $\mathbf{\Sigma}_{K,\underline{f}}$, the estimators $\widehat{f}_0, \ldots, \widehat{f}_K$ are asymptotically also independent. In particular, the CLTs in Theorem 2.5 imply $\sqrt{I}$-consistency of $\widehat{f}_j$ for f_j for each fixed j.

Using the results in Theorem 2.5, by a direct application of the delta method, we can easily derive CLTs also for (sufficiently smooth) functions of $\widehat{\underline{f}}_K = (\widehat{f}_0, \widehat{f}_1, \ldots, \widehat{f}_K)'$. Specifically, as stated in the following corollary, joint asymptotic normality results hold for products of $\widehat{f}_j$'s. Picking-up the representation of $\widehat{C}_{i,I}$ in (2.10), these results will turn out to be crucial to derive asymptotic theory also for the best estimate of the reserve $\widehat{R}_{I,n}$ in Section 2.4.

Corollary 2.6 (Asymptotic normality for products of $\widehat{f}_j$'s) *Suppose the assumptions of Theorem 2.5 hold. Then, as $I \to \infty$, the following holds:*

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

$$\sqrt{I}\left(\prod_{j=i}^{K}\widehat{f}_j - \prod_{j=i}^{K} f_j\right) \xrightarrow{d} \mathcal{N}\left(0, \sum_{j=i}^{K}\frac{\sigma_j^2}{\mu_j}\prod_{l=i,l\neq j}^{K} f_l^2\right).$$

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

$$\sqrt{I}\begin{pmatrix}\prod_{j=i}^{K}\widehat{f}_j - \prod_{j=i}^{K} f_j \\ i = 0, \ldots, K\end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\prod f_j}\right),$$

where $\mathbf{\Sigma}_{K,\prod f_j} = J_h(\underline{f}_K)\mathbf{\Sigma}_{K,\underline{f}}J_h(\underline{f}_K)' = (\mathbf{\Sigma}_{K,\prod f_j}(i_1,i_2))_{i_1,i_2=0,\ldots,K}$ *is a* $(K+1)\times(K+1)$ *covariance matrix with entries*

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

for $i_1,i_2=0,\ldots,K$. $\mathbf{\Sigma}_{K,\underline{f}}$ *and* $J_h(\underline{f}_K)$ *are defined in Theorem 2.5(ii) and* (2.74).

In the appendix, corresponding to the results in Theorem 2.5 and Corollary 2.6, we also provide asymptotic theory conditional on the diagonal $\mathcal{Q}_{I,\infty}$ to be defined in (2.34) under the asymptotic framework of Section 2.4. In this regard, Theorem 2.24 and Corollary 2.25 allow a *conditional* asymptotic treatment of the estimation uncertainty part of the predictive root of the reserve.

2.3.2 Asymptotic normality for $\widehat{\sigma}_j^2$

Similar to the limiting results in Section 2.3.1, we want to prove a CLT also for $\widehat{\sigma}_j^2$. For this purpose, in addition to Assumptions 2.2 and 2.3, we have to impose also assumptions on higher-order (conditional) moments of the individual development factors to establish asymptotic normality.

Assumption 2.7 (Higher-order conditional moments of individual development factors) *Suppose that, conditional on* C_{ij}, *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 for* $j=0,\ldots,I-1$ *such that both*

$$\kappa_j^{(3)} := E\left(C_{i,j}^2 E\left((F_{i,j}-f_j)^3|C_{i,j}\right)\right) \quad \text{and} \quad \kappa_j^{(4)} := E\left(C_{i,j}^2 E\left((F_{i,j}-f_j)^4|C_{i,j}\right)\right) \tag{2.24}$$

exist and are finite, respectively.

Note that the definitions of the quantities $\kappa_j^{(3)}$ and $\kappa_j^{(4)}$ above are in terms of the central third and fourth conditional moments $E((F_{i,j}-f_j)^3|C_{i,j})$ and $E((F_{i,j}-f_j)^4|C_{i,j})$, respectively, which allows a more convenient representation of the limiting variance in Theorem 2.9 below. However, using Assumptions 2.2 and 2.3, it is always possible to represent the *non-central* third and fourth conditional moments in terms of the *central* conditional moments. Precisely, we have

$$E(F_{i,j}^3|C_{i,j}) = f_j^3 + 3f_j\frac{\sigma_j^2}{C_{i,j}} + E\left((F_{i,j}-f_j)^3|C_{i,j}\right), \tag{2.25}$$

$$E(F_{i,j}^4|C_{i,j}) = f_j^4 + 6f_j^2\frac{\sigma_j^2}{C_{i,j}} + 4f_jE\left((F_{i,j}-f_j)^3|C_{i,j}\right) + E\left((F_{i,j}-f_j)^4|C_{i,j}\right). \tag{2.26}$$

In the following example, we illustrate the general representation of the third and fourth conditional moments in (2.25) and (2.26) for several important parametric families of distributions and provide explicit formulas for the corresponding central conditional moments for Gaussian, gamma and log-normal distributions.

Example 2.8 (Central third and fourth conditional moments for parametric families)

a) If $F_{i,j}|C_{i,j} \sim \mathcal{N}(f_j, \frac{\sigma^2}{C_{i,j}})$, Equations (2.18) *as well as* (2.25) *and* (2.26) *hold, where*

$$E\left((F_{i,j}-f_j)^3|C_{i,j}\right) = 0 \quad \textit{and} \quad E\left((F_{i,j}-f_j)^4|C_{i,j}\right) = 3\frac{\sigma_j^4}{C_{i,j}^2}.$$

b) If $F_{i,j}|C_{i,j} \sim Gamma(\alpha, \beta)$ with $\alpha = f_j^2\frac{C_{i,j}}{\sigma_j^2}$ and $\beta = f_j\frac{C_{i,j}}{\sigma_j^2}$, Equations (2.18) *as well as* (2.25) *and* (2.26) *hold, where*

$$E\left((F_{i,j}-f_j)^3|C_{i,j}\right) = 2\frac{\sigma_j^4}{f_jC_{i,j}^2} \quad \textit{and} \quad E\left((F_{i,j}-f_j)^4|C_{i,j}\right) = 3\frac{\sigma_j^4}{C_{i,j}^2} + 6\frac{\sigma_j^6}{C_{i,j}^3f_j^2}.$$

c) If $F_{i,j}|C_{i,j} \sim log\mathcal{N}(\mu, \sigma^2)$ with $\mu = \log(f_j^2/(\frac{\sigma_j^2}{C_{i,j}} + f_j^2)^{1/2})$ and $\sigma^2 = \log(1 + \frac{\sigma_j^2}{C_{i,j}f_j^2})$, Equations (2.18) *as well as* (2.25) *and* (2.26) *hold, where*

$$\begin{aligned} E\left((F_{i,j}-f_j)^3|C_{i,j}\right) &= 3\frac{\sigma_j^4}{f_jC_{i,j}^2} + \frac{\sigma_j^6}{f_j^3C_{i,j}^3}, \\ E\left((F_{i,j}-f_j)^4|C_{i,j}\right) &= 3\frac{\sigma_j^4}{C_{i,j}^2} + 16\frac{\sigma_j^6}{C_{i,j}^3f_j^2} + 15\frac{\sigma_j^8}{C_{i,j}^4f_j^4} + 6\frac{\sigma_j^{10}}{C_{i,j}^5f_j^6} + \frac{\sigma_j^{12}}{C_{i,j}^6f_j^8}. \end{aligned}$$

Now, Assumption 2.7 allows to derive the limiting variance of $\widehat{\sigma}_j^2$ for each fixed j and to state a CLT in the following theorem.

Theorem 2.9 (Asymptotic Normality of $\widehat{\sigma}_j^2$) *Suppose Assumptions 2.2, 2.3, and 2.7 hold. Then, as $I \to \infty$, for each fixed j, $j \in \mathbb{N}_0$, we have*

$$\sqrt{I-j}\left(\widehat{\sigma}_j^2 - \sigma_j^2\right) \xrightarrow{d} \mathcal{N}\left(0, \Psi_j\right),$$

where $\Psi_j = \kappa_j^{(4)} - \sigma_j^4$.

In comparison to the CLTs for $\widehat{f}_j$ and (smooth functions of) $\underline{\widehat{f}}_K$ in Theorem 2.5 and Corollary 2.6, respectively, conditions imposed for the first two (conditional) moments in Assumption 2.1 and 2.3 are not sufficient to prove asymptotic normality of $\widehat{\sigma}_j^2$ in Theorem 2.9. Assumption 2.7 requires $\kappa_j^{(3)}$ and $\kappa_j^{(4)}$ to be finite, which is equivalent to have that

$C_{i,j}^2 E((F_{i,j}-f_j)^3|C_{i,j})$ and $C_{i,j}^2 E((F_{i,j}-f_j)^4|C_{i,j})$ are integrable. Let $\nu_j^{(k)} := E(C_{i,j}^{-k}) < \infty$ denote the kth reciprocal moment of $C_{i,j}$. Then, in view of Example 2.8, Assumption 2.7 does always hold for case a), but additionally requires $\nu_j^{(1)}$ to be finite for case b), and even $\nu_j^{(4)}$ to be finite for case c).

In the following example, we illustrate how the different parametric distributions affect the limiting variance Ψ_j of $\widehat{\sigma}_j^2$ derived in Theorem 2.9.

Example 2.10 (On Ψ_j for parametric families) *Let the assumptions of Theorem 2.9 hold.*

a) If $F_{i,j}|C_{i,j} \sim \mathcal{N}(f_j, \frac{\sigma_j^2}{C_{i,j}})$, we have $\Psi_j = 2\sigma_j^4$.

b) If $F_{i,j}|C_{i,j} \sim Gamma(\alpha, \beta)$ with $\alpha = f_j^2 \frac{C_{i,j}}{\sigma_j^2}$ and $\beta = f_j \frac{C_{i,j}}{\sigma_j^2}$, we have

$$\Psi_j = 2\sigma_j^4 + \frac{6\sigma_j^6}{f_j^2}\nu_j^{(1)}.$$

c) If $F_{i,j}|C_{i,j} \sim log\mathcal{N}(\mu, \sigma^2)$ with $\mu = \log(f_j^2/(\frac{\sigma_j^2}{C_{i,j}} + f_j^2)^{1/2})$ and $\sigma^2 = \log(1 + \frac{\sigma_j^2}{C_{i,j} f_j^2})$, we have

$$\Psi_j = 2\sigma_j^4 + \frac{16\sigma_j^6}{f_j^2}\nu_j^{(1)} + \frac{15\sigma_j^8}{f_j^4}\nu_j^{(2)} + \frac{6\sigma_j^{10}}{f_j^6}\nu_j^{(3)} + \frac{\sigma_j^{12}}{f_j^8}\nu_j^{(4)}.$$

It is worth noting that we included a conditional normal distribution in Examples 2.8 and 2.10 although it is *not* covered by Assumption 2.3 as it allows for negative support for $F_{i,j}$ conditional on $C_{i,j}$. Nevertheless, it is often used in practice together with a truncation step to cope with issues caused by individual development factors estimated negative.

The reciprocal moments $\nu_{j+1}^{(k)}$, $k = 1, 2, 3, 4$ do not have explicit expressions, but they can always be represented as

$$\nu_{j+1}^{(k)} = E\left(C_{i,j+1}^{-k}\right) = E\left(C_{i,j}^{-k} E\left(F_{i,j}^{-k}|C_{i,j}\right)\right). \tag{2.27}$$

For the gamma distribution in Example 2.10, $E\left(F_{i,j}^{-k}|C_{i,j}\right)$ exists if $\alpha = f_j^2 \frac{C_{i,j}}{\sigma_j^2} > k$ holds, and for the log-normal distribution, $E(F_{i,j}^{-k}|C_{i,j})$ does always exist. The reciprocal moments can be approximated by a Taylor expansion (of order two) leading to $\nu_{j+1}^{(k)} \approx \frac{1}{\mu_{j+1}^k} + \frac{k(k+1)\tau_{j+1}^2}{2\mu_{j+1}^{k+2}}$.

2.4 Asymptotic Theory for Mack's model: Reserve Risk

For non-life insurance companies, an accurate estimation of the reserves is crucial for pricing future policies and for the assessment of their solvency. For this purpose, the CLM is widely used to forecast future claims (summarized in R_I) for reserving based on the *point prediction* $\widehat{R}_I$ (best estimate) defined in (2.12). In addition, by prudential regulation, insurance companies need to measure their *reserve risk*, i.e. that their reserve will be not sufficient to pay for all outstanding claims. Hence, insurers have strong interest in accurate approximations of the distribution of the stochastic (unobserved) reserve R_I defined in (2.4) centered around its best estimate $\widehat{R}_I$ defined in (2.12). In the following, we call this difference $R_I - \widehat{R}_I$ the *predictive root of the reserve*. In particular, high quantiles of the distribution of $R_I - \widehat{R}_I$ can serve to construct asymptotic theory-backed approximations for the value-at-risk or other risk measures.

2.4.1 Sequences of development factors and variance parameters

In comparison to the asymptotic theory established in Section 2.3, where only an arbitrary large, but still a *fixed* number of estimators $\widehat{f}_j$ and $\widehat{\sigma}_j^2$ are considered, we have to impose some regularity conditions on the whole sequences of development factors $(f_j, j \in \mathbb{N}_0)$ and variance parameters $(\sigma_j^2, j \in \mathbb{N}_0)$.

Assumption 2.11 (Development factors and variance parameters) *Letting $I \to \infty$ 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 $f_j \to 1$ as $j \to \infty$ 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 $\sigma_j^2 \to 0$ as $j \to \infty$ 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 2.11 are rather mild and cover short-term and long-term insurance lines. In practice, each claim has a finite, but possibly unknown horizon until it is finally settled. The time horizons of claim developments vary by the insurance lines, which are usually categorized in short-term and long-term. Whereas most claims in short-tailed insurance businesses are usually notified and/or settled in a short period after the date of exposure and/or occurrence, claims in long-tailed insurance lines may be settled a long time after the insurance policy has already expired.

Assumption 2.11 allows to derive several useful properties for the cumulative claims $C_{i,j}$ for fixed i and increasing j as summarized in the following lemma.

Lemma 2.12 (Properties of $(\mu_j, j \in \mathbb{N}_0)$ and $(\tau_j^2, j \in \mathbb{N}_0)$) *Suppose Assumptions 2.2, 2.3, and 2.11 hold and let $\mu_j = E(C_{i,j})$ and $\tau_j^2 = Var(C_{i,j})$ as defined in* (2.20) *and* (2.21). *Then, both sequences $(\mu_j, j \in \mathbb{N}_0)$ and $(\tau_j^2, j \in \mathbb{N}_0)$ are non-negative and monotonically non-decreasing, that is, we have $\mu_{j+1} \geq \mu_j \geq 1$ as well as $\tau_{j+1}^2 \geq \tau_j^2 > 0$ for all $j \in \mathbb{N}_0$. Moreover, both sequences are converging and, for $j \to \infty$, we have*

$$\mu_j \to \mu_\infty := \mu_0 \prod_{k=0}^{\infty} f_k < \infty, \tag{2.28}$$

$$\tau_j^2 \to \tau_\infty^2 := \tau_0^2 \prod_{k=0}^{\infty} f_k^2 + \mu_0 \sum_{l=0}^{\infty} \left(\prod_{m=0}^{l-1} f_m \right) \sigma_l^2 \left(\prod_{n=l+1}^{\infty} f_n^2 \right) < \infty \tag{2.29}$$

and, for $h \in \mathbb{N}$ fixed,

$$Cov(C_{i,j}, C_{i,j+h}) = \left(\prod_{k=j}^{j+h-1} f_k \right) \tau_j^2 \to \tau_\infty^2. \tag{2.30}$$

2.4.2 Asymptotic framework for reserve prediction

First, let us recap the stochastic prediction model of R_I and the point prediction $\widehat{R}_I$ leading to the predictive root $R_I - \widehat{R}_I$. Having observed the upper loss triangle $\mathcal{D}_I$ defined in (2.1), the last cumulative claims $C_{i,I-i}$ observed in calendar year I are those on the diagonal $\mathcal{Q}_I$ defined in (2.2). Hence, starting from the diagonal $\mathcal{Q}_I$, the predictive root of the reserve is defined as the difference of the stochastic (unobserved) reserve R_I and its best estimate $\widehat{R}_I$, that is

$$R_I - \widehat{R}_I = \sum_{i=0}^{I} C_{i,I-i} \left(\prod_{j=I-i}^{I-1} F_{i,j} - \prod_{j=I-i}^{I-1} \widehat{f}_j \right) = \sum_{i=0}^{I} C_{I-i,i} \left(\prod_{j=i}^{I-1} F_{I-i,j} - \prod_{j=i}^{I-1} \widehat{f}_j \right), \tag{2.31}$$

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

With the loss triangle $\mathcal{D}_I$ at hand, a conditional asymptotic analysis of $R_I - \widehat{R}_I$ is of interest. For this purpose, as common for predictive inference (see e.g. Paparoditis and Shang (2021) for a recent reference), we shall employ a different asymptotic framework in this section. In comparison to the seemingly more "natural" asymptotic framework for $I \to \infty$ employed in Section 2.3, where increasing I means adding new diagonals $\mathcal{Q}_{I+h} = \{C_{I-i,i} | i = 0, \dots, I+h\}$, $h \geq 1$ to the loss triangle $\mathcal{D}_I$ (see the upper part of Table 3), we shall keep the *latest* cumulative claims in $\mathcal{D}_I$, that is, $\mathcal{Q}_I$, fixed and let $\mathcal{D}_I$ grow

	Development Year j						
Accident Year i	0	1	2	$\cdots$	I	$I+1$	$I+2$
0	$C_{0,0}$	$C_{0,1}$		$\cdots$	$C_{0,I}$	$C_{0,I+1}$	$C_{0,I+2}$
1	$C_{1,0}$			$\cdots$		$C_{1,I+1}$	
2							
		$\cdots$					
	$\cdots$		$\cdots$				
$\cdots$	$\cdots$						
I	$C_{I,0}$						
$I+1$	$C_{I+1,0}$						
$I+2$	$C_{I+2,0}$						

	Development Year j						
Accident Year i	0	1	2	$\cdots$	I	$I+1$	$I+2$
-2	$C_{-2,0}$			$\cdots$		$C_{-2,I+1}$	$C_{-2,I+2}$
-1	$C_{-1,0}$			$\cdots$		$C_{-1,I+1}$	
0	$C_{0,0}$	$C_{0,1}$			$C_{0,I}$		
1	$C_{1,0}$	$\cdots$					
	$\cdots$		$\cdots$				
$\cdots$	$\cdots$						
$I-1$	$C_{I-1,0}$						
I	$C_{I,0}$						

Table 3: 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.

instead by adding new rows of cumulative claims $\{C_{-h,i}|i = 0, \ldots, I + h\}$, $h \geq 1$ (see the lower part of Table 3). However, by extending Assumptions 2.2 and 2.3 also to negative indices i, both versions of sequences of growing loss triangles indicated in Table 3 are equal in distribution. Hence, as a main consequence, all asymptotic results derived in Section 2.3 remain valid without any further restriction.

Hence, in what follows, all asymptotic results are derived under the assumption that we have observed a loss triangle of the form

$$\mathcal{D}_{I,n} := \{C_{i,j}|i = -n, \ldots, I, \; j = 0, \ldots, I + n, \; -n \leq i + j \leq I\}. \tag{2.32}$$

with diagonal

$$\mathcal{Q}_{I,n} := \{C_{I-i,i}|i = 0, \ldots, I + n\} \tag{2.33}$$

in which we view I as *fixed* and let $n \to \infty$ leading to

$$\mathcal{D}_{I,\infty} := \{C_{i,j} | i \in \mathbb{Z}, i \leq I,\ j \in \mathbb{N}_0,\ i+j \leq I\} \quad \text{and} \quad \mathcal{Q}_{I,\infty} := \{C_{I-i,i} | i \in \mathbb{N}_0\}\,. \tag{2.34}$$

In accordance to (2.31), in this asymptotic setup of $n \to \infty$, the predictive root of the reserve is denoted by $R_{I,n} - \widehat{R}_{I,n}$, which can be decomposed in 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

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

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

$$\widehat{f}_{j,n} := \frac{\sum\limits_{i=-n}^{I-j-1} C_{i,j+1}}{\sum\limits_{i=-n}^{I-j-1} C_{i,j}}, \tag{2.35}$$

according to (2.9).

In the following Sections 2.4.3, 2.4.4 and 2.4.5, we derive unconditional and conditional asymptotic theory for $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$ as well as joint results.

2.4.3 Asymptotics for reserve prediction: process uncertainty

In this section, we consider the first term $(R_{I,n} - \widehat{R}_{I,n})_1$, which corresponds to the *process uncertainty* inherent in the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$. We derive asymptotic theory for $(R_{I,n} - \widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,\infty}$ as well as the unconditional limiting distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ for $n \to \infty$. Note that $(R_{I,n} - \widehat{R}_{I,n})_1 | \mathcal{D}_{I,\infty} = (R_{I,n} - \widehat{R}_{I,n})_1 | \mathcal{Q}_{I,\infty}$ holds.

Theorem 2.13 (Unconditional and conditional asymptotics for $(\boldsymbol{R}_{I,n} - \widehat{\boldsymbol{R}}_{I,n})_1$)

Suppose Assumptions 2.2, 2.3 and 2.11 hold. Then, as $n \to \infty$, unconditionally as well as conditionally on $\mathcal{Q}_{I,\infty}$ (or on $\mathcal{D}_{I,\infty}$), $(R_{I,n} - \widehat{R}_{I,n})_1$ converges in L_2-sense to the *non-degenerate random variable* $(R_{I,\infty} - \widehat{R}_{I,\infty})_1$. *That is, we have*

$$E\left(\left((R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,\infty} - \widehat{R}_{I,\infty})_1\right)^2\right) \to 0, \tag{2.36}$$

$$E\left(\left((R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,\infty} - \widehat{R}_{I,\infty})_1\right)^2 |\mathcal{Q}_{I,\infty}\right) \xrightarrow{p} 0, \tag{2.37}$$

where

$$(R_{I,\infty} - \widehat{R}_{I,\infty})_1 := \sum_{i=0}^{\infty} C_{I-i,i}\left(\prod_{j=i}^{\infty} F_{I-i,j} - \prod_{j=i}^{\infty} f_j\right) \sim \mathcal{G}_1. \tag{2.38}$$

Unconditionally and conditionally on $\mathcal{Q}_{I,\infty}$, *the (limiting) distribution* $\mathcal{G}_1$ *has mean zero, i.e.* $E((R_{I,\infty} - \widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}) = 0$ *and* $E((R_{I,\infty} - \widehat{R}_{I,\infty})_1) = 0$, *and variances*

$$Var\left((R_{I,\infty} - \widehat{R}_{I,\infty})_1\right) = \sum_{i=0}^{\infty} \mu_i \sum_{j=i}^{\infty}\left(\prod_{k=i}^{j-1} f_k\right)\sigma_j^2\left(\prod_{l=j+1}^{\infty} f_l^2\right) < \infty, \tag{2.39}$$

$$Var\left((R_{I,\infty} - \widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}\right) = \sum_{i=0}^{\infty} C_{I-i,i} \sum_{j=i}^{\infty}\left(\prod_{k=i}^{j-1} f_k\right)\sigma_j^2\left(\prod_{l=j+1}^{\infty} f_l^2\right) = O_P(1). \tag{2.40}$$

Remark 2.14 (On the limiting distributions $\mathcal{G}_1$ and $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$) *The conditional limiting distribution* $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ *as well as the unconditional limiting distribution* $\mathcal{G}_1$ *will be typically non-Gaussian. In particular, in the setup of Theorem 2.13, it is neither possible to show asymptotic normality by employing a suitable CLT nor it is possible to prove asymptotic normality wrong. This is because, unconditionally and conditionally, both the Lindeberg Condition and the Feller Condition do not hold for* $(R_{I,n} - \widehat{R}_{I,n})_1$ *in the framework of Theorem 2.13.*
Nevertheless, both $\mathcal{G}_1$ *and* $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ *could be Gaussian, if all summands* $C_{I-i,i}(\prod_{j=i}^{I+n-1} F_{I-i,j} - \prod_{j=i}^{I+n-1} f_j)$ *were, unconditionally or conditionally on* $\mathcal{Q}_{I,\infty}$, *jointly Gaussian, respectively. However, as these summands rely on products of (dependent) random variables, it is unclear under which conditions imposed on the* $F_{i,j}$*'s this would be the case.*

The following example illustrates the previous remark.

Example 2.15 (Non-Gaussianity of $\mathcal{G}_1$ and $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$) *Suppose Assumptions 2.2, 2.3 and 2.11 hold. Let* $f_0 \geq 1$, $\sigma_0^2 \in (0,\infty)$ *as well as* $f_j = 1$ *and* $\sigma_j^2 = 0$ *for all* $j \in \mathbb{N}$. *Further, suppose that* $F_{i,j}|C_{i,j} \sim Gamma(\alpha, \beta)$ *with* $\alpha = f_j^2 \frac{C_{i,j}}{\sigma_j^2}$ *and* $\beta = f_j \frac{C_{i,j}}{\sigma_j^2}$ *for all* $i \in \mathbb{Z}$, $i \leq I$ *and* $j = 0$. *Then, as we have* $F_{i,j} = f_j = 1$ *a.s. for all* $i \in \mathbb{Z}$, $i \leq I$ *and* $j \in \mathbb{N}$, *in this setup, we get*

$$\left(R_{I,\infty} - \widehat{R}_{I,\infty}\right)_1 = \sum_{i=0}^{\infty} C_{I-i,i}\left(\prod_{j=i}^{\infty} F_{I-i,j} - \prod_{j=i}^{\infty} f_j\right) = C_{I,0}\left(F_{I,0} - f_0\right).$$

In general, neither unconditionally nor conditionally, $C_{I,0}\left(F_{I,0}-f_0\right)$ *follows a Gaussian distribution. Conditional on* $\mathcal{Q}_{I,\infty}$ *(i.e. conditional on* $C_{I,0}$*),* $C_{I,0}\left(F_{I,0}-f_0\right)$ *follows a centered gamma distribution with variance* $C_{I,0}\sigma_0^2$*, skewness* $\frac{2\sigma_0}{f_0\sqrt{C_{I,0}}}$ *and excess of kurtosis* $\frac{6\sigma_0^2}{f_0^2C_{I,0}}$.

In the example above, only one summand of $\left(R_{I,\infty}-\widehat{R}_{I,\infty}\right)_1$ remains. Nevertheless, this setup is covered by Assumptions 2.2, 2.3 and 2.11. The arguments for non-Gaussianity maintain also for other sequences $(f_j, j\in\mathbb{N}_0)$ and $(\sigma_j^2, j\in\mathbb{N}_0)$ that fulfill Assumption 2.11.

2.4.4 Asymptotics for reserve prediction: estimation uncertainty

In this section, we consider the second term $(R_{I,n}-\widehat{R}_{I,n})_2$, which corresponds to the *estimation uncertainty* inherent in the predictive root of the reserve $R_{I,n}-\widehat{R}_{I,n}$. In comparison to the conditional and unconditional L_2-limiting theory derived for the first part $(R_{I,n}-\widehat{R}_{I,n})_1$ in Theorem 2.13, the derivation of asymptotic results for $(R_{I,n}-\widehat{R}_{I,n})_2$ is rather different and also much more cumbersome. For instance, to obtain non-degenerate limiting distributions, we have to inflate $(R_{I,n}-\widehat{R}_{I,n})_2$ by $\sqrt{I+n+1}$. Moreover, $(R_{I,n}-\widehat{R}_{I,n})_2$ is measurable with respect to $\mathcal{D}_{I,\infty}$, but not with respect to $\mathcal{Q}_{I,\infty}$. Hence, the derivation of the unconditional and conditional asymptotic theory for $(R_{I,n}-\widehat{R}_{I,n})_2$ requires additional assumptions, but also different techniques of proof.

In the following Sections 2.4.4.1 and 2.4.4.2, unconditional and conditional asymptotic theory for $(R_{I,n}-\widehat{R}_{I,n})_2$ will be addressed separately.

2.4.4.1 Unconditional asymptotic theory for estimation uncertainty

For the derivation of the unconditional limiting distribution of $\sqrt{I+n+1}(R_{I,n}-\widehat{R}_{I,n})_2$ for $n\to\infty$, we make use of the CLTs established in Section 2.3, in particular, from Corollary 2.6. Recall that all asymptotic normality results from Section 2.3 remain valid also under the different notion of the asymptotic framework in Section 2.4. For the derivation of asymptotic theory, we have to impose some additional regularity conditions on the stochastic properties of the individual development factors $F_{i,j}$'s to strengthen Assumptions 2.3 and 2.11.

Assumption 2.16 (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 $j \to \infty$ such that $\sum\limits_{j=0}^{\infty}(j+1)^2\frac{\sigma_j^2}{\epsilon^j} < \infty$.*

Now, for $n \to \infty$, this allows to derive the limiting distribution of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ in the following theorem.

Theorem 2.17 (Unconditional asymptotic theory for $(\boldsymbol{R_{I,n}} - \widehat{\boldsymbol{R}}_{\boldsymbol{I,n}})_2$) *Suppose Assumptions 2.2, 2.3, 2.11 and 2.16 hold. Then, as $n \to \infty$, $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ converges in distribution to a non-degenerate limiting distribution $\mathcal{G}_2$. That is, we have*

$$\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2 \xrightarrow{d} \langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty \rangle \sim \mathcal{G}_2, \tag{2.41}$$

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

$$Cov(Y_{i_1}, Y_{i_2}) = \lim_{K\to\infty} \mathbf{\Sigma}_{K,\prod f_j}(i_1, i_2) = \sum_{j=\max(i_1,i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{\substack{l=\max(i_1,i_2)\\ l\neq j}}^{\infty} f_l^2 \prod_{m=\min(i_1,i_2)}^{\max(i_1,i_2)-1} f_m \tag{2.42}$$

for $i_1, i_2 \in \mathbb{N}_0$, where $\mathbf{\Sigma}_{K,\prod f_j}(i_1, i_2)$ is defined in Corollary 2.6. Here, the two random sequences $\mathcal{Q}_{I,\infty}$ and $\mathbf{Y}_\infty$ are stochastically independent, and the limiting distribution $\mathcal{G}_2$ has mean zero and (finite) variance

$$\begin{aligned} &Var\left(\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty \rangle\right) \qquad\qquad (2.43)\\ =&\sum_{i=0}^{\infty}(\tau_i^2+\mu_i^2)\sum_{j=i}^{\infty}\frac{\sigma_j^2}{\mu_j}\prod_{\substack{l=i\\ l\neq j}}^{\infty}f_l^2 + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty}\mu_{i_1}\mu_{i_2}\sum_{j=\max(i_1,i_2)}^{\infty}\frac{\sigma_j^2}{\mu_j}\prod_{\substack{l=\max(i_1,i_2)\\ l\neq j}}^{\infty}f_l^2\prod_{m=\min(i_1,i_2)}^{\max(i_1,i_2)-1}f_m. \end{aligned}$$

The limiting distribution $\mathcal{G}_2$ will be generally non-Gaussian as $\mathcal{G}_2$ is the distribution of an inner product of some independent, but not identically distributed sequence $\mathcal{Q}_{I,\infty}$ (with unspecified distribution) and a (dependent) Gaussian sequence $\mathbf{Y}_\infty$. In the Example 2.21 below, which picks up the setup of Example 2.15, the unconditional limiting distribution $\mathcal{G}_2$ is illustrated together with its conditional version.

2.4.4.2 Conditional asymptotic theory for estimation uncertainty

Next, we derive the limiting distribution of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ conditional on $\mathcal{Q}_{I,\infty}$ for $n \to \infty$ by making use of the conditional versions of the CLTs derived in Section 2.3, which can be found in the appendix. In particular, we employ Corollary 2.25 which contains a conditional version of Corollary 2.6 established under the asymptotic framework of Section 2.4. For this purpose, we further decompose $(R_{I,n} - \widehat{R}_{I,n})_2$ to get

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

where $(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ is measurable wrt $\mathcal{Q}_{I,\infty}$ and $f_{j,n}(\mathcal{Q}_{I,\infty}) := \mu^{(1)}_{j+1,n}(\mathcal{Q}_{I,\infty})/\mu^{(2)}_{j,n}(\mathcal{Q}_{I,\infty})$ with

$$\begin{aligned}
\mu^{(1)}_{j+1,n}(\mathcal{Q}_{I,\infty}) &:= E\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j+1} | \mathcal{Q}_{I,\infty} \right) = \frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} E(C_{i,j+1}|C_{i,I-i}), \\
\mu^{(2)}_{j,n}(\mathcal{Q}_{I,\infty}) &:= E\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} | \mathcal{Q}_{I,\infty} \right) = \frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} E(C_{i,j}|C_{i,I-i}).
\end{aligned}$$

In addition to the assumptions imposed for the derivation of the unconditional asymptotics in Theorem 2.17, we require also a regularity condition for the *backward* conditional distribution of cumulative claim $C_{i,j}$ given $C_{i,j+1}$.

Assumption 2.18 (Backward conditional moments) *Assumptions 2.2, 2.3, 2.11 and 2.16 are fulfilled such that, for all $K \in \mathbb{N}_0$, $k \geq 0$ and $j, j_1, j_2 \in \{0, \ldots, K\}$, $j_1 \leq j_2$, we have*

$$|E(C_{i,j}|C_{i,j+k}) - E(C_{i,j}|C_{i,j+k+1})| \leq a_k X_i, \tag{2.44}$$

$$|Cov(C_{i,j_1}, C_{i,j_2}|C_{i,j_2+k}) - Cov(C_{i,j_1}, C_{i,j_2}|C_{i,j_2+k+1})| \leq b_k Y_i, \tag{2.45}$$

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 $\sum_{j=0}^{\infty}(j+1)^2 b_j < \infty$.

Note that the above assumption is required, as Mack's model is designed to generate loss triangles in a rather simple *forward* way starting with development year 0 (first column in Table 2) and then, independently for each row i, the whole loss triangle is easily generated column by column according to (2.17). That is, by recursively multiplying a cumulative claim $C_{i,j}$ with individual development factors $F_{i,j}$, $F_{i,j+1}$, etc. , we get $C_{i,j+1}$, $C_{i,j+2}$, etc. By construction, the conditional distribution of an individual development factor $F_{i,j}$ depends on the realization of $C_{i,j}$ (see Assumptions 2.1 and 2.3). Hence, it is easy to calculate *forward* conditional means $E(C_{i,j+1}|C_{i,j})$ and variances $Var(C_{i,j+1}|C_{i,j})$, but it is

not straightforward to calculate *backward* conditional means $E(C_{i,j}|C_{i,j+1})$ and variances $Var(C_{i,j}|C_{i,j+1})$.
The following example illustrates the backward conditional moments for a special case based on the uniform distribution.

Example 2.19 (Forward and backward conditional moments) *For $j \in \mathbb{N}$, let C_j denote a cumulative claim and F_j the corresponding individual development factor leading to the next claim $C_{j+1} = C_j F_j$. Suppose that C_j is uniformly distributed with $C_j \sim U(\mu_j - \sqrt{3\tau_j^2}, \mu_j + \sqrt{3\tau_j^2})$ such that $E(C_j) = \mu_j$ and $Var(C_j) = \tau_j^2$, and F_j, conditional on C_j, is also uniformly distributed with $F_j|C_j \sim U(f_j - \sqrt{3\sigma_j^2/C_j}, f_j + \sqrt{3\sigma_j^2/C_j})$ such that $E(F_j|C_j) = f_j$ and $Var(F_j|C_j) = \frac{\sigma_j^2}{C_j}$. Further, we assume that $\mu_j - \sqrt{3\tau_j^2} > 0$ as well as $f_j - \sqrt{3\sigma_j^2/C_j} > 0$ for all $C_j \in [\mu_j - \sqrt{3\tau_j^2}, \mu_j + \sqrt{3\tau_j^2}]$. Note that this setup is covered by Assumption 2.3. Then, for the backward conditional mean and variance, we get*

$$E\left(C_j|C_{j+1}\right) = \frac{C_{j+1}}{f_j} + O_P\left(\sigma_j\right) \quad \text{and} \quad Var\left(C_j|C_{j+1}\right) = \frac{\sigma_j^2 C_{j+1}}{f_j^3} + O_P\left(\sigma_j^3\right). \tag{2.46}$$

Now, with Assumption 2.18 in place, this allows to derive asymptotic theory for both parts of $(R_{I,n} - \widehat{R}_{I,n})_2$. Precisely, in the following theorem, we show that $(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$, which is measurable wrt $\mathcal{Q}_{I,\infty}$, converges unconditionally to some limiting distribution $\mathcal{G}_2^{(1)}$ as well as asymptotic normality of $(R_{I,n} - \widehat{R}_{I,n})_2^{(2)}$ conditionally on $\mathcal{Q}_{I,\infty}$.

Theorem 2.20 (Asymptotic theory for $(\boldsymbol{R_{I,n}} - \widehat{\boldsymbol{R}}_{\boldsymbol{I,n}})_{\boldsymbol{2}}$ conditional on $\boldsymbol{\mathcal{Q}_{I,\infty}}$) *Suppose Assumptions 2.2, 2.3, 2.11, 2.16 and 2.18 hold. Then, as $n \to \infty$, 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

$$\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(1)} \xrightarrow{d} \left\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty^{(1)} \right\rangle \sim \mathcal{G}_2^{(1)}, \tag{2.47}$$

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

$$Cov(Y_{i_1}^{(1)}, Y_{i_2}^{(1)}) = \lim_{K \to \infty} \mathbf{\Sigma}_{K,\prod f_j}^{(1)}(i_1, i_2) \tag{2.48}$$

for $i_1, i_2 \in \mathbb{N}_0$, where $\mathbf{\Sigma}_{K,\prod f_j}^{(1)}(i_1, i_2)$ is defined in Corollary 2.25. Here, the two random sequences $\mathcal{Q}_{I,\infty}$ and $\mathbf{Y}_\infty^{(1)}$ are stochastically independent.

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

$$\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(2)}|\mathcal{Q}_{I,\infty} \xrightarrow{d} \left\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty^{(2)} \right\rangle \sim \mathcal{G}_2^{(2)}|\mathcal{Q}_{I,\infty}, \tag{2.49}$$

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

$$\Xi(\mathcal{Q}_{I,\infty}) = \lim_{K\to\infty} \mathcal{Q}_{I,K-I}\mathbf{\Sigma}^{(2)}_{K,\prod f_j}\mathcal{Q}'_{I,K-I} = \lim_{K\to\infty} \mathcal{Q}_{I,K-I}(\mathbf{\Sigma}_{K,\prod f_j} - \mathbf{\Sigma}^{(1)}_{K,\prod f_j})\mathcal{Q}'_{I,K-I}, \tag{2.50}$$

where $\mathbf{\Sigma}_{K,\prod f_j}$ as well as $\mathbf{\Sigma}^{(1)}_{K,\prod f_j}$ and $\mathbf{\Sigma}^{(2)}_{K,\prod f_j}$ are defined in Corollaries 2.6 and 2.25, respectively.

The following example picks up the setup of Example 2.15 and illustrates the unconditional limiting distributions $\mathcal{G}_2$ and $\mathcal{G}_2^{(1)}$ from Theorem 2.17 and Theorem 2.20(i), respectively, as well as the conditional limiting distribution $\mathcal{G}_2^{(2)}|\mathcal{Q}_{I,\infty}$ from Theorem 2.20(ii).

Example 2.21 (On $\mathcal{G}_2$, $\mathcal{G}_2^{(1)}$ and $\mathcal{G}_2^{(2)}|\mathcal{Q}_{I,\infty}$)

(i) Suppose that Assumptions 2.2, 2.3, 2.11 and 2.16 hold in the setup of Example 2.15. Then, we have

$$\sqrt{I+n+1}\left(R_{I,n} - \widehat{R}_{I,n}\right)_2 = C_{I,0}\left(\sqrt{I+n+1}\left(f_0 - \widehat{f}_0\right)\right),$$

where $\sqrt{I+n+1}\left(f_0 - \widehat{f}_0\right)$ is asymptotically Gaussian with mean zero and variance $\frac{\sigma_0^2}{\mu_0}$ according to Theorem 2.5. In total, the limiting distribution of $C_{I,0}(\sqrt{I+n+1}(f_0 - \widehat{f}_0))$ is non-Gaussian with mean zero and limiting variance $(\tau_0^2 + \mu_0^2)\frac{\sigma_0^2}{\mu_0}$.

(ii) If, additionally, Assumption 2.18 holds, we have

$$\begin{aligned}&\sqrt{I+n+1}\left(R_{I,n} - \widehat{R}_{I,n}\right)_2\\ =&C_{I,0}\left(\sqrt{I+n+1}\left(f_0 - f_{0,n}(\mathcal{Q}_{I,\infty})\right)\right) + C_{I,0}\left(\sqrt{I+n+1}\left(f_{0,n}(\mathcal{Q}_{I,\infty}) - \widehat{f}_0\right)\right),\end{aligned}$$

where $\sqrt{I+n+1}\left(f_0 - f_{0,n}(\mathcal{Q}_{I,\infty})\right)$ is asymptotically Gaussian with mean zero and variance $\sigma^2_{f_0,1}$ according to Theorem 2.24.

In total, the limiting distribution of $C_{I,0}\left(\sqrt{I+n+1}\left(f_0 - f_{0,n}(\mathcal{Q}_{I,\infty})\right)\right)$ is non-Gaussian with mean zero and limiting variance $(\tau_0^2 + \mu_0^2)\sigma^2_{f_0,1}$. Moreover, conditional on $\mathcal{Q}_{I,\infty}$,

$C_{I,0}\left(\sqrt{I+n+1}\left(f_{0,n}(\mathcal{Q}_{I,\infty})-\widehat{f}_0\right)\right)$ *is asymptotically Gaussian with mean zero and variance* $C_{I,0}^2\left(\frac{\sigma_0^2}{\mu_0}-\sigma_{f_0,1}^2\right)$.

2.4.5 Joint asymptotics for reserve prediction

After having established the limiting conditional and unconditional distributions for both parts $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$ of the predictive root of the reserve $R_{I,n}-\widehat{R}_{I,n}$ separately in Theorems 2.13, 2.17 and 2.20, we are now concerned with their joint asymptotics. Note that L_2-convergence in Theorem 2.13 implies convergence in distribution.

Theorem 2.22 (Joint asymptotics for $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$)

(i) Suppose the assumptions of Theorems 2.13 and 2.17 hold. Then, unconditionally, $(R_{I,n}-\widehat{R}_{I,n})_1$ *and* $(R_{I,n}-\widehat{R}_{I,n})_2$ *are uncorrelated. Hence, we have*

$$\begin{pmatrix} \left(R_{I,n}-\widehat{R}_{I,n}\right)_1 \\ \sqrt{I+n+1}\left(\left(R_{I,n}-\widehat{R}_{I,n}\right)_2\right) \end{pmatrix} \xrightarrow{d} \mathcal{G}, \tag{2.51}$$

where $\mathcal{G}$ *is a bivariate distribution with marginals* $\mathcal{G}_1$ *and* $\mathcal{G}_2$ *as given in Theorems 2.13 and 2.17 and diagonal* (2×2) *covariance matrix with variances from* (2.39) *and* (2.43) *on the diagonal, respectively.*

(ii) Suppose the assumptions of Theorems 2.13 and 2.20 hold. Then, conditionally on $\mathcal{Q}_{I,\infty}$, $(R_{I,n}-\widehat{R}_{I,n})_1$ *and* $(R_{I,n}-\widehat{R}_{I,n})_2$ *are stochastically independent. Hence, we have*

$$\begin{pmatrix} \left(R_{I,n}-\widehat{R}_{I,n}\right)_1 \\ \sqrt{I+n+1}\left(\left(R_{I,n}-\widehat{R}_{I,n}\right)_2\right) \end{pmatrix} |\mathcal{Q}_{I,\infty} \xrightarrow{d} \mathcal{G}|\mathcal{Q}_{I,\infty}, \tag{2.52}$$

where $\mathcal{G}|\mathcal{Q}_{I,\infty}$ *is a bivariate conditional distribution with conditionally independent marginals* $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ *and* $\mathcal{G}_2|\mathcal{Q}_{I,\infty}$, *where* $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ *is given in Theorem 2.13 and* $\mathcal{G}_2|\mathcal{Q}_{I,\infty}$ *is the conditional Gaussian distribution* $\mathcal{G}_2^{(2)}|\mathcal{Q}_{I,\infty}$ *from Theorem 2.20(ii) plus a realization of* $\mathcal{G}_2^{(1)}$ *from Theorem 2.20(i).*

Now, with Theorem 2.22 in place, we can also state the overall unconditional and conditional limiting distributions of the predicitve root of the reserve $R_{I,n}-\widehat{R}_{I,n}$. In this regard, note that Theorems 2.17 and 2.20 require the inflation of $(R_{I,n}-\widehat{R}_{I,n})_2$ with $\sqrt{I+n+1}$ to establish convergence towards non-degenerate limiting distributions. As this is not the case for $(R_{I,n}-\widehat{R}_{I,n})_1$ in Theorem 2.13, the latter part $(R_{I,n}-\widehat{R}_{I,n})_1$ corresponding to the process uncertainty will asymptotically dominate, which gives the following result.

Corollary 2.23 (Asymptotics for $R_{I,n} - \widehat{R}_{I,n}$)

(i) Suppose the assumptions of Theorems 2.13 and 2.17 hold. Then, unconditionally, $R_{I,n} - \widehat{R}_{I,n}$ converges in distribution to $\mathcal{G}_1$. That is, we have

$$R_{I,n} - \widehat{R}_{I,n} = (R_{I,n} - \widehat{R}_{I,n})_1 + (R_{I,n} - \widehat{R}_{I,n})_2 \xrightarrow{d} \mathcal{G}_1. \tag{2.53}$$

(ii) Suppose the assumptions of Theorems 2.13 and 2.20 hold. Then, conditionally on $\mathcal{Q}_{I,\infty}$, $R_{I,n} - \widehat{R}_{I,n}$ converges in distribution to $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$. That is, we have

$$R_{I,n} - \widehat{R}_{I,n} = (R_{I,n} - \widehat{R}_{I,n})_1 + (R_{I,n} - \widehat{R}_{I,n})_2|\mathcal{Q}_{I,\infty} \xrightarrow{d} \mathcal{G}_1|\mathcal{Q}_{I,\infty}. \tag{2.54}$$

Based on the (unconditional) uncorrelatedness and (conditional) independence of $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$ from Theorem 4.12, respectively, the corresponding variances of the predictive root $R_{I,n} - \widehat{R}_{I,n}$ decompose into two additive terms that capture the process uncertainty and estimation uncertainty. That is, unconditionally, we have $Var(R_{I,n} - \widehat{R}_{I,n}) = Var(R_{I,n}) + Var(\widehat{R}_{I,n})$ and, conditional on $\mathcal{Q}_{I,\infty}$, we get $Var(R_{I,n} - \widehat{R}_{I,n}|\mathcal{Q}_{I,\infty}) = Var(R_{I,n}|\mathcal{Q}_{I,\infty}) + Var(\widehat{R}_{I,n}|\mathcal{Q}_{I,\infty})$.

