Mathematical Statistics
Stockholm University

The collective reserving model

Felix Wahl[†], Mathias Lindholm[†] and Richard Verrall[‡]

Research Report 2018:19

ISSN 1650-0377

Postal address:
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:
http://www.math.su.se

Mathematical Statistics
Stockholm University
Research Report **2018:19**,
http://www.math.su.se

The collective reserving model

Felix Wahl† Mathias Lindholm† Richard Verrall‡

September 2018

Abstract

This paper sets out a model for analysing claims development data, which we call the collective reserving model (CRM). The model is defined on the individual claim level and it produces separate IBNR and RBNS reserve estimators at the collective level *without* using any approximations. The CRM is based on ideas from a paper by Verrall, Nielsen and Jessen (VNJ) from 2010 in which a model is proposed that relies on a claim giving rise to a single payment. This is generalised by the CRM to the case of multiple payments per claim. All predictors of outstanding claims payments for the VNJ model are shown to hold for this new model. Moreover, the quasi-Poisson GLM estimation framework will be applicable as well. Furthermore, analytical expressions for the variance of the total outstanding claims payments are given, with a subdivision on IBNR and RBNS claims. To quantify the effect of allowing only one payment per claim, the model is related and compared to the VNJ model, in particular by looking at variance inequalities. The double chain ladder (DCL) method is discussed as an estimation method for this new model and it is shown that both the GLM- and DCL-based estimators are consistent in terms of an exposure measure. Lastly, both of these methods are shown to asymptotically reproduce the regular chain ladder reserve estimator, motivating the chain ladder technique as a large-exposure approximation of this model.

JEL: G22.

Keywords: Stochastic claims reserving; risk; solvency; chain ladder.

1 Introduction

There has been much research carried out in the area of triangle-based reserving methods, with one of the most famous methods being the chain ladder technique. The chain ladder technique is a simple algorithm used to predict outstanding claims payments based on historical payment patterns. Initially it was not based on any stochastic model, and for this reason, considerable effort has been made trying to motivate its use. Models that produce the same estimators as the chain ladder technique can be found in, for instance, Mack (1991) and Renshaw and Verrall (1998). In Mack (1991), the maximum likelihood estimators of a multiplicative Poisson model and the estimators from the chain ladder technique were shown to coincide. In Renshaw and Verrall (1998), another prominent triangle-based model, the over-dispersed Poisson model, was introduced and also shown to have estimators which

†Department of Mathematics, Stockholm University, Sweden
‡Cass Business School, City, University of London, U.K.

coincide with the chain ladder technique. Further, Mack (1993) presented a model often referred to as the distribution-free chain ladder model, which built on ideas from Schnieper (1991).

Instead of looking for a model that motivates the chain ladder technique, other authors start from a suitable stochastic model given some assumptions on the data generating process, and derive reserve estimators based on this. A few articles have even moved away from aggregated data and developed models at the level of individual claims; see for instance Norberg (1993, 1999), and more recently, Antonio and Plat (2014). In Antonio and Plat (2014), a likelihood approach is taken based on the framework of position dependent marked Poisson processes and recurrent events, and a complete simulation procedure is detailed based on this to project future losses. Others use aggregated data but based on models built from the ground up, such as Bühlmann et al. (1980), Norberg (1986) and Verrall et al. (2010).

In this paper, we present a new model for stochastic claims reserving which provides a connection between detailed assumptions for individual claims and a model for data aggregated into triangles. The model is motivated by the starting point for all actuarial modelling: risk theory based on Poisson-generated events. This new model is related to the approach set out in Verrall et al. (2010) in that it is based on detailed assumptions on individual claims, but is then modelled by an over-dispersed Poisson distribution on the aggregate level. In the proposed model, each individual claim is assumed to follow a discrete-time inhomogeneous Poisson process and the number of claims is assumed to be generated by an over-dispersed Poisson distribution. It is helpful to give our new model a name, and we decided on calling it the collective reserving model (CRM). The reason for choosing this name is that it emphasises the similarity in the approach with collective risk theory: we start with events generated by Poisson processes and derive distributions for a portfolio of risks as a whole *without* approximations. For a detailed discussion on collective risk models, see Mikosch (2009) (Part I).

The model of Verrall et al. (2010) relies on the assumption that each claim will give rise to a single payment (which can be zero). The model proposed in this paper, the CRM, generalises this to the case of multiple (including zero) payments per claim. We will show that all predictors of outstanding payments from Verrall et al. (2010) hold for this new model, and, moreover, the quasi-Poisson generalized linear model (GLM) estimation framework will be applicable as well. There will, however, be subtle differences between the models. To name one; the CRM presented here can be represented as an over-dispersed Poisson distribution, while in Verrall et al. (2010) this is an approximation. Therefore the aggregate modelling of this new model will be exact, as opposed to the model in Verrall et al. (2010). This and other differences will be discussed in this paper. Furthermore, we give analytical expressions for the variance of the total outstanding claims payments, with a subdivision on IBNR and RBNS claims.

Further, to quantify the practical implications of using the simplifying assumption of one payment from Verrall et al. (2010), we will relate and compare the CRM to the model in Verrall et al. (2010), especially by looking at second moments. This is done in two ways. First by relating the models through a theoretical construct that amounts to replacing a series of payments with a single payment whose first two moments are equal to those of the underlying series, and then comparing the resulting model variances. Secondly, by comparing the estimators of the variances obtained by plugging in the parameter estimators.

Parameter estimation is, as mentioned, primarily done in the same way as in Verrall et al. (2010) and Miranda et al. (2011). Another estimation approach is the double chain ladder method as described in Miranda et al. (2012). We show in this paper that both of these methods yield consistent estimators and that the reserves of the methods are asymptotically close. The asymptotic argument is based on an exposure measure and is therefore not dependent on the size of the claims triangles. Furthermore, we show that both of these methods asymptotically reproduce the standard chain ladder reserve estimator, thereby motivating it as a large-exposure approximation of the collective reserving model.

Finally, we give a brief discussion on predictive uncertainty. To quantify predictive uncertainty, bootstrap methods are often used. See for instance Pinheiro et al. (2003), Björkwall et al. (2009) and England and Verrall (1999) for bootstrap methods applied to reserving. In Miranda et al. (2011) this is discussed in detail for the model of Verrall et al. (2010), setting up all the necessary components for a parametric bootstrap, all of which will apply, with slight modification, for the model presented in the present paper. To add to this, we will present a different approach to account for predictive uncertainty, based on recent developments in semi-analytical approximations of the mean squared error of prediction (MSEP) in Lindholm et al. (2018). For similar results for the chain ladder model, see Diers et al. (2016), Röhr (2016) and Buchwalder et al. (2006). This approximation of the MSEP is, naturally, based on the variances of the total outstanding claims payments, and we will see that the MSEP differs between the CRM and the model of Verrall et al. (2010) only as much as their variances differ. Moreover, the described approach only relies on quantities obtained from a standard quasi-Poisson GLM fit using e.g. `R`.

The remainder of the paper is organized as follows: In Section 2 a brief description of the notation and relevant data used for the models is given. Section 3 introduces the CRM and several theoretical results, such as moments and model fitting, and how the model relates to the chain ladder technique and the DCL method. In Section 4 the CRM is compared to the model of Verrall et al. (2010) in terms of variance inequalities, and in Section 5 the semi-analytical approximation of the MSEP is given. Finally, a numerical example, which compares the two models and benchmarks them against the chain ladder model, is presented in Section 6.

2 Data and notation

This section introduces the notation and relevant data used throughout this paper. Let m be the total number of accident years and let d be the maximal delay with which claims payments are made after being reported. Using this, let X_{ij} denote the incremental aggregated claims payments during development year $j \in \{0, \dots, m+d-1\}$, and let N_{ij} denote the incremental number of claims incurred in development year $j \in \{0, \dots, m-1\}$, both from accidents during accident year $i \in \{1, \dots, m\}$. The data can be formatted in the following rectangles:

$$\begin{array}{ccc} N_{10} & \dots & N_{1,m-1} \\ N_{20} & \dots & N_{2,m-1} \\ \vdots & & \vdots \\ N_{m0} & \dots & N_{m,m-1} \end{array} \qquad \begin{array}{ccc} X_{10} & \dots & X_{1,m+d-1} \\ X_{20} & \dots & X_{2,m+d-1} \\ \vdots & & \vdots \\ X_{m0} & \dots & X_{m,m+d-1} \end{array}$$

The data known "today" and the data used for prediction consists of the X_{ij} and N_{ij} where $(i,j) \in \mathcal{A}_0$ where

$$\mathcal{A}_0 := \{(i,j) \in \mathbb{N} \times \mathbb{N}_0 : i + j \leq m\}.$$

That is, data on the following form:

$$\begin{array}{ccccc} N_{10} & N_{11} & \dots & N_{1,m-2} & N_{1,m-1} \\ N_{20} & N_{21} & \dots & N_{2,m-2} & \\ \vdots & \ddots & & & \\ N_{m0} & & & & \end{array} \qquad \begin{array}{ccccc} X_{10} & X_{11} & \dots & X_{1,m-2} & X_{1,m-1} \\ X_{20} & X_{21} & \dots & X_{2,m-2} & \\ \vdots & \ddots & & & \\ X_{m0} & & & & \end{array}$$

For estimation of parameters needed for the variance of the total outstanding claims payments we will need a triangle consisting of the total number of payments made in $(i,j) \in \mathcal{A}_0$, denoted by N_{ij}^{paid}.

The lower right trapezoid that is to be predicted consists of the X_{ij}s with indices in the set

$$\mathcal{A}_0^* := \{(i,j) \in \mathbb{N} \times \mathbb{N}_0 : i \leq m,\ j \leq m-1+d,\ i+j \geq m+1\}.$$

If we only consider predictions confined to a regular lower right triangle, such as the predictions made with the chain ladder technique, the index set for predictions would be

$$\tilde{\mathcal{A}}_0^* := \{(i,j) \in \mathcal{A}_0^* : j \leq m-1\}.$$

It will also be useful to define the filtration

$$\mathcal{N}_0 := \sigma\{N_{ij} : (i,j) \in \mathcal{A}_0\},$$

corresponding to the information generated by the number of claims observed up until "today".

Lastly, in this paper we will define the CRM and the model of Verrall et al. (2010), referred to as the VNJ model. Superscripts "(CRM)" and "(VNJ)" will be used to indicate which of these models certain quantities are referring to.

3 A model with multiple payments per claim

The model introduced in Verrall et al. (2010) relies on a claim giving rise to a single payment. Given this assumption, the authors show how predictors of the total outstanding claims payments are calculated and describe how the model may be fitted as a quasi-Poisson GLM. A formal definition of the model from Verrall et al. (2010), henceforth referred to as the VNJ model, together with a summary of relevant properties, can be found in Appendix A.

The single payment assumption is in many cases unrealistic. Inspired by the work in Verrall et al. (2010) and Miranda et al. (2011, 2012), we will in this section introduce a new model, referred to as the collective reserving model (CRM). The CRM will allow for multiple payments per claim, but still possess the desirable features of the VNJ model discussed in the aforementioned papers.

Following the introduction of the new model, we will in Section 3.1 make a simplifying assumption and calculate expectations and variances of claim amounts. Thereafter, we will discuss parameter estimation and show consistency of parameter estimators, primarily in terms of an exposure measure, before moving on to giving separate predictors of outstanding IBNR and RBNS claims payments with corresponding variances. In Section 3.2 we will discuss the double chain ladder technique as an estimation method for the CRM and present corresponding consistency results. Finally, we motivate the chain ladder technique by an asymptotic argument.

The new multiple payments model is defined as follows:

CRM: The collective reserving model. *Let the total number of claims incurred in accident year i that are reported j periods later, N_{ij}, follow an over-dispersed Poisson distribution according to $N_{ij} \sim \mathrm{ODP}(\nu_{ij}, \phi)$, i.e.*

$$\mathbb{E}[N_{ij}] = \nu_{ij},$$
$$\mathrm{Var}(N_{ij}) = \phi\mathbb{E}[N_{ij}],$$

where all N_{ij} are assumed to be independent. Further, each claim will produce a series of payments according to a discrete-time Poisson process, that is

$$N_{ijk}^{\mathrm{paid},(CRM)} \mid N_{ij} \sim \mathrm{Po}(\lambda_k N_{ij}), \quad \lambda_k \geq 0,\ k = 0, \dots, d,$$

where $N_{ijk}^{\mathrm{paid},(CRM)}$ denotes the number of payments stemming from claims reported in (i,j), paid k periods later. Similarly, let $\widetilde{Y}_{ijkl}$ denote the lth payment from claims reported in (i,j), paid k periods later, where all $\widetilde{Y}_{ijkl}$ are assumed to be independent over all l for fixed (i,j,k) with

$$\mathbb{E}[\widetilde{Y}_{ijkl}] = \mu_{ijk},$$
$$\mathrm{Var}(\widetilde{Y}_{ijkl}) = \sigma_{ijk}^2.$$

Note that by introducing $N_{ij}^{\mathrm{paid},(\mathrm{CRM})}$ according to

$$N_{ij}^{\mathrm{paid},(\mathrm{CRM})} := \sum_{k=0}^{j\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})},$$

i.e. the number of payments in (i,j), it follows that

$$N_{ij}^{\mathrm{paid},(\mathrm{CRM})} | \{N_{il}\}_{l\leq j} \sim \mathrm{Po}\left(\sum_{k=0}^{j\wedge d} \lambda_k N_{i,j-k}\right).$$

Further, the total aggregated incremental payments in (i,j) may be written according to

$$X_{ij}^{(\mathrm{CRM})} := \sum_{k=0}^{j\wedge d} X_{i,j-k,k}^{(\mathrm{CRM})} := \sum_{k=0}^{j\wedge d} \sum_{l=0}^{N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}} \widetilde{Y}_{i,j-k,k,l}, \tag{1}$$

where the superscript (CRM) refers to the collective reserving model. This implies the following payment dynamics of the CRM:

Proposition 1. *For the CRM it holds that*

$$\mathbb{E}[X_{ij}^{(CRM)}|\mathcal{N}_0] = \sum_{k=0}^{j\wedge d} \mathbb{E}[X_{i,j-k,k}^{(CRM)} \mid \mathcal{N}_0] = \sum_{k=0}^{j\wedge d} \lambda_k \mu_{i,j-k,k} N_{i,j-k}, \tag{2}$$

and that

$$\mathrm{Var}(X_{ij}^{(CRM)}|\mathcal{N}_0) = \sum_{k=0}^{j\wedge d} \mathrm{Var}(X_{i,j-k,k}^{(CRM)}|\mathcal{N}_0) = \sum_{k=0}^{j\wedge d} \frac{\sigma_{i,j-k,k}^2 + \mu_{i,j-k,k}^2}{\mu_{i,j-k,k}} \mathbb{E}[X_{i,j-k,k}^{(CRM)} \mid \mathcal{N}_0]. \tag{3}$$

Moreover, all $X_{ij}^{(CRM)}$ and $X_{ijk}^{(CRM)}$ are independent.

