Article

Valuation of Non-Life Liabilities from Claims Triangles

Mathias Lindholm [iD]**, Filip Lindskog** * [iD] **and Felix Wahl**

Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden; lindholm@math.su.se (M.L.); felixw@math.su.se (F.W.)
* Correspondence: lindskog@math.su.se; Tel.: +46-8-164564

Abstract: This paper provides a complete program for the valuation of aggregate non-life insurance liability cash flows based on claims triangle data. The valuation is fully consistent with the principle of valuation by considering the costs associated with a transfer of the liability to a so-called reference undertaking subject to capital requirements throughout the runoff of the liability cash flow. The valuation program includes complete details on parameter estimation, bias correction and conservative estimation of the value of the liability under partial information. The latter is based on a new approach to the estimation of mean squared error of claims reserve prediction.

1. Introduction

The purpose of the present paper is to describe how the cost-of-capital approach from Engsner et al. (2017), which builds on the approach to market-consistent liability valuation from Möhr (2011), may be implemented using non-life claims triangle data covering both premium and reserve risk. In Engsner et al. (2017), general formulas for cost-of-capital valuation of liability cash flows are provided. For this to become practically relevant in a non-life insurance context, we need to define cash flow dynamics that are consistent with those seen in actual non-life insurance applications.

The undoubtedly most common claims triangle reserving method used in practice is the (distribution-free) chain-ladder method defined in Mack (1993a). This is a pure reserving method, and there is from that perspective no need to define how the model is initiated, since reserving amounts to estimating outstanding costs for already incurred claims. This is in contrast to premium risk, where there is a need to describe how not yet incurred claims will generate future payments. Further, the pan-European insurance regulation Solvency II is constructed to measure, amongst other things, non-life insurance risk expressed in terms of premium and reserve risk. Concerning reserve risk, the Solvency II regulation's so-called standard formula is based on a one-year re-reserving risk based on a time series version of the classical chain-ladder method; see Merz and Wüthrich (2008). Regarding premium risk, there is no single standard method; for more on this, see, e.g., Gisler (2009); Ohlsson and Lauzeningks (2009). One can here note that these methods do not necessarily have a one-to-one correspondence with the assumptions underlying, e.g., the standard chain-ladder method. We can also remark that there are connections between the consistent multi-period valuation approach from Engsner et al. (2017) and the simplified inconsistent multi-period valuation approach provided by EIOPA; see the definition of the so-called "risk margin" in (European Commission 2015, Article 37).

Moreover, the valuation formulas from Engsner et al. (2017) are based on the payment process dynamics and the resulting capital costs from repeated capital requirements. In practice, we need to define a specific model, estimate parameters and assess both process and estimation uncertainty. Further, in Engsner et al. (2017), it is shown that the general cost-of-capital valuation formulas become analytically tractable when the underlying class of models describing the payment dynamics are

Gaussian. Due to this, we give two examples of Gaussian time series payment process models. These claims' payment generating processes will, in turn, induce natural predictors for outstanding costs stemming from incurred, as well as not incurred claims. The Gaussian models introduced in Section 4 are merely two examples of reasonable models, which in many aspects are similar to the standard chain-ladder method, but which allow for premium risk calculations. For more on time series versions of the chain-ladder method, see, e.g., Buchwalder et al. (2006); Kremer (1984) and (Wüthrich and Merz 2008, chp. 3) and the references therein. Furthermore, for the models introduced in Section 4, we derive parameter estimators using conditional weighted least squares and show that the natural predictors for outstanding payments using plug-in estimation are unbiased. Since the valuation formulas are non-linear functions of the model parameters, plug-in estimators of liability values will be biased. As a consequence of this, we derive an approximate plug-in bias correction for the valuation formulas, and based on that, we approximate the parameter uncertainty using a type of resampled pseudo estimator; see Section 3. These pseudo estimators share many features with the approach to approximate prediction uncertainty in Mack (1993a) and Buchwalder et al. (2006). One can also note that the Gaussian payment processes introduced in Section 4 will induce claims reserving methods. By using the idea of resampled pseudo estimators, we may calculate a type of mean squared error of prediction (MSEP) for the induced reserve estimators, which are consistent with the plug-in bias correction for the valuation formula. In fact, in Lindholm et al. (2017), it is shown that the idea of using resampled pseudo estimators may be used to retrieve the standard MSEP approximation for the chain-ladder method derived in Mack (1993a).

Examples of papers relating to risk-margin-like calculations, apart from the ones mentioned above, are Salzmann and Wüthrich (2010), Wüthrich et al. (2011) and Ferriero (2016). It is demonstrated in Salzmann and Wüthrich (2010) and commented on in Wüthrich et al. (2011) that a cost-of-capital approach to risk-margin calculations leads to evaluations of path-dependent multi-period risk measures that in most cases can neither analytically, nor numerically be calculated. Various proxies are presented in Salzmann and Wüthrich (2010). In Wüthrich et al. (2011), an alternative, computationally more attractive risk margin is defined as the difference between a risk-adjusted reserve and a classical reserve and analyzed using probability-distortion techniques and a Bayesian chain-ladder method. In Ferriero (2016), reserve losses are modeled directly using a stochastic process designed to replicate actuary behavior in repeated re-estimation of reserves. An approximation is made in Ferriero (2016) that allows the risk margin to be approximated without having to solve a backward recursion. In the current paper, we define the risk margin using the multi-period cost-of-capital approach presented in Engsner et al. (2017), in line with how it is defined in Möhr (2011) and Salzmann and Wüthrich (2010). However, by assuming that aggregate paid amounts for different accident and development periods have Gaussian joint distributions, we are in the exceptional situation where the risk margin can be calculated analytically as a solution to a backward recursion. As with any model assumptions, the extent to which specific claims triangle data are consistent with Gaussian model assumptions should be assessed. See Section 5 for further details. In situations where it is likely that aggregate paid amounts are essentially due to a few large claims, Gaussian model assumptions are not suitable.

To conclude, we provide a detailed procedure for how to carry out a consistent multi-period cost-of-capital liability valuation based on claims triangles. This is done for two Gaussian cash flow models including parameter estimators, assessment of parameter uncertainty and a full description of the underlying claims reserving method, which is implied by the cash flow model.

The paper is organized as follows: general results on cost-of-capital valuation for Gaussian cash flows are given in Section 2, which is followed by Section 3, where necessary results on estimation and methods to deal with parameter uncertainty are described. In Section 4, two specific Gaussian cash flow models are introduced and analyzed in detail, and we conclude with a short section on model evaluation (see Section 5), together with numerical illustrations (see Section 6). In order to make the exposition more comprehensive, all proofs are collected in the Appendix.

2. Valuation of Non-Life Insurance Cash Flows

In order to present our approach to the valuation of non-life insurance liabilities, we first need to define the liability cash flows as stochastic processes of a certain kind. We take a rolling one-year risk-free bond as the basis for the numéraire process with value one at the current time. In what follows, unless stated otherwise, all future cash flows are discounted by this numéraire or, equivalently, denoted in units of the numéraire process. Similarly, using historic one-year discount factors from data on one-year risk-free bonds, we may adjust historic cash flow data to be expressed in units of the numéraire process. This data adjustment implies that predictions based on historic payment data are predictions adjusted for the time-value of money. Other choices of traded numéraire processes are possible.

Consider K lines of business providing one-year insurance products: the buyer of the insurance product is not entitled to economic compensation for accidents later than twelve months from the time the product is purchased. For $k \in \{1,\ldots,K\} =: \mathcal{K}$, let $I_{ij}^{(k)}$ denote the incremental claims payment during development year $j \in \{1,\ldots,J^{(k)}\} =: \mathcal{J}^{(k)}$, and let $C_{ij}^{(k)}$ denote the cumulative payments during development years $1,\ldots,j \in \mathcal{J}^{(k)}$ from accidents during accident year $i \in \{i_0^{(k)},\ldots,J^{(k)}+1\} =: \mathcal{I}^{(k)}$, where $i_0^{(k)} \leq 1$. For $t \in \{i_0^{(k)} - J^{(k)},\ldots,J^{(k)}\}$, let:

$$D_t^{(k)} := \{I_{ij}^{(k)} : (i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}, i+j \leq J^{(k)}+1+t\}$$

denote the data from business line k that are known at time t. Calendar year $t = 0$ corresponds to the current time. Such a dataset is commonly referred to as a claims triangle, although claims trapezium may be more appropriate. Notice that accident years $i_0^{(k)},\ldots,1$ are fully developed. Notice also that accidents during accident year $J^{(k)}+1$ have not yet occurred. Including accident year $J^{(k)}+1$ in the setup allows for so-called premium risk. Omitting accident year $J^{(k)}+1$ implies that only so-called reserve risk can be studied in the setup. Let:

$$\mathcal{F}_t^{(k)} := \sigma(D_t^{(k)}), \quad \mathcal{F}_t := \sigma(\cup_{k \in \mathcal{K}} D_t^{(k)}),$$

where we set:

$$D_t^{(k)} := \begin{cases} \varnothing & \text{if } t < i_0^{(k)} - J^{(k)}, \\ D_{J^{(k)}}^{(k)} & \text{if } t > J^{(k)}. \end{cases}$$

The overall aim is to assign a value to the future liability cash flow from both active contracts (non-incurred claims) and claims from contracts that may still generate claims (incurred, but not reported and reported, but not settled claims). Set $T = \max\{J^{(k)} : k \in \mathcal{K}\}$ and:

$$X_t^{o,(k)} := \begin{cases} \sum_{u=t+1}^{J^{(k)}+1} I_{u,J^{(k)}-u+t+1}^{(k)}, & t = 1,\ldots,J^{(k)}, \\ 0, & t = J^{(k+1)}+1,\ldots,T. \end{cases} \tag{1}$$

Notice that $i = J^{(k)}+1$ corresponds to non-incurred claims from still active contracts. Further, note that $X_t^{o,(k)}$ corresponds to the payments during a specific year for all accident years simultaneously, i.e., $X_t^{o,(k)}$ corresponds to a diagonal in an incremental claims payment triangle. Furthermore, set:

$$X_t^o := \sum_{k=1}^K X_t^{o,(k)}, \quad R^{(k)} := \sum_{t=1}^T X_t^{o,(k)}, \quad R := \sum_{k=1}^K R^{(k)}.$$

For each k, if accident year $J^{(k)} + 1$ is omitted, then $R^{(k)}$ is the classical claims reserve. If accident year $J^{(k)} + 1$ is included, then the setting allows for analysis of both so-called reserve risk and premium risk.

Let $X^{o,(k)}$ and X^o denote the column vectors with components $X_t^{o,(k)}$ and X_t^o, respectively. The value of the liability cash flow X^o will be defined as the value of a static replicating portfolio and the value of a position in the numéraire (one-year bond) needed to handle the residual liability cash flow from imperfect replication by the static replicating portfolio. In general, the choice of replicating portfolio depends on whether the liability cash flow depends on values or cash flows from traded financial assets. For the liability cash flows considered here, we assume that there are no clear dependencies on financial asset values. We choose the replicating portfolio as the one generating the deterministic cash flow $\mathbb{E}[X^o \mid \mathcal{F}_0]$. This cash flow is obtained by, for each $t \in \{1, \dots, T\}$, buying $\mathbb{E}[X_t^o \mid \mathcal{F}_0]$ units of the numéraire (one-year bonds that are rolled forward) at Time 0 and then selling them at time t. The market price of the static replicating portfolio is:

$$\sum_{t=1}^{T} \mathbb{E}[X_t^o \mid \mathcal{F}_0] = \mathbb{E}[R \mid \mathcal{F}_0].$$

Let $X := X^o - \mathbb{E}[X^o \mid \mathcal{F}_0]$ be the residual cash flow from imperfect replication. Given a risk-based regulatory framework, such as the one currently in place, the residual cash flow X will give rise to capital requirements. From the costs associated with meeting these capital requirements, a value V_0 at Time 0 will be assigned to X. The basis for determining V_0 is to consider a hypothetical transfer of the liability cash flow X^o and the replicating portfolio to a separate entity, a so-called reference undertaking, whose sole purpose is the handle the runoff of the liability cash flow. The procedure for determining V_0 is as follows.

Let V_t be the value of the residual cash flow from time $t + 1$ to time T as seen from Time 0. In particular, $V_T = 0$ since there is no cash flow beyond time T. At time t, the reference undertaking is required to hold the solvency capital $\rho_t = \rho_t(-X_{t+1} - V_{t+1})$, where ρ_t is the conditional monetary risk measure value-at-risk $\mathrm{VaR}_{t,p}$ or expected shortfall $\mathrm{ES}_{t,p}$, conditional on $\mathcal{F}_t$. Conditioning on $\mathcal{F}_t$ means that we only consider information from payment data. The reason for this choice is that we aim for a non-life valuation method that can be easily applied by anyone having access to yearly aggregated claims payment triangle data.

At time t, the reference undertaking has capital V_t that equals the value of the residual cash flow $(X_s)_{s=t+1}^T$ and asks a capital provider, with limited liability, to provide the capital $\rho_t - V_t$ ensuring solvency at time t. Upon providing the capital $\rho_t - V_t$ at time t, the capital provider receives any surplus $\max(\rho_t - X_{t+1} - V_{t+1}, 0)$ at time $t + 1$. However, capital is only provided if the expected return is good enough, if:

$$\mathbb{E}[\max(\rho_t - X_{t+1} - V_{t+1}, 0) \mid \mathcal{F}_t] \geq (1 + \eta_t)(\rho_t - V_t).$$

The value V_t is defined as the value for which capital is provided, but such that capital providers do not obtain better-than-required expected returns. That is, the inequality above is replaced by an equality. In general, η_t is $\mathcal{F}_t$-measurable. Here, in order to derive closed-form valuation formulas, we assume that η_t is $\mathcal{F}_0$-measurable. That is, we assume that future requirements on the expected return on capital from capital providers are known at present time.

