Stochastic loss reserving using individual information model with over-dispersed Poisson

Zhigao Wang, Xianyi Wu & Chunjuan Qiu

STATISTICAL THEORY AND RELATED FIELDS
2022, VOL. 6, NO. 2, 114–128
https://doi.org/10.1080/24754269.2021.1898181

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Stochastic loss reserving using individual information model with over-dispersed Poisson

Zhigao Wang, Xianyi Wu and Chunjuan Qiu

School of Statistics, East China Normal University, Shanghai, People's Republic of China

ABSTRACT

For stochastic loss reserving, we propose an individual information model (IIM) which accommodates not only individual/micro data consisting of incurring times, reporting developments, settlement developments as well as payments of individual claims but also heterogeneity among policies. We give over-dispersed Poisson assumption about the moments of reporting developments and payments of every individual claims. Model estimation is conducted under quasi-likelihood theory. Analytic expressions are derived for the expectation and variance of outstanding liabilities, given historical observations. We utilise conditional mean square error of prediction (MSEP) to measure the accuracy of loss reserving and also theoretically prove that when risk portfolio size is large enough, IIM shows a higher prediction accuracy than individual/micro data model (IDM) in predicting the outstanding liabilities, if the heterogeneity indeed influences claims developments and otherwise IIM is asymptotically equivalent to IDM. Some simulations are conducted to investigate the conditional MSEPs for IIM and IDM. A real data analysis is performed basing on real observations in health insurance.

ARTICLE HISTORY

Received 20 August 2020
Revised 28 December 2020
Accepted 1 March 2021

KEYWORDS

Risk management; loss reserving; individual information model; over-dispersed Poisson; quasi-likelihood; MSEP

1. Introduction

In the background of stochastic reserving, loss reserving is referred to a procedure to predict incurred outstanding liabilities in general insurance companies. It is well known that chain-ladder method proposed by Mack (1993) and its related versions can be easily performed by using pencil and paper because of the simple aggregate data structure called run-off triangle and hence are popular in practice. However, as England and Verrall (2002) mentioned, the advantages of aggregate data models are at the cost of prediction accuracy because of information loss caused by simply aggregating individual or micro data, which records incurring time, reporting time, settlement time as well as payment processes of individual claims. In risk management of insurance companies, with the modern computer technology, it is urgent for actuaries to explore the usage of related information to improve the accuracy in predicting the liabilities, which also attracts increasing interests of many scholars from actuarial science. Antonio and Plat (2014), Pigeon et al. (2013, 2014) demonstrated by an empirical analysis that loss reserving based on individual data had more prediction accuracy than aggregate models. Huang, Qiu, Wu, Zhou (2015), Huang, Qiu, Wu (2015) and Huang et al. (2016) revealed that individual loss reserving had more accuracy than methods using aggregate data in sense that the former produced a smaller mean square error.

A small stream of earlier literature about IDMs, for example, Arjas (1989) and Norberg (1993, 1999) formulated a probabilistic framework for the developments of individual claims. Most recently, Yu and He (2016) modelled the individual claim development process by marked Cox processes (also known as double stochastic processes). As we all know, it is challenging to acquire analytic expressions for the moments of outstanding liabilities under continuous-time IDMs. Perhaps partly for this reason, there is a great deal of work that has been done under discrete-time IDMs, see, e.g., Pigeon et al. (2013, 2014), Verrall et al. (2010), Huang, Qiu, Wu, Zhou (2015), Huang, Qiu, Wu (2015) and Huang et al. (2016). Zhao and Zhou (2010) considered the R-delays so as to predict the incurred but not settled outstanding liabilities. Unfortunately, IDMs also confront information loss caused by neglecting individual information, i.e., information from policy or policyholder. It is not clear so far how much accuracy in predicting the outstanding liabilities is sacrificed, when the individual information is neglected. In the present paper, we will explore how much improvement in the accuracy that will be measured by conditional MSEP can be achieved by incorporating the useful individual information into modelling under discrete time framework similar as Huang, Qiu, Wu, Zhou (2015), Huang, Qiu, Wu (2015) and Huang et al. (2016). Besides, we avoid the strong Poisson distribution assumption for

CONTACT Zhigao Wang wangzhigao2015@163.com

the number of individual claims assumed in Huang, Qiu, Wu, Zhou (2015), Huang, Qiu, Wu (2015) and Huang et al. (2016) and instead extend to weak assumptions about the first two moments so that parameters estimation can be conducted under quasi-likelihood theory (cf. McCullagh & Nelder, 1989).

The conditional MSEP is broadly used to compare different models for loss reserving. It is well known that the conditional MSEP is the sum of process variance caused by the randomness of outstanding liabilities and estimation error originating from uncertainty of parameters estimators. It is theoretically feasible to estimate it by bootstrap method. There were some examples which discussed the MSEP under collective models – for instance, Mack (1993), Mack (2000) (comparing three methods – Benktander, Bornhuetter–Ferguson and chain-ladder under the criteria of MSEP), Alai et al. (2009, 2010) (Bornhuetter–Ferguson method under generalised linear model) and Wüthrich and Merz (2008) (comprehensive summary of the details of methods based on aggregate data). Besides, Lindholm et al. (2020) introduced a semi-analytic approximation method to estimate the conditional MSEP, where the method is illustrated by loss reserving based on aggregate data. Examples that have applied the approximation method are Wahl (2019), who computed explicit moments of outstanding liabilities by applying discretisation scheme under the framework of Antonio and Plat (2014), and Wahl et al. (2019) who modelled individual data on aggregate level. In the present paper, we also use the approximation method for the MSEP, which is derived under IIM, because of its simplification.

The paper is organised as follows. In Section 2, we describe the data structure and display the mathematical expression of outstanding liabilities caused by a risk portfolio at a given evaluation date. In Section 3, we separately model reporting developments, settlement delays and payments of claims and in each part, we formulate the model assumptions as well as estimation for the model parameters. Section 4 mainly derives the formulas of loss reserve and conditional variance of outstanding liabilities given historical observations, and studies the improvement of accuracy achieved by IIM with respect to IDM. Section 5 reports some simulation results and a real data analysis. Section 6 concludes the paper with a few remarks.