The proof is given in Appendix B. A problem with the general formulation of Proposition 1 is that without additional structure, parameters with indices in $\mathcal{A}_0^*$, i.e. the lower right trapezoid including the tail, cannot be estimated based on the data with indices in $\mathcal{A}_0$. An example of structure that makes estimation possible is setting $\mu_{ijk} = \mu_i$ and $\sigma_{ijk}^2 = \sigma_i^2$. There is always the possibility of extrapolation, but we will refrain from going into this topic. In the next section we introduce a simplifying assumption that will deal with this problem.

3.1 The homogeneous CRM

In this section, we will make a simplifying assumption on the payment distribution, after which we will calculate expectations and variances of claim amounts. Following this, we will discuss parameter estimation and show consistency of parameter estimators, primarily in terms of an exposure measure. Then we move on to give separate predictors of outstanding IBNR and RBNS claims payments with corresponding variances.

The following assumption is a natural starting point for modelling:

Assumption 1. *All $\widetilde{Y}_{ijkl}$s are independent with*

$$\begin{aligned}\mathbb{E}[\widetilde{Y}_{ijkl}] &= \mu,\\ \mathrm{Var}(\widetilde{Y}_{ijkl}) &= \sigma^2.\end{aligned}$$

This simplifying assumption more or less amounts to payments being i.i.d. This is, however, unlikely to be valid in practise, and in further developments of the CRM, this assumption can likely be relaxed. Given Assumption 1 we obtain the following corollary to Proposition 1:

Corollary 1. *Given Assumption 1, it holds that*

$$\mathbb{E}[X_{ij}^{(CRM)}|\mathcal{N}_0] = \mu \sum_{k=0}^{j\wedge d} \lambda_k N_{i,j-k}, \tag{4}$$

and that

$$\mathrm{Var}(X_{ij}^{(CRM)}|\mathcal{N}_0) = \frac{\sigma^2+\mu^2}{\mu}\mathbb{E}[X_{ij}^{(CRM)}|\mathcal{N}_0]. \tag{5}$$

That is, $X_{ij}^{(CRM)}$ may be represented as an over-dispersed Poisson distribution with mean $\mathbb{E}[X_{ij}^{(CRM)}|\mathcal{N}_0]$ and over-dispersion parameter

$$\varphi := \frac{\sigma^2+\mu^2}{\mu}.$$

Moreover, all $X_{ij}^{(CRM)}$ are independent.

Note that in Corollary 1 we have equality, whereas there is only *approximate* equality for the variance of the VNJ model, see (24) in Appendix A. Henceforth, unless stated otherwise, all $\widetilde{Y}_{ijkl}$s will be assumed to follow Assumption 1 to allow for easy comparison to results from Verrall et al. (2010) and Miranda et al. (2011, 2012). Moreover, the CRM will refer to the homogeneous CRM in accordance with Assumption 1.

Further, note that if we let

$$\Lambda := \sum_{k=0}^{d} \lambda_k,$$

$\pi_k = \lambda_k/\Lambda$ and $\tilde{\mu} = \mu\Lambda$, it follows that

$$\begin{aligned}\mathbb{E}[X_{ij}^{(\mathrm{CRM})}|\mathcal{N}_0] &= \mu \sum_{k=0}^{j\wedge d} \lambda_k N_{i,j-k}\\ &= \tilde{\mu} \sum_{k=0}^{j\wedge d} \pi_k N_{i,j-k}.\end{aligned}$$

Here $\tilde{\mu}$ can be interpreted as the expected total claim amount for a single claim and π_k corresponds to the distribution of the payments over time from reporting since $\pi_0+\ldots+\pi_d = 1$. Further, note that, in agreement with Verrall et al. (2010), see Appendix A, the CRM may be fitted using standard GLM theory, replacing

$$\psi_k := \mu p_k$$

from Verrall et al. (2010) with

$$\tilde{\psi}_k := \tilde{\mu}\pi_k,$$

where

$$\sum_{k=0}^{d} \pi_k = \sum_{k=0}^{d} p_k = 1.$$

That is, in our situation

$$\mathbb{E}[X_{ij}^{(\text{CRM})}|\mathcal{N}_0] = \sum_{k=0}^{j\wedge d} \tilde{\psi}_k N_{i,j-k}.$$

Since these two parameters, ψ_k and $\tilde{\psi}_k$, will be estimated in the same way, we drop the tilde and refer to both as ψ_k. The above means that the parameters ψ_k in the CRM are estimated by $\hat{\psi}_k$ which maximize the quasi-likelihood

$$l^p(\boldsymbol{\psi};\mathcal{F}_0) = \sum_{i,j\in\mathcal{A}_0}\left[X_{ij}\log\left(\sum_{k=0}^{j\wedge d}\psi_k N_{i,j-k}\right) - \sum_{k=0}^{j\wedge d}\psi_k N_{i,j-k}\right].$$

Moreover, by assuming that $\mathbb{E}[N_{ij}] = \alpha_i\beta_j$, estimation and prediction of the count variables N_{ij} may be carried out using the classical chain ladder technique, see for instance Mack (1991). Thus, all expressions for predictors of outstanding payments from Verrall et al. (2010) will hold for the CRM as well. In particular, the predictor of outstanding claims payments can be sub-divided into an IBNR and an RBNS part, which we will return to after a short discussion on parameter estimation.

Note that by using the above parametrisation we have that

$$\sum_{k=0}^{d}\psi_k = \tilde{\mu}\sum_{k=0}^{d}\pi_k = \tilde{\mu} = \mu\Lambda,$$

that is, unless we have explicit knowledge of $\{N_{ij}^{\text{paid},(\text{CRM})}\}$ we will not be able to identify μ and Λ. Since we may express all predictors of outstanding payments in terms of ψ_k, this only poses an issue when interpreting model output and, as we will see below, when estimating the variance of outstanding claims payments. From a practical perspective, however, it does not seem controversial to assume knowledge of $\{N_{ij}^{\text{paid},(\text{CRM})}\}$, which would allow for standard estimation of μ by, for instance,

$$\hat{\mu} = \frac{1}{|\mathcal{A}_0|}\sum_{i,j\in\mathcal{A}_0}\frac{X_{ij}}{N_{ij}^{\text{paid},(\text{CRM})}},$$

where $|\mathcal{A}_0| = m(m+1)/2$ is the size of the observed triangles. Using this we can then estimate the rest of the parameters by

$$\begin{cases} \hat{\pi}_k &= \frac{\hat{\psi}_k}{\sum_{j=0}^{d}\hat{\psi}_j}, \\ \hat{\Lambda} &= \frac{\sum_{j=0}^{d}\hat{\psi}_j}{\hat{\mu}}, \\ \hat{\lambda}_k &= \frac{\hat{\psi}_k}{\hat{\mu}}, \\ \hat{\sigma}^2 &= \hat{\varphi}\hat{\mu} - (\hat{\mu})^2, \end{cases}$$

where the over-dispersion parameter φ is estimated using (11) in Miranda et al. (2011), i.e.

$$\hat{\varphi} = \frac{1}{|\mathcal{A}_0| - (d+1)} \sum_{i,j \in \mathcal{A}_0} \frac{\left(X_{ij} - \sum_{k=0}^{j\wedge d} \hat{\psi}_k N_{i,j-k}\right)^2}{\sum_{k=0}^{j\wedge d} \hat{\psi}_k N_{i,j-k}}.$$

Further, assuming that $\nu_{ij} = \alpha_i \beta_j$, ϕ is estimated by

$$\hat{\phi} = \frac{1}{|\mathcal{A}_0| - m(m-1)} \sum_{i,j \in \mathcal{A}_0} \frac{(N_{ij} - \hat{\nu}_{ij})^2}{\hat{\nu}_{ij}}.$$

Remark 1. *Note that $\hat{\mu}$ is needed for identification of μ and Λ – the CRM and the VNJ model use the same $\hat{\psi}_k$s given the original quasi-Poisson GLM-based estimation procedure in Verrall et al. (2010). Therefore, the predictors of outstanding claims payments, given by* (4) *and* (22)*, will be equal given estimation by plugging in the $\hat{\psi}_k$s.*

Now, given the above parameter estimators, we can state the following proposition dealing with consistency:

Proposition 2. *The maximum quasi-likelihood estimator $\hat{\boldsymbol{\psi}}$ of the CRM and the VNJ model satisfy*

$$(i) \qquad \hat{\boldsymbol{\psi}} \xrightarrow{p} \boldsymbol{\psi}$$

as $n \to \infty$ when estimating over

$$\bar{\mathcal{A}}_0 := \mathcal{A}_0 \cup \{(i,j) : i = -n, \dots, 0, j = 0, \dots, m-1\}.$$

Moreover, if w_i is an exposure measure such that $\mathbb{E}[N_{ij}] = w_i \nu_{ij}$, then

$$(ii) \qquad \hat{\boldsymbol{\psi}} \xrightarrow{p} \boldsymbol{\psi}$$

as $w_i \to \infty$ for all $i = 1, \dots, m$.

The proof of Proposition 2 is given in Appendix B. From this proposition it is possible to show convergence of parameter estimators that are functions of the $\hat{\psi}_k$s. As an example, it can be shown using Slutsky's theorem that $\hat{\pi}_k \xrightarrow{p} \pi_k$, but this will not be further discussed in this paper.

Note that the interesting part of Proposition 2 is part (ii) since this deals with convergence in terms of an exposure measure while keeping the size of the triangle fixed. Part (i) relies on the claims triangle growing height wise, which is merely a technical affirmation of the estimation procedure being sound.

Regarding the split of future payments into an RBNS and IBNR part, let

$$N_{ij}^{\mathcal{R},(\mathrm{CRM})} := \sum_{k=j-(m-i)}^{j\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}, \ (i,j) \in \mathcal{A}_0^*,$$

corresponding to the total number of RBNS claims payments in (i,j), and let

$$N_{ij}^{\mathcal{I},(\mathrm{CRM})} := \sum_{k=0}^{j-(m-i)-1} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}, \ (i,j) \in \mathcal{A}_0^*,$$

denote the corresponding IBNR claims payments. Thus, the total RBNS claims payments in $(i,j) \in \mathcal{A}_0^*$ is given by

$$X_{ij}^{\mathcal{R},(\mathrm{CRM})} := \sum_{l=0}^{N_{ij}^{\mathcal{R},(\mathrm{CRM})}} Y_{ijl}^{\mathcal{R},(\mathrm{CRM})},\ (i,j) \in \mathcal{A}_0^*,$$

where $Y_{ijl}^{\mathcal{R},(\mathrm{CRM})}$ is the lth payment in (i,j) stemming from RBNS claims. That is, for a fixed (i,j), it is an enumeration of all $\widetilde{Y}_{i,j-k,k,p}$ with $k \geq j-(m-i)$. Similarly, the total IBNR claims payments in $(i,j) \in \mathcal{A}_0^*$ is given by

$$X_{ij}^{\mathcal{I},(\mathrm{CRM})} := \sum_{l=0}^{N_{ij}^{\mathcal{I},(\mathrm{CRM})}} Y_{ijl}^{\mathcal{I},(\mathrm{CRM})},\ (i,j) \in \mathcal{A}_0^*,$$

where again, $Y_{ijl}^{\mathcal{I},(\mathrm{CRM})}$ is the lth payment in (i,j) stemming from IBNR claims. That is, for a fixed (i,j) it is an enumeration of all $\widetilde{Y}_{i,j-k,k,p}$ with $k < j-(m-i)$. From this, we finally get that the total outstanding claims payments stemming from RBNS claims for accident year i is

$$R_i^{\mathcal{R},(\mathrm{CRM})} := \sum_{j=m-i+1}^{m-1+d} X_{ij}^{\mathcal{R},(\mathrm{CRM})}, \tag{6}$$

and the corresponding quantity for IBNR claims for accident year i is

$$R_i^{\mathcal{I},(\mathrm{CRM})} := \sum_{j=m-i+1}^{m-1+d} X_{ij}^{\mathcal{I},(\mathrm{CRM})}. \tag{7}$$

Hence, based on relations (6) and (7) it is clear that the expected claims payments stemming from RBNS and IBNR claims are

$$h_i^{\mathcal{R},(\mathrm{CRM})}(\boldsymbol{\psi};\mathcal{N}_0) := \mathbb{E}\left[R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}_0\right] = \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-m+i}^{j\wedge d} \psi_k N_{i,j-k}, \tag{8}$$

$$h_i^{\mathcal{I},(\mathrm{CRM})}(\boldsymbol{\psi},\boldsymbol{\nu}) := \mathbb{E}\left[R_i^{\mathcal{I},(\mathrm{CRM})} \mid \mathcal{N}_0\right] = \sum_{j=m-i+1}^{m-1+d} \sum_{k=0\vee(j-m+1)}^{(j-m+i-1)\wedge d} \psi_k \nu_{i,j-k}, \tag{9}$$

respectively. Using this we denote the plug-in estimators of the RBNS and IBNR claims reserves by

$$h_i^{\mathcal{R},(\mathrm{CRM})}(\hat{\boldsymbol{\psi}};\mathcal{N}_0)$$

and

$$h_i^{\mathcal{I},(\mathrm{CRM})}(\hat{\boldsymbol{\psi}},\hat{\boldsymbol{\nu}}),$$

i.e. the estimators acquired from simply plugging in the parameter estimators. Note that this is not to be confused with non-parametric plug-in estimation, also referred to as the "plug-in principle", see e.g. Efron and Tibshirani (1994) and Friedman et al. (2001).