From Propositions 1 and 4 and Remark 5 in Engsner et al. (2017), it follows that:

$$\begin{aligned} V_t &= W_t(X_{t+1} + V_{t+1}) \\ &= W_t \circ \cdots \circ W_{T-1}(X_{t+1} + \cdots + X_T), \end{aligned} \tag{2}$$

where $\circ$ denotes composition and:

$$W_s(Y) := \rho_s(-Y) - \frac{1}{1+\eta_s}\mathbb{E}\left[\max(\rho_s(-Y) - Y, 0) \mid \mathcal{F}_s\right]. \tag{3}$$

With $L^1(\mathcal{F}_t)$ denoting the space of $\mathcal{F}_t$-measurable random variables Y with $\mathbb{E}\left[|Y|\right] < \infty$, $W_s : L^1(\mathcal{F}_{s+1}) \to L^1(\mathcal{F}_s)$ is a mapping satisfying:

$$\text{if } \lambda \in L^1(\mathcal{F}_s) \text{ and } Y \in L^1(\mathcal{F}_{s+1}), \text{ then } W_s(\lambda + Y) = \lambda + W_s(Y),$$
$$\text{if } Y, \widetilde{Y} \in L^1(\mathcal{F}_{s+1}) \text{ and } Y \leq \widetilde{Y}, \text{ then } W_s(Y) \leq W_s(\widetilde{Y}),$$
$$W_s(0) = 0,$$

properties that are inherited from the risk measure ρ_s.

Definition 1. *The value of the non-life insurance liability cash flow X^o is:*

$$L_0 := \mathbb{E}\left[R \mid \mathcal{F}_0\right] + V_0, \tag{4}$$

where V_0 is given by Equations (2) and (3).

Notice that V_0 and L_0 are conceptually-consistent alternatives to the ill-defined Solvency II quantities risk margin (see (European Commission 2015, Article 37)) and technical provisions, respectively. Thus, $\mathbb{E}\left[R \mid \mathcal{F}_0\right]$ corresponds to the Solvency II best estimate (BE). We want to emphasize that the Solvency II risk margin is not properly defined mathematically. For further details, see, e.g., p. 316 in Möhr (2011).

Notice that the definition of the value L_0 of the liability cash flow X^o is intimately connected to the particular choice of static replicating portfolio considered here. The procedure for determining V_0 is independent of the choice of replicating portfolio, although the value of V_0 depends on the replicating portfolio via the residual cash flow X.

In general, no explicit expression can be derived for V_0. However, one important exception is when all cash flows are jointly Gaussian and the cost-of-capital rates η_t are $\mathcal{F}_0$-measurable.

Proposition 1. *Assume that the joint distribution of the random variables:*

$$\{I_{ij}^{(k)}; k \in \mathcal{K}, (i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}\}$$

is Gaussian and, for all t, η_t is $\mathcal{F}_0$-measurable. Then:

$$V_0 = \sum_{t=1}^{T} c(p, \eta_t)\left(\text{Var}(R \mid \mathcal{F}_{t-1}) - \text{Var}(R \mid \mathcal{F}_t)\right)^{1/2}, \tag{5}$$

where:

$$c(p, \eta_t) := r(p) - \frac{1}{1+\eta_t}\left(r(p)\Phi(r(p)) + \varphi(r(p))\right)$$

with:

$$r(p) = \begin{cases} \Phi^{-1}(1-p) & \text{if } \rho_t = \text{VaR}_{t,p}, \\ \frac{1}{p}\varphi(\Phi^{-1}(1-p)) & \text{if } \rho_t = \text{ES}_{t,p}, \end{cases}$$

where $\text{VaR}_{t,p}$ and $\text{ES}_{t,p}$ are defined in Remark 1.

The proof is given in the Appendix.

Notice that Equation (5) makes sense due to the Gaussian model assumption: $\mathrm{Var}(R \mid \mathcal{F}_t)$ is a constant and therefore measurable with respect to the trivial σ-field, which is a subset of $\mathcal{F}_0$.

Remark 1. *The choices* $\rho_t = \mathrm{VaR}_{t,0.005}$ *and* $\eta_t = 0.06$ *correspond to choices made in the Solvency II regulation, whereas* $\rho_t = \mathrm{ES}_{t,0.01}$ *correspond to the Swiss Solvency Test. For an* $\mathcal{F}_{t+1}$-*measurable (discounted) value* Z, *the conditional monetary risk measures* $\mathrm{VaR}_{t,u}(Z)$ *and* $\mathrm{ES}_{t,u}(Z)$ *are defined in terms of the conditional distribution function* $F_{t,-Z}(x) := \mathbb{P}(-Z \leq x \mid \mathcal{F}_t)$ *and quantile function* $F_{t,-Z}^{-1}(1-u) := \min\{m \in \mathbb{R} : F_{t,-Z}(m) \geq 1-u\}$ *as:*

$$\mathrm{VaR}_{t,u}(Z) := F_{t,-Z}^{-1}(1-u),$$

$$\mathrm{ES}_{t,u}(Z) := \frac{1}{u} \int_0^u \mathrm{VaR}_{t,v}(Z) dv.$$

Remark 2. *From the variance decomposition formula,*

$$\mathrm{Var}(R \mid \mathcal{F}_{t-1}) = \mathrm{Var}(\mathbb{E}\left[R \mid \mathcal{F}_t\right] \mid \mathcal{F}_{t-1}) + \mathbb{E}\left[\mathrm{Var}(R \mid \mathcal{F}_t) \mid \mathcal{F}_{t-1}\right].$$

With $\mathrm{CDR}_t := \mathbb{E}\left[R \mid \mathcal{F}_{t-1}\right] - \mathbb{E}\left[R \mid \mathcal{F}_t\right]$ *and since* $\mathrm{Var}(R \mid \mathcal{F}_t)$ *is independent of* $\mathcal{F}_{t-1}$ *under the Gaussian model assumption,*

$$\mathrm{Var}(\mathrm{CDR}_t \mid \mathcal{F}_{t-1}) = \mathrm{Var}(R \mid \mathcal{F}_{t-1}) - \mathrm{Var}(R \mid \mathcal{F}_t).$$

Hence, Equation (4) can be expressed as:

$$L_0 = \mathbb{E}\left[R \mid \mathcal{F}_0\right] + \sum_{t=1}^{T} c(p, \eta_t)\, \mathrm{Var}(\mathrm{CDR}_t \mid \mathcal{F}_{t-1})^{1/2}.$$

If we have a Gaussian model for the cash flow X^o, then its value L_0 is determined by Equation (5) and can be estimated based on suitable estimates of the mean vector and covariance matrix of X^o.

However, the typical situations for the valuation of an aggregate cash flow are models, or at least reserving methods, for the cash flows $X^{o,(k)}$, $k \in \mathcal{K}$. The dependence between the cash flows $X^{o,(k)}$ for different business lines is not well-understood, and independence is often assumed.

Proposition 2. *Suppose the assumptions of Proposition 1 hold and, for $k_1, k_2 \in \mathcal{K}$, that* $\mathrm{Cor}(R^{(k_1)}, R^{(k_2)} \mid \mathcal{F}_0) \geq 0$. *Then:*

$$V_0 \leq \frac{A^{1/2} T}{\left(\sum_{s=1}^{T} c(p, \eta_s)^{-2} \right)^{1/2}} =: V_0^+, \tag{6}$$

where:

$$A := \sum_{k_1, k_2 = 1}^{K} \mathrm{Cor}(R^{(k_1)}, R^{(k_2)} \mid \mathcal{F}_0)\, \mathrm{Var}(R^{(k_1)} \mid \mathcal{F}_0^{(k_1)})^{1/2}\, \mathrm{Var}(R^{(k_2)} \mid \mathcal{F}_0^{(k_2)})^{1/2}.$$

In particular, if $\eta_t = \eta_0$ *for all t, then* $V_0 \leq c(p, \eta_0)(AT)^{1/2}$.

The proof of Proposition 2 is given in the Appendix.

The natural application of Proposition 2 is conservative estimation of the value V_0 of the residual liability cash flow in the case of only partial information. Conservative estimates of the conditional

variances $\mathrm{Var}(R^{(k)} \mid \mathcal{F}_0^{(k)})$ may be given by the estimates of conditional mean-squared errors of predictions (MSEPs):

$$\mathbb{E}\left[(R^{(k)} - \widehat{R}^{(k)})^2 \mid \mathcal{F}_0^{(k)} \right] = \mathrm{Var}(R^{(k)} \mid \mathcal{F}_0^{(k)}) + \mathbb{E}\left[\left(\widehat{R}^{(k)} - \mathbb{E}\left[R^{(k)} \mid \mathcal{F}_0^{(k)} \right] \right)^2 \mid \mathcal{F}_0^{(k)} \right].$$

Estimation of MSEP is discussed in detail in later sections. The correlations $\mathrm{Cor}(R^{(k_1)}, R^{(k_2)} \mid \mathcal{F}_0)$ cannot be reliably estimated based on claims triangle data. Proposition 2 provides a flexible tool for assessing the effect of dependence between claims triangles in terms of exogenously-given correlations in line with standard formulas found in regulatory frameworks.

3. Estimation and Parameter Uncertainty

Up to this point, we have discussed liability valuation given general Gaussian cash flow models when all, unspecified, model parameters are fully known. In practice, we will use a specific Gaussian cash flow model to value a certain liability. For this particular model, we will need to estimate an unknown parameter vector $\boldsymbol{\theta}$. Based on the estimator $\widehat{\boldsymbol{\theta}}$, a natural estimator of the liability value according to Proposition 1 is the plug-in estimator $\widehat{L}_0 := L_0(\widehat{\boldsymbol{\theta}})$. However, the estimator $\widehat{L}_0$ is typically biased, since it in general is a non-linear function in terms of the components of $\widehat{\boldsymbol{\theta}}$, even though individual parameter estimators may be unbiased. Hence, there is in general a need for a bias correction of $\widehat{L}_0$.

In this section, we will present parameter estimation, estimation of the bias of $\widehat{L}_0$ and, finally, estimation of mean squared error of reserve prediction allowing for conservative estimation of L_0 under partial information.

3.1. Parameter Estimation

Here, we present the estimation of model parameters based on the information, $\mathcal{F}_0$, available at Time 0. The Gaussian cash flow models, which will be defined in Section 4, will all be such that the cash flows of different accident years are independent, and for a fixed accident year, the cash flows, incremental or cumulative, are given by dynamic linear regression (autoregressive) models. In common for these models is that it is possible to represent them as a sequence of conditional general linear models, where the sequence runs over development years. That is, the state of our conditional payment process in development year j for all relevant accident years i for the k-th line of business, denoted $\mathbf{Y}_j^{(k)}$, conditional on the information available up to $l \le j-1$ satisfying:

$$\mathbf{Y}_j^{(k)} = \mathbf{A}_{j-1}^{(k)} \boldsymbol{\xi}_j^{(k)} + \mathbf{e}_j^{(k)}.$$

By "relevant accident years", we mean that only payments from accident years i such that $\mathbf{Y}_j^{(k)}$ is $\mathcal{F}_0$-measurable appear in $\mathbf{Y}_j^{(k)}$. Here, $\mathbf{A}_{j-1}^{(k)}$ denotes a matrix containing relevant (functions of) $Y_{i,l}^{(k)}$'s with $l \le j-1$; $\boldsymbol{\xi}_j^{(k)}$ is the parameter vector of regression coefficients for the j-th conditional linear model; and $\mathbf{e}_j^{(k)}$ is multivariate Gaussian with zero mean and covariance matrix $(\sigma_j^{(k)})^2 \boldsymbol{\Sigma}_j^{(k)}$, where $\boldsymbol{\Sigma}_j^{(k)}$ is some known (deterministic) weight matrix. Moreover, the $\mathbf{e}_j^{(k)}$'s are independent across indices j and k. The independence w.r.t. k is due to all model parameters being estimated separately for each line of business k, hence neglecting any potential dependencies between lines of business. Further, notice that $\mathbf{A}_{j-1}^{(k)}$ is defined in terms of vectors $\mathbf{Y}_l^{(k)}, l \le j-1$, which in the linear model context here are treated as covariates.

From general theory on linear models, we know that the weighted least squares estimator of $\xi_j^{(k)}$ is given by:

$$\widehat{\xi}_j^{(k)} = ((\mathbf{A}_{j-1}^{(k)})'(\mathbf{\Sigma}_j^{(k)})^{-1}\mathbf{A}_{j-1}^{(k)})^{-1}(\mathbf{A}_{j-1}^{(k)})'(\mathbf{\Sigma}_j^{(k)})^{-1}\mathbf{Y}_j^{(k)}. \tag{7}$$

Define:

$$B_{j-1}^{(k)} := \left\{ I_{il}^{(k)} \in D_0^{(k)} : l \leq j-1 \right\} \tag{8}$$

and notice that $\widehat{\xi}_j^{(k)}$ is an unbiased estimator of $\xi_j^{(k)}$, (Fahrmeir et al. 2013, p. 181), in the sense

$$\mathbb{E}\left[\widehat{\xi}_j^{(k)} \mid B_{j-1}^{(k)}\right] = \mathbb{E}\left[\widehat{\xi}_j^{(k)} \mid \mathbf{A}_{j-1}^{(k)}\right] = \xi_j^{(k)}.$$

As for standard general linear models, Equation (7) does not provide an estimator for $(\sigma_j^{(k)})^2$. A natural estimator of $(\sigma_j^{(k)})^2$ is:

$$\widehat{(\sigma_j^{(k)})^2} = \frac{1}{n}(\mathbf{Y}_j^{(k)} - \mathbf{A}_{j-1}^{(k)}\widehat{\xi}_j^{(k)})'(\mathbf{\Sigma}_j^{(k)})^{-1}(\mathbf{Y}_j^{(k)} - \mathbf{A}_{j-1}^{(k)}\widehat{\xi}_j^{(k)}),$$

where n corresponds to the degrees of freedom and $\widehat{(\sigma_j^{(k)})^2}$ is independent of $\widehat{\xi}_j^{(k)}$. We will throughout use an unbiased estimator of $(\sigma_j^{(k)})^2$. For ease of exposition, we collect all parameters for the model for the k-th line of business in $\boldsymbol{\theta}_j^{(k)} := (\xi_j^{(k)}, (\sigma_j^{(k)})^2)$, and we consequently have that:

$$\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right] = \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)} \mid \mathbf{A}_{j-1}^{(k)}\right] = \boldsymbol{\theta}_j^{(k)}.$$

Regarding the estimation uncertainty, we know from standard theory on general linear models and weighted least squares estimation that:

$$\text{Cov}(\widehat{\xi}_j^{(k)} \mid \mathbf{A}_{j-1}^{(k)}) = (\sigma_j^{(k)})^2((\mathbf{A}_{j-1}^{(k)})'(\mathbf{\Sigma}_j^{(k)})^{-1}\mathbf{A}_{j-1}^{(k)})^{-1}; \tag{9}$$

see, e.g., (Fahrmeir et al. 2013, p. 181). Moreover, using that $\widehat{(\sigma_j^{(k)})^2}$ is a quadratic form of a multivariate Gaussian vector together with standard matrix algebra gives:

$$\text{Var}\left(\widehat{(\sigma_j^{(k)})^2} \mid \mathbf{A}_{j-1}^{(k)}\right) = \frac{2}{n}(\sigma_j^{(k)})^4. \tag{10}$$

The above together with the fact that $\widehat{\xi}_j^{(k)}$ and $\widehat{(\sigma_j^{(k)})^2}$ are independent fully specifies $\text{Cov}(\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}) = \text{Cov}(\widehat{\boldsymbol{\theta}}_j^{(k)} \mid \mathbf{A}_{j-1}^{(k)})$. Moreover, for $l < j$,

$$\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_l^{(k)}\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right] = \widehat{\boldsymbol{\theta}}_l^{(k)}\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right] = \widehat{\boldsymbol{\theta}}_l^{(k)}\boldsymbol{\theta}_j^{(k)},$$

which together with:

$$\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_l^{(k)} \mid B_{j-1}^{(k)}\right]\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right] = \widehat{\boldsymbol{\theta}}_l^{(k)}\boldsymbol{\theta}_j^{(k)},$$

imply that:

$$\text{Cov}(\widehat{\boldsymbol{\theta}}_l^{(k)}, \widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}) = 0. \tag{11}$$

This is a property that will be used below.