2. Data structure

Claim events incurred by some policy are usually reported to the insurer in some time periods (reporting delays) after their occurrence and the reported claims are finally settled with some time lags (settlement delays) between their reports and final settlements. Before going further, it is necessary to discuss the supports of reporting and settlement delays. In the following assumption, we assume that there exists maximum reporting delay D^r and settlement delay D^s. Actually, there are basically two cases for the supports of the delays: finite and infinite. It would be *a* known *priori* (generally read from the items of the insurance contracts) if the supports are finite or infinite before any loss reserving is taken care of. Even for the case the delays take unrestricted values, if the probability to take values over certain limits is quite small, one can safely assume a capped delays by cutting off the tails with probability small enough. As a result, the assumption of capped delays is reasonable in many real insurance businesses, especially for such insurance without very much high claims payments. An example is the general health insurance. The assumption of capped delays has been extensively adopted in such traditional methods as chain-ladder algorithm. If the tails cannot be safely cut off, however, the models such as the one proposed in Crevecoeur et al. (2019) or some others would be more suitable. From the statistical point of view, for their distributions to be reasonably estimated with observations over a finite number of years, at least one of the two assumptions is necessary: they take only a finite number of values with arbitrary probabilities (but subject to normalisation) or countably infinitely many values but with their distribution functions identified by finite many parameters. Whatever the case, the number of unknown parameters that need to be estimated must be finite. Here the former is taken, whereas Crevecoeur et al. (2019), for example, took the latter.

Then we specify the data structure used in our model. It is in discrete time version as, e.g., Huang, Qiu, Wu (2015) did. Typically, the data for modelling is organised through periods with fixed length such as 1 year, one season or 1 month depending on lines of business. Conventionally, those periods are referred to '(accident) years'. This is also a way widely adopted by insurers to predict the incurred outstanding liabilities in practice. Specifically, the whole observation horizon is made of n accident years and loss reserving is evaluated at the end of nth accident year. In year i, $i = 1, 2, \ldots, n$, there are m_i insurance policies, each of which is coded by (i, k), $k = 1, 2, \ldots, m_i$.

Every individual (i, k) is associated with a random risk exposure r_{ik} and d-dimensional vector of covariate $\boldsymbol{x}_{ik}$ whose first entry is 1 and other entries indicating the individual information/features that influence the developments of individual claims. The developments of claims incurred by individual (i, k) are detailed as follows.

(1) The reporting developments of claims are recorded by N^r_{iku}, $u = 0, 1, \ldots, D^r$, where N^r_{iku} is the number of claims which are incurred in year i and reported in year $i + u$.
(2) For N^r_{iku} claims, $u = 0, 1, \ldots, D^r$, their settlement developments are tracked by N_{ikuv}, $v =$

$0, 1, \ldots, D^s$, where N_{ikuv} is the number of claims which are reported in year $i+u$ and settled in year $i+u+v$.

(3) Payments for each claim are assumed to be paid for only once at its final settlement. For N_{ikuv} claims, $u = 0, 1, \ldots, D^r$, $v = 0, 1, \ldots, D^s$, we use Y_{ikuvl}, $l = 1, 2, \ldots, N_{ikuv}$ to record corresponding payments.

Then the random element associated with individual (i, k) is denoted by

$$\{r_{ik}, \boldsymbol{x}_{ik}; \{N^r_{iku}; (N_{ikuv}; (Y_{ikuvl})_{l=1}^{N_{ikuv}})_{v=0}^{D^s}\}_{u=0}^{D^r}\},$$

$i = 1, \ldots, n, k = 1, \ldots, m_i$, which are i.i.d. from the population

$$\{r, \boldsymbol{x}; \{N^r_u; (N_{uv}; (Y_{uvl})_{l=1}^{N_{uv}})_{v=0}^{D^s}\}_{u=0}^{D^r}\},$$

which can be considered as a complete observation of a representative policy in year i.

Following conventional terms, a claim, which has been reported to the insurer but not settled, is known as RBNS claim and a claim, which has been incurred but not reported to the insurer, is known as IBNR claim. For accident year i, the individual observed data is as follows.

(1) The reporting developments of a representative policy in year i are truncated in sense that we can only observe $\mathscr{F}^r = \{N^r_0, N^r_1, \ldots, N^r_{D^r_i}\}$, where

$$D^r_i = D^r \wedge (n - i), \tag{1}$$

represents the largest reporting delays of the reported claims in accident year i.

(2) For N^r_u reported claims, $u = 0, 1, \ldots, D^r_i$, their settlement developments are censored in sense that we can observe $\{N_{u0}, N_{u1}, \ldots, N_{uD^s_{i+u}}, N^{rbns}_u\}$, where

$$D^s_{i+u} = D^s \wedge (n - i - u), \tag{2}$$

is the largest settlement delays of settled claims with reporting delay u in accident year i and the number $N^{rbns}_u := \sum_{v=n-i-u+1}^{D^s} N_{uv}$, which is the number of RBNS claims with reporting delays u. Note that $N^{rbns}_u = 0$ if $n - i - u \geq D^s$. Denote by $\mathscr{F}^s = \bigcup_{u=0}^{D^r_i}\{N_{u0}, N_{u1}, \ldots, N_{uD^s_{i+u}}, N^{rbns}_u\}$.

(3) For N_{uv} settled claims, $u = 0, 1, \ldots, D^r_i, v = 0, 1, \ldots, D^s_{i+u}$, the observed payments for them are gathered in set $\{Y_{uv1}, Y_{uv2}, \ldots, Y_{uvN_{uv}}\}$. Denote $\mathscr{F}^p = \bigcup_{u=0}^{D^r_i} \bigcup_{v=0}^{D^s_{i+u}} \{Y_{uv1}, Y_{uv2}, \ldots, Y_{uvN_{uv}}\}$.

Then individual observation $\mathscr{F}^o$ is the union of $\{r, \boldsymbol{x}\}$, $\mathscr{F}^r$, $\mathscr{F}^s$ and $\mathscr{F}^p$, that is

$$\mathscr{F}^o = \{r, \boldsymbol{x}\} \cup \mathscr{F}^r \cup \mathscr{F}^s \cup \mathscr{F}^p$$

and the historical observations of all policies in the portfolio, denoted by $\mathscr{F}^{uo}$, is just the union of policy-specified observation that is $\mathscr{F}^{uo} = \bigcup_{i=1}^n \bigcup_{k=1}^{m_i} \mathscr{F}^o_{ik}$, where $\mathscr{F}^o_{ik}$ is the policy-specified realisations of $\mathscr{F}^o$ in year i that is

$$\mathscr{F}^o_{ik} = \{r_{ik}, \boldsymbol{x}_{ik}\} \cup \mathscr{F}^r_{ik} \cup \mathscr{F}^s_{ik} \cup \mathscr{F}^p_{ik}.$$

It is well known that RBNS and IBNR claims of the risk portfolio naturally result in outstanding liabilities to the insurer. Specifically, the total of future payments for all the RBNS and IBNR claims can be represented as

$$R := \sum_{i=1}^n R^{rbns}_i + \sum_{i=1}^n R^{ibnr}_i, \tag{3}$$

where

$$R^{rbns}_i = \sum_{k=1}^{m_i} \sum_{u=0}^{D^r_i} \sum_{v=n-i-u+1}^{D^s} \sum_{l=1}^{N_{ikuv}} Y_{ikuvl} \quad \text{and}$$

$$R^{ibnr}_i = \sum_{k=1}^{m_i} \sum_{u=n-i+1}^{D^r} \sum_{v=0}^{D^s} \sum_{l=1}^{N_{ikuv}} Y_{ikuvl},$$

are RBNS and IBNR liabilities incurred in year i, respectively. Thoroughly, we take the convention $\sum_{j=j_1}^{j_2} \cdot = 0$ if $j_1 > j_2$.