Note that we have not defined $\hat{\boldsymbol{\nu}}$. If we assume that $\mathbb{E}[N_{ij}] = \alpha_i\beta_j$, then it would be the vector of the $\hat{\alpha}_i$s and $\hat{\beta}_j$s from applying the chain ladder technique to the triangle of incurred claims.

To end this section we present a proposition giving the variance of the outstanding claims payments, both in total but also the variances stemming from IBNR and RBNS claims separately. This proposition will be useful later when we discuss the semi-analytical approximation of the MSEP, since part of this quantity is the variance of the outstanding claims payments. The proof is given in Appendix B. For later comparison, the analogous results for the VNJ model are presented in Appendix A, (Proposition 11).

Proposition 3. *Given Assumption 1, the variance of the total outstanding claims payments is given by*

$$\begin{aligned}\mathrm{Var}(R^{(CRM)} \mid \mathcal{N}_0) &= \mathrm{Var}\left(\sum_{i=1}^{m} R_i^{\mathcal{R},(CRM)} + R_i^{\mathcal{I},(CRM)} \,\middle|\, \mathcal{N}_0\right) \\ &= \sum_{i=1}^{m} \mathrm{Var}(R_i^{\mathcal{R},(CRM)} \mid \mathcal{N}_0) + \sum_{i=2}^{m} \mathrm{Var}(R_i^{\mathcal{I},(CRM)}),\end{aligned}$$

where

$$\mathrm{Var}(R_i^{\mathcal{R},(CRM)} \mid \mathcal{N}_0) = \sigma^2 \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} \lambda_k N_{i,j-k}, \tag{10}$$

and

$$\mathrm{Var}(R_i^{\mathcal{I},(CRM)}) = (\sigma^2 + \mu^2(1+\Lambda\phi))\Lambda \sum_{j=m-i+1}^{m-1} \nu_{i,j}. \tag{11}$$

3.2 A connection to the double chain ladder method and the chain ladder technique

In this section, we will discuss the double chain ladder (DCL) technique as an estimation method for the CRM. By letting $\mathbb{E}[N_{ij}] = \alpha_i\beta_j$ with $\sum_{j=0}^{m-1} \beta_j = 1$, the CRM fulfils Assumptions M1-M3 in Miranda et al. (2012). Therefore a number of the results in that paper apply to the CRM. Most importantly, it is possible to use the DCL technique as an estimation method. While we have already shown that the GLM-based estimators are consistent, one might ask if the same thing holds for the DCL estimators in terms of an exposure measure. In this section, we will show this to be the case. Moreover, we will show that the plug-in reserve estimator of the CRM, suitably normalised, converges to the same limiting object independent of the estimation method used (GLM or DCL). Further, this limiting object turns out to be the same as the asymptotic normalised reserve of the chain ladder technique, motivating its use as a large-exposure approximation of the CRM.

To summarise, this section aims to motivate:

- The use of the DCL estimators for the CRM.
- The chain ladder technique as a large-exposure approximation of the CRM.

The DCL method uses the chain ladder technique separately on the two triangles of claims payments and claim counts. Equivalently, it fits a Poisson-GLM with multiplicative row and column effects, see Mack (1991). It then solves for the parameters used in prediction by exploiting a relationship between the row and column effects of the two triangles. To show that the CRM fits into this structure we state the following proposition:

Proposition 4. *Assume that*

$$\mathbb{E}[N_{ij}] = \alpha_i\beta_j,$$

with

$$\sum_{j=0}^{d} \beta_j = 1.$$

It then holds that

$$\mathbb{E}[X_{ij}^{(CRM)}] = \tilde{\alpha}_i \tilde{\beta}_j,$$

where

$$\tilde{\alpha}_i := c\tilde{\mu}\alpha_i,$$

and

$$\tilde{\beta}_j := \frac{1}{c} \sum_{k=0}^{j \wedge d} \lambda_k \beta_{j-k}.$$

for some $c \in \mathbb{R}$.

The proof of Proposition 4 is a direct consequence of Corollary 1. Proposition 4 shows that the aggregated incremental payments has multiplicative row and column effects in the CRM. Moreover, it shows how the row and column effects of the triangle of claim counts, α_i and β_j, relate to the corresponding ones for the triangle of claims payments, $\tilde{\alpha}_i$ and $\tilde{\beta}_j$.

We will now collect the row and column effects of the two triangles into two vectors $\boldsymbol{\theta}$ and $\tilde{\boldsymbol{\theta}}$, with corresponding chain ladder technique estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\tilde{\boldsymbol{\theta}}}$. For these estimators it is possible to show the following proposition dealing with consistency in terms of exposure:

Proposition 5. *Assuming that*

$$\begin{aligned} \mathbb{E}[N_{ij}] &= w_i \bar{\alpha}_i \beta_j, \\ \mathrm{Var}(N_{ij}) &= \phi \mathbb{E}[N_{ij}], \end{aligned}$$

it holds that

$$\begin{aligned} \hat{\boldsymbol{\theta}} &\xrightarrow{p} \boldsymbol{\theta}, \\ \hat{\tilde{\boldsymbol{\theta}}} &\xrightarrow{p} \tilde{\boldsymbol{\theta}}, \end{aligned}$$

as $w_i \to \infty$ *for all* $i = 1, \ldots, m$.

The proof of Proposition 5 follows from a slight, but straightforward, modification of the proof of Proposition 2 part (ii).

Assuming that inflation is already accounted for, i.e. γ_i from Miranda et al. (2012), there are two more parameters defining the DCL method, $(\pi_0^{\mathrm{D}}, \ldots \pi_{m-1}^{\mathrm{D}})$ and μ^{D}, here with the superscript D to distinguish them from the already defined parameters π_k and μ. These are estimated by equations (7) and and (9) in Miranda et al. (2012), i.e. by solving for the π_ks in

$$\begin{pmatrix} \hat{\tilde{\beta}}_0 \\ \vdots \\ \vdots \\ \hat{\tilde{\beta}}_{m-1} \end{pmatrix} = \begin{pmatrix} \hat{\beta}_0 & 0 & \ldots & 0 \\ \hat{\beta}_1 & \hat{\beta}_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \hat{\beta}_0 & \ldots & \hat{\beta}_1 & \hat{\beta}_0 \end{pmatrix} \begin{pmatrix} \pi_0 \\ \vdots \\ \vdots \\ \pi_{m-1} \end{pmatrix},$$

yielding estimators $\hat{\pi}_k^{\mathrm{D}}$, and by calculating

$$\hat{\mu}^{\mathrm{D}} = \frac{\hat{\tilde{\alpha}}_1}{\hat{\alpha}_1}.$$

Using these estimators it is possible to construct estimators of the ψ_ks by

$$\hat{\psi}_k^{\mathrm{D}} := \hat{\mu}^{\mathrm{D}} \hat{\pi}_k^{\mathrm{D}}.$$

Now the following proposition, dealing with consistency of the DCL estimators of the ψ_ks, can be stated:

Proposition 6. *For the DCL estimators $\hat{\pi}_k^D$ and $\hat{\mu}^D$ it holds for $k = 1, \ldots, m-1$ that*

$$\hat{\pi}_k^D \xrightarrow{p} \frac{1}{c}\lambda_k 1_{\{k \le d\}},$$
$$\hat{\mu}^D \xrightarrow{p} c\mu,$$

as $w_i \to \infty$ for some $c \in \mathbb{R}$. From which it is clear that

$$\hat{\psi}_k^D := \mu^D \hat{\pi}_k^D \xrightarrow{p} \mu\lambda_k 1_{\{k \le d\}} =: \psi_k,$$

for $k = 1, \ldots, m-1$.

The proof of Proposition 6 can be found in Appendix B. Proposition 6 states that the DCL estimators $\hat{\psi}_k^{\mathrm{D}}$ of ψ_k are consistent in terms of exposure. Therefore, the DCL method yields consistent estimators of the parameters ψ_k in the CRM, motivating its use.

Now that we know that the DCL estimation procedure yields consistent estimators, we move on to show that the reserve estimates acquired using plug-in of either the GLM or the DCL-based estimators are asymptotically close. And further, that they are asymptotically close to the reserve estimator of the chain ladder technique.

In Miranda et al. (2012) it is shown that using the DCL estimators and replacing N_{ij} in (8) by $\hat{\alpha}_i \hat{\beta}_j$, yields the chain ladder reserve estimator based on only the triangle of claims payments. For the CRM, this is equivalent to basing prediction on $\mathbb{E}[X_{ij}]$ and not on $\mathbb{E}[X_{ij} | \mathcal{N}_0]$. Note that for the IBNR part this makes no difference and we only have to investigate the RBNS part. The following proposition shows that these two methods are equivalent in the limit as the exposure tends to infinity:

Proposition 7. *Let $h_i^{\mathcal{R},(CRM)}(\boldsymbol{\psi}; \mathcal{N}_0)$ be defined according to (8) and assume that $\mathbb{E}[N_{ij}] = w_i \bar{\alpha}_i \beta_j$, where w_i is an exposure measure. For any $\delta > 0$ it holds that*

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{w_i}\left|\mathbb{E}[R_i^{\mathcal{R},(CRM)} | \mathcal{N}_0] - \mathbb{E}[R_i^{\mathcal{R},(CRM)}]\right| > \delta\right) &= \mathbb{P}\left(\left|\frac{1}{w_i} h_i^{\mathcal{R},(CRM)}(\boldsymbol{\psi}; \mathcal{N}_0) - \bar{r}_i\right| > \delta\right) \\
&\le \frac{\bar{v}_i}{\delta^2 w_i} \to 0,
\end{aligned}$$

as $w_i \to \infty$, where

$$\bar{r}_i := \frac{1}{w_i}\mathbb{E}[h_i^{\mathcal{R},(CRM)}(\boldsymbol{\psi}; \mathcal{N}_0)] = \mu \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j \wedge d} \lambda_k \bar{\alpha}_i \beta_{j-k},$$

$$\bar{v}_i := \frac{1}{w_i}\operatorname{Var}(h_i^{\mathcal{R},(CRM)}(\boldsymbol{\psi}; \mathcal{N}_0)) = \phi\mu^2 \sum_{j=0}^{m-i} \left(\sum_{k=m-i+1-j}^{d \wedge (m-1+d-j)} \lambda_k\right)^2 \bar{\alpha}_i \beta_j.$$

The proof of Proposition 7 is given in Appendix B, from which it is clear that the same result follows verbatim for the VNJ model. Proposition 7 states that

$$\frac{1}{w_i} h_i^{\mathcal{R},(\mathrm{CRM})}(\boldsymbol{\theta}; \mathcal{N}_0) \xrightarrow{p} \bar{r}_i,$$

as $w_i \to \infty$, i.e. that the normalised conditional expectation of the outstanding claims payments stemming from RBNS claims converges to the normalised unconditional ones. Since the corresponding IBNR quantities are the same for finite exposure this implies that the conditional expectation of the total outstanding claims payments converges to the unconditional one.