Remark 3. *Note that the variance estimator* $\widehat{(\sigma_j^{(k)})^2}$ *above is unbiased. In practice, making it unbiased corresponds to disregarding some observed data. In a claims triangle setup, data are typically scarce from the beginning, which makes it worth considering the tradeoff between bias and the lack of data.*

Remark 4. *Above, we have exploited the fact that the time series models defining the cash flows may be represented as a sequence of conditional general linear models and used conditional weighted least squares estimation. For more on the relation between weighted least squares estimation and maximum likelihood estimation for general linear models, see, e.g., (Fahrmeir et al. 2013, chp. 4.1.2 and chp. 5.1.2). For more on estimation procedures based on both conditional, as well as unconditional approaches for ARMA models, see, e.g., (Christensen 1991, chp. V.5) and (Brockwell and Davis 1987, chp. 8).*

3.2. Bias and MSEP

Given the $\mathcal{F}_0$-measurable parameter estimator $\widehat{\boldsymbol{\theta}}$, the plug-in estimator of $L_0 = L_0(\boldsymbol{\theta})$ is $\widehat{L}_0 := L_0(\widehat{\boldsymbol{\theta}})$. Notice that $\widehat{L}_0$ is a $\mathcal{F}_0$-measurable function of the $\mathcal{F}_0$-measurable estimator $\widehat{\boldsymbol{\theta}}$, but for bias correction, we are only interested in correcting for the estimation uncertainty, not the fact that we, e.g., have observed a particular amount of realized payments, which will form the basis for the prediction of future outstanding payments. Notice also that the bias of $\widehat{L}_0$,

$$\mathbb{E}\left[L_0(\widehat{\boldsymbol{\theta}}) - L_0(\boldsymbol{\theta}) \mid \mathcal{F}_0\right] = L_0(\widehat{\boldsymbol{\theta}}) - L_0(\boldsymbol{\theta}),$$

is zero if one uses a naive plug-in estimator and therefore useless; see Mack (1993a).

In order to obtain a relevant estimator of the bias of $\widehat{L}_0$, we propose a method that has similarities with the methods in Buchwalder et al. (2006) and Mack (1993a). We introduce a random vector $\widehat{\boldsymbol{\theta}}^*$ representing parameter uncertainty, which can be regarded as a pseudo estimator. We require that, conditional on past payments, $\widehat{\boldsymbol{\theta}}^*$ is independent of future payments. Formally,

$$\text{conditional on } \mathcal{F}_0, \ \widehat{\boldsymbol{\theta}}^* \text{ is independent of } \sigma\left(\cup_{k \in \mathcal{K}} \left\{D_{J^{(k)}}^{(k)} \setminus D_0^{(k)}\right\}\right). \tag{12}$$

Moreover, the conditional mean and covariance:

$$\boldsymbol{\mu}_{\mathcal{F}_0}(\boldsymbol{\theta}) := \mathbb{E}\left[\widehat{\boldsymbol{\theta}}^* \mid \mathcal{F}_0\right], \quad \boldsymbol{\Sigma}_{\mathcal{F}_0}(\boldsymbol{\theta}) := \text{Cov}\left(\widehat{\boldsymbol{\theta}}^* \mid \mathcal{F}_0\right) \tag{13}$$

are assumed to exist finitely and will be chosen to reflect estimation uncertainty in sequential estimation of $\boldsymbol{\theta}$. In the present paper, we will use $\widehat{\boldsymbol{\theta}}$ based on (conditional) weighted least squares; see Section 3.1 and further details in Section 4. Based on the calculations in Section 3.1 concerning the properties of the estimator $\widehat{\boldsymbol{\theta}}$, we require that $\widehat{\boldsymbol{\theta}}^*$ is chosen such that $\widehat{\boldsymbol{\theta}}_j^{(1),*}, \ldots, \widehat{\boldsymbol{\theta}}_j^{(K),*}$ are independent,

$$\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k),*} \mid \mathcal{F}_0\right] = \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right] \quad \left(= \boldsymbol{\theta}_j^{(k)}\right),$$

$$\text{Cov}(\widehat{\boldsymbol{\theta}}_j^{(k),*} \mid \mathcal{F}_0) = \text{Cov}(\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}),$$

$$\text{Cov}(\widehat{\boldsymbol{\theta}}_l^{(k),*}, \widehat{\boldsymbol{\theta}}_j^{(k),*} \mid \mathcal{F}_0) = \text{Cov}(\widehat{\boldsymbol{\theta}}_l^{(k)}, \widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}) \quad \left(= 0 \text{ for } l \neq j\right).$$

Notice that $\mathrm{Cov}(\widehat{\boldsymbol{\theta}}^* \mid \mathcal{F}_0)$ is block diagonal with blocks $\mathrm{Cov}(\widehat{\boldsymbol{\theta}}^{(k),*} \mid \mathcal{F}_0)$ that are again, due to Equation (11), block diagonal with blocks $\mathrm{Cov}(\widehat{\boldsymbol{\theta}}_j^{(k),*} \mid \mathcal{F}_0)$. Thus, we have fully specified $\boldsymbol{\mu}_{\mathcal{F}_0}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}_{\mathcal{F}_0}(\boldsymbol{\theta})$.

Remark 5. *Note that the definition of $\widehat{\boldsymbol{\theta}}^*$ shares many features with, e.g., Buchwalder et al. (2006), as well as England and Verrall (1999).*

Remark 6. *Note that trivially, it holds that for $l < j$:*

$$\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_l^{(k)}\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}, B_{l-1}^{(k)}\right] = \mathbb{E}\left[\widehat{\boldsymbol{\theta}}_l^{(k)}\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right],$$

that is, the sub-matrices defining $\boldsymbol{\Sigma}_{\mathcal{F}_0}(\boldsymbol{\theta})$ are not chosen w.r.t. arbitrary conditionings on $B_j^{(k)}$'s.

Using the pseudo estimator $\widehat{\boldsymbol{\theta}}^*$, we define the bias, conditional on $\mathcal{F}_0$, of $\widehat{L}_0$ with respect to plug-in parameter uncertainty as:

$$\mathrm{Bias}_{L_0}(\widehat{\boldsymbol{\theta}}^*) := \mathbb{E}\left[L_0(\widehat{\boldsymbol{\theta}}^*) - L_0(\boldsymbol{\theta}) \mid \mathcal{F}_0\right].$$

Now, let $h(\boldsymbol{\theta}) := \mathbb{E}\left[R \mid \mathcal{F}_0\right]$ and:

$$u_t(\boldsymbol{\theta}) := \left(\mathrm{Var}(R \mid \mathcal{F}_{t-1}) - \mathrm{Var}(R \mid \mathcal{F}_t)\right)^{1/2}, \tag{14}$$

which gives:

$$\mathrm{Bias}_{L_0}(\widehat{\boldsymbol{\theta}}^*) = \mathbb{E}\left[h(\widehat{\boldsymbol{\theta}}^*) - h(\boldsymbol{\theta}) + \sum_{t=1}^{T} c(p, \eta_t)(u_t(\widehat{\boldsymbol{\theta}}^*) - u_t(\boldsymbol{\theta})) \mid \mathcal{F}_0\right]$$

$$= \mathrm{Bias}_h(\widehat{\boldsymbol{\theta}}^*) + \sum_{t=1}^{T} c(p, \eta_t)\,\mathrm{Bias}_{u_t}(\widehat{\boldsymbol{\theta}}^*).$$

The bias of $\widehat{L}_0$ as an estimator of L_0 is due to the u_t's being nonlinear functions of the components of its vector-valued argument. The requirement $\mathrm{Bias}_h(\widehat{\boldsymbol{\theta}}^*) = 0$ is easily met in the dynamic linear-regression-type models we will consider. What remains is to quantify $\mathrm{Bias}_{u_t}(\widehat{\boldsymbol{\theta}}^*)$ such that it leads to accurate estimation $\widehat{\mathrm{Bias}}_{L_0}(\widehat{\boldsymbol{\theta}}^*)$ of the bias of $\widehat{L}_0$. The bias $\mathrm{Bias}_{u_t}(\widehat{\boldsymbol{\theta}}^*)$ is approximated by approximating u_t by its second degree Taylor expansion:

$$\mathrm{Bias}_{u_t}(\widehat{\boldsymbol{\theta}}^*) \approx \mathbb{E}\left[(\widehat{\boldsymbol{\theta}}^* - \boldsymbol{\theta})'\nabla u_t(\boldsymbol{\theta}) + \frac{1}{2}(\widehat{\boldsymbol{\theta}}^* - \boldsymbol{\theta})'H_t(\boldsymbol{\theta})(\widehat{\boldsymbol{\theta}}^* - \boldsymbol{\theta}) \mid \mathcal{F}_0\right]$$

$$=: \mathrm{Bias}_{u_t}^H(\widehat{\boldsymbol{\theta}}^*),$$

where $H_t(\boldsymbol{\theta})$ denotes the Hessian of $u_t(\boldsymbol{\theta})$. Some matrix algebra yields the following proposition:

Proposition 3. *If $\widehat{\boldsymbol{\theta}}^*$ satisfies Equation (13), then:*

$$\mathrm{Bias}_{u_t}^H(\widehat{\boldsymbol{\theta}}^*) = (\boldsymbol{\mu}_{\mathcal{F}_0}(\boldsymbol{\theta}) - \boldsymbol{\theta})'\nabla u_t(\boldsymbol{\theta}) + \frac{1}{2}(\boldsymbol{\mu}_{\mathcal{F}_0}(\boldsymbol{\theta}) - \boldsymbol{\theta})'H_t(\boldsymbol{\theta})(\boldsymbol{\mu}_{\mathcal{F}_0}(\boldsymbol{\theta}) - \boldsymbol{\theta}) + \frac{1}{2}\mathrm{tr}(H_t(\boldsymbol{\theta})\boldsymbol{\Sigma}_{\mathcal{F}_0}(\boldsymbol{\theta})).$$

Remark 7. *Notice that $\mathrm{Bias}_{u_t}^H(\widehat{\boldsymbol{\theta}}^*)$ in Proposition 3 is a $\mathcal{F}_0$-measurable function of $\boldsymbol{\theta}$. We suggest plug-in estimation of $\mathrm{Bias}_{u_t}^H(\widehat{\boldsymbol{\theta}}^*)$ with plug-in $\widehat{\boldsymbol{\theta}}$ in place of $\boldsymbol{\theta}$. The resulting quantity is then used to correct the estimate*

$\widehat{L}_0$ of the liability value L_0 from Proposition 1. Further, notice that if $\boldsymbol{\mu}_{\mathcal{F}_0}(\boldsymbol{\theta}) = \boldsymbol{\theta}$, then the estimator of $\text{Bias}^H_{u_t}(\widehat{\boldsymbol{\theta}}^*)$ simplifies to:

$$\widehat{\text{Bias}}^H_{u_t}(\widehat{\boldsymbol{\theta}}^*) = \frac{1}{2}\text{tr}(H_t(\widehat{\boldsymbol{\theta}})\boldsymbol{\Sigma}_{\mathcal{F}_0}(\widehat{\boldsymbol{\theta}})).$$

We now turn to the estimation of MSEP. The natural plug-in estimator of outstanding claims costs is given by $\widehat{R}^{(k)} := h(\widehat{\boldsymbol{\theta}})$. Following the definition in Mack (1993a), the mean squared error of prediction (MSEP) for the claims reserve for the k-th line of business is given by:

$$\text{MSEP}_{\mathcal{F}_0}(\widehat{R}^{(k)}) := \mathbb{E}\left[(R^{(k)} - \widehat{R}^{(k)})^2 \mid \mathcal{F}_0\right]$$

which may be re-written as:

$$\text{MSEP}_{\mathcal{F}_0}(\widehat{R}^{(k)}) = \text{Var}(R^{(k)} \mid \mathcal{F}_0) + \mathbb{E}\left[\left(\widehat{R}^{(k)} - \mathbb{E}\left[R^{(k)}|\mathcal{F}_0\right]\right)^2 |\mathcal{F}_0\right]$$

$$= \text{Var}(R^{(k)} \mid \mathcal{F}_0) + \mathbb{E}\left[\left(h^{(k)}(\widehat{\boldsymbol{\theta}}) - h^{(k)}(\boldsymbol{\theta})\right)^2 |\mathcal{F}_0\right].$$

As in Mack (1993a) and Buchwalder et al. (2006), the first term in the MSEP expression corresponds to the "process variance", which may be approximated using a plug-in estimator. The second term, corresponding to parameter uncertainty or "estimation error", will collapse to zero if a plug-in estimation is used. Considerable effort has been made on how to estimate the parameter uncertainty; see, e.g., Mack (1993a) and Buchwalder et al. (2006) and the references therein, which rely on various ways to condition on data. In the present paper, we will adopt the procedure described for the bias correction in Proposition 3 in order to ascertain consistency between the approximations. Based on the above argumentation, we introduce the following MSEP:

$$\text{MSEP}^*_{\mathcal{F}_0}(\widehat{R}^{(k)}) := \mathbb{E}\left[(R^{(k)} - h^{(k)}(\widehat{\boldsymbol{\theta}}^*))^2 \mid \mathcal{F}_0\right] \tag{15}$$

$$= \text{Var}(R^{(k)} \mid \mathcal{F}_0) + \mathbb{E}\left[\left(h^{(k)}(\widehat{\boldsymbol{\theta}}^*) - h^{(k)}(\boldsymbol{\theta})\right)^2 \mid \mathcal{F}_0\right], \tag{16}$$

which is close to (Buchwalder et al. 2006, Approach 3) and Mack (1993a). The equality between Equations (15) and (16) is assured by Equation (12).