While the calculation and asymptotic theory of $Var(R_{I,n})$ and $Var(R_{I,n}|\mathcal{Q}_{I,\infty})$ is straightforward (see the proof of Theorem 2.13), the calculation of $Var(\widehat{R}_{I,n})$ and $Var(\widehat{R}_{I,n}|\mathcal{Q}_{I,\infty})$ is more cumbersome and its asymptotic treatment relies on limiting results for the estimation of the development factors.

By Corollary 2.23, we can conclude that asymptotic normality of the (predictive root of the) reserve does *not* hold, which casts the common practice to use a normal approximation for the reserve in Mack's model into doubt.

In the following section, we illustrate the non-Gaussianity of the conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,\infty}$. In particular, we demonstrate how the scaling of the claim sizes affects the deviation from Gaussianity. Moreover, we show that the shape of $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ does depend on the specific distribution of the $F_{i,j}$'s. In addition, we illustrate the non-Gaussianity of the unconditional limiting distribution of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$.

2.5 Simulation Study

In this section, we illustrate our findings from Section 2.4 by simulating several parameter scenarios to generate cumulative claims according to Mack's model.

2.5.1 Simulations for $(R_{I,n} - \widehat{R}_{I,n})_1 | \mathcal{Q}_{I,n}$

In the setup of Section 2.4, 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 case and for different parameters scenarios to be specified below, we generate representative diagonals $\mathcal{Q}_{I,n}$ to condition on first. This is done by simulating $M = 500$ loss triangles $\mathcal{D}_{I,n}^{(m)} = \{C_{i,j}^{(m)} | i = -n, \ldots, I,\ j = 0, \ldots, I+n,\ -n \le i+j \le I\}$, $m = 1, \ldots, 500$ by generating the entries in the first columns $C_{\bullet,0}$ (independently) from a uniform distribution and the individual developments factors $F_{i,j}$ from a conditional gamma distribution. Eventually, this gives diagonals $\mathcal{Q}_{I,n}^{(m)}$, $m = 1, \ldots, 500$. Note that as illustrated in Table 3 (lower panel) $\mathcal{D}_{I,k}^{(m)}$ is always a subset of $\mathcal{D}_{I,n}^{(m)}$ for all $0 \le k \le n$. Recall that we are interested in the distributions of $(R_{I,n} - \widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}$.

Then, for each m, that is, given the diagonal $\mathcal{Q}_{I,n}^{(m)}$, we simulate $L = 5000$ lower triangles $\{C_{i,j}^{(l)} | i = -n, \ldots, I,\ j = 0, \ldots, I+n,\ i+j > I\}$, $l = 1, \ldots, 5000$ using

(i) a conditional gamma distribution

(ii) a conditional log-normal distribution

(iii) a conditional left-tail truncated normal distribution (truncated at 0.1)

In all three cases, we compute the first part of the predictive roots of the reserve corresponding to the process uncertainty, i.e.

$$(R_{I,n} - \widehat{R}_{I,n})_1^{(l)}, \quad l = 1, \ldots 5000. \tag{2.55}$$

Hence, for all three distributions (i), (ii) and (iii) and for each m with $m = 1, \ldots, 500$, respectively, this leads to simulated distributions of $(R_{I,n} - \widehat{R}_{I,n})_1^{(m)}$ given the diagonal $\mathcal{Q}_{I,n}^{(m)}$.

Now, we specify the parameters for the different scenarios to be simulated. In the first scenario, we consider the setup in Example 2.15, where $F_{I,0}$ follows either a gamma, a log-normal or a truncated normal distribution given $C_{I,0}$. Here, we set $f_0 = 1.39$, $\sigma_0^2 = 509,518$, and $C_{I,0}$ is uniformly distributed on $[120 \times 10^4, 350 \times 10^4]$ for $I = 10$ and $n = 0$ with $f_j = 1$ and $\sigma_j^2 = 0$ for all $j > 0$. Figure 3 summarizes the simulation results, where we show boxplots of skewness and kurtosis based on the $M = 500$ simulated distributions, where the red line indicates the benchmark skewness and kurtosis of a normal distribution, as well as five (arbitrarily chosen) density plots. The first row of panels refers to the gamma, the second one to the log-normal, and the third one to the truncated normal distribution. Moreover, for each of the $M = 500$ simulated distributions, we applied the Kolmogorov-Smirnov test for normality with mean zero and variance $C_{I,0}^{(m)} \sigma_0^2$ of level

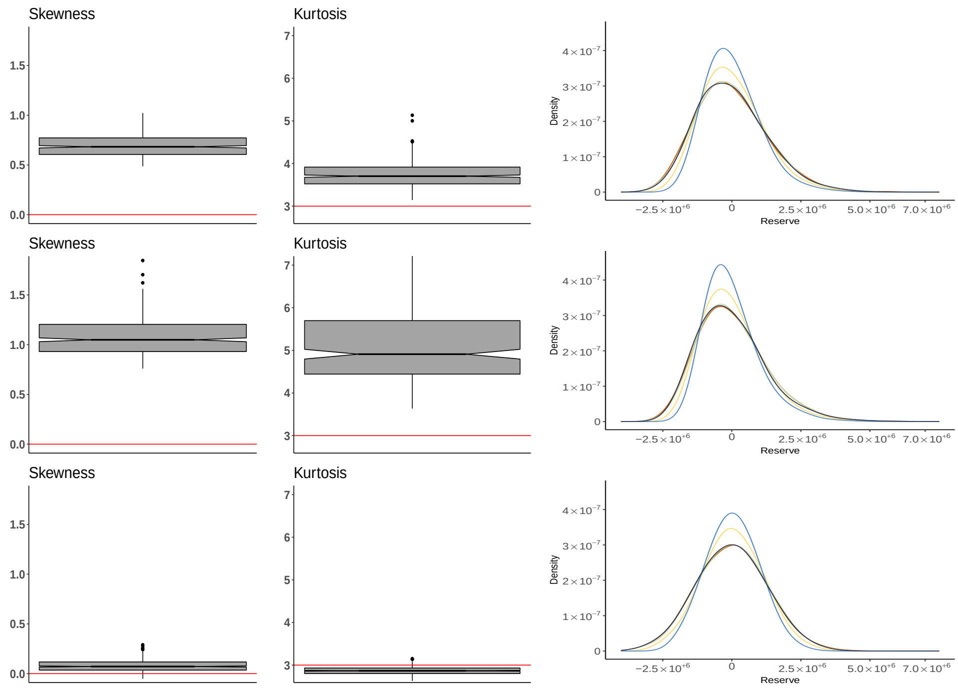

Figure 3: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $I = 10$ and $n = 0$ for the setup of Example 2.15, where $F_{I,0}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

$\alpha = 5\%$. For the gamma and for the log-normal distribution, the null is rejected for all $M = 500$ samples. By contrast, for the truncated normal distribution, the null is rejected only in about 17% out of $M = 500$ samples.

In the second scenario, we consider a more general and practically more realistic setup, where $f_j > 1$ and $\sigma_j^2 > 0$ for all $j = 0, \ldots, I+n-1$ such that f_j and σ_j^2 decrease to 1 and 0, respectively. Precisely, we use exponentially decreasing sequences $(f_j)_{j=0,\ldots I+n-1}$ and $(\sigma_j^2)_{j=0,\ldots I+n-1}$ with $f_j = 1 + e^{-1-0.7j}$ and $\sigma_j^2 = 509,518 \cdot e^{-1.3-2.7j}$ for $j = 0, \ldots, I+n-1$ for $I = 10$ and $n \in \{0, 10, 20, 30, 40\}$. Further, we distinguish between two different setups a) and b), where the parameter settings are exactly the same in both cases, but the first column $C_{\bullet,0} = (C_{-n,0}, \ldots, C_{I,0})'$ of the (upper) loss triangle is uniformly distributed on $[120 \times 10^6, 350 \times 10^6]$ in case a) and on $[120 \times 10^4, 350 \times 10^4]$ in case b). Otherwise, we use the same approach as described above to simulate the distributions of $(R_{I,n} - \widehat{R}_{I,n})_1^{(m)}$ given the diagonals $\mathcal{Q}_{I,n}^{(m)}$ for $m = 1, \ldots, 500$. Similar to what is reported in Figure 3, we show boxplots of skewness and kurtosis as well as density plots for both settings a) and b) in Figures 4 and 5 for $I = 10$ and $n = 10$ and for all three different distributions in (i), (ii) and (iii). Plots for $n \in \{0, 20, 30, 40\}$ can be found in the appendix (see Figures 8- 15). Moreover, in both setups a) and b), we applied the Kolmogorov-Smirnov test of

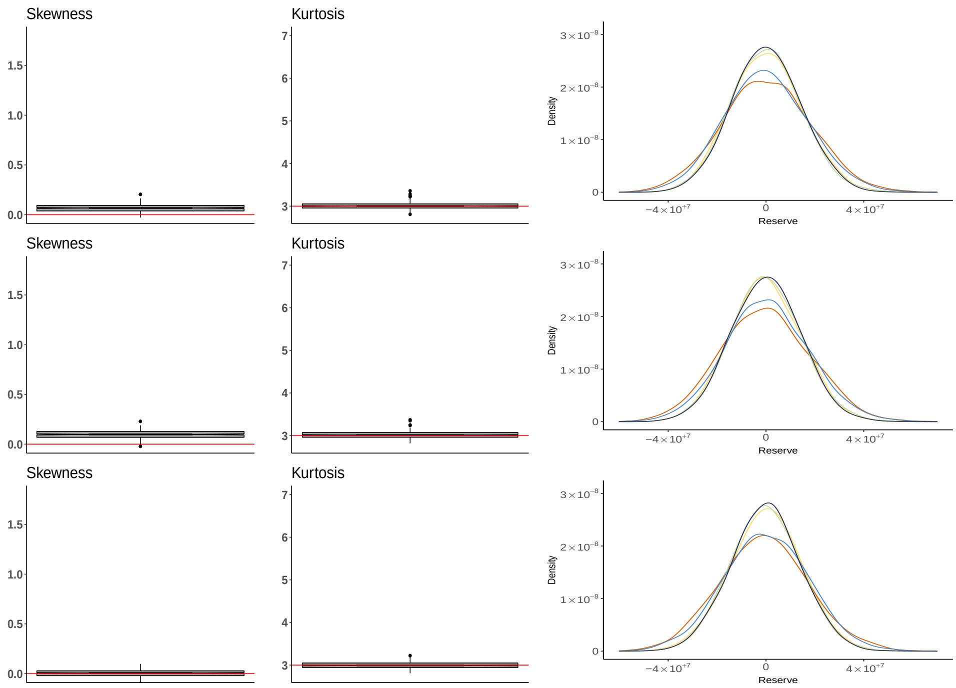

Figure 4: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 10$ and $I = 10$ for the setup of a), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

level $\alpha = 5\%$. For setup a), it fails to reject the null hypothesis of a Gaussian distribution for about 92% out of $M = 500$ samples, if the gamma distribution is used, for about 87% in the case of a log-normal, and for about 94% for a truncated normal distribution. The picture is essentially the same for all $n \in \{0, 10, 20, 30, 40\}$. In comparison, for setup b), the test does always reject the null for the gamma and for log-normal distribution, but only in about 13% out of $M = 500$ samples for the truncated normal distribution. Again the results are pretty similar for all $n \in \{0, 10, 20, 30, 40\}$.

These findings from Figures 4 and 5 can be explained by a property of the gamma and the log-normal distribution. Both tend to 'lose' their skewness and excess of kurtosis for $\frac{C_{i,j}}{\sigma_j^2}$ growing large in this parameter setting. Hence, as the range for the entries of the first column in setup a) is $[120 \times 10^6, 350 \times 10^6]$ with $[120 \times 10^4, 350 \times 10^4]$ for setup b), we observe more skewness and more excess kurtosis in b) in comparison to a). In particular, this demonstrates that the distribution of the (asymptotically dominating) $(R_{I,n} - \widehat{R}_{I,n})_1$, given $\mathcal{Q}_{I,n}$, generally does depend on the distribution (family) of the individual development factors also for large (effective) number of accident years $I + n + 1$.

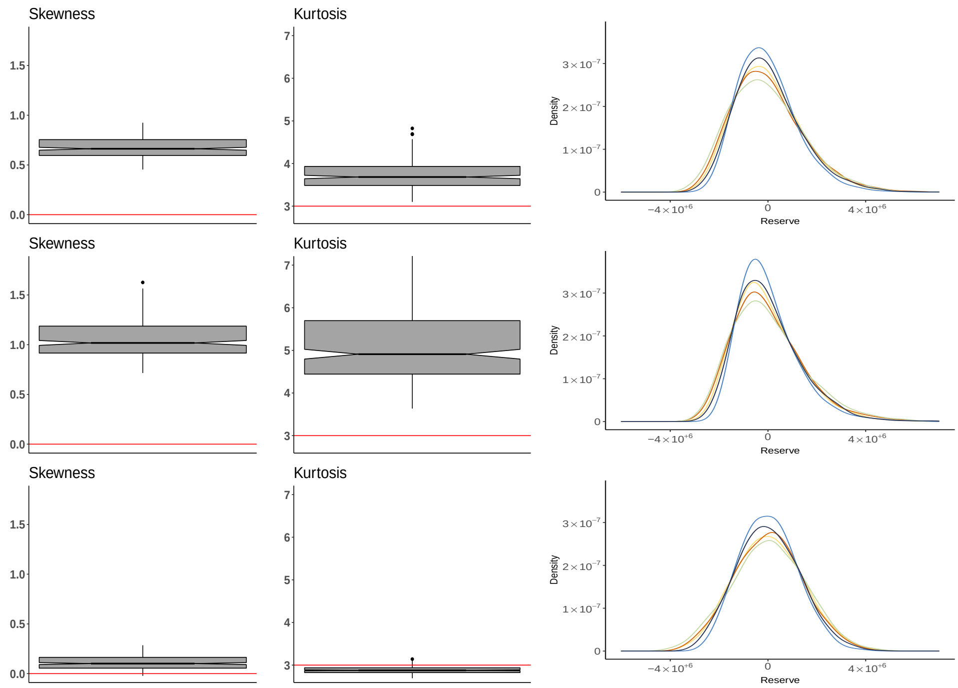

Figure 5: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 10$ and $I = 10$ for the setup of b), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

2.5.2 Simulations for $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$

To simulate the (unconditional) distribution of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$, we consider again the two scenarios a) and b) as explained in Section 2.5.1 and proceed as follows. For $I = 10$ and $n \in \{0, 10, 20, 30, 40\}$, we simulate $M = 5000$ upper loss triangles for the three cases of a conditional gamma, log-normal and truncated normal distribution, respectively. For each case, we calculate $\widehat{f}_j$ for $j = 0, \ldots, I-1+n$. Next, for all three distributions, we approximate the (finite sample) distribution of $\sqrt{I+n+1}(R_{I,n} - R_{I,n})_2$ by the empirical distribution of $\sqrt{I+n+1}(R_{I,n} - R_{I,n})_2^{(m)}$, $m = 1, \ldots, 5000$. In Figure 6 and 7, we show the simulated distributions for $I = 10$ and $n \in \{0, 10, 20, 30, 40\}$ and for scenario a) and b), respectively. Again, for level $\alpha = 5\%$, we use the Kolmogorov-Smirnov test for normality with mean zero and variance derived in (2.43). For both scenarios a) and b), for all n and for all distributions, the test does always reject the null hypothesis.

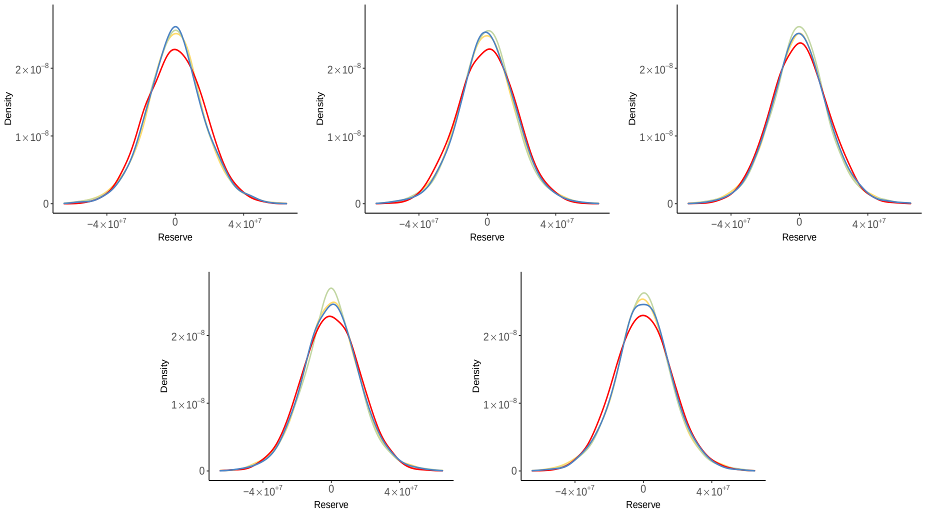

Figure 6: Density plots for simulated unconditional distributions of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ given the upper loss triangles in the setup of a) $I = 10$ and $n \in \{0, 10, 20, 30, 40\}$ (from left to right) for (conditional) gamma (yellow), log-normal (green) and truncated normal (blue), normal approximation (red).

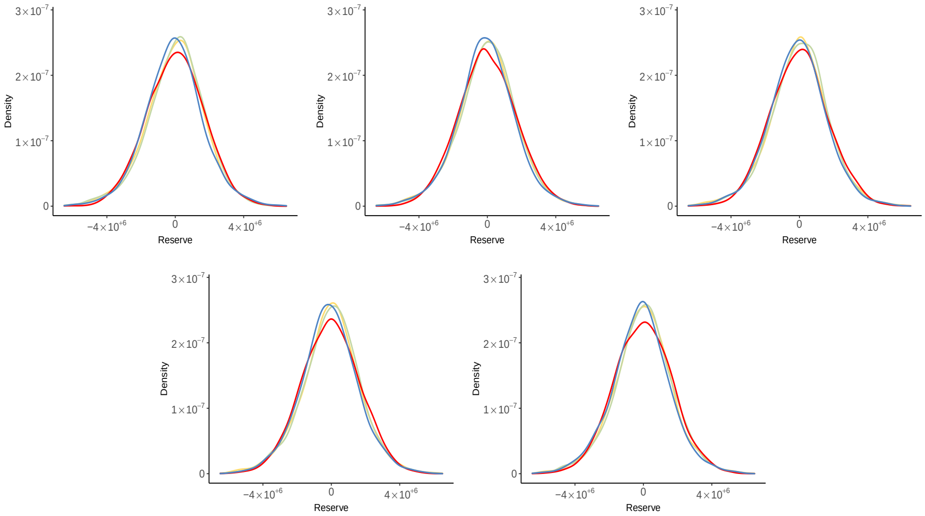

Figure 7: Density plots for simulated (unconditional) distributions of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ given the upper loss triangles in the setup of b) $I = 10$ and $n \in \{0, 10, 20, 30, 40\}$ (from left to right) for (conditional) gamma (yellow), log-normal (green) and truncated normal (blue), normal approximation (red).

2.6 Conclusion

We propose a general and fully described stochastic framework which allows to derive conditional and unconditional asymptotic theory for Mack's model. First, for increasing number of accident years, we establish unconditional and conditional versions of central limit theorems for (smooth functions of) the parameter estimators in Mack's model, which allows for asymptotic inference for the development factors and variance parameters. Moreover, these results enable us to derive also unconditional and conditional limiting distributions for the predictive root of the reserve. For this purpose, the reserve risk is split into two random parts that carry the process uncertainty and the estimation uncertainty, respectively.

It turns out that, unconditionally, but also when conditioning on either the whole observed loss triangle or on its diagonal, the limiting distribution of the process uncertainty part will be usually non-Gaussian. When properly inflated, the estimation uncertainty part is measurable with respect to the loss triangle and, unconditionally, turns out to asymptotically non-Gaussian as well. By contrast, when conditioning only on the diagonal, this results in a Gaussian limit for the estimation uncertainty part. Altogether, as the process uncertainty part dominates asymptotically, and in contrast to common practice, this leads overall to a non-Gaussian limiting distribution for the reserve in both cases.

Our findings are illustrated by simulations, where we demonstrate, according to our established theory, that the limiting distribution of the reserve might deviate substantially from a Gaussian distribution. Also we show that the shape of the limiting distribution does depend on the specific parameter setting in Mack's model and on the conditional distribution of the individual development factors.

Appendix

2.7 Proofs of Section 2.2

2.7.1 Proof of Lemma 2.4

From Assumption 2.2, we know $E(C_{i,0}) = \mu_0$ and $Var(C_{i,0}) = \tau_0^2$ for all $i = 0, \dots, I$. Hence, let $i \in \{0, \dots, I\}$ and $j \in \{1, \dots, I\}$. By the law of iterated expectation and using (2.17), (2.18) and (2.19), we get immediately

$$E(C_{i,j}) = E(E(F_{i,j-1}C_{i,j-1}|C_{i,j-1})) = E(C_{i,j-1}E(F_{i,j-1}|C_{i,j-1})) = f_{j-1}E(C_{i,j-1}). \quad (2.56)$$

Applying (2.56) recursively and using Assumption 2.2 and 2.3, we get

$$E(C_{i,j}) = \mu_0 \prod_{k=0}^{j-1} f_k = \mu_j. \quad (2.57)$$

Similarly, for the second moment of $C_{i,j}$, we get

$$\begin{aligned} E(C_{i,j}^2) &= E(E(F_{i,j-1}^2 C_{i,j-1}^2|C_{i,j-1})) = E(C_{i,j-1}^2 E(F_{i,j-1}^2|C_{i,j-1})) \\ &= E\left(C_{i,j-1}^2\left(Var(F_{i,j-1}|C_{i,j-1}) + (E(F_{i,j-1}|C_{i,j-1}))^2\right)\right) \\ &= E\left(C_{i,j-1}^2\left(\frac{\sigma_{j-1}^2}{C_{i,j-1}} + f_{j-1}^2\right)\right) = \sigma_{j-1}^2\mu_{j-1} + f_{j-1}^2 E\left(C_{i,j-1}^2\right). \end{aligned} \quad (2.58)$$

Recursively plugging-in (2.58) and using $E(C_{i,0}^2) = \tau_0^2 + \mu_0^2$, leads to

$$E(C_{i,j}^2) = (\tau_0^2 + \mu_0^2)\prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1}\left(\prod_{m=0}^{l-1} f_m\right)\sigma_l^2\left(\prod_{n=l+1}^{j-1} f_n^2\right). \quad (2.59)$$

Together with (2.57), this leads to

$$Var(C_{i,j}) = (\tau_0^2 + \mu_0^2)\prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1}\left(\prod_{m=0}^{l-1} f_m\right)\sigma_l^2\left(\prod_{n=l+1}^{j-1} f_n^2\right) - \left(\mu_0 \prod_{k=0}^{j-1} f_k\right)^2 \quad (2.60)$$

$$= \tau_0^2 \prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1}\left(\prod_{m=0}^{l-1} f_m\right)\sigma_l^2\left(\prod_{n=l+1}^{j-1} f_n^2\right) = \tau_j^2. \quad (2.61)$$

Now, to derive the formula for the covariance $Cov(C_{i,j_1}, C_{i,j_2})$, we consider the mixed moment $E(C_{i,j_1}C_{i,j_2})$ and we can assume wlog that $j_1 \geq j_2$ holds. Then, we get

$$\begin{aligned}
E(C_{i,j_1}C_{i,j_2}) &= E(E(F_{i,j_1-1}C_{i,j_1-1}C_{i,j_2}|C_{i,j_1-1},\ldots,C_{i,j_2})) \\
&= E(C_{i,j_1-1}C_{i,j_2}E(F_{i,j_1-1}|C_{i,j_1-1},\ldots,C_{i,j_2})) \\
&= E(C_{i,j_1-1}C_{i,j_2}E(F_{i,j_1-1}|C_{i,j_1-1})) \\
&= f_{j_1-1}E(C_{i,j_1-1}C_{i,j_2}) \\
&= \left(\prod_{k=j_2}^{j_1-1} f_k\right) E(C_{i,j_2}^2) \\
&= \left(\prod_{k=j_2}^{j_1-1} f_k\right) \left(\tau_{j_2}^2 + \mu_{j_2}^2\right).
\end{aligned}$$

Together with $(\prod_{k=j_2}^{j_1-1} f_k)\mu_{j_2} = \mu_{j_1}$ by construction, we get

$$Cov(C_{i,j_1}, C_{i,j_2}) = \left(\prod_{k=j_2}^{j_1-1} f_k\right) \tau_{j_2}^2 \tag{2.62}$$

for $j_1 \geq j_2$ and, in general,

$$Cov(C_{i,j_1}, C_{i,j_2}) = \left(\prod_{k=\min(j_1,j_2)}^{\max(j_1,j_2)-1} f_k\right) \tau_{\min(j_1,j_2)}^2. \tag{2.63}$$

As C_{i_1,j_1} and C_{i_2,j_2} are stochastically independent whenever $i_1 \neq i_2$ holds, we get immediately $Cov(C_{i_1,j_1}, C_{i_2,j_2}) = 0$, $i_1 \neq i_2$. □

2.8 Proofs of Section 2.3

2.8.1 Proof of Theorem 2.5

As part (i) is contained as a special case, we have to prove part (ii). Let $K \in \mathbb{N}_0$ be fixed, that is, we consider $\underline{\widehat{f}}_K = (\widehat{f}_0, \widehat{f}_1, \ldots, \widehat{f}_K)'$ as an estimator for $\underline{f}_K = (f_0, f_1, \ldots, f_K)'$, where

$$\widehat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}. \tag{2.64}$$

Note that the sums in numerator and denominator of $\widehat{f}_j$ have different length depending on j. Hence, it is convenient to approximate $\widehat{f}_j$ by $\widetilde{f}_{j,K}$, where

$$\widetilde{f}_{j,K} = \frac{\sum_{i=0}^{I-K-1} C_{i,j+1}}{\sum_{i=0}^{I-K-1} C_{i,j}} = \frac{\frac{1}{I-K}\sum_{i=0}^{I-K-1} C_{i,j+1}}{\frac{1}{I-K}\sum_{i=0}^{I-K-1} C_{i,j}}. \tag{2.65}$$

Due to $\sqrt{I-j}/\sqrt{I-K} \to 1$ as $I \to \infty$, because $j \in \{0,\ldots,K\}$ with $K \in \mathbb{N}_0$ fixed, and as

$$\begin{aligned}
\widetilde{f}_{j,K} - \widehat{f}_j &= \frac{\sum_{i=0}^{I-K-1} C_{i,j+1}}{\sum_{i=0}^{I-K-1} C_{i,j}} - \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \\
&= \frac{\sum_{i=0}^{I-K-1} C_{i,j+1} - \sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-K-1} C_{i,j}} + \left(\sum_{i=0}^{I-j-1} C_{i,j+1}\right) \frac{\sum_{i=0}^{I-K-1} C_{i,j} - \sum_{i=0}^{I-j-1} C_{i,j}}{\left(\sum_{i=0}^{I-K-1} C_{i,j}\right)\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)} \\
&= \frac{\frac{1}{I-K}\sum_{i=I-j}^{I-K-1} C_{i,j+1}}{\frac{1}{I-K}\sum_{i=0}^{I-K-1} C_{i,j}} + \left(\frac{1}{I-j}\sum_{i=0}^{I-j-1} C_{i,j+1}\right) \frac{\frac{1}{I-K}\sum_{i=I-j}^{I-K-1} C_{i,j}}{\left(\frac{1}{I-K}\sum_{i=0}^{I-K-1} C_{i,j}\right)\left(\frac{1}{I-j}\sum_{i=0}^{I-j-1} C_{i,j}\right)} \\
&= O_P\left(\frac{1}{I}\right),
\end{aligned}$$

we get

$$J^{1/2}\left(\underline{\widehat{f}}_K - \underline{f}_K\right) = \sqrt{I-K}\left(\underline{\widetilde{f}}_K - \underline{f}_K\right) + O_P\left(\frac{1}{\sqrt{I}}\right), \tag{2.66}$$

where $\underline{\widetilde{f}}_K = (\widetilde{f}_{0,K}, \widetilde{f}_{1,K}, \ldots, \widetilde{f}_{K,K})'$.
It remains to prove asymptotic normality of $\sqrt{I-K}\left(\underline{\widetilde{f}}_K - \underline{f}_K\right)$. For this purpose, we apply the delta method and define the function $g : \mathbb{R}_+^{K+2} \to \mathbb{R}_+^{K+1}$, where $\mathbb{R}_+ = \{x \in \mathbb{R} | x > 0\}$, by $g(\underline{x}) = (g_0(\underline{x}), \ldots, g_K(\underline{x}))$, $\underline{x} = (x_0, \ldots, x_{K+1})' \in \mathbb{R}_+^{K+2}$ and

$$g_j(\underline{x}) = \frac{x_{j+1}}{x_j}, \quad j = 0, \ldots, K. \tag{2.67}$$

Then, we have $\underline{\widetilde{f}}_K = g(\underline{\overline{C}}_K)$ and $\underline{f}_K = g(\underline{\mu}_K)$, where

$$\underline{\overline{C}}_K = \frac{1}{I-K}\sum_{i=0}^{I-K-1} \begin{pmatrix} C_{i,0} \\ \vdots \\ C_{i,K+1} \end{pmatrix} \quad \text{and} \quad \underline{\mu}_K = \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_{K+1} \end{pmatrix}.$$

Recall that, by Assumptions 2.2 and 2.3, the vectors $\underline{C}_{i,K} = (C_{i,0}, \ldots, C_{i,K+1})'$, $i = 0, \ldots, I$ are i.i.d. with mean vector $\underline{\mu}_K$ and variance-covariance matrix

$$\mathbf{\Sigma}_{K,\underline{C}} = Cov(\underline{C}_{i,K}) = \begin{pmatrix} \left(\prod_{k=\min(j_1,j_2)}^{\max(j_1,j_2)-1} f_k\right) \tau^2_{\min(j_1,j_2)} \\ j_1, j_2 = 0, \ldots, K+1 \end{pmatrix}. \tag{2.68}$$

Hence, by a direct application of the Lindeberg-Lévy CLT, we get

$$\sqrt{I-K}\left(\overline{C}_K - \underline{\mu}_K\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma}_{K,\underline{C}}). \tag{2.69}$$

Now, to prove asymptotic normality also of

$$\sqrt{I-K}\left(\underline{\widetilde{f}}_K - \underline{f}_K\right) = \sqrt{I-K}\left(g(\overline{C}_K) - g(\underline{\mu}_K)\right), \tag{2.70}$$

we apply the delta method. Using

$$\frac{\partial g_j(\underline{x})}{\partial x_i} = \begin{cases} -\frac{x_{j+1}}{x_j^2}, & i = j \\ \frac{1}{x_j}, & i = j+1 \\ 0, & \text{otherwise} \end{cases} \tag{2.71}$$

for $i = 0, \ldots, K+1$ and $j = 0, \ldots, K$, the $(K+1) \times (K+2)$ Jacobian $J_g(\underline{x})$ of g becomes

$$J_g(\underline{x}) = \begin{pmatrix} -\frac{x_1}{x_0^2} & \frac{1}{x_0} & 0 & \cdots & 0 \\ 0 & -\frac{x_2}{x_1^2} & \frac{1}{x_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{x_{K+1}}{x_K^2} & \frac{1}{x_K} \end{pmatrix}. \tag{2.72}$$

Finally, using $\frac{\mu_{j+1}}{\mu_j} = f_j$, we get

$$\left(J_g(\underline{\mu}_K)\mathbf{\Sigma}_{K,\underline{C}} J_g(\underline{\mu}_K)'\right)_{j,j} = \left(-\frac{f_j}{\mu_j}, \frac{1}{\mu_j}\right) \begin{pmatrix} \tau_j^2 & f_j \tau_j^2 \\ f_j \tau_j^2 & \tau_{j+1}^2 \end{pmatrix} \left(-\frac{f_j}{\mu_j}, \frac{1}{\mu_j}\right)' = \frac{-\tau_j^2 f_j^2}{\mu_j^2} + \frac{\tau_{j+1}^2}{\mu_j^2} = \frac{\sigma_j^2}{\mu_j}$$

for all $j = 0, \ldots, K$. Similarly, for all $j_1, j_2 \in \{0, \ldots, K\}$ with $j_1 > j_2$, we get

$$\begin{aligned} &\left(J_g(\underline{\mu}_K)\mathbf{\Sigma}_{K,\underline{C}} J_g(\underline{\mu}_K)'\right)_{j_1,j_2} \\ &= \left(-\frac{f_{j_1}}{\mu_{j_1}}, \frac{1}{\mu_{j_1}}\right) \begin{pmatrix} \left(\prod_{k=j_2}^{j_1-1} f_k\right)\tau_{j_2}^2 & \left(\prod_{k=j_2+1}^{j_1-1} f_k\right)\tau_{j_2+1}^2 \\ \left(\prod_{k=j_2}^{j_1} f_k\right)\tau_{j_2}^2 & \left(\prod_{k=j_2+1}^{j_1} f_k\right)\tau_{j_2+1}^2 \end{pmatrix} \left(-\frac{f_{j_2}}{\mu_{j_2}}, \frac{1}{\mu_{j_2}}\right)' = 0. \end{aligned}$$

Altogether, this completes the proof of

$$J^{1/2}\left(\underline{\widehat{f}}_K - \underline{f}_K\right) \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\underline{f}}\right).$$

□

2.8.2 Proof of Corollary 2.6

As part (i) is contained as a special case, we have to prove part (ii). We define the function $h : \mathbb{R}_+^{K+1} \to \mathbb{R}_+^{K+1}$ with $h(\underline{x}) = (h_0(\underline{x}), \ldots, h_K(\underline{x}))$ for $\underline{x} = (x_0, \ldots, x_{K+1})' \in \mathbb{R}_+^{K+2}$ by

$$h_i(\underline{x}) = \prod_{l=i}^{K} x_l \quad \text{with} \quad \frac{\partial h_i(\underline{x})}{\partial x_k} = \begin{cases} 0, & k < i \\ \prod_{l=i,l\neq k}^{K} x_l, & k \geq i \end{cases} \tag{2.73}$$

for $k, i = 0, \ldots, K$. Hence, the $(K+1) \times (K+1)$ Jacobian $J_h(\underline{x})$ of h becomes upper triangular and we have

$$J_h(\underline{x}) = \begin{pmatrix} \prod_{l=0,l\neq 0}^{K} x_l & \prod_{l=0,l\neq 1}^{K} x_l & \cdots & \prod_{l=0,l\neq K}^{K} x_l \\ 0 & \prod_{l=1,l\neq 1}^{K} x_l & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \prod_{l=K,l\neq K}^{K} x_l \end{pmatrix}. \tag{2.74}$$

Finally, using the delta method, we get

$$\left(J_h(\underline{f}_K)\mathbf{\Sigma}_{K,\underline{f}} J_h(\underline{f}_K)'\right)_{i,i} = \sum_{j=i}^{K} \left(\prod_{l=i,l\neq j}^{K} f_l\right) \frac{\sigma_j^2}{\mu_j} \left(\prod_{l=i,l\neq j}^{K} f_l\right) = \sum_{j=i}^{K} \frac{\sigma_j^2}{\mu_j} \prod_{l=i,l\neq j}^{K} f_l^2$$

for all $i = 0, \ldots, K$. Similarly, for all $i_1, i_2 \in \{0, \ldots, K\}$ with $i_1 > i_2$, we get

$$\begin{aligned}
\left(J_h(\underline{f}_K)\mathbf{\Sigma}_{K,\underline{f}} J_h(\underline{f}_K)'\right)_{i_1,i_2} &= \sum_{j=i_1}^{K} \left(\prod_{l=i_1,l\neq j}^{K} f_l\right) \frac{\sigma_j^2}{\mu_j} \left(\prod_{m=i_2,m\neq j}^{K} f_m\right) \\
&= \sum_{j=i_1}^{K} \left(\prod_{l=i_1,l\neq j}^{K} f_l\right) \frac{\sigma_j^2}{\mu_j} \left(\prod_{m=i_2}^{i_1-1} f_m\right) \left(\prod_{m=i_1,m\neq j}^{K} f_m\right) \\
&= \sum_{j=i_1}^{K} \frac{\sigma_j^2}{\mu_j} \left(\prod_{l=i_1,l\neq j}^{K} f_l^2\right) \left(\prod_{m=i_2}^{i_1-1} f_m\right).
\end{aligned}$$

Together, this completes the proof of

$$\sqrt{I} \begin{pmatrix} \prod_{j=i}^{K} \widehat{f}_j - \prod_{j=i}^{K} f_j \\ i = 0, \ldots, K \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\prod f_j}\right),$$

since $\sqrt{I-K}/\sqrt{I} \to 1$ as $I \to \infty$ and $K \in \mathbb{N}_0$ fixed. □

2.8.3 Proof of Theorem 2.9

For any fixed j, the estimator $\widehat{\sigma}_j^2$ can be represented as follows. Precisely, by plugging-in for $\widehat{f}_j$ and due to $F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$, we get

$$\begin{aligned}\widehat{\sigma}_j^2 &= \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}(F_{i,j}-\widehat{f}_j)^2 \\ &= \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}F_{i,j}^2 - \frac{\left(\frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j+1}\right)^2}{\frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}}.\end{aligned}$$

Hence, by defining the function $v : \mathbb{R}_+^3 \to \mathbb{R}_+$ by $v(a,b,c) = a - \frac{c^2}{b}$, we have

$$\widehat{\sigma}_j^2 = v\left(\frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}F_{i,j}^2, \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}, \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j+1}\right). \tag{2.75}$$

As v is differentiable on $\mathbb{R}_+^3$, we can calculate the first partial derivatives. Using

$$\frac{\partial v(a,b,c)}{\partial a} = 1, \quad \frac{\partial v(a,b,c)}{\partial b} = \frac{c^2}{b^2} \quad \text{and} \quad \frac{\partial v(a,b,c)}{\partial c} = -2\frac{c}{b}, \tag{2.76}$$

its Jacobian $J_v(a,b,c)$ becomes

$$J_v(a,b,c) = \left(1, \frac{c^2}{b^2}, -2\frac{c}{b}\right). \tag{2.77}$$

Hence, making use of the delta method, it remains to prove a CLT for

$$\underline{\overline{B}}_j = \frac{1}{I-j-1}\sum_{i=0}^{I-j-1}\begin{pmatrix} C_{i,j}F_{i,j}^2 \\ C_{i,j} \\ C_{i,j+1}\end{pmatrix} = \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} \underline{B}_{i,j}.$$

For this purpose, note that, for each fixed j, the random vectors $\underline{B}_{i,j}$, $i = 0, \ldots, I-j-1$ are i.i.d. with mean vector $\underline{\mu}_B := E((C_{i,j}F_{i,j}^2, C_{i,j}, C_{i,j+1})') = (\mu_j f_j^2 + \sigma_j^2, \mu_j, \mu_{j+1})'$ due to Lemma 2.4 and

$$E(C_{i,j}F_{i,j}^2) = E(E(C_{i,j}F_{i,j}^2|C_{i,j})) = E(C_{i,j}E(F_{i,j}^2|C_{i,j})) = E\left(C_{i,j}\left(\frac{\sigma_j^2}{C_{i,j}} + f_j^2\right)\right) \tag{2.78}$$

$$= \sigma_j^2 + E\left(C_{i,j}\right) f_j^2 = \mu_j f_j^2 + \sigma_j^2. \tag{2.79}$$

Its variance-covariance matrix becomes

$$\mathbf{\Sigma}_B = Cov(\underline{B}_{i,j}) = \begin{pmatrix} Var(C_{i,j}F_{i,j}^2) & Cov(C_{i,j}F_{i,j}^2, C_{i,j}) & Cov(C_{i,j}F_{i,j}^2, C_{i,j+1}) \\ Cov(C_{i,j}, C_{i,j}F_{i,j}^2) & \tau_j^2 & f_j\tau_j^2 \\ Cov(C_{i,j+1}, C_{i,j}F_{i,j}^2) & f_j\tau_j^2 & \tau_{j+1}^2 \end{pmatrix}. \tag{2.80}$$

To calculate $Var(C_{i,j}F_{i,j}^2)$, using (2.24) and (2.26), we get

$$\begin{aligned} E(C_{i,j}^2F_{i,j}^4) &= E(E(C_{i,j}^2F_{i,j}^4|C_{i,j})) = E(C_{i,j}^2E(F_{i,j}^4|C_{i,j})) \\ &= E\left(C_{i,j}^2\left(f_j^4 + 6\frac{\sigma_j^2}{C_{i,j}} + 4f_jE((F_{i,j}-f_j)^3|C_{i,j}) + E((F_{i,j}-f_j)^4|C_{i,j})\right)\right) \\ &= (\tau_j^2+\mu_j^2)f_j^4 + 6\mu_jf_j^2\sigma_j^2 + 4f_j\kappa_j^{(3)} + \kappa_j^{(4)}, \end{aligned}$$

which, together with $E(C_{i,j}F_{i,j}^2) = \mu_jf_j^2 + \sigma_j^2$, leads to

$$\begin{aligned} Var(C_{i,j}F_{i,j}^2) &= (\tau_j^2+\mu_j^2)f_j^4 + 6\mu_jf_j^2\sigma_j^2 + 4f_j\kappa_j^{(3)} + \kappa_j^{(4)} - (\mu_jf_j^2+\sigma_j^2)^2 \\ &= \tau_j^2f_j^4 + 4\mu_jf_j^2\sigma_j^2 + 4f_j\kappa_j^{(3)} + \kappa_j^{(4)} - \sigma_j^4. \end{aligned}$$

Similarly, we have

$$E(C_{i,j}^2F_{i,j}^2) = E(E(C_{i,j}^2F_{i,j}^2|C_{i,j})) = E\left(C_{i,j}^2\left(\frac{\sigma_j^2}{C_{i,j}} + f_j^2\right)\right) \tag{2.81}$$

$$= E\left(C_{i,j}\right)\sigma_j^2 + E\left(C_{i,j}^2\right)f_j^2 = \mu_j\sigma_j^2 + \left(\tau_j^2+\mu_j^2\right)f_j^2 \tag{2.82}$$

leading to

$$Cov(C_{i,j}F_{i,j}^2, C_{i,j}) = \mu_j\sigma_j^2 + \left(\tau_j^2+\mu_j^2\right)f_j^2 - (\mu_jf_j^2+\sigma_j^2)\mu_j = \tau_j^2f_j^2. \tag{2.83}$$

Further, using $C_{i,j+1} = C_{i,j}F_{i,j}$, (2.24) and (2.25), we get

$$\begin{aligned} E\left(C_{i,j}F_{i,j}^2C_{i,j+1}\right) &= E\left(C_{i,j}^2F_{i,j}^3\right) = E\left(E(C_{i,j}^2F_{i,j}^3|C_{i,j})\right) = E\left(C_{i,j}^2E(F_{i,j}^3|C_{i,j})\right) \\ &= E\left(C_{i,j}^2\left(f_j^3 + 3f_j\frac{\sigma_j^2}{C_{i,j}} + E\left((F_{i,j}-f_j)^3|C_{i,j}\right)\right)\right) \\ &= \left(\tau_j^2+\mu_j^2\right)f_j^3 + 3\mu_jf_j\sigma_j^2 + \kappa_j^{(3)} \end{aligned}$$

leading to

$$\begin{aligned} Cov(C_{i,j}F_{i,j}^2, C_{i,j+1}) &= (\tau_j^2+\mu_j^2)f_j^3 + 3\mu_jf_j\sigma_j^2 + \kappa_j^{(3)} - (\mu_jf_j^2+\sigma_j^2)\mu_{j+1} \\ &= \tau_j^2f_j^3 + 2\mu_jf_j\sigma_j^2 + \kappa_j^{(3)}. \end{aligned}$$

By an application of the Lindeberg-Lévy CLT, we get

$$\sqrt{I-j}\left(\underline{\overline{B}}_j - \underline{\mu}_B\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma}_B) \tag{2.84}$$

and by the delta method, using $J_v(\underline{\mu}_B) = (1, \frac{\mu_{j+1}^2}{\mu_j^2}, -\frac{2\mu_{j+1}}{\mu_j}) = (1, f_j^2, -2f_j)$, this leads to

$$\sqrt{I-j}\left(\widehat{\sigma}_j^2 - \sigma_j^2\right) = \sqrt{I-j}\left(v(\underline{\overline{B}}_j) - v(\underline{\mu}_B)\right) \xrightarrow{d} \mathcal{N}(0, J_v(\underline{\mu}_B)\mathbf{\Sigma}_B J_v(\underline{\mu}_B)'). \tag{2.85}$$

Finally, for the limiting variance, by direct calculation, we get

$$\begin{aligned}
&J_v(\underline{\mu}_B)\mathbf{\Sigma}_B J_v(\underline{\mu}_B)' \\
&= (1, f_j^2, -2f_j)\begin{pmatrix} Var(C_{i,j}F_{i,j}^2) & Cov(C_{i,j}F_{i,j}^2, C_{i,j}) & Cov(C_{i,j}F_{i,j}^2, C_{i,j+1}) \\ Cov(C_{i,j}, C_{i,j}F_{i,j}^2) & \tau_j^2 & f_j\tau_j^2 \\ Cov(C_{i,j+1}, C_{i,j}F_{i,j}^2) & f_j\tau_j^2 & \tau_{j+1}^2 \end{pmatrix}\begin{pmatrix} 1 \\ f_j^2 \\ -2f_j \end{pmatrix} \\
&= \Psi_j.
\end{aligned}$$

□

2.9 Conditional versions of the CLTs from Section 2.3

In the following, we establish conditional versions of the CLTs from Theorem 2.5 under the asymptotic framework of Section 2.4.