Remark 2. *Note that Proposition 7 gives convergence when we consider $X_{ij}^{\mathcal{R},(CRM)}$ with $(i,j) \in \mathcal{A}_0^*$ and not $\tilde{\mathcal{A}}_0^*$, i.e. in the complete lower right trapezoid (including the tail) and not just the regular lower right triangle (without the tail), but it is straightforward to modify the proof to this case.*

In practice, it is necessary to estimate the parameters and use these when performing prediction. Since $\bar{r}_i := \bar{r}_i(\boldsymbol{\psi};\boldsymbol{\theta})$ the plug-in estimator of $\bar{r}_i$ is given by $\bar{r}_i(\hat{\boldsymbol{\psi}};\hat{\boldsymbol{\theta}})$. An application of Slutsky's theorem together with Proposition 7 yields the following corollary:

Corollary 2. *The following convergences hold as $w_i \to \infty$:*

$$
\begin{aligned}
\frac{1}{w_i} h_i^{\mathcal{R},(CRM)}(\hat{\boldsymbol{\psi}};\mathcal{N}_0) &\xrightarrow{p} \bar{r}_i, \\
\frac{1}{w_i} h_i^{\mathcal{R},(CRM)}(\hat{\boldsymbol{\psi}}^D;\mathcal{N}_0) &\xrightarrow{p} \bar{r}_i, \\
\bar{r}_i(\hat{\boldsymbol{\psi}};\hat{\boldsymbol{\theta}}) &\xrightarrow{p} \bar{r}_i, \\
\bar{r}_i(\hat{\boldsymbol{\psi}}^D;\hat{\boldsymbol{\theta}}) &\xrightarrow{p} \bar{r}_i.
\end{aligned}
$$

Corollary 2 states that plug-in prediction of the outstanding claims payments from RBNS claims, based on either conditional or unconditional expectations, and either GLM or DCL-based estimators, converge to the same limiting object. Using $\bar{r}_i$ with plug-in of the DCL estimators $\hat{\boldsymbol{\psi}}^{\mathrm{D}}$ and $\hat{\boldsymbol{\theta}}$ together with the corresponding IBNR part, both restricted to the lower right triangle without the tail $\tilde{\mathcal{A}}_0^*$ (see Remark 2) yields the reserve estimator of the chain ladder technique, see Miranda et al. (2012). It is thus clear that the chain ladder technique can be motivated as a large-exposure approximation of the CRM without the tail.

4 Comparison of the CRM and the VNJ model

As pointed out in Section 3, the VNJ model assumes that a claim can only generate a single payment, whereas the CRM allows for multiple payments per claim. We will now relate the two approaches by making a detailed comparison of the models. The main purpose of this paper is to set out the theory underlying the CRM, but it is also of interest to see the modelling consequences of using the simplified single payment model. In Section 4.1 we relate the two models using a theoretical construct and investigate the relationship between model variances. Section 4.2 is concerned with comparing estimated variances based on plug-in estimation.

4.1 Comparison in terms of model dynamics

In this section, we will introduce a parametrisation of the VNJ model corresponding to a series of payments being exchanged for a single payment. This single payment will have the same first two moments as those of the underlying series of payments coming from the CRM. This is a straightforward way of connecting the two models, which will allow for comparison of model dynamics. Note, however, that the two models are fundamentally different. Given that a claim has already given rise to one payment, the CRM assumes that there are possible future payments that are unknown today. In the VNJ model, however, these are considered known since they are part of the already occurred payment.

The parametrisation introduced below will guarantee that the expected outstanding claims payments are the same between the two models, but the variances will naturally be different. We will now investigate the effect of approximating the CRM by the VNJ model in the above described manner on the ordering of the variances.

To make the connection between the models more precise, let N_{ij}^{tot} denote a random variable which is distributed in the same way as the number of payments of a single claim incurred in (i, j) given the payment dynamics of the CRM, i.e.

$$N_{ij}^{\text{tot}} \sim \text{Po}\left(\Lambda\right).$$

Now, define the payment distribution of the VNJ model in the following way:

$$Y_{ijk}^{(\text{VNJ})} \stackrel{d}{=} \sum_{l=0}^{N_{ij}^{\text{tot}}} Y_{ijl}^{(\text{CRM})} \tag{12}$$

where $Y_{ijl}^{(\text{CRM})}$ is the lth payment in (i, j) in the CRM. That is, for a fixed (i, j) it is an enumeration of all $\widetilde{Y}_{i,j-k,k,p}$. Moreover, let the payments in the VNJ model be distributed according to the intensities of the Poisson payment process of the CRM, i.e.

$$p_k := \frac{\lambda_k}{\Lambda} = \pi_k. \tag{13}$$

That is, by this definition, $Y_{ijk}^{(\text{VNJ})}$ corresponds, in distribution, to the replacement of a series of payments with a single aggregated payment, which is distributed in time according to probabilities proportional to the Poisson-intensities. For the VNJ model, parametrised in this way, it holds that

$$\mathbb{E}[Y_{ijk}^{(\text{VNJ})}] = \mu\Lambda = \tilde{\mu}, \tag{14}$$

and

$$\text{Var}(Y_{ijk}^{(\text{VNJ})}) = (\mu^2 + \sigma^2)\Lambda = \tilde{\sigma}^2. \tag{15}$$

Note that the meanings of π_k and p_k differ: The π_ks describe how payments, generated by a claim in the CRM, are distributed w.r.t. time since reporting, whereas the p_ks describe the distribution of when in time a single total payment is paid w.r.t. time since reporting. That is, we replace a Poisson number of smaller payments with a single larger payment. This single larger payment is paid k periods since reporting with a probability proportional to the kth Poisson intensity.

Given the above definition we can now state the following result relating the two models:

Proposition 8. *For $X_{ij}^{(VNJ)}$ defined in terms of $Y_{ijk}^{(VNJ)}$ in* (12) *such that* (13), (14) *and* (15) *are fulfilled, it holds that*

$$\mathbb{E}[X_{ij}^{(VNJ)}|\mathcal{N}_0] = \mathbb{E}[X_{ij}^{(CRM)}|\mathcal{N}_0],$$

and

$$\text{Var}(X_{ij}^{(VNJ)}|\mathcal{N}_0) \geq \text{Var}(X_{ij}^{(CRM)}|\mathcal{N}_0).$$

Further, it holds that

$$\mathrm{Var}(R_i^{\mathcal{R},(VNJ)} \mid \mathcal{N}_0) \geq \mathrm{Var}(R_i^{\mathcal{R},(CRM)} \mid \mathcal{N}_0),$$

and

$$\mathrm{Var}(R_i^{\mathcal{I},(VNJ)} \mid \mathcal{N}_0) = \mathrm{Var}(R_i^{\mathcal{I},(CRM)} \mid \mathcal{N}_0),$$

i.e.

$$\mathrm{Var}(R^{(VNJ)} \mid \mathcal{N}_0) \geq \mathrm{Var}(R^{(CRM)} \mid \mathcal{N}_0).$$

The proof of Proposition 8 is a direct consequence of the parametrization (13)-(15) together with simple algebraic manipulations of the expectations and variances found in (5), (23), and in Proposition 3 and 11. Proposition 8 states that the variances of the outstanding payments from IBNR claims are equal for the two models. This is no surprise, since in the IBNR case we sum over the whole future of a claim that is currently unreported. In this case it does not matter how in time the payments are distributed, only that the moments of the total payments are equal. For the RBNS part, the variance is always greater in the VNJ model. Part of the reason for this is that there is more of an "all or nothing" situation in the VNJ model. That is, since we are not conditioning on whether a payment has occurred or not, it is possible that a payment has already occurred for a claim, and therefore there will be 0 outstanding payments coming from that claim. Comparatively, in the CRM there is more of a steady stream of smaller payments.

The above comparison is based on a theoretical construct. The models are fundamentally different and exist on completely different probability spaces. In the next section we look at a comparison that is more practically relevant. We look at the ordering of the estimated variances based on plug-in of the parameter estimates.

4.2 Comparison in terms of the effect of parameter estimation

In Section 4.1 we discussed the effect of replacing a series of payments with a single payment whose first two moments are equal to those of the underlying series of payments. The conclusion stated in Proposition 8 is that the process variance is larger in the approximating single payment model, the VNJ model, compared with the multiple payments model, the CRM. In this section we discuss the effects of fitting the two different models to the same data, regardless of what process has generated it, and how this affects the size of the estimated variances. In Section 3, when parameter estimation was discussed, it was noted that both the CRM and the VNJ model will use the same $\hat{\psi}_k$s, see Remark 1. Therefore, both models will produce the same reserve estimates. This, however, does not imply that the variances are related according to the theoretical construct of Section 4.1. By using the parameter estimates from Section 3, it is possible to show the following:

Proposition 9. *Let* $\widehat{\mathrm{Var}}_{(i)}(\cdot|\mathcal{N}_0) = \mathrm{Var}(\cdot|\mathcal{N}_0)(\hat{\boldsymbol{\theta}}^{(i)})$ *denote the plug-in variance estimator using the parameter estimators of model i,* $i \in \{CRM,\ VNJ\}$*. It then holds that*

$$\widehat{\mathrm{Var}}_{(VNJ)}(X_{ij}^{(VNJ)} \mid \mathcal{N}_0) \leq \widehat{\mathrm{Var}}_{(CRM)}(X_{ij}^{(CRM)} \mid \mathcal{N}_0),$$

and

$$\widehat{\mathrm{Var}}_{(VNJ)}(R_i^{\mathcal{I},(VNJ)} \mid \mathcal{N}_0) \leq \widehat{\mathrm{Var}}_{(CRM)}(R_i^{\mathcal{I},(CRM)} \mid \mathcal{N}_0).$$

Moreover,

$$\widehat{\mathrm{Var}}_{(VNJ)}(R_i^{\mathcal{R},(VNJ)} \mid \mathcal{N}_0) \leq \widehat{\mathrm{Var}}_{(CRM)}(R_i^{\mathcal{R},(CRM)} \mid \mathcal{N}_0),$$

if and only if

$$\hat{\Lambda} \geq \frac{\sum_{j=0}^{m-i} N_{ij}\hat{q}_j}{\sum_{j=0}^{m-i} N_{ij}\hat{q}_j^2} =: \hat{\Lambda}_i^{Bound,1}, \tag{16}$$

and

$$\widehat{\mathrm{Var}}_{(VNJ)}(R_i^{(VNJ)} \mid \mathcal{N}_0) \leq \widehat{\mathrm{Var}}_{(CRM)}(R_i^{(CRM)} \mid \mathcal{N}_0), \tag{17}$$

if and only if

$$\hat{\Lambda} \geq \frac{\sum_{j=0}^{m-i} N_{ij}\hat{q}_j}{\sum_{j=0}^{m-i} N_{ij}\hat{q}_j^2 + \sum_{j=m-i+1}^{m-1} \hat{\nu}_{ij}} =: \hat{\Lambda}_i^{Bound,2}, \tag{18}$$

where

$$\hat{q}_j = \sum_{k=m-i+1-j}^{d\wedge(m-1+d-j)} \hat{p}_k. \tag{19}$$

The proof of Proposition 9 can be found in Appendix B. The first thing we can conclude from Proposition 9 is that we will always estimate the variance of a single cell to be larger in the CRM. Moreover, the same conclusion follows for the total (per accident year) outstanding payments stemming from IBNR claims. The RBNS part is not as straightforward, and the ordering depends on the estimated payment delay distribution in a complex way. The variance of the total outstanding claims payments per accident year is even less straightforward to interpret since it depends on the estimated values of the ν_{ij}s.

We can, however, note that $\hat{\Lambda}_i^{\text{Bound},1} \geq 1$ since $0 \leq \hat{q}_j \leq 1$. Therefore, if the estimated expected number of payments per claim is smaller than 1, then the RBNS variance of the CRM is guaranteed to be smaller than that of the VNJ model.

We will refrain from going further into interpreting Proposition 9 and simply conclude that the bounds $\hat{\Lambda}_i^{\text{Bound},1}$ and $\hat{\Lambda}_i^{\text{Bound},2}$ depend on estimated parameters, making it difficult to say beforehand what ordering will hold. Moreover, the ordering might be different for different accident years. In Section 6 we illustrate the bounds in a real world example based on the data used in Verrall et al. (2010).

As a final note, in Section 5 we will see that the ordering of the estimated variances of the total outstanding claims payments also applies to the ordering of the approximations of the MSEP according to Lindholm et al. (2018), see Corollary 3.

5 Assessment of prediction uncertainty

To assess prediction uncertainty, including estimation error, Miranda et al. (2011) described how to bootstrap the VNJ model, and this can also be applied to the CRM by exchanging the multinomial distribution for a Poisson distribution. Since there is not much to add to this method of estimating the predictive uncertainty, we will discuss another method based on semi-analytical approximations of the conditional MSEP.