3. Model specification

This section separately specifies the models for the reporting developments, settlement developments and payments of claims. In each part, we first give model assumptions and then detail the parameter estimations under both IIM and IDM. The model assumptions in this section are all given under the condition that risk exposure r and covariates $\boldsymbol{x}$ are known.

3.1. Modelling reporting developments of claims

Model assumption for reporting developments of claims is given as follows. It is mainly about the first and second moments of reporting developments of claims. The assumption involves vectors of parameters $\boldsymbol{\beta}, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_{D^r}$, which are all d-dimensional vector.

Assumption 3.1: *For an individual with $(r, \boldsymbol{x})$, assume that $N^r_u, u = 0, 1, \ldots, D^r$ are independent, $\mathbb{E}[N^r_u \mid r, \boldsymbol{x}] = r\lambda_u$ and $\mathrm{Var}(N^r_u \mid r, \boldsymbol{x}) = \phi r\lambda_u$, where $\lambda_u = \lambda p_u$ with $\lambda = \exp(\boldsymbol{x}'\boldsymbol{\beta})$ and $p_u(\boldsymbol{\pi}; \boldsymbol{x}) = \frac{\exp(\boldsymbol{x}'\boldsymbol{\pi}_u)}{\sum_{j=0}^{D^r}\exp(\boldsymbol{x}'\boldsymbol{\pi}_j)}$, $\boldsymbol{\pi}_0 = \boldsymbol{0}$ as well as $\boldsymbol{\pi}' = (\boldsymbol{\pi}'_1, \boldsymbol{\pi}'_2, \ldots, \boldsymbol{\pi}'_{D^r})$.*

Remark 3.1: In order to make $\boldsymbol{\pi}$ be reasonably estimated, the condition $n > D^r$ is necessary.

By independence among policies and assumption above, one can construct the quasi-likelihood function of reported claims as follows,

$$Q^r(\boldsymbol{\beta},\boldsymbol{\pi}) = \frac{1}{\phi}\sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{u=0}^{D_i^r}(N_{iku}^r \log\lambda_{iku} - r_{ik}\lambda_{iku}), \tag{4}$$

where λ_{iku} is policy-specified quantities of λ_u that is

$$\lambda_{iku} = \lambda_{ik}p_{iku} = \exp(\boldsymbol{x}_{ik}'\boldsymbol{\beta})\cdot\frac{\exp(\boldsymbol{x}_{ik}'\boldsymbol{\pi}_u)}{\sum_{j=0}^{D^r}\exp(\boldsymbol{x}_{ik}'\boldsymbol{\pi}_j)}.$$

One can refer to McCullagh and Nelder (1989) for more details about quasi-likelihood theory. Similar to maximum likelihood estimation, parameters $(\boldsymbol{\beta},\boldsymbol{\pi})$ can be estimated by maximising $Q^r(\boldsymbol{\beta},\boldsymbol{\pi})$ with respect to the parameters. Denote by $\boldsymbol{\lambda} = \text{vec}(r_{ik}\lambda_{iku},\ i = 1,\dots,n,\ k = 1,\dots,m_i,\ u = 0,\dots,D_i^r)$ and stack N_{iku}^rs as a vector $\boldsymbol{N}^r$ such that entry N_{iku} is corresponding to $r_{ik}\lambda_{iku}$ in vector $\boldsymbol{\lambda}$. The quasi-score function, i.e., partial derivatives of $Q^r(\boldsymbol{\beta},\boldsymbol{\pi})$ with respect to the parameters is

$$\nabla Q^r(\boldsymbol{\beta},\boldsymbol{\pi}) := \frac{\partial Q^r(\boldsymbol{\beta},\boldsymbol{\pi})}{\partial(\boldsymbol{\beta}',\boldsymbol{\pi}')'} = \frac{1}{\phi}X^{r\prime}\text{diag}(\boldsymbol{\lambda})^{-1}(\boldsymbol{N}^r - \boldsymbol{\lambda}), \tag{5}$$

where $X^r = \frac{\partial\boldsymbol{\lambda}}{\partial(\boldsymbol{\beta}',\boldsymbol{\pi}')}$. To determine the block entries of X^r, one needs the unit vector $\boldsymbol{\delta}_s$ with 1 at component s (any positive integer) and $\boldsymbol{\delta}_0 = \boldsymbol{0}$, of which dimensions can be read from context, and the following partial derivatives

$$r_{ik}\cdot\frac{\partial\lambda_{iku}}{\partial(\boldsymbol{\beta}',\boldsymbol{\pi}')} = r_{ik}\lambda_{iku}\begin{pmatrix}1\\ \boldsymbol{\delta}_u - \boldsymbol{p}_{ik}\end{pmatrix}'\otimes\boldsymbol{x}_{ik}',$$

where $\boldsymbol{p}_{ik} = (p_{ik1},p_{ik2},\dots,p_{ikD^r})'$ and $\otimes$ is the Kronecker product.

The covariance matrix of $\nabla Q^r(\boldsymbol{\beta},\boldsymbol{\pi})$, which is also the negative expected value of $\frac{\partial\nabla Q^r(\boldsymbol{\beta},\boldsymbol{\pi})}{\partial(\boldsymbol{\beta}',\boldsymbol{\pi}')}$, is

$$I^r(\boldsymbol{\beta},\boldsymbol{\pi}) = \frac{1}{\phi}X^{r\prime}\text{diag}(\boldsymbol{\lambda})^{-1}X^r. \tag{6}$$