Moreover, if we assume that the plug-in estimator is an unbiased estimator of the claims reserve, the following approximation to MSEP* is obtained by repeating the steps of Proposition 3:

$$\text{MSEP}^{*,\nabla}_{\mathcal{F}_0}(\widehat{R}^{(k)}) := \text{Var}(R^{(k)} \mid \mathcal{F}_0) + \nabla h^{(k)}(\boldsymbol{\theta})'\boldsymbol{\Sigma}_{\mathcal{F}_0}(\boldsymbol{\theta})\nabla h^{(k)}(\boldsymbol{\theta}). \tag{17}$$

In Lindholm et al. (2017), these relationships are explored in more detail with respect to the results of Buchwalder et al. (2006) and Mack (1993a). More specifically, it is shown that Equation (17) using plug-in estimation coincides with the standard MSEP calculation for the chain-ladder method given in Mack (1993a).

Remark 8. *Note that Proposition 2 may be approximated using MSEP* as an approximation of the variances needed in the specification of the quantity A, which also captures estimation error.*

4. Gaussian Cash Flow Models

In the present section, we will describe two different Gaussian cash flow models and their parameter and reserve estimators. Further, the effect of estimation uncertainty is analyzed, and in particular, we derive bias corrections for the cost-of-capital valuation using plug-in estimators and calculate the approximate mean squared errors of prediction for expected future costs. For more on

time-series-reserving methods, see, e.g., Buchwalder et al. (2006); Kremer (1984); Merz and Wüthrich (2008). One can also note that, e.g., Buchwalder et al. (2006) provide a time-series model with Gaussian errors, which does not yield a Gaussian process. In the present paper, we need the entire cash flow model to be Gaussian for the valuation formulas from Section 2 to be applicable, thus disqualifying the model from Buchwalder et al. (2006).

4.1. An Autoregressive Model for Incremental Amounts

Let $v_i^{(k)}$ denote the premium volume for accident year i (or a similar exposure measure). It is assumed that all $v_i^{(k)}$ are known constants. Let $\widetilde{I}_{ij}^{(k)} := I_{ij}^{(k)} / v_i^{(k)}$ denote the normalized incremental claims payments.

We assume the following stochastic model for incremental claims payments:

$$\widetilde{I}_{ij}^{(k)} = \alpha_j^{(k)} + \beta_j^{(k)} \widetilde{I}_{i,j-1}^{(k)} + \frac{\sigma_j^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{ij}^{(k)}, \tag{18}$$

where $\widetilde{I}_{i0} := 0$ and for every k, the $\epsilon_{ij}^{(k)}$, $(i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}$, are mutually independent and standard normally distributed.

The incremental Gaussian cash flow model defined by Equation (18) implies the following basic properties:

Proposition 4. *For the model in Equation* (18),

$$\mathbb{E}\left[R^{(k)} \mid \mathcal{F}_0^{(k)}\right] = \sum_{t=1}^{T} \mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0^{(k)}\right],$$

where:

$$\mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0^{(k)}\right] = \sum_{u=t+1}^{J^{(k)}+1} v_u^{(k)} \left(\widetilde{I}_{u,J^{(k)}-u+1}^{(k)} \prod_{l=J^{(k)}-u+2}^{J^{(k)}-u+t+1} \beta_l^{(k)} + \sum_{l=J^{(k)}-u+2}^{J^{(k)}-u+t+1} \alpha_l^{(k)} \prod_{m=l+1}^{J^{(k)}-u+t+1} \beta_m^{(k)}\right),$$

$$\mathrm{Var}(R^{(k)} \mid \mathcal{F}_0^{(k)}) = \sum_{i=2}^{J^{(k)}+1} v_i^{(k)} \sum_{l=J^{(k)}-i+2}^{J^{(k)}} (\sigma_l^{(k)})^2 \left(1 + \sum_{j=l+1}^{J^{(k)}} \prod_{m=l+1}^{j} \beta_m^{(k)}\right)^2$$

and:

$$\mathrm{Var}(R^{(k)} \mid \mathcal{F}_{t-1}^{(k)}) - \mathrm{Var}(R^{(k)} \mid \mathcal{F}_t^{(k)}) = \sum_{i=t+1}^{J^{(k)}+1} v_i^{(k)} \sigma_{J-i+1+t}^2 \left(1 + \sum_{j=J-i+2+t}^{J^{(k)}} \prod_{m=J^{(k)}-i+2+t}^{j} \beta_m^{(k)}\right)^2.$$

If $(\mathcal{F}_t^{(1)}), \ldots, (\mathcal{F}_t^{(K)})$ *are independent, then*

$$\mathbb{E}\left[X_t^o \mid \mathcal{F}_0\right] = \sum_{k=1}^{K} \mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0^{(k)}\right],$$

$$\mathrm{Var}(R \mid \mathcal{F}_{t-1}) - \mathrm{Var}(R \mid \mathcal{F}_t) = \sum_{k=1}^{K} \sum_{i=t+1}^{J^{(k)}+1} v_i^{(k)} \sigma_{J-i+1+t}^2 \left(1 + \sum_{j=J-i+2+t}^{J^{(k)}} \prod_{m=J^{(k)}-i+2+t}^{j} \beta_m^{(k)}\right)^2.$$

The proof is presented in the Appendix.

Note that Proposition 4 contains all information needed in order to be able to value a liability following the dynamics described by Equation (18) when all parameters are assumed to be known. In practice, we need to estimate the parameters α_j, β_j and σ_j in order to be able to use the valuation

formulas. Given parameter estimates, the valuation formulas are then evaluated based on plug-in estimates. Further, from Proposition 4, we see that the natural plug-in reserve estimator:

$$\widehat{R}^{(k)} := \widehat{\mathbb{E}}[R^{(k)} | \mathcal{F}_0^{(k)}]$$

$$= \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{u=t+1}^{J^{(k)}+1} v_u^{(k)} \left(\widetilde{I}_{u,J^{(k)}-u+1}^{(k)} \prod_{l=J^{(k)}-u+2}^{J^{(k)}-u+t+1} \widehat{\beta}_l^{(k)} + \sum_{l=J^{(k)}-u+2}^{J^{(k)}-u+t+1} \widehat{\alpha}_l^{(k)} \prod_{m=l+1}^{J^{(k)}-u+t+1} \widehat{\beta}_m^{(k)} \right)$$

is unbiased given that $\widehat{\alpha}_j^{(k)}$ and $\widehat{\beta}_j^{(k)}$ are unbiased and that $\widehat{\alpha}_l^{(k)}$ and $\widehat{\beta}_m^{(k)}$ are uncorrelated for $l \neq m$. Proposition 6 below affirms that this indeed is the case.

Moreover, since $\mathrm{Var}(R^{(k)} | \mathcal{F}_t^{(k)})$ is a non-linear function in terms of $\beta_j^{(k)}$, the liability valuation formula using plug-in estimates will not be unbiased. Below, we will present an approximate bias correction for the valuation formula based on the incremental Gaussian cash flow model from Equation (18).

Remark 9. *As commented upon above, the valuation formula is a non-linear function in terms of model parameters. In order for the natural reserve estimator to be unbiased, it is necessary that the $\widehat{\alpha}_j$ and $\widehat{\beta}_j$ are unbiased. Due to the valuation formula's non-linearity, it is, however, less relevant to require unbiased estimation of the σ_j^2, since the entire valuation formula most likely will need a bias correction regardless of whether the $\widehat{\sigma}_j^2$ are unbiased or not.*

We will now proceed with parameter estimation.

4.1.1. Parameter Estimators and Predictors

As described in Section 3.1 above, the incremental model defined by Equation (18) may be expressed in terms of a sequence of conditional general linear models defined as follows: Let:

$$\begin{bmatrix} \widetilde{I}_{i_0,j}^{(k)} \\ \vdots \\ \widetilde{I}_{J^{(k)}-j+1,j}^{(k)} \end{bmatrix} = \begin{bmatrix} 1 & \widetilde{I}_{i_0,j-1}^{(k)} \\ \vdots & \vdots \\ 1 & \widetilde{I}_{J^{(k)}-j+1,j-1}^{(k)} \end{bmatrix} \begin{bmatrix} \alpha_j \\ \beta_j \end{bmatrix} + \begin{bmatrix} e_{i_0,j}^{(k)} \\ \vdots \\ e_{J^{(k)}-j+1,j}^{(k)} \end{bmatrix}$$

compactly as $\widetilde{\mathbf{I}}_j^{(k)} = \mathbf{A}_{j-1}^{(k)} \boldsymbol{\xi}_j^{(k)} + \mathbf{e}_j^{(k)}$ where $\mathbf{e}_j^{(k)}$ has zero mean and covariance matrix $(\sigma_j^{(k)})^2 \boldsymbol{\Sigma}_j^{(k)}$ with $\boldsymbol{\Sigma}_j^{(k)} = \mathrm{diag}(1/v_{i_0}^{(k)}, \ldots, 1/v_J^{(k)})$, where all $v_j^{(k)}$ are known. Furthermore, recall that this is a model that is conditional on $\widetilde{\mathbf{I}}_{j-1}^{(k)}$, which is put in $\mathbf{A}_{j-1}^{(k)}$, here treated as a set of covariates.

By using the results from Section 3.1 based on the specification above, the parameter estimators for the incremental model may be written out explicitly as follows:

Proposition 5. *The (conditional) weighted least squares estimators for model Equation (18) are given by:*

$$\widehat{\beta}_j^{(k)} = \frac{1}{c^{(k),j}} \left(\sum_{i=i_0}^{J^{(k)}-j+1} \frac{v_i^{(k)}}{v^{(k),j}} \widetilde{I}_{ij}^{(k)} \widetilde{I}_{i,j-1}^{(k)} - \bar{I}_{j-1}^{(k),j} \bar{I}_j^{(k),j} \right),$$

$$\widehat{\alpha}_j^{(k)} = \bar{I}_j^{(k),j} - \widehat{\beta}_j^{(k)} \bar{I}_{j-1}^{(k),j},$$

and:

$$\widehat{(\sigma_j^{(k)})^2} = \frac{1}{J^{(k)} - i_0 - j} \sum_{i=i_0}^{J^{(k)}-j+1} v_i^{(k)} (\widetilde{I}_{ij}^{(k)} - \widehat{\alpha}_j^{(k)} - \widehat{\beta}_j^{(k)} \widetilde{I}_{i,j-1}^{(k)})^2,$$

given that $J^{(k)} - i_0 - j \geq 2$ and:

$$\widehat{(\sigma_1^{(k)})^2} = \frac{1}{J^{(k)} - i_0 + 1} \sum_{i=i_0}^{J^{(k)}} v_i^{(k)} (\tilde{I}_{i1}^{(k)} - \widehat{\alpha}_1^{(k)})^2,$$

where $v^{(k),j} = \sum_{i=i_0}^{J^{(k)}-j+1} v_i^{(k)}$ and:

$$c^{(k),j} = \sum_{i=i_0}^{J^{(k)}-j+1} \frac{v_i^{(k),j}}{v^{(k),j}} (\tilde{I}_{i,j-1}^{(k),j})^2 - (\bar{I}_{j-1}^{(k),j})^2,$$

$$\bar{I}_j^{(k),j} = \sum_{i=i_0}^{J^{(k)}-j+1} \frac{v_i^{(k)}}{v^{(k),j}} \tilde{I}_{i,j}^{(k)},$$

$$\bar{I}_{j-1}^{(k),j} = \sum_{i=i_0}^{J^{(k)}-j+1} \frac{v_i^{(k)}}{v^{(k),j}} \tilde{I}_{i,j-1}^{(k)}.$$

Further, as commented on in Section 3.1, all $\widehat{\theta}_j^{(k)}$'s are unbiased, but it is possible to say more:

Proposition 6. *The estimators $\widehat{\theta}_j^{(k)}$, $j \in \mathcal{J}^{(k)}$, are unbiased and uncorrelated.*

The proof is given in the Appendix and is based on the tower property conditioning on $B_l^{(k)} := \{I_{ij}^{(k)} : (i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}, i+j \leq J^{(k)}+1, j \leq l\}$.

In the beginning of Section 4.1, plug-in estimation of outstanding costs was discussed. That is, based on the observed data D_0:

$$\widehat{I}_{ij}^{(k)} := \widehat{\mathbb{E}}[I_{ij}^{(k)} \mid D_0]$$

$$= I_{i,J^{(k)}-i+1} \prod_{l=J^{(k)}-i+2}^{j} \widehat{\beta}_l^{(k)} + \sum_{l=J^{(k)}-i+2}^{j} \widehat{\alpha}_l^{(k)} \prod_{m=l+1}^{j} \widehat{\beta}_m^{(k)}. \tag{19}$$

Moreover, we set:

$$\widehat{X}_t^{o,(k)} := \begin{cases} \sum_{u=t+1}^{J^{(k)}+1} \widehat{I}_{u,J^{(k)}-u+t+1}^{(k)}, & t = 1, \ldots, J^{(k)}, \\ 0, & t = J^{(k+1)} + 1, \ldots, T. \end{cases}$$

and $\widehat{X}^o := \sum_{k=1}^{K} \widehat{X}^{o,(k)}$.

Proposition 7. $\mathbb{E}\left[\widehat{I}_{ij}^{(k)}\right] = \mathbb{E}\left[I_{ij}^{(k)}\right]$ *for $(i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}$ with $i + j > J^{(k)} + 1$. Moreover, $\mathbb{E}\left[\widehat{X}^o\right] = \mathbb{E}[X^o]$.*

The proof is given in the Appendix.

We may now move on to treat parameter uncertainty with respect to liability valuation and claims reserve (MSEP).