Let us define the conditional MSEP. Given a random variable X, a filtration $\mathcal{F}_0$ and an $\mathcal{F}_0$-measurable predictor of X, $\hat{X}$, we define the conditional MSEP as follows:

$$\begin{aligned}\mathrm{MSEP}(X, \hat{X}) &= \mathbb{E}[(X - \hat{X})^2 | \mathcal{F}_0] \\ &= \mathrm{Var}(X|\mathcal{F}_0) + (E[X|\mathcal{F}_0] - \hat{X})^2.\end{aligned}$$

With a bootstrap approach in mind, it is possible to consider semi-analytical approximations following, for example, Lindholm et al. (2018), which gives an estimator of the mean squared error of prediction (MSEP). See also Diers et al. (2016), Röhr (2016) and Buchwalder et al. (2006) for similar approaches, particularly applied to the distribution-free chain ladder model. In Lindholm et al. (2018) the following approximation of the conditional MSEP is provided:

$$\widehat{\mathrm{MSEP}}(X, \hat{X}) = \widehat{\mathrm{Var}}(X \mid \mathcal{F}_0) + \nabla\mathbb{E}[X|\mathcal{F}_0](\hat{\boldsymbol{\theta}}; \mathcal{F}_0)'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})\nabla\mathbb{E}[X|\mathcal{F}_0](\hat{\boldsymbol{\theta}}; \mathcal{F}_0),$$

where the gradient of the expectation is taken with respect to the parameter vector $\boldsymbol{\theta}$ and the hats on the variance and covariance denote plug-in estimation. For the CRM this translates to

$$\begin{aligned}\widehat{\mathrm{MSEP}}(R^{(\mathrm{CRM})}, \hat{R}^{(\mathrm{CRM})}) =& \widehat{\mathrm{Var}}(R^{(\mathrm{CRM})} \mid \mathcal{N}_0) \\ &+ \nabla h^{(\mathrm{CRM})}(\hat{\boldsymbol{\theta}}; \mathcal{N}_0)'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})\nabla h^{(\mathrm{CRM})}(\hat{\boldsymbol{\theta}}; \mathcal{N}_0),\end{aligned} \tag{20}$$

where

$$h^{(\mathrm{CRM})}(\boldsymbol{\theta}; \mathcal{N}_0) = \mathbb{E}[R^{\mathcal{R},(\mathrm{CRM})} + R^{\mathcal{I},(\mathrm{CRM})} \mid \mathcal{N}_0],$$

and $\hat{\boldsymbol{\theta}}$ is a vector containing the parameter estimators $\hat{\boldsymbol{\psi}}$ and $\hat{\boldsymbol{\nu}}$. The approximation is based on a re-sampling/bootstrap argument together with a Taylor approximation. Moreover, the result relies on (i) that $\mathbb{E}[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}$ (unbiased estimators), (ii) that it is possible to calculate the variance and (iii) that it is possible to calculate $\mathrm{Cov}(\hat{\boldsymbol{\theta}})$. For both the CRM and the VNJ model (ii) is possible, see Proposition 3 and 11. Concerning (i), if the estimator is biased, then there is a need for the following bias correction term:

$$\mathrm{Bias}(\hat{\boldsymbol{\theta}})'\nabla h^{(\mathrm{CRM})}(\hat{\boldsymbol{\theta}}; \mathcal{N}_0)\nabla h^{(\mathrm{CRM})}(\hat{\boldsymbol{\theta}}; \mathcal{N}_0)'\mathrm{Bias}(\hat{\boldsymbol{\theta}}),$$

see Lindholm et al. (2018). Now, it is the case that the maximum quasi-likelihood estimator will (likely) be biased. Since we cannot quantify this bias, and since the estimators are consistent in terms of exposure (and therefore asymptotically unbiased), we can approximate the MSEP by disregarding this, likely, small contribution.

We still need to be able to calculate the covariance matrix of the estimator. This is not an easy task since we are optimizing a quasi-likelihood without a closed form solution. A solution to this would be to approximate the covariance matrix by the Hessian of the quasi-likelihood, evaluated at the maximum quasi-likelihood estimate. The rationale being that this is a consistent estimator of the expected Fisher information, which in a conventional maximum likelihood setting is the asymptotic variance. Here, this is not the case, but it will serve as an approximation.

We stress that, since we are already performing a GLM fitting on the incurred claims triangle and on the aggregated claims payments triangle, we would have easy access to the Hessians and therefore this approximation of the conditional MSEP is easily calculated.

A final note is that the Hessian of the quasi-likelihood, evaluated at the maximum quasi-likelihood estimate, will be the same for the CRM and the VNJ model. Therefore, the approximated MSEP will differ between them only as much as the variances differ. The following corollary to Proposition 9 states this:

Corollary 3. *Let $\hat{q}_j$ be as in* (19)*, then*

$$\widehat{\mathrm{MSEP}}(R^{(VNJ)}, \hat{R}^{(VNJ)}) \leq \widehat{\mathrm{MSEP}}(R^{(CRM)}, \hat{R}^{(CRM)}),$$

if and only if (18) *holds.*

It should be noted though that this is in terms of the approximation of the MSEP. The actual MSEP will likely differ in more than the variance.

We end this section by including a proposition giving all the necessary gradients needed to calculate (20). The proof is a simple exercise in taking derivatives and keeping track of indices and is therefore left out.

Proposition 10. *The gradients needed to calculate* (20) *are:*

$$\frac{\partial}{\partial \alpha_l} h_i^{\mathcal{I},(CRM)}(\theta;\mathcal{N}_0) = 1_{\{l=i\}} \sum_{j=m-l+1}^{m-1+d} \sum_{k=0\vee(j-m+1)}^{(j-m+l-1)\wedge d} \psi_k \beta_{j-k},$$

$$\frac{\partial}{\partial \beta_l} h_i^{\mathcal{I},(CRM)}(\theta;\mathcal{N}_0) = 1_{\{m-i+1\leq l\}} \alpha_i \sum_{k=(m-i+1-l)\vee 0}^{(m-1+d-l)\wedge d} \psi_k,$$

$$\frac{\partial}{\partial \psi_l} h_i^{\mathcal{I},(CRM)}(\theta;\mathcal{N}_0) = \sum_{k=m-i+1}^{m-1} \nu_{ik},$$

$$\frac{\partial}{\partial \psi_l} h_i^{\mathcal{R},(CRM)}(\theta;\mathcal{N}_0) = \sum_{k=(m-i+1-l)\vee 0}^{m-i} N_{ik}.$$

6 Numerical example

The main contribution of this paper is the theoretical underpinning. It is however useful to have an illustration to see how the CRM compares in a benchmarking situation. In this section we will look at a numerical example to illustrate how the CRM compares to the VNJ model and to other benchmarking models, such as the distribution-free chain ladder model. Moreover, some results from this paper will be illustrated in this numerical example.

We will make use of the same data as Verrall et al. (2010) and Miranda et al. (2011, 2012) consisting of a portfolio of motor third party liability policies from the general insurer Royal & Sun Alliance. The data are in two triangles consisting of aggregate claims payments and total incurred claims as specified in Section 2. The size of the triangles is $m = 10$ and they can be seen in Tables 1 and 2. Since we do not have access to the number of payments made we cannot separate Λ from μ in the CRM when doing estimation. Therefore we will have to make an assumption about the size of Λ. For simplicity we choose $\Lambda = 1$. Moreover, we have to choose the number of delay parameters. We choose $d = 7$ to keep in line with the numerical example of Verrall et al. (2010).

In Table 3 we show the point predictions of the RBNS, IBNR and the total outstanding claims payments according to (8) and (9) for the CRM and the VNJ model. For comparison we have included the predictions from the chain ladder technique. The predictions are split by accident year with the total at the bottom. To illustrate how the tail affects the result we have included the predictions from the CRM restricted to $\tilde{\mathcal{A}}_0^*$, i.e. the lower right triangle excluding the tail. We see that, in this particular case, there is little difference between doing this and including the tail. Further, there is not a huge difference between the prediction from the chain ladder technique and the predictions from the CRM and the VNJ model, which is in line with the conclusions of Section 3.2 that the chain ladder technique is a large-exposure approximation of the CRM.

In Table 4, RBNS, IBNR and total standard deviations are shown. For comparison we have included the corresponding quantities for the distribution-free chain ladder model (CLM) and an ODP model with multiplicative row and column effects. Looking at the total, it can

Table 1: Run-off triangle of aggregated payments, X_{ij}.

$i \backslash j$	0	1	2	3	4	5	6	7	8	9
1	451,288	339,519	333,371	144,988	93,243	45,511	25,217	20,406	31,482	1,729
2	448,627	512,882	168,467	130,674	560,44	33,397	56,071	26,522	14,346	
3	693,574	497,737	202,272	120,753	125,046	37,154	27,608	17,864		
4	652,043	546,406	244,474	200,896	106,802	106,753	63,688			
5	566,082	503,970	217,838	145,181	165,519	91,313				
6	606,606	562,543	227,374	153,551	132,743					
7	536,976	472,525	154,205	150,564						
8	554,833	590,880	300,964							
9	537,238	701,111								
10	684,944									

Table 2: Run-off triangle of number of reported claims, N_{ij}.

$i \backslash j$	0	1	2	3	4	5	6	7	8	9
1	6,238	831	49	7	1	1	2	1	2	3
2	7,773	1,381	23	4	1	3	1	1	3	
3	10,306	1,093	17	5	2	0	2	2		
4	9,639	995	17	6	1	5	4			
5	9,511	1,386	39	4	6	5				
6	10,023	1,342	31	16	9					
7	9,834	1,424	59	24						
8	10,899	1,503	84							
9	11,954	1,704								
10	10,989									

Table 3: The reserve estimates split into IBNR and RBNS claims together with the total reserve estimate for each accident year $i = 1, \ldots, 10$. Moreover, the total reserve estimate based on the lower right triangle $\tilde{\mathcal{A}}_0^*$ is added to illustrate the effect of the tail. The reserve estimates from the chain ladder technique are added as a benchmark.

i	IBNR	RBNS	Total	Cut Total	Chain Ladder
1	0	556	556	0	0
2	628	605	1,233	539	1,685
3	1,350	4,514	5,863	5,010	29,379
4	1,510	43,623	45,133	44,231	60,638
5	1,967	94,526	96,493	95,575	101,158
6	2,579	171,633	174,212	173,217	173,802
7	3,168	299,136	302,304	301,327	249,349
8	5,349	509,334	514,684	513,662	475,992
9	14,280	852,144	866,424	865,301	763,919
10	254,499	1,135,678	1,390,177	1,389,152	1,459,860
Total	285,329	3,111,750	3,397,079	3,388,014	3,315,779

be seen that the CRM and the VNJ model have similar estimates, while the ODP model has a smaller variance and the CLM has a larger variance.

Table 4: The row-wise conditional standard deviations of the total outstanding claims payments stemming from IBNR and RBNS claims together with their total for the CRM and the VNJ model. The corresponding quantities from the chain ladder model (CLM) and an ODP model with multiplicative row and column effects are added as benchmarks.

i	CRM^{IBNR}	CRM^{RBNS}	$\text{CRM}^{\text{Total}}$	VNJ^{IBNR}	VNJ^{RBNS}	$\text{VNJ}^{\text{Total}}$	ODP	CLM
1	0	3,099	3,099	0	3,107	3,107	0	0
2	3,464	3,232	4,738	3,449	3,241	4,733	4,250	4,848
3	5,078	8,827	10,184	5,057	8,866	10,207	17,748	14,047
4	5,372	27,441	27,962	5,349	27,565	28,080	25,498	16,643
5	6,131	40,394	40,857	6,105	40,572	41,029	32,933	25,708
6	7,020	54,431	54,882	6,990	54,661	55,106	43,168	40,761
7	7,781	71,859	72,279	7,748	72,138	72,553	51,706	51,324
8	10,111	93,766	94,310	10,068	94,092	94,629	71,439	63,787
9	16,520	121,283	122,403	16,449	121,621	122,728	90,502	135,237
10	69,740	140,014	156,422	69,443	140,254	156,504	125,110	234,965
Total	73,843	231,765	243,244	73,529	232,406	243,760	188,550	288,133

The estimated payment parameters, based on the assumption that $\hat{\Lambda} = 1$, are $\hat{\mu} = 162.41$ and $\hat{\sigma}^2 =$ 2,803,491. Further, Table 5 shows the estimates $(\hat{\pi}_0, \dots, \hat{\pi}_7)$. Based on these, and the $\hat{\nu}_{ij}$s, we can calculate the ratios in (16) and (18), corresponding to the bounds of $\hat{\Lambda}$ for which the estimated RBNS and total variances change ordering, respectively. These can be seen in Table 6. Concerning the RBNS variances, the first bounds are quite a bit greater than the assumed $\hat{\Lambda} = 1$. For half of the accident years to have a greater estimated RBNS variance in the CRM, compared to the VNJ model, we would need $\hat{\Lambda} \approx 4$. With the assumed $\hat{\Lambda} = 1$, the only accident year where the estimated total variance is larger for the CRM than for the VNJ model is year 2. As Proposition 9 states, the estimated IBNR variances are always greater for the CRM.

Table 5: Quasi-maximum likelihood estimates of the parameters π_k for $k = 0, \dots, d$.

j	0	1	2	3	4	5	6	7
$\hat{\pi}_j$	0.36	0.29	0.11	0.09	0.07	0.04	0.03	0.02

Table 6: Ratios $\hat{\Lambda}_{\text{Bound},1}$ and $\hat{\Lambda}_{\text{Bound},2}$ as defined by (16) and (18), respectively.

i	$\hat{\Lambda}_{\text{Bound},1}$	$\hat{\Lambda}_{\text{Bound},2}$
1	2.18	2.18
2	2.51	0.70
3	16.24	2.77
4	28.08	14.24
5	16.53	12.30
6	9.85	8.58
7	5.84	5.50
8	3.84	3.69
9	2.45	2.36
10	1.57	1.16

Finally, we compare the MSEP of the CRM, the VNJ model and the distribution-free chain ladder model given in Mack (1993). These can be seen in Table 7. The MSEP is smaller for the CRM and the VNJ model than for the chain ladder model. This stems partly from the variances being smaller for those models, but it also indicates that the estimation error is of reasonable size. One could have imagined that the number of parameters necessary to estimate in the CRM and the VNJ model, would have led to a large estimation error.

Table 7: Approximated root MSEP for the CRM and the VNJ model with the MSEP for the Chain Ladder model according to Mack (1993) as a reference.