The parameters $(\boldsymbol{\beta},\boldsymbol{\pi})$ are estimated by Newton–Raphson with Fisher scoring starting with initials $(\boldsymbol{\beta}^{\text{old}},\boldsymbol{\pi}^{\text{old}})$ and updating estimated parameters in the following way:

$$\begin{aligned}(\boldsymbol{\beta}^{\text{new}\prime},\boldsymbol{\pi}^{\text{new}\prime})' &= (\boldsymbol{\beta}^{\text{old}\prime},\boldsymbol{\pi}^{\text{old}\prime})'\\ &\quad + (X_0^{r\prime}\text{diag}(\boldsymbol{\lambda}_0)^{-1}X_0^r)^{-1}X_0^{r\prime}\text{diag}(\boldsymbol{\lambda}_0)^{-1}(\boldsymbol{N}^r - \boldsymbol{\lambda}_0),\end{aligned}$$

where X_0^r and $\boldsymbol{\lambda}_0$ are obtained by replacing $(\boldsymbol{\beta},\boldsymbol{\pi})$ with $(\boldsymbol{\beta}^{\text{old}},\boldsymbol{\pi}^{\text{old}})$. Write the estimated parameters as $(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\pi}})$. To estimate dispersion parameter ϕ, we adopt conventional method–moment estimation that is,

$$\hat{\phi} = \frac{1}{\sum_{u=0}^{D^r}\sum_{i=1}^{n-u}m_i - (D^r+1)p}$$

$$\times\sum_{u=0}^{D^r}\sum_{i=1}^{n-u}\sum_{k=1}^{m_i}\frac{(N_{iku}^r - r_{ik}\hat{\lambda}_{iku})^2}{r_{ik}\hat{\lambda}_{iku}},$$

where $\hat{\lambda}_{iku}$s are plug-in estimates of λ_{iku} that is

$$\hat{\lambda}_{iku} = \exp(\boldsymbol{x}_{ik}'\hat{\boldsymbol{\beta}})\cdot\frac{\exp(\boldsymbol{x}_{ik}'\hat{\boldsymbol{\pi}}_u)}{\sum_{j=0}^{D^r}\exp(\boldsymbol{x}_{ik}'\hat{\boldsymbol{\pi}}_j)}.$$

IDM considers that policy's feature information has no effect on reporting developments that is the coefficients of $x_1,x_2,\dots,x_{d-1}$ are thought to be zero. Obviously, IDM is a misspecified model if the feature information indeed influence those developments. Therefore, in IDM, λ and p_u are thought to keep fixed among all policies and then λ_u is same for the policies. By maximising function Q^r in (4) with respect to λ_u, one can obtain that

$$\hat{\lambda}_u = \frac{\tilde{N}_u^r}{r_{(u)}},\quad u = 0,1,\dots,D^r, \tag{7}$$

where $\tilde{N}_u^r = \sum_{i=1}^{n-u}\sum_{i=1}^{m_i}N_{iku}^r$ representing total number of reported claims with reporting delay u and $r_{(u)} = \sum_{i=1}^{n-u}\sum_{i=1}^{m_i}r_{ik}$ meaning total exposures in the first $n-u$ years.

3.2. Modelling settlement delays

In IIM, the settlement developments of individual claims after their reporting to the insurer have the following assumption. The assumption involves vectors of parameters $\boldsymbol{\rho}_1,\boldsymbol{\rho}_2,\dots,\boldsymbol{\rho}_{D^r}$, which are all d-dimensional vector.

Assumption 3.2: *Assume that given N_u^r, $(N_{u0},N_{u1},\dots,N_{uD^s})$ follows multinomial distribution with parameters N_u^r and $(q_0,\dots,q_{D^s})$, where*

$$q_v(\boldsymbol{\rho};\boldsymbol{x}) = \frac{\exp(\boldsymbol{x}'\boldsymbol{\rho}_v)}{\sum_{j=0}^{D^s}\exp(\boldsymbol{x}'\boldsymbol{\rho}_j)},\quad v = 0,\dots,D^s,$$

with $\boldsymbol{\rho}_0 = \boldsymbol{0}$, as well as $\boldsymbol{\rho}' = (\boldsymbol{\rho}_1',\boldsymbol{\rho}_2',\dots,\boldsymbol{\rho}_{D^s}')$ and the tuples $(N_{u0},N_{u1},\dots,N_{uD^s})$, $u = 0,\dots,D^r$ are independent.

Remark 3.2: Similar to the condition in Remark 3.1, the condition $n > D^s$ is necessary to make $\boldsymbol{\rho}$ be reasonably estimated. Therefore, it is enough to assume $n > \max(D^r,D^s)$.

For N_u^r ($u \le D_i^r$) reported claims of representative policy in year i, one can only observe $N_{u0},N_{u1},\dots,N_{u,D_{i+u}^s}$ and $N_u^{rbns} := \sum_{v=n-i-u+1}^{D^s}N_{uv}$ (the number of RBNS claims with settlement delays no less than $n-i-u$), where $N_u^{rbns} = 0$ if $u \le n-i-D^s$. According to the assumption above, the individual log-likelihood of settlement developments is

$$Q^{ios}(\boldsymbol{\rho}) = \sum_{v=0}^{D_i^s}\sum_{u=0}^{D_{i+v}^r} N_{uv}\log q_v + \sum_{u=0}^{D_i^r} N_u^{rbns}\log \bar{Q}_{n-i-u}, \tag{8}$$

where $\bar{Q}_v := \sum_{s=v+1}^{D^s} q_s$ is the tail probability of settlement delays no less than v. Obviously, an alternative form of term in the last term in the first line of (8) is $\sum_{u=(n-i-D^s+1)_+}^{D_i^r}$. Further, if we write $N_v^s = \sum_{u=0}^{D_{i+v}^r} N_{uv}$, which means number of settled claims with settlement delay v, (8) becomes

$$Q^{ios}(\boldsymbol{\rho}) = \sum_{v=0}^{D_i^s} N_v^s \log q_v + \sum_{u=0}^{D_i^r} N_u^{rbns}\log \bar{Q}_{n-i-u}. \tag{9}$$

To estimate $\boldsymbol{\rho}$ by Newton–Raphson with Fisher scoring, we need the identities in the following proposition.