4.1.2. Bias of the Liability Valuation Using Plug-In Estimators

The procedure for approximating the plug-in bias of the valuation formula from Section 3 is fully defined up to the specific parametrization of the model being studied. In the Appendix, the details

are presented for the incremental model, and the results for bias correction are given in Proposition 8 below. If we introduce:

$$\{\nabla u_t(\boldsymbol{\theta})\}_{\theta_j} := \frac{\partial}{\partial \theta_j} u_t(\boldsymbol{\theta}), \quad \{H_t(\boldsymbol{\theta})\}_{\theta_j} := \frac{\partial^2}{\partial(\theta_j)^2} u_t(\boldsymbol{\theta}),$$

we may state the approximate bias correction as follows:

Proposition 8. *The plug-in estimation bias for the incremental Gaussian model defined by Equation* (18) *is approximated by:*

$$\text{Bias}_{L_0}^H(\widehat{\boldsymbol{\theta}}^*) = \sum_{t=1}^{T} \frac{c(p, \eta_t)}{2} \sum_{k=1}^{K} \sum_{j=t+1}^{J^{(k)}} (\sigma_j^{(k)})^2 \left(\frac{\{H_t(\boldsymbol{\theta})\}_{\beta_j^{(k)}}}{c^{(k),j}} \sum_{i=i_0}^{J^{(k)}-j+1} v_i^{(k)} + \{H_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} \frac{2(\sigma_j^{(k)})^2}{J^{(k)} - i_0 - j} \right),$$

where:

$$\{\nabla u_t(\boldsymbol{\theta})\}_{\beta_j^{(k)}} = 1_{\{t+2 \leq j \leq J^{(k)}\}} [u_t(\boldsymbol{\theta})]^{-1} \times \sum_{n=t+1}^{j-1} v_{J^{(k)}-n+1+t}^{(k)} (\sigma_n^{(k)})^2 \left(1 + \sum_{l=n+1}^{J^{(k)}} \prod_{m=n+1}^{l} \beta_m^{(k)} \right) \sum_{l=j}^{J^{(k)}} \prod_{\substack{m=n+1 \\ m \neq j}}^{l} \beta_m^{(k)}$$

$$\{\nabla u_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} = \frac{1}{2} 1_{\{t+1 \leq j \leq J^{(k)}\}} [u_t(\boldsymbol{\theta})]^{-1} v_{J^{(k)}-j+1+t}^{(k)} \left(1 + \sum_{l=j+1}^{J^{(k)}} \prod_{m=j+1}^{l} \beta_m^{(k)} \right)^2,$$

and:

$$\{H_t(\boldsymbol{\theta})\}_{\beta_j} = 1_{\{t+2 \leq j \leq J^{(k)}\}} [u_t(\boldsymbol{\theta})]^{-1} \sum_{n=t+1}^{j-1} v_{J^{(k)}-n+1+t}^{(k)} (\sigma_n^{(k)})^2 \left(\sum_{l=j}^{J^{(k)}} \prod_{\substack{m=n+1 \\ m \neq j}}^{l} \beta_m^{(k)} \right)^2 - [u_t(\boldsymbol{\theta})]^{-1} [\{\nabla u_t(\boldsymbol{\theta})\}_{\beta_j}]^2$$

$$\{H_t(\boldsymbol{\theta})\}_{\sigma_j^2} = -[u_t(\boldsymbol{\theta})]^{-1} [\{\nabla u_t(\boldsymbol{\theta})\}_{\sigma_j^2}]^2.$$

As discussed above, the plug-in estimation bias from Proposition 8 is again a function in terms of the unknown model parameters $\boldsymbol{\theta}$ and is estimated using $\widehat{\boldsymbol{\theta}}$.

The proof of Proposition 8 is given in the Appendix.

4.1.3. Mean Squared Error of Prediction

As pointed out in Section 3.2, it is shown in Lindholm et al. (2017) that the proposed MSEP calculation given by Equation (17) for the classical claims reserve, that is when not considering premium risk, will coincide with that of Mack (1993a) for the chain-ladder method. Due to the dependence between $\widehat{\alpha}_j$ and $\widehat{\beta}_j$, these calculations turn out to be very complicated. We will now provide explicit formulas for the classical claims reserve for a single line of business, but leaving out the explicit dependence on a particular line of business "k". This is in accordance with the situation in Mack (1993a). Further, due to the complexity of the calculations, we only provide explicit formulas for the MSEP (= MSEP$^{*,\nabla}$) when all $\alpha_j \equiv 0$:

Proposition 9. *The MSEP for the classical claims reserve for the autoregressive incremental model given by Equation* (18) *with $\alpha_j \equiv 0, j = 1, \ldots, J$, may be approximated according to:*

$$\text{MSEP}_{\mathcal{F}_0}(\widehat{R}) = \sum_{i=2}^{J} \text{MSEP}_{\mathcal{F}_0}(\widehat{R}_i) + 2 \sum_{2 \leq i < i' \leq J} I_{i,J-i+1} I_{i',J-i'+1} G_{i,i'},$$

where:

$$\text{MSEP}_{\mathcal{F}_0}(\widehat{R}_i) = v_i \sum_{l=J-i+2}^{J} \widehat{\sigma}_l^2 \left(1 + \sum_{j=l+1}^{J} \prod_{m=l+1}^{j} \widehat{\beta}_m\right)^2 + I_{i,J-i+1}^2 \sum_{j=J-i+2}^{J} \sum_{l=J-i+2}^{j} (\widehat{s}_l^{J-i+2,j})^2$$

$$+ 2I_{i,J-i+1}^2 \sum_{j<j'} \sum_{l=J-i+2}^{j} \sum_{l'=J-i+2}^{j'} 1_{\{l=l'\}} \widehat{s}_l^{J-i+2,j} \widehat{s}_{l'}^{J-i+2,j'}$$

and:

$$G_{i,i'} = \sum_{j=J-i+2}^{J} \sum_{j'=J-i'+2}^{J} \sum_{l=J-i+2}^{j} \sum_{l'=J-i'+2}^{j'} 1_{\{l=l'\}} \widehat{s}_l^{J-i+2,j} \widehat{s}_{l'}^{J-i+2,j'},$$

where:

$$\widehat{s}_l^{J-i+2,j} = \left(\prod_{m=J-i+2}^{j} \widehat{\beta}_m\right) \sqrt{\widehat{\text{Var}}(\widehat{\beta}_l | B_{l-1})} / \widehat{\beta}_l.$$

The proof follows the steps of Mack (1993a) closely and is given in the Appendix.

Remark 10. *The proof of Proposition 9 in the case when $\alpha_j \neq 0$ turns out to be very tedious following Mack (1993a). It is however possible to calculate the approximate MSEP following Equation (17), but not very informative. For the case $\alpha_j \equiv 0$, Lindholm et al. (2017) shows that the results coincide with those of Mack (1993a).*

Remark 11. *Above, we provide formulas for the classical claims reserve. If we want to include premium risk, as well, this is straightforward to do using Equation (17) when including accident year $J^{(k)} + 1$. In the Appendix, $\Sigma_{\mathcal{F}_0}(\boldsymbol{\theta})$ is given in the derivation of Proposition 8, and $\nabla h^{(k)}(\boldsymbol{\theta})$ is easily calculated, since it is multilinear in terms of the components of $\boldsymbol{\theta}$. We do not write out these calculations explicitly, due to these becoming even more involved than the expressions above, and do not provide any further insights.*

4.2. An Autoregressive Model for Cumulative Amounts

We assume the following model for cumulative claims payments $C_{ij}^{(k)} := I_{i1}^{(k)} + \cdots + I_{ij}^{(k)}$:

$$\widetilde{C}_{ij}^{(k)} := C_{ij}^{(k)} / v_i^{(k)},$$

$$\widetilde{C}_{i1} := \alpha_1^{(k)} + \frac{\sigma_1^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{i1}^{(k)},$$

$$\widetilde{C}_{ij}^{(k)} := \gamma_{j-1}^{(k)} \widetilde{C}_{i,j-1}^{(k)} + \frac{\sigma_j^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{ij}^{(k)}, \quad j \geq 2, \tag{20}$$

where, for every k, the $\epsilon_{ij}^{(k)}$, $(i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}$, are mutually independent and standard normally distributed. This model is consistent with the chain-ladder prediction method, with a structure for the conditional variances that ensures that all incremental and cumulative payments are jointly Gaussian.

Remark 12. *Note that the actual reserve estimator based on model Equation (20) will not coincide with that of the standard chain-;adder method; see further Section 4.2.1. Further, note that for the standard chain-ladder method, there is no need to initiate the data-generating process, i.e., $\widetilde{I}_{i1}$ is always observed, but in the present paper, this step is needed in order to capture premium risk.*

Remark 13. *The techniques used for the derivation of the basic properties of the cumulative model will be identical to those used in Section 4.1 for the incremental model. Further, for the cumulative model, the actual calculations will be simpler than for the incremental model, and due to this, most calculations in this section will be kept very brief.*

Proposition 10. *For the model in Equation (20),*

$$\mathbb{E}\left[R^{(k)} \mid \mathcal{F}_0^{(k)}\right] = \sum_{t=1}^{T} \mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0^{(k)}\right],$$

where:

$$\mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0^{(k)}\right] = \sum_{i=1+t}^{J^{(k)}} C_{i,J^{(k)}-i+1}^{(k)} \prod_{m=J^{(k)}-i+1}^{J^{(k)}-i+t-1} \gamma_m^{(k)}(\gamma_{J^{(k)}-i+t}^{(k)} - 1)$$

$$+ a_1^{(k)} I\{t=1\} + a_1^{(k)} I\{t \geq 2\} \prod_{m=1}^{t-2} \gamma_m^{(k)}(\gamma_{t-1}^{(k)} - 1),$$

$$\mathrm{Var}(R^{(k)} \mid \mathcal{F}_0^{(k)}) = \sum_{i=2}^{J^{(k)}+1} v_i^{(k)} \sum_{j=J^{(k)}-i+2}^{J^{(k)}} (\sigma_j^{(k)})^2 \prod_{m=j}^{J^{(k)}-1} (\gamma_m^{(k)})^2,$$

$$\mathrm{Var}(R^{(k)} \mid \mathcal{F}_{t-1}^{(k)}) - \mathrm{Var}(R^{(k)} \mid \mathcal{F}_t^{(k)}) = \sum_{i=1+t}^{J^{(k)}+1} v_i^{(k)} (\sigma_{J^{(k)}-i+1+t}^{(k)})^2 \prod_{m=J^{(k)}-i+1+t}^{J^{(k)}-1} (\gamma_m^{(k)})^2.$$

If $(\mathcal{F}_t^{(1)}), \ldots, (\mathcal{F}_t^{(K)})$ *are independent, then* $\mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0^{(k)}\right] = \mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0\right]$ *and:*

$$\mathrm{Var}(R \mid \mathcal{F}_{t-1}) - \mathrm{Var}(R \mid \mathcal{F}_t) = \sum_{k=1}^{K} \sum_{i=1+t}^{J^{(k)}+1} v_i^{(k)} (\sigma_{J^{(k)}-i+1+t}^{(k)})^2 \prod_{m=J^{(k)}-i+1+t}^{J^{(k)}-1} (\gamma_m^{(k)})^2.$$

The proof is given in the Appendix.

One can here note that for the classical claims reserve, i.e., when omitting year $J^{(k)} + 1$, it holds that:

$$\mathbb{E}\left[R^{(k)} \mid \mathcal{F}_0^{(k)}\right] = \sum_{i=2}^{J^{(k)}} v_i^{(k)} \widetilde{C}_{i,J-i+1}^{(k)} \left(\prod_{j=i}^{J^{(k)}-1} \gamma_j^{(k)} - 1\right),$$

which suggests the following natural reserve estimator:

$$\widehat{R}^{(k)} := \widehat{\mathbb{E}}[R^{(k)} \mid \mathcal{F}_0^{(k)}]$$

$$= \sum_{i=2}^{J^{(k)}} v_i^{(k)} \widetilde{C}_{i,J-i+1}^{(k)} \left(\prod_{j=i}^{J^{(k)}-1} \widehat{\gamma}_j^{(k)} - 1\right).$$

4.2.1. Parameter Estimators and Predictors

In order to be able to use the valuation formulas, we need to obtain parameter estimators. Following Section 4.1.1, making the obvious changes to arrive at a sequence of (conditional) general linear models describing model Equation (20), we get the following proposition:

Proposition 11. *The (conditional) weighted least squares estimators of the parameters* $\alpha_1^{(k)}, \gamma_j^{(k)}, j = 1, \ldots, J^{(k)} - 1$, *from model Equation* (20), *are given by:*

$$\widehat{\alpha}_1^{(k)} = \frac{\sum_{i=i_0}^{J} v_i^{(k)} \widetilde{C}_{i1}^{(k)}}{\sum_{i=i_0}^{J} v_i^{(k)}},$$

$$\widehat{\gamma}_{j-1}^{(k)} = \frac{\sum_{i=i_0}^{J-j+1} v_i^{(k)} \widetilde{C}_{ij}^{(k)} \widetilde{C}_{i,j-1}^{(k)}}{\sum_{i=i_0}^{J-j+1} v_i^{(k)} (\widetilde{C}_{i,j-1}^{(k)})^2}, \quad j = 2, \ldots, J^{(k)},$$

and:

$$\widehat{(\sigma_1^{(k)})^2} = \frac{1}{J^{(k)} - i_0} \sum_{i=i_0}^{J^{(k)}} v_i^{(k)} (\widetilde{C}_{i1}^{(k)} - \widehat{\alpha}_1^{(k)})^2,$$

$$\widehat{(\sigma_j^{(k)})^2} = \frac{1}{J^{(k)} - j - i_0} \sum_{i=i_0}^{J^{(k)}-j+1} v_i^{(k)} (\widetilde{C}_{ij}^{(k)} - \widehat{\gamma}_{j-1}^{(k)} \widetilde{C}_{i,j-1}^{(k)})^2, \quad j = 2, \ldots, J^{(k)},$$

given that $J^{(k)} - i_0 - j \geq 2$. *Further,* $\widehat{\alpha}_1^{(k)}$ *and all* $\widehat{\gamma}_j^{(k)}$ *and* $\widehat{(\sigma_j^{(k)})^2}$ *are uncorrelated, and* $\widehat{\alpha}_1^{(k)}$ *and all* $\widehat{\gamma}_j^{(k)}$ *and* $\widehat{(\sigma_j^{(k)})^2}$ *are unbiased.*

The proof of Proposition 11 follows the corresponding proofs for the incremental model verbatim.