Theorem 2.24 (Asymptotic normality of $\widehat{f}_j$ conditionally on $\mathcal{Q}_{I,\infty}$) *Suppose Assumptions 2.2, 2.3, 2.11, 2.16 and 2.18 are satisfied. Then, as $n \to \infty$, the following holds:*

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

$$\sqrt{I+n-j}\,(f_{j,n}(\mathcal{Q}_{I,\infty}) - f_j)) \xrightarrow{d} \mathcal{N}\left(0, \sigma_{f_j,1}^2\right),$$

where $f_{j,n}(\mathcal{Q}_{I,\infty})$ is defined in Section 2.4.4.2 and the variance is

$$\sigma_{f_j,1}^2 = \frac{f_j^2 E\left(E(C_{i,j}|C_{i,\infty})^2\right) - 2f_j E\left(E(C_{i,j}|C_{i,\infty})E(C_{i,j+1}|C_{i,\infty})\right) + E\left(E(C_{i,j+1}|C_{i,\infty})^2\right)}{\mu_j^2}.$$

(ii) For each fixed $K \in \mathbb{N}_0$, let $\underline{f}_K = (f_0, f_1, \ldots, f_K)'$. Then, unconditionally, we have

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

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

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

where $\mathbf{\Sigma}^{(1)}_{K,\underline{C}}$ *is defined in* (2.94), *has entries*

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

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

(iii) For each fixed $j\in\mathbb{N}_0=\{0,1,2,\ldots\}$, *conditionally on* $\mathcal{Q}_{I,\infty}$, *we have*

$$\sqrt{I+n-j}\left(\widehat{f}_{j,n}-f_{j,n}(\mathcal{Q}_{I,\infty})\right)|\mathcal{Q}_{I,\infty}\xrightarrow{d}\mathcal{N}\left(0,\sigma^2_{f_j,2}\right),$$

where the variance is $\sigma^2_{f_j,2}=\frac{\sigma_j^2}{\mu_j}-\sigma^2_{f_j,1}$.

(iv) For each fixed $K\in\mathbb{N}_0$, *let* $\widehat{\underline{f}}_{K,n}=(\widehat{f}_{0,n},\widehat{f}_{1,n},\ldots,\widehat{f}_{K,n})'$ *and define* $\underline{f}_{K,n}(\mathcal{Q}_{I,\infty}) = (f_{0,n}(\mathcal{Q}_{I,\infty}), f_{1,n}(\mathcal{Q}_{I,\infty}),\ldots,f_{K,n}(\mathcal{Q}_{I,\infty}))'$. *Then, conditionally on* $\mathcal{Q}_{I,\infty}$, *we have*

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

where the variance-covariance matrix

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

where $\mathbf{\Sigma}^{(2)}_{K,\underline{C}}$ *is defined in* (2.89), *has entries* $\mathbf{\Sigma}^{(2)}_{K,\underline{f}}(j,j)=\sigma^2_{f_j,2}=\frac{\sigma_j^2}{\mu_j}-\sigma^2_{f_j,1}$ *for* $j=0,\ldots,K$ *and* $\mathbf{\Sigma}^{(2)}_{K,\underline{f}}(j_1,j_2)=-\mathbf{\Sigma}^{(1)}_{K,\underline{f}}(j_1,j_2)$ *for* $j_1,j_2=0,\ldots,K$, $j_1\neq j_2$.

2.9.1 Proof of Theorem 2.24

As in the proof of Theorem 2.5, (i) and (iii) are contained as special cases, and we have to prove parts (ii) and (iv). In contrast to Theorem 2.5, we have to prove unconditional and conditional CLTs under the asymptotic framework of Section 2.4.

Similar to the proof of Theorem 2.5, for fixed $K \in \mathbb{N}_0$, we consider $\underline{\widehat{f}}_{K,n} = (\widehat{f}_{0,n}, \widehat{f}_{1,n}, \ldots, \widehat{f}_{K,n})'$ and $\underline{\widetilde{f}}_{K,n} = (\widetilde{f}_{0,K,n}, \widetilde{f}_{1,K,n}, \ldots, \widetilde{f}_{K,K,n})'$, where $\widetilde{f}_{j,K,n}$ approximates $\widehat{f}_{j,n}$ with

$$\widehat{f}_{j,n} = \frac{\sum_{i=-n}^{I-j-1} C_{i,j+1}}{\sum_{i=-n}^{I-j-1} C_{i,j}} \quad \text{and} \quad \widetilde{f}_{j,K,n} = \frac{\sum_{i=-n}^{I-K-1} C_{i,j+1}}{\sum_{i=-n}^{I-K-1} C_{i,j}} = \frac{\frac{1}{I-K}\sum_{i=-n}^{I-K-1} C_{i,j+1}}{\frac{1}{I-K}\sum_{i=-n}^{I-K-1} C_{i,j}}. \tag{2.86}$$

For their difference, we have

$$\widetilde{f}_{j,K,n} - \widehat{f}_{j,n} = \frac{\frac{1}{I+n-K}\sum_{i=I-j}^{I-K-1} C_{i,j+1}}{\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} C_{i,j}} + \frac{\left(\frac{1}{I+n-j}\sum_{i=-n}^{I+n-j-1} C_{i,j+1}\right)\left(\frac{1}{I+n-K}\sum_{i=I-j}^{I-K-1} C_{i,j}\right)}{\left(\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} C_{i,j}\right)\left(\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}\right)}.$$

Further, due to $C_{i,j} \geq 0$ for all i, j, we have that $\frac{1}{b-a+1}\sum_{i=a}^{b} C_{i,j} = O_P(1)$ unconditionally and also conditionally on $\mathcal{Q}_{I,\infty}$ for all $a \leq b$ and all j. Hence, as K and I are fixed, we get $\widetilde{f}_{j,K,n} - \widehat{f}_{j,n} = O_P\left(\frac{1}{n}\right)$ conditionally on $\mathcal{Q}_{I,\infty}$ as $n \to \infty$.

We continue with showing with part (iv). Consequently, conditionally on $\mathcal{Q}_{I,\infty}$, and due to $\sqrt{I+n-j}/\sqrt{I+n-K} \to 1$ as $n \to \infty$, it remains to show

$$\sqrt{I+n-K}\left(\underline{\widetilde{f}}_{K,n} - \underline{\widetilde{f}}_{K,n}(\mathcal{Q}_{I,\infty})\right)|\mathcal{Q}_{I,\infty} \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}^{(2)}_{K,\underline{f}}\right), \tag{2.87}$$

where $\underline{\widetilde{f}}_{K,n}(\mathcal{Q}_{I,\infty}) = (\widetilde{f}_{0,K,n}(\mathcal{Q}_{I,\infty}), \widetilde{f}_{1,K,n}(\mathcal{Q}_{I,\infty}), \ldots, \widetilde{f}_{K,K,n}(\mathcal{Q}_{I,\infty}))'$ and

$$\widetilde{f}_{j,n}(\mathcal{Q}_{I,\infty}) := \widetilde{\mu}^{(1)}_{j+1,n}(\mathcal{Q}_{I,\infty})/\widetilde{\mu}^{(2)}_{j,n}(\mathcal{Q}_{I,\infty})$$

with

$$\begin{aligned}\widetilde{\mu}^{(1)}_{j+1,n}(\mathcal{Q}_{I,\infty}) &:= E\left(\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} C_{i,j+1}|\mathcal{Q}_{I,\infty}\right) = \frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} E(C_{i,j+1}|C_{i,I-i}),\\ \widetilde{\mu}^{(2)}_{j,n}(\mathcal{Q}_{I,\infty}) &:= E\left(\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} C_{i,j}|\mathcal{Q}_{I,\infty}\right) = \frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} E(C_{i,j}|C_{i,I-i}).\end{aligned}$$

To prove (2.87), we have to apply the same delta method argument as in Theorem 2.5, and it is sufficient to show

$$\sqrt{I+n-K}\left(\underline{\overline{C}}_{K,n} - \underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty})\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma}^{(2)}_{K,\underline{C}}), \tag{2.88}$$

where

$$\underline{\overline{C}}_{K,n} = \frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\begin{pmatrix} C_{i,0} \\ \vdots \\ C_{i,K+1}\end{pmatrix} \quad \text{and} \quad \underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) = \begin{pmatrix}\widetilde{\mu}_{0,n}(\mathcal{Q}_{I,\infty}) \\ \vdots \\ \widetilde{\mu}_{K+1,n}(\mathcal{Q}_{I,\infty})\end{pmatrix}.$$

Recall that, by Assumptions 2.2 and 2.3, the vectors $\underline{C}_{i,K} = (C_{i,0}, \ldots, C_{i,K+1})'$, $i = -n, \ldots, I$ are still independent when conditioning on $\mathcal{Q}_{I,\infty}$ with conditional mean vector $\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty})$, but not identically distributed anymore. Hence, we make use of the Lyapunov CLT, to get asymptotic normality of $\sqrt{I+n-K}(\underline{\overline{C}}_{K,n} - \underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}))$. Using Assumption 2.18, the Lyapunov condition follows from

$$
\begin{aligned}
&\frac{1}{(I+n-K)^{(2+\delta)/2}} \sum_{i=-n}^{I-K} E\left(|C_{i,j} - E(C_{i,j}|\mathcal{Q}_{I,\infty})|^{2+\delta}\right) \\
=&\frac{1}{(I+n-K)^{(2+\delta)/2}} \\
&\times \sum_{i=-n}^{I-K} E\left(\left|C_{i,j} - \left(E(C_{i,j}|C_{i,j}) + \sum_{k=0}^{I-i-j-1} (E(C_{i,j}|C_{i,j+k+1}) - E(C_{i,j}|C_{i,j+k}))\right)\right|^{2+\delta}\right) \\
=&\frac{1}{(I+n-K)^{(2+\delta)/2}} \sum_{i=-n}^{I-K} E\left(\left|\sum_{k=0}^{I-i-j-1} (E(C_{i,j}|C_{i,j+k}) - E(C_{i,j}|C_{i,j+k+1}))\right|^{2+\delta}\right) \\
=&\frac{1}{(I+n-K)^{(2+\delta)/2}} \sum_{i=-n}^{I-K} \left\|\sum_{k=0}^{I-i-j-1} (E(C_{i,j}|C_{i,j+k}) - E(C_{i,j}|C_{i,j+k+1}))\right\|_{2+\delta}^{2+\delta} \\
\leq&\frac{1}{(I+n-K)^{(2+\delta)/2}} \sum_{i=-n}^{I-K} \left(\sum_{k=0}^{I-i-j-1} \|(E(C_{i,j}|C_{i,j+k}) - E(C_{i,j}|C_{i,j+k+1}))\|_{2+\delta}\right)^{2+\delta} \\
\leq&\frac{1}{(I+n-K)^{(2+\delta)/2}} \sum_{i=-n}^{I-K} \left(\sum_{k=0}^{I-i-j-1} a_k \left(E(X_i^{2+\delta})\right)^{1/(2+\delta)}\right)^{2+\delta} \\
=&\frac{1}{(I+n-K)^{(2+\delta)/2}} \sum_{i=-n}^{I-K} \left(\sum_{k=0}^{I-i-j-1} a_k\right)^{2+\delta} E(X_i^{2+\delta}) = O\left(\frac{1}{(I+n-K)^{\delta/2}}\right) = o(1)
\end{aligned}
$$

for all $j \in \{0, \ldots, K+1\}$.

And for the variance, we get

$$
\begin{aligned}
&Var\left(\sqrt{I+n-K}\left(\underline{\overline{C}}_{K,n} - \underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty})\right)|\mathcal{Q}_{I,\infty}\right) \\
=&\frac{1}{I+n-K} \sum_{i=-n}^{I-K-1} Var\left(\underline{C}_{i,K}|\mathcal{Q}_{I,\infty}\right) = \frac{1}{I+n-K} \sum_{i=-n}^{I-K-1} Var\left(\underline{C}_{i,K}|C_{i,I-i}\right).
\end{aligned}
$$

Note that the conditional variances in the last right-hand side are independent, but not identically distributed as the 'gaps' between $\underline{C}_{i,K}$ and $C_{i,I-i}$ vary with index i. However, the last right-hand can be written as

$$
\begin{aligned}
&\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\left(Var\left(\underline{C}_{i,K}|C_{i,I-i}\right)\pm Var\left(\underline{C}_{i,K}|C_{i,I-i+1}\right)\pm Var\left(\underline{C}_{i,K}|C_{i,I-i+2}\right)\pm\cdots\right)\\
=&\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}Var\left(\underline{C}_{i,K}|C_{i,\infty}\right)\\
&+\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\sum_{k=I-i}^{\infty}\left(Var\left(\underline{C}_{i,K}|C_{i,k}\right)-Var\left(\underline{C}_{i,K}|C_{i,k+1}\right)\right),
\end{aligned}
$$

where $C_{i,\infty}=C_{i,0}\prod_{k=0}^{\infty}F_{i,k}$ with $C_{i,j}\xrightarrow{p}C_{i,\infty}$. Now, using Assumption 2.18, the second term on the last right-hand side is of order $O_P\left(\frac{1}{I+n-K}\right)$ and hence asymptotically negligible, because

$$
\begin{aligned}
&E\left(\left\|\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\sum_{k=I-i}^{\infty}\left(Var\left(\underline{C}_{i,K}|C_{i,k}\right)-Var\left(\underline{C}_{i,K}|C_{i,k+1}\right)\right)\right\|\right)\\
\leq&\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\sum_{k=I-i}^{\infty}E\left(\left\|Var\left(\underline{C}_{i,K}|C_{i,k}\right)-Var\left(\underline{C}_{i,K}|C_{i,k+1}\right)\right\|\right)\\
\leq&\frac{const.}{I+n-K}\sum_{i=-n}^{I-K-1}\sum_{k=I-i}^{\infty}b_kE\left(Y_i\right)\leq const.\frac{1}{I+n-K}\left(\sum_{i=-n}^{I-K-1}\frac{1}{(I-i+1)^2}\sum_{k=I-i}^{\infty}(k+1)^2b_k\right)\\
\leq& const.\frac{1}{I+n-K}.
\end{aligned}
$$

Finally, note that the first term $\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}Var\left(\underline{C}_{i,K}|C_{i,\infty}\right)$ is a sample average over (absolutely integrable) i.i.d. random variables $(Var\left(\underline{C}_{i,K}|C_{i,\infty}\right), i\in\mathbb{Z}, i\leq I-K-1)$, which converges in probability to $E(Var\left(\underline{C}_{i,K}|C_{i,\infty}\right))$ due to the weak law of large numbers (WLLN). Hence, we get

$$
\mathbf{\Sigma}^{(2)}_{K,\underline{C}}=E(Var\left(\underline{C}_{i,K}|C_{i,\infty}\right)). \tag{2.89}
$$

Now, let us take a closer look at $\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty})$. By definition, we have

$$
\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty})=E\left(\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\underline{C}_{i,K}|\mathcal{Q}_{I,\infty}\right)=\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}E\left(\underline{C}_{i,K}|C_{i,I-i}\right),
$$

which can be written as

$$
\begin{aligned}
&\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\left(E\left(\underline{C}_{i,K}|C_{i,I-i}\right)\pm E\left(\underline{C}_{i,K}|C_{i,I-i+1}\right)\pm E\left(\underline{C}_{i,K}|C_{i,I-i+2}\right)\pm\cdots\right)\\
=&\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}E\left(\underline{C}_{i,K}|C_{i,\infty}\right)+\frac{1}{I+n-K}\sum_{i=-n}^{I-K-1}\sum_{k=I-i}^{\infty}\left(E\left(\underline{C}_{i,K}|C_{i,k}\right)-E\left(\underline{C}_{i,K}|C_{i,k+1}\right)\right),
\end{aligned}
$$

where the second term on the last right-hand side is of order $O_P\left(\frac{1}{I+n-K}\right)$ and hence asymptotically negligible (even when multiplied with $\sqrt{I+n-K}$) due to Assumption 2.18. The first term on the last right-hand is a sample average over (absolutely integrable) i.i.d. random variables $(E\left(\underline{C}_{i,K}|C_{i,\infty}\right), i \in \mathbb{Z}, i \leq I-K-1)$, which converges in probability to $E\left(E\left(\underline{C}_{i,K}|C_{i,\infty}\right)\right) = E\left(\underline{C}_{i,K}\right) = \underline{\mu}_K$ due to a WLLN. Note that the latter argument is not valid when multiplied with $\sqrt{I+n-K}$!

Now, applying the same delta method argument as in the proof of Theorem 2.5, we obtain the (conditional on $\mathcal{Q}_{I,\infty}$) limiting variance for $\sqrt{I+n-K}\left(\underline{\widetilde{f}}_{K,n} - \underline{\widetilde{f}}_{K,n}(\mathcal{Q}_{I,\infty})\right)$ as

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

where we used that $\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) \xrightarrow{p} \underline{\mu}_K$ and $J_g(\cdot)$ is defined in (2.72).

Now, let us consider part (ii). Using similar arguments, it remains to show that, unconditionally, it holds

$$\sqrt{I+n-K}\left(\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) - \underline{\mu}_K\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma}^{(1)}_{K,\underline{C}}), \tag{2.90}$$

where

$$\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) = \begin{pmatrix} \widetilde{\mu}_{0,n}(\mathcal{Q}_{I,\infty}) \\ \vdots \\ \widetilde{\mu}_{K+1,n}(\mathcal{Q}_{I,\infty}) \end{pmatrix} \quad \text{and} \quad \underline{\mu}_K = \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_{K+1} \end{pmatrix}.$$

As already shown above, we have $\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) = \frac{1}{I+n-K}\sum_{i=-n}^{I-K-1} E\left(\underline{C}_{i,K}|C_{i,\infty}\right) + O_P\left(\frac{1}{I+n-K}\right)$ leading to

$$\sqrt{I+n-K}\left(\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) - \underline{\mu}_K\right) \tag{2.91}$$

$$= \frac{1}{\sqrt{I+n-K}}\sum_{i=-n}^{I-K-1}\left(E\left(\underline{C}_{i,K}|C_{i,\infty}\right) - \underline{\mu}_K\right) + O_P\left(\frac{1}{\sqrt{I+n-K}}\right). \tag{2.92}$$

As $(E(\underline{C}_{i,K}|C_{i,\infty}), i \in \mathbb{Z}, i \leq I)$ is a sequence of (square integrable) i.i.d random variables, an application of the Lindeberg-Lévy CLT results in

$$\frac{1}{\sqrt{I+n-K}}\sum_{i=-n}^{I-K-1}\left(E\left(\underline{C}_{i,K}|C_{i,\infty}\right) - \underline{\mu}_K\right) \xrightarrow{d} \mathcal{N}(0, \mathbf{\Sigma}^{(1)}_{K,\underline{C}}), \tag{2.93}$$

where $\mathbf{\Sigma}^{(1)}_{K,\underline{C}} = Var(E(\underline{C}_{i,K}|C_{i,\infty}))$. Applying again the same delta method argument as in the proof of Theorem 2.5 used already above, we obtain the (unconditional) limiting variance for $\sqrt{I+n-K}\left(\underline{f}_{K,n}(\mathcal{Q}_{I,\infty})-\underline{f}_{K}\right)$ as

$$\mathbf{\Sigma}^{(1)}_{K,\underline{f}} = J_g\left(\underline{\mu}_K\right)\mathbf{\Sigma}^{(1)}_{K,\underline{C}} J_g\left(\underline{\mu}_K\right)'. \tag{2.94}$$

Note that, due to the law of total variance, we have

$$\mathbf{\Sigma}_{K,\underline{C}} = \mathbf{\Sigma}^{(1)}_{K,\underline{C}} + \mathbf{\Sigma}^{(2)}_{K,\underline{C}} \quad \text{and} \quad \mathbf{\Sigma}_{K,\underline{f}} = \mathbf{\Sigma}^{(1)}_{K,\underline{f}} + \mathbf{\Sigma}^{(2)}_{K,\underline{f}}, \tag{2.95}$$

where $\mathbf{\Sigma}_{K,\underline{f}} = diag\left(\frac{\sigma_0^2}{\mu_0},\frac{\sigma_1^2}{\mu_1},\ldots,\frac{\sigma_K^2}{\mu_K}\right)$. □

Furthermore, we can formulate a conditional version of Corollary 2.6.

Corollary 2.25 (Asymptotic normality for products of $\widehat{f}_j$'s conditionally on $\mathcal{Q}_{I,\infty}$) *Suppose the assumptions of Theorem 2.24 hold. Then, as $n\to\infty$, the following holds:*

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

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

where $f_{j,n}(\mathcal{Q}_{I,\infty})$ as defined in Theorem 2.24 and

$$\sigma^2_{\prod f_j,1} = \left(0,\ldots,0,\prod_{\substack{j=i\\ j\neq i}}^{K} f_j,\ \prod_{\substack{j=i\\ j\neq i+1}}^{K} f_j,\ldots,\prod_{\substack{j=i\\ j\neq K}}^{K} f_j\right)\mathbf{\Sigma}^{(1)}_{K,\underline{f}}\left(0,\ldots,0,\prod_{\substack{j=i\\ j\neq i}}^{K} f_j,\ \prod_{\substack{j=i\\ j\neq i+1}}^{K} f_j,\ldots,\prod_{\substack{j=i\\ j\neq K}}^{K} f_j\right)'.$$

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

$$\sqrt{I+n+1}\begin{pmatrix}\prod_{j=i}^{K} f_{j,n}(\mathcal{Q}_{I,\infty}) - \prod_{j=i}^{K} f_{j,n}\\ i=0,\ldots,K\end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0,\mathbf{\Sigma}^{(1)}_{K,\prod f_j}\right),$$

where $\mathbf{\Sigma}^{(1)}_{K,\prod f_j} = J_h(\underline{f}_K)\mathbf{\Sigma}^{(1)}_{K,\underline{f}} J_h(\underline{f}_K)'$ with $J_h(\cdot)$ as defined in (2.74).

(iii) For each fixed $K\in\mathbb{N}_0$ and $i=0,\ldots,K$, conditionally on $\mathcal{Q}_{I,\infty}$, we have

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

where $\sigma^2_{\prod f_j,2} = \sum_{j=i}^{K}\frac{\sigma_j^2}{\mu_j}\prod_{l=i,l\neq j}^{K} f_l^2 - \sigma^2_{\prod f_j,1}$.

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

$$\sqrt{I+n+1}\begin{pmatrix}\prod_{j=i}^K \widehat{f}_{j,n} - \prod_{j=i}^K f_{j,n}(\mathcal{Q}_{I,\infty}) \\ i=0,\ldots,K\end{pmatrix} |\mathcal{Q}_{I,\infty} \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}^{(2)}_{K,\prod f_j}\right),$$

where $\mathbf{\Sigma}^{(2)}_{K,\prod f_j} = \mathbf{\Sigma}_{K,\prod f_j} - \mathbf{\Sigma}^{(1)}_{K,\prod f_j}$, where $\mathbf{\Sigma}_{K,\prod f_j}$ is defined in Corollary 2.6(ii).

2.9.2 Proof of Corollary 2.25

As (i) and (iii) are contained in (ii) and (iv), respectively, it remains to show (ii) and (iv). As in the proof of Corollary 2.6, the same delta method argument leads to the limiting variances

$$\mathbf{\Sigma}^{(1)}_{K,\prod f_j} = J_h\left(\underline{f}_K\right)\mathbf{\Sigma}^{(1)}_{K,\underline{f}} J_h\left(\underline{f}_K\right)' \quad \text{and} \quad \mathbf{\Sigma}^{(2)}_{K,\prod f_j} = J_h\left(\underline{f}_K\right)\mathbf{\Sigma}^{(2)}_{K,\underline{f}} J_h\left(\underline{f}_K\right)'$$

for (ii) and (iv), respectively, where $J_h(\cdot)$ is defined in (2.74). Here, we used that, by continuous mapping theorem, $\underline{\widetilde{\mu}}_{K,n}(\mathcal{Q}_{I,\infty}) \xrightarrow{p} \underline{\mu}_K$ implies also $\underline{f}_{K,n}(\mathcal{Q}_{I,\infty}) \xrightarrow{p} \underline{f}_K$. □

2.10 Proofs of Section 2.4

2.10.1 Proof of Lemma 2.12

By Assumption 2.2, we know that $\mu_0 \geq 1$ and, by Assumption 2.3, we have that $f_j \geq 1$ such that $\mu_{j+1} \geq \mu_j \geq 1$. Similarly, as all summands of

$$\tau_j^2 = \tau_0^2 \prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1}\left(\prod_{m=0}^{l-1} f_m\right)\sigma_l^2 \left(\prod_{n=l+1}^{j-1} f_n^2\right) \tag{2.96}$$

are non-negative and $\tau_0^2 > 0$ by Assumption 2.2, we have $\tau_{j+1}^2 \geq \tau_j^2 > 0$. Furthermore, by Assumption 2.11, we have convergence of $\mu_j = \mu_0 \prod_{k=0}^{j-1} f_k$ for $j \to \infty$. Similarly, as $(\tau_j^2, j \in \mathbb{N}_0)$ is monotonically non-decreasing, and we have

$$\tau_j^2 \leq \tau_0^2 \left(\prod_{k=0}^{\infty} f_k\right)^2 + \mu_0 \left(\prod_{m=0}^{\infty} f_m\right)^2 \sum_{l=0}^{\infty} \sigma_l^2 < \infty \tag{2.97}$$

again by Assumption 2.11, we have also convergence of $\tau_j^2 \to \tau_\infty^2$ for $j \to \infty$. Finally, for $h \in \mathbb{N}$ and using (2.22), we get

$$Cov(C_{i,j}, C_{i,j+h}) = \left(\prod_{k=j}^{j+h-1} f_k \right) \tau_j^2 \to \tau_\infty^2, \tag{2.98}$$

because $f_k \to 1$ as $k \to \infty$ such that also $\prod_{k=j}^{j+h-1} f_k \to 1$ as $j \to \infty$ as well as $\tau_j^2 \to \tau_\infty^2$ as $j \to \infty$ from above. □

2.10.2 Proof of Theorem 2.13

First, as $((R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,\infty} - \widehat{R}_{I,\infty})_1)^2 \geq 0$, unconditional L_2-convergence in (2.36) implies also conditional L_2-convergence in (2.37). Further, the L_2-space of all square-integrable random variables is a Hilbert space and, hence, complete. Consequently, it is sufficient to show that $((R_{I,n} - \widehat{R}_{I,n})_1, n \in \mathbb{N}_0)$ is a Cauchy sequence in (unconditional) L_2-sense. By letting $n \to \infty$ for $(R_{I,n} - \widehat{R}_{I,n})_1$ naturally leads to the L_2-limit $(R_{I,\infty} - \widehat{R}_{I,\infty})_1$ as defined in (2.38). Precisely, we have to show that for all $\epsilon > 0$, there exists an $n_0 \in \mathbb{N}_0$ such that for all $m, n \geq n_0$, we have

$$E\left(\left((R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,m} - \widehat{R}_{I,m})_1 \right)^2 \right) < \epsilon. \tag{2.99}$$

Without loss of generality, let $m \leq n$ in the following. Then, we have

$$\begin{aligned}
&(R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,m} - \widehat{R}_{I,m})_1 \\
=& \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,j} - \prod_{j=i}^{I+n-1} f_j \right) - \sum_{i=0}^{I+m} C_{I-i,i} \left(\prod_{j=i}^{I+m-1} F_{I-i,j} - \prod_{j=i}^{I+m-1} f_j \right) \\
=& \sum_{i=0}^{I+m} C_{I-i,i} \left(\prod_{j=i}^{I+m-1} F_{I-i,j} \left(\prod_{l=I+m}^{I+n-1} F_{I-i,l} - 1 \right) - \prod_{j=i}^{I+m-1} f_j \left(\prod_{l=I+m}^{I+n-1} f_l - 1 \right) \right) \\
&+ \sum_{i=I+m+1}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,j} - \prod_{j=i}^{I+n-1} f_j \right) \\
=& A_{1,I,m,n} + A_{2,I,m,n}.
\end{aligned}$$

Hence, using $(a+b)^2 \leq 2(a^2 + b^2)$ for $a, b \in \mathbb{R}$, it remains to show that

$$\lim_{m\to\infty} \lim_{n\to\infty} E(A_{1,I,m,n}^2) = 0 \quad \text{and} \quad \lim_{m\to\infty} \lim_{n\to\infty} E(A_{2,I,m,n}^2) = 0. \tag{2.100}$$

We begin with the first term $A_{1,I,m,n}$. For convenient notation, let $E_{\mathcal{Q}}(\cdot) = E(\cdot|\mathcal{Q}_{I,\infty})$ and $Var_{\mathcal{Q}}(\cdot) = Var(\cdot|\mathcal{Q}_{I,\infty})$ denote expectation and variance, respectively, conditional on $\mathcal{Q}_{I,\infty}$. First, we calculate the conditional second moment $E_{\mathcal{Q}}(A_{1,I,m,n}^2)$, which easily allows to calculate also $E(A_{1,I,m,n}^2) = E(E_{\mathcal{Q}}(A_{1,I,m,n}^2))$ afterwards. Further, as we already have $E_{\mathcal{Q}}(A_{1,I,m,n}) = 0$, it holds $E_{\mathcal{Q}}(A_{1,I,m,n}^2) = Var_{\mathcal{Q}}(A_{1,I,m,n})$. Using stochastic independence over accident years i, the summands of $A_{1,I,m,n}$ are also independent, leading to

$$\begin{aligned}
&Var_{\mathcal{Q}}(A_{1,I,m,n})\\
&= \sum_{i=0}^{I+m} C_{I-i,i}^2 E_{\mathcal{Q}}\left(\left(\prod_{j=i}^{I+m-1} F_{I-i,j}\left(\prod_{l=I+m}^{I+n-1} F_{I-i,l} - 1\right) - \prod_{j=i}^{I+m-1} f_j \left(\prod_{l=I+m}^{I+n-1} f_l - 1\right)\right)^2\right)\\
&= \sum_{i=0}^{I+m} C_{I-i,i}^2 \left[E_{\mathcal{Q}}\left(\left(\prod_{j=i}^{I+m-1} F_{I-i,j}\left(\prod_{l=I+m}^{I+n-1} F_{I-i,l} - 1\right)\right)^2\right) - \left(\prod_{j=i}^{I+m-1} f_j \left(\prod_{l=I+m}^{I+n-1} f_l - 1\right)\right)^2\right].
\end{aligned}$$

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

$$\begin{aligned}
&\sum_{i=0}^{I+m} C_{I-i,i}^2 E_{\mathcal{Q}}\left(\left(\prod_{j=i}^{I+m-1} F_{I-i,j}\left(\prod_{l=I+m}^{I+n-1} F_{I-i,l} - 1\right)\right)^2\right)\\
&= \sum_{i=0}^{I+m} C_{I-i,i}^2 E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-1} F_{I-i,j}^2 - 2\left(\prod_{j=i}^{I+m-1} F_{I-i,j}^2\right)\left(\prod_{l=I+m}^{I+n-1} F_{I-i,l}\right) + \prod_{j=i}^{I+m-1} F_{I-i,j}^2\right). \quad (2.101)
\end{aligned}$$

For the first expectation on the last right-hand side of (2.101), due to $F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$, we get

$$\begin{aligned}
E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-1} F_{I-i,j}^2\right) &= E_{\mathcal{Q}}\left(\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^2\right) E_{\mathcal{Q}}\left(F_{I-i,I+n-1}^2 | C_{I-i,i}, \ldots, C_{I-i,I+n-1}\right)\right)\\
&= E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^2\left(\frac{\sigma_{I+n-1}^2}{C_{I-i,I+n-1}} + f_{I+n-1}^2\right)\right)\\
&= E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^2 \frac{1}{C_{I-i,I+n-1}}\right)\sigma_{I+n-1}^2 + E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^2\right) f_{I+n-1}^2\\
&= E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-2} F_{I-i,j}\right)\frac{\sigma_{I+n-1}^2}{C_{I-i,i}} + E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^2\right) f_{I+n-1}^2\\
&= \frac{\prod_{j=i}^{I+n-2} f_j \sigma_{I+n-1}^2}{C_{I-i,i}} + E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-2} F_{I-i,j}^2\right) f_{I+n-1}^2.
\end{aligned}$$

By recursively plugging-in, we get

$$E_{\mathcal{Q}}\left(\prod_{j=i}^{I+n-1} F_{I-i,j}^2\right) = \frac{1}{C_{I-i,i}} \sum_{k=i}^{I+n-1}\left(\prod_{j=i}^{k-1} f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+n-1} f_h^2\right) + \prod_{j=i}^{I+n-1} f_j^2. \quad (2.102)$$

Similarly, for the second expectation in (2.101), we get

$$E_{\mathcal{Q}}\left(-2\left(\prod_{j=i}^{I+m-1}F_{I-i,j}^2\right)\left(\prod_{l=I+m}^{I+n-1}F_{I-i,l}\right)\right)$$
$$=-2\left(\frac{1}{C_{I-i,i}}\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)+\prod_{j=i}^{I+m-1}f_j^2\right)\left(\prod_{l=I+m}^{I+n-1}f_l\right)$$

and for the third one, we have

$$E_{\mathcal{Q}}\left(\prod_{j=i}^{I+m-1}F_{I-i,j}^2\right)=\frac{1}{C_{I-i,i}}\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)+\prod_{j=i}^{I+m-1}f_j^2. \tag{2.103}$$

Altogether, for all $m \le n$, this leads to

$$\begin{aligned}
&Var_{\mathcal{Q}}(A_{1,I,m,n})\\
&=\sum_{i=0}^{I+m}C_{I-i,i}\Bigg[\sum_{k=i}^{I+n-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+n-1}f_h^2\right)-2\left(\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\right)\left(\prod_{l=I+m}^{I+n-1}f_l\right)\\
&\qquad+\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\Bigg].
\end{aligned}$$

Next, taking the (unconditional) expectation of $Var_{\mathcal{Q}}(A_{1,I,m,n})$, gives

$$\begin{aligned}
&E(A_{1,I,m,n}^2)\\
&=\sum_{i=0}^{I+m}\mu_i\Bigg[\sum_{k=i}^{I+n-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+n-1}f_h^2\right)-2\left(\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\right)\left(\prod_{l=I+m}^{I+n-1}f_l\right)\\
&\qquad+\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\Bigg]\\
&\underset{n\to\infty}{\to}\sum_{i=0}^{I+m}\mu_i\Bigg[\sum_{k=i}^{\infty}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{\infty}f_h^2\right)-2\left(\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\right)\left(\prod_{l=I+m}^{\infty}f_l\right)\\
&\qquad+\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\Bigg].
\end{aligned}$$

The last right-hand side can be written as

$$\begin{aligned}\sum_{i=0}^{I+m}\mu_i\Bigg[&\sum_{k=i}^{\infty}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{\infty}f_h^2\right)\left(1-\prod_{l=I+m}^{\infty}f_l\right)\\&+\sum_{k=I+m}^{\infty}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{\infty}f_h^2\right)\left(\prod_{l=I+m}^{\infty}f_l\right)\\&+\left(\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\right)\left(\prod_{h=I+m}^{\infty}f_h^2-1\right)\left(\prod_{l=I+m}^{\infty}f_l\right)\\&+\left(\sum_{k=i}^{I+m-1}\left(\prod_{j=i}^{k-1}f_j\right)\sigma_k^2\left(\prod_{h=k+1}^{I+m-1}f_h^2\right)\right)\left(1-\prod_{l=I+m}^{\infty}f_l\right)\Bigg],\end{aligned}$$

which can be bounded by

$$\begin{aligned}&\sum_{i=0}^{I+m}\mu_i\Bigg[\left(\sum_{k=i}^{\infty}\sigma_k^2\right)\mu_\infty^2\left(1-\prod_{l=I+m}^{\infty}f_l\right)+\left(\sum_{k=I+m}^{\infty}\sigma_k^2\right)\mu_\infty^3\\&\qquad+\left(\sum_{k=i}^{I+m-1}\sigma_k^2\right)\mu_\infty^2\left(\prod_{h=I+m}^{\infty}f_h^2-1\right)+\left(\sum_{k=i}^{I+m-1}\sigma_k^2\right)\mu_\infty^2\left(1-\prod_{l=I+m}^{\infty}f_l\right)\Bigg]\\&\underset{m\to\infty}{\longrightarrow}0,\end{aligned}$$

because $\prod_{l=I+m}^{\infty}f_l\to 1$ and $\prod_{l=I+m}^{\infty}f_l^2\to 1$ as $m\to\infty$ by Assumption 2.11, as well as

$$\sum_{i=0}^{I+m}\mu_i\sum_{k=i}^{\infty}\sigma_k^2\le\sum_{i=0}^{\infty}\mu_i\sum_{k=i}^{\infty}\sigma_k^2=\sum_{k=0}^{\infty}\left(\frac{1}{k+1}\sum_{i=0}^{k}\mu_i\right)(k+1)\sigma_k^2\le M\sum_{k=0}^{\infty}(k+1)\sigma_k^2<\infty$$

and

$$\begin{aligned}\sum_{i=0}^{I+m}\mu_i\sum_{k=I+m}^{\infty}\sigma_k^2&=\frac{1}{I+m+1}\sum_{i=0}^{I+m}\mu_i(I+m+1)\sum_{k=I+m}^{\infty}\sigma_k^2\\&\le\left(\frac{1}{I+m+1}\sum_{i=0}^{I+m}\mu_i\right)\sum_{k=I+m}^{\infty}(k+1)\sigma_k^2\underset{m\to\infty}{\longrightarrow}0\end{aligned}$$

due $\sum_{j=0}^{\infty}(j+1)^2\sigma_j^2<\infty$ by Assumption 2.11. Here, we used that the sequence $(\mu_j, j\in\mathbb{N}_0)$ is converging by (2.28), which implies that the sequence $(\frac{1}{j+1}\sum_{k=0}^{j}\mu_k)_{j\in\mathbb{N}_0}$ is also converging and there exists a constant $M<\infty$ such that $\frac{1}{j+1}\sum_{k=0}^{j}\mu_k<M$ for all $j\in\mathbb{N}_0$. This completes the first part of (2.100).
For the second part, we also have $E_{\mathcal{Q}}(A_{2,I,m,n})=0$ and, by using similar arguments, we

have

$$Var_{\mathcal{Q}}(A_{2,I,m,n}) = \sum_{i=I+m+1}^{I+n} C_{I-i,i} \sum_{j=i}^{I+n-1} \left(\prod_{h=i}^{j-1} f_h \right) \sigma_j^2 \left(\prod_{l=j+1}^{I+n-1} f_l^2 \right)$$

and taking the (unconditional) expectation of $Var_{\mathcal{Q}}(A_{2,I,m,n})$ leads to

$$\begin{aligned} E(A_{2,I,m,n}^2) &= \sum_{i=I+m+1}^{I+n} \mu_i \sum_{j=i}^{I+n-1} \left(\prod_{h=i}^{j-1} f_h \right) \sigma_j^2 \left(\prod_{l=j+1}^{I+n-1} f_l^2 \right) \\ &\underset{n\to\infty}{\longrightarrow} \sum_{i=I+m+1}^{\infty} \mu_i \sum_{j=i}^{\infty} \left(\prod_{h=i}^{j-1} f_h \right) \sigma_j^2 \left(\prod_{l=j+1}^{\infty} f_l^2 \right). \end{aligned}$$

The last right-hand side can be bounded by

$$\mu_\infty^2 \sum_{i=I+m+1}^{\infty} \sum_{j=i}^{\infty} \sigma_j^2 = \mu_\infty^2 \sum_{j=I+m+1}^{\infty} (j+1-(I+m+1))\sigma_j^2 \leq \mu_\infty^2 \sum_{j=I+m+1}^{\infty} (j+1)\sigma_j^2 \underset{m\to\infty}{\longrightarrow} 0.$$

Now, as $E_{\mathcal{Q}}((R_{I,n} - \widehat{R}_{I,n})_1) = 0$ for all n, we also get $E_{\mathcal{Q}}((R_{I,\infty} - \widehat{R}_{I,\infty})_1) = 0$ as well as $E((R_{I,\infty} - \widehat{R}_{I,\infty})_1) = 0$. And for the conditional variance of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,\infty}$, we have

$$Var_{\mathcal{Q}}\left((R_{I,n} - \widehat{R}_{I,n})_1\right) = \sum_{i=0}^{I+n} C_{I-i,i} \sum_{j=i}^{I+n-1} \left(\prod_{k=i}^{j-1} f_k \right) \sigma_j^2 \left(\prod_{l=j+1}^{I+n-1} f_l^2 \right),$$

which is indeed bounded in probability, due to

$$\begin{aligned} E\left(\left| \sum_{i=0}^{I+n} C_{I-i,i} \sum_{j=i}^{I+n-1} \sigma_j^2 \right| \right) \mu_\infty^2 &= \left(\sum_{i=0}^{I+n} \mu_i \sum_{j=i}^{I+n-1} \sigma_j^2 \right) \mu_\infty^2 = \left(\sum_{j=0}^{I+n-1} \sigma_j^2 \sum_{i=0}^{j} \mu_i \right) \mu_\infty^2 \\ &= \left(\sum_{j=0}^{I+n-1} (j+1)\sigma_j^2 \left(\frac{1}{j+1} \sum_{i=0}^{j} \mu_i \right) \right) \mu_\infty^2 < \infty. \end{aligned}$$

Hence, for $n \to \infty$, we have

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

Using $E(Var_{\mathcal{Q}}\left((R_{I,n}-\widehat{R}_{I,n})_1\right))=Var\left((R_{I,n}-\widehat{R}_{I,n})_1\right)$, for the unconditional variance of $(R_{I,n}-\widehat{R}_{I,n})_1$, we get

$$
\begin{aligned}
Var\left((R_{I,n}-\widehat{R}_{I,n})_1\right) &= \sum_{i=0}^{I+n}\mu_i\sum_{j=i}^{I+n-1}\left(\prod_{k=i}^{j-1}f_k\right)\sigma_j^2\left(\prod_{l=j+1}^{I+n-1}f_l^2\right)\\
&\xrightarrow{p}\sum_{i=0}^{\infty}\mu_i\sum_{j=i}^{\infty}\left(\prod_{k=i}^{j-1}f_k\right)\sigma_j^2\left(\prod_{l=j+1}^{\infty}f_l^2\right)<\infty.
\end{aligned}
$$

Finally, the non-degeneracy of the limiting distributions is implied by the boundedness away from zero of $Var_{\mathcal{Q}}\left((R_{I,n}-\widehat{R}_{I,n})_1\right)$, which is obtained by taking the first summands of both sums over i and j in the formula for the variance $Var_{\mathcal{Q}}((R_{I,n}-\widehat{R}_{I,n})_1)$ as a lower bound and by bounding all f_j's from below by 1. This leads to

$$
Var_{\mathcal{Q}}\left((R_{I,n}-\widehat{R}_{I,n})_1\right)\geq C_{I,0}\sigma_0^2>0
$$

as $C_{i,0}\geq 1$ for all i by Assumption 2.2 and $\sigma_0^2>0$ by Assumption 2.11. $\square$

2.10.3 Proof of Theorem 2.17

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

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

where $\widehat{f}_{j,n}$ is defined in (2.35). Hence, to prove the theorem, it suffices to show that, a) for all $K\in\mathbb{N}_0$, $B_{1,K,I,n}\xrightarrow{d}B_{1,K}$ as $n\to\infty$, where $B_{1,K}\sim\mathcal{G}_{2,K}$ for some distribution $\mathcal{G}_{2,K}$, b)

$B_{1,K} \xrightarrow{d} \mathcal{G}_2$ as $K \to \infty$, and c) that, for all $\epsilon > 0$, we have

$$\lim_{K\to\infty} \limsup_{n\to\infty} P\left(|B_{2,K,I,n}| > \epsilon\right) = 0 \quad \text{and} \quad \lim_{K\to\infty} \limsup_{n\to\infty} P\left(|B_{3,K,I,n}| > \epsilon\right) = 0. \tag{2.104}$$

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

$$B_{1,K,I,n} = \sqrt{I+n+1} \sum_{i=0}^{K} C_{I-i,i} \left(\prod_{j=i}^{K} f_j - \prod_{j=i}^{K} \widehat{f}_{j,n} \right), \tag{2.105}$$

where

$$\widehat{f}_{j,n} = \frac{\sum\limits_{k=-n}^{I-j-1} C_{k,j+1}}{\sum\limits_{k=-n}^{I-j-1} C_{k,j}}. \tag{2.106}$$

A closer inspection of (2.105) and (2.106) shows that for each $i = 1, \ldots, K$ the diagonal elements $C_{I-i,i}$, $i \in \{1, \ldots, K\}$ show twice: as weights in the sum in (2.105), but also as the last summand in the numerator of $\widehat{f}_{i-1,n}$ in (2.106). Hence, for each $i = 1, \ldots, K$, $C_{I-i,i}$ and $\widehat{f}_{i-1,n}$ are *not* independent. To solve this dependence, we replace $\widehat{f}_{j,n}$ in (2.105) by

$$\widetilde{f}_{j,n,K} = \frac{\sum\limits_{i=-n}^{I-K-1} C_{i,j+1}}{\sum\limits_{i=-n}^{I-K-1} C_{i,j}} \tag{2.107}$$

defined similarly as $\widetilde{f}_{j,K}$ in (2.65) such that $(C_{I-i,i}, i = 0, \ldots, K)$ and $(\widetilde{f}_{j,n,K}, j = 0, \ldots, K)$ are independent. More precisely, we write

$$\begin{aligned}
&B_{1,K,I,n} \\
&= \sqrt{I+n+1} \sum_{i=0}^{K} C_{I-i,i} \left(\prod_{j=i}^{K} f_j - \prod_{j=i}^{K} \widetilde{f}_{j,n,K} \right) + \sqrt{I+n+1} \sum_{i=0}^{K} C_{I-i,i} \left(\prod_{j=i}^{K} \widetilde{f}_{j,n,K} - \prod_{j=i}^{K} \widehat{f}_{j,n} \right) \\
&= \sum_{i=0}^{K} C_{I-i,i} \sqrt{I+n+1} \left(\prod_{j=i}^{K} f_j - \prod_{j=i}^{K} \widetilde{f}_{j,n,K} \right) + \sum_{i=0}^{K} C_{I-i,i} \sqrt{I+n+1} \left(\prod_{j=i}^{K} \widetilde{f}_{j,n,K} - \prod_{j=i}^{K} \widehat{f}_{j,n} \right) \\
&=: B_{1,K,I,n,1} + B_{1,K,I,n,2}.
\end{aligned}$$

Although the results in Section 2.3 are obtained in a different asymptotic framework, all results still hold under the asymptotic framework of Section 2.4. Hence, for $n \to \infty$, Corollary 2.6 leads to

$$B_{1,K,I,1} \xrightarrow{d} \langle \mathcal{Q}_{I,K-I}, \mathbf{Y}_K \rangle \,,$$

where $\mathcal{Q}_{I,K-I} = \{C_{I-i,i} | i = 0, \ldots, I + n + (K - I - n) = K\}$, $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^{K+1}$, and $\mathbf{Y}_K = (Y_i, i = 0, \ldots, K)$ is a $(K + 1)$-dimensional multivariate normally distributed random variable with $\mathbf{Y}_K \sim \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\prod f_j}\right)$. As $(C_{I-i,i} | i = 0, \ldots, K)$ and $(\widetilde{f}_{j,K}, j = 0, \ldots, K)$ are independent by construction, $\mathcal{Q}_{I,K-I}$ and $\mathbf{Y}_K$ are also independent. Now, let's turn to

$$B_{1,K,I,n,2} = \sqrt{I + n + 1} \sum_{i=0}^{K} C_{I-i,i} \left(\prod_{j=i}^{K} \widetilde{f}_{j,n,K} - \prod_{j=i}^{K} \widehat{f}_{j,n} \right).$$

As K is finite, it remains to consider some arbitrary summand of $B_{1,K,I,n,2}$, i.e.

$$C_{I-i,i} \left(\prod_{j=i}^{K} \widetilde{f}_{j,n,K} - \prod_{j=i}^{K} \widehat{f}_{j,n} \right).$$

Further, as $C_{I-i,i} = O_P(1)$ by Lemma 2.4 and because of $\widetilde{f}_{j,n,K} - \widehat{f}_{j,n} = O_P\left(\frac{1}{I+n+1}\right)$ for all $j = 0, \ldots, K$ as shown in the proof of Theorem 2.5, and again as K is finite, we get also $\prod_{j=i}^{K} \widetilde{f}_{j,n,K} - \prod_{j=i}^{K} \widehat{f}_{j,n} = O_P\left(\frac{1}{I+n+1}\right)$. Together, this proves $B_{1,K,I,n,2} = O_P\left(\frac{1}{\sqrt{I+n+1}}\right)$ such that

$$B_{1,K,I,n} = B_{1,K,I,n,1} + O_P\left(\frac{1}{\sqrt{I + n + 1}}\right) \xrightarrow{d} \langle \mathcal{Q}_{I,K-I}, \mathbf{Y}_K \rangle = \mathcal{G}_{2,K},$$

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

$$Cov(Y_{i_1}, Y_{i_2}) = \lim_{K \to \infty} \mathbf{\Sigma}_{K,\prod f_j}(i_1, i_2) = \sum_{j=\max(i_1,i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{l=\max(i_1,i_2), l \neq j}^{\infty} f_l^2 \prod_{m=\min(i_1,i_2)}^{\max(i_1,i_2)-1} f_m \quad (2.108)$$

for $i_1, i_2 \in \mathbb{N}_0$.