Proposition 3.1: *The gradient of $Q^{ios}(\boldsymbol{\rho})$ with respect to $\boldsymbol{\rho}$ is*

$$\frac{\partial Q^{ios}(\boldsymbol{\rho})}{\boldsymbol{\rho}} = \left[\boldsymbol{N}_i^s + \sum_{u=0}^{D_i^r}\frac{N_u^{rbns}}{\bar{Q}_{n-i-u}}\begin{pmatrix}\boldsymbol{0}\\ \bar{\boldsymbol{q}}_{n-i-u}\end{pmatrix} - N^r\boldsymbol{q}\right]\otimes \boldsymbol{x},$$

and conditional expectation of Hessian matrix of $Q^{ios}(\boldsymbol{\rho})$ given $(r,\boldsymbol{x})$ is

$$\mathbb{E}\left[\frac{\partial^2 Q^{ios}(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}\boldsymbol{\rho}'}\middle| r,\boldsymbol{x}\right] = r\left[\sum_{u=0}^{n-i-D^s}\lambda_u(\mathrm{diag}(\boldsymbol{q})-\boldsymbol{q}\boldsymbol{q}') + \sum_{v=(n-i-D^r)_+}^{(D^s-1)\wedge(n-i)}\lambda_{n-i-v} \times\begin{pmatrix}\mathrm{diag}(\boldsymbol{q}_v)-\boldsymbol{q}_v\boldsymbol{q}_v' & -\boldsymbol{q}_v\bar{\boldsymbol{q}}_v'\\ -\bar{\boldsymbol{q}}_v\boldsymbol{q}_v' & \frac{Q_v}{\bar{Q}_v}\bar{\boldsymbol{q}}_v\bar{\boldsymbol{q}}_v'\end{pmatrix}\right]\otimes\boldsymbol{x}\boldsymbol{x}', \tag{10}$$

where $\boldsymbol{N}_i^s = (N_1^s, N_2^s,\ldots,N_{D_i^s}^s)'$, $N^r = \sum_{u=0}^{D_i^r} N_u^r$ and

$$\boldsymbol{q} = (q_1,q_2,\ldots,q_{D^s})',\quad \boldsymbol{q}_v = (q_1,q_2,\ldots,q_v)',$$
$$\bar{\boldsymbol{q}}_v = (q_{v+1},q_{v+2},\ldots,q_{D^s})'.$$

We estimate $\boldsymbol{\rho}$ by maximising overall log-likelihood function $Q^s(\boldsymbol{\rho})$ which is the summation of individual log-likelihood $Q_{ik}^{ios}(\boldsymbol{\rho})$, that is $\hat{\boldsymbol{\rho}}$ is obtained as follows:

$$\hat{\boldsymbol{\rho}} = \arg\max_{\boldsymbol{\rho}} Q^s(\boldsymbol{\rho}),$$

where

$$Q^s(\boldsymbol{\rho}) = \sum_{v=0}^{D^s}\sum_{i=1}^{n-v}\sum_{k=1}^{m_i} N_{ikv}^s\log q_{ikv} + \sum_{u=0}^{D^r}\sum_{i=(n-u-D^s)_++1}^{n-u}\sum_{k=1}^{m_i} N_{iku}^{rbns}\log\bar{Q}_{ik,n-i-u}.$$

To obtain $\hat{\boldsymbol{\rho}}$, similar as previous section, we use Newton–Raphson with Fisher scoring which needs the following gradients $\nabla Q^s(\boldsymbol{\rho})$ and its covariance matrix $I^s(\boldsymbol{\rho})$, where

$$\nabla Q^s(\boldsymbol{\rho}) := \frac{\partial Q^s(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}} = \sum_{i=1}^{n}\sum_{k=1}^{m_i}\frac{\partial Q_{ik}^{ios}(\boldsymbol{\rho})}{\boldsymbol{\rho}},$$
$$I^s(\boldsymbol{\rho}) := \sum_{i=1}^{n}\sum_{k=1}^{m_i}\mathbb{E}\left[\frac{\partial^2 Q_{ik}^{ios}(\boldsymbol{\rho})}{\partial\boldsymbol{\rho}\boldsymbol{\rho}'}\middle| r_{ik},\boldsymbol{x}_{ik}\right]. \tag{11}$$

In IDM, similar as λ_u in the section above, probabilities $(q_0,\ldots,q_{D^s})$ are thought to keep fixed among all policies that is $(q_0,\ldots,q_{D^s})$ is independent of $\boldsymbol{x}$. By MLE again, we have

$$\hat{q}_0 = \hat{h}_0 \text{ and } \hat{q}_v = \hat{h}_v\prod_{s=0}^{v-1}(1-\hat{h}_s),\quad v=0,1,\ldots,D^s, \tag{12}$$

where $\hat{h}_v = \frac{\tilde{N}_v^s}{\sum_{t=v}^{D^s}\tilde{N}_t^s + \sum_{t=v+1}^{D^s} G_t}$ with

$$\tilde{N}_v^s = \sum_{i=1}^{n-v}\sum_{k=1}^{m_i} N_{ikv}^s \text{ and } vG_v = \sum_{i=1}^{n-v+1}\sum_{k=1}^{m_i} N_{ik,n-i-v+1}^{rbns}.$$

3.3. Modelling claim payments

We give some assumptions about payments of individual claims as follows. The assumptions involve a $(d+D^r+D^s)$-dimensional vector of parameters $\boldsymbol{\gamma}$.

Assumption 3.3: *Claim payments Y_{uvl}, $u=0,\ldots,D^r$, $v=0,\ldots,D^s$, $l=1,\ldots,N_{uv}$ are independent, independent of N_{uv}; $u=0,\ldots,D^r$, $v=0,\ldots,D^s$ and also assume that conditional mean and variance satisfy*

$$\mathbb{E}[Y_{uvl}|\boldsymbol{x}] = \mu_{uv}, \mathrm{Var}(Y_{uvl}|\boldsymbol{x}) = \phi^p\mu_{uv}$$

with $\mu_{uv} = \exp(\boldsymbol{x}_{uv}'\boldsymbol{\gamma})$, where $\boldsymbol{x}_{uv} = (\boldsymbol{x}',\boldsymbol{\delta}_u',\boldsymbol{\delta}_v')'$ is a $(d+D^r+D^s)$-dimensional vector of covariates.

Arrange all settled payments of the risk portfolio into the set $\{(Y_l,\tilde{\boldsymbol{x}}_l), l=1,2,\ldots,N^{ts}\}$, where $\tilde{\boldsymbol{x}}_l$ is covariate associated with payments Y_l and N^{ts} is the total number of settled claims. Construct quasi-likelihood by independence among policies and assumption above,

$$Q^p(\boldsymbol{\gamma}) = \frac{1}{\phi^p}\sum_{l=1}^{N^{ts}}(Y_l\log\mu_l - \mu_l), \tag{13}$$

where $\mu_l = \exp(\tilde{\boldsymbol{x}}_l'\boldsymbol{\gamma})$. Denote $\boldsymbol{\mu} = (\mu_1,\ldots,\mu_{N^{ts}})'$ and $\boldsymbol{Y} = (Y_1,\ldots,Y_{N^{ts}})'$. The quasi-score function–

partial derivatives of $Q^p(\boldsymbol{\gamma})$ with respect to the parameters is

$$\dot{Q}^p(\boldsymbol{\gamma}) := \frac{\partial Q^p(\boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = \frac{1}{\phi^p}\tilde{X}'(\boldsymbol{Y}-\boldsymbol{\mu}), \qquad (14)$$

where $\tilde{X} = (\tilde{\boldsymbol{x}}_1, \ldots, \tilde{\boldsymbol{x}}_{N^{ts}})'$.