4.2.2. Bias of the Liability Valuation Using Plug-In Estimators

By repeating the same steps as for the incremental model, it is possible to arrive at the following approximation of the plug-in estimation bias from Proposition 8:

Proposition 12. *The plug-in estimation bias for the cumulative Gaussian model defined by Equation* (20) *is approximated by:*

$$\text{Bias}_{L_0}^H(\widehat{\boldsymbol{\theta}}^*) = \sum_{t=1}^{T} \frac{c(p, \eta_t)}{2} \sum_{k=1}^{K} \sum_{j=t+1}^{J^{(k)}} (\sigma_j^{(k)})^2 \left(\frac{\{H_t(\boldsymbol{\theta})\}_{\gamma_{j-1}^{(k)}}}{\sum_{i=i_0}^{J^{(k)}-j+1} v_i^{(k)} (\widetilde{C}_{i,j-1}^{(k)})^2} + \{H_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} \frac{2(\sigma_j^{(k)})^2}{J^{(k)} - i_0 - j} \right),$$

where:

$$\{\nabla u_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}} = 1_{\{t+1 \leq j\}} [u_t(\boldsymbol{\theta})]^{-1} \sum_{u=t+1}^{j} v_{J^{(k)}-u+t+1}^{(k)} (\sigma_u^{(k)})^2 \gamma_j^{(k)} \prod_{\substack{m=u \\ m \neq j}}^{J^{(k)}-1} (\gamma_m^{(k)})^2,$$

$$\{\nabla u_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} = 1_{\{t+1 \leq j\}} \frac{1}{2} [u_t(\boldsymbol{\theta}]^{-1} v_{J^{(k)}-j+t+1}^{(k)} \prod_{m=j}^{J^{(k)}-1} (\gamma_m^{(k)})^2,$$

and:

$$\{H_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}} = \{\nabla u_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}} \left(\frac{1}{\gamma_j^{(k)}} - \frac{\{\nabla u_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}}}{u_t(\boldsymbol{\theta})} \right),$$

$$\{H_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} = -[u_t(\boldsymbol{\theta})]^{-1} [\{\nabla u_t(\boldsymbol{\theta})\}_{\sigma_j^2}]^2.$$

As before, the plug-in estimation bias from Proposition 12 is again a function in terms of the unknown model parameters $\boldsymbol{\theta}$ and is estimated using $\widehat{\boldsymbol{\theta}}$.

4.2.3. Mean Squared Error of Prediction

In relation to the approach in Mack (1993a) for approximating the classical claims reserve MSEP, it is perhaps not surprising that the cumulative model from Equation (20) will prove easier to treat than the incremental model. Again, we will focus on the situation in Mack (1993a), that is a single line of business (without explicit reference to line of business "*k*") and no premium risk included (accident year $J + 1$ is omitted). The following approximation of MSEP (=MSEP$^{*,\nabla}$) for the cumulative payment model is obtained by directly applying the methods of Mack (1993a):

Proposition 13. *The MSEP for the autoregressive cumulative model given by Equation* (20) *may be approximated according to:*

$$\text{MSEP}_{\mathcal{F}_0}(\widehat{R}) = \sum_{i=2}^{J} \text{MSEP}_{\mathcal{F}_0}(\widehat{R}_i) + 2 \sum_{2 \leq i < i' \leq J} C_{i,J-i+1} C_{i',J-i'+1} G_{i,i'},$$

where:

$$\text{MSEP}_{\mathcal{F}_0}(\widehat{R}_i) = v_i \sum_{j=J-i+2}^{J} \widehat{\sigma}_j^2 \prod_{l=j}^{J-1} \widehat{\gamma}_l^2 + C_{i,J-i+1}^2 \prod_{j=J-i+2}^{J} \widehat{\gamma}_j \sum_{l=J-i+1}^{J-1} \frac{\widehat{\sigma}_{l+1}^2 / \widehat{\gamma}_l^2}{\sum_{m=i_0}^{J-l} v_m \widetilde{C}_{m,l}^2},$$

$$G_{i,i'} = \sum_{l=J-i+1}^{J-1} \widehat{\gamma}_{J-i'+1} \cdots \widehat{\gamma}_{J-i} \widehat{\gamma}_{J-i+1}^2 \cdots \widehat{\gamma}_{l-1}^2 \left(\frac{\widehat{\sigma}_{l+1}^2}{\sum_{m=i_0}^{J-l} v_m \widetilde{C}_{m,l}^2} \right) \widehat{\gamma}_{l+1} \cdots \widehat{\gamma}_{J-1}.$$

The proof of Proposition 13 is identical with that of Theorem 3 and the Corollary to Theorem 3 in Mack (1993a), given the parameter estimators for model Equation (20). This is a consequence of the underlying model structure being consistent with that of the standard chain-ladder method. The proof is hence omitted, but the steps follow those of the proof of Proposition 9.

Again, as described in Remark 11, it is straightforward to calculate the MSEP for the expected future costs, including premium risk, but these calculations are omitted.

5. Some Comments on Model Validation: In-Sample and Out-Of-Sample Performance

The focus of the present paper is to describe how to carry out a fully-data-driven cost-of-capital valuation based on claims triangles. Above, we have introduced two Gaussian models for which explicit calculations are possible including estimation and treating parameter uncertainty. Due to that, we have introduced two different models with different (maximal) numbers of parameters; there is a need to assess both the in-sample and the out-of-sample performance of the models.

As commented upon above, both models may be treated as a series of Gaussian linear models. Thus, it is recommended to carry out standard statistical residual analysis and tests of model assumptions for linear models. A detailed description on how this analysis may be implemented for the chain-ladder method is described in Mack (1993b); see also (Wüthrich and Merz 2008, chp. 11).

Further, regarding model selection, including significance tests for parameters, one can note that the two model classes are nested in the classical sense, which means that it is possible to carry out standard likelihood ratio tests (LRT), but other metrics such as AIC/BIC may be considered. Due to that, the models studied in the present paper are Gaussian, and there are no difficulties in determining (conditional) likelihoods. Given this, it is possible to analyze, e.g., LRT and AIC/BIC.

Moreover, since the models' actual purpose is to be used for out-of-sample prediction, it is reasonable to try to assess the out-of-sample performance. A straight forward out-of-sample performance is to exclude one or two observed diagonals and analyze the corresponding out-of-sample prediction performance. In general, it often is not recommended to exclude more than one or two years of data due to the lack of representative yearly data.

6. Numerical Example

In this section, we illustrate the incremental model in Equation (18) and the cumulative model in Equation (20) and demonstrate the implementation of the results in this paper.

Recall from Section 2 that the models and methods presented assume that historical data and future cash flows are adjusted for the time value of money. The aim of the numerical example in this section is simply to clarify and demonstrate the use of the framework presented earlier. In order to keep the numerical example brief and focus on methodological aspects, we ignore discounting and preprocessing of historical data here.

The data we use are the run-off triangle from Table 1 in Mack (1993a). In order to be able to use the unbiased estimators of the σ^2, we remove the last two columns of this triangle. We could use maximum likelihood estimation or some form of extrapolation of the σ^2, but the purpose here is not to compare methods for estimating the tail variances; we thus simply remove the last two columns. The data are fairly similar between accident years, and for illustrative purposes, it should suffice to set the weights equal to one for all accident years.

If we fit the incremental and cumulative models to the data, it is recommended to analyze standardized residuals and linearity for each development year and model. This is illustrated in Figure 1, where the standardized residuals for the second development year are shown for both models. There, we see that the residuals of the cumulative model have a systematic pattern implying that the cumulative linear structure defined by the model in Equation (20) does not capture the dynamics of the data well. This is an artifact of the lack of an intercept. However, the small sample size hinders us from making any definitive claims.

Before we proceed with calculating the cost-of-capital margin, etc., we make a brief assessment of the model performance given ideal circumstances. This is done by fitting the incremental and cumulative models to the data, considering these fits as the true models. Given the fitted parameters, we then simulate complete rectangles of data on which we re-fit the two models and plot $R - \widehat{R}$. The result of this can be seen in Figure 2. Both models seem to be performing similarly regardless of which model is simulated.
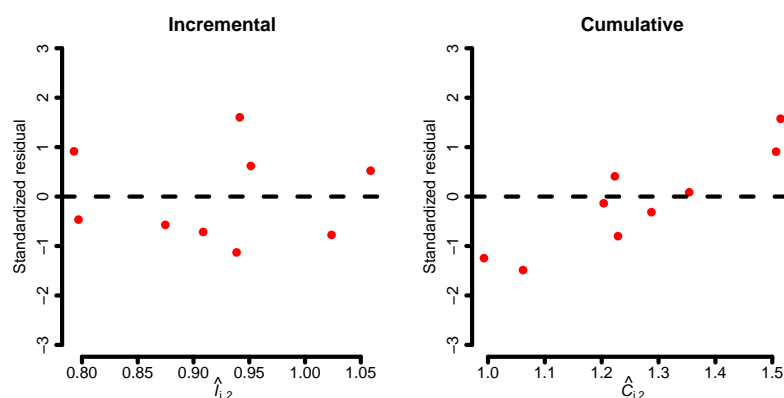

Figure 1. Standardized residuals for the second development year for the incremental model (left) and for the cumulative model (right). Fitted values are in millions.

With the aim of illustrating the bias correction of the cost-of-capital margin, we next simulate new triangles from the "true" models and fit the correctly specified model to each. That is, the incremental model is fitted to data simulated from the incremental model, and the cumulative model is fitted to data simulated from the cumulative model. Then, we estimate the cost-of-capital margin V_0 (see Equation (5)) and its bias correction for each simulated triangle using the cost-of-capital rate $\eta_0 = 6\%$, which is in alignment with the current Solvency II directive. This can then be compared to the true cost-of-capital margin, which we assume known since we consider the models fitted to the original

data to be the true models. The results can be seen in Figure 3. The bias correction does seem to work quite well. It should be noted that the expectation part of the liability value $L_0 := \mathbb{E}\left[R \mid \mathcal{F}_0\right] + V_0$, i.e., the best estimate (BE), in the current example is much larger than the cost-of-capital margin. Thus, the bias correction will, in this situation, have a small impact on the estimated liability value; see Table 1.
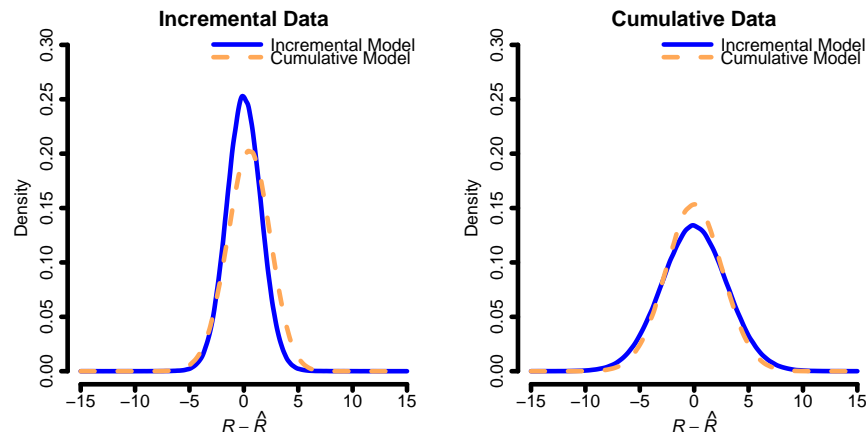

Figure 2. Distribution of the prediction error for the incremental and cumulative models when fitted to incremental data (left) and cumulative data (right). Values are in millions.
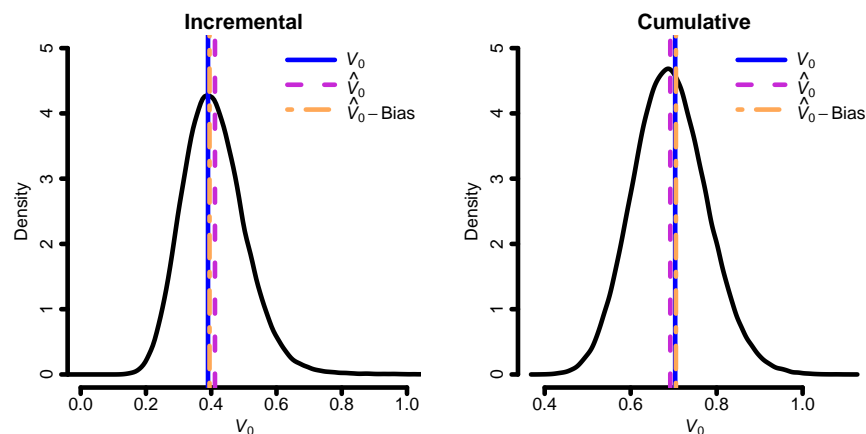

Figure 3. Distribution of the cost-of-capital margin for the incremental model (left) and the cumulative model (right). Values are in millions.

Proposition 2 provides an upper bound on V_0, denoted V_0^+; see Equation (6). Moreover, as commented on in Section 2, the cost-of-capital margin is related to the Solvency II Risk-Margin (RM). In (EIOPA 2015, Guideline 61), the following approximation of RM is suggested:

$$\mathrm{RM} = \mathrm{CoC} \sum_{t=0}^{T} \frac{\mathrm{BE}(t)}{\mathrm{BE}(0)} \mathrm{SCR}(0),$$

where $\mathrm{BE}(t) = \sum_{s=t+1}^{T} \mathbb{E}\left[X_s^o \mid \mathcal{F}_0\right]$ in our setting. The Solvency II regulation provides one-year non-life insurance risk standard deviations in the range of 10–20%, depending on the line of business, which form the basis for the SCR calculations. For the dataset from Mack (1993a), it is not clear from which line of business the data stem. Thus, we use:

$$\mathrm{SCR}(0) = 3\sqrt{\mathrm{Var}(\mathrm{CDR}_1 \mid \mathcal{F}_0)}.$$

Note however that the V_0 calculations are based on the standard deviations of the one-year claims development results, which are roughly 4% and 8% of the BE for the incremental and cumulative models, respectively. These values should be compared to the ones above, which are in the range of 10–20%. Consequently, we expect V_0 and RM to be considerably lower than the corresponding standard formula results. In Table 1, we have summarized V_0, V_0^+ and RM. Furthermore, the table also includes L_0, BE, $\mathrm{Var}(\mathrm{BE} \mid \mathcal{F}_0)^{1/2}$ and RMSEP. Notice that when including premium risk, we cannot calculate the Solvency II RM without making further assumptions on premium volumes. For the part of Table 1 with premium risk, we have set premium volumes consistent with weights equal to one. Furthermore, note that the cost-of-capital margin cannot be calculated for the chain-ladder model, nor can we calculate any quantity involving premium risk since there is no initiation in the first column. The RMSEP for the chain-ladder model is calculated according to Mack (1993a), and $\mathrm{Var}(\mathrm{CDR}_1 \mid \mathcal{F}_0)^{1/2}$ is calculated according to Merz and Wüthrich (2008). Hence, it is possible to calculate RM for the chain-ladder model without premium risk. Turning to the results, from Table 1, it is seen that V_0 and V_0^+ are of comparable order and that the corresponding RM values are lower than V_0^+ and of a size comparable to V_0. The cumulative model has almost double the cost-of-capital margin compared to the incremental model. This can most likely be explained by the fact that the intercepts in the incremental model catch some of the variation that is put into the variance parameters of the cumulative model. This is also seen in $\mathrm{Var}(\mathrm{BE} \mid \mathcal{F}_0)^{1/2}$ and RMSEP.