As $\mathcal{Q}_{I,\infty}$ and $\mathbf{Y}_\infty$ are stochastically independent, the variance of $\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty \rangle$ computes to

$$
\begin{aligned}
&Var\left(\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty \rangle\right) = \sum_{i=0}^{\infty} Var(C_{I-i,i}Y_i) + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty} Cov(C_{I-i_1,i_1}Y_{i_1}, C_{I-i_2,i_2}Y_{i_2})\\
&= \sum_{i=0}^{\infty} \left(Var(C_{I-i,i}) + E(C_{I-i,i})^2\right) Var(Y_i) + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty} E(C_{I-i_1,i_1})E(C_{I-i_2,i_2})Cov(Y_{i_1}, Y_{i_2})\\
&= \sum_{i=0}^{\infty}(\tau_i^2 + \mu_i^2) \sum_{j=i}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{\substack{l=i\\ l\neq j}}^{\infty} f_l^2 + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty} \mu_{i_1}\mu_{i_2} \sum_{j=\max(i_1,i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{\substack{l=\max(i_1,i_2)\\ l\neq j}}^{\infty} f_l^2 \prod_{m=\min(i_1,i_2)}^{\max(i_1,i_2)-1} f_m\\
&= \sum_{j=0}^{\infty} \frac{\sigma_j^2}{\mu_j f_j^2} \left(\sum_{i=0}^{j} \left(\mu_i^2 + \tau_i^2\right) \prod_{k=i}^{\infty} f_k^2 \right) + 2\sum_{j=1}^{\infty} \frac{\sigma_j^2}{\mu_j} \left(\sum_{i_1=1}^{j} \mu_{i_1} \sum_{i_2=0}^{i_1-1} \mu_{i_2} \prod_{l=i_2}^{i_1-1} f_l \right) \prod_{k=j}^{\infty} f_k^2\\
&\leq \sum_{j=0}^{\infty} \frac{\sigma_j^2}{\mu_j}(j+1)\left(\mu_\infty^2 + \tau_\infty^2\right) + \sum_{j=1}^{\infty} \frac{\sigma_j^2}{\mu_j} j(j+1)\mu_\infty^2 < \infty.
\end{aligned}
$$

We continue with showing part c) for $B_{2,K,I,n}$. Using the unbaisedness and uncorrelatedness of the $\widehat{f}_{j,n}$'s and the stochastic independence of $C_{I-i,i}$ and $\widehat{f}_{j,n}$, $j \geq i$, we have $E(B_{2,K,I,n}) = 0$ by construction. Note that, for the variance $Var(B_{2,K,I,n})$, in contrast to the calculations for $A_{1,I,m,n}$ in the proof of Theorem 2.13, we cannot exploit the stochastic independence over accident years here as numerator and denominator of $\widehat{f}_{j,n}$ sum cumulative claims of several rows of $\mathcal{C}_{I,\infty} = \{C_{i,j} | i \in \mathbb{Z}, i \leq I, j \in \mathbb{N}_0\}$. However, for any fixed $K \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$ large enough such that $K < I + n - 1$, we get

$$
Var(B_{2,K,I,n}) \tag{2.109}
$$

$$
= (I+n+1)\times
$$

$$
\sum_{i_1,i_2=0}^{K} Cov\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1} \widehat{f}_{j_1,n} - \prod_{j_1=i_1}^{K} \widehat{f}_{j_1,n} \right), C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1} \widehat{f}_{j_2,n} - \prod_{j_2=i_2}^{K} \widehat{f}_{j_2,n} \right) \right) \tag{2.110}
$$

$$
\leq 2(I+n+1)\times
$$

$$
\sum_{i_2=0}^{K} \sum_{i_1=0}^{i_2} Cov\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1} \widehat{f}_{j_1,n} - \prod_{j_1=i_1}^{K} \widehat{f}_{j_1,n} \right), C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1} \widehat{f}_{j_2,n} - \prod_{j_2=i_2}^{K} \widehat{f}_{j_2,n} \right) \right). \tag{2.111}
$$

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

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

Further, let $\mathcal{B}_{I,n}(k)=\{C_{i,j}|i=-n,\ldots,I,\ j=0,\ldots,k,\ i+j\leq I\}$ denote all elements of $\mathcal{D}_{I,n}$ up to its kth column. Let us consider the first term on the last right-hand side. By recursively applying the law of iterated expectation, we get

$$\begin{aligned}
&E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}^2_{j_2,n}\right)\right)\\
&=E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-2}\widehat{f}^2_{j_2,n}\right)E\left(\widehat{f}^2_{I+n-1,n}|\mathcal{B}_{I,n}(I+n-1)\right)\right)\\
&=E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-2}\widehat{f}^2_{j_2,n}\right)\left(\frac{\sigma^2_{I+n-1}}{\sum_{k=-n}^{I-(I+n-1)-1}C_{k,I+n-1}}+f^2_{I+n-1}\right)\right)\\
&=\sum_{j_4=0}^{I+n-K-2}E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K}\widehat{f}^2_{j_2,n}\right)\left(\prod_{j_3=K+1}^{I+n-1-(j_4+1)}\widehat{f}^2_{j_3,n}\right)\right.\\
&\qquad\left.\times\frac{\sigma^2_{I+n-1-j_4}}{\sum\limits_{k=-n}^{I-(I+n-1-j_4)-1}C_{k,I+n-1-j_4}}\right)\left(\prod_{j_5=I+n-1-j_4+1}^{I+n-1}f^2_{j_5}\right)\\
&\qquad+E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K}\widehat{f}^2_{j_2,n}\right)\right)\left(\prod_{j_3=K+1}^{I+n-1}f^2_{j_3}\right)
\end{aligned}$$

due to, for all $c\in\{0,\ldots,I+n-1\}$,

$$E\left(\widehat{f}^2_{c,n}|\mathcal{B}_{I,n}(c)\right)=\frac{\sigma^2_c}{\sum_{k=-n}^{I-c-1}C_{k,c}}+f^2_c. \tag{2.112}$$

Altogether, using similar arguments to (partially) calculate also the other expectations above and due to $E(C_{I-i_1,i_1}(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}-\prod_{j_1=i_1}^{K}\widehat{f}_{j_1,n}))=\mu_{i_1}(\prod_{j_1=i_1}^{I+n-1}f_{j_1}-\prod_{j_1=i_1}^{K}f_{j_1})$ as well as $E(C_{I-i_2,i_2}(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}-\prod_{j_2=i_2}^{K}\widehat{f}_{j_2,n}))=\mu_{i_2}(\prod_{j_2=i_2}^{I+n-1}f_{j_2}-\prod_{j_2=i_2}^{K}f_{j_2})$, we get for one covariance summand on the right-hand side of (2.111) (inflated with $I+n+1$)

$$(I+n+1)\cdot Cov\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}-\prod_{j_1=i_1}^{K}\widehat{f}_{j_1,n}\right),C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}-\prod_{j_2=i_2}^{K}\widehat{f}_{j_2,n}\right)\right)$$

$$\begin{aligned}=&(I+n+1)\times\\ &\left[E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K}\widehat{f}^2_{j_2,n}\right)\right)\left\{\left(\prod_{j_3=K+1}^{I+n-1}f^2_{j_3}\right)-2\left(\prod_{j_3=K+1}^{I+n-1}f_{j_3}\right)+1\right\}\right.\\ &\left.\qquad-\mu_{i_1}\left(\prod_{j_1=i_1}^{I+n-1}f_{j_1}-\prod_{j_1=i_1}^{K}f_{j_1}\right)\mu_{i_2}\left(\prod_{j_2=i_2}^{I+n-1}f_{j_2}-\prod_{j_2=i_2}^{K}f_{j_2}\right)\right] \qquad (2.113)\\ &+(I+n+1)\sum_{j_4=0}^{I+n-K-2}E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K}\widehat{f}^2_{j_2,n}\right)\left(\prod_{j_3=K+1}^{I+n-1-(j_4+1)}\widehat{f}^2_{j_3,n}\right)\right.\\ &\left.\qquad\times\frac{\sigma^2_{I+n-1-j_4}}{\sum_{k=-n}^{I-(I+n-1-j_4)-1}C_{k,I+n-1-j_4}}\right)\left(\prod_{j_5=I+n-1-j_4+1}^{I+n-1}f^2_{j_5}\right). \qquad (2.114)\end{aligned}$$

Next, we will consider the two terms above separately. Starting with (2.113), which can be expressed as

$$(I+n+1)\left[E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K}\widehat{f}^2_{j_2,n}\right)\right)-\mu_{i_1}\left(\prod_{j_1=i_1}^{i_2-1}f_{j_1}\right)\mu_{i_2}\left(\prod_{j_2=i_2}^{K}f^2_{j_2}\right)\right] \qquad (2.115)$$

$$\times\left(\prod_{j_3=K+1}^{I+n-1}f_{j_3}-1\right)^2,$$

we use the same technique as above and make recursive use of (2.112). Then, similar to the calculations above, the first factor in (2.115) becomes

$$(I+n+1)\sum_{j_4=0}^{K-i_2}E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K-(j_4+1)}\widehat{f}^2_{j_2,n}\right)\frac{\sigma^2_{K-j_4}}{\sum\limits_{k=-n}^{I-(K-j_4)-1}C_{k,K-j_4}}\right)$$

$$\times\left(\prod_{j_5=K-j_4+1}^{K}f^2_{j_5}\right),$$

which can be bounded by

$$\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K-(j_4+1)}\widehat{f}^2_{j_2,n}\right)\frac{(I+n+1)\sigma^2_{K-j_4}}{\sum_{k=-n}^{I-(K-j_4)-1}C_{k,K-j_4}}\right)$$
$$=\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\right.$$
$$\left.\times\left(\prod_{j_2=i_2}^{K-(j_4+1)}\widehat{f}^2_{j_2,n}\right)\frac{(I+n+1-(I+n-(K-j_4)))\sigma^2_{K-j_4}}{\sum_{k=-n}^{I-(K-j_4)-1}C_{k,K-j_4}}\right)$$
$$+\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{K-(j_4+1)}\widehat{f}^2_{j_2,n}\right)\frac{\sigma^2_{K-j_4}}{\frac{1}{I+n-(K-j_4)}\sum_{k=-n}^{I-(K-j_4)-1}C_{k,K-j_4}}\right)$$
$$=\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{i_2-1+j_4}\widehat{f}^2_{j_2,n}\right)\frac{(i_2+j_4+1)\sigma^2_{i_2+j_4}}{\sum_{k=-n}^{I-(i_2+j_4)-1}C_{k,i_2+j_4}}\right)$$
$$+\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{i_2-1+j_4}\widehat{f}^2_{j_2,n}\right)\frac{\sigma^2_{i_2+j_4}}{\frac{1}{I+n-(i_2+j_4)}\sum_{k=-n}^{I-(i_2+j_4)-1}C_{k,i_2+j_4}}\right),$$

where we reversed the summation order of j_4 in the last step. The leading term of the last right-hand side (with respect to $n \to \infty$) is the second one, which we will consider next in more detail. Using Assumption 2.16, $\frac{1}{I+n-(i_2+j_4)}\sum_{k=-n}^{I-(i_2+j_4)-1}C_{k,i_2+j_4}$ can be bounded from below by $\epsilon^{i_2+j_4}$

$$\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{i_2-1+j_4}\widehat{f}^2_{j_2,n}\right)\frac{\sigma^2_{i_2+j_4}}{\epsilon^{i_2+j_4}}\right) \tag{2.116}$$
$$\leq\mu_\infty^2 \sum_{j_4=0}^{K-i_2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\right)\left(\prod_{j_2=i_2}^{i_2-1+j_4}\left(f^2_{j_2}+\frac{\sigma^2_{j_2}}{\epsilon^{j_2}}\right)\right)\frac{\sigma^2_{i_2+j_4}}{\epsilon^{i_2+j_4}}. \tag{2.117}$$

Now, we compute the expectation on the last right-hand side above. We get

$$E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\right)$$
$$=E\left(E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}|\mathcal{B}_{I,n}(i_2-1)\right)\right)$$
$$=E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-2}\widehat{f}_{j_1,n}\right)E\left(\widehat{f}_{i_2-1,n}C_{I-i_2,i_2}|\mathcal{B}_{I,n}(i_2-1)\right)\right)$$

and the last interior conditional expectation can be bounded as follows

$$\begin{aligned}
&\frac{1}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}} E\left(\sum_{r=-n}^{I-i_2} C_{r,i_2} C_{I-i_2,i_2} | \mathcal{B}_{I,n}(i_2-1)\right)\\
&= \frac{1}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}} E\left(\sum_{r=-n}^{I-i_2-1} C_{r,i_2} C_{I-i_2,i_2} + C_{I-i_2,i_2}^2 | \mathcal{B}_{I,n}(i_2-1)\right)\\
&= \frac{1}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}} \left(\sum_{r=-n}^{I-i_2-1} E\left(C_{r,i_2} C_{I-i_2,i_2} | \mathcal{B}_{I,n}(i_2-1)\right) + E\left(C_{I-i_2,i_2}^2 | \mathcal{B}_{I,n}(i_2-1)\right)\right)\\
&= \frac{1}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}} \left(\sum_{r=-n}^{I-i_2-1} E\left(C_{r,i_2} | \mathcal{B}_{I,n}(i_2-1)\right) E\left(C_{I-i_2,i_2} | \mathcal{B}_{I,n}(i_2-1)\right)\right.\\
&\qquad\qquad \left. + E\left(C_{I-i_2,i_2}^2 | \mathcal{B}_{I,n}(i_2-1)\right)\right)\\
&= \frac{1}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}} \left(\sum_{r=-n}^{I-i_2-1} f_{i_2-1} C_{r,i_2-1} f_{i_2-1} C_{I-i_2,i_2-1} + \sigma_{i_2-1}^2 C_{I-i_2,i_2-1} + \left(f_{i_2-1} C_{I-i_2,i_2-1}\right)^2\right)\\
&= \frac{1}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}} \left(f_{i_2-1}^2 \left(\sum_{r=-n}^{I-i_2} C_{r,i_2-1}\right) C_{I-i_2,i_2-1} + \sigma_{i_2-1}^2 C_{I-i_2,i_2-1}\right)\\
&= \left(f_{i_2-1}^2 + \frac{\sigma_{i_2-1}^2}{\sum_{r=-n}^{I-i_2} C_{r,i_2-1}}\right) C_{I-i_2,i_2-1}\\
&\leq \left(f_{i_2-1}^2 + \frac{\sigma_{i_2-1}^2}{(I+n-i_2+1)\epsilon^{i_2-1}}\right) C_{I-i_2,i_2-1}\\
&\leq \left(f_{i_2-1}^2 + \frac{\sigma_{i_2-1}^2}{\epsilon^{i_2-1}}\right) C_{I-i_2,i_2-1}
\end{aligned}$$

as $I - i_2 + 1 \geq 1$ due to $i_2 \in \{1, \ldots, K\}$. Continuing this way, we get rid of $\prod_{j_1=i_1}^{i_2-1} \widehat{f}_{j_1,n}$ in the expectation in (2.117), which then can be bounded by

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

Together with (2.111), the corresponding expression can be bounded by

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

where we used that $\prod_{j=0}^{\infty}x_j<\infty$ if and only if $\sum_{j=0}^{\infty}(x_j-1)<\infty$ for $x_j\ge 1$ for all j, and we have

$$
\begin{aligned}
\sum_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}-1\right)&=\sum_{j_3=0}^{\infty}(f_{j_3}^2-1)+\sum_{j_3=0}^{\infty}\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}=\sum_{j_3=0}^{\infty}(f_{j_3}-1)(f_{j_3}+1)+\sum_{j_3=0}^{\infty}\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\\
&\le\sup_{j\in\mathbb{N}}(f_j+1)\sum_{j_3=0}^{\infty}(f_{j_3}-1)+\sum_{j_3=0}^{\infty}\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}<\infty.
\end{aligned}
$$

Finally, we have

$$
\sum_{i_2=0}^{K}(i_2+1)\sum_{j_4=i_2}^{K}\frac{\sigma_{j_4}^2}{\epsilon^{j_4}}=\sum_{j_4=0}^{K}\frac{\sigma_{j_4}^2}{\epsilon^{j_4}}\sum_{i_2=0}^{j_4}(i_2+1)\le\sum_{j_4=0}^{K}(j_4+1)^2\frac{\sigma_{j_4}^2}{\epsilon^{j_4}}, \tag{2.118}
$$

which is bounded for $K\to\infty$ by Assumption 2.16. Together with the second factor in (2.115), for which we have

$$
\left(\prod_{j_3=K+1}^{I+n-1}f_{j_3}-1\right)^2\underset{n\to\infty}{\to}\left(\prod_{j_3=K+1}^{\infty}f_{j_3}-1\right)^2\underset{K\to\infty}{\to}0, \tag{2.119}
$$

the corresponding expression in (2.109) vanishes. This completes the first part of c) for $B_{2,K,I,n}$. Now, using similar arguments to deal with the first term in (2.113), we continue with the second term in (2.114) leading to

$$\begin{aligned}
&(I+n+1) \\
&\times \sum_{j_4=0}^{I+n-K-2} \left[E\left(C_{I-i_1,i_1} \left(\prod_{j_1=i_1}^{i_2-1} \widehat{f}_{j_1,n} \right) C_{I-i_2,i_2} \left(\prod_{j_3=i_2}^{I+n-1-(j_4+1)} \widehat{f}^2_{j_3,n} \right) \frac{\sigma^2_{I+n-1-j_4}}{\sum_{k=-n}^{I-(I+n-1-j_4)-1} C_{k,I+n-1-j_4}} \right) \right. \\
&\qquad \left. \times \left(\prod_{j_5=I+n-1-j_4+1}^{I+n-1} f^2_{j_5} \right) \right] \\
&\leq \mu^2_\infty \sum_{j_4=0}^{I+n-K-2} E\left(C_{I-i_1,i_1} \left(\prod_{j_1=i_1}^{i_2-1} \widehat{f}_{j_1,n} \right) C_{I-i_2,i_2} \left(\prod_{j_3=i_2}^{K+j_4} \widehat{f}^2_{j_3,n} \right) \frac{(K+j_4+2)\sigma^2_{K+j_4+1}}{\sum_{k=-n}^{I-(K+j_4+1)-1} C_{k,K+j_4+1}} \right) \\
&+ \mu^2_\infty \sum_{j_4=0}^{I+n-K-2} E\left(C_{I-i_1,i_1} \left(\prod_{j_1=i_1}^{i_2-1} \widehat{f}_{j_1,n} \right) C_{I-i_2,i_2} \left(\prod_{j_3=i_2}^{K+j_4} \widehat{f}^2_{j_3,n} \right) \right. \\
&\qquad \left. \times \frac{\sigma^2_{K+j_4+1}}{\frac{1}{I+n-(K+j_4+1)} \sum_{k=-n}^{I-(K+j_4+1)-1} C_{k,K+j_4+1}} \right),
\end{aligned}$$

where we reversed again the summation order of j_4 in the last step. The leading term of the last right-hand side is the second one, which we will consider next in more detail. Using Assumption 2.16, similar to above, $\frac{1}{I+n-(K+j_4+1)} \sum_{k=-n}^{I-(K+j_4+1)-1} C_{k,K+j_4+1}$ can be bounded from below by ϵ^{K+j_4+1} such that the leading term from above can be bounded by

$$\begin{aligned}
&\mu^2_\infty \sum_{j_4=0}^{I+n-K-2} E\left(C_{I-i_1,i_1} \left(\prod_{j_1=i_1}^{i_2-1} \widehat{f}_{j_1,n} \right) C_{I-i_2,i_2} \left(\prod_{j_3=i_2}^{K+j_4} \widehat{f}^2_{j_3,n} \right) \frac{\sigma^2_{K+j_4+1}}{\epsilon^{K+j_4+1}} \right) \\
\leq &\mu^2_\infty (\tau^2_\infty + \mu^2_\infty) \sum_{j_4=0}^{I+n-K-2} \left(\prod_{j_3=i_1}^{K+j_4} \left(f^2_{j_3} + \frac{\sigma^2_{j_3}}{\epsilon^{j_3}} \right) \right) \frac{\sigma^2_{K+j_4+1}}{\epsilon^{K+j_4+1}}.
\end{aligned}$$

Together with (2.111), the corresponding expression can be bounded by

$$\begin{aligned}
&2\mu^2_\infty (\tau^2_\infty + \mu^2_\infty) \sum_{i_2=0}^{K} \sum_{i_1=0}^{i_2} \sum_{j_4=0}^{I+n-K-2} \left(\prod_{j_3=i_1}^{K+j_4} \left(f^2_{j_3} + \frac{\sigma^2_{j_3}}{\epsilon^{j_3}} \right) \right) \frac{\sigma^2_{K+j_4+1}}{\epsilon^{K+j_4+1}} \\
\leq &2\mu^2_\infty (\tau^2_\infty + \mu^2_\infty) \left(\prod_{j_3=0}^{\infty} \left(f^2_{j_3} + \frac{\sigma^2_{j_3}}{\epsilon^{j_3}} \right) \right) \sum_{j_4=K+1}^{I+n-1} (j_4+1)^2 \frac{\sigma^2_{j_4}}{\epsilon^{j_4}} \\
\xrightarrow[n\to\infty]{} &2\mu^2_\infty (\tau^2_\infty + \mu^2_\infty) \left(\prod_{j_3=0}^{\infty} \left(f^2_{j_3} + \frac{\sigma^2_{j_3}}{\epsilon^{j_3}} \right) \right) \sum_{j_4=K+1}^{\infty} (j_4+1)^2 \frac{\sigma^2_{j_4}}{\epsilon^{j_4}},
\end{aligned}$$

which is finite for each fixed K. Finally, letting $K \to \infty$, we get

$$2\mu_\infty^2(\tau_\infty^2+\mu_\infty^2)\left(\prod_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\right)\right)\sum_{j_4=K+1}^{\infty}(j_4+1)^2\frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \longrightarrow 0.$$

Similarly, for showing part c) for $B_{3,K,I,n}$, we get $E(B_{3,K,I,n})=0$. Further, for any fixed $K \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$ large enough such that $K < I+n-1$, we get

$$Var(B_{3,K,I,n}) \tag{2.120}$$

$$=(I+n+1)\sum_{i_1,i_2=K+1}^{I+n-1} Cov\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}\right), C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}\right)\right) \tag{2.121}$$

$$\leq 2(I+n+1)\sum_{i_2=K+1}^{I+n-1}\sum_{i_1=K+1}^{i_2} Cov\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}\right), C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}\right)\right). \tag{2.122}$$

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

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

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

$$\begin{aligned}
&(I+n+1)Cov\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}\right), C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}\right)\right)\\
&=(I+n+1)\Bigg[E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}\right)C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}\right)\right)\\
&\qquad -E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1}\widehat{f}_{j_1,n}\right)\right)E\left(C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1}\widehat{f}_{j_2,n}\right)\right)\Bigg]\\
&=(I+n+1)\\
&\times\sum_{j_4=0}^{I+n-K-2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1}\right)C_{I-i_2,i_2}\left(\prod_{j_3=i_2}^{I+n-1-(j_4+1)}\widehat{f}_{j_3}^2\right)\frac{\sigma_{I+n-1-j_4}^2}{\sum_{k=-n}^{I-(I+n-1-j_4)-1}C_{k,I+n-1-j_4}}\right)\\
&\times\left(\prod_{j_5=I+n-1-j_4+1}^{I+n-1} f_{j_5}^2\right).
\end{aligned}$$

Note that $i_1, i_2 > K$ and K is fixed. This remaining covariance term has the same structure as the one in the covariance term for $B_{2,K,I,n}$. Hence, this term is in the same way bounded.

The corresponding expression can be bounded by

$$\mu_\infty^2 \sum_{j_4=0}^{I+n-K-2} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{i_2-1}\widehat{f}_{j_1}\right)C_{I-i_2,i_2}\left(\prod_{j_3=i_2}^{K+j_4}\widehat{f}_{j_3}^2\right)\frac{\sigma_{K+j_4+1}^2}{\epsilon^{K+j_4+1}}\right) \tag{2.123}$$

$$\leq \mu_\infty^2(\tau_\infty^2+\mu_\infty^2)\sum_{j_4=0}^{I+n-K-2}\left(\prod_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\right)\right)\frac{\sigma_{K+j_4+1}^2}{\epsilon^{K+j_4+1}}, \tag{2.124}$$

where we reversed the summation over j_4. Together with (2.122), the corresponding expression can be bounded by

$$2\mu_\infty^2(\tau_\infty^2+\mu_\infty^2)\left(\prod_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\right)\right)\sum_{i_2=K+1}^{I+n-1}\sum_{i_1=K+1}^{i_2}\sum_{j_4=0}^{I+n-K-2}\frac{\sigma_{K+j_4+1}^2}{\epsilon^{K+j_4+1}} \tag{2.125}$$

$$\leq 2\mu_\infty^2(\tau_\infty^2+\mu_\infty^2)\left(\prod_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\right)\right)\sum_{j_4=K+1}^{I+n-1}(j_4+1)^2\frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \tag{2.126}$$

$$\xrightarrow[n\to\infty]{} 2\mu_\infty^2(\tau_\infty^2+\mu_\infty^2)\left(\prod_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\right)\right)\sum_{j_4=K+1}^{\infty}(j_4+1)^2\frac{\sigma_{j_4}^2}{\epsilon^{j_4}}, \tag{2.127}$$

which is finite for each fixed K. Finally, letting $K \to \infty$, we get

$$2\mu_\infty^2(\tau_\infty^2+\mu_\infty^2)\left(\prod_{j_3=0}^{\infty}\left(f_{j_3}^2+\frac{\sigma_{j_3}^2}{\epsilon^{j_3}}\right)\right)\sum_{j_4=K+1}^{\infty}(j_4+1)^2\frac{\sigma_{j_4}^2}{\epsilon^{j_4}} \longrightarrow 0.$$

□

2.10.4 Proof of Example 2.19

By construction, the marginal pdf of C_j is

$$g_{C_j}(x) = \frac{1}{2\sqrt{3\tau_j^2}} 1_{[\mu_j - \sqrt{3\tau_j^2}, \mu_j + \sqrt{3\tau_j^2}]}(x)$$

and, given $C_j = x$, F_j has the conditional pdf

$$g_{F_j|C_j}(y|x) = \frac{1}{2\sqrt{3\sigma_j^2/x}} 1_{[f_j - \sqrt{3\sigma_j^2/x}, f_j + \sqrt{3\sigma_j^2/x}]}(y).$$

Hence, (C_j, F_j) has the joint pdf

$$\begin{aligned} g_{C_j,F_j}(x,y) &= g_{F_j|C_j}(y|x) g_{C_j}(x) \\ &= \frac{1}{2\sqrt{3\tau_j^2}} 1_{[\mu_j - \sqrt{3\tau_j^2}, \mu_j + \sqrt{3\tau_j^2}]}(x) \frac{1}{2\sqrt{3\sigma_j^2}} 1_{[f_j - \sqrt{3\sigma_j^2}, f_j + \sqrt{3\sigma_j^2}]}(y) \\ &= \frac{1}{12\sqrt{\tau_j^2 \sigma_j^2/x}} 1_{[\mu_j - \sqrt{3\tau_j^2}, \mu_j + \sqrt{3\tau_j^2}]}(x) 1_{[f_j - \sqrt{3\sigma_j^2/x}, f_j + \sqrt{3\sigma_j^2/x}]}(y). \end{aligned}$$

As interest is in the joint pdf of (C_j, C_{j+1}), we can use the density transformation theorem, and

$$\begin{pmatrix} C_j \\ C_{j+1} \end{pmatrix} = \begin{pmatrix} C_j \\ C_j F_j \end{pmatrix} = H(C_j, F_j),$$

where $H(u,v) = (u, uv)'$ and $H^{-1}(u,v) = (u, v/u)'$ to obtain also the joint pdf of (C_j, C_{j+1}). Precisely, we have

$$\begin{aligned} g_{C_j,C_{j+1}}(u,v) &= g_{C_j,F_j}(H^{-1}(u,v)) \cdot \left| det\left(\frac{dH^{-1}(m,n)}{d(m,n)}\right)_{|(m,n)=(u,v)} \right| \\ &= \frac{1}{12\sqrt{\tau_j^2 \sigma_j^2/u}} 1_{[\mu_j - \sqrt{3\tau_j^2}, \mu_j + \sqrt{3\tau_j^2}]}(u) 1_{[f_j - \sqrt{3\sigma_j^2/u}, f_j + \sqrt{3\sigma_j^2/u}]}\left(\frac{v}{u}\right) \cdot \left|\frac{1}{u}\right| \end{aligned}$$

due to $H^{-1}(m,n) = (m, n/m)'$ and

$$det\left(\frac{dH^{-1}(m,n)}{d(m,n)}\right) = \begin{pmatrix} \frac{dm}{dm} & \frac{dm}{dn} \\ \frac{d(n/m)}{dm} & \frac{d(n/m)}{dm} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{n}{m^2} & \frac{1}{m} \end{pmatrix} = \frac{1}{m}.$$

As $u > 0$ by construction, we have $|1/u| = 1/u$. Further, due to

$$f_j - \sqrt{3\sigma_j^2/u} \leq \frac{v}{u} \Leftrightarrow u \leq \left(\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2} + \frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2$$

and

$$\frac{v}{u} \leq f_j + \sqrt{3\sigma_j^2/u} \Leftrightarrow u \geq \left(\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2} - \frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2,$$

we can re-write the second indicator in $g_{C_j,C_{j+1}}(u,v)$ as

$$1_{\left[\left(\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2} - \frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2, \left(\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2} + \frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2 \right]}(u).$$

Hence, by introducing the notation

$$\begin{aligned} &a := \mu_j - \sqrt{3\tau_j^2}, \quad b := \mu_j + \sqrt{3\tau_j^2}, \\ &c(v) := \left(\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2} - \frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2, \quad d(v) := \left(\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2} + \frac{\sqrt{3\sigma_j^2}}{2f_j} \right)^2, \end{aligned}$$

we get

$$g_{C_j,C_{j+1}}(u,v) = \frac{1}{12\sqrt{\tau_j^2\sigma_j^2 u}} 1_{[a,b]}(u) 1_{[c(v),d(v)]}(u). \tag{2.128}$$

Now, as $g_{C_j|C_{j+1}}(u|v) = g_{C_j,C_{j+1}}(u,v)/g_{C_{j+1}}(v)$, we have to calculate $g_{C_{j+1}}(v)$ next. It is obtained by solving the integral

$$\int_{-\infty}^{\infty} \frac{1}{12\sqrt{\tau_j^2\sigma_j^2 u}} 1_{[a,b]}(u) 1_{[c(v),d(v)]}(u) du.$$

Then, we have $g_{C_{j+1}}(v) = 0$ if v is such that either $b < c(v)$ or $d(v) < a$ hold. Otherwise, that is if $b \geq c(v)$ and $d(v) \geq a$, we have

$$\begin{aligned} g_{C_{j+1}}(v) &= \int_{\max(a,c(v))}^{\min(b,d(v))} \frac{1}{12\sqrt{\tau_j^2\sigma_j^2 u}} du = \frac{1}{12\sqrt{\tau_j^2\sigma_j^2}} \int_{\max(a,c(v))}^{\min(b,d(v))} u^{-1/2} du \\ &= \frac{1}{12\sqrt{\tau_j^2\sigma_j^2}} [2u^{1/2}]_{\max(a,c(v))}^{\min(b,d(v))} = \frac{1}{6\sqrt{\tau_j^2\sigma_j^2}} \left(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))} \right). \end{aligned}$$

Together, for all v such that $b \geq c(v)$ and $d(v) \geq a$, we get

$$\begin{aligned} g_{C_j|C_{j+1}}(u|v) &= \frac{\frac{1}{12\sqrt{\tau_j^2\sigma_j^2 u}} 1_{[a,b]}(u) 1_{[c(v),d(v)]}(u)}{\frac{1}{6\sqrt{\tau_j^2\sigma_j^2}}\left(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}\right)} \\ &= \frac{1_{[\max(a,c(v)),\min(b,d(v))]}(u)}{2\sqrt{u}\left(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}\right)}. \end{aligned}$$

The latter formula allows to compute the conditional expectation $E(C_j|C_{j+1} = v)$ by solving the integral

$$\begin{aligned} &\int_{\max(a,c(v))}^{\min(b,d(v))} u \frac{1}{2\sqrt{u}} \frac{1}{\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}} du \\ &= \frac{1}{2\left(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}\right)} \int_{\max(a,c(v))}^{\min(b,d(v))} \sqrt{u} du \\ &= \frac{1}{2\left(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}\right)} \frac{2}{3}\left((\min(b,d(v)))^{3/2} - (\max(a,c(v)))^{3/2}\right) \\ &= \frac{1}{3}\left(\min(b,d(v)) + \sqrt{\min(b,d(v))\max(a,c(v))} + \max(a,c(v))\right). \end{aligned}$$

By going through all possible cases, it is easy to see that the last right-hand side can always be bounded by $d(v)$ from above and by $c(v)$ from below, that is, we have

$$c(v) \leq E\left(C_j|C_{j+1} = v\right) \leq d(v).$$

This leads to

$$\begin{aligned} c(v) =& \frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2 - 2\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2}\frac{\sqrt{3\sigma_j^2}}{2f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2 \\ =& \frac{v}{f_j} + O(\sigma_j^2) + O\left(\sqrt{v\sigma_j^2}\right) + O(\sigma_j^2) \end{aligned}$$

and

$$\begin{aligned} d(v) =& \frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2 + 2\sqrt{\frac{v}{f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2}\frac{\sqrt{3\sigma_j^2}}{2f_j} + \left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2 \\ =& \frac{v}{f_j} + O(\sigma_j^2) + O\left(\sqrt{v\sigma_j^2}\right) + O(\sigma_j^2). \end{aligned}$$

Altogether, this leads to

$$O\left(\sqrt{v\sigma_j^2}\right) \leq E\left(C_j|C_{j+1}=v\right) - \frac{v}{f_j} \leq O\left(\sqrt{v\sigma_j^2}\right)$$

and

$$\left|E\left(C_j|C_{j+1}\right) - \frac{C_{j+1}}{f_j}\right| \leq O\left(\sqrt{C_{j+1}\sigma_j^2}\right) = O_P(\sigma_j).$$

Now, let us consider the conditional variance $Var(C_j|C_{j+1}=v) = E(C_j^2|C_{j+1}=v) - (E(C_j|C_{j+1}=v))^2$. For computing $E(C_j^2|C_{j+1}=v)$, we have to solve the integral

$$\begin{aligned}
&\int_{\max(a,c(v))}^{\min(b,d(v))} u^2 \frac{1}{2\sqrt{u}} \frac{1}{\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}} du \\
&= \frac{1}{2\left(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))}\right)} \int_{\max(a,c(v))}^{\min(b,d(v))} u^{3/2} du \\
&= \frac{1}{2(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))})} \frac{2}{5} \left(\min(b,d(v))^{5/2} - \max(a,c(v))^{5/2}\right) \\
&= \frac{\min(b,d(v))^{5/2} - \max(a,c(v))^{5/2}}{5(\sqrt{\min(b,d(v))} - \sqrt{\max(a,c(v))})} \\
&= \frac{1}{5}\Big(\min(b,d(v))^2 + \min(b,d(v))^{3/2}\max(a,c(v))^{1/2} + \min(b,d(v))\max(a,c(v)) \\
&\qquad + \min(b,d(v))^{1/2}\max(a,c(v))^{3/2} + \max(a,c(v))^2\Big).
\end{aligned}$$

Then, altogether, the conditional variance computes to

$$\begin{aligned}
&Var(C_j|C_{j+1}=v) = E(C_j^2|C_{j+1}=v) - E(C_j|C_{j+1}=v)^2 \\
&= \frac{1}{45}\left(\sqrt{max(a,c(v))} - \sqrt{min(b,d(v))}\right)^2 \\
&\times \left(7\sqrt{max(a,c(v))}\sqrt{min(b,d(v))} + 4max(a,c(v)) + 4min(b,d(v)))\right).
\end{aligned}$$

Using that the variance is bounded by 0 from below and, going through all possible cases, from above by

$$0 \leq Var(C_j|C_{j+1}=v) \leq \frac{1}{45}\left(\sqrt{d(v)} - \sqrt{c(v)}\right)^2 \left(7d(v) + 4d(v) + 4d(v)\right),$$

where the upper bound is equal to

$$\frac{1}{45}\left(2\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2 15\left(\frac{v}{f_j}+\left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2+2\sqrt{\frac{v}{f_j}+\left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2}\frac{\sqrt{3\sigma_j^2}}{2f_j}+\left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2\right)$$
$$=\left(\frac{\sigma_j^2}{f_j^2}\right)\left(\frac{v}{f_j}+\left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2+\sqrt{\frac{v}{f_j}+\left(\frac{\sqrt{3\sigma_j^2}}{f_j}\right)^2}\frac{\sqrt{3\sigma_j^2}}{2f_j}+\left(\frac{\sqrt{3\sigma_j^2}}{2f_j}\right)^2\right).$$

Altogether, this leads to

$$Var(C_j|C_{j+1}=v)\leq\frac{\sigma_j^2 v}{f_j^3}+O(\sigma_j^4)+O\left(\sqrt{v\sigma_j^2}\sigma_j^2\right)$$

and

$$Var(C_j|C_{j+1})\leq\frac{\sigma_j^2 v}{f_j^3}+O_P\left(\sigma_j^3\right).$$

□

2.10.5 Proof of Theorem 2.20

Following the technique of proof in Theorem 2.17 and using Theorem 2.24 and Corollary 2.25 instead of Theorem 2.5 and Corollary 2.6, respectively, leads to the claimed results. Note that, making use of the law of total variance, asymptotic negligibility of the remainder terms $B_{2,K,I,n}$ and $B_{3,K,I,n}$ also conditional on $\mathcal{Q}_{I,\infty}$ follows directly from the corresponding unconditional result, which was already established in the proof of Theorem 2.17. □

2.10.6 Proof of Theorem 2.22

By adapting the technique of proof of Theorem 2.17 also for the proof of Theorem 2.13, it is straightforward to treat the two terms $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$ also jointly.