The covariance matrix of $\dot{Q}(\boldsymbol{\gamma})$, which is also the negative expected value of $\partial\dot{Q}(\boldsymbol{\gamma})/\partial\boldsymbol{\gamma}'$, is

$$I^p = \frac{1}{\phi^p}\tilde{X}'\text{diag}(\boldsymbol{\mu})\tilde{X}. \qquad (15)$$

The parameters $\boldsymbol{\gamma}$ are estimated by iteratively reweighted least square (IRLS) algorithm, which is as follows,

(1) Initialise $\hat{\boldsymbol{\gamma}} = \boldsymbol{\gamma}_0$ such that $\hat{\mu}_l = \exp(\tilde{x}_l'\hat{\boldsymbol{\gamma}})$ and $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \hat{\mu}_2, \ldots, \hat{\mu}_{N^{ts}})$, where $\boldsymbol{\gamma}_0$ is usually zero vector.
(2) Compute adjusted response $z_l = y_l - \hat{\mu}_l + \tilde{X}\hat{\boldsymbol{\gamma}}$.
(3) Update $\hat{\boldsymbol{\gamma}}$ by what follows,

$$\hat{\boldsymbol{\gamma}} = (\tilde{X}'\text{diag}(\hat{\boldsymbol{\mu}})\tilde{X})^{-1}\tilde{X}'\text{diag}(\hat{\boldsymbol{\mu}})Z,$$

where $Z = (z_1, z_2, \ldots, z_{N^{ts}})$, and then $\hat{\mu}_l = \exp(\tilde{x}_l'\hat{\boldsymbol{\gamma}})$.

To estimate dispersion parameter ϕ^p, we also adopt conventional method–moment estimation that is,

$$\hat{\phi}^p = \frac{1}{N^{ts} - (p + D^r + D^s)} \sum_{l=1}^{N^{ts}} \frac{(Y_l - \hat{\mu}_l)^2}{\hat{\mu}_l},$$

where $\hat{\mu}_l = \exp(\tilde{\boldsymbol{x}}_l'\hat{\boldsymbol{\gamma}})$.

In IDM, the coefficients of covariates about individual features are considered to be zero, i.e., $\gamma_1 = \cdots = \gamma_{d-1} = 0$, and μ_ls only depended on reporting and settlement delays, which means it just needs to estimate $\boldsymbol{\gamma}^{ID} := (\gamma_0, \gamma_d, \ldots, \gamma_{d-1+D^r}, \ldots, \gamma_{d-1+D^r+D^s})'$ by the similar procedure as stated above. Therefore, estimator $\hat{\boldsymbol{\gamma}}^{ID}$ for $\boldsymbol{\gamma}^{ID}$ is a maximiser of the Q^p, which is the function of $\boldsymbol{\gamma}^{ID}$ that is under $\gamma_1 = \cdots = \gamma_{d-1} = 0$, and μ_l in Equation (13) is independent of individual information and only takes one of the following forms:

$$\mu_{uv}^{ID} = \exp((1, \boldsymbol{\delta}_u', \boldsymbol{\delta}_v')\boldsymbol{\gamma}^{ID}),$$
$$u = 0, 1, \ldots, D^r, \quad v = 0, 1, \ldots, D^s. \qquad (16)$$

Then the estimate of μ_{uv}^{ID} under IDM is denoted by $\hat{\mu}_{uv} := \exp((1, \boldsymbol{\delta}_u', \boldsymbol{\delta}_v')\hat{\boldsymbol{\gamma}}^{ID})$, which is a policy-free estimate.

4. Prediction for outstanding liabilities

In this section, the terminologies 'loss reserve' and 'loss reserving' are precisely specified, measurement of accuracy of loss reserving is then discussed and we also shows the improvement of accuracy of loss reserving basing on IIM with respect to IDM.

4.1. Loss reserve and loss reserving

Recalling the total outstanding liability R defined in (3), by 'loss reserve', we refer to the projection

$$R_m = R_m(\boldsymbol{\theta}) = \mathbb{E}[R|\mathscr{F}^{uo}] \qquad (17)$$

of R on the observations $\mathscr{F}^{uo}$ by the evaluation date n, where the subscript 'm' indicates portfolio size, since loss reserve is based on specific risk portfolio. One can see that R_m is a function of unknown parameters $\boldsymbol{\theta} := (\boldsymbol{\beta}', \boldsymbol{\pi}', \boldsymbol{\rho}', \boldsymbol{\gamma}')'$ and hence it needs to be estimated.

To derive moments about outstanding liabilities R and conditional variance of R, the following quantities are needed. For $u = 0, 1, \ldots, D^r, v = 0, 1, \ldots, D^s$, denote by

$$\tilde{\mu}_{uv} = \frac{\sum_{t=v}^{D^s} q_t\mu_{ut}}{\sum_{t=v}^{D^s} q_t} \quad \text{and} \quad \tilde{\mu}_{uv}^s = \frac{\sum_{t=v}^{D^s} q_t\mu_{ut}^2}{\sum_{t=v}^{D^s} q_t}, \qquad (18)$$

where $\tilde{\mu}_{uv}$ is conditional moment of claim payments given $\boldsymbol{x}$, reporting delays u and settlement delays no less than v, so that corresponding policy-specified quantities are

$$\tilde{\mu}_{ikuv} = \frac{\sum_{t=v}^{D^s} q_{ikt}\mu_{ikut}}{\sum_{t=v}^{D^s} q_{ikt}} \text{ and } \tilde{\mu}_{ikuv}^s = \frac{\sum_{t=v}^{D^s} q_{ikt}\mu_{ikut}^2}{\sum_{t=v}^{D^s} q_{ikt}}. \qquad (19)$$

Then we derive the following theorem which provides formulas to compute not only the loss reserve R_m but also variance of outstanding liabilities R given observations $\mathscr{F}^{uo}$.