Table 1. Summary table. Values are expressed in millions. RM, risk-margin; BE, best estimate.

Model	V_0	V_0^+	RM	L_0	BE	$\mathrm{Var}(\mathrm{BE} \mid \mathcal{F}_0)^{1/2}$	RMSEP
			With premium risk				
I	0.39	0.44	0.31	18.47	18.08	1.09	1.58
C	0.70	0.87	0.83	19.94	19.24	2.12	2.67
			Without premium risk				
I	0.31	0.38	0.25	13.69	13.38	0.93	1.33
C	0.51	0.67	0.54	15.03	14.52	1.64	2.06
CL	–	–	0.71	–	14.77	1.18	1.51

To conclude, we have analyzed the methods presented in the current paper based on data from Mack (1993a), which illustrate that the current valuation framework will result in comparable results with the formalism of the Solvency II directive and that the introduced models seem appropriate.

Author Contributions: The authors have contributed equally to this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs

Appendix A.1. Proof of Proposition 1

Proof. The statement follows from minor modifications of the proof of Proposition 6 in Engsner et al. (2017). Here, different values $c(p, \eta_t)$ appear instead of a common value $c = c(p, \eta_0)$. Moreover, since the cash flow of the replicating portfolio is $\mathcal{F}_0$-measurable, $\mathrm{Var}(\sum_{s=1}^T X_s \mid \mathcal{F}_t) = \mathrm{Var}(\sum_{s=1}^T X_s^o \mid \mathcal{F}_t) = \mathrm{Var}(R \mid \mathcal{F}_t)$. □

Appendix A.2. Proof of Proposition 2

Proof. The relation $\mathbb{E}\left[\mathrm{Var}(Y \mid \mathcal{G}) \mid \mathcal{F}\right] \geq \mathbb{E}\left[\mathrm{Var}(Y \mid \mathcal{H}) \mid \mathcal{F}\right]$ holds for σ-fields $\mathcal{F} \subset \mathcal{G} \subset \mathcal{H}$ and a random variable Y with finite variance. For Gaussian models, the conditional variance expressions correspond to variances of conditional distributions of a multivariate Gaussian random vector,

where Y is one component, and the σ-fields correspond to subsets of components. In particular, with $\mathcal{F} := \sigma(\Omega, \varnothing)$, $\mathcal{G} := \mathcal{F}_0^{(k)}$, $\mathcal{H} := \mathcal{F}_0$ and $Y := R^{(k)}$, $\mathrm{Var}(R^{(k)} \mid \mathcal{F}_0) \leq \mathrm{Var}(R^{(k)} \mid \mathcal{F}_0^{(k)})$. Hence,

$$
\begin{aligned}
\mathrm{Var}(R \mid \mathcal{F}_0) &= \sum_{k_1,k_2=1}^{K} \mathrm{Cor}(R^{(k_1)}, R^{(k_2)} \mid \mathcal{F}_0)\, \mathrm{Var}(R^{(k_1)} \mid \mathcal{F}_0)^{1/2}\, \mathrm{Var}(R^{(k_2)} \mid \mathcal{F}_0)^{1/2} \\
&\leq \sum_{k_1,k_2=1}^{K} \mathrm{Cor}(R^{(k_1)}, R^{(k_2)} \mid \mathcal{F}_0)\, \mathrm{Var}(R^{(k_1)} \mid \mathcal{F}_0^{(k_1)})^{1/2}\, \mathrm{Var}(R^{(k_2)} \mid \mathcal{F}_0^{(k_2)})^{1/2}.
\end{aligned}
$$

For a given $c \in (0, \infty)^T$ and $A \in (0, \infty)$, solving for $a \in \mathbb{R}^T$ the standard convex optimization problem:

$$
\text{maximize } \sum_{t=1}^{T} c_t a_t^{1/2},
$$

$$
\text{subject to } \sum_{t=1}^{T} a_t \leq A,
$$

yields the unique maximizer and maximum:

$$
a_t = A \frac{c_t^{-2}}{\sum_{s=1}^{T} c_s^{-2}} \quad \text{and} \quad \frac{A^{1/2} T}{\left(\sum_{s=1}^{T} c_s^{-2} \right)^{1/2}}
$$

Notice that the expression Equation (5) for V_0 is of the form $\sum_{t=1}^{T} c_t a_t^{1/2}$ with $a_t = \mathrm{Var}(R \mid \mathcal{F}_{t-1}) - \mathrm{Var}(R \mid \mathcal{F}_t)$. For these a_t's, $\sum_{t=1}^{T} a_t = \mathrm{Var}(R \mid \mathcal{F}_0) - \mathrm{Var}(R \mid \mathcal{F}_T) = \mathrm{Var}(R \mid \mathcal{F}_0)$. □

Appendix A.3. Proof of Proposition 4

Proof. Notice that for $t = 1, \ldots, J^{(k)}$,

$$
\mathbb{E}\left[X_t^{o,(k)} \mid \mathcal{F}_0 \right] = \sum_{u=t+1}^{J^{(k)}+1} v_u^{(k)} \mathbb{E}\left[\widetilde{I}_{u,J^{(k)}-u+t+1}^{(k)} \mid \widetilde{I}_{u,J^{(k)}-u+1}^{(k)} \right],
$$

where:

$$
\mathbb{E}\left[\widetilde{I}_{u,J^{(k)}-u+t+1}^{(k)} \mid \widetilde{I}_{u,J^{(k)}-u+1}^{(k)} \right] = \widetilde{I}_{u,J^{(k)}-u+1}^{(k)} \prod_{l=J^{(k)}-u+2}^{J^{(k)}-u+t+1} \beta_l^{(k)} + \sum_{l=J^{(k)}-u+2}^{J^{(k)}-u+t+1} \alpha_l^{(k)} \prod_{m=l+1}^{J^{(k)}-u+t+1} \beta_m^{(k)}.
$$

It remains to determine the conditional variances. Due to independence between accident years:

$$
\begin{aligned}
\mathrm{Var}(R^{(k)} \mid \mathcal{F}_t^{(k)}) &= \sum_{i=t+2}^{J^{(k)}+1} (v_i^{(k)})^2 \, \mathrm{Var}\left(\sum_{j=J^{(k)}-i+2+t}^{J^{(k)}} I_{i,j}^{(k)} \mid \mathcal{F}_t^{(k)} \right) \\
&= \sum_{i=t+2}^{J^{(k)}+1} (v_i^{(k)})^2 \, \mathrm{Var}\left(\sum_{j=J^{(k)}-i+2+t}^{J^{(k)}} \sum_{l=J^{(k)}-i+2}^{j} \frac{\sigma_l^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{i,l}^{(k)} \prod_{m=l+1}^{j} \beta_m^{(k)} \right) \\
&= \sum_{i=t+2}^{J^{(k)}+1} (v_i^{(k)})^2 \, \mathrm{Var}\left(\sum_{l=J^{(k)}-i+2+t}^{J^{(k)}} \sum_{j=l}^{J^{(k)}} \frac{\sigma_l^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{i,l}^{(k)} \prod_{m=l+1}^{j} \beta_m^{(k)} \right) \\
&= \sum_{i=t+2}^{J^{(k)}+1} v_i^{(k)} \sum_{l=J^{(k)}-i+2+t}^{J^{(k)}} (\sigma_l^{(k)})^2 \left(1 + \sum_{j=l+1}^{J^{(k)}} \prod_{m=l+1}^{j} \beta_m^{(k)} \right)^2.
\end{aligned}
$$

Under the assumption $\text{Cov}(R^{(k_1)}, R^{(k_2)} \mid \mathcal{F}_t) = 0$ for all $k_1 \neq k_2$ and all t,

$$\text{Var}(R \mid \mathcal{F}_{t-1}) - \text{Var}(R \mid \mathcal{F}_t) = \sum_{k=1}^{K} \sum_{i=t+1}^{J^{(k)}+1} v_i^{(k)} \sigma_{J-i+1+t}^2 \left(1 + \sum_{j=J-i+2+t}^{J^{(k)}} \prod_{m=J^{(k)}-i+2+t}^{j} \beta_m^{(k)}\right)^2.$$

□

Appendix A.4. Proof of Proposition 6

Proof. Let:

$$B_l^{(k)} := \{I_{ij}^{(k)} : (i,j) \in \mathcal{I}^{(k)} \times \mathcal{J}^{(k)}, i+j \leq J^{(k)}+1, j \leq l\}.$$

First, notice that:

$$\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)}\right] = \mathbb{E}\left[\mathbb{E}\left[\widehat{\boldsymbol{\theta}}_j^{(k)} \mid B_{j-1}^{(k)}\right]\right] = \boldsymbol{\theta}_j^{(k)}.$$

Secondly, with $a, b \in \{1, 2\}$, and $l < m$,

$$\mathbb{E}\left[\widehat{\theta}_l^{(a),(k)}\widehat{\theta}_m^{(b),(k)}\right] = \mathbb{E}\left[\mathbb{E}\left[\widehat{\theta}_l^{(a),(k)}\widehat{\theta}_m^{(b),(k)} \mid B_{m-1}^{(k)}\right]\right] = \mathbb{E}\left[\widehat{\theta}_l^{(a),(k)}\mathbb{E}\left[\widehat{\theta}_m^{(b),(k)} \mid B_{m-1}^{(k)}\right]\right]$$
$$= \mathbb{E}\left[\widehat{\theta}_l^{(a),(k)}\right]\theta_m^{(b),(k)} = \theta_l^{(a),(k)}\theta_m^{(b),(k)}$$

and similarly for products with more factors. □

Appendix A.5. Proof of Proposition 7

Proof.

$$\mathbb{E}\left[\widehat{I}_{ij}^{(k)}\right] = \mathbb{E}\left[\mathbb{E}\left[\widehat{I}_{ij}^{(k)} \mid B_{J-i+1}^{(k)}\right]\right] = \mathbb{E}\left[I_{i,J-i+1}^{(k)} \prod_{l=J^{(k)}-i+2}^{j} \beta_l^{(k)} + \sum_{l=J^{(k)}-i+2}^{j} \alpha_l^{(k)} \prod_{m=l+1}^{j} \beta_m^{(k)}\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[I_{ij}^{(k)} \mid B_{J-i+1}^{(k)}\right]\right] = \mathbb{E}\left[I_{ij}^{(k)}\right].$$

Since each component of $\widehat{X}^o$ is a linear combination of predictors $\widehat{I}_{ij}^{(k)}$, the second statement follows immediately from the first statement. □

Appendix A.6. Proof of Proposition 8

Proof. In order to arrive at the conclusion of Proposition 8, we need to specify $\text{Cov}(\widehat{\boldsymbol{\theta}}^* \mid \mathcal{F}_0)$ and calculate $\nabla u_t(\boldsymbol{\theta})$ and $H_t(\boldsymbol{\theta})$.

Based on the matrix expressions presented in Section 3.1 and Section 3.2, it is straightforward to deduce that:

$$\text{Cov}(\widehat{\boldsymbol{\theta}}_j \mid \mathbf{A}_{j-1}) = \sigma_j^2 \begin{pmatrix} \frac{1}{c}\sum_{i=i_0}^{J-j+1} v_i \widetilde{I}_{i,j-1}^2 & -\frac{1}{c}\sum_{i=i_0}^{J-j+1} v_i \widetilde{I}_{i,j-1} & 0 \\ -\frac{1}{c}\sum_{i=i_0}^{J-j+1} v_i \widetilde{I}_{i,j-1} & \frac{1}{c}\sum_{i=i_0}^{J-j+1} v_i & 0 \\ 0 & 0 & \frac{2\sigma_j^2}{J-j-i_0+1} \end{pmatrix},$$

where we have omitted the notation for corresponding to the line of business. As described in Section 3.2, $\text{Cov}(\widehat{\boldsymbol{\theta}}_j^* \mid \mathcal{F}_0)$ is obtained by (independent) stacking of $\text{Cov}(\widehat{\boldsymbol{\theta}}_j^* \mid \mathcal{F}_0) := \text{Cov}(\widehat{\boldsymbol{\theta}}_j \mid \mathbf{A}_{j-1})$.