Whereas the stochastic independence of $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$ conditionally on $\mathcal{Q}_{I,\infty}$ follows immediately from the recursive construction of the loss triangle in Mack's model, for the unconditional case, the claimed uncorrelatedness follows from

$$
\begin{aligned}
&Cov\left((R_{I,n}-\widehat{R}_{I,n})_1, \sqrt{I+n+1}(R_{I,n}-\widehat{R}_{I,n})_2\right)\\
=&Cov\left(\sum_{i=0}^{I+n} C_{I-i,i}\left(\prod_{j=i}^{I+n-1} F_{I-i,j}-\prod_{j=i}^{I+n-1} f_j\right), \sqrt{I+n+1}\sum_{i=0}^{I+n} C_{I-i,i}\left(\prod_{j=i}^{I+n-1} f_j-\prod_{j=i}^{I+n-1} \widehat{f}_{j,n}\right)\right)\\
=&\sqrt{I+n+1}\sum_{i_1,i_2=0}^{I+n} E\left(C_{I-i_1,i_1}\left(\prod_{j_1=i_1}^{I+n-1} F_{I-i_1,j_1}-\prod_{j_1=i_1}^{I+n-1} f_{j_1}\right) C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1} f_{j_2}-\prod_{j_2=i_2}^{I+n-1} \widehat{f}_{j_2,n}\right)\right)\\
=&\sqrt{I+n+1}\\
&\times\sum_{i_1,i_2=0}^{I+n} E\left(C_{I-i_1,i_1}C_{I-i_2,i_2}\left(\prod_{j_2=i_2}^{I+n-1} f_{j_2}-\prod_{j_2=i_2}^{I+n-1} \widehat{f}_{j_2,n}\right) E\left(\left(\prod_{j_1=i_1}^{I+n-1} F_{I-i_1,j_1}-\prod_{j_1=i_1}^{I+n-1} f_{j_1}\right)\middle|\mathcal{D}_{I,\infty}\right)\right)\\
=&0,
\end{aligned}
$$

because the conditional expectation on the last right-hand side is zero. □

2.11 Additional simulation results

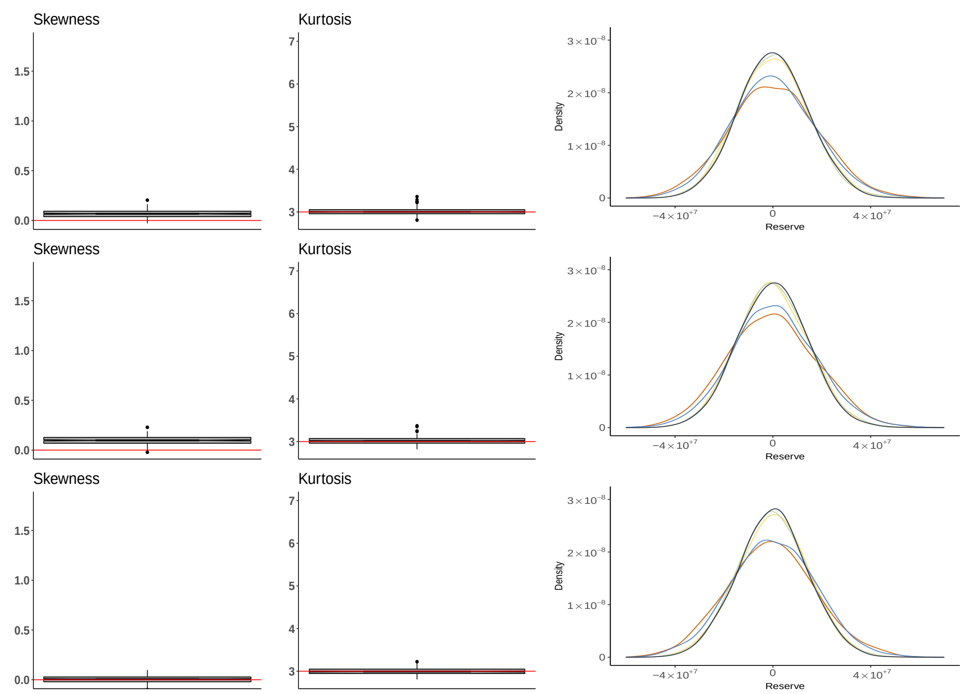

Figure 8: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n}-\widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n=0$ and $I=10$ for the setup of a), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

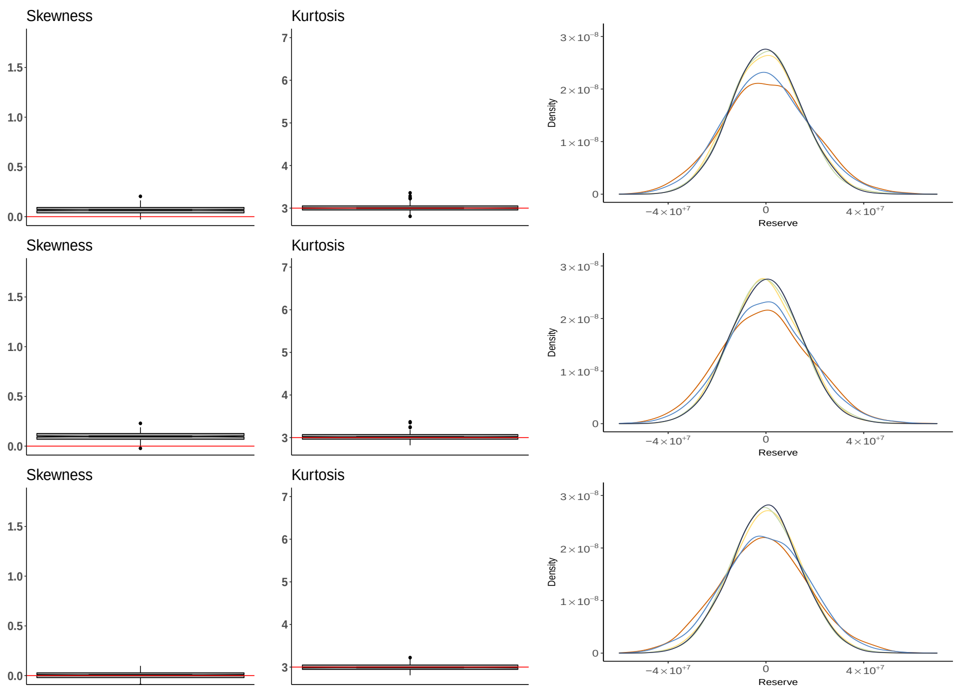

Figure 9: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 20$ and $I = 10$ for the setup of a), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

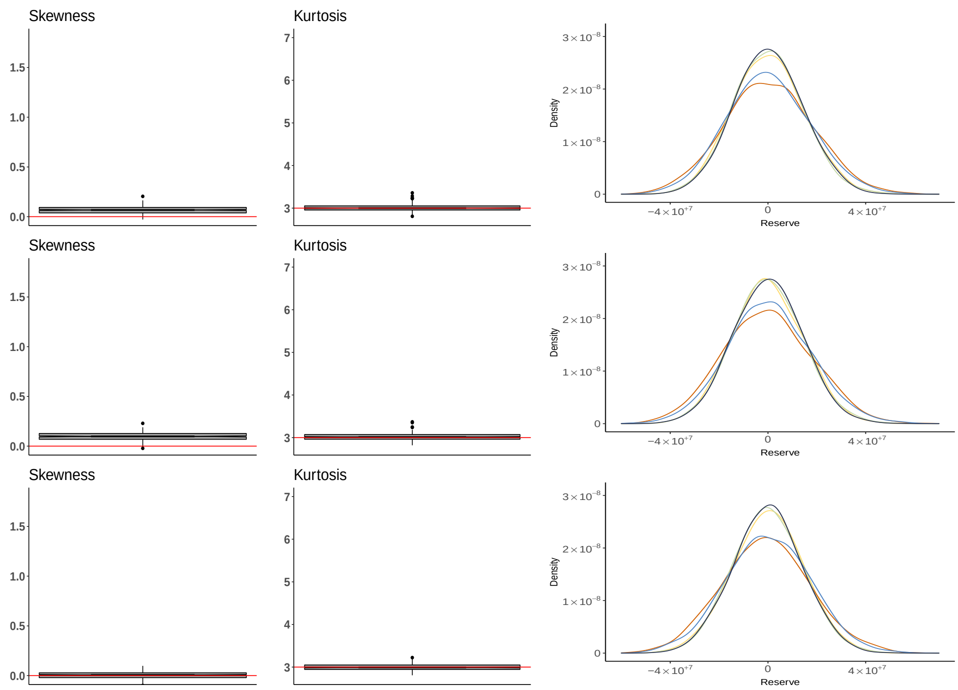

Figure 10: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 30$ and $I = 10$ for the setup of a), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

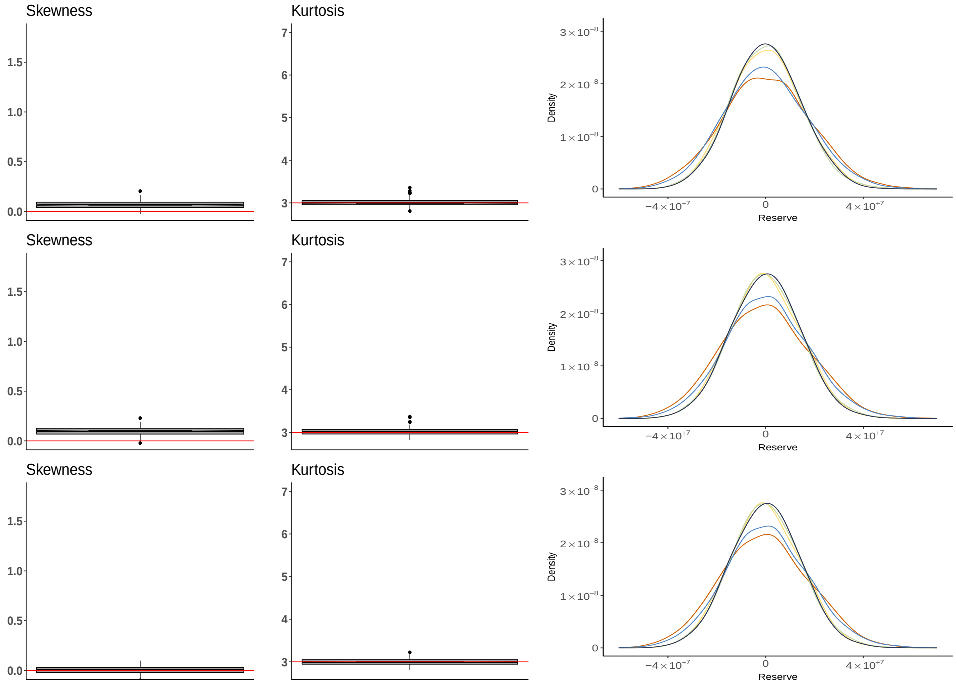

Figure 11: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 40$ and $I = 10$ for the setup of a), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

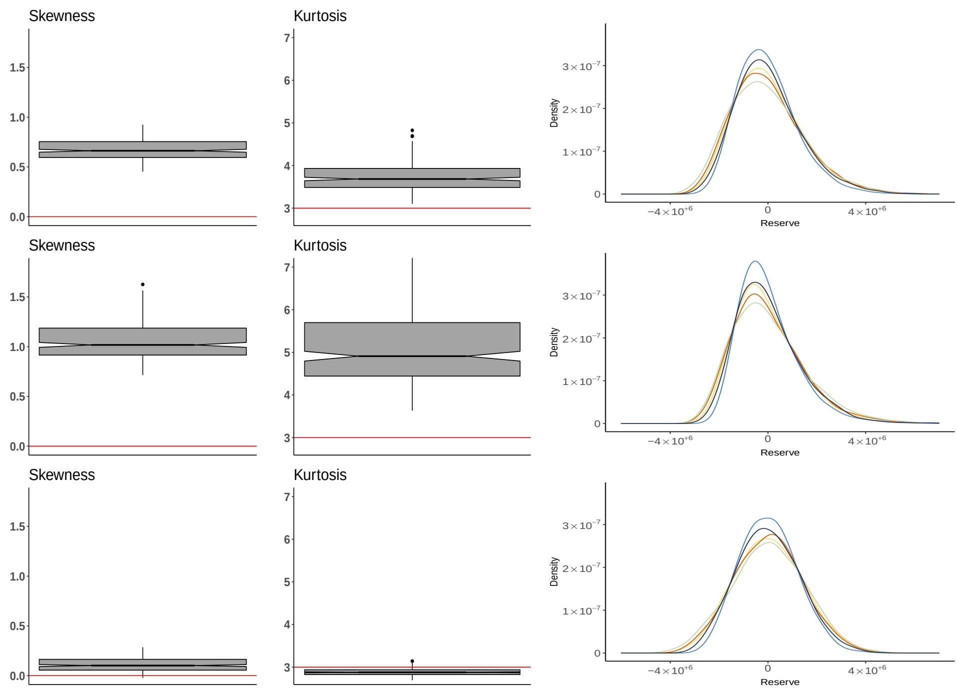

Figure 12: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 0$ and $I = 10$ for the setup of b), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

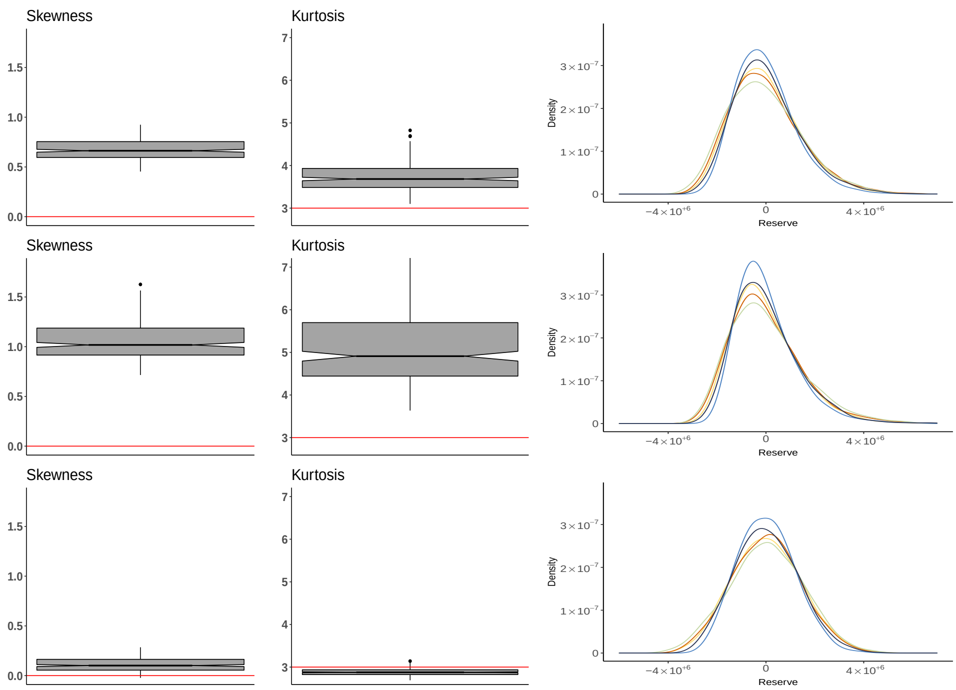

Figure 13: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 20$ and $I = 10$ for the setup of b), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

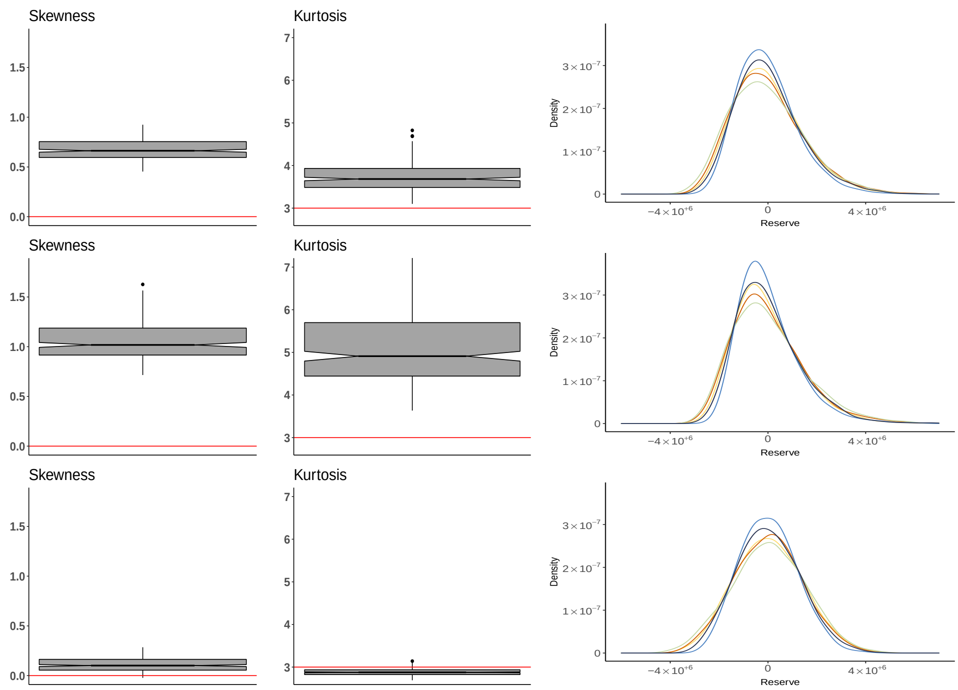

Figure 14: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 30$ and $I = 10$ for the setup of b), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

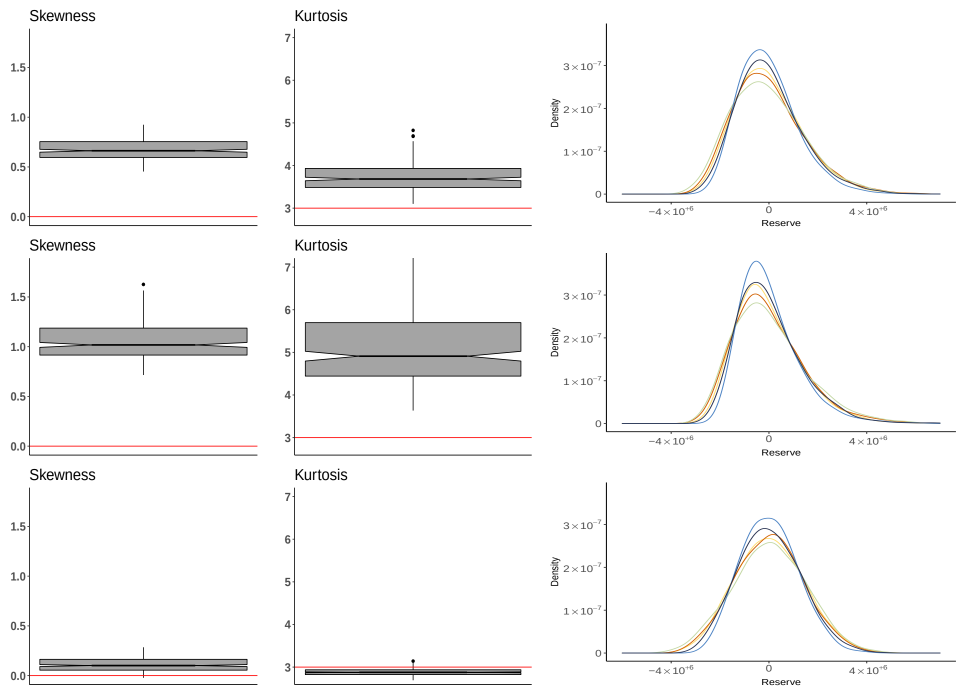

Figure 15: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated conditional distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}_{I,n}$ for $n = 40$ and $I = 10$ for the setup of b), where $F_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom). The red line indicates the benchmark skewness and kurtosis of a normal distribution.

3 Chapter

Bootstrap consistency for the Mack bootstrap

This chapter will be submitted for publication in due course.

Abstract
Mack's distribution-free chain ladder reserving model belongs to the most popular approaches in non-life insurance mathematics. While 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, which combines a residual-based non-parametric resampling step together with a parametric bootstrap. Although it is widely used in applications to estimate the reserve risk, no theoretical bootstrap consistency results exist that justify this approach.

In this paper, to fill this gap in the literature, we adopt the theoretical framework proposed by Steinmetz and Jentsch (2022) to derive asymptotic theory in Mack's model. By splitting the reserve risk into two additive parts corresponding to process and estimation uncertainty, it 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 the Mack bootstrap if the parametric family of distributions of the individual development factors is correctly specified in Mack's bootstrap proposal. 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 Mack 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-type bootstrap performs superior to the original Mack bootstrap in finite samples.

Keywords: Bootstrap consistency, loss reserving, Mack's model, Mack bootstrap, predictive inference

MSC subject classification: Primary 62E20, 62F40; secondary 62F12, 62P05

3.1 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 and to construct the (best) estimate of the reserve using a set of development factors and variance parameters. 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.

Also, frameworks based on general linear models (GLMs) considered in Renshaw and Verrall (1998), which make e.g. use of over-dispersed Poisson and Log-normal distributions, have been proposed for the calculation of the first two moments 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, since it does not allow 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 best estimate of the reserve does not suffice to pay for all outstanding claims. To get such insights, the knowledge about the *whole* distribution or at least of high quantiles of the reserve is inevitable. For this purpose, England and Verrall (2006) proposed the Mack bootstrap which equips Mack's model with a tailor-made bootstrap procedure that makes use of the recursive structure of CLMs and relies on additional parametric assumptions. Alternative bootstrap procedures for GLM-based setups have been addressed also in England (2002) and England and Verrall (1999, 2006) and Pinheiro et al. (2003). Björkwall et al. (2009) review these bootstrap techniques and suggest alternative non-parametric and parametric bootstrap procedures without providing any consistency results. Similarly, Björkwall et al. (2010) suggest also a bootstrap technique for the separation method, which takes calendar year effects into account. In recent years, bootstrap-based approaches have been favored by many actuaries, because such methods usually produce distributions that appear to be plausible in 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. However, the existing literature lacks a theoretical framework and rigorous asymptotic results that would allow to identify such scenarios a priori. In this regard, more 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. Alternatively, 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 more robust bootstrap procedure in a GLM setting based on M-estimators using influence functions and suggest to use weighted bootstrap resampling. Peters et al. (2010) compare the Mack bootstrap with a Bayesian bootstrap and show that their Bayesian approach gives the same results.

Nevertheless, a deeper and mathematically rigorous understanding of the original Mack bootstrap method and its underlying stochastic model is desirable to be able to justify the application of the Mack bootstrap. Recently, to enable a rigorous asymptotic treatment of Mack's model, Steinmetz and Jentsch (2022) proposed a suitable theoretical (stochastic and asymptotic) framework, which allows the derivation of conditional as well as unconditional asymptotic distributions of the reserve in Mack's model. Precisely, they split the reserve (centered around its best estimate) into two random additive parts, that carry the *process uncertainty* and the *estimation uncertainty*, respectively. This allows to derive limiting distributions for both parts of the reserve, when conditioning on the latest observed cumulative claims, but also unconditionally. In this regard, when addressing the question of bootstrap consistency for risk reserving, which is generally a prediction task, the *conditional* limiting distributions are crucial and serve as a benchmark for the corresponding Mack bootstrap distributions. Whereas the conditional limiting distribution of the second part, which corresponds to the estimation uncertainty, will be Gaussian under mild regularity conditions and when properly inflated, the conditional limiting distribution of the first part corresponding to the process uncertainty, 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 stochastic and asymptotic framework introduced in Steinmetz and Jentsch (2022) to rigorously investigate the long-standing question of Mack bootstrap consistency. Our contributions are twofold. First, to fill this gap in the literature, we derive bootstrap asymptotic theory for both parts of the (centered) Mack bootstrap reserve, that correspond to process uncertainty and estimation uncertainty, respectively. 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 proposal. 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 Mack bootstrap. Second, motivated 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 3.2 introduces the required notation and assumptions for the CLM, discusses parameter estimation in Mack's model, and provides the asymptotic and stochastic framework as introduced in Steinmetz and Jentsch (2022). In Section 3.3, we introduce the Mack bootstrap approach as originally proposed by England and Verrall (2006) and discuss its construction. In Section 3.4, we summarize the (conditional) asymptotic results from Steinmetz and Jentsch (2022) for the process uncertainty and estimation uncertainty terms in Section 3.4.1, which will serve as benchmark results for the Mack bootstrap results. Then, in Section 3.4.2, we derive bootstrap asymptotic theory for both parts of the (centered) Mack bootstrap reserve corresponding to process uncertainty and estimation uncertainty, respectively. For this purpose, asymptotic bootstrap theory for (smooth functions of) bootstrap development factor estimators has to be established, which might be of independent interest and can be found in the appendix. In Section 3.5, based on the findings from the previous section, we propose an alternative Mack-type bootstrap and derive its asymptotic properties in Section 3.6. We illustrate our findings in simulations in Section 3.7 and show that the newly proposed alternative Mack-type bootstrap performs superior to the original Mack bootstrap in finite samples. Section 3.8 concludes. All proofs, auxiliary results and additional simulations are deferred to the appendix.

3.2 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 4). 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

$$\mathcal{D}_I = \{C_{i,j} | i, j = 0, \dots, I,\ 0 \leq i + j \leq I\}. \tag{3.1}$$

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

Accident Year i	Development Year j					
	0	1	2	$\cdots$	$I-1$	I
0	$C_{0,0}$	$C_{0,1}$			$C_{0,I-1}$	$C_{0,I}$
1	$C_{1,0}$	$\cdots$			$C_{1,I-1}$	$C_{1,I}$
	$\cdots$		$\cdots$			
$\cdots$	$\cdots$					$\cdots$
				$\cdots$		
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$		$\cdots$		
I	$C_{I,0}$	$C_{I,1}$				$C_{I,I}$

Table 4: 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).

denote these diagonals by

$$\mathcal{Q}_k = \{C_{k-i,i} | i = 0, \dots, k\}. \tag{3.2}$$

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 4) summarizes the latest cumulative claim amounts collected in year I.

For the theoretical analysis of the prediction of the outstanding (unobserved) claims, it is useful to augment the (observed) *upper* loss triangle $\mathcal{D}_I$ by an unobserved *lower* triangle

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

that contains all future claims that have *not* been observed (yet) up to time I (green triangle in Table 4). The resulting cumulative claim matrix is denoted by $\mathcal{C}_I = (C_{i,j})_{i,j=0,\dots,I} = \mathcal{D}_I \cup \mathcal{D}_I^c$. 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

$$R_I = \sum_{i=0}^{I} R_{i,I}, \tag{3.3}$$

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 $C_{i,I}$ based on the observed upper triangle $\mathcal{D}_I$. Starting from the last observed claim $C_{i,I-i}$, this is done by predicting sequentially all future, yet (at time I)

Accident Year i	Development Year j: 0	1	2	$\cdots$	I	$I+1$	$I+2$
0	$C_{0,0}$	$C_{0,1}$		$\cdots$	$C_{0,I}$	$C_{0,I+1}$	$C_{0,I+2}$
1	$C_{1,0}$			$\cdots$		$C_{1,I+1}$	
2							
		$\cdots$					
	$\cdots$			$\cdots$			
$\cdots$	$\cdots$						
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$	$C_{I-1,2}$				
I	$C_{I,0}$	$C_{I,1}$	$C_{I,2}$				
$I+1$	$C_{I+1,0}$	$C_{I+1,1}$					
$I+2$	$C_{I+2,0}$						

Accident Year i	Development Year j: 0	1	2	$\cdots$	I	$I+1$	$I+2$
-2	$C_{-2,0}$			$\cdots$		$C_{-2,I+1}$	$C_{-2,I+2}$
-1	$C_{-1,0}$			$\cdots$		$C_{-1,I+1}$	
0	$C_{0,0}$	$C_{0,1}$		$\cdots$	$C_{0,I}$		
1	$C_{1,0}$	$\cdots$		$C_{1,I-1}$			
	$\cdots$		$\cdots$	$\cdots$			
$\cdots$	$\cdots$						
$I-1$	$C_{I-1,0}$	$C_{I-1,1}$					
I	$C_{I,0}$						

Table 5: 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 (based on Steinmetz and Jentsch (2022)).

unobserved claims $\{C_{i,j}|j = I - i + 1, \dots, I\}$. That is, by doing this for all accident years $i = 0, \dots, I$, the whole unobserved lower loss triangle $\mathcal{D}_I^c$ 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 meaningful limiting 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 is described in the following.

3.2.1 Asymptotic framework for reserve prediction

With the loss triangle $\mathcal{D}_I$ at hand, a conditional asymptotic analysis given the diagonal $\mathcal{Q}_I$, 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 for $I \to \infty$, where increasing I means adding new *diagonals* $\mathcal{Q}_{I+h} = \{C_{I-i,i}|i = 0, \dots, I+h\}$, $h \geq 1$ to the loss triangle $\mathcal{D}_I$ (see Table 5, 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 $\mathcal{D}_I$, that is, $\mathcal{Q}_I$, fixed and let $\mathcal{D}_I$ grow by adding new *rows* of cumulative claims $\{C_{-h,i}|i = 0, \dots, I+h\}$, $h \geq 1$ (Table 5, lower panel). Nevertheless, both versions of differently growing loss triangles indicated in Table 5 are equal in distribution.

Hence, in what follows, all asymptotic results are derived under the framework that we observe a sequence of (upper) loss triangles

$$\mathcal{D}_{I,n} = \{C_{i,j}|i = -n, \dots, I,\ j = 0, \dots, I+n,\ -n \leq i+j \leq I\}, \quad n \in \mathbb{N}_0, \tag{3.4}$$

where $\mathbb{N}_0 = \{0, 1, 2, \dots\}$, with (main) diagonals

$$\mathcal{Q}_{I,n} = \{C_{I-i,i}|i = 0, \dots, I+n\}, \quad n \in \mathbb{N}_0. \tag{3.5}$$

Note that $\mathcal{D}_{I,0} = \mathcal{D}_I$, $\mathcal{Q}_{I,0} = \mathcal{Q}_I$ and $\mathcal{D}_{I,n}$ (and $\mathcal{Q}_{I,n}$) is obtained by sequentially adding n rows of lengths $I+2, I+3, \dots, I+n+1$, respectively, on top to $\mathcal{D}_I$ (see Table 5, lower panel). As before, for all $n \in \mathbb{N}_0$, we augment the (observed) *upper* loss triangle $\mathcal{D}_{I,n}$ by an unobserved *lower* triangle $\mathcal{D}_{I,n}^c = \{C_{i,j}|i = -n, \dots, I,\ j = 0, \dots, I+n,\ i+j > I\}$ that contains all future claims that have not been observed (yet) up to time I. Further, according to (3.3), the aggregated total amount of the reserve is denoted by

$$R_{I,n} = \sum_{i=-n}^{I} R_{i,I+n}, \quad n \in \mathbb{N}_0, \tag{3.6}$$

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 3.3, we let $n \to \infty$ to derive the limiting distribution of the reserve in Section 3.4. According to (3.4) and (3.5), the limiting upper loss triangle $\mathcal{D}_{I,\infty}$ and the limiting diagonal $\mathcal{Q}_{I,\infty}$ are defined by

$$\mathcal{D}_{I,\infty} = \{C_{i,j}|i \in \mathbb{Z}, i \leq I,\ j \in \mathbb{N}_0,\ i+j \leq I\} \quad \text{and} \quad \mathcal{Q}_{I,\infty} = \{C_{I-i,i}|i \in \mathbb{N}_0\}, \tag{3.7}$$

respectively.

3.2.2 Mack's distribution-free chain ladder reserving

The distribution-free chain ladder model proposed by Mack (1993), often denoted simply as Mack's model, is widely used in practice to determine both the mean and the variance of the reserve. By adopting the notion of the asymptotic framework described in Section 3.2.1, the conditions of Mack's model originally proposed in Mack (1993) can be summarized as follows.

Assumption 3.1 (Mack's Model) *For any $n \in \mathbb{N}_0$, let $\mathcal{C}_{I,n} = (C_{i,j}, i = -n, \dots, I,\ j = 0, \dots, I+n)$ denote random variables on some probability space $(\Omega, \mathcal{A}, P)$ and suppose the following holds:*

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

$$E(C_{i,j+1}|C_{i,j}) = f_j C_{i,j} \tag{3.8}$$

for $i = -n, \dots, I$ and $j = 0, \dots, I+n-1$.

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

$$Var(C_{i,j+1}|C_{i,j}) = \sigma_j^2 C_{i,j} \tag{3.9}$$

for $i = -n, \dots, I$ and $j = 0, \dots, I+n-1$.

(iii) The cumulative payments are stochastically independent over the accident years $i = -n, \dots, I$, that is, the cumulative claim matrix $\mathcal{C}_{I,n}$ 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 $\mathcal{D}_{I,n}$, 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 $\mathcal{D}_{I,n}$. As proposed by the CLM, the development factors $f_0, \dots, f_{I+n-1}$ can be (consistently) estimated by $\widehat{f}_{0,n}, \dots, \widehat{f}_{I+n-1,n}$, where

$$\widehat{f}_{j,n} = \frac{\sum_{i=-n}^{I-j-1} C_{i,j+1}}{\sum_{i=-n}^{I-j-1} C_{i,j}}, \quad j = 0, \dots, I+n-1. \tag{3.10}$$

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

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

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

$$\widehat{R}_{i,I+n} = \widehat{C}_{i,I+n} - C_{i,I-i} = C_{i,I-i} \left(\prod_{j=I-i}^{I+n-1} \widehat{f}_{j,n} - 1 \right), \quad i = -n, \ldots, I, \tag{3.11}$$

and the best estimate $\widehat{R}_{I,n}$ of the total reserve $R_{I,n}$ defined in (3.6) computes to

$$\widehat{R}_{I,n} = \sum_{i=-n}^{I} \widehat{R}_{i,I+n} \tag{3.12}$$

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

$$\widehat{\sigma}_{j,n}^2 = \frac{1}{I+n-j-1} \sum_{i=-n}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_{j,n} \right)^2, \quad j = 0, \ldots, I+n-2, \tag{3.13}$$

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

Of particular interest is the difference of the stochastic (unobserved) reserve $R_{I,n}$ and its best estimate $\widehat{R}_{I,n}$ (based on the observed data $\mathcal{D}_{I,n}$), which is denoted as the *predictive root of the reserve* in the following. That is, by combining (3.6) and (3.12), it computes to

$$R_{I,n} - \widehat{R}_{I,n} = \sum_{i=-n}^{I} \left(R_{i,I+n} - \widehat{R}_{i,I+n} \right) = \sum_{i=-n}^{I} \left(C_{i,I+n} - \widehat{C}_{i,I+n} \right). \tag{3.14}$$

By adopting the notion of the asymptotic framework from Section 3.2.1, we define the mean-squared error of prediction (MSEP) of $\widehat{R}_{I,n}$ given $\mathcal{D}_{I,n}$ by $MSEP(\widehat{R}_{I,n}|\mathcal{D}_{I,n}) = E((R_{I,n} - \widehat{R}_{I,n})^2|\mathcal{D}_{I,n})$ going back to Mack (1993), who derived the formula

$$MSEP(\widehat{R}_{I,n}|\mathcal{D}_{I,n}) = \sum_{i=-n}^{I} \left(\widehat{C}_{i,I+n}^2 \sum_{j=I-i}^{I+n-1} \frac{\widehat{\sigma}_{j,n}^2}{\widehat{f}_{j,n}^2} \frac{1}{\widehat{C}_{i,j}} + \widehat{C}_{i,I+n}^2 \sum_{j=I-i}^{I+n-1} \frac{\widehat{\sigma}_{j,n}^2}{\widehat{f}_{j,n}^2} \frac{1}{\sum_{k=-n}^{I-j-1} C_{k,j}} \right) \tag{3.15}$$

$$+ 2 \sum_{\substack{i,l=-n \\ i<l}}^{I} \left(\widehat{C}_{i,I+n} \widehat{C}_{l,I+n} \sum_{j=I-i}^{I+n-1} \frac{\widehat{\sigma}_{j,n}^2}{\widehat{f}_{j,n}^2} \frac{1}{\sum_{k=-n}^{I-j-1} C_{k,j}} \right), \tag{3.16}$$

where $\widehat{C}_{i,j} = C_{i,I-i} \widehat{f}_{I-i,n} \cdots \widehat{f}_{j-1,n}$ for $j > I-i$ are the estimated values of the future claims $C_{i,j}$ and $\widehat{C}_{i,I-i} = C_{i,I-i}$. In the above, the summands of the first sum on the right-hand side in (3.15) consist of two terms corresponding to the *process variance* and *estimation variance* (of parameter estimates) of $R_{i,I+n}$, respectively. The second expression in (3.16) reflects the linear dependence between $\widehat{R}_{i,I+n}$ and $\widehat{R}_{l,I+n}$, $i \neq l$ and contains their covariances. The term contributes also to the *estimation variance*. Alternatively, the MSEP of $\widehat{R}_{I,n}$ given $\mathcal{D}_{I,n}$ can be rewritten as

$$\begin{aligned}
&MSEP\left(\widehat{R}_{I,n}|\mathcal{D}_{I,n}\right)\\
&= \sum_{i=-n}^{I}\left(C_{i,I-i}\sum_{j=I-i}^{I+n-1}\widehat{\sigma}^2_{j,n}\prod_{k=I-i}^{j-1}\widehat{f}_{k,n}\prod_{l=j+1}^{I+n-1}\widehat{f}^2_{l,n} + C^2_{i,I-i}\sum_{j=I-i}^{I+n-1}\frac{\widehat{\sigma}^2_{j,n}}{\sum_{k=-n}^{I-j-1}C_{k,j}}\prod_{\substack{l=I-i\\ l\neq j}}^{I+n-1}\widehat{f}^2_{l,n}\right)\\
&\quad + 2\sum_{\substack{i,l=-n\\ i<l}}^{I}\left(C_{i,I-i}C_{l,I-l}\sum_{j=I-i}^{I+n-1}\frac{\widehat{\sigma}^2_{j,n}}{\sum_{k=-n}^{I-j-1}C_{k,j}}\prod_{m=I-l}^{I-i-1}\widehat{f}_{m,n}\prod_{\substack{p=I-i\\ p\neq j}}^{I+n-1}\widehat{f}^2_{p,n}\right).
\end{aligned}$$

While Mack's formula reflects the negative correlation of $\widehat{f}^2_{j,n}$ and $\widehat{f}^2_{k,n}$ for $j \neq k$, Buchwalder et al. (2006) derived another formula for the MSEP of the reserve, where they modeled $C_{i,j}$ using an $AR(1)$ time series, which leads to independence of $\widehat{f}_{j,n}$ and $\widehat{f}_{k,n}$ for $j \neq k$. Hence, the formulas for the MSEP of Buchwalder et al. (2006) differs from the formulas above, but in application the differences are small.

Mack's model setup as originally proposed in Mack (1993) and presented above allows to estimate the MSEP of $\widehat{R}_{I,n}$ given $\mathcal{D}_{I,n}$, but this will be generally not sufficient to determine the whole distribution or high quantiles of the reserve, which are of particular interest to approximate the *reserve risk* as e.g. the value-at-risk. While a common solution to such problems is the derivation of asymptotic theory to approximate the unknown (finite sample) distributions, Mack's conditions summarized in Assumption 3.1 are not sufficient to establish limiting distributions. For this purpose, as summarized in the next section, Steinmetz and Jentsch (2022) proposed a suitable stochastic model which slightly strengthens Mack's model assumptions and enables the derivation of (conditional and unconditional) limiting distributions for the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$ defined in (3.14).

3.2.3 A fully-described stochastic framework of Mack's Model

Following Steinmetz and Jentsch (2022, Section 2.2), to establish a stochastic framework sufficient to be able to derive asymptotic theory for parameter estimators $\widehat{f}_{j,n}$ and $\widehat{\sigma}^2_{j,n}$, 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 3.2, 3.3, and 3.5 on the stochastic mechanism that generates the cumulative claim matrix $\mathcal{C}_{I,n}$, and Assumption 3.4 on the sequences of development factors and of variance parameters. They resemble the Assumptions 2.2, 2.3 and 3.3 in Steinmetz and Jentsch (2022), when adopted to the asymptotic framework of Section 3.2.1, as well as Assumption 4.1 in Steinmetz and Jentsch (2022), respectively. This framework also allows to rigorously investigate bootstrap consistency properties of the Mack bootstrap in Section 3.4.

The first assumption addresses the initial claims, i.e. the first column $C_{\bullet,0} = (C_{-n,0}, \dots, C_{I,0})'$ of $\mathcal{C}_{I,n}$ (and of $\mathcal{D}_{I,n}$).

Assumption 3.2 (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* $\tau_0^2 := Var(C_{i,0}) \in (0, \infty)$*.*

Note that the independence between the initial claims is a common assumption which is also a direct consequence of Assumption 3.1 (iii). In addition, Assumption 3.2 also imposes an identical distribution for the initial claims. In practice, the condition on the support $[1, \infty)$ of $C_{i,0}$ is not restrictive and can be relaxed to the condition that $C_{i,0}$ is bounded away from zero.

In view of the multiplicative structure of $E(C_{i,j+1}|C_{i,j})$ in (3.8), suppose that the (random) cumulative claims $C_{i,j+1}$, $i = -n, \dots, I$, $j = 0, \dots, I+n-1$, are recursively defined by

$$C_{i,j+1} = C_{i,j} F_{i,j} = C_{i,0} \prod_{k=0}^{j} F_{i,k}, \tag{3.17}$$

where the *individual development factors* $F_{i,j}$, which satisfy $F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$ by construction, are assumed to fulfill the following condition.

Assumption 3.3 (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 conditional mean and conditional variance*

$$E(F_{i,j}|C_{i,j}) = f_j \quad \text{and} \quad Var(F_{i,j}|C_{i,j}) = \frac{\sigma_j^2}{C_{i,j}}. \tag{3.18}$$

Note that Mack's original model setup in Assumption 2.1 is implied by Assumptions 3.2 and 3.3 together. Also note that the stochastic mechanism determined by (3.17) and Assumption 3.3 are assumed for the whole cumulative claim matrix $\mathcal{C}_{I,n}$. However, recall that only those $C_{i,j}$ in $\mathcal{C}_{I,n}$ are observed up to year I that are contained in the upper loss triangle $\mathcal{D}_{I,n}$. Hence, by using the multiplicative relationship in the first identity of (3.17), we have also *perfect* knowledge of the individual development factors $F_{i,j}$, $i=-n,\dots,I-1$, $j=0,\dots,I-i-1$.

According to Lemma 2.4 in Steinmetz and Jentsch (2022), the stochastic framework determined by Assumptions 3.2 and 3.3 allows to derive formulas for the (unconditional) means and variances of $C_{i,j}$, $i=-n,\dots,I$, $j=0,\dots,I+n$ leading to

$$E(C_{i,j}) = \mu_0 \prod_{k=0}^{j-1} f_k =: \mu_j, \tag{3.19}$$

$$Var(C_{i,j}) = \tau_0^2 \prod_{k=0}^{j-1} f_k^2 + \mu_0 \sum_{l=0}^{j-1} \sigma_l^2 \prod_{n=l+1}^{j-1} f_n^2 \prod_{m=0}^{l-1} f_m =: \tau_j^2, \tag{3.20}$$

where μ_0 and τ_0^2 are defined in Assumption 3.2. Together with Assumption 3.4 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 $\mu_j \to \mu_\infty$ and $\tau_j^2 \to \tau_\infty^2$ as $j\to\infty$, 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 3.4 (Development Factors and Variance Parameters) *Letting $n\to\infty$ in the setup of Assumptions 3.2 and 3.3 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 $f_j\to 1$ as $j\to\infty$ such that $\prod\limits_{j=0}^{\infty} f_j < \infty$, which is equivalent to $\sum\limits_{j=0}^{\infty}(f_j-1)<\infty$.

(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 $\sigma_j^2\to 0$ as $j\to\infty$ such that $\sum\limits_{j=0}^{\infty}(j+1)^2\sigma_j^2<\infty$.

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 3.4 are rather mild. In practice, each claim has a finite, but possibly unknown horizon until it is finally settled. The time horizons of claim developments vary by the insurance lines, which are usually categorized in short-term and long-term. 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 $\widehat{f}_{j,n}$ for $n\to\infty$.

For the derivation of a similar CLT result for $\widehat{\sigma}_{j,n}^2$ for fixed j, according to Steinmetz and Jentsch (2022, Section 3.2), the following additional assumption is imposed. However, although the distributional properties of $\widehat{\sigma}_{j,n}^2$ do not show asymptotically in the distribution of the reserve, $\sqrt{I+n}$-consistency of $\widehat{\sigma}_{j,n}^2$ as obtained in Steinmetz and Jentsch (2022, Theorem 3.5) is required for establishing the bootstrap asymptotic theory in Section 3.4.

Assumption 3.5 (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 for $j = 0, \ldots, I+n-1$ such that both*

$$\kappa_j^{(3)} = E(C_{i,j}^2 E((F_{i,j} - f_j)^3|C_{i,j})) \quad \text{and} \quad \kappa_j^{(4)} = E(C_{i,j}^2 E((F_{i,j} - f_j)^4|C_{i,j})) \tag{3.21}$$

exist and are finite, respectively.

Using the recursive stochastic model for the claims in (3.17), conditional on $\mathcal{Q}_{I,n}$, the reserve $R_{I,n}$ can be written as

$$R_{I,n} = \sum_{i=-n}^{I} C_{i,I-i} \left(\prod_{j=I-i}^{I+n-1} F_{i,j} - 1 \right). \tag{3.22}$$

Hence, by plugging-in (3.11) and (3.22), the predictive root of the reserve from (3.14) becomes

$$R_{I,n} - \widehat{R}_{I,n} = \sum_{i=-n}^{I} C_{i,I-i} \left(\prod_{j=I-i}^{I+n-1} F_{i,j} - \prod_{j=I-i}^{I+n-1} \widehat{f}_{j,n} \right) = \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,i} - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right), \tag{3.23}$$

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

3.3 Mack's Bootstrap Scheme

The Mack bootstrap proposal, introduced by England and Verrall (2006), equips Mack's model with a resampling procedure to estimate the whole distribution of the (predicted) reserve. The resulting Mack bootstrap is very popular and widely used in practice as it describes a rather simple to implement algorithm to estimate the *reserve risk* such as e.g. the value-at-risk, by estimating high quantiles of the distribution of the reserve.

As proposed by England and Verrall (2006), to mimic the distribution of the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$, the Mack bootstrap constructs a certain bootstrap version $R_{I,n}^* - \widehat{R}_{I,n}$ of it. On the one hand, this bootstrap predictive root relies on the *best estimate* of the reserve and centers $R_{I,n}^*$ around $\widehat{R}_{I,n}$. This is motivated by the definition of the reserve risk, which captures the risk that the best estimate of the reserve is not sufficient to pay for all outstanding claims. On the other hand, it constructs a (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 $F^*_{i,j}$, $j = 0, \ldots, I+n-1$, $i = -n, \ldots, I-j-1$, that is, for the *upper triangle*, in order to get bootstrap development factor estimators $\widehat{f}^*_{j,n}$, $j = 0, \ldots, I+n-1$.

(ii) Second, the bootstrap development factor estimators $\widehat{f}^*_{j,n}$ from (i) together with a *parametric* bootstrap (see Step 5 below) are used to construct also bootstrap individual development factors $F^*_{i,j}$, $i = -n, \ldots, I$, $j = 0, \ldots, I+n-1$ and $i+j \geq I$, that is for the *lower triangle*. For this purpose, a parametric family of (conditional) bootstrap distributions (such as e.g. a gamma or a log-normal distribution) has to be chosen. In contrast to the non-parametric approach in (i), the parametric approach is favored here to assure that $F^*_{i,j} > 0$ (a.s.).

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 $R_{I,n} - \widehat{R}_{I,n}$ given $\mathcal{Q}_{I,n}$ by the conditional bootstrap distribution of $R^*_{I,n} - \widehat{R}_{I,n}$ given $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$.

3.3.1 Mack's Bootstrap Algorithm

With the upper triangle $\mathcal{D}_{I,n}$ at hand, Mack's bootstrap algorithm is defined as follows:

Step 1. Estimate the development factors f_j and the variance parameters σ^2_j from $\mathcal{D}_{I,n}$ by computing $\widehat{f}_{j,n}$ and $\widehat{\sigma}^2_{j,n}$ for $j = 0, \ldots, I+n-1$ as defined in (3.10) and (3.13), respectively.

Step 2. For all $j = 0, \ldots, I+n-1$ with $\widehat{\sigma}^2_{j,n} > 0$, compute 'residuals'

$$\widehat{r}_{i,j} = \frac{\sqrt{C_{i,j}}(F_{i,j} - \widehat{f}_{j,n})}{\widehat{\sigma}_{j,n}} \tag{3.24}$$

for $i = -n, \ldots, I-j-1$. Re-center and re-scale these $\widehat{r}_{i,j}$'s to get

$$\widetilde{r}_{i,j} = \frac{1}{s}\left(\widehat{r}_{i,j} - \overline{r}\right),$$

where[1]

$$\overline{r} = \frac{2}{(I+n+1)(I+n)-2} \sum_{k=0}^{I+n-2} \sum_{l=-n}^{I-k-1} \widehat{r}_{l,k}, \tag{3.25}$$

$$s^2 = \frac{2}{(I+n+1)(I+n)-2} \sum_{k=0}^{I+n-2} \sum_{l=-n}^{I-k-1} \left(\widehat{r}_{l,k} - \overline{r}\right)^2. \tag{3.26}$$

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

Step 4. Define the bootstrap individual development factors

$$F^*_{i,j} = \widehat{f}_{j,n} + \frac{\widehat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} r^*_{i,j} \tag{3.27}$$

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

$$\widehat{f}^*_{j,n} = \frac{\sum_{i=-n}^{I-j-1} C_{i,j} F^*_{i,j}}{\sum_{i=-n}^{I-j-1} C_{i,j}} = \widehat{f}_{j,n} + \frac{\widehat{\sigma}_{j,n} \sum_{i=-n}^{I-j-1} \sqrt{C_{i,j}} r^*_{i,j}}{\sum_{i=-n}^{I-j-1} C_{i,j}} \tag{3.28}$$

for $j = 0, \ldots, I+n-1$.

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

$$E^*(F^*_{i,j} | C^*_{i,j}, \mathcal{F}^*_{I,n}) = \widehat{f}^*_{j,n}, \quad Var^*(F^*_{i,j} | C^*_{i,j}, \mathcal{F}^*_{I,n}) = \frac{\widehat{\sigma}^2_{j,n}}{C^*_{i,j}}$$

for $i = -n, \ldots, I,\ j = I-i, \ldots, I+n-1$, where $E^*(\cdot) := E^*(\cdot | \mathcal{D}_{I,n})$, $Var^*(\cdot) := Var^*(\cdot | \mathcal{D}_{I,n})$, etc. denote the Mack bootstrap mean, Mack bootstrap variance, etc., respectively, that is, conditional on the data $\mathcal{D}_{I,n}$. Then, given $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$, 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

[1] Note that $\widehat{\sigma}^2_{I+n-1,n} = 0$ by construction such that $\widehat{r}_{-n,I+n-1}$ is excluded in (3.24) such that (at most) $(I+n+1)(I+n)/2 - 1 = ((I+n+1)(I+n)-2)/2$ residuals can be computed. If $\widehat{\sigma}^2_{j,n} = 0$ holds also for other j, the corresponding residuals are excluded in (3.24) as well and the formulas for $\overline{r}$ and s in (3.25) and (3.26), respectively, have to be adjusted accordingly. In the following, for notational convenience, we assume that only $\widehat{\sigma}^2_{I+n-1,n} = 0$ and $\widehat{\sigma}^2_{j,n} > 0$ holds for all $j = 0, \ldots, I+n-2$ and all $n \in \mathbb{N}_0$.

$i = -n, \ldots, I$ using the recursion

$$C^*_{i,j+1} = C^*_{i,j} F^*_{i,j} \tag{3.29}$$

for $j = I - i, \ldots, I + n - 1$.