Theorem 4.1: *Under the model formulated by Assumptions* 3.1–3.3, *the loss reserve is*

$$R_m(\boldsymbol{\theta}) = \sum_{v=1}^{D^s}\sum_{u=0}^{D_v^r}\sum_{k=1}^{m_{n-v-u+1}} N_{(n-v-u+1)ku}^{rbns}\tilde{\mu}_{(n-v-u+1)kuv} + \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}\lambda_{iku}\tilde{\mu}_{iku0}, \qquad (20)$$

and the variance of R given observations $\mathscr{F}^{uo}$ *is*

$$\text{Var}(R \mid \mathscr{F}^{uo}) = \sum_{v=1}^{D^s}\sum_{u=0}^{D_v^r}\sum_{k=1}^{m_{n-v-u+1}} N_{(n-v-u+1)ku}^{rbns} \times \left(\frac{\tilde{\mu}_{n-v-u+1,kuv}^s}{\bar{Q}_{n-v-u+1,k,v-1}} - \tilde{\mu}_{n-v-u+1,kuv}^2 + \phi^p\tilde{\mu}_{n-v-u+1,kuv}\right) + \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}\lambda_{iku} \times (\tilde{\mu}_{iku0}^s + (\phi - 1)\tilde{\mu}_{iku0}^2 + \phi^p\tilde{\mu}_{iku0}).$$

It can be clearly seen that loss reserve R_m depends on not only the information from observed data in terms of the number of RBNS claims and policy's feature information but also unknown parameters $\boldsymbol{\theta}$, which results in the need for estimating R_m. Accordingly, the term 'loss/claims reserving' is used for certain reasonable estimate of the loss reserve. Formally, after getting certain reasonable estimates $\hat{\boldsymbol{\theta}}$ of the unknown parameters from the observed data, as, for example, what has been done in the previous section, we have the following theorem.

Theorem 4.2: *By loss reserving we refer to the (random) quantity*

$$\hat{R}_{II} = \sum_{v=1}^{D^s}\sum_{u=0}^{D^r_v}\sum_{k=1}^{m_{n-v-u+1}} N^{rbns}_{(n-v-u+1)ku}\hat{\hat{\mu}}_{(n-v-u+1)kuv} + \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}\hat{\lambda}_{iku}\hat{\hat{\mu}}_{iku0}, \quad (21)$$

where $\hat{\hat{\mu}}_{ikuv}$s and $\hat{\lambda}_{iku}$s are obtained by substituting unknown parameters with their estimates.

According to the theorem above, it is easy to obtain loss reserving under IDM by simply replacing $\hat{\mu}_{ikuv}$s and $\hat{\lambda}_{iku}$s in (21) with policy-free estimates $\hat{\mu}_{uv}$s and $\hat{\lambda}_u$s, respectively. Specifically, to distinguish two different estimates for reserve, we use symbol $\hat{R}_{ID}$ to indicate loss reserving under IDM, which is

$$\hat{R}_{ID} = \sum_{v=1}^{D^s}\sum_{u=0}^{D^r_v}\tilde{N}^{rbns}_{n-v-u+1,u}\hat{\hat{\mu}}_{uv} + \sum_{u=1}^{D^r} r_{[u]}\hat{\lambda}_u\hat{\hat{\mu}}_{u0}, \quad (22)$$

where $\tilde{N}^{rbns}_{n-v-u+1,u} = \sum_{k=1}^{m_{n-v-u+1}} N^{rbns}_{n-v-u+1,ku}$, $r_{[u]} = \sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}$ and $\hat{\hat{\mu}}_{uv} = \frac{\sum_{t=v}^{D^s}\hat{q}_t\hat{\mu}_{ut}}{\sum_{t=v}^{D^s}\hat{q}_t}$.

4.2. Measurement of prediction accuracy

It is essential to measure accuracy of loss reserving and especially accuracy improvement of loss reserving by considering useful individual information with respect to the one without this information. To measure the prediction accuracy of some reserve estimate $\hat{R}$, which is $\mathscr{F}^{uo}$ measurable, a natural idea is conditional mean square error of prediction (MSEP) which is defined as

$$\begin{aligned}\text{MSEP}(R,\hat{R}) &= \mathbb{E}[(R-\hat{R})^2|\mathscr{F}^{uo}] \\ &= \text{Var}(R|\mathscr{F}^{uo}) + (\mathbb{E}[R|\mathscr{F}^{uo}] - \hat{R})^2. \quad (23)\end{aligned}$$

For loss reserving $\hat{R}_{II}$, which includes individual information, and $\hat{R}_{ID}$ without individual information, their MSEPs are MSEP$(R,\hat{R}_{II})$ and MSEP$(R,\hat{R}_{ID})$, respectively. To measure the difference in prediction accuracy of $\hat{R}_{II}$ and $\hat{R}_{ID}$, we use the following ratio:

$$\begin{aligned}M^r &= \frac{\text{MSEP}(R,\hat{R}_{II})}{\text{MSEP}(R,\hat{R}_{ID})} \\ &= \frac{\text{Var}(R|\mathscr{F}^{uo}) + (\mathbb{E}[R|\mathscr{F}^{uo}] - \hat{R}_{II})^2}{\text{Var}(R|\mathscr{F}^{uo}) + (\mathbb{E}[R|\mathscr{F}^{uo}] - \hat{R}_{ID})^2}. \quad (24)\end{aligned}$$

It is well known that individual information model performs better in terms of prediction accuracy than individual data model, if $M^r < 1$, but it is hard to compute M^r with unknown parameters. Fortunately, we can compare M^r and number 1 when portfolio size m is large enough. It is notable that individual data model is nested in individual information model. Then we have the following theorem under some regular conditions (see Van der Vaart, 2000), which illustrates the advantages of individual information model over individual data model.

Theorem 4.3: *When portfolio size m tends to infinity, $M^r \xrightarrow{P} 1$, where $\xrightarrow{P}$ means converging in probability, if the individual data model is true, that is the coefficients of $x_1, x_2, \ldots, x_{d-1}$ are zero. Otherwise,*

$$\begin{aligned}\frac{1}{m}(\hat{R}_{ID} - R_m(\boldsymbol{\theta})) &\xrightarrow{P} \Delta \\ &= \sum_{v=1}^{D^s}\sum_{u=0}^{D^r_v}\kappa_{n-v-u+1}\mathbb{E}[r\lambda_u\bar{Q}_{v-1}(\check{\tilde{\mu}}_{uv} - \tilde{\mu}_{uv})] \\ &\quad + \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\kappa_i\mathbb{E}[r\lambda_u(\check{\tilde{\mu}}_{u0} - \tilde{\mu}_{u0})], \quad (25)\end{aligned}$$

where $\check{\tilde{\mu}}_{uv} = \sum_{s=v}^{D^s}\check{q}_s\check{\mu}_{us}/\sum_{s=v}^{D^s}\check{q}_s$ with

$$\check{q}_v = \check{h}_v\prod_{s=0}^{v-1}(1-\check{h}_s),$$

$$\check{h}_v = \frac{\sum_{i=1}^{n-v}\sum_{u=0}^{D^r_{i+v}}\kappa_i\mathbb{E}\left[r\lambda_u q_v\right]}{\sum_{i=1}^{n-v}\sum_{u=0}^{D^r_{i+v}}\kappa_i\mathbb{E}\left[r\lambda_u\bar{Q}_{v-1}\right]},$$