Turning to the calculation of $\nabla u_t(\boldsymbol{\theta})$ and $H_t(\boldsymbol{\theta})$ for the incremental model given by Equation (18), we have:

$$u_t(\boldsymbol{\theta}) = \left(\sum_{k=1}^{K} \sum_{i=t+1}^{J^{(k)}} v_i^{(k)} (\sigma_{J^{(k)}-i+1+t}^{(k)})^2 \left(1 + \sum_{l=J^{(k)}-i+2+t}^{J^{(k)}} \prod_{m=J^{(k)}-i+2+t}^{l} \beta_m^{(k)} \right)^2 \right)^{1/2},$$

which immediately gives us that the following components are the only non-zero contributions:

$$\{\nabla u_t(\boldsymbol{\theta})\}_{\beta_j^{(k)}} = \mathbb{1}_{\{t+2 \leq j \leq J^{(k)}\}} [u_t(\boldsymbol{\theta})]^{-1} \sum_{n=t+1}^{j-1} v_{J^{(k)}-n+1+t}^{(k)} (\sigma_n^{(k)})^2 \left(1 + \sum_{l=n+1}^{J^{(k)}} \prod_{m=n+1}^{l} \beta_m^{(k)} \right) \sum_{l=j}^{J^{(k)}} \prod_{\substack{m=n+1 \\ m \neq j}}^{l} \beta_m^{(k)}$$

$$\{\nabla u_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} = \frac{1}{2} \mathbb{1}_{\{t+1 \leq j \leq J^{(k)}\}} [u_t(\boldsymbol{\theta})]^{-1} v_{J^{(k)}-j+1+t}^{(k)} \left(1 + \sum_{l=j+1}^{J^{(k)}} \prod_{m=j+1}^{l} \beta_m^{(k)} \right)^2,$$

and so:

$$\{H_t(\boldsymbol{\theta})\}_{\beta_j} = \mathbb{1}_{\{t+2 \leq j \leq J^{(k)}\}} [u_t(\boldsymbol{\theta})]^{-1} \sum_{n=t+1}^{j-1} v_{J^{(k)}-n+1+t}^{(k)} (\sigma_n^{(k)})^2 \left(\sum_{l=j}^{J^{(k)}} \prod_{\substack{m=n+1 \\ m \neq j}}^{l} \beta_m^{(k)} \right)^2 - [u_t(\boldsymbol{\theta})]^{-1} [\{\nabla u_t(\boldsymbol{\theta})\}_{\beta_j}]^2$$

$$\{H_t(\boldsymbol{\theta})\}_{\sigma_j^2} = - [u_t(\boldsymbol{\theta})]^{-1} [\{\nabla u_t(\boldsymbol{\theta})\}_{\sigma_j^2}]^2.$$

By combining the above, we arrive at the desired result. □

Appendix A.7. Proof of Proposition 9

Proof. As in the proof of Theorem 3 in Mack (1993a), the process variance is already calculated (see Proposition 4), and consequently, its estimator is again a straightforward plug-in estimator. The complicated part is to take care of the estimation error. First, consider a single accident year i and the natural reserve estimator $\widehat{R}_i = I_{i,J-i+1} \sum_{j=J-i+2}^{J} \prod_{l=J-i+2}^{j} \widehat{\beta}_l$, which gives us:

$$\widehat{R}_i - \mathbb{E}[R_i \mid \mathcal{F}_0] = I_{i,J-i+1} \left(\sum_{j=J-i+2}^{J} \prod_{l=J-i+2}^{j} \widehat{\beta}_l - \sum_{j=J-i+2}^{J} \prod_{l=J-i+2}^{j} \beta_l \right).$$

If we follow the steps of Mack (1993a), we start by introducing:

$$F^{J-i+2,j} = \sum_{l=J-i+2}^{j} S_l^{J-i+1,j},$$

where:

$$S_l^{J-i+2,j} = \widehat{\beta}_{J-i+2} \cdots \widehat{\beta}_{l-1} (\widehat{\beta}_l - \beta_l) \beta_{l+1} \cdots \beta_j,$$

that is:

$$\widehat{R}_i - \mathbb{E}[R_i|\mathcal{F}_0] = I_{i,J-i+1} \sum_{j=J-i+2}^{J} F^{J-i+2,j},$$

and in particular:

$$\left(\sum_{j=J-i+2}^{J} F^{J-i+2,j} \right)^2 = \sum_{j=J-i+2}^{J} \sum_{l=J-i+2}^{j} (S_l^{J-i+2,j})^2 + 2 \sum_{j<j'} \sum_{l=J-i+2}^{j} \sum_{l'=J-i+2}^{j'} S_l^{J-i+2,j} S_{l'}^{J-i+2,j'}.$$

Further, the approximation idea in Mack (1993a) is then:

$$(S_l^{J-i+2,j})^2 \approx \mathbb{E}\left[(S_l^{J-i+2,j})^2 \mid \mathcal{F}_{l-1}\right]$$

$$= \widehat{\beta}_{J-i+2}^2 \cdots \widehat{\beta}_{l-1}^2 \widehat{\beta}_l^2 \widehat{\beta}_{l+1}^2 \cdots \widehat{\beta}_j^2 \widehat{\mathrm{Var}}(\widehat{\beta}_l \mid \mathcal{F}_{l-1})/\widehat{\beta}_l^2$$

$$= \left(\prod_{m=J-i+2}^{j} \widehat{\beta}_m^2\right) \widehat{\mathrm{Var}}(\widehat{\beta}_l \mid \mathcal{F}_{l-1})/\widehat{\beta}_l^2,$$

and:

$$S_l^{J-i+2,j} S_{l'}^{J-i+2,j'} \approx \mathbb{E}\left[S_l^{J-i+2,j} S_{l'}^{J-i+2,j'} \mid \mathcal{F}_{l'-1}\right] = \begin{cases} 0 & \text{if } l < l', \\ \widehat{s}_l^{J-i+2,j} \widehat{s}_l^{J-i+2,j'} & \text{if } l = l', \end{cases}$$

where:

$$\widehat{s}_l^{J-i+2,j} = \left(\prod_{m=J-i+2}^{j} \widehat{\beta}_m\right) \sqrt{\widehat{\mathrm{Var}}(\widehat{\beta}_l \mid \mathcal{F}_{l-1})}/\widehat{\beta}_l.$$

Thus, by combining the above, it holds that:

$$\left(\sum_{j=J-i+2}^{J} F^{J-i+2,j}\right)^2 \approx \sum_{j=J-i+2}^{J} \sum_{l=J-i+2}^{j} (\widehat{s}_l^{J-i+2,j})^2 + 2 \sum_{j<j'} \sum_{l=J-i+2}^{j} \sum_{l'=J-i+2}^{j'} 1_{\{l=l'\}} \widehat{s}_l^{J-i+2,j} \widehat{s}_{l'}^{J-i+2,j'}.$$

The MSEP for the total reserve R is obtained in the same way as for the individual accident years, but there is a need to treat terms that describe the dependence between two accident years i and i', that is:

$$G_{i,i'} := \left(\sum_{j=J-i+2}^{J} F^{J-i+2,j}\right)\left(\sum_{j'=J-i'+2}^{J} F^{J-i'+2,j'}\right),$$

which follow analogously from the single accident year situation, i.e.,

$$G_{i,i'} \approx \sum_{j=J-i+2}^{J} \sum_{j'=J-i'+2}^{J} \sum_{l=J-i+2}^{j} \sum_{l'=J-i'+2}^{j'} 1_{\{l=l'\}} \widehat{s}_l^{J-i+2,j} \widehat{s}_{l'}^{J-i+2,j'}.$$

$\square$

Appendix A.8. Proof of Proposition 10

Proof. Notice that from:

$$\widetilde{C}_{iJ} = \frac{\sigma_J^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{iJ}^{(k)} + \gamma_{J-1} \frac{\sigma_{J-1}^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{i,J-1}^{(k)} + \cdots + \gamma_{J-1} \cdots \gamma_{J-i+2+t} \frac{\sigma_{J-i+2+t}^{(k)}}{\sqrt{v_i^{(k)}}} \epsilon_{i,J-i+2+t}^{(k)}$$

$$+ \gamma_{J-1} \cdots \gamma_{J-i+2+t} \widetilde{C}_{i,J-i+1+t}$$

it follows that:

$$
\begin{aligned}
\mathrm{Var}(C_{iJ}^{(k)} \mid \mathcal{F}_t) &= \mathrm{Var}(C_{iJ}^{(k)} - C_{i,J-i+1+t}^{(k)} \mid \mathcal{F}_t) \\
&= v_i^{(k)} \sum_{j=J^{(k)}-i+2+t}^{J^{(k)}} \sigma_j^2 \prod_{m=j}^{J^{(k)}-1} \gamma_m^2,
\end{aligned}
$$

$$
\mathrm{Var}(R^{(k)} \mid \mathcal{F}_t) = \sum_{i=2+t}^{J^{(k)}+1} v_i^{(k)} \sum_{j=J^{(k)}-i+2+t}^{J^{(k)}} \sigma_j^2 \prod_{m=j}^{J^{(k)}-1} \gamma_m^2,
$$

$$
\mathrm{Var}(R^{(k)} \mid \mathcal{F}_{t-1}) - \mathrm{Var}(R^{(k)} \mid \mathcal{F}_t) = \sum_{i=1+t}^{J^{(k)}+1} v_i^{(k)} \sigma_{J^{(k)}-i+1+t}^2 \prod_{m=J^{(k)}-i+1+t}^{J^{(k)}-1} \gamma_m^2,
$$

from which the conclusion follows.　□

Appendix A.9. Proof of Proposition 12

Proof. The steps for the cumulative model will again be identical to those for the incremental model, but simpler. For further details, we refer to Proof Appendix A.6.

To start off, $\mathrm{Cov}(\widehat{\boldsymbol{\theta}}^* \mid \mathcal{F}_0)$ is constructed based on $\mathrm{Cov}(\widehat{\boldsymbol{\theta}}_j^* \mid \mathcal{F}_0) := \mathrm{Cov}(\widehat{\boldsymbol{\theta}}_j \mid \mathbf{A}_{j-1})$, where:

$$
\mathrm{Cov}(\widehat{\boldsymbol{\theta}}_j \mid \mathbf{A}_{j-1}) = \sigma_j^2 \begin{pmatrix} \frac{1}{\sum_{i=i_0}^{J-j+1} v_i \widetilde{C}_{i,j-1}^2} & 0 \\ 0 & \frac{2\sigma_j^2}{J-j-i_0} \end{pmatrix},
$$

for $j = 2, \ldots, J$. For the cumulative model, we have:

$$
u_t(\boldsymbol{\theta}) = \left(\sum_{k=1}^{K} \sum_{i=t+1}^{J^{(k)}+1} v_i^{(k)} (\sigma_{J^{(k)}-i+t+1}^{(k)})^2 \prod_{m=J^{(k)}-i+t+1}^{J^{(k)}-1} (\gamma_m^{(k)})^2 \right)^{1/2},
$$

which gives us that the only non-zero elements of $\nabla u_t(\boldsymbol{\theta})$ and $H_t(\boldsymbol{\theta})$ turn out to be:

$$
\{\nabla u_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}} = 1_{\{t+1\leq j\}} [u_t(\boldsymbol{\theta})]^{-1} \sum_{n=t+1}^{j} v_{J^{(k)}-n+t+1}^{(k)} (\sigma_n^{(k)})^2 \gamma_j^{(k)} \prod_{\substack{m=n \\ m\neq j}}^{J^{(k)}-1} (\gamma_m^{(k)})^2
$$

$$
\{\nabla u_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} = 1_{\{t+1\leq j\}} \frac{1}{2} [u_t(\boldsymbol{\theta})]^{-1} v_{J^{(k)}-j+t+1}^{(k)} \prod_{m=j}^{J^{(k)}-1} (\gamma_m^{(k)})^2,
$$

and consequently:

$$
\{H_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}} = \{\nabla u_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}} \left(\frac{1}{\gamma_j^{(k)}} - \frac{\{\nabla u_t(\boldsymbol{\theta})\}_{\gamma_j^{(k)}}}{u_t(\boldsymbol{\theta})} \right)
$$

$$
\{H_t(\boldsymbol{\theta})\}_{(\sigma_j^{(k)})^2} = - [u_t(\boldsymbol{\theta})]^{-1} [\{\nabla u_t(\boldsymbol{\theta})\}_{\sigma_j^2}]^2.
$$

Given the above, the results of the proposition follows.　□

Brockwell, Peter J., and Richard A. Davis. 1987. *Time Series: Theory and Methods*. Springer Series in Statistics. New York: Springer.

Buchwalder, Markus, Hans Bühlmann, Michael Merz, and Mario V. Wüthrich. 2006. The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *Astin Bulletin* 36: 521–42.

Christensen, Ronald. 1991. *Linear Models for Multivariate, Time Series, and Spatial Data*. Berlin: Springer Science & Business Media.

EIOPA. 2015. Guidlines on the valuation of technical provisions. *Official Publication of EIOPA*. EIOPA-BoS-14/166 EN.

England, Peter, and Richard Verrall. 1999. Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: Mathematics and Economics* 25: 281–93.

Engsner, Hampus, Mathias Lindholm, and Filip Lindskog. 2017. Insurance valuation: A computable multi-period cost-of-capital approach. *Insurance: Mathematics and Economics* 72: 250–64.

European Commission. 2015. Commission Delegated Regulation (EU) 2015/35 of 10 October 2014. *Official Journal of the European Union 58*.

Fahrmeir, Ludwig, Thomas Kneib, Stefan Lang, and Brian Marx. 2013. *Regression: Models, Methods and Applications*. Berlin: Springer Science & Business Media.

Ferriero, Alessandro. 2016. Solvency capital estimation, reserving cycle and ultimate risk. *Insurance: Mathematics and Economics* 68: 162–68.

Gisler, Alois. 2009. The insurance risk in the SST and in Solvency II: Modelling and parameter estimation. *SSRN Manuscript*. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2704364 (accessed on 15 July 2017).

Kremer, Erhard. 1984. A class of autoregressive models for predicting the final claims amount. *Insurance: Mathematics and Economics* 3: 111–19.

Lindholm, Mathias, Filip Lindskog, and Felix Wahl. 2017. Comments on the classical chain-ladder MSEP calculation. Unpublished manuscript.

Mack, Thomas. 1993a. Distribution-free calculation of the standard error of chain ladder reserve estimates. *Astin Bulletin* 23: 213–25.

Mack, Thomas. 1993b. Measuring the variability of chain ladder reserve estimates. *Casualty Actuarial Society Forum* 1: 101–83.

Merz, Michael, and Mario V. Wüthrich. 2008. Modelling the claims development result for solvency purposes. *Casualty Actuarial Society E-Forum Fall 2008*: 542–68.

Möhr, C. 2011. Market-consistent valuation of insurance liabilities by cost of capital. *Astin Bulletin* 41: 315–41.

Ohlsson, Esbjörn, and Jan Lauzeningks. 2009. The one-year non-life insurance risk. *Insurance: Mathematics and Economics* 45: 203–8.

Salzmann, Robert, and Mario V. Wüthrich. 2010. Cost-of-capital margin for a general insurance liability runoff. *Astin Bulletin* 40: 415–51.

Wüthrich, Mario V., Paul Embrechts, and Andreas Tsanakas. 2011. Risk margin for a non-life insurance run-off. *Statistics & Risk Modeling with Applications in Finance and Insurance* 28: 299–317.

Wüthrich, Mario V., and Michael Merz. 2008. *Stochastic Claims Reserving Methods in Insurance*. Hoboken: John Wiley & Sons, vol. 435.