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

$$R^*_{I,n} - \widehat{R}_{I,n} = \sum_{i=0}^{I+n} C^*_{I-i,i} \left(\prod_{j=i}^{I+n-1} F^*_{I-i,i} - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right). \tag{3.30}$$

Step 7. Repeat Steps 3 - 6 above B times, where B is large, to get bootstrap predictive roots $(R^*_{I,n} - \widehat{R}_{I,n})^{(b)}$, $b = 1, \ldots, B$, 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

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

Remark 3.6 (On Mack's bootstrap proposal)

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

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

(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 $\widehat{f}^*_{j,n}$ and $\widehat{f}^*_{k,n}$ defined in (3.28) *are independent for $j \neq k$ conditional on $\mathcal{D}_{I,n}$. This is on contrast to the development factor estimators $\widehat{f}_{j,n}$ and $\widehat{f}_{k,n}$, which are asymptotically independent for $j \neq k$, but only uncorrelated in finite samples such that $E(\widehat{f}^2_{j,n}\widehat{f}^2_{k,n}) < 0$ for $j \neq k$. Whether the development factors are independent is also reflected in the formula for the MSEP of the reserve. Mack's formula takes into account the uncorrelatedness of $\widehat{f}_{j,n}$ and $\widehat{f}_{k,n}$, whereas in the formula of the MSEP by Buchwalder et al. (2006) the estimates of the development factors are independent.*

(iii) *The non-parametric bootstrap used to construct the $\widehat{f}^*_{j,n}$'s in Step 4 uses residuals, but according to Assumption 3.2 and 3.3, the claims $C_{i,j}$ are defined recursively such that there are no errors in the Mack model setup 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 $\widehat{f}_{j,n}$'s.*

(iv) *While $\widehat{f}^*_{j,n} > 0$ is not guaranteed in finite samples by the non-parametric bootstrap proposal in Step 4, the parametric bootstrap used to construct the $F^*_{i,j}$'s in Step 5 assures that all individual development factors $F^*_{i,j}$ are a.s. positive. However, the generation of $F^*_{i,j}$ requires $\widehat{f}^*_{j,n} > 0$ to hold.*

(v) *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 and, consequently, the $\widehat{f}^*_{j,n}$'s can be used. A fully non-parametric approach that uses the non-parametric bootstrap from Step 4 also in Step 5 is thinkable, but suffers from issues arising from potentially negative $F^*_{i,j}$'s leading to a reduced finite sample performance.*

3.4 Asymptotic Theory for the Mack Bootstrap

Although the Mack bootstrap as proposed by England and Verrall (2006) and described in Section 3.3 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 3.2, 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 (3.23), 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

$$R_{I,n} - \widehat{R}_{I,n} = \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} F_{I-i,j} - \prod_{j=i}^{I+n-1} f_j \right) + \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} f_j - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right)$$
$$=: \left(R_{I,n} - \widehat{R}_{I,n}\right)_1 + \left(R_{I,n} - \widehat{R}_{I,n}\right)_2, \tag{3.31}$$

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 (3.30), by subtracting and adding $\sum\limits_{i=0}^{I+n} C^*_{I-i,i} \prod_{j=i}^{I+n-1} \widehat{f}^*_{j,n}$, we get

$$R^*_{I,n} - \widehat{R}_{I,n} = \sum_{i=0}^{I+n} C^*_{I-i,i} \left(\prod_{j=i}^{I+n-1} F^*_{I-i,j} - \prod_{j=i}^{I+n-1} \widehat{f}^*_{j,n} \right) + \sum_{i=0}^{I+n} C^*_{I-i,i} \left(\prod_{j=i}^{I+n-1} \widehat{f}^*_{j,n} - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right)$$
$$=: \left(R^*_{I,n} - \widehat{R}_{I,n}\right)_1 + \left(R^*_{I,n} - \widehat{R}_{I,n}\right)_2, \tag{3.32}$$

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 (3.31) and (3.32), it is instructive to 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, as well as jointly. They will serve as valuable benchmark distributions to enable a meaningful investigation of the consistency properties of the Mack bootstrap in Section 3.4.2. 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 3.4.1.

3.4.1 Conditional asymptotics for the predictive root of the reserve

Based on the same stochastic and asymptotic framework used throughout this paper, Steinmetz and Jentsch (2022) established asymptotic theory for both parts of the predictive root of the reserve $R_{I,n} - \widehat{R}_{I,n}$, that is, for $(R_{I,n} - \widehat{R}_{I,n})_1$ and $(R_{I,n} - \widehat{R}_{I,n})_2$ separately, as well as jointly for $R_{I,n} - \widehat{R}_{I,n}$. As risk reserving relies on the *prediction* of $R_{I,n}$ using $\widehat{R}_{I,n}$, which is computed from the observed upper loss triangle $\mathcal{D}_{I,n}$ which consists of all cumulative claims up to the diagonal $\mathcal{Q}_{I,n}$, the main interest is in asymptotic theory for $(R_{I,n} - \widehat{R}_{I,n})_1$, $(R_{I,n} - \widehat{R}_{I,n})_2$, and $R_{I,n} - \widehat{R}_{I,n}$ conditional on $\mathcal{Q}_{I,n}$, respectively.

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 3.4.1.1, for the estimation uncertainty term $(R_{I,n} - \widehat{R}_{I,n})_2$ in Section 3.4.1.2, and jointly for $R_{I,n} - \widehat{R}_{I,n}$ in Section 3.4.1.3, respectively.

3.4.1.1 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 3.7 (Asymptotics for $(R_{I,n} - \widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}$) *Suppose Assumptions 3.2, 3.3 and 3.4 hold. Then, as $n \to \infty$, 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*

$$E\left(\left((R_{I,n} - \widehat{R}_{I,n})_1 - (R_{I,\infty} - \widehat{R}_{I,\infty})_1\right)^2 |\mathcal{Q}_{I,n}\right) \xrightarrow{p} 0, \tag{3.33}$$

where

$$(R_{I,\infty} - \widehat{R}_{I,\infty})_1 := \sum_{i=0}^{\infty} C_{I-i,i}\left(\prod_{j=i}^{\infty} F_{I-i,j} - \prod_{j=i}^{\infty} f_j\right) \sim \mathcal{G}_1. \tag{3.34}$$

Conditional on $\mathcal{Q}_{I,\infty}$, the (limiting) distribution $\mathcal{G}_1$ has mean zero, i.e. $E((R_{I,\infty} - \widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}) = 0$, and variance

$$Var\left((R_{I,\infty} - \widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}\right) = \sum_{i=0}^{\infty} C_{I-i,i}\sum_{j=i}^{\infty}\left(\prod_{k=i}^{j-1} f_k\right)\sigma_j^2\left(\prod_{l=j+1}^{\infty} f_l^2\right) = O_P(1). \tag{3.35}$$

The (conditional) L_2-convergence result in Theorem 3.7 immediately implies also (conditional) convergence in distribution. That is, for $n \to \infty$, we have

$$(R_{I,n} - \widehat{R}_{I,n})_1|\mathcal{Q}_{I,n} \xrightarrow{d} (R_{I,\infty} - \widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty} \sim \mathcal{G}_1|\mathcal{Q}_{I,\infty}. \tag{3.36}$$

Moreover, according to Theorem 3.7 (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 $F_{i,j}|C_{i,j}$.

3.4.1.2 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 3.7, 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 $\mathcal{Q}_{I,n}$[2], it is instructive to further decompose $(R_{I,n} - \widehat{R}_{I,n})_2$ to get

$$\begin{aligned}
&(R_{I,n} - \widehat{R}_{I,n})_2 \\
&= \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} f_j - \prod_{j=i}^{I+n-1} f_{j,n}(\mathcal{Q}_{I,n}) \right) + \sum_{i=0}^{I+n} C_{I-i,i} \left(\prod_{j=i}^{I+n-1} f_{j,n}(\mathcal{Q}_{I,n}) - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right) \\
&= (R_{I,n} - \widehat{R}_{I,n})_2^{(1)} + (R_{I,n} - \widehat{R}_{I,n})_2^{(2)}, \qquad (3.37)
\end{aligned}$$

where $(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ is measurable with respect to $\mathcal{Q}_{I,n}$ 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

$$\begin{aligned}
\mu_{j+1,n}^{(1)}(\mathcal{Q}_{I,n}) &:= E\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j+1} | \mathcal{Q}_{I,n} \right) = \frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} E(C_{i,j+1} | C_{i,I-i}), \\
\mu_{j,n}^{(2)}(\mathcal{Q}_{I,n}) &:= E\left(\frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} C_{i,j} | \mathcal{Q}_{I,n} \right) = \frac{1}{I+n-j} \sum_{i=-n}^{I-j-1} E(C_{i,j} | C_{i,I-i}).
\end{aligned}$$

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 3.8 and 3.9 below, which resemble Assumptions 4.6 and 4.8 in Steinmetz and Jentsch (2022).

Assumption 3.8 (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 $j \to \infty$ such that $\sum_{j=0}^{\infty} (j+1)^2 \frac{\sigma_j^2}{\epsilon^j} < \infty$.*

In addition to the condition on the support and the variance parameters in Assumption 3.8, a regularity condition for the *backward* conditional distribution of cumulative claim $C_{i,j}$ given $C_{i,j+1}$ is required.

[2]Note that conditioning on $\mathcal{Q}_{I,n}$ or $\mathcal{Q}_{I,\infty}$ is equivalent.

Assumption 3.9 (Backward conditional moments) *Assumptions 3.2, 3.3, 3.4 and 3.8 are fulfilled such that, for all* $K \in \mathbb{N}_0$, $k \geq 0$ *and* $j, j_1, j_2 \in \{0, \ldots, K\}$, $j_1 \leq j_2$, *we have*

$$|E(C_{i,j}|C_{i,j+k}) - E(C_{i,j}|C_{i,j+k+1})| \leq a_k X_i, \tag{3.38}$$

$$|Cov(C_{i,j_1}, C_{i,j_2}|C_{i,j_2+k}) - Cov(C_{i,j_1}, C_{i,j_2}|C_{i,j_2+k+1})| \leq b_k Y_i, \tag{3.39}$$

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 (3.17), 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 $Var(C_{i,j}|C_{i,j+1})$; see Example 4.9 in Steinmetz and Jentsch (2022). Hence, Assumption 3.9 is required to control the backward conditional distributions of cumulative claims.

Based on Theorem 4.10 from 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 3.8 and 3.9 together allow to establish asymptotic normality of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(2)}$ conditional on $\mathcal{Q}_{I,n}$.

Theorem 3.10 (Asymptotics for $(\boldsymbol{R_{I,n}} - \widehat{\boldsymbol{R}}_{\boldsymbol{I,n}})_{\mathbf{2}}$ conditional on $\boldsymbol{\mathcal{Q}_{I,n}}$) *Suppose Assumptions 3.2, 3.3, 3.4, 3.8 and 3.9 hold. Then, as* $n \to \infty$, *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*

$$\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2^{(1)} \xrightarrow{d} \left\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty^{(1)} \right\rangle \sim \mathcal{G}_2^{(1)}, \tag{3.40}$$

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

$$Cov(Y_{i_1}^{(1)}, Y_{i_2}^{(1)}) = \lim_{K \to \infty} \mathbf{\Sigma}_{K,\prod f_j}^{(1)}(i_1, i_2) \tag{3.41}$$

for $i_1, i_2 \in \mathbb{N}_0$, *where* $\mathbf{\Sigma}_{K,\prod f_j}^{(1)}(i_1, i_2)$ *is defined in Corollary 3.28. Here, the two random sequences* $\mathcal{Q}_{I,\infty}$ *and* $\mathbf{Y}_\infty^{(1)}$ *are stochastically independent.*

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

$$\sqrt{I+n+1}(R_{I,n}-\widehat{R}_{I,n})_2^{(2)}|\mathcal{Q}_{I,n} \xrightarrow{d} \left\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty^{(2)} \right\rangle |\mathcal{Q}_{I,\infty} \sim \mathcal{G}_2^{(2)}|\mathcal{Q}_{I,\infty}, \tag{3.42}$$

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

$$\Xi(\mathcal{Q}_{I,\infty}) = \lim_{K\to\infty} \mathcal{Q}_{I,K-I}\mathbf{\Sigma}^{(2)}_{K,\prod f_j}\mathcal{Q}'_{I,K-I} = \lim_{K\to\infty} \mathcal{Q}_{I,K-I}(\mathbf{\Sigma}_{K,\prod f_j} - \mathbf{\Sigma}^{(1)}_{K,\prod f_j})\mathcal{Q}'_{I,K-I}, \tag{3.43}$$

where $\mathbf{\Sigma}_{K,\prod f_j}$ as well as $\mathbf{\Sigma}^{(1)}_{K,\prod f_j}$ and $\mathbf{\Sigma}^{(2)}_{K,\prod f_j}$ are defined in Corollary 3.28.

According to Theorem 3.10(ii), in contrast to $\mathcal{G}_1|\mathcal{Q}_{I,n}$ in Theorem 3.7, the conditional limiting distribution $\mathcal{G}_2^{(2)}|\mathcal{Q}_{I,\infty}$ will be Gaussian. Together with Theorem 3.10(i), conditional on $\mathcal{Q}_{I,\infty}$ and inflated with $\sqrt{I+n+1}$, the estimation uncertain term $\sqrt{I+n+1}(R_{I,n}-\widehat{R}_{I,n})_2$ will be Gaussian with mean $\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty^{(2)} \rangle$ and variance $\Xi(\mathcal{Q}_{I,\infty})$.

3.4.1.3 Conditional asymptotics for the whole predictive root of the reserve

By combining the limiting conditional distributions derived separately for both parts $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$, joint asymptotic results can also be established. Based on Theorem 4.12 and Corollary 4.13 from Steinmetz and Jentsch (2022), the following theorem provides the limiting distribution of the whole predictive root of the reserve $R_{I,n}-\widehat{R}_{I,n}$ conditional on $\mathcal{Q}_{I,n}$.

Theorem 3.11 (Asymptotics for $R_{I,n}-\widehat{R}_{I,n}$ conditional on $\mathcal{Q}_{I,n}$) *Suppose the assumptions of Theorems 3.7 and 3.10 hold. Then, conditional on $\mathcal{Q}_{I,n}$, $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$ are stochastically independent, and $R_{I,n}-\widehat{R}_{I,n}|\mathcal{Q}_{I,n}$ converges in distribution to $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$. That is, we have*

$$R_{I,n}-\widehat{R}_{I,n}|\mathcal{Q}_{I,n} = (R_{I,n}-\widehat{R}_{I,n})_1 + (R_{I,n}-\widehat{R}_{I,n})_2|\mathcal{Q}_{I,n} \xrightarrow{d} \mathcal{G}_1|\mathcal{Q}_{I,\infty}. \tag{3.44}$$

According to Theorem 3.10, $(R_{I,n}-\widehat{R}_{I,n})_2$ requires an inflation factor $\sqrt{I+n+1}$ to establish convergence towards a non-degenerate limiting distribution. As this is not the case for $(R_{I,n}-\widehat{R}_{I,n})_1$ in Theorem 3.7, the latter term $(R_{I,n}-\widehat{R}_{I,n})_1$ corresponding to the process uncertainty will asymptotically dominate the predictive root $R_{I,n}-\widehat{R}_{I,n}$.

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 $F_{i,j}|C_{i,j}$.

3.4.2 Conditional bootstrap asymptotics for the Mack bootstrap predictive root of the reserve

In view of the decomposition $R_{I,n}-\widehat{R}_{I,n}=(R_{I,n}-\widehat{R}_{I,n})_1+(R_{I,n}-\widehat{R}_{I,n})_2$ of the predictive root of the reserve in (3.31) and the conditional limiting distributions of $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$ gathered in Section 3.4.1, it is instructive to consider the corresponding Mack bootstrap expressions $(R^*_{I,n}-\widehat{R}_{I,n})_1$ and $(R^*_{I,n}-\widehat{R}_{I,n})_2$ from (3.32) and check whether they are correctly mimicking such limiting distributions. While $(R_{I,n}-\widehat{R}_{I,n})_1$ and $(R_{I,n}-\widehat{R}_{I,n})_2$ are asymptotically analyzed conditional on $\mathcal{Q}_{I,n}$, the bootstrap quantities $(R^*_{I,n}-\widehat{R}_{I,n})_1$ and $(R^*_{I,n}-\widehat{R}_{I,n})_2$ have to be considered conditional on $\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}$, but also (as common in the bootstrap literature) conditional on $\mathcal{D}_{I,n}$, that is, on the available cumulative claim data.

3.4.2.1 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 3.12 (Parametric family of (conditional) distributions of $\boldsymbol{F_{i,j}}$) *The (conditional) distribution of $F_{i,j}|C_{i,j},\ i\in\mathbb{Z},\ i\leq I,\ j\in\mathbb{N}_0$, belongs to a parametric family of distributions $\mathbf{H}$, that fulfills the following properties:*

(i) A distribution $\mathcal{H}\in\mathbf{H}$ is uniquely specified by its first two (conditional) moments. That is, for all $i\in\mathbb{Z},\ i\leq I,\ j\in\mathbb{N}_0$, the conditional distribution of $F_{i,j}|C_{i,j}=c$ is (almost surely) uniquely determined by its conditional mean $E(F_{i,j}|C_{i,j}=c)=f_j$ and its conditional variance $Var(F_{i,j}|C_{i,j}=c)=\frac{\sigma_j^2}{c}$ according to (3.18)*.*

(ii) The distributions $\mathcal{H}\in\mathbf{H}$ are continuous in the parameters 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 3.7 is generally non-Gaussian and depending on the (conditional) distribution (family) of the individual development factors, we require also that the bootstrap individual development factors $F^*_{i,j}$ for $i\in\mathbb{Z},\ i\leq I,\ j\in\mathbb{N}_0$ and

$i+j \geq I$, that is, for the *lower* triangle, follow the true parametric family of (conditional) distributions as the $F_{i,j}$'s according to Assumption 3.12.

Assumption 3.13 ((Conditional) distributions of $F^*_{i,j}$ in lower triangle) *The (conditional) distribution of $F^*_{i,j}|C^*_{i,j}, \mathcal{D}_{I,n}, \mathcal{F}^*_{I,n}$, for $i \in \mathbb{Z}$, $i \leq I$, $j \in \mathbb{N}_0$ and $i+j \geq I$ in Step 5 of the Mack Bootstrap Scheme in Section 3.3.1 belongs to the true parametric family of (conditional) distributions $\mathbf{H}$ used to generate $F_{i,j}|C_{i,j}$ according to Assumption 3.12. That is, we have $F^*_{i,j}|(C^*_{i,j}=x, \widehat{f}^*_{j,n}=y, \widehat{\sigma}^2_{j,n}=z) \overset{d}{=} F_{i,j}|(C_{i,j}=x, f_j=y, \sigma^2_j=z)$ for all $(x,y,z)' \in (0,\infty)$.*

Together with the assumptions imposed in Theorem 3.7, the Assumptions 3.12 and 3.13 allow to prove the following theorem.

Theorem 3.14 (Bootstrap asymptotics for $(R^*_{I,n}-\widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$) *Suppose Assumptions 3.2, 3.3, 3.4, 3.5, 3.12, and 3.13 hold. Then, as $n \to \infty$, conditionally on $\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, $(R^*_{I,n}-\widehat{R}_{I,n})_1$ converges in distribution to $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ in probability, which is the (limiting) distribution of $(R_{I,\infty}-\widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}$ according to* (3.36) *described in Theorem 3.7. Moreover, for all $n \in \mathbb{N}_0$, it holds $E^*((R^*_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n})=0$ and, for $n \to \infty$, we have*

$$Var^*\left((R^*_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}\right) \longrightarrow Var\left((R_{I,\infty}-\widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}\right) \tag{3.45}$$

in probability, where $Var((R_{I,\infty}-\widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty})=O_P(1)$ as given in (3.35). *Consequently, as $n \to \infty$, we have*

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

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 $d_2(G,H)=d_2(X,Y)=inf(E||X-Y||^2)^{\frac{1}{2}}$, where the infimum is taken over all joint distributions (X,Y) with marginals $X \sim G$ and $Y \sim H$.

3.4.2.2 Conditional bootstrap asymptotics for reserve prediction: estimation uncertainty

In view of the decomposition $(R_{I,n}-\widehat{R}_{I,n})_2=(R_{I,n}-\widehat{R}_{I,n})^{(1)}_2+(R_{I,n}-\widehat{R}_{I,n})^{(2)}_2$ in (3.37), for the derivation of corresponding bootstrap asymptotic theory, it is *seemingly* instructive to further decompose also its bootstrap counterpart $(R^*_{I,n}-\widehat{R}_{I,n})_2$ 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 $\widehat{f}^*_{j,n}$ in

(3.28), we get

$$
\begin{aligned}
&(R^*_{I,n}-\widehat{R}_{I,n})_2\\
=&\sum_{i=0}^{I+n} C^*_{I-i,i}\left(\prod_{j=i}^{I+n-1}\widehat{f}^*_{j,n}-\prod_{j=i}^{I+n-1}f^*_{j,n}(\mathcal{Q}_{I,n})\right)+\sum_{i=0}^{I+n} C^*_{I-i,i}\left(\prod_{j=i}^{I+n-1}f^*_{j,n}(\mathcal{Q}_{I,n})-\prod_{j=i}^{I+n-1}\widehat{f}_{j,n}\right)\\
=&(R^*_{I,n}-\widehat{R}_{I,n})_2^{(1)}+(R^*_{I,n}-\widehat{R}_{I,n})_2^{(2)},
\end{aligned}
\tag{3.46}
$$

where $f^*_{j,n}(\mathcal{Q}_{I,n}) := \mu^{*(1)}_{j+1,n}(\mathcal{Q}_{I,n})/\mu^{*(2)}_{j,n}(\mathcal{Q}_{I,n})$ with

$$
\begin{aligned}
\mu^{*(1)}_{j+1,n}(\mathcal{Q}_{I,n}) &:= E^*\left(\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}F^*_{i,j}|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}\right)\\
&=\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} E^*(C_{i,j}F^*_{i,j}|C^*_{i,I-i}=C_{i,I-i}),\\
\mu^{*(2)}_{j,n}(\mathcal{Q}_{I,n}) &:= E^*\left(\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}\right)=\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}.
\end{aligned}
$$

Now, taking a closer look at $\mu^{*(1)}_{j+1,n}(\mathcal{Q}_{I,n})$, we get

$$
\begin{aligned}
\mu^{*(1)}_{j+1,n}(\mathcal{Q}_{I,n}) &= \frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}E^*(F^*_{i,j}|C^*_{i,I-i}=C_{i,I-i}) = \frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}E^*(F^*_{i,j})\\
&=\left(\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1} C_{i,j}\right)\widehat{f}_{j,n},
\end{aligned}
$$

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 $\mathcal{D}^*_{I,n}$. Altogether, using $E^*(F^*_{i,j})=\widehat{f}_{j,n}$, we get

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

leading to $(R^*_{I,n}-\widehat{R}_{I,n})_2^{(2)}=0$ such that $(R^*_{I,n}-\widehat{R}_{I,n})_2=(R^*_{I,n}-\widehat{R}_{I,n})_2^{(1)}$. Hence, in contrast to the limiting behavior of $\sqrt{I+n+1}(R_{I,n}-\widehat{R}_{I,n})_2$, which was derived after further decomposing $(R_{I,n}-\widehat{R}_{I,n})_2$ into two additive parts $(R_{I,n}-\widehat{R}_{I,n})_2^{(1)}$ and $(R_{I,n}-\widehat{R}_{I,n})_2^{(2)}$ as stated in Theorem 3.10, such an analogous decomposition of $(R^*_{I,n}-\widehat{R}_{I,n})_2$ does *not* exist. However, letting $n\to\infty$, it remains to check the limiting properties of $(R^*_{I,n}-\widehat{R}_{I,n})_2$, which are presented in the following theorem.

For this purpose, it is important to note that, in contrast to the derivation of the *conditional* limiting result obtained in Theorem 3.10(ii), which relies on conditional CLTs for the development factor estimators $\widehat{f}_{j,n}$ as stated in Steinmetz and Jentsch (2022, Appendix C), the derivation of the limiting properties of $(R^*_{I,n} - \widehat{R}_{I,n})_2$ rely on *unconditional* bootstrap CLTs for the Mack bootstrap development factor estimators $\widehat{f}^*_{j,n}$, that is, *without* conditioning on $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$. Nevertheless, to prove asymptotic normality for the $\widehat{f}^*_{j,n}$'s by justifying a Lyapunov condition, we have to impose additional regularity conditions on the estimators for the development factors and variance parameters. This is required to control the behavior of the non-parametric bootstrap in Step 3 of the Mack bootstrap in Section 3.3.1, which leads to bootstrap errors $r^*_{i,j}$ that are drawn randomly from *all* (re-centered and re-scaled) residuals $\widetilde{r}_{i,j}$, that is, for all $j = 0, \ldots, I+n-2$ and $i = -n, \ldots, I-j-1$.

Assumption 3.15 (Uniform boundedness condition) *For $n \to \infty$, suppose that the development factor estimators $\widehat{f}_{j,n}$, $j = 0, \ldots, I+n-1$ and the variance parameter estimators $\widehat{\sigma}^2_{j,n}$, $j = 0, \ldots, I+n-2$ fulfill the uniform boundedness conditions*

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

Moreover, for $\kappa_j^{(4)}$ defined in (3.21)*, suppose that $((\kappa_j^{(4)}/\sigma_j^4), j \in \mathbb{N}_0)$ is a bounded sequence.*

The assumption above allows to state the following asymptotic result for the Mack bootstrap estimation uncertainty part.

Theorem 3.16 (Bootstrap asymptotics for $(\boldsymbol{R}^*_{I,n} - \widehat{\boldsymbol{R}}_{I,n})_2$ conditional on $\boldsymbol{\mathcal{Q}}^*_{I,n} = \boldsymbol{\mathcal{Q}}_{I,n}$ and $\boldsymbol{\mathcal{D}}_{I,n}$) *Suppose Assumptions 3.2, 3.3, 3.5, 3.4, 3.8 and 3.15 hold. Then, as $n \to \infty$, conditionally on $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, $\sqrt{I+n+1}(R^*_{I,n} - \widehat{R}_{I,n})_2$ converges in distribution to $\widetilde{\mathcal{G}}_2|\mathcal{Q}_{I,\infty}$ in probability, where $\widetilde{\mathcal{G}}_2|\mathcal{Q}_{I,\infty} \sim \mathcal{N}(0, \widetilde{\Xi}(\mathcal{Q}_{I,\infty}))|\mathcal{Q}_{I,\infty}$ is a conditional Gaussian distribution with conditional mean zero and conditional variance*

$$\widetilde{\Xi}(\mathcal{Q}_{I,\infty}) = \lim_{K\to\infty} \mathcal{Q}_{I,K-I} \boldsymbol{\Sigma}_{K,\prod f_j} \mathcal{Q}'_{I,K-I}, \tag{3.47}$$

where $\boldsymbol{\Sigma}_{K,\prod f_j} = \boldsymbol{\Sigma}^{(1)}_{K,\prod f_j} + \boldsymbol{\Sigma}^{(2)}_{K,\prod f_j}$ is defined in Corollary 3.28 and

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

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

$$d_2\left(\mathcal{L}\left(\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2|\mathcal{Q}_{I,n}\right), \mathcal{L}^*\left(\sqrt{I+n+1}(R^*_{I,n} - \widehat{R}_{I,n})_2|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right)\right) \not\longrightarrow 0$$

in probability. This is because the limiting normal distribution of $\sqrt{I+n+1}(R^*_{I,n} - \widehat{R}_{I,n})_2$ *conditional on* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ *and* $\mathcal{D}_{I,n}$ *deviates in its (zero) mean and its variance* $\widetilde{\Xi}(\mathcal{Q}_{I,\infty})$ *from the limiting distribution of* $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ *conditional on* $\mathcal{Q}_{I,n}$*, which has a mean* $\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}^{(2)}_\infty \rangle$ *and variance* $\Xi(\mathcal{Q}_{I,\infty})$ *according to Theorem 3.10.*

3.4.2.3 Conditional bootstrap asymptotics for the whole predictive root of the reserve

As in Section 3.4.1.3, we can combine the results on the limiting distributions for $(R^*_{I,n} - \widehat{R}_{I,n})_1$ and $(R^*_{I,n} - \widehat{R}_{I,n})_2$ from Theorems 3.14 and 3.16, respectively, to get the limiting bootstrap distribution of the whole bootstrap predictive root of the reserve $R^*_{I,n} - \widehat{R}_{I,n}$ conditional on $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$.

Theorem 3.17 (Bootstrap asymptotics for $\boldsymbol{R^*_{I,n} - \widehat{R}_{I,n}}$ conditional on $\boldsymbol{\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}}$ and $\boldsymbol{\mathcal{D}_{I,n}}$) *Suppose the assumptions of Theorems 3.14 and 3.16 hold. Then, conditional on* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ *and* $\mathcal{D}_{I,n}$, $(R^*_{I,n} - \widehat{R}_{I,n})_1$ *and* $(R^*_{I,n} - \widehat{R}_{I,n})_2$ *are uncorrelated, and* $R^*_{I,n} - \widehat{R}_{I,n} | (\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n})$ *converges in distribution to* $\mathcal{G}_1 | \mathcal{Q}_{I,\infty}$*. That is, we have*

$$R^*_{I,n} - \widehat{R}_{I,n} | \left(\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}\right) = (R^*_{I,n} - \widehat{R}_{I,n})_1 + (R^*_{I,n} - \widehat{R}_{I,n})_2 | \left(\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}\right)$$
$$\xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I,\infty}$$

in probability.

As already observed in Theorem 3.10 for the estimation uncertainty term $(R_{I,n} - \widehat{R}_{I,n})_2$, 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 $(R_{I,n} - \widehat{R}_{I,n})_1$ in Theorem 3.7 and its bootstrap version in Theorem 3.14, the process uncertainty terms will asymptotically dominate the predictive roots $R_{I,n} - \widehat{R}_{I,n}$ and $R^*_{I,n} - \widehat{R}_{I,n}$. Hence, although the limiting bootstrap distribution of $\sqrt{I+n+1}(R^*_{I,n} - \widehat{R}_{I,n})_2$ conditional on $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$ in Theorem 3.16 does *not* correctly mimic the corresponding limiting behavior of $\sqrt{I+n+1}(R_{I,n} - \widehat{R}_{I,n})_2$ conditional on $\mathcal{Q}_{I,n}$ in Theorem 3.10, the whole bootstrap predictive root $R^*_{I,n} - \widehat{R}_{I,n}$ still mimics the limiting distribution of the predictive root $R_{I,n} - \widehat{R}_{I,n}$ correctly.

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

This motivates the construction of an alternative Mack-type bootstrap proposed in the following section. We conclude this section with some remarks.

Remark 3.18 (On the asymptotic results for the Mack bootstrap)

(i) *A closer inspection of the decompositions in* (3.31) *and* (3.32) *reveals some inconsistencies:*

- *While a term based on products of development parameters* f_j *is added to and subtracted from* $R_{I,n} - \widehat{R}_{I,n}$ *to get* $(R_{I,n} - \widehat{R}_{I,n})_1$ *and* $(R_{I,n} - \widehat{R}_{I,n})_2$*, a term using products of bootstrap development factor estimators* $\widehat{f}^*_{j,n}$ *instead of the more natural choice of development factor estimators* $\widehat{f}_{j,n}$ *is added to and subtracted from* $R^*_{I,n} - \widehat{R}_{I,n}$ *to get* $(R^*_{I,n} - \widehat{R}_{I,n})_1$ *and* $(R^*_{I,n} - \widehat{R}_{I,n})_2$*.*

- *Consequently, while* $(R_{I,n} - \widehat{R}_{I,n})_1$ *relies on products of individual development factors* $F_{i,j}$ *centered around products of development parameters* f_j*, its Mack bootstrap version* $(R^*_{I,n} - \widehat{R}_{I,n})_1$ *relies on bootstrap individual development factors* $F^*_{i,j}$*, which are not naturally centered around development factor estimators* $\widehat{f}_{j,n}$*, but around bootstrap quantities* $\widehat{f}^*_{j,n}$*.*

- *Moreover, while* $(R_{I,n} - \widehat{R}_{I,n})_2$ *relies on differences between products of development parameters* f_j *and products of their estimators* $\widehat{f}_{j,n}$*, its Mack bootstrap version* $(R^*_{I,n} - \widehat{R}_{I,n})_2$ *relies on differences between products of bootstrap development factor estimators* $\widehat{f}^*_{j,n}$ *and products of estimators* $\widehat{f}_{j,n}$*. Hence, the sign of* $(R^*_{I,n} - \widehat{R}_{I,n})_2$ *is flipped in comparison to* $(R_{I,n} - \widehat{R}_{I,n})_2$*. 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* $(R^*_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ *and* $(R^*_{I,n} - \widehat{R}_{I,n})_2^{(2)}$ *in the seemingly natural decomposition of the bootstrap estimation uncertainty term in* (3.46) *are switched in comparison to* $(R_{I,n} - \widehat{R}_{I,n})_2^{(1)}$ *and* $(R_{I,n} - \widehat{R}_{I,n})_2^{(2)}$*.*

(ii) *The bootstrap consistency result for the Mack bootstrap process uncertainty part conditional on* $\mathcal{Q}_{I,n}$ *in Theorem 3.14 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 3.3.1. Otherwise, only the first and second moments of the conditional distribution will be correctly mimicked asymptotically, but in general not the whole distribution and, consequently, also not the quantiles.*

(iii) *The uniform boundedness conditions in Assumption 3.15 are required to establish a Lyapunov condition for bootstrap CLTs in Theorem 3.25, because the Mack bootstrap draws bootstrap errors* $r^*_{i,j}$ *from residuals computed from all columns in* $\mathcal{D}_{I,n}$*.*

(iv) *In contrast to* $(R_{I,n} - \widehat{R}_{I,n})_1$ *and* $(R_{I,n} - \widehat{R}_{I,n})_2$*, which are independent conditional on* $\mathcal{Q}_{I,n}$*, both parts* $(R^*_{I,n} - \widehat{R}_{I,n})_1$ *and* $(R^*_{I,n} - \widehat{R}_{I,n})_2$ *of the bootstrap predictive root are in general only uncorrelated conditional on* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ *and* $\mathcal{D}_{I,n}$.

(v) *The bootstrap inconsistency result for the Mack bootstrap estimation uncertainty part conditional on* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ *established in Theorem 3.16 is due to the fact that the bootstrap approach in Step 4 is not taking the condition* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ *into account.*

Hence, the (always larger!) variance-covariance matrix $\mathbf{\Sigma}_{K,\prod f_j}$ *shows in the conditional limiting distribution instead of* $\mathbf{\Sigma}^{(2)}_{K,\prod f_j}$ *obtained in Theorem 3.10. Moreover, a decomposition of* $(R^*_{I,n} - \widehat{R}_{I,n})_2$ *resembling the decomposition of* $(R_{I,n} - \widehat{R}_{I,n})_2$ *in* (3.37) *does not exist. Consequently, the behavior of* $(R_{I,n} - \widehat{R}_{I,n})^{(1)}_2$ *is also not correctly mimicked.*

(vi) *The requirement of a bootstrap procedure to not only mimic the asymptotically dominating part of the (conditional) predictive distribution that captures the prediction (i.e. process) uncertainty (i.e. asymptotic validity), but also the asymptotically negligible part capturing the uncertainty due to model parameter estimation is closely related top the concept coined asymptotic pertinence in Pan and Politis (2016a) for time series prediction, which is also discussed by Beutner et al. (2021) from a different perspective. Pan and Politis (2016a) argue that asymptotic validity of predictive inference is a fundamental property, but capturing the uncertainty due to model estimation is beneficial in finite samples.*

(vii) *The discussion above motivates an alternative notion of a Mack-type bootstrap to be introduced in Section 3.5 that will be designed to eliminate the raised issues. In particular, it respects the conditioning on* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ *and generates a bootstrap loss triangle* $\mathcal{D}^*_{I,n}$ *in a backward manner starting from the diagonal* $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$*. See e.g. Paparoditis and Shang (2021) for bootstrap predictive inference in a functional time series setup.*

3.5 An alternative Mack-type Bootstrap Scheme

According to the findings and the discussion in Section 3.4, 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 $R_{I,n} - \widehat{R}_{I,n}$ using an *alternative* bootstrap predictive root of the reserve $R^+_{I,n} - \widehat{R}^+_{I,n}$ to be defined below. To distinguish it from the original Mack bootstrap proposal in Section 3.3, we denote all related bootstrap quantities and operations with a "+" instead of a "$*$". This novel approach deviates from the original Mack bootstrap scheme in several ways:

(i) First, given the loss triangle $\mathcal{D}_{I,n}$ and conditional on the bootstrap diagonal $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$, where $\mathcal{Q}^+_{I,n} = \{C^+_{I-i,i} | i = 0, \ldots, I+n\}$, a recursive *backward* bootstrap approach is employed to generate a whole bootstrap *upper triangle*

$$\mathcal{D}^+_{I,n} = \left\{ C^+_{i,j} | i = -n, \ldots, I,\ j = 0, \ldots, I+n,\ -n \leq i+j \leq I \right\}.$$

Then, bootstrap estimators for the development factors $\widehat{f}^+_{j,n}$, $j = 0, \ldots, I+n-1$ are computed according to formula (3.10), but based on the bootstrap upper loss triangle $\mathcal{D}^+_{I,n}$.

(ii) Second, instead of the bootstrap development factor estimators $\widehat{f}^+_{j,n}$ computed from $\mathcal{D}^+_{I,n}$, the development factor estimators $\widehat{f}_{j,n}$ computed from $\mathcal{D}_{I,n}$ are used for a *parametric* bootstrap to construct bootstrap individual development factors $F^+_{i,j}$, $i = -n, \ldots, I$, $j = 0, \ldots, I+n-1$ and $i+j \geq I$, that is for the *lower triangle*. This allows to construct also $R^+_{I,n}$.

(iii) Third, for the construction of the bootstrap predictive root of the reserve $R^+_{I,n} - \widehat{R}^+_{I,n}$, the bootstrap reserve $R^+_{I,n}$ is *not* centered around its best estimate $\widehat{R}_{I,n}$, but around a suitable bootstrap version $\widehat{R}^+_{I,n}$.

Finally, analogous to the original Mack bootstrap, the alternative Mack bootstrap is employed to estimate the conditional distribution of $R_{I,n} - \widehat{R}_{I,n}$ given $\mathcal{Q}_{I,n}$ is estimated by the conditional bootstrap distribution of $R^+_{I,n} - \widehat{R}^+_{I,n}$ given $\mathcal{D}_{I,n}$ *and* $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$.

3.5.1 An alternative Mack-type Bootstrap Algorithm

With the upper triangle $\mathcal{D}_{I,n}$ 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 $\widehat{f}_{j,n}$ and $\widehat{\sigma}^2_{j,n}$ for $j = 0, \ldots, I+n-1$ as defined in (3.10) and (3.13), respectively.

Step 2. Choose a parametric family for the (conditional) bootstrap distributions of the *backward* individual development factors $G^+_{i,j}$, $j = 0, \ldots, I+n-1$ and $i = -n, \ldots, I-j-1$ given $C^+_{i,j+1}$ and $\mathcal{D}_{I,n}$ such that $G^+_{i,j} > 0$ a.s. holds with mean $E^+(G^+_{i,j}|C^+_{i,j+1}) = \widehat{f}^{-1}_{j,n}$ and variance $Var^+(G^+_{i,j}|C^+_{i,j+1}) = \frac{\widehat{\sigma}^2_{j,n}}{C^+_{i,j+1}}$ for $i = -n, \ldots, I$, $j = 0, \ldots, I+n-1$ and $i + j < I$, where $E^+(\cdot) = E^+(\cdot|\mathcal{D}_{I,n})$ and $Var^+(\cdot) = Var^+(\cdot|\mathcal{D}_{I,n})$. Then, conditional on $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$, generate *backwards* a bootstrap loss triangle $\mathcal{D}^+_{I,n}$ using the recursion

$$C^+_{i,j} = C^+_{i,j+1} G^+_{i,j}$$

for $j = 0, \ldots, I+n-1$ and $i = -n, \ldots, I-j-1$.

Step 3. Compute bootstrap estimators for the development factors f_j by computing $\widehat{f}^+_{j,n}$ for $j = 0, \ldots, I+n-1$, which are defined analogously to $\widehat{f}_{j,n}$ defined in (3.10), but $\widehat{f}^+_{j,n}$ is calculated from the bootstrap loss triangle $\mathcal{D}^+_{I,n}$. That is, we compute

$$\widehat{f}^+_{j,n} = \frac{\sum\limits_{i=-n}^{I-j-1} C^+_{i,j+1}}{\sum\limits_{i=-n}^{I-j-1} C^+_{i,j}} = \frac{\sum\limits_{i=-n}^{I-j-1} C^+_{i,j+1}}{\sum\limits_{i=-n}^{I-j-1} C^+_{i,j+1} G^+_{i,j}} \tag{3.48}$$

for $j = 0, \ldots, I+n-1$.

Step 4. Choose a parametric family for the (conditional) bootstrap distributions of $F^+_{i,j}$ given $C^+_{i,j}$ and $\mathcal{D}_{I,n}$ such that $F^+_{i,j} > 0$ a.s. with

$$E^+(F^+_{i,j}|C^+_{i,j}) = \widehat{f}_{j,n} \quad Var^+(F^+_{i,j}|C^+_{i,j}) = \frac{\widehat{\sigma}^2_{j,n}}{C^+_{i,j}}$$

for $i = -n, \ldots, I$, $j = 0, \ldots, I+n-1$ and $i + j \geq I$.
Then, given $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$, 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 $i = -n, \ldots, I$ using the *forward* recursion

$$C^+_{i,j+1} = C^+_{i,j} F^+_{i,j} \tag{3.49}$$

for $\quad j = I-i, \ldots, I+n-1.$

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

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

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

$$\widehat{R}^+_{I,n} = \sum_{i=-n}^{I} C^+_{i,I-i} \prod_{j=I-i}^{I+n-1} \widehat{f}^+_{j,n} = \sum_{i=0}^{I+n} C^+_{I-i,i} \prod_{j=i}^{I+n-1} \widehat{f}^+_{j,n}. \tag{3.50}$$

Step 6. Repeat Steps 2 - 5 above B times, where B is large, to get $(R^+_{I,n} - \widehat{R}^+_{I,n})^{(b)}$, $b = 1, \ldots, B$ 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 3.19 (On the alternative Mack-type bootstrap)

(i) In comparison to the Mack bootstrap from Section 3.3.1, 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 backward 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 in a non-parametric way similar to Step 2 - 4 in Section 3.3.1 or using the parametric family of distributions used in Step 5 in Section 3.3.1. 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 $Var(C_{i,j}|C_{i,j+1})$, respectively. This will be reflected by Assumption 3.21 to be imposed in Section 3.6.

3.6 Asymptotic Theory for the alternative Mack Bootstrap

By adopting the general strategy of Section 3.4 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 into two additive parts that account for the prediction error and the estimation error, respectively. 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

$$R^+_{I,n} - \widehat{R}^+_{I,n} = \sum_{i=0}^{I+n} C^+_{I-i,i} \left(\prod_{j=i}^{I+n-1} F^+_{I-i,j} - \prod_{j=i}^{I+n-1} \widehat{f}_{j,n} \right) + \sum_{i=0}^{I+n} C^+_{I-i,i} \left(\prod_{j=i}^{I+n-1} \widehat{f}_{j,n} - \prod_{j=i}^{I+n-1} \widehat{f}^+_{j,n} \right)$$
$$=: \left(R^+_{I,n} - \widehat{R}^+_{I,n}\right)_1 + \left(R^+_{I,n} - \widehat{R}^+_{I,n}\right)_2, \tag{3.51}$$

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.

3.6.1 Conditional bootstrap asymptotics for the alternative Mack bootstrap predictive root of the reserve

As in Section 3.4.2 for the original Mack bootstrap, conditional on $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, we have to check whether the alternative Mack bootstrap expressions $(R^+_{I,n} - \widehat{R}^+_{I,n})_1$ and $(R^+_{I,n} - \widehat{R}^+_{I,n})_2$ from (3.32) are correctly mimicking 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, as summarized in Section 3.4.1.

3.6.1.1 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 $(R^*_{I,n} - \widehat{R}_{I,n})_1$ in two aspects. On the one hand, 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, on the other hand, $\prod_{j=i}^{I+n-1} F^+_{I-i,j}$ 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, as the proof to establish the asymptotic distribution in Theorem 3.14 exclusively relies on $\widehat{f}_{j,n} - f_j = O_P((I+n-1)^{-1/2})$, $\widehat{f}^*_{j,n} - \widehat{f}_{j,n} = O_{P^*}((I+n-1)^{-1/2})$ and $\widehat{\sigma}^2_{j,n} - \sigma^2_j = O_P((I+n-1)^{-1/2})$ for all fixed $j \in \mathbb{N}_0$, by using the same arguments, we get immediately the same result also for the process uncertainty part $(R^+_{I,n} - \widehat{R}^+_{I,n})_1$ of the alternative Mack bootstrap.

Theorem 3.20 (Bootstrap asymptotics for $(R^+_{I,n} - \widehat{R}^+_{I,n})_1$ conditional on $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$) *Suppose Assumptions 3.2, 3.3, 3.4, 3.5, 3.12 and 3.13 (for $F^+_{i,j}$ instead of $F^*_{i,j}$) hold. Then, as $n \to \infty$, conditionally on $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, $(R^+_{I,n} - \widehat{R}^+_{I,n})_1$*

converges in distribution to $\mathcal{G}_1|\mathcal{Q}_{I,\infty}$ *in probability, which is the (limiting) distribution of* $(R_{I,\infty}-\widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}$ *according to* (3.36) *described in Theorem 3.7. Moreover, for all* $n\in\mathbb{N}_0$*, it holds* $E^+((R^+_{I,n}-\widehat{R}^+_{I,n})_1|\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n})=0$ *and, for* $n\to\infty$*, we have*

$$Var^+\left((R^+_{I,n}-\widehat{R}^+_{I,n})_1|\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n}\right)\longrightarrow Var\left((R_{I,\infty}-\widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty}\right) \tag{3.52}$$

in probability, where $Var((R_{I,\infty}-\widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty})=O_P(1)$ *as given in* (3.35)*. Consequently, as* $n\to\infty$*, we have*

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

in probability.

3.6.1.2 Conditional bootstrap asymptotics for reserve prediction: estimation uncertainty

In view of the decomposition $(R_{I,n}-\widehat{R}_{I,n})_2=(R_{I,n}-\widehat{R}_{I,n})_2^{(1)}+(R_{I,n}-\widehat{R}_{I,n})_2^{(2)}$ in (3.37), conditional on $\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, its alternative Mack bootstrap counterpart $(R^+_{I,n}-\widehat{R}^+_{I,n})_2$ can be also decomposed further. That is, we have

$$\begin{aligned}&(R^+_{I,n}-\widehat{R}^+_{I,n})_2\\ =&\sum_{i=0}^{I+n}C^+_{I-i,i}\left(\prod_{j=i}^{I+n-1}\widehat{f}_{j,n}-\prod_{j=i}^{I+n-1}f^+_{j,n}(\mathcal{Q}_{I,n})\right)+\sum_{i=0}^{I+n}C^+_{I-i,i}\left(\prod_{j=i}^{I+n-1}f^+_{j,n}(\mathcal{Q}_{I,n})-\prod_{j=i}^{I+n-1}\widehat{f}^+_{j,n}\right)\\ =&(R^+_{I,n}-\widehat{R}^+_{I,n})_2^{(1)}+(R^+_{I,n}-\widehat{R}^+_{I,n})_2^{(2)},\end{aligned} \tag{3.53}$$

where $(R^+_{I,n}-\widehat{R}^+_{I,n})_2^{(1)}$ is measurable wrt $\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$ and

$$f^+_{j,n}(\mathcal{Q}_{I,n}):=\mu^{+(1)}_{j+1,n}(\mathcal{Q}_{I,n})/\mu^{+(2)}_{j,n}(\mathcal{Q}_{I,n}) \tag{3.54}$$

with

$$\begin{aligned}\mu^{+(1)}_{j+1,n}(\mathcal{Q}_{I,n})&:=E^+\left(\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1}C^+_{i,j+1}|\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n}\right)\\ &=\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1}E^+(C^+_{i,j+1}|C^+_{i,I-i}=C_{i,I-i}),\\ \mu^{+(2)}_{j,n}(\mathcal{Q}_{I,n})&:=E^+\left(\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1}C^+_{i,j}|\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n}\right)\\ &=\frac{1}{I+n-j}\sum_{i=-n}^{I-j-1}E^+(C^+_{i,j}|C^+_{i,I-i}=C_{i,I-i}).\end{aligned}$$

Besides a correctly chosen and sufficiently smooth parametric family of (conditional) bootstrap distributions of the individual development factors for the *lower triangle*, the derivation of (conditional) bootstrap asymptotic theory and consistency results for $(R^+_{I,n} - \widehat{R}^+_{I,n})_2$ requires additional assumptions for the *backward* individual development factors $G^+_{i,j}$ from Step 3 in Section 3.5.1.