CRM	VNJ	Mack CL
337,742	338,114	351,784

7 Conclusions

In this paper, we have introduced a model that generalises the model presented in Verrall et al. (2010). This model has the advantage that it allows for multiple payments per claim and zero claims. Since predictors of outstanding payments are the same and model fitting can be done in the same way, this model allows the use of results from Verrall et al. (2010) while feeling confident that the common case of multiple payments is still taken into account. In addition, we have provided variance expressions for total outstanding payments stemming from both IBNR and RBNS claims. This together with the semi-analytical approximation of the conditional MSEP given in Section 5 provides the possibility to conveniently assess predictive uncertainty.

Another interesting result presented in this paper is that the double chain ladder method gives consistent estimators, motivating its use. This together with the fact that it is easily implemented, speaks in its favour. Moreover, as was stated in the introduction, many articles have sought to motivate the chain ladder technique. In this paper we have provided a result stating that the chain ladder technique reserve estimate is asymptotically the same as that of the model of Verrall et al. (2010) and the model introduced in the present paper, restricted to the lower right triangle, regardless of estimation method used (GLM or DCL). This motivates the chain ladder technique as a large-exposure approximation of the model of Verrall et al. (2010) and the model of the present paper.

References

Antonio, K. and Plat, R. (2014). Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, 2014(7):649–669.

Björkwall, S., Hössjer, O., and Ohlsson, E. (2009). Non-parametric and parametric bootstrap techniques for age-to-age development factor methods in stochastic claims reserving. *Scandinavian Actuarial Journal*, 2009(4):306–331.

Buchwalder, M., Bühlmann, H., Merz, M., and Wüthrich, M. V. (2006). The mean square error of prediction in the chain ladder reserving method (mack and murphy revisited). *ASTIN Bulletin: The Journal of the IAA*, 36(2):521–542.

Bühlmann, H., Schnieper, R., and Straub, E. (1980). Claims reserves in casualty insurance based on a probabilistic model. *Bulletin of the Swiss Association of Actuaries*, 80:21–45.

Diers, D., Linde, M., and Hahn, L. (2016). Addendum to 'the multi-year non-life insurance risk in the additive reserving model'[insurance math. econom. 52 (3)(2013) 590–598]: Quantification of multi-year non-life insurance risk in chain ladder reserving models. *Insurance: Mathematics and Economics*, 67:187–199.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

England, P. and Verrall, R. (1999). Analytic and bootstrap estimates of prediction errors in claims reserving. *Insurance: mathematics and economics*, 25(3):281–293.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The elements of statistical learning*, volume 1. Springer series in statistics New York, NY, USA:.

Lehmann, E. L. and Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.

Lindholm, M., Lindskog, F., and Wahl, F. (2018). Estimation of conditional mean squared error of prediction for claims triangle reserving. *Unpublished manuscript*.

Mack, T. (1991). A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin: The Journal of the IAA*, 21(1):93–109.

Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin: The Journal of the IAA*, 23(2):213–225.

Mikosch, T. (2009). *Non-life insurance mathematics: an introduction with the Poisson process*. Springer Science & Business Media.

Miranda, M. D. M., Nielsen, B., Nielsen, J. P., and Verrall, R. (2011). Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers. *ASTIN Bulletin: The Journal of the IAA*, 41(1):107–129.

Miranda, M. D. M., Nielsen, J. P., and Verrall, R. (2012). Double chain ladder. *ASTIN Bulletin: The Journal of the IAA*, 42(1):59–76.

Norberg, R. (1986). A contribution to modelling of IBNR claims. *Scandinavian Actuarial Journal*, 1986(3-4):155–203.

Norberg, R. (1993). Prediction of outstanding liabilities in non-life insurance 1. *ASTIN Bulletin: The Journal of the IAA*, 23(1):95–115.

Norberg, R. (1999). Prediction of outstanding liabilities ii. model variations and extensions. *ASTIN Bulletin: The Journal of the IAA*, 29(1):5–25.

Pinheiro, P. J., Andrade e Silva, J. M., and de Lourdes Centeno, M. (2003). Bootstrap methodology in claim reserving. *Journal of Risk and Insurance*, 70(4):701–714.

Renshaw, A. E. and Verrall, R. J. (1998). A stochastic model underlying the chain-ladder technique. *British Actuarial Journal*, 4(4):903–923.

Röhr, A. (2016). Chain ladder and error propagation. *ASTIN Bulletin: The Journal of the IAA*, 46(2):293–330.

Schnieper, R. (1991). Separating true IBNR and IBNER claims 1. *ASTIN Bulletin: The Journal of the IAA*, 21(1):111–127.

Schweder, T. and Hjort, N. L. (2016). *Confidence, likelihood, probability*, volume 41. Cambridge University Press.

Verrall, R., Nielsen, J. P., and Jessen, A. H. (2010). Prediction of RBNS and IBNR claims using claim amounts and claim counts. *ASTIN Bulletin: The Journal of the IAA*, 40(2):871–887.

A Summarising properties for the Single payment model introduced in Verrall et al. (2010)

Verrall et al. (2010) introduced the following model:

The VNJ model. *Let the total number of claims incurred in accident year i that are reported j periods later, N_{ij}, follow an over-dispersed Poisson distribution according to $N_{ij} \sim \mathrm{ODP}(\nu_{ij}, \phi)$, i.e.*

$$\mathbb{E}[N_{ij}] = \nu_{ij},$$
$$\mathrm{Var}(N_{ij}) = \phi\mathbb{E}[N_{ij}],$$

where all N_{ij} are assumed to be independent. Further, each claim will produce a single payment, paid with a delay after reporting according to a multinomial distribution. That is, for each (i,j), the payments from the N_{ij} claims are distributed in time according to

$$(N_{ij0}^{\mathrm{paid},(VNJ)}, \ldots, N_{ijd}^{\mathrm{paid},(VNJ)}) \sim \mathrm{Mult}(N_{ij}; p_0, \ldots, p_d), \quad \sum_{i=0}^{d} p_i = 1,$$

where $N_{ijk}^{\mathrm{paid},(VNJ)}$ denotes the number of payments stemming from claims reported in (i,j), paid k periods later. Moreover, let Y_{ijk} denote the kth payment in (i,j), where all Y_{ijk} are assumed to be independent with

$$\mathbb{E}[Y_{ijk}] = \mu,$$
$$\mathrm{Var}(Y_{ijk}) = \sigma^2.$$

Note that the VNJ model consists of two parts: (i) a model for claim counts N_{ij} and (ii) a model for how payments are distributed given observed claim counts. An appealing feature of the VNJ model is that by assuming that

$$\nu_{ij} := \alpha_i\beta_j, \quad \sum_{j=0}^{m-1} \beta_j = 1, \tag{21}$$

these parameters may be estimated using the standard Chain-Ladder technique, see e.g. Renshaw and Verrall (1998). Concerning estimation of the parameters of the payment part, part (ii), of the VNJ model, it follows by construction that the total amount paid in $(i,j) \in \mathcal{A}_0$ is given by

$$X_{ij}^{(\mathrm{VNJ})} := \sum_{k=0}^{N_{ij}^{\mathrm{paid},(\mathrm{VNJ})}} Y_{ijk},$$

where "(VNJ)" refers to "the VNJ model" and

$$N_{ij}^{\mathrm{paid},(\mathrm{VNJ})} := \sum_{k=0}^{j\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{VNJ})}.$$

In particular it follows that

$$\mathbb{E}[X_{ij}^{(\mathrm{VNJ})} \mid \mathcal{N}_0] = \mu \sum_{k=0}^{j\wedge d} p_k N_{i,j-k}, \tag{22}$$

$$\mathrm{Var}(X_{ij}^{(\mathrm{VNJ})} \mid \mathcal{N}_0) = \sigma^2 \sum_{k=0}^{j\wedge d} p_k N_{i,j-k} + \mu^2 \sum_{k=0}^{j\wedge d} N_{i,j-k} p_k (1-p_k) \tag{23}$$

$$\approx \varphi \mathbb{E}[X_{ij}^{(\mathrm{VNJ})} | \mathcal{N}_0], \tag{24}$$

where $\varphi = (\sigma^2 + \mu^2)/\mu$. That is, conditional on $\mathcal{N}_0$, X_{ij} may be, approximately, defined in terms of a quasi-Poisson or over-dispersed Poisson generalized linear model, whose parameters are easy to estimate using standard statistical software, e.g. R, by introducing the alternative parametrisation $\psi_k = \mu p_k$ and using that $\mu = \sum_{k=0}^{d} \psi_k$ together with $\sigma^2 + \mu^2 = \varphi\mu$:

$$\begin{cases} \hat{p}_k & = \frac{\hat{\psi}_k}{\sum_{j=0}^{d} \hat{\psi}_j}, \\ \hat{\mu}^{(\mathrm{VNJ})} & = \sum_{j=0}^{d} \hat{\psi}_j, \\ (\hat{\sigma}^{(\mathrm{VNJ})})^2 & = \hat{\varphi}\hat{\mu}^{(\mathrm{VNJ})} - (\hat{\mu}^{(\mathrm{VNJ})})^2. \end{cases}$$

Further, if we let $R_i^{\mathcal{R},(\mathrm{VNJ})}$ and $R_i^{\mathcal{I},(\mathrm{VNJ})}$ be defined in accordance with (6) and (7), then

$$\mathbb{E}\left[R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0\right] = \mu \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} p_k N_{i,j-k}, \tag{25}$$

$$\mathbb{E}\left[R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}_0\right] = \mu \sum_{j=m-i+1}^{m-1+d} \sum_{k=0}^{(j-(m-i+1))\wedge d} p_k \nu_{i,j-k}. \tag{26}$$

Note that in (26) we have used that $\mathbb{E}[N_{i,j-k} \mid \mathcal{N}_0] = \mathbb{E}[N_{i,j-k}] = \nu_{i,j-k}$ due to construction. The reserve estimators used in practice are hence obtained by replacing all parameters in (25) and (26) with their corresponding estimates. That is, the ν_{ij} are estimated using the chain ladder technique and all other parameters are estimated using the above described quasi-likelihood approach.

Further, the variance for the VNJ model, not found in Verrall et al. (2010), is given in the following proposition:

Proposition 11. *The variance of the total outstanding claims payments in the VNJ model is given by*

$$\begin{aligned} \mathrm{Var}(R^{(VNJ)} \mid \mathcal{N}_0) &= \mathrm{Var}\left(\sum_{i=1}^{m} R_i^{\mathcal{R},(VNJ)} + R_i^{\mathcal{I},(VNJ)} \,\middle|\, \mathcal{N}_0\right) \\ &= \sum_{i=1}^{m} \mathrm{Var}(R_i^{\mathcal{R},(VNJ)} \mid \mathcal{N}_0) + \sum_{i=2}^{m} \mathrm{Var}(R_i^{\mathcal{I},(VNJ)}), \end{aligned}$$

where

$$\begin{aligned} \mathrm{Var}(R_i^{\mathcal{R},(VNJ)} \mid \mathcal{N}_0) = \sigma^2 &\sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} p_k N_{i,j-k} \\ &+ \mu^2 \sum_{j=0}^{m-i} N_{ij} q_j (1-q_j), \end{aligned} \tag{27}$$

where

$$q_j = \sum_{k=m-i+1-j}^{d\wedge(m-1+d-j)} p_k,$$

and where

$$\mathrm{Var}(R_i^{\mathcal{I},(VNJ)}) = (\sigma^2 + \phi\mu^2) \sum_{j=m-i+1}^{m-1} \nu_{i,j}. \tag{28}$$

The proof of Proposition 11 is given in Appendix B.

B Proofs

Proof of Proposition 1. By towered expectations we have

$$\begin{aligned}
\mathbb{E}[X_{ij}^{(\mathrm{CRM})}|\mathcal{N}_0] &= \sum_{k=0}^{j\wedge d} \mathbb{E}[X_{i,j-k,k}^{(\mathrm{CRM})} \mid \mathcal{N}_0] \\
&= \sum_{k=0}^{j\wedge d} \mu_{i,j-k,k} E[N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}]|\mathcal{N}_0] \\
&= \sum_{k=0}^{j\wedge d} \lambda_k \mu_{i,j-k,k} N_{i,j-k},
\end{aligned}$$

and by variance decomposition we have

$$\begin{aligned}
\mathrm{Var}(X_{ij}^{(\mathrm{CRM})}|\mathcal{N}_0) &= \sum_{k=0}^{j\wedge d} \left(\sigma_{i,j-k,k}^2 \mathbb{E}[N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}|\mathcal{N}_0] + \mu_{i,j-k,k}^2 \,\mathrm{Var}(N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}|\mathcal{N}_0)\right) \\
&= \sum_{k=0}^{j\wedge d} (\sigma_{i,j-k,k}^2 + \mu_{i,j-k,k}^2)\lambda_k N_{i,j-k} \\
&= \sum_{k=0}^{j\wedge d} \frac{\sigma_{i,j-k,k}^2 + \mu_{i,j-k,k}^2}{\mu_{i,j-k,k}} \mathbb{E}[X_{i,j-k,k}^{(\mathrm{CRM})}|\mathcal{N}_0].
\end{aligned}$$

The independence between all $X_{ij}^{(\mathrm{CRM})}$ and $X_{ijk}^{(\mathrm{CRM})}$ follows trivially from the independence between all $N_{ijk}^{\mathrm{paid},(\mathrm{CRM})}$ and $\widetilde{Y}_{ijkl}$. □

Proof of Proposition 2 Part (i). We can split the quasi-likelihood into two parts according to:

$$l^p(\boldsymbol{\theta}; \mathcal{F}_0) = l_{tri}^p(\boldsymbol{\theta}; \mathcal{F}_0) + l_n^p(\boldsymbol{\theta}; \mathcal{F}_0),$$

where

$$\begin{aligned}
l_{tri}^p(\boldsymbol{\theta}; \mathcal{F}_0) &= \sum_{i=1}^{m} \sum_{j=0}^{m-i+1} \left[X_{ij} \log\left(\sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k}\right) - \sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k} \right], \\
l_n^p(\boldsymbol{\theta}; \mathcal{F}_0) &= \sum_{i=-n}^{0} \sum_{j=0}^{m} \left[X_{ij} \log\left(\sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k}\right) - \sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k} \right].
\end{aligned}$$

This function is a sum of compositions of a mapping of the form $\eta \mapsto a\log(\eta) - \eta$ (which is concave for $a \geq 0$) and the mapping $\boldsymbol{\theta} \mapsto \sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k}$ (which is also concave). Since compositions of concave functions are concave, l^p is concave in $\boldsymbol{\theta}$. Define $M_n(\boldsymbol{\theta}) := \frac{1}{n} l^p(\boldsymbol{\theta}; \mathcal{F}_0)$, then

$$M_n(\boldsymbol{\theta}) \xrightarrow{p} \sum_{j=0}^{m} \mathbb{E}\left[X_{ij} \log\left(\sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k}\right) - \sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k} \right] =: M(\boldsymbol{\theta}),$$

as $n \to \infty$ for all $\boldsymbol{\theta} \in \Theta$, where Θ is the parameter space (either $\mathbb{R}$ or $\mathbb{R}_+$ in our case), since $\frac{1}{n} l_n^p(\boldsymbol{\theta}; \mathcal{F}_0) \xrightarrow{p} M(\boldsymbol{\theta})$ by the weak law of large numbers and $\frac{1}{n} l_{tri}^p(\boldsymbol{\theta}; \mathcal{F}_0) \xrightarrow{p} 0$ trivially. Now, by iterated expectations, we know that

$$M(\boldsymbol{\theta}) = \sum_{j=0}^{m} \mathbb{E}\left[\left(\sum_{k=1}^{j\wedge d} \psi_k N_{i,j-k}\right) \log\left(\sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k}\right) - \sum_{k=1}^{j\wedge d} \theta_k N_{i,j-k} \right].$$

This function is again concave and has the unique optimum $\boldsymbol{\psi}$. By (Schweder and Hjort, 2016, Lemma A.2) it now follows that $\hat{\boldsymbol{\psi}}$ converges in probability to $\boldsymbol{\psi}$.

$\square$

Proof of Proposition 2 Part (ii). We first note that $\frac{N_{ij}}{w_i} \xrightarrow{p} \nu_{ij}$ as $w_i \to \infty$ since, for any $\delta > 0$,

$$\mathbb{P}\left(\left|\frac{N_{ij}}{w_i} - \nu_{ij}\right| > \delta\right) \leq \frac{\mathrm{Var}(N_{ij})}{\delta^2 w_i^2} = \frac{\phi \nu_{ij}}{\delta^2 w_i} \to 0,$$

as $w_i \to \infty$. Moreover, for any $\delta > 0$,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{X_{ij}}{w_i} - \sum_{k=0}^{j \wedge d} \psi_k \nu_{i,j-k}\right| > \delta\right) &\leq \frac{\mathrm{Var}(X_{ij})}{\delta^2 w_i^2} \\ &= \frac{1}{\delta^2 w_i} \sum_{k=0}^{j \wedge d} \lambda_k \nu_{i,j-k} \left(\mu^2(1+\phi\lambda_k) + \sigma^2\right) \to 0,\end{aligned}$$

as $w_i \to \infty$, and therefore $\frac{X_{ij}}{w_i} \xrightarrow{p} \sum_{k=0}^{j \wedge d} \psi_k \nu_{i,j-k}$ as $w_i \to \infty$.

Furthermore, we have that

$$\frac{N_{ij}}{\sum_i w_i} = \frac{w_i}{\sum_i w_i} \frac{N_{ij}}{w_i} \xrightarrow{p} c_i \nu_{ij},$$

where $c_i \in [0, 1]$.

Now, let

$$\begin{aligned}M_n(\boldsymbol{\theta}) &:= \frac{1}{\sum_i w_i} \sum_{i,j \in \mathcal{I}} \left[X_{ij} \log\left(\frac{\sum_{k=0}^{j \wedge d} \theta_k N_{i,j-k}}{w_i}\right) - \sum_{k=0}^{j \wedge d} \theta_k N_{i,j-k}\right] \\ &\propto \frac{l^p(\boldsymbol{\theta}; \mathcal{F}_0)}{\sum_i w_i},\end{aligned}$$

which is nothing but the normalized quasi-likelihood with a constant function of the w_is added to it. It will be concave and have the same optimum as the quasi-likelihood. From the above it is clear that

$$M_n(\boldsymbol{\theta}) \xrightarrow{p} \sum_{i,j \in \mathcal{I}} c_i \left[\left(\sum_{k=0}^{j \wedge d} \psi_k \nu_{i,j-k}\right) \log\left(\sum_{k=0}^{j \wedge d} \theta_k \nu_{i,j-k}\right) - \sum_{k=0}^{j \wedge d} \theta_k \nu_{i,j-k}\right] =: M(\boldsymbol{\theta})$$

and that the optimum of M is $\boldsymbol{\psi}$. It now follows, again by (Schweder and Hjort, 2016, Lemma A.2), that $\hat{\boldsymbol{\psi}}$ converges in probability to $\boldsymbol{\psi}$. $\square$

Proof of Proposition 11. Since all accident years, as well as IBNR and RBNS claims payments, are independent it suffices to analyse $R_i^{(\mathrm{VNJ})}$. Further, by expanding $R_i^{(\mathrm{VNJ})}$ we get that

$$\begin{aligned}\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0) &= \mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} X_{ij}^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0\right) \\ &= \mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} \sum_{l=0}^{N_{ij}^{\mathcal{R},(\mathrm{VNJ})}} Y_{ijl}^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0\right).\end{aligned}$$

Moreover, if we set $\mathcal{N}^{\mathcal{R},(\mathrm{VNJ})} := \sigma\{N_{ij}^{\mathcal{R},(\mathrm{VNJ})} : (i,j) \in \mathcal{A}_0^*\}$ we can use variance decomposition which yields

$$\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0) = \mathbb{E}[\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}) \mid \mathcal{N}_0] + \mathrm{Var}(\mathbb{E}[R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}] \mid \mathcal{N}_0),$$

where

$$\begin{aligned}
\mathbb{E}[\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}) \mid \mathcal{N}_0] &= \mathbb{E}\left[\sum_{j=m-i+1}^{m-1+d} \mathrm{Var}\left(\sum_{l=0}^{N_{ij}^{\mathcal{R},(\mathrm{VNJ})}} Y_{ijl}^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}\right) \mid \mathcal{N}_0\right] \\
&= \mathbb{E}\left[\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{R},(\mathrm{VNJ})} \sigma^2 \mid \mathcal{N}_0\right] \\
&= \sigma^2 \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(mi-i)}^{j\wedge d} p_k N_{i,j-k},
\end{aligned}$$

where the first equality follows from that all $Y_{ijl}^{\mathcal{R},(\mathrm{VNJ})}$ are independent together with the conditioning on $\mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}$. Similarly it follows that

$$\begin{aligned}
\mathrm{Var}(\mathbb{E}[R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}] \mid \mathcal{N}_0) &= \mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} \mathbb{E}\left[\sum_{l=0}^{N_{ij}^{\mathcal{R},(\mathrm{VNJ})}} Y_{ijl}^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}\right] \mid \mathcal{N}_0\right) \\
&= \mu^2\, \mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0\right),
\end{aligned}$$

where we may make use of that

$$\begin{aligned}
\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{R},(\mathrm{VNJ})} &= \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{VNJ})} \\
&= \sum_{l=0}^{m-i} \sum_{v=m-i+1-l}^{d\wedge(m-1+d-l)} N_{i,l,v}^{\mathrm{paid},(\mathrm{VNJ})} \qquad (29)
\end{aligned}$$

which combined with the above gives us

$$\begin{aligned}
\mathrm{Var}(\mathbb{E}[R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{VNJ})}] \mid \mathcal{N}_0) &= \mu^2\, \mathrm{Var}\left(\sum_{l=0}^{m-i} \sum_{v=m-i+1-l}^{d\wedge(m-1+d-l)} N_{i,l,v}^{\mathrm{paid},(\mathrm{VNJ})} \mid \mathcal{N}_0\right) \\
&= \mu^2 \sum_{l=0}^{m-i} N_{il} q_l(1-q_l),
\end{aligned}$$

where

$$q_l = \sum_{v=m-i+1-l}^{d\wedge(m-1+d-l)} p_v.$$

Analogously we may, for $i \geq 2$, calculate

$$\begin{aligned}
\mathrm{Var}(R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}_0) &= \mathrm{Var}(R_i^{\mathcal{I},(\mathrm{VNJ})}) \\
&= \mathbb{E}[\mathrm{Var}(R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{VNJ})})] + \mathrm{Var}(\mathbb{E}[R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{VNJ})}]).
\end{aligned}$$

The first term is given by

$$\begin{aligned}
\mathbb{E}[\mathrm{Var}(R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{VNJ})})] &= \sigma^2 \sum_{j=m-i+1}^{m-1+d} \sum_{k=0}^{(j-(m-i+d))\wedge d} p_k \mathbb{E}[N_{i,j-k}] \\
&= \sigma^2 \sum_{j=m-i+1}^{m-1+d} \sum_{k=0}^{(j-(m-i+d))\wedge d} p_k \nu_{i,j-k} \\
&= \sigma^2 \sum_{j=m-i+1}^{m-1+d} \nu_{ij},
\end{aligned}$$

where the last equality follows from that we for each ν_{ij} will sum over all p_k where $\sum p_k = 1$, i.e. we sum over all possible future payment periods. Further, the second term may be simplified to

$$\mathrm{Var}(\mathbb{E}[R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{VNJ})}]) = \mu^2 \,\mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{I},(\mathrm{VNJ})}\right)$$

and we may, similarly to (29), make use of an alternative re-indexation which gives us that

$$\begin{aligned}
\sum_{j=m-i+1}^{m-1+d} \sum_{k=0}^{(j-(m-i+d))\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{VNJ})} &= \sum_{l=m-i+1}^{m-1} \sum_{v=0}^{d} N_{i,l,v}^{\mathrm{paid},(\mathrm{VNJ})} \\
&= \sum_{l=m-i+1}^{m-1} N_{il},
\end{aligned}$$

where the last equality is due to the multinomial assumption. By combining the above we get

$$\begin{aligned}
\mathrm{Var}(\mathbb{E}[R_i^{\mathcal{I},(\mathrm{VNJ})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{VNJ})}]) &= \mu^2 \,\mathrm{Var}\left(\sum_{j=m-i+1}^{m-1} N_{ij}\right) \\
&= \mu^2 \phi \sum_{j=m-i+1}^{m-1} \nu_{ij},
\end{aligned}$$

which finally gives us that

$$\mathrm{Var}(R_i^{\mathcal{I},(\mathrm{VNJ})}) = (\sigma^2 + \mu^2\phi) \sum_{j=m-i+1}^{m-1} \nu_{ij}, \ i \geq 2.$$

□

Proof of Proposition 3. The proof of the variance of the outstanding payments stemming from RBNS claims is identical with that of Proposition 11, although the assumption of a Poisson number of payments per claim introduces independence between disjoint time intervals. Thus,

$$\begin{aligned}
\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}_0) =& \mathbb{E}[\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{CRM})}) \mid \mathcal{N}_0] \\
&+ \mathrm{Var}(\mathbb{E}[R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{CRM})}] \mid \mathcal{N}_0)
\end{aligned}$$

where

$$\mathbb{E}[\mathrm{Var}(R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{CRM})}) \mid \mathcal{N}_0] = \mathbb{E}\left[\mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} \sum_{l=0}^{N_{ij}^{\mathcal{R},(\mathrm{CRM})}} Y_{ijl}^{\mathcal{R}} \mid \mathcal{N}^{\mathcal{R},(\mathrm{CRM})}\right) \mid \mathcal{N}_0\right]$$

$$= \sigma^2 \sum_{j=m-i+1}^{m-1+d} \mathbb{E}\left[\sum_{k=j-(m-i)}^{j\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})} \mid \mathcal{N}_0\right]$$

$$= \sigma^2 \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} \lambda_k N_{i,j-k},$$

and

$$\mathrm{Var}(\mathbb{E}[R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}^{\mathcal{R},(\mathrm{CRM})}] \mid \mathcal{N}_0) = \mu^2 \,\mathrm{Var}\left(\sum_{k=j-(m-i)}^{j\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})} \mid \mathcal{N}_0\right)$$

$$= \mu^2 \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} \lambda_k N_{i,j-k},$$

which concludes the proof for the RBNS reserve.