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) (see also Theorem 3.27 in the appendix) are correctly mimicked.

Assumption 3.21 (Consistent estimation of backward moments) *For $n \to \infty$, suppose that the (conditional) bootstrap distributions of the backward individual development factors $G^+_{i,j}$, $j = 0, \ldots, I+n-1$ and $i = -n, \ldots, I-j-1$ given $C^+_{i,j+1}$ and $\mathcal{D}_{I,n}$ are chosen in Step 2 in Section 3.5.1 such that the following holds:*

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

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

where $J_n^{1/2} = diag\left(\sqrt{I+n-j}, j = 0, \ldots, K\right)$ is a diagonal $(K+1) \times (K+1)$ matrix of inflation factors and the variance-covariance matrix $\mathbf{\Sigma}^{(1)}_{K,\underline{f}}$ is defined in Theorem 3.27.

(ii) For each fixed $K \in \mathbb{N}_0$, let $\underline{\widehat{f}}^+_{K,n} = (\widehat{f}^+_{0,n}, \widehat{f}^+_{1,n}, \ldots, \widehat{f}^+_{K,n})'$. Then, conditional on $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, we have

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

in probability, where the variance-covariance matrix $\mathbf{\Sigma}^{(2)}_{K,\underline{f}}$ is defined in Theorem 3.27.

In concordance to the derivation of the conditional limiting result obtained in Theorem 3.10(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 3.21 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 $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$. Precisely, while $(R^+_{I,n} - \widehat{R}^+_{I,n})^{(1)}_2$ is measurable with respect to $\mathcal{D}_{I,n}$, Assumption 3.21 allows to establish asymptotic normality of $\sqrt{I+n+1}(R^+_{I,n} - \widehat{R}^+_{I,n})^{(2)}_2$ conditional on $\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$.

Theorem 3.22 (Bootstrap asymptotics for $(\boldsymbol{R}_{\boldsymbol{I,n}}^{+} - \widehat{\boldsymbol{R}}_{\boldsymbol{I,n}}^{+})_2$ conditional on $\boldsymbol{\mathcal{Q}}_{\boldsymbol{I,n}}^{+} = \boldsymbol{\mathcal{Q}}_{\boldsymbol{I,n}}$ and $\boldsymbol{\mathcal{D}}_{\boldsymbol{I,n}}$) *Suppose Assumptions 3.2, 3.3, 3.5, 3.4, 3.8 and 3.21 hold. Then, as $n \to \infty$, the following holds:*

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

$$\sqrt{I+n+1}(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_2^{(1)} | \left(\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}\right) \xrightarrow{d} \left\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_{\infty}^{(1)} \right\rangle \sim \mathcal{G}_2^{(1)}, \tag{3.55}$$

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

$$Cov(Y_{i_1}^{(1)}, Y_{i_2}^{(1)}) = \lim_{K \to \infty} \boldsymbol{\Sigma}_{K,\prod f_j}^{(1)}(i_1, i_2) \tag{3.56}$$

for $i_1, i_2 \in \mathbb{N}_0$, where $\boldsymbol{\Sigma}_{K,\prod f_j}^{(1)}(i_1, i_2)$ is defined in Corollary 3.28. Here, the two random sequences $\mathcal{Q}_{I,\infty}$ and $\mathbf{Y}_{\infty}^{(1)}$ are stochastically independent.

(ii) Conditionally on $\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, $\sqrt{I+n+1}(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_2^{(2)}$ converges in distribution to $\mathcal{G}_2 | \mathcal{Q}_{I,\infty}$ in probability, where $\mathcal{G}_2 | \mathcal{Q}_{I,\infty} \sim \mathcal{N}(0, \Xi(\mathcal{Q}_{I,\infty})) | \mathcal{Q}_{I,\infty}$ is the (conditional) limiting distribution obtained in Theorem 3.10(ii).

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

$$d_K\left(\mathcal{L}\left((R_{I,n} - \widehat{R}_{I,n})_2 | \mathcal{Q}_{I,n}\right), \mathcal{L}^{+}\left((R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_2 | \mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}\right)\right) \longrightarrow 0$$

in probability.

3.6.1.3 Conditional bootstrap asymptotics for the whole predictive root of the reserve

As in Section 3.4.1.3, we can combine the results on the limiting distributions for $(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_1$ and $(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_2$ from Theorems 3.20 and 3.22, respectively, to get the limiting bootstrap distribution for the whole bootstrap predictive root of the reserve $R_{I,n}^{+} - \widehat{R}_{I,n}^{+}$ conditional on $\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$.

Theorem 3.23 (Bootstrap asymptotics for $\boldsymbol{R}_{\boldsymbol{I,n}}^{+} - \widehat{\boldsymbol{R}}_{\boldsymbol{I,n}}^{+}$ conditional on $\boldsymbol{\mathcal{Q}}_{\boldsymbol{I,n}}^{+} = \boldsymbol{\mathcal{Q}}_{\boldsymbol{I,n}}$ and $\boldsymbol{\mathcal{D}}_{\boldsymbol{I,n}}$) *Suppose the assumptions of Theorems 3.20 and 3.22 hold. Then, conditional on $\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, $(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_1$ and $(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_2$ are stochastically independent, and $R_{I,n}^{+} - \widehat{R}_{I,n}^{+}$ converges in distribution to $\mathcal{G}_1 | \mathcal{Q}_{I,\infty}$ in probability. That is, we have*

$$\begin{aligned} R_{I,n}^{+} - \widehat{R}_{I,n}^{+} | \left(\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}\right) &= (R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_1 + (R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_2 | \left(\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}\right) \\ &\xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I,\infty} \end{aligned}$$

in probability.

According to the discussion below Theorem 3.17 and in view of the concepts of *asymptotic validity* and *asymptotic pertinence* of a bootstrap prediction approach as discussed in Pan and Politis (2016a), the alternative Mack bootstrap can be regarded as asymptotically valid *and* asymptotically pertinent under the stated conditions.

Remark 3.24 (Backward vs. forward bootstrapping) *While a backward bootstrap approach appears to be natural in time series setups addressed in Pan and Politis (2016a), they also propagate a somewhat simpler forward bootstrap approach 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 CLM setup considered here, this is not the case and a (fixed-design) forward bootstrap as proposed by England and Verrall (2006) does not correctly capture the conditional limiting distribution of the estimation uncertainty part, while a backward bootstrap is capable of doing this.

3.7 Simulation Study

In this section, we compare the original Mack bootstrap from Section 3.3 and the alternative Mack bootstrap from Section 3.5 to illustrate our theoretical findings from Sections 3.4 and 3.6 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 3.5.1, but coincides otherwise with the alternative Mack bootstrap. The inclusion of this intermediate Mack-type bootstrap allows to disentangle the effects caused by the backward resampling proposed in Step 2 and by the different centering used in Step 5 of Section 3.5.1 on the finite sample performance. Both aspects constitute the deviance of the alternative Mack bootstrap from the original Mack bootstrap.

3.7.1 Simulation setup

To assure comparability, we pick up the simulation setup employed in Steinmetz and Jentsch (2022, Section 5). That is, in the notion of the asymptotic framework introduced in Section 3.2.1, let $I = 10$ and choose $n \in \{0, 10, 20, 30, 40\}$ leading to the effective number of accident years $I + n + 1 \in \{11, 21, 31, 41, 51\}$. For each n and for different parameter scenarios to be specified below, we generate $M = 500$ loss triangles $\mathcal{D}_{I,n}^{(m)} = \{C_{i,j}^{(m)} | i = -n, \dots, I,\ j = 0, \dots, I + n,\ -n \leq i + j \leq I\}$, $m = 1, \dots, 500$, having diagonals $\mathcal{Q}_{I,n}^{(m)}$ by generating the entries in their first columns $C_{\bullet,0}^{(m)}$ (independently) from a uniform distribution and the individual developments factors $F_{i,j}$ from a

(DGP1) a conditional gamma distribution,

(DGP2) a conditional log-normal distribution,

(DGP3) a conditional left-tail truncated normal distribution (truncated at 0.1).

In all scenarios, the development factors and the variance parameters are specified to fulfill $f_j > 1$ and $\sigma_j^2 > 0$ for all $j = 0, \ldots, I+n-1$ such that f_j and σ_j^2 decrease 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 $\sigma_j^2 = 509,518 \cdot e^{-1-0.7j}$. Further, we distinguish between two different Setups a) and b), where the parameter settings are exactly the same in both cases, but the first column $C_{\bullet,0} = (C_{-n,0}, \ldots, C_{I,0})'$ of the (upper) loss triangle is uniformly distributed on $[120 \times 10^6, 350 \times 10^6]$ in case a) and on $[120 \times 10^4, 350 \times 10^4]$ in case b).

In the following, to evaluate the performance of all bootstrap procedures under study, for each diagonal $\mathcal{Q}_{I,n}^{(m)}$, $m = 1, \ldots, 500$, we would like to know the exact distribution $R_{I,n}^{(m)} - \widehat{R}_{I,n}^{(m)}$ conditional on $\mathcal{Q}_{I,n}^{(m)}$. However, although knowing exactly the stochastic mechanism to generate a loss triangle $\mathcal{D}_{I,n}$, it is not straightforward at all to simulate $R_{I,n}^{(m)} - \widehat{R}_{I,n}^{(m)} | \mathcal{Q}_{I,n}^{(m)}$. This is because $\widehat{R}_{I,n}^{(m)}$ requires a *backward* generation of a loss triangle $\mathcal{D}_{I,n}$ starting with $\mathcal{Q}_{I,n}^{(m)}$. Hence, as a workaround, we simulate instead the distribution of the "true" predictive root $(R_{I,n}^{(m)} - \widehat{R}_{I,n}^{(m)})$ conditional on $\mathcal{Q}_{I,n}^{(m)}$ using a Monte Carlo simulation with $B = 10,000$, since we know the true underlying parametric family of distributions of the individual development factors for each observed triangle $\mathcal{D}_{I,n}^{(m)}$ and the true parameters for the simulation of $R_{I,n}^{(m)}$ for each setup (DGP1)-(DGP3) such that

$$F_{i,j} | C_{i,j} \sim \left(f_j, \frac{\sigma_j^2}{C_{i,j}} \right) \quad \text{for } j = I-i, \ldots, I+n-1 \text{ and } i = -n, \ldots I. \tag{3.57}$$

Next, for each setup (DGP1)-(DGP3) above and for each loss triangle $\mathcal{D}_{I,n}^{(m)}$, $m = 1, \ldots, 500$, we perform three different Mack-type bootstraps based on 10,000 bootstrap replications each to estimate the conditional distributions of the predictive roots of the reserve. That is, we apply the following three bootstrap approaches:

(oMB) original Mack bootstrap (from Section 3.3),

(aMB) alternative Mack-type bootstrap (from Section 3.5),

(iMB) intermediate Mack-type bootstrap (using a *forward* bootstrap in Step 2 of Section 3.5).

The third intermediate Mack-type bootstrap is included to be able 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 3.5.1 on the finite sample performance. For this purpose, we introduce a novel centering term $\widehat{R}_{I,n}^{++}$ defined by

$$\widehat{R}_{I,n}^{++} = \sum_{i=0}^{I+n} C_{I-i,i}^{+} \prod_{j=i}^{I+n-1} \widehat{f}_{j,n}^{*}, \tag{3.58}$$

which deviates from $\widehat{R}_{I,n}^{+}$ in (3.50) as it relies on $\widehat{f}_{j,n}^{*}$ defined in (3.28), but based on (parametrically generated)

$$F_{i,j}^{*}|C_{i,j}^{*}, \mathcal{D}_{I,n} \sim \left(\widehat{f}_{j,n}, \frac{\widehat{\sigma}_{j,n}^{2}}{C_{i,j}^{*}}\right), \tag{3.59}$$

instead of $\widehat{f}_{j,n}^{+}$ defined in (3.48). This choice of the centering term still resembles the decomposition in (3.51), that shares the (sign) properties of (3.31), which is not the case for (3.32). 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 distribution families in (DGP1)-(DGP3) for all three bootstrap approaches, respectively. Finally, knowing the true parametric family of distributions of the $F_{i,j}$'s, which allows to simulate the "true" predictive root of the reserve, we compare the simulation results to investigate the effect of a misspecified parametric family of distributions to generate $R_{I,n}^{*}$ and $R_{I,n}^{+}$.

In Appendix 3.13, we also provide corresponding simulation results that compare just the distribution of the first (i.e. the process uncertainty) parts of the bootstrap predictive roots $(R_{I,n}^{*} - \widehat{R}_{I,n})_1$ and $(R_{I,n}^{+} - \widehat{R}_{I,n}^{+})_1$ conditional on $\mathcal{Q}_{I,n}^{*} = \mathcal{Q}_{I,n}$ or $\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, respectively, with the distribution of $(R_{I,n} - \widehat{R}_{I,n})_1$ conditional on $\mathcal{Q}_{I,n}$. Note that this distribution is straightforward to simulate. As expected, in view of Theorems 3.14 and 3.20, we find essentially no differences between the two construction principles.

3.7.2 Simulation results

First, we consider the bootstrap variances of the bootstrap predictive roots of the reserves obtained for the three Mack-type bootstraps under study. For both Setups a) and b), we find that the alternative Mack-type bootstrap variance is always 1-5 percentage points smaller than the bootstrap variances obtained for the other two approaches, which do not differ much (less than 1 percentage point). This result perfectly agrees to the findings of Theorem 3.16, where the (conditional) variance $\widetilde{\Xi}(\mathcal{Q}_{I,\infty})$, which is mimicked by the original Mack bootstrap and by the intermediate Mack-type bootstrap, is generally larger than the variance $\Xi(\mathcal{Q}_{I,\infty})$ found in Theorem 3.10, which is mimicked by the alternative Mack bootstrap correctly according to Theorem 3.22.

Next, we consider the whole distributions of the bootstrap predictive roots $R_{I,n}^{*(m)} - \widehat{R}_{I,n}^{(m)}$, $R_{I,n}^{+(m)} - \widehat{R}_{I,n}^{+(m)}$ and $R_{I,n}^{+(m)} - \widehat{R}_{I,n}^{++(m)}$ conditional on $\mathcal{Q}_{I,n}^{*} = \mathcal{Q}_{I,n}$ or $\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$, respectively, for $m = 1, \ldots, 500$. Using the Kolmogorov-Smirnov test of level $\alpha = 5\%$ to test the null hypotheses

$$H_0^{*} : \mathcal{L}^{*}\left(R_{I,n}^{*(m)} - \widehat{R}_{I,n}^{(m)} | (\mathcal{Q}_{I,n}^{*(m)} = \mathcal{Q}_{I,n}^{(m)}, \mathcal{D}_{I,n}^{(m)})\right) = \mathcal{L}\left(R_{I,n}^{(m)} - \widehat{R}_{I,n}^{(m)} | \mathcal{Q}_{I,n}^{(m)}\right)$$

$$H_0^{+} : \mathcal{L}^{+}\left(R_{I,n}^{+(m)} - \widehat{R}_{I,n}^{+(m)} | (\mathcal{Q}_{I,n}^{+(m)} = \mathcal{Q}_{I,n}^{(m)}, \mathcal{D}_{I,n}^{(m)})\right) = \mathcal{L}\left(R_{I,n}^{(m)} - \widehat{R}_{I,n}^{(m)} | \mathcal{Q}_{I,n}^{(m)}\right)$$

$$H_0^{++} : \mathcal{L}^{++}\left(R_{I,n}^{+(m)} - \widehat{R}_{I,n}^{++(m)} | (\mathcal{Q}_{I,n}^{+(m)} = \mathcal{Q}_{I,n}^{(m)}, \mathcal{D}_{I,n}^{(m)})\right) = \mathcal{L}\left(R_{I,n}^{(m)} - \widehat{R}_{I,n}^{(m)} | \mathcal{Q}_{I,n}^{(m)}\right)$$

for $m = 1, \ldots, 500$. The resulting percentages of *failed rejections of the null hypotheses* for all three bootstrap approaches, for different n and different families of distributions are summarized in Tables 6 and 7 for Setups a) and b), respectively. While the percentages increase for growing n for all bootstraps and in both Setups a) and b), the alternative Mack-type bootstrap consistently achieves percentages that are always higher by 1-3 percentage points in comparison to the percentages of the two other bootstraps, which turn out to be quite similar throughout. Moreover, the percentages obtained for Setup a) are higher than for Setup b), and choosing the true distributional family for $F_{i,j}^{*}$ appears to be less important than for Setup b). In particular, when choosing a log-normal distribution instead of a truncated normal distribution or vice versa leads to the lowest percentages of failed rejections for Setup b).

Instead of considering the supremum of the difference between the empirical bootstrap cdf and the true cdf of the predictive root of the reserve, we also consider the average over the $M = 500$ simulations of the squared mean of the deviation of the bootstrap distribution given $\mathcal{Q}_{I,n}^{*} = \mathcal{Q}_{I,n}$ or $\mathcal{Q}_{I,n}^{+} = \mathcal{Q}_{I,n}$, respectively, and $\mathcal{D}_{I,n}$ and its true distribution given $\mathcal{Q}_{I,n}$. Therefore, we calculate the mean squared error of each simulation for $b = 1, \ldots, 10,000$ and then consider the root of the overall mean of the mean squared error (RMMSE) over all simulations $M = 500$, that is,

$$RMMSE_{oMB} = \sqrt{\frac{1}{500}\sum_{m=1}^{500}\frac{1}{10,000}\sum_{b=1}^{10,000}\left((R_{I,n}^{*(b)(m)} - \widehat{R}_{I,n}^{(m)}) - (R_{I,n}^{(b)(m)} - \widehat{R}_{I,n}^{(m)})\right)^2}, \quad (3.60)$$

where $(R_{I,n}^{*(b)(m)} - \widehat{R}_{I,n}^{(m)})$ represents the bth ordered Mack-type bootstrap predictive root and $(R_{I,n}^{(b)(m)} - \widehat{R}_{I,n}^{(m)})$ the bth ordered true simulated predictive root for the mth simulation for $m = 1, \ldots, 500$. Similarly, we calculate $RMMSE_{aMB}$ and $RMMSE_{iMB}$ for the alternative Mack bootstrap and for the intermediate Mack bootstrap, respectively.

The results obtained for all $RMMSE$s are summarized for all three bootstrap approaches in Tables 8 and 9, respectively, for Setups a) and b). For increasing n, the $RMMSE$s are decreasing for all bootstrap approaches in both setups, while the alternative Mack bootstrap has, in most cases, the smallest $RMMSE$ in comparison to the intermediate and the original Mack bootstraps.

chosen dist.		**gamma**			**log-normal**			**trunc. normal**		
true dist.	n	oMB	aM	iM	oMB	aM	iM	oMB	aM	iM
gamma	0	0.21	0.22	0.21	0.30	0.33	0.29	0.29	0.28	0.21
	10	0.38	0.49	0.38	0.47	0.48	0.43	0.37	0.41	0.38
	20	0.47	0.56	0.47	0.51	0.56	0.51	0.54	0.53	0.49
	30	0.58	0.64	0.58	0.56	0.61	0.55	0.60	0.65	0.59
	40	0.66	0.70	0.66	0.61	0.66	0.60	0.72	0.76	0.70
log-normal	0	0.20	0.22	0.20	0.27	0.25	0.25	0.24	0.24	0.22
	10	0.37	0.38	0.37	0.38	0.40	0.37	0.37	0.41	0.36
	20	0.45	0.49	0.45	0.48	0.55	0.52	0.45	0.51	0.49
	30	0.51	0.55	0.51	0.57	0.60	0.55	0.51	0.56	0.54
	40	0.57	0.62	0.57	0.60	0.63	0.60	0.63	0.66	0.64
trunc. normal	0	0.13	0.16	0.13	0.30	0.33	0.29	0.24	0.27	0.20
	10	0.42	0.45	0.42	0.44	0.45	0.43	0.38	0.47	0.42
	20	0.54	0.58	0.54	0.53	0.55	0.52	0.58	0.62	0.59
	30	0.57	0.63	0.57	0.57	0.60	0.57	0.67	0.70	0.68
	40	0.60	0.66	0.60	0.57	0.62	0.59	0.75	0.78	0.76

Table 6: Percentages of failed rejections for Kolmogorov-Smirnov tests of level $\alpha = 5\%$ for the null hypotheses H_0^*, H_0^+ and H_0^{++}, respectively, for the original Mack bootstrap (oMB), the alternative Mack bootstrap (aMB) and the intermediate Mack bootstrap (iMB) for different parametric families of distributions of $F_{i,j}^*$ for $i + j \geq I$, for $I = 10$ and different n in Setup a).

chosen dist.		gamma			log-normal			trunc. normal		
true dist.	n	oMB	aM	iM	oMB	aM	iM	oMB	aM	iM
gamma	0	0.10	0.11	0.09	0.11	0.14	0.12	0.09	0.10	0.10
	10	0.19	0.21	0.18	0.23	0.25	0.23	0.16	0.17	0.17
	20	0.28	0.35	0.32	0.36	0.40	0.37	0.21	0.22	0.22
	30	0.38	0.42	0.40	0.40	0.44	0.39	0.28	0.28	0.28
	40	0.52	0.56	0.51	0.51	0.53	0.50	0.41	0.41	0.41
log-normal	0	0.07	0.10	0.08	0.09	0.10	0.08	0.07	0.08	0.08
	10	0.16	0.19	0.17	0.20	0.22	0.20	0.16	0.15	0.15
	20	0.31	0.33	0.30	0.22	0.26	0.23	0.22	0.22	0.22
	30	0.34	0.38	0.33	0.30	0.33	0.30	0.27	0.27	0.27
	40	0.39	0.42	0.39	0.45	0.47	0.44	0.29	0.28	0.28
trunc. normal	0	0.09	0.11	0.08	0.15	0.17	0.16	0.24	0.23	0.23
	10	0.16	0.18	0.15	0.21	0.24	0.20	0.34	0.34	0.34
	20	0.28	0.30	0.27	0.26	0.30	0.27	0.41	0.42	0.42
	30	0.36	0.39	0.35	0.31	0.35	0.32	0.56	0.55	0.55
	40	0.43	0.45	0.42	0.36	0.39	0.36	0.61	0.62	0.60

Table 7: Percentages of failed rejections for Kolmogorov-Smirnov tests of level $\alpha = 5\%$ for the null hypotheses H_0^*, H_0^+ and H_0^{++}, respectively, for the original Mack bootstrap (oMB), the alternative Mack bootstrap (aMB) and the intermediate Mack bootstrap (iMB) for different parametric families of distributions of $F_{i,j}^*$ for $i + j \geq I$, for $I = 10$ and different n in Setup b).

chosen distribution		gamma			log-normal			trunc. normal		
true distribution	n	oMB	aM	iM	oMB	aM	iM	oMB	aM	iM
gamma	0	99.720	99.706	99.600	99.244	99.202	99.967	98.210	98.156	98.822
	10	94.120	93.786	94.320	94.659	94.376	94.885	97.713	97.239	97.854
	20	92.800	92.438	92.760	93.024	92.738	93.048	95.153	94.935	95.296
	30	86.660	85.957	86.120	86.483	85.935	86.143	88.281	87.613	87.790
	40	81.910	81.667	81.510	84.329	81.215	83.828	84.832	83.421	84.990
log-normal	0	99.070	99.080	99.873	99.787	97.971	98.538	98.197	98.112	98.623
	10	93.790	93.587	94.040	94.558	94.389	94.720	97.669	97.389	97.867
	20	92.150	91.996	92.314	90.629	89.843	90.179	95.680	95.365	95.657
	30	86.390	85.607	85.805	87.071	84.159	85.457	88.669	87.945	88.147
	40	81.510	81.226	81.420	82.191	81.522	82.955	83.895	83.665	84.530
trunc. normal	0	96.770	93.993	94.836	97.694	97.519	98.403	93.974	93.450	94.670
	10	94.970	91.815	95.185	93.988	93.774	94.298	92.498	92.216	92.630
	20	90.740	89.767	90.745	91.329	90.016	90.045	89.120	88.406	89.117
	30	86.540	85.786	86.033	86.193	85.567	85.784	87.758	86.896	87.120
	40	82.650	82.277	83.865	81.090	81.374	82.839	83.343	82.455	82.840

Table 8: Root of the overall mean of the mean squared error (RMMSE) ($\times 10^{-3}$) for different Mack-type bootstraps, different distributional assumptions and different n and $I = 10$ for Setup a)

chosen distribution		gamma			log-normal			trunc. normal		
true distribution	n	oMB	aM	iM	oMB	aM	iM	oMB	aM	iM
gamma	0	9.881	9.874	9.961	9.925	9.850	9.919	9.841	9.822	9.849
	10	9.644	9.582	9.680	9.722	9.650	9.742	9.414	9.328	9.442
	20	9.479	9.317	9.476	9.459	9.362	9.477	9.254	9.118	9.265
	30	8.799	8.623	8.754	8.757	8.598	8.692	9.194	8.871	8.648
	40	8.452	8.449	8.544	8.513	8.431	8.534	8.706	8.556	8.589
log-normal	0	9.990	9.914	9.982	9.983	9.893	9.961	9.868	9.857	9.827
	10	9.831	9.752	9.803	9.652	9.635	9.682	9.626	9.537	9.636
	20	9.457	9.399	9.468	9.342	9.243	9.354	9.339	9.292	9.355
	30	8.959	8.933	8.998	8.771	8.764	8.775	8.712	8.654	8.658
	40	8.656	8.584	8.643	8.348	8.325	8.346	8.249	8.184	8.199
trunc. normal	0	9.830	9.789	9.843	9.894	9.876	9.997	9.881	9.874	9.910
	10	9.520	9.513	9.524	9.676	9.583	9.517	9.543	9.474	9.575
	20	9.167	9.156	9.234	9.234	9.227	9.234	9.344	9.323	9.345
	30	8.745	8.692	8.698	8.672	8.660	8.687	8.683	8.630	9.143
	40	8.388	8.356	8.498	8.414	8.376	8.497	8.388	8.321	8.367

Table 9: Root of the overall mean of the mean squared error (RMMSE) ($\times 10^{-3}$) for different Mack-type bootstraps, different distributional assumptions and different n and $I = 10$ for Setup b)

3.8 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 to investigate 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 bootstrap if the parametric family of distributions of the individual development factors is correctly specified in Mack's bootstrap. Otherwise, this will generally not be the case. In contrast, the corresponding (conditional) distribution of the estimation uncertainty part is generally not correctly captured by the Mack 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 (2016a) 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 indeed be asymptotically valid and pertinent. Our findings are illustrated by simulations, which show that the alternative Mack-type bootstrap performs 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.

Appendix

3.9 Auxiliary results for Section 3.4

3.9.1 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 3.2.1.

Theorem 3.25 (Asymptotic normality of $\widehat{f}^*_{j,n}$ conditional on $\mathcal{D}_{I,n}$) *Suppose Assumptions 3.2, 3.3, 3.5 and 3.15 are satisfied and let $\widehat{f}^*_{j,n}$, $j = 0, \ldots, I+n-1$ be defined as in* (3.28) *according to the Mack bootstrap scheme of Section 3.3. Then, as $n \to \infty$, the following holds:*

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

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

where "$\xrightarrow{d}$" denotes convergence in distribution.

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

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

where $J^{1/2} = diag\left(\sqrt{I+n+1-j}, j = 0, \ldots, K\right)$ is a diagonal $(K+1) \times (K+1)$ matrix of inflation factors and $\mathbf{\Sigma}_{K,\underline{f}} = J_g(\underline{\mu}_K)\mathbf{\Sigma}_{K,\underline{C}} J_g(\underline{\mu}_K)' = diag\left(\frac{\sigma_0^2}{\mu_0}, \frac{\sigma_1^2}{\mu_1}, \ldots, \frac{\sigma_K^2}{\mu_K}\right)$ is a diagonal $(K+1) \times (K+1)$ covariance matrix, where

$$\mathbf{\Sigma}_{K,\underline{C}} = Cov(\underline{C}_{i,K}) = \begin{pmatrix} \left(\prod_{k=\min(j_1,j_2)}^{\max(j_1,j_2)-1} f_k\right) \tau^2_{\min(j_1,j_2)} \\ j_1, j_2 = 0, \ldots, K+1 \end{pmatrix} \tag{3.61}$$

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

$$J_g(\underline{x}) = \begin{pmatrix} -\frac{x_1}{x_0^2} & \frac{1}{x_0} & 0 & \cdots & 0 \\ 0 & -\frac{x_2}{x_1^2} & \frac{1}{x_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{x_{K+1}}{x_K^2} & \frac{1}{x_K} \end{pmatrix} \tag{3.62}$$

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 3.25 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

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

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 3.2.1.

Corollary 3.26 (Asymptotic normality for products of $\widehat{f}^*_{j,n}$'s conditional on $\mathcal{D}_{I,n}$) *Suppose the assumptions of Theorem 3.25 hold. Then, as $n \to \infty$, the following holds:*

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

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

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

$$\sqrt{I}\begin{pmatrix} \prod_{j=i}^K \widehat{f}^*_{j,n} - \prod_{j=i}^K \widehat{f}_{j,n} \\ i = 0,\dots,K \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\prod f_j}\right) \quad \textit{in probability,}$$

where $\mathbf{\Sigma}_{K,\prod f_j} = J_h(\underline{f}_K)\mathbf{\Sigma}_{K,\underline{f}} J_h(\underline{f}_K)' = (\mathbf{\Sigma}_{K,\prod f_j}(i_1,i_2))_{i_1,i_2=0,\dots,K}$ *is a* $(K+1)\times(K+1)$ *covariance matrix with entries*

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

for $i_1, i_2 = 0, \ldots, K$. Here, $\mathbf{\Sigma}_{K,\underline{f}}$ is defined in Theorem 3.25(ii) and

$$J_h(\underline{x}) = \begin{pmatrix} \prod_{l=0,l\neq 0}^{K} x_l & \prod_{l=0,l\neq 1}^{K} x_l & \cdots & \prod_{l=0,l\neq K}^{K} x_l \\ 0 & \prod_{l=1,l\neq 1}^{K} x_l & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \prod_{l=K,l\neq K}^{K} x_l \end{pmatrix} \tag{3.63}$$

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

3.9.2 Proof of Theorem 3.25

By construction of the Mack bootstrap estimators $\widehat{f}^*_{j,n}$, $j = 0, \ldots, I+n-1$ according to (3.28), for each fixed $K \in \mathbb{N}_0$, the $K+1$ estimators $\widehat{f}^*_{0,n}, \widehat{f}^*_{1,n}, \ldots, \widehat{f}^*_{K,n}$ are independent conditional on $\mathcal{D}_{I,n}$. Hence, it is actually sufficient to prove part (i). For any fixed j and from (1.8) and (3.28), 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, $(Z_{i,n}, i = -n, \ldots, I-j-1,\ n \in \mathbb{N}_0)$ forms a triangular array of random variables that are independent conditional on $\mathcal{D}_{I,n}$, 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 $\mathcal{D}_{I,n}$, we get

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

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

$$E^*(F^*_{i,j}) = E^*\left(\widehat{f}_{j,n} + \frac{\widehat{\sigma}_{j,n}}{\sqrt{C_{i,j}}}r^*_{i,j}\right) = \widehat{f}_{j,n} + \frac{\widehat{\sigma}_{j,n}}{\sqrt{C_{i,j}}}E^*\left(r^*_{i,j}\right) = \widehat{f}_{j,n}$$

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

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

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

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

such that

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

and, altogether,

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

Letting $n \to \infty$, making use of Assumption 3.5, we get $\widehat{\sigma}^2_{j,n} \to \sigma^2_j$ 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^2_j according to (3.19) and (3.20), 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

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

Due to measurability of all $C_{i,j}$'s with respect to $\mathcal{D}_{I,n}$, we get

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

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

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

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

$$E^* \left(\left(F^*_{i,j} - \widehat{f}_{j,n} \right)^4 \right) = E^* \left(\left(\frac{\widehat{\sigma}_{j,n}}{\sqrt{C_{i,j}}} r^*_{i,j} \right)^4 \right) = \frac{\widehat{\sigma}^4_{j,n}}{C^2_{i,j}} E^* \left(r^{*4}_{i,j} \right)$$

leading to

$$\frac{1}{I+n-K} \sum_{i=-n}^{I-K-1} C^4_{i,j} E^* \left(\left(F^*_{i,j} - \widehat{f}_{j,n} \right)^4 \right) = \frac{\widehat{\sigma}^4_{j,n}}{I+n-K} \sum_{i=-n}^{I-K-1} C^2_{i,j} E^* \left(r^{*4}_{i,j} \right).$$

Further, as $\frac{\widehat{\sigma}^4_{j,n}}{I+n-K} \sum_{i=-n}^{I-K-1} C^2_{i,j} = O_P(1)$, it remains to show that $E^* \left(r^{*4}_{i,j} \right) = O_P(1)$ holds as well. By construction, we have[3]

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

In the following, suppose for convenience that $\widetilde{r}_{t,s} = \widehat{r}_{t,s}$. However, the arguments for $\widetilde{r}_{t,s}$ including re-centering (and re-scaling) are essentially the same, but tedious and lengthy.

[3]Note that we implicitly assume that *only* $\widehat{\sigma}^2_{I+n-1,n}$ is estimated as zero; see Section 3.3.1.

In this case, by plugging-in for $\widetilde{r}_{t,s}$, we get

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

By Assumption 3.15, for $n \to \infty$, we have

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

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

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

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 (3.21) are assumed to form a bounded sequence $((\kappa_j^{(4)}/\sigma_j^4), j \in \mathbb{N}_0)$ again according to Assumption 3.15. □

3.9.3 Proof of Corollary 3.26

The proof follows from an application of the delta method and Theorem 3.25 and is completely analogous to the proof of Corollary 3.2 in Steinmetz and Jentsch (2022). □

3.10 Proofs of Section 3.4

3.10.1 Proof of Theorem 3.14

As the (conditional) L_2-convergence result in Theorem 3.7 implies the (conditional) convergence in distribution in (3.36), for $n \to \infty$, it remains to show

$$(R^*_{I,n} - \widehat{R}_{I,n})_1 | \left(\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}\right) \xrightarrow{d} \mathcal{G}_1 | \mathcal{Q}_{I,\infty} \tag{3.64}$$

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

$$Var^*\left((R^*_{I,n} - \widehat{R}_{I,n})_1 | \mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) \to Var\left((R_{I,\infty} - \widehat{R}_{I,\infty})_1 | \mathcal{Q}_{I,\infty}\right) \tag{3.65}$$

in probability, respectively.

Nevertheless, the asymptotic theory for $(R^*_{I,n} - \widehat{R}_{I,n})_1$ 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 $(R^*_{I,n} - \widehat{R}_{I,n})_1$ 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

$$\begin{aligned}
(R^*_{I,n} - \widehat{R}_{I,n})_1 =& \sum_{i=0}^{I+n} C^*_{I-i,i} \left(\prod_{j=i}^{I+n-1} F^*_{I-i,j} - \prod_{j=i}^{I+n-1} \widehat{f}^*_{j,n} \right) \\
=& \sum_{i=0}^{K} C^*_{I-i,i} \left(\prod_{j=i}^{K} F^*_{I-i,j} - \prod_{j=i}^{K} \widehat{f}^*_{j,n} \right) \\
&+ \sum_{i=0}^{K} C^*_{I-i,i} \left(\prod_{j=i}^{K} F^*_{I-i,j} \left(\prod_{l=K+1}^{I+n-1} F^*_{I-i,l} - 1 \right) - \prod_{j=i}^{K} \widehat{f}^*_{j,n} \left(\prod_{l=K+1}^{I+n-1} \widehat{f}^*_{l,n} - 1 \right) \right) \\
&+ \sum_{i=K+1}^{I+n} C^*_{I-i,i} \left(\prod_{j=i}^{I+n-1} F^*_{I-i,j} - \prod_{j=i}^{I+n-1} \widehat{f}^*_{j,n} \right) \\
=:& A^*_{1,K,I,n} + A^*_{2,K,I,n} + A^*_{3,K,I,n}.
\end{aligned}$$

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 $n \to \infty$ 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 $K \to \infty$, and c) that, for all $\epsilon > 0$, we have

$$\lim_{K\to\infty} \limsup_{n\to\infty} P^*\left(|A^*_{2,K,I,n}| > \epsilon | \mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) = 0 \quad \text{and} \tag{3.66}$$
$$\lim_{K\to\infty} \limsup_{n\to\infty} P^*\left(|A^*_{3,K,I,n}| > \epsilon | \mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) = 0. \tag{3.67}$$

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^2_j and $C^*_{i,j}$, $\widehat{f}^*_{j,n}$, $\widehat{\sigma}^2_{j,n}$, respectively, by Assumption 3.12. Hence, as $\widehat{f}_{j,n} - f_j = O_P((I+n-1)^{-1/2})$, $\widehat{f}^*_{j,n} - \widehat{f}_{j,n} = O_{P^*}((I+n-1)^{-1/2})$ and $\widehat{\sigma}^2_{j,n} - \sigma^2_j = O_P((I+n-1)^{-1/2})$ holds for all fixed $j \in \mathbb{N}_0$, we can conclude that, for all fixed $K \in \mathbb{N}_0$ and as $n \to \infty$, that

$$\sum_{i=0}^{K} C^*_{I-i,i} \left(\prod_{j=i}^{K} F^*_{I-i,j} - \prod_{j=i}^{K} \widehat{f}^*_{j,n}\right) |(\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}) \xrightarrow{d} \sum_{i=0}^{K} C_{I-i,i} \left(\prod_{j=i}^{K} F_{I-i,j} - \prod_{j=i}^{K} f_j\right) |\mathcal{Q}_{I,\infty} \tag{3.68}$$

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 $K \to \infty$, we get immediately

$$\sum_{i=0}^{K} C_{I-i,i} \left(\prod_{j=i}^{K} F_{I-i,j} - \prod_{j=i}^{K} f_j\right) |\mathcal{Q}_{I,\infty} \xrightarrow{d} \sum_{i=0}^{\infty} C_{I-i,i} \left(\prod_{j=i}^{\infty} F_{I-i,j} - \prod_{j=i}^{\infty} f_j\right) |\mathcal{Q}_{I,\infty} \sim \mathcal{G}_1|\mathcal{Q}_{I,\infty}, \tag{3.69}$$

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

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

due to

$$\begin{aligned}
E^*\left(\prod_{j=i}^{K} F^*_{I-i,j} - \prod_{j=i}^{K} \widehat{f}^*_{j,n} \middle| \mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{F}^*_{I,n}\right) &= E^*\left(\prod_{j=i}^{K} F^*_{I-i,j}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{F}^*_{I,n}\right) - \prod_{j=i}^{K} \widehat{f}^*_{j,n} \\
&= \prod_{j=i}^{K} \widehat{f}^*_{j,n} - \prod_{j=i}^{K} \widehat{f}^*_{j,n} = 0
\end{aligned} \tag{3.70}$$

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 (3.70), we get for the variance

$$\begin{aligned}
Var^*(A^*_{1,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}) =& E^*\left(Var^*(A^*_{1,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{F}^*_{I,n})|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) \\
&+ Var^*\left(E^*(A^*_{1,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{F}^*_{I,n})|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) \\
=& E^*\left(\sum_{i=0}^{K} C_{I-i,i} \sum_{j=i}^{K} \left(\prod_{k=i}^{j-1} \widehat{f}^*_{k,n}\right) \widehat{\sigma}^2_{j,n} \left(\prod_{l=j+1}^{K} \widehat{f}^{*2}_{l,n}\right) |\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) \\
=& \sum_{i=0}^{K} C_{I-i,i} \sum_{j=i}^{K} \widehat{\sigma}^2_{j,n} E^*\left(\left(\prod_{k=i}^{j-1} \widehat{f}^*_{k,n}\right)\left(\prod_{l=j+1}^{K} \widehat{f}^{*2}_{l,n}\right)\right) \qquad (3.71)
\end{aligned}$$

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

$$Var^*(A^*_{1,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{F}^*_{I,n}) = \sum_{i=0}^{K} C_{I-i,i} \sum_{j=i}^{K} \left(\prod_{k=i}^{j-1} \widehat{f}^*_{k,n}\right) \widehat{\sigma}^2_{j,n} \left(\prod_{l=j+1}^{K} \widehat{f}^{*2}_{l,n}\right)$$

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 $\widehat{\sigma}^2_{j,n}$ with respect to $\mathcal{D}_{I,n}$. Now, using similar arguments as in Steinmetz and Jentsch (2022, Theorem 4.7) and exploiting the fact that the $\widehat{f}^*_{k,n}$'s are stochastically independent conditional on $\mathcal{D}_{I,n}$, for the expectation in (3.71), we get

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

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

$$E^*\left(\widehat{f}^{*2}_{c,n}|\mathcal{B}_{I,n}(c)\right) = \frac{\widehat{\sigma}^2_{c,n}}{\sum_{k=-n}^{I-c-1} C_{k,c}} + \widehat{f}^2_{c,n}, \qquad (3.72)$$

where $\mathcal{B}_{I,n}(k) = \{C_{i,j}|i = -n, \ldots, I,\ j = 0, \ldots, k,\ i + j \leq I + n\}$ denotes all elements of $\mathcal{D}_{I,n}$ up to its kth column, and because of $\widehat{\sigma}^2_{l,n} \to \sigma^2_l$ in probability for all $l \in \{0, \ldots, K\}$ 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

$$Var^*(A^*_{1,K,I,n}|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}) = \sum_{i=0}^{K} C_{I-i,i} \sum_{j=i}^{K} \left(\prod_{k=i}^{j-1} \widehat{f}_{k,n}\right) \widehat{\sigma}^2_{j,n} \left(\prod_{l=j+1}^{K} \widehat{f}^2_{l,n}\right) + O_P\left(\frac{1}{I+n}\right) \tag{3.73}$$

$$\xrightarrow{p} \sum_{i=0}^{K} C_{I-i,i} \sum_{j=i}^{K} \left(\prod_{k=i}^{j-1} f_k\right) \sigma^2_j \left(\prod_{l=j+1}^{K} f^2_l\right) \tag{3.74}$$

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

$$\sum_{i=0}^{K} C_{I-i,i} \sum_{j=i}^{K} \left(\prod_{k=i}^{j-1} f_k\right) \sigma^2_j \left(\prod_{l=j+1}^{K} f^2_l\right) \longrightarrow \sum_{i=0}^{\infty} C_{I-i,i} \sum_{j=i}^{\infty} \left(\prod_{k=i}^{j-1} f_k\right) \sigma^2_j \left(\prod_{l=j+1}^{\infty} f^2_l\right), \tag{3.75}$$

which equals $Var((R_{I,\infty} - \widehat{R}_{I,\infty})_1|\mathcal{Q}_{I,\infty})$. 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 $E^*(A^*_{2,K,I,n}|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}) = 0$ due to $E^*(A^*_{2,K,I,n}|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n}) = 0$ and, for the variance, we have $Var^*(A^*_{2,K,I,n}|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}) = E^*((A^*_{2,K,I,n})^2|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}) = E^*(E^*((A^*_{2,K,I,n})^2|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n})|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n})$. For the inner expectation, using stochastic independence over accident years leading to stochastic independent summands of $A^*_{2,K,I,n}$ (conditional on $\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n}$, $\mathcal{D}_{I,n}$ and $\mathcal{F}^*_{I,n}$), we get

$$\begin{aligned}
&E^*((A^*_{2,K,I,n})^2|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n}) \\
&= \sum_{i=0}^{K} C^2_{I-i,i} E^*\left(\left(\prod_{j=i}^{K} F^*_{I-i,j}\left(\prod_{l=K+1}^{I+n-1} F^*_{I-i,l} - 1\right) - \prod_{j=i}^{K} \widehat{f}^*_{j,n}\left(\prod_{l=K+1}^{I+n-1} \widehat{f}^*_{l,n} - 1\right)\right)^2 \middle| \mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n}\right) \\
&= \sum_{i=0}^{K} C^2_{I-i,i} \\
&\times \left[E^*\left(\left(\prod_{j=i}^{K} F^*_{I-i,j}\left(\prod_{l=K+1}^{I+n-1} F^*_{I-i,l} - 1\right)\right)^2 \middle| \mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n}\right) - \left(\prod_{j=i}^{K} \widehat{f}^*_{j,n}\left(\prod_{l=K+1}^{I+n-1} \widehat{f}^*_{l,n} - 1\right)\right)^2\right].
\end{aligned}$$

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

$$\begin{aligned}
&\sum_{i=0}^{K} C^2_{I-i,i} E^*\left(\left(\prod_{j=i}^{K} F^*_{I-i,j}\left(\prod_{l=K+1}^{I+n-1} F^*_{I-i,l} - 1\right)\right)^2 \middle| \mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n}\right) \\
&= \sum_{i=0}^{K} C^2_{I-i,i} E^*\left(\prod_{j=i}^{I+n-1} F^{*2}_{I-i,j} - 2\left(\prod_{j=i}^{K} F^{*2}_{I-i,j}\right)\left(\prod_{l=K+1}^{I+n-1} F^*_{I-i,l}\right) + \prod_{j=i}^{K} F^{*2}_{I-i,j} \middle| \mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n},\mathcal{F}^*_{I,n}\right).
\end{aligned} \tag{3.76}$$

Using linearity of expectations, for the first expectation on the last right-hand side of (3.76), due to $F^*_{i,j} = \frac{C^*_{i,j+1}}{C^*_{i,j}}$ and $C^*_{I-i,i} = C_{I-i,i}$, we get

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

By recursively plugging-in, we get

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

Similarly, for the second expectation in (3.76), we get

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

and for the third one, we have

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

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

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

Plugging-in and making use of the fact that the $\widehat{f}^*_{j,n}$'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}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n})\\
&= \sum_{i=0}^{K} C_{I-i,i}\left[\sum_{k=i}^{I+n-1}\widehat{\sigma}^2_{k,n}\left(\prod_{j=i}^{k-1}E^*\left(\widehat{f}^*_{j,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right)\right)\left(\prod_{h=k+1}^{I+n-1}E^*\left(\widehat{f}^{*2}_{h,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right)\right)\right.\\
&\qquad - 2\sum_{k=i}^{K}\widehat{\sigma}^2_{k,n}\left(\prod_{j=i}^{k-1}E^*\left(\widehat{f}^*_{j,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right)\right)\left(\prod_{h=k+1}^{K}E^*\left(\widehat{f}^{*2}_{h,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right)\right)\\
&\qquad \times\left(\prod_{l=K+1}^{I+n-1}E^*\left(\widehat{f}^*_{l,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right)\right)\\
&\qquad \left. + \sum_{k=i}^{K}\widehat{\sigma}^2_{k,n}\left(\prod_{j=i}^{k-1}E^*\left(\widehat{f}^*_{j,n}\right)\right)\left(\prod_{h=k+1}^{K}E^*\left(\widehat{f}^{*2}_{h,n}\right)\right)\right]\\
&= \sum_{i=0}^{K} C_{I-i,i}\left[\sum_{k=i}^{I+n-1}\widehat{\sigma}^2_{k,n}\left(\prod_{j=i}^{k-1}\widehat{f}_{j,n}\right)\left(\prod_{h=k+1}^{I+n-1}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\sum_{p=-n}^{I-h-1}C_{p,h}}\right)\right)\right.\\
&\qquad - 2\sum_{k=i}^{K}\widehat{\sigma}^2_{k,n}\left(\prod_{j=i}^{k-1}\widehat{f}_{j,n}\right)\left(\prod_{h=k+1}^{K}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\sum_{p=-n}^{I-h-1}C_{p,h}}\right)\right)\left(\prod_{l=K+1}^{I+n-1}\widehat{f}_{l,n}\right)\\
&\qquad \left. + \sum_{k=i}^{K}\widehat{\sigma}^2_{k,n}\left(\prod_{j=i}^{k-1}\widehat{f}_{j,n}\right)\left(\prod_{h=k+1}^{K}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\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}\widehat{f}_{j,n}\right)\widehat{\sigma}^2_{k,n}\left(\prod_{h=k+1}^{K}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\sum_{p=-n}^{I-h-1}C_{p,h}}\right)\right)\right.\\
&\qquad \times\left(\left(\prod_{h=K+1}^{I+n-1}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\sum_{p=-n}^{I-h-1}C_{p,h}}\right)\right) - \left(\prod_{l=K+1}^{I+n-1}\widehat{f}_{l,n}\right)\right)\\
&\qquad + \sum_{k=K+1}^{I+n-1}\left(\prod_{j=i}^{k-1}\widehat{f}_{j,n}\right)\widehat{\sigma}^2_{k,n}\left(\prod_{h=k+1}^{I+n-1}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\sum_{p=-n}^{I-h-1}C_{p,h}}\right)\right)\\
&\qquad \left. + \sum_{k=i}^{K}\left(\prod_{j=i}^{k-1}\widehat{f}_{j,n}\right)\widehat{\sigma}^2_{k,n}\left(\prod_{h=k+1}^{K}\left(\widehat{f}^2_{h,n} + \frac{\widehat{\sigma}^2_{h,n}}{\sum_{p=-n}^{I-h-1}C_{p,h}}\right)\right)\left(1 - \prod_{l=K+1}^{I+n-1}\widehat{f}_{l,n}\right)\right]
\end{aligned}
$$

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

$$
E^*\left(\widehat{f}^{*2}_{j,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) = E^*\left(\widehat{f}^{*2}_{j,n}\right) = \widehat{f}^2_{j,n} + \frac{\widehat{\sigma}^2_{j,n}}{\sum_{p=-n}^{I-j-1}C_{p,j}}
$$

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

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

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

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

Next, using the law of iterated expectations and

$$
E\left[\widehat{f}^2_{c,n}+\frac{\widehat{\sigma}^2_{c,n}}{\epsilon^c}|\mathcal{B}_{I,n}(c)\right]=\frac{\sigma^2_c}{\sum_{k=-n}^{I-c-1}C_{k,c}}+f^2_c+\frac{\sigma^2_c}{\epsilon^c}\leq f^2_c+\frac{2\sigma^2_c}{\epsilon^c}
$$

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

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

and, similarly, for the second term, we obtain

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

Together, this term becomes

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

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

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

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

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

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

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

by Assumptions 3.4 and 3.8. Now, letting also $K \to \infty$, the term $\sum_{i=0}^{K} \sum_{k=i}^{K} \sigma_k^2$ also remains bounded due to

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

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

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

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

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

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

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

which vanishes for $K \to \infty$.

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

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

While, for $n \to \infty$, 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 $\prod_{l=K+1}^{\infty} f_l \to 1$ for $K \to \infty$, the first factor above can be bounded by

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

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

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

which is bounded as $\sum_{j=0}^{K}(j+1)\sigma_j^2 \to \sum_{j=0}^{\infty}(j+1)\sigma_j^2 < \infty$ for $K \to \infty$. 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

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

and

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

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 $n \to \infty$ and that its bound converges to zero as $K \to \infty$. By plugging-in and using similar arguments as above for $A^*_{2,K,I,n}$, we get

$$\begin{aligned} E(Var^*(A^*_{3,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n})) &\leq \sum_{i=K+1}^{I+n} \mu_i \sum_{j=i}^{I+n-1} \left(\prod_{h=i}^{j-1} f_h\right) \sigma^2_j \left(\prod_{l=j+1}^{I+n-1} \left(f^2_l + \frac{2\sigma^2_l}{\epsilon^l}\right)\right) \\ &\leq \mu^2_\infty \left(\prod_{l=0}^{\infty} \left(f^2_l + \frac{2\sigma^2_l}{\epsilon^l}\right)\right) \sum_{i=K+1}^{I+n} \sum_{j=i}^{I+n-1} \sigma^2_j, \end{aligned}$$

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

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

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

3.10.2 Proof of Theorem 3.16

Similar to the proof of Theorem 3.14 for the conditional limiting behavior of $(R^*_{I,n} - \widehat{R}_{I,n})_1$ and to the proof of Theorem 4.7 in Steinmetz and Jentsch (2022) for the (unconditional!) limiting behavior of $(R_{I,n} - \widehat{R}_{I,n})_2$, we decompose $(R^*_{I,n} - \widehat{R}_{I,n})_2$ 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 $(R^*_{I,n} - \widehat{R}_{I,n})_2$ with $\sqrt{I+n+1}$, we get

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

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

$$
\begin{aligned}
&\lim_{K\to\infty}\limsup_{n\to\infty} P^*\left(|B^*_{2,K,I,n}| > \epsilon|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) = 0 \quad \text{and} \\
&\lim_{K\to\infty}\limsup_{n\to\infty} P^*\left(|A^*_{3,K,I,n}| > \epsilon|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}\right) = 0.
\end{aligned}
$$

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

$$
B^*_{1,K,I,n} = \sqrt{I+n+1}\sum_{i=0}^{K} C^*_{I-i,i}\left(\prod_{j=i}^{K}\widehat{f}^*_{j,n} - \prod_{j=i}^{K}\widehat{f}_{j,n}\right), \tag{3.77}
$$

where

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

In contrast to the situation in the proof of Theorem 4.7 in Steinmetz and Jentsch (2022), where all $\widehat{f}_{j,n}$'s are indeed affected by conditioning on $\mathcal{Q}_{I,n}$, here, conditional on $\mathcal{D}_{I,n}$, all $\widehat{f}^*_{j,n}$'s are *independent* of the condition $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$. Hence, for $n \to \infty$, the (unconditional!) asymptotic bootstrap theory derived in Theorem 3.25 and Corollary 3.26 leads to

$$B^*_{1,K,I,n} | (\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}) \xrightarrow{d} \langle \mathcal{Q}_{I,K-I}, \mathbf{Y}_K \rangle \, | \mathcal{Q}_{I,\infty}$$

in probability, where $\mathcal{Q}_{I,K-I} = \{C_{I-i,i} | i = 0, \ldots, I + (K-I) = K\}$, $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product in $\mathbb{R}^{K+1}$, and $\mathbf{Y}_K = (Y_i, i = 0, \ldots, K)$ is a $(K+1)$-dimensional multivariate normally distributed random variable with $\mathbf{Y}_K \sim \mathcal{N}\left(0, \mathbf{\Sigma}_{K,\prod f_j}\right)$ with $\mathbf{\Sigma}_{K,\prod f_j}$ defined in Corollary 3.26.

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

$$Cov(Y_{i_1}, Y_{i_2}) = \lim_{K \to \infty} \mathbf{\Sigma}_{K,\prod f_j}(i_1, i_2) = \sum_{j=\max(i_1,i_2)}^{\infty} \frac{\sigma_j^2}{\mu_j} \prod_{l=\max(i_1,i_2), l \neq j}^{\infty} f_l^2 \prod_{m=\min(i_1,i_2)}^{\max(i_1,i_2)-1} f_m \quad (3.78)$$

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

$$
\begin{aligned}
&Var\left(\langle \mathcal{Q}_{I,\infty}, \mathbf{Y}_\infty\rangle \,|\mathcal{Q}_{I,\infty}\right) = \sum_{i=0}^{\infty} Var(C_{I-i,i}Y_i|\mathcal{Q}_{I,\infty}) + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty} Cov(C_{I-i_1,i_1}Y_{i_1}, C_{I-i_2,i_2}Y_{i_2}|\mathcal{Q}_{I,\infty})\\
&= \sum_{i=0}^{\infty} C_{I-i,i}^2 Var(Y_i) + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty} C_{I-i_1,i_1}C_{I-i_2,i_2}Cov(Y_{i_1}, Y_{i_2})\\
&= \sum_{i=0}^{\infty} C_{I-i,i}^2 \sum_{j=i}^{\infty}\frac{\sigma_j^2}{\mu_j}\prod_{\substack{l=i\\ l\neq j}}^{\infty} f_l^2 + \sum_{\substack{i_1,i_2=0\\ i_1\neq i_2}}^{\infty} C_{I-i_1,i_1}C_{I-i_2,i_2} \sum_{j=\max(i_1,i_2)}^{\infty}\frac{\sigma_j^2}{\mu_j}\prod_{\substack{l=\max(i_1,i_2)\\ l\neq j}}^{\infty} f_l^2 \prod_{m=\min(i_1,i_2)}^{\max(i_1,i_2)-1} f_m\\
&= \sum_{j=0}^{\infty}\frac{\sigma_j^2}{\mu_j f_j^2}\left(\sum_{i=0}^{j} C_{I-i,i}^2\prod_{k=i}^{\infty} f_k^2\right) + 2\sum_{j=1}^{\infty}\frac{\sigma_j^2}{\mu_j}\left(\sum_{i_1=1}^{j} C_{I-i_1,i_1}\sum_{i_2=0}^{i_1-1} C_{I-i_2,i_2}\prod_{l=i_2}^{i_1-1} f_l\right)\prod_{k=j}^{\infty} f_k^2\\
&= \sum_{j=0}^{\infty}\frac{\sigma_j^2}{\mu_j f_j^2}\left(\frac{j+1}{j+1}\sum_{i=0}^{j} C_{I-i,i}^2\prod_{k=i}^{\infty} f_k^2\right) + 2\sum_{j=1}^{\infty}\frac{\sigma_j^2}{\mu_j}\left(\frac{j}{j}\sum_{i_1=1}^{j} C_{I-i_1,i_1}\frac{i_1}{i_1}\sum_{i_2=0}^{i_1-1} C_{I-i_2,i_2}\prod_{l=i_2}^{i_1-1} f_l\right)\prod_{k=j}^{\infty} f_k^2\\
&\leq \mu_\infty^2\sum_{j=0}^{\infty}\frac{\sigma_j^2}{\mu_j f_j^2}(j+1)\left(\frac{1}{j+1}\sum_{i=0}^{j} C_{I-i,i}^2\right) + 2\mu_\infty^2\sum_{j=1}^{\infty}\frac{\sigma_j^2}{\mu_j}(j+1)^2\left(\frac{1}{j}\sum_{i_1=1}^{j} C_{I-i_1,i_1}\frac{1}{i_1}\sum_{i_2=0}^{i_1-1} C_{I-i_2,i_2}\right)\\
&= O_P(1)
\end{aligned}
$$

due to $\sum_{j=0}^{\infty}(j+1)^2\sigma_j^2 < \infty$ by Assumption 3.4.

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

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

Using the unbiasedness of $\widehat{f}_{j,n}^*$ conditional on $\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n}$ for $\widehat{f}_{j,n}$, that is,

$$E^*(\widehat{f}_{j,n}^*|\mathcal{Q}_{I,n}^* = \mathcal{Q}_{I,n}) = E^*(\widehat{f}_{j,n}^*) = \widehat{f}_{j,n}$$

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

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

using that, conditional on $\mathcal{D}_{I,n}$, the $\widehat{f}^*_{j,n}$'s are independent of the condition $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$.

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

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

since $\widehat{f}^*_{j,n}$ and $\widehat{f}^*_{k,n}$ are independent for $j \neq k$ and $j,k \in \{0,\ldots,I+n-1\}$ conditional on $\mathcal{D}_{I,n}$.

Similarly, we have

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

leading to

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

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

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

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 $C_{i,j}>\epsilon^j$, $E(\widehat{f}_{c,n}|\mathcal{B}_{I,n}(c))=f_c$, $E(\widehat{\sigma}^2_{c,n}|\mathcal{B}_{I,n}(c))=\sigma^2_c$ as well as

$$E\left(\widehat{f}^2_{c,n}|\mathcal{B}_{I,n}(c)\right)=\frac{\sigma^2_c}{\sum_{k=-n}^{I-c-1}C_{k,c}}+f^2_c\leq\frac{\sigma^2_c}{(I+n-c)\epsilon^c}+f^2_c\leq\frac{\sigma^2_c}{\epsilon^c}+f^2_c \tag{3.79}$$

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

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

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

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

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

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

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

which can be bounded further by

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

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

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

for $K\to\infty$, 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 3.8. Similarly, using that $\prod_{j=K+1}^{\infty}f_j\to 1$ and $\prod_{j=K+1}^{\infty}(f_j^2+\frac{2\sigma_j^2}{\epsilon^j})\to 1$ for $K\to\infty$, 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

$$
\begin{aligned}
&B_{3,K,I,n}^*|(\mathcal{Q}_{I,n}^*=\mathcal{Q}_{I,n},\mathcal{D}_{I,n})\\
&=\sqrt{I+n+1}\sum_{i=K+1}^{I+n}C_{I-i,i}^*\left(\prod_{j=i}^{I+n-1}\widehat{f}_{j,n}^*-\prod_{j=i}^{I+n-1}\widehat{f}_{j,n}\right)|(\mathcal{Q}_{I,n}^*=\mathcal{Q}_{I,n},\mathcal{D}_{I,n}).
\end{aligned}
$$

By the same arguments as used above for $B^*_{2,K,I,n}$, we get $E^*(B^*_{3,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}) = 0$ 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

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

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

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

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 $E^*(\widehat{f}^{*2}_{c,n})$ (for $B^*_{2,K,I,n}$), we obtain

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

$E(\widehat{f}_{c,n}|\mathcal{B}_{I,n}(c)) = f_c$, $E(\widehat{\sigma}^2_{c,n}|\mathcal{B}_{I,n}(c)) = \sigma^2_c$ as well as (3.79), we get

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

such that the leading term of $E(Var^*(B^*_{3,K,I,n}|\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}))$ becomes

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

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

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

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 $\sum_{l=1}^{j-K} l\mu_{l+K} \leq const. j^2$. Consequently, we have

$$\sum_{j=1+K}^{I+n-1}\frac{\sigma^2_j}{\epsilon^j}\sum_{l=1}^{j-K} l\mu_{l+K} \leq const. \sum_{j=1+K}^{\infty} j^2\frac{\sigma^2_j}{\epsilon^j} \to 0$$

as $K\to\infty$ by Assumption 3.8. □

3.10.3 Proof of Theorem 3.17

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

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

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

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

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

3.11 Proofs of Section 3.6

3.11.1 Proof of Theorem 3.20

Following the technique of proof in Theorem 3.14 and using $\widehat{f}_{j,n}-f_j=O_P((I+n-1)^{-1/2})$, $\widehat{f}^*_{j,n}-\widehat{f}_{j,n}=O_{P^*}((I+n-1)^{-1/2})$ and $\widehat{\sigma}^2_{j,n}-\sigma^2_j=O_P((I+n-1)^{-1/2})$ leads to the same limiting result also for the process uncertainty part $(R^+_{I,n}-\widehat{R}^+_{I,n})_1$ of the alternative Mack bootstrap. □

3.11.2 Proof of Theorem 3.22

Following the technique of proof in Theorem 3.16 and exploiting the limiting properties from Assumption 3.21, we get the claimed asymptotic results. □

3.11.3 Proof of Theorem 3.23

Based on the results established in Theorems 3.20 and 3.22, the arguments are completely analogous to those used in the proof of Theorem 3.17. □

3.12 Conditional versions of the CLTs from Steinmetz and Jentsch (2022)

For the sake of completeness, in Theorem 3.27 and Corollary 3.28 below, we summarize the results from Theorem C.1(ii,iv) and Corollary C.2(ii,iv) in Steinmetz and Jentsch (2022).

Theorem 3.27 (Asymptotic normality of $\widehat{f}_j$ conditionally on $\mathcal{Q}_{I,n}$; Theorem C.1(ii,iv) in Steinmetz and Jentsch (2022)) *Suppose Assumptions 3.2, 3.3, 3.4, 3.8 and 3.9 are satisfied. Then, as $n\to\infty$, the following holds:*

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

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

Then, unconditionally, we have

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

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

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

where $\boldsymbol{\Sigma}^{(1)}_{K,\underline{C}}$ is defined in (3.80), has entries

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

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

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

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

where the variance-covariance matrix

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

where $\boldsymbol{\Sigma}^{(2)}_{K,\underline{C}}$ *is defined in* (3.81)*, has entries* $\boldsymbol{\Sigma}^{(2)}_{K,\underline{f}}(j,j) = \sigma^2_{f_j,2} = \frac{\sigma_j^2}{\mu_j} - \sigma^2_{f_j,1}$ *for* $j = 0, \ldots, K$ *and* $\boldsymbol{\Sigma}^{(2)}_{K,\underline{f}}(j_1, j_2) = -\boldsymbol{\Sigma}^{(1)}_{K,\underline{f}}(j_1, j_2)$ *for* $j_1, j_2 = 0, \ldots, K$*,* $j_1 \neq j_2$*.*

We obtain

$$\boldsymbol{\Sigma}^{(1)}_{K,\underline{f}} = J_g \left(\underline{\mu}_K \right) \boldsymbol{\Sigma}^{(1)}_{K,\underline{C}} J_g \left(\underline{\mu}_K \right)' \tag{3.80}$$

and

$$\boldsymbol{\Sigma}^{(2)}_{K,\underline{C}} = E(Var(\underline{C}_{i,K} | C_{i,\infty})), \tag{3.81}$$

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

$$\boldsymbol{\Sigma}_{K,\underline{C}} = \boldsymbol{\Sigma}^{(1)}_{K,\underline{C}} + \boldsymbol{\Sigma}^{(2)}_{K,\underline{C}} \quad \text{and} \quad \boldsymbol{\Sigma}_{K,\underline{f}} = \boldsymbol{\Sigma}^{(1)}_{K,\underline{f}} + \boldsymbol{\Sigma}^{(2)}_{K,\underline{f}}, \tag{3.82}$$

where $\boldsymbol{\Sigma}_{K,\underline{f}} = diag \left(\frac{\sigma_0^2}{\mu_0}, \frac{\sigma_1^2}{\mu_1}, \ldots, \frac{\sigma_K^2}{\mu_K} \right)$.

Corollary 3.28 (Asymptotic normality for products of $\widehat{f}_{j,n}$'s conditionally on $\mathcal{Q}_{I,n}$; Corollary C.2(ii,iv) in Steinmetz and Jentsch (2022)) *Suppose the assumptions of Theorem 3.27 hold. Then, as* $n \to \infty$*, the following holds:*

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

$$\sqrt{I+n+1} \begin{pmatrix} \prod_{j=i}^K f_{j,n}(\mathcal{Q}_{I,n}) - \prod_{j=i}^K f_{j,n} \\ i = 0, \ldots, K \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(0, \boldsymbol{\Sigma}^{(1)}_{K,\prod f_j} \right)$$

where $\boldsymbol{\Sigma}^{(1)}_{K,\prod f_j} = J_h(\underline{f}_K) \boldsymbol{\Sigma}^{(1)}_{K,\underline{f}} J_h(\underline{f}_K)'$ *with* $J_h(\cdot)$ *as defined in* (3.63)*.*

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

$$\sqrt{I+n+1} \begin{pmatrix} \prod_{j=i}^K \widehat{f}_{j,n} - \prod_{j=i}^K f_{j,n}(\mathcal{Q}_{I,n}) \\ i = 0, \ldots, K \end{pmatrix} | \mathcal{Q}_{I,n} \xrightarrow{d} \mathcal{N} \left(0, \boldsymbol{\Sigma}^{(2)}_{K,\prod f_j} \right)$$

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

3.13 Additional simulation results

Note that the first parts of the alternative Mack bootstrap predictive root of the reserve and the intermediate Mack bootstrap predictive root of the reserve are equal. Hence, the following findings hold for both approaches.

Moreover, in both Setups a) and b) for the different distributional assumptions, we applied the Kolmogorov-Smirnov test of level $\alpha=5\%$ to test $(R^{*(m)}_{I,n} - \widehat{R}^{(m)}_{I,n})_1$ given $\mathcal{Q}^{(m)*}_{I,n} = \mathcal{Q}^{(m)}_{I,n}$ and $\mathcal{D}^{(m)}_{I,n}$ for $m=1,\ldots,500$ is normally distributed with zero mean and variance as in (3.35).

For Setup a), it fails to reject the null hypothesis of a Gaussian distribution for about 92% out of $M=500$ samples, if the gamma distribution is used, for about 87% in the case of a log-normal, and for about 95% for a truncated normal distribution to generate the lower bootstrap triangle. The picture is essentially the same for all $n \in \{0, 10, 20, 30, 40\}$. In comparison, for Setup b), the test does always reject the null for the gamma and for log-normal distribution, but only in about 28% out of $M=500$ for the truncated normal distribution. The results are pretty similar for all $n \in \{0, 10, 20, 30, 40\}$.

These findings can be explained by a property of the gamma and the log-normal distribution. Both tend to 'lose' their skewness and excess of kurtosis for $\frac{C^*_{i,j}}{\sigma^2_{j,n}}$ growing large in this parameter setting. Hence, as the range for the entries of the first column in Setup a) is $[120 \times 10^6, 350 \times 10^6]$ with $[120 \times 10^4, 350 \times 10^4]$ for Setup b), we observe more skewness and more excess kurtosis in b) in comparison to a). In particular, this demonstrates that the distribution of the (asymptotically dominating) process uncertainty terms $(R^*_{I,n} - \widehat{R}_{I,n})_1 | (\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n})$ and $(R^+_{I,n} - \widehat{R}^+_{I,n})_1 | (\mathcal{Q}^+_{I,n} = \mathcal{Q}_{I,n}, \mathcal{D}_{I,n})$, respectively, generally does depend on the distribution (family) of the individual development factors also for large (effective) number of accident years $I+n+1$.

As a summary, we show boxplots of skewness and kurtosis as well as arbitrarily chosen density plots for both settings a) and b) in Figures 16 and 17 for $I=10$ and $n=10$ and for all three different distribution assumptions in (DGP1), (DGP2), (DGP3) generated by the original Mack bootstrap. The results do not change for the alternative Mack bootstrap.

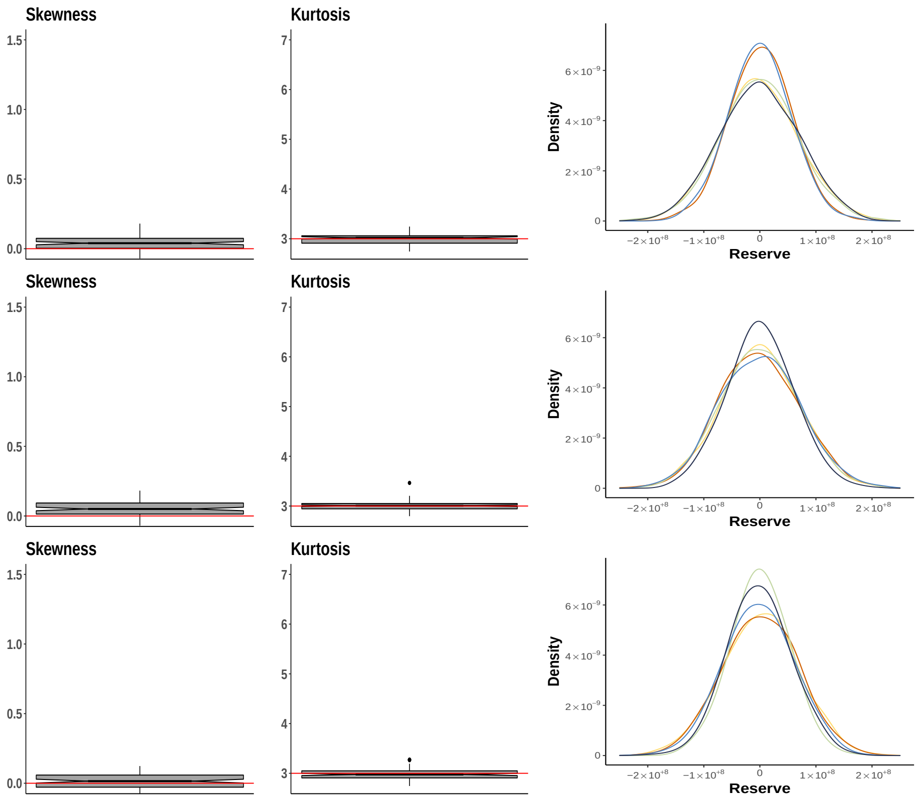

Figure 16: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated Mack type bootstrap conditional distribution of $(R^*_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$ for $n=10$ and $I=10$ for the Setup of a), where $F^*_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom).

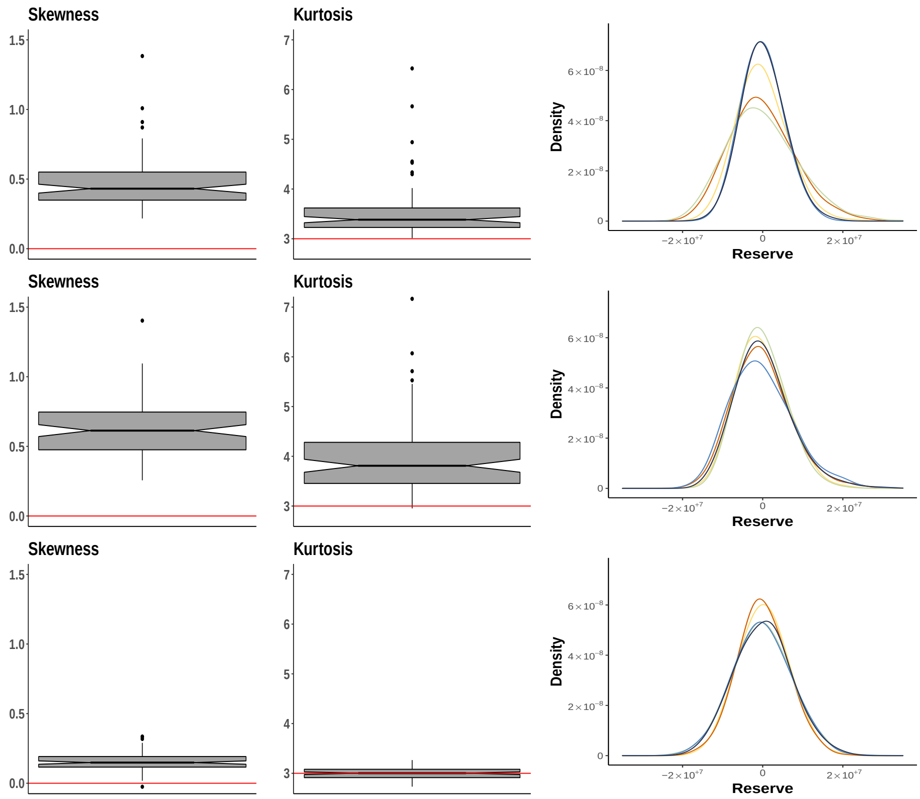

Figure 17: Boxplots of skewness and kurtosis as well as five arbitrarily selected density plots for the simulated Mack type bootstrap conditional distribution of $(R^*_{I,n} - \widehat{R}_{I,n})_1$ given $\mathcal{Q}^*_{I,n} = \mathcal{Q}_{I,n}$ and $\mathcal{D}_{I,n}$ for $n=10$ and $I=10$ for the Setup of b), where $F^*_{i,j}$ follows a (conditional) gamma (top), log-normal (center) and truncated normal distribution (bottom).

Next we compare the bootstrap distribution of $(R_{I,n}^{*(m)}-\widehat{R}_{I,n}^{(m)})_1|(\mathcal{Q}_{I,n}^{(m)*}=\mathcal{Q}_{I,n}^{(m)},\mathcal{D}_{I,n}^{(m)})$ and $(R_{I,n}^{+(m)}-\widehat{R}_{I,n}^{+(m)})_1^{(m)}|(\mathcal{Q}_{I,n}^{(m)+}=\mathcal{Q}_{I,n}^{(m)},\mathcal{D}_{I,n}^{(m)})$, respectively, to the distribution $(R_{I,n}^{(m)}-\widehat{R}_{I,n}^{(m)})_1|\mathcal{Q}_{I,n}^{(m)}$ obtained by Monte Carlo Simulation for $m=1,\ldots,500$. We apply the Kolmogorov-Smirnov test of level $\alpha=5\%$.

Tables 10 and 11 summarize the results for Setup a) and b), respectively, for the original Mack and alternative Mack bootstrap. The results of the original Mack and the alternative Mack bootstrap do not differ.

In general, for increasing n the percentages of fail to reject the null hypothesis increase. If we choose the true underlying distribution, we fail to reject the null hypothesis more frequently than if we choose the wrong distribution. For Setup b) it is more important to choose the true underlying distribution compared to Setup b). If the underlying distribution of the individual development factors is skewed, the chosen distribution for $F_{i,j}^*$ and $F_{i,j}^+$, respectively, for the lower triangle should be skewed. For example, if we choose a gamma distribution for $F_{i,j}^*$ instead of a log-normal distribution as true distributional family of $F_{i,j}^*$, the percentage to fail to reject the null hypothesis is higher compared to if we choose a truncated normal distribution, e.g., for $n=40$, we get that 69% out of $M=500$ fail to reject the null hypothesis assuming a gamma distribution compared to 31% using a truncated normal distribution (cf. Table 11). Also, if the true underlying distribution is a truncated normal distribution and we choose a gamma, then 50% out of $M=500$ fail to reject the null hypothesis or to assume a log-normal distribution, then 39% and if we choose the true underlying distribution, then 84% for $n=40$ (cf. Table 11).

For Setup a) the effect of choosing the wrong distribution is not as high as for b). We can explain this with the property of the gamma and the log-normal distribution. Both tend to 'lose' their skewness and excess of kurtosis for $\frac{C_{i,j}^*}{\widehat{\sigma}_{j,n}^2}$ growing large in this parameter setting.

chosen distribution		**gamma**		**log-normal**		**trunc. normal**	
true distribution	n	oMB	aMB	oMB	aMB	oMB	aMB
gamma	0	0.57	0.55	0.45	0.50	0.49	0.51
	10	0.66	0.69	0.65	0.70	0.58	0.71
	20	0.73	0.72	0.72	0.79	0.68	0.80
	30	0.75	0.73	0.75	0.83	0.72	0.81
	40	0.80	0.76	0.79	0.87	0.79	0.83
log-normal	0	0.44	0.47	0.57	0.56	0.45	0.47
	10	0.60	0.61	0.69	0.68	0.60	0.62
	20	0.69	0.73	0.78	0.77	0.70	0.64
	30	0.70	0.77	0.83	0.80	0.78	0.70
	40	0.81	0.80	0.89	0.85	0.80	0.73
trunc. normal	0	0.46	0.49	0.45	0.52	0.50	0.48
	10	0.62	0.63	0.71	0.67	0.59	0.57
	20	0.67	0.68	0.78	0.75	0.71	0.71
	30	0.72	0.71	0.80	0.80	0.75	0.73
	40	0.76	0.73	0.82	0.82	0.84	0.85

Table 10: Process Uncertainty: Percentages of failed rejections for Kolmogorov-Smirnov tests of level α=5% for the null hypotheses $\mathcal{L}^*((R^*_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n})=\mathcal{L}((R_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}_{I,n})$ and $\mathcal{L}^*((R^+_{I,n}-\widehat{R}^+_{I,n})_1|\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n})=\mathcal{L}((R_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}_{I,n})$, respectively, for the original Mack bootstrap (oMB) and the alternative Mack bootstrap (aMB) for different parametric families of distributions of $F^*_{i,j}$ for $i+j\geq I$, for I=10 and different n in Setup a).

chosen distribution		**gamma**		**log-normal**		**trunc. normal**	
true distribution	n	oMB	aMB	oMB	aMB	oMB	aMB
gamma	0	0.52	0.48	0.40	0.35	0.34	0.34
	10	0.66	0.63	0.44	0.53	0.49	0.47
	20	0.77	0.72	0.66	0.66	0.58	0.51
	30	0.80	0.76	0.70	0.68	0.60	0.53
	40	0.83	0.80	0.71	0.75	0.61	0.57
log-normal	0	0.32	0.33	0.44	0.41	0.18	0.18
	10	0.50	0.54	0.55	0.50	0.21	0.22
	20	0.60	0.55	0.65	0.63	0.23	0.25
	30	0.61	0.60	0.68	0.72	0.29	0.28
	40	0.69	0.67	0.78	0.75	0.31	0.30
trunc. normal	0	0.30	0.33	0.21	0.26	0.49	0.51
	10	0.40	0.45	0.29	0.33	0.64	0.60
	20	0.43	0.49	0.39	0.35	0.78	0.78
	30	0.45	0.54	0.40	0.38	0.80	0.85
	40	0.50	0.59	0.42	0.45	0.84	0.87

Table 11: Process Uncertainty: Percentages of failed rejections for Kolmogorov-Smirnov tests of level α=5% for the null hypotheses $\mathcal{L}^*((R^*_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}^*_{I,n}=\mathcal{Q}_{I,n})=\mathcal{L}((R_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}_{I,n})$ and $\mathcal{L}^*((R^+_{I,n}-\widehat{R}^+_{I,n})_1|\mathcal{Q}^+_{I,n}=\mathcal{Q}_{I,n})=\mathcal{L}((R_{I,n}-\widehat{R}_{I,n})_1|\mathcal{Q}_{I,n})$, respectively, for the original Mack bootstrap (oMB) and the alternative Mack bootstrap (aMB) for different parametric families of distributions of $F^*_{i,j}$ for $i+j\geq I$, for I=10 and different n in Setup b).

4 Conclusion and Outlook

In this dissertation, we derived the limit unconditional and conditional distribution of the predictive root of the reserve based on Mack's model and showed bootstrap consistency results.

We proposed a general and fully described stochastic framework based on Mack's model and a suitable asymptotic framework for a loss triangle. We derived for an increasing number of accident years CLTs for (smooth functions of) the parameter estimators in Mack's model, which allows for asymptotic inference for the development factors and variance parameters.
We obtained the limit unconditional and conditional distribution of the predictive root of the reserve. Therefore, we split the predictive root of the reserve into two additive random parts covering the process and estimation uncertainty, whereas the estimation uncertainty is asymptotically negligible.
Unconditionally, but also conditional on either the whole observed upper loss triangle or the last observed diagonal we showed that the process uncertainty part is non-Gaussian distributed and the limiting distribution depends on the parametric family of individual development factors. By contrast, the parameter uncertainty part is measurable with respect to the observed loss triangle. We derived that the estimation uncertainty part has to be inflated properly and follows only conditional on the last diagonal of the loss triangle a normal distribution. We proved that the unconditional limit distribution of the estimation uncertainty is non-Gaussian.
Altogether we got that the predictive root of the reserve has unconditional and conditional the same limit distribution as the process uncertainty part which is in general a non-Gaussian distribution. This cast the common practice to use a normal approximation together with the estimators of the moments of the reserve by Mack's model into doubt.

We used the derived conditional limit distributions of the process uncertainty and the estimation uncertainty, as a benchmark and compared them with their bootstrap versions. For the Mack bootstrap proposed by England and Verrall (2006) we showed that the bootstrap approach is only valid. We showed that the Mack bootstrap mimics correctly the (condional) distribution process uncertainty part if the parametric family of distributions of the individual development factors is correctly specified in the bootstrap proposal. In

contrast, we showed that the parameter uncertainty part is not captured correctly by the Mack bootstrap. The limit conditional variance of the Mack bootstrap parameter uncertainty part is bigger compared to the conditional variance of the parameter uncertainty part. Hence, the Mack bootstrap does not correctly mimic the (conditional) distribution of the parameter uncertainty part. Using these results together with the fact that the estimation uncertainty part is asymptotically negligible, we derived that the (conditional) bootstrap distribution of the predictive root of the reserve is correctly mimicked and called valid if the parametric family of distributions of the individual development factors is correctly specified in Mack's bootstrap. Asymptotic validity of a bootstrap approach is a fundamental property, but especially for finite samples it is beneficial to capture also the uncertainty due to model estimation correctly.
Therefore, we proposed a new more natural alternative Mack bootstrap, that is centered differently. The new bootstrap proposal generates recursively backward based on the diagonal $\mathcal{Q}^{+}_{I,n}=\mathcal{Q}_{I,n}$ new upper loss triangles. Bootstrap estimators for the development factors are derived based on the bootstrap-generated loss triangles. The conditional variance of the backward bootstrap estimation uncertainty part is asymptotically the same as the variance of the estimation part conditional on the last observed diagonal. In addition, the conditional estimation uncertainty distribution is correctly mimicked by the new bootstrap.
We showed that the new alternative Mack bootstrap is able to mimic the limit (conditional) distribution of the process uncertainty and the predictive root of the reserve, if the parametric family of distributions of the individual development factors is correctly specified, and in addition, it is able to mimic the (conditional) limit distribution of the estimation uncertainty correctly. Therefore, the new alternative Mack bootstrap is valid and especially pertinent, since it captures in addition correctly the (asymptotic negligible) estimation uncertainty part.

In the future, we want to apply the alternative Mack bootstrap not only to simulated data but also to real-world loss data and analyze its performance in comparison to the original Mack bootstrap.

Also, the assumptions and the proposed framework can be used to define a model for quarterly data, since the claims data are collected quarterly. A quarterly model will lead to a bigger set of residuals that can be used for a non-parametric bootstrap step. Also, patterns of the development of claims can be easier identified in quarterly data. Additionally, it would be interesting to analyze the performance of the bootstrap based on quarterly data.

In addition, we are interested in the results for the limit distribution of the predictive root if we assume that claims $C_{i,0}$ are not identical, but still independently distributed. In practice the $C_{i,0}$ may not the identically distributed since the underwriting of the line of business may have changed over the years or the business line is growing over the years.

Mack's model takes a so called ultimate view of the reserve risk. The ultimate view considers the uncertainty of future claims between the latest observed cumulative claim for an accident year i and its ultimate, i.e., estimating the uncertainty of the claims until they finally settled. For the one-year view perspective, we look at the uncertainty of the best estimate reserve made at time I and after one year, i.e., $I+1$. Therefore, we look at the difference between the best estimate of the reserve made at time I and the best estimate made at $I+1$ considering the claim payments made during this year. The difference is called claims development result (CDR).
For future work we want to apply the general and fully stochastic model to the one year view for the reserve risk to derive the (conditional) limit distribution of the CDR and to derive bootstrap consistency results by the tailor-made one-year bootstrap.
Also, we want to adjust the new alternative Mack bootstrap algorithm for the one-year view, and derive bootstrap consistency results.

The concept of the predictive root of the reserve could be applied together with the asymptotic framework of the loss triangles to other prediction models for reserving, e.g., reserving models based on general linear models. Their asymptotic limit distributions could be derived and used as a benchmark to derive bootstrap consistency results since also for the GLM type models England (2002) and England and Verrall (1999, 2006) proposed a tailor-made bootstrap approach. It would be interesting to check the proposed bootstrap type for validity and pertinence. If required, a new bootstrap type can be proposed again that would be pertinent, not only valid.

Taken together, we showed that the limit distribution of the reserve is in general non-Gaussian and casts the common practice to use a normal approximation for the reserve in Mack's model into doubt. In addition, we showed that the Mack bootstrap only mimics under mild assumptions the process uncertainty part correctly but not the estimation uncertainty. Therefore, we suggest an alternative Mack bootstrap.

References

Beutner, E., Heinemann, A., and Smeekes, S. (2021). A justification of conditional confidence intervals. In: *Electronic Journal of Statistics* 15.1:pp. 2517–2565.

Björkwall, S., Hössjer, O., and Ohlsson, E. (2009). Non-parametric and parametric bootstrap techniques for age-to-age development factor methods in stochastic claims reserving. In: *Scandinavian Actuarial Journal* 2009.4:pp. 306–331.

Björkwall, S., Hössjer, O., and Ohlsson, E. (2010). Bootstrapping the separation method in claims reserving. In: *ASTIN Bulletin: The Journal of the IAA* 40.2:pp. 845–869.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods.* Springer, New York.

Bruce, N., Chen, C., Dunne, G., Hinder, I., McMurrough, T., Meyers, G., White, A., and Wright, T. (2008). *Best Estimates and Reserving Uncertainity.* Actuarial Profession General Insurance (GIRO) Convention. Available at: `https://www.actuaries.org.uk/system/files/documents/pdf/bhprizegibson.pdf`.

Buchwalder, M., Bühlmann, H., Merz, M., and Wüthrich, M. V. (2006). The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited). In: *ASTIN Bulletin: The Journal of the IAA* 36.2:pp. 521–542.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. In: *Annals of Statistics* 7:pp. 1–26.

England, P. D. (2002). Addendum to “Analytic and bootstrap estimates of prediction errors in claims reserving”. In: *Insurance: Mathematics and Economics* 31.3:pp. 461–466.

England, P. D. and Verrall, R. J. (1999). Analytic and bootstrap estimates of prediction errors in claims reserving. In: *Insurance: Mathematics and Economics* 25.3:pp. 281–293.

England, P. D. and Verrall, R. J. (2002). Stochastic claims reserving in general insurance. In: *British Actuarial Journal* 8.3:pp. 443–518.

England, P. D. and Verrall, R. J. (2006). Predictive distributions of outstanding liabilities in general insurance. In: *Annals of Actuarial Science* 1.2:pp. 221–270.

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., and Wright, T. (2007). *Best Estimates and Reserving Uncertainity.* Actuarial Profession General Insurance (GIRO) Convention. Available at: `https://www.actuaries.org.uk/system/files/documents/pdf/bhprizegibson.pdf`.

Gisler, A. (2019). The Reserve Uncertainties in the Chain Ladder Model of Mack Revisited. In: *ASTIN Bulletin: The Journal of the IAA* 49.3:pp. 787–821.

Gisler, A. (2021). Estimation error and bootstrapping in the chain-ladder model of Mack. In: *European Actuarial Journal* 11.1:pp. 269–283.

Harnau, J. and Nielsen, B. (2018). Over-dispersed age-period-cohort models. In: *Journal of the American Statistical Association* 113.524:pp. 1722–1732.

Hartl, T. (2010). “Bootstrapping generalized linear models for development triangles using deviance residuals”. In: *CAS E–Forum Fall.*

Hindley, D. (2017). *Claims Reserving in General Insurance*. Cambridge University Press.

Kuang, D. and Nielsen, B. (2020). Generalized log-normal chain-ladder. In: *Scandinavian Actuarial Journal* 2020.6:pp. 553–576.

Lindholm, M., Lindskog, F., and Wahl, F. (2020). Estimation of conditional mean squared error of prediction for claims reserving. In: *Annals of Actuarial Science* 14.1:pp. 93–128.

Mack, T. (1993). Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates. In: *ASTIN Bulletin: The Journal of the International Actuarial Association* 23.02:pp. 213–225.

Mack, T. (1994). Which stochastic model is underlying the chain ladder method? In: *Insurance: Mathematics and Economics* 15.2-3:pp. 133–138.

Mack, T., Quarg, G., and Braun, C. (2006). The Mean Square Error of Prediction in the Chain Ladder Reserving Method – A Comment. In: *ASTIN Bulletin: The Journal of the IAA* 36.2:pp. 543–552.

Pan, L. and Politis, D. N. (2016a). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. In: *Journal of Statistical Planning and Inference* 177:pp. 1–27.

Pan, L. and Politis, D. N. (2016b). Bootstrap prediction intervals for Markov processes. In: *Computational Statistics & Data Analysis* 100:pp. 467–494.

Paparoditis, E. and Shang, H. L. (2021). Bootstrap prediction bands for functional time series. In: *Journal of the American Statistical Association.*

Peremans, K., Segaert, P., Van Aelst, S., and Verdonck, T. (2017). Robust bootstrap procedures for the chain-ladder method. In: *Scandinavian Actuarial Journal* 2017.10:pp. 870–897.

Pešta, M. and Hudecová, Š. (2012). Asymptotic consistency and inconsistency of the chain ladder. In: *Insurance: Mathematics and Economics* 51.2:pp. 472–479.

Peters, G. W., Wüthrich, M. V., and Shevchenko, P. V. (2010). Chain ladder method: Bayesian bootstrap versus classical bootstrap. In: *Insurance: Mathematics and Economics* 47.1:pp. 36–51.

Pinheiro, P. JR, Andrade e Silva, J. M., and Lourdes Centeno, M. de (2003). Bootstrap methodology in claim reserving. In: *Journal of Risk and Insurance* 70.4:pp. 701–714.

Renshaw, A. E. and Verrall, R. J. (1998). A Stochastic Model Underlying the Chain-Ladder Technique. In: *British Actuarial Journal* 4.4:pp. 903–923.

Siegenthaler, F. (2021). The ultimate prediction uncertainty in the chain-ladder model of Mack: An unbiased estimator. In: *Available at SSRN 3812966.*

Steinmetz, J. and Jentsch, C. (2022). Asymptotic theory for Mack’s model. In: *Insurance: Mathematics and Economics* 107:pp. 223–268.

Tee, L., Käärik, M., and Viin, R. (2017). On comparison of stochastic reserving methods with bootstrapping. In: *Risks* 5.1:p. 2.

Verdonck, T. and Debruyne, M. (2011). The influence of individual claims on the chain-ladder estimates: Analysis and diagnostic tool. In: *Insurance: Mathematics and Economics* 48.1:pp. 85–98.

Wüthrich, M. V. and Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. Vol. 435. John Wiley & Sons.

Eidesstattliche Versicherung

Ich, Julia Steinmetz (Matrikelnummer: 215880), versichere hiermit an Eides statt, dass ich die vorliegende Doktorarbeit mit dem Titel "Statistical Inference for the Reserve Risk" selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt so wie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Dortmund, February 10, 2023

Julia Steinmetz

Belehrung

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz - HG -). Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft. Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen. Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Dortmund, February 10, 2023

Julia Steinmetz