Concerning the IBNR reserve, we may again exploit the independence between payments in different time intervals, which gives us that

$$\mathrm{Var}(R_i^{\mathcal{I},(\mathrm{CRM})}) = \mathbb{E}[\mathrm{Var}(R_i^{\mathcal{I},(\mathrm{CRM})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{CRM})})] + \mathrm{Var}(\mathbb{E}[R_i^{\mathcal{I},(\mathrm{CRM})} \mid \mathcal{N}^{\mathcal{I},(\mathrm{CRM})}])$$

$$= \sigma^2 \mathbb{E}\left[\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{I},(\mathrm{CRM})}\right] + \mu^2 \,\mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{I},(\mathrm{CRM})}\right).$$

By letting $\mathcal{N}_0^* := \sigma\{N_{ij} : (i,j) \in \mathcal{A}_0^*\}$ we can use towered expectation to get

$$\mathbb{E}\left[\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{I},(\mathrm{CRM})}\right] = \mathbb{E}\left[\sum_{j=m-i+1}^{m-1+d} \sum_{k=0\vee(j-m+1)}^{(j-(m-i+1))\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})}\right]$$

$$= \mathbb{E}\left[\mathbb{E}\left[\sum_{j=m-i+1}^{m-1+d} \sum_{k=0\vee(j-m+1)}^{(j-(m-i+1))\wedge d} N_{i,j-k,k}^{\mathrm{paid},(\mathrm{CRM})} \mid \mathcal{N}_0^*\right]\right]$$

$$= \sum_{j=m-i+1}^{m-1+d} \sum_{k=0\vee(j-m+1)}^{(j-(m-i+1))\wedge d} \lambda_k \nu_{i,j-k}$$

$$= \Lambda \sum_{j=m-i+1}^{m-1} \nu_{ij},$$

since the sum spans all possible future payment periods for each ν_{ij}. Following the IBNR calculations in the proof of Proposition 11, where the same argument together with another variance decomposition w.r.t. $\mathcal{N}_0^*$ yields

$$\mathrm{Var}\left(\sum_{j=m-i+1}^{m-1+d} N_{ij}^{\mathcal{I},(\mathrm{CRM})}\right) = \mathbb{E}\left[\Lambda \sum_{j=m-i+1}^{m-1} N_{ij}\right] + \mathrm{Var}\left(\Lambda \sum_{j=m-i+1}^{m-1} N_{ij}\right)$$

$$= \Lambda(1+\Lambda\phi) \sum_{j=m-i+1}^{m-1} \nu_{ij},$$

which finishes the proof. $\square$

Proof of Proposition 6. By Lemma 5.2 in Lehmann and Casella (2006), as $w_i \to \infty$, $\hat{\pi}_k^{\mathrm{D}}$ converges in probability to $\pi_k^{\mathrm{D},asy}$, denoting the solution of

$$\begin{pmatrix} \tilde{\beta}_0 \\ \vdots \\ \vdots \\ \tilde{\beta}_{m-1} \end{pmatrix} = \begin{pmatrix} \beta_0 & 0 & \dots & 0 \\ \beta_1 & \beta_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \beta_0 & \dots & \beta_1 & \beta_0 \end{pmatrix} \begin{pmatrix} \pi_0 \\ \vdots \\ \vdots \\ \pi_{m-1} \end{pmatrix}.$$

Row-wise this is

$$\tilde{\beta}_j = \sum_{k=0}^{j} \pi_k^{\mathrm{D},asy} \beta_{j-k}.$$

Noting that $\tilde{\beta}_j = \frac{1}{c}\sum_{k=0}^{j\wedge d} \lambda_k \beta_{j-k}$ and solving for $\pi_k^{\mathrm{D},asy}$ row by row, starting at $j = 0$, we get

$$\pi_k^{\mathrm{D},asy} = \frac{1}{c}\lambda_k 1_{\{k\leq d\}},$$

i.e.

$$\hat{\pi}_k^{\mathrm{D}} \xrightarrow{p} \frac{1}{c}\lambda_k 1_{\{k\leq d\}},$$

as $w_i \to \infty$.

For the parameter μ^{D} we note that

$$\hat{\mu}^{\mathrm{D}} = \frac{\hat{\tilde{\alpha}}_1}{\hat{\alpha}_1} \xrightarrow{p} \frac{\tilde{\alpha}_1}{\alpha_1} = c\mu,$$

as $w_i \to \infty$ by Proposition 5. Therefore, for $0 \leq k \leq d$,

$$\hat{\psi}_k^{\mathrm{D}} := \hat{\mu}^{\mathrm{D}}\hat{\pi}_k^{\mathrm{D}} \xrightarrow{p} \mu\lambda_k =: \psi_k,$$

as $w_i \to \infty$. $\square$

Proof of Proposition 7. Note that $\bar{r}_i$ is a direct consequence of (8) together with the assumed parametrisation of $\nu_{ij} = w_i\bar{\alpha}_i\beta_j$. Concerning $\bar{v}_i$, we can rewrite (8) as

$$h_i^{\mathcal{R},(\mathrm{CRM})}(\boldsymbol{\theta};\mathcal{N}_0) = \mu \sum_{j=0}^{m-i} \sum_{k=m-i+1-j}^{d\wedge(m-1+d-j)} \lambda_k N_{i,j},$$

and therefore

$$\bar{v}_i := \frac{1}{w_i}\,\mathrm{Var}(h_i^{\mathcal{R},(\mathrm{VNJ})}(\boldsymbol{\theta};\mathcal{N}_0)) = \phi\mu^2 \sum_{j=0}^{m-i} \left(\sum_{k=m-i+1-j}^{d\wedge(m-1+d-j)} \lambda_k \right)^2 \bar{\alpha}_i\beta_j.$$

By Chebyshev's inequality it holds for $\delta > 0$ that

$$\mathbb{P}\left(\left| \frac{1}{w_i} h_i^{\mathcal{R},(\mathrm{CRM})}(\boldsymbol{\theta};\mathcal{N}_0) - \bar{r}_i \right| > \delta \right) \leq \frac{\bar{v}_i}{\delta^2 w_i} \to 0,$$

as $w_i \to \infty$. $\square$

Proof of Proposition 9. For the CRM we have

$$\begin{aligned}\widehat{\mathrm{Var}}_{(\mathrm{CRM})}(X_{ij}^{(\mathrm{CRM})}|\mathcal{N}_0) &= ((\hat{\sigma}^{(\mathrm{CRM})})^2 + \hat{\mu}^2)\sum_{k=0}^{j\wedge d}\hat{\lambda}_k N_{i,j-k} \\ &= \hat{\varphi}\hat{\mu}\hat{\Lambda}\sum_{k=0}^{j\wedge d}\hat{\pi}_k N_{i,j-k},\end{aligned}$$

and for the VNJ model we have

$$\begin{aligned}\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(X_{ij}^{(\mathrm{VNJ})}|\mathcal{N}_0) &= (\hat{\sigma}^{(\mathrm{VNJ})})^2\sum_{k=0}^{j\wedge d}\hat{p}_k N_{i,j-k} + (\hat{\mu}^{(\mathrm{VNJ})})^2\sum_{k=0}^{j\wedge d}N_{i,j-k}\hat{p}_k(1-\hat{p}_k) \\ &= (\hat{\varphi}\hat{\Lambda}\hat{\mu} - \hat{\Lambda}^2\hat{\mu}^2 + \hat{\Lambda}^2\hat{\mu}^2)\sum_{k=0}^{j\wedge d}\hat{p}_k N_{i,j-k} - \hat{\mu}^2\hat{\Lambda}^2\sum_{k=0}^{j\wedge d}N_{i,j-k}\hat{p}_k^2 \\ &= \hat{\varphi}\hat{\Lambda}\hat{\mu}\sum_{k=0}^{j\wedge d}\hat{p}_k N_{i,j-k} - \hat{\mu}^2\hat{\Lambda}^2\sum_{k=0}^{j\wedge d}N_{i,j-k}\hat{p}_k^2.\end{aligned}$$

Therefore,

$$\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(X_{ij}^{(\mathrm{VNJ})}|\mathcal{N}_0) - \widehat{\mathrm{Var}}_{(\mathrm{CRM})}(X_{ij}^{(\mathrm{CRM})}|\mathcal{N}_0) = -\hat{\mu}^2\hat{\Lambda}^2\sum_{k=0}^{j\wedge d}N_{i,j-k}\hat{p}_k^2 \le 0,$$

and thus first part follows.

We now move on to the variances of the total outstanding claims stemming from either IBNR or RBNS claims. We start with the RBNS variance. First of all, inserting the parameter estimators of the CRM into (10) yields

$$\begin{aligned}\widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}_0) &= (\hat{\sigma}^{(\mathrm{CRM})})^2\sum_{j=m-i+1}^{m-1+d}\sum_{k=j-(m-i)}^{j\wedge d}\hat{\lambda}_k N_{i,j-k} \\ &= \hat{\Lambda}\left(\hat{\varphi}\hat{\mu} - \hat{\mu}^2\right)\sum_{j=m-i+1}^{m-1+d}\sum_{k=j-(m-i)}^{j\wedge d}\frac{1}{\hat{\mu}\hat{\Lambda}}\hat{\psi}_k N_{i,j-k}.\end{aligned}$$

In the same way, inserting the parameter estimators of the VNJ model into (27) yields

$$\begin{aligned}&\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0) \\ &= (\hat{\sigma}^{(\mathrm{VNJ})})^2\sum_{j=m-i+1}^{m-1+d}\sum_{k=j-(m-i)}^{j\wedge d}\hat{p}_k N_{i,j-k} + (\hat{\mu}^{(\mathrm{VNJ})})^2\sum_{j=0}^{m-i}N_{ij}\hat{q}_j(1-\hat{q}_j) \\ &= \left(\hat{\varphi}\hat{\mu}\hat{\Lambda} - \hat{\mu}^2\hat{\Lambda}^2\right)\sum_{j=m-i+1}^{m-1+d}\sum_{k=j-(m-i)}^{j\wedge d}\frac{1}{\hat{\mu}\hat{\Lambda}}\hat{\psi}_k N_{i,j-k} + \hat{\mu}^2\hat{\Lambda}^2\sum_{j=0}^{m-i}N_{ij}\hat{q}_j(1-\hat{q}_j).\end{aligned}$$

By noting that

$$\sum_{j=0}^{m-i}N_{ij}\hat{q}_j(1-\hat{q}_j) = \sum_{j=0}^{m-i}N_{ij}\hat{q}_j - \sum_{j=0}^{m-i}N_{ij}\hat{q}_j^2,$$

and

$$
\begin{aligned}
\sum_{j=0}^{m-i} N_{ij}\hat{q}_j &= \sum_{j=0}^{m-i} \sum_{k=m-i+1-j}^{d\wedge(m-1+d-j)} \hat{p}_k N_{ij} \\
&= \sum_{j=m-i+1}^{m-1+d} \sum_{k=j-(m-i)}^{j\wedge d} \hat{p}_k N_{i,j-k},
\end{aligned}
$$

we finally get that

$$
\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0) - \widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}_0) = \hat{\mu}^2\hat{\Lambda}\left(\sum_{j=0}^{m-i} N_{ij}\hat{q}_j - \hat{\Lambda}\sum_{j=0}^{m-i} N_{ij}\hat{q}_j^2\right).
$$

From which it is clear that

$$
\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{R},(\mathrm{VNJ})} \mid \mathcal{N}_0) \leq \widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{R},(\mathrm{CRM})} \mid \mathcal{N}_0)
$$

if and only if (16) holds.

For the IBNR variance we go about things in the same way. Inserting the parameter estimators of the CRM into (11) yields

$$
\begin{aligned}
\widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{I},(\mathrm{CRM})}) &= \left((\hat{\sigma}^{(\mathrm{CRM})})^2 + \hat{\mu}^2(1+\hat{\Lambda}\hat{\phi})\right)\hat{\Lambda}\sum_{j=m-i+1}^{m-1} \hat{\nu}_{ij} \\
&= \left(\hat{\varphi}\hat{\mu} + \hat{\mu}^2\hat{\Lambda}\hat{\phi}\right)\hat{\Lambda}\sum_{j=m-i+1}^{m-1} \hat{\nu}_{ij},
\end{aligned}
$$

and inserting the parameter estimators of the VNJ model into (28) yields

$$
\begin{aligned}
\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{I},(\mathrm{VNJ})}) &= \left((\hat{\sigma}^{(\mathrm{VNJ})})^2 + \hat{\phi}(\hat{\mu}^{(\mathrm{VNJ})})^2\right)\sum_{j=m-i+1}^{m-1} \hat{\nu}_{ij} \\
&= \left(\hat{\varphi}\hat{\mu} - \hat{\mu}^2\hat{\Lambda} + \hat{\phi}\hat{\mu}^2\hat{\Lambda}\right)\hat{\Lambda}\sum_{j=m-i+1}^{m-1} \hat{\nu}_{ij}.
\end{aligned}
$$

Therefore we can conclude that,

$$
\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{I},(\mathrm{VNJ})}) - \widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{I},(\mathrm{CRM})}) = -\hat{\mu}^2\hat{\Lambda}^2\sum_{j=m-i+1}^{m-1} \hat{\nu}_{ij} \leq 0.
$$

Finally, the last part of the proposition follows in the same way as the first part, but instead looking at the difference

$$
\left(\widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{I},(\mathrm{VNJ})}) + \widehat{\mathrm{Var}}_{(\mathrm{VNJ})}(R_i^{\mathcal{R},(\mathrm{VNJ})})\right) - \left(\widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{I},(\mathrm{CRM})}) + \widehat{\mathrm{Var}}_{(\mathrm{CRM})}(R_i^{\mathcal{R},(\mathrm{CRM})})\right).
$$

This calculation is left out of the proof since it is a tedious repetition of the above calculations with a slight modification. □