$$\check{\mu}_{uv} = \exp((1,\boldsymbol{\delta}'_u,\boldsymbol{\delta}'_v)\check{\boldsymbol{\gamma}}^{ID}), \quad \textit{and}$$

$$\check{\boldsymbol{\gamma}}^{ID} = \underset{\boldsymbol{\gamma}^{ID}}{\text{Argmax}}\sum_{u=0}^{D^r}\sum_{v=0}^{D^s}\sum_{i=1}^{n-u-v}\kappa_i\mathbb{E} \times [r\lambda_u q_v(\mu_{uv}\log\mu^{ID}_{uv} - \mu^{ID}_{uv})], \quad \mu^{ID}_{uv} \text{ in (16)}, \quad (26)$$

and if the asymptotic bias $\Delta \neq 0$, $M^r \xrightarrow{P} 0$.

The theorem above shows that IIM is asymptotically equivalent to IDM, if IDM is true and otherwise the former has higher prediction accuracy than the latter when portfolio size is large enough. One can intuitively understand that as portfolio size tends to infinity, both models can capture all the information included in observations when IDM holds true, since IIM is a generalised version of IDM. However, individual data model fails to capture the effects of policy's feature information and thus leads to greater bias when IIM holds true.

An important issue one concerns is how much prediction accuracy of loss reserving $\hat{R}_{II}$ can be improved, if IIM holds true, in a fixed risk portfolio that is one cares about actual value of M^r under true IIM. However, there are unknown parameters $\boldsymbol{\theta}$ in $\mathrm{Var}(R|\mathscr{F}^{uo})$ and $\mathbb{E}[R|\mathscr{F}^{uo}]$. An approximation method that comes to one's mind is substituting estimated parameters $\hat{\boldsymbol{\theta}}$ to them, which however needs to further take estimation error of $\hat{\boldsymbol{\theta}}$ into account. We directly use the method named semi-analytical approximation for $\mathrm{MSEP}(R, \hat{R})$ (One can refer to Lindholm et al. (2020) for more details), which is also discussed in Wahl (2019) under micro data model. Then the approximations for $\mathrm{MSEP}(R, \hat{R}_{II})$ and $\mathrm{MSEP}(R, \hat{R}_{ID})$ are

$$\begin{aligned}\widehat{\mathrm{MSEP}}(R, \hat{R}_{II}) &= \mathrm{Var}(R|\mathscr{F}^{uo})(\hat{\boldsymbol{\theta}}) \\ &\quad + \nabla R_m(\hat{\boldsymbol{\theta}})'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})\nabla R_m(\hat{\boldsymbol{\theta}}), \\ \widehat{\mathrm{MSEP}}(R, \hat{R}_{ID}) &= \mathrm{Var}(R|\mathscr{F}^{uo})(\hat{\boldsymbol{\theta}}) \\ &\quad + \nabla R_m(\hat{\boldsymbol{\theta}})'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})\nabla R_m(\hat{\boldsymbol{\theta}}) \\ &\quad + (\hat{R}_{II} - \hat{R}_{ID})^2, \end{aligned} \tag{27}$$

so that

$$\hat{M}^r = \frac{\mathrm{Var}(R\,|\,\mathscr{F}^{uo})(\hat{\boldsymbol{\theta}}) + \nabla R_m(\hat{\boldsymbol{\theta}})'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})\nabla R_m(\hat{\boldsymbol{\theta}})}{\mathrm{Var}(R|\mathscr{F}^{uo})(\hat{\boldsymbol{\theta}}) + \nabla R_m(\hat{\boldsymbol{\theta}})'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})\nabla R_m(\hat{\boldsymbol{\theta}}) + (\hat{R}_{II} - \hat{R}_{ID})^2}, \tag{28}$$

where $\nabla R_m(\hat{\boldsymbol{\theta}})$ is the gradient of loss reserve $R_m(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ computed at $\hat{\boldsymbol{\theta}}$ and $\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}})$ is asymptotic covariance of $\hat{\boldsymbol{\theta}}$. It is easily known that

$$\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\theta}}) = \mathrm{diag}(\hat{\phi}(\hat{X}^{r\prime}\mathrm{diag}(\hat{\boldsymbol{\lambda}})^{-1}\hat{X}^r)^{-1}, (I^s(\hat{\boldsymbol{\rho}}))^{-1}, \hat{\phi}^p(\tilde{X}'\tilde{X})^{-1}),$$

where $\hat{X}^r$ and $\hat{\boldsymbol{\lambda}}$ are plug-in estimates and $I^s(\hat{\boldsymbol{\rho}})$ is obtained by inserting $\hat{\boldsymbol{\rho}}$ into $I^s(\boldsymbol{\rho})$ in Equation (11). One can refer to Chapter 9 in McCullagh and Nelder (1989) for more details.

Proposition 4.4: *The gradient of $R_m(\boldsymbol{\theta})$ with respect to $(\boldsymbol{\beta}', \boldsymbol{\pi}')'$ is*

$$\frac{\partial R_m(\boldsymbol{\theta})}{\partial(\boldsymbol{\beta}', \boldsymbol{\pi}')'} = \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}\lambda_{iku}\tilde{\mu}_{iku0}\begin{pmatrix}1\\ \boldsymbol{\delta}_u - \boldsymbol{p}_{ik}\end{pmatrix}\otimes \boldsymbol{x}_{ik},$$

the gradient of $R_m(\boldsymbol{\theta})$ with respect to $\boldsymbol{\rho}$ is

$$\begin{aligned}\frac{\partial R_m(\boldsymbol{\theta})}{\partial \boldsymbol{\rho}} &= \sum_{v=1}^{D^s}\sum_{u=0}^{D^r_v}\sum_{k=1}^{m_{i_{uv}}} \frac{N^{rbns}_{i_{uv}ku}}{\bar{Q}_{i_{uv}k,v-1}}\begin{pmatrix}\mathbf{0}\\ \boldsymbol{q}^{mu}_{iku,v-1}\end{pmatrix}\otimes \boldsymbol{x}_{i_{uv}k} \\ &\quad + \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}\lambda_{iku} \\ &\quad \times [\mathrm{diag}(\boldsymbol{q}_{ik})\overline{\boldsymbol{\mu}}_{iku0} - \tilde{\mu}_{iku0}\boldsymbol{q}_{ik}]\otimes \boldsymbol{x}_{ik},\end{aligned}$$

where $i_{uv} = n - v - u + 1$,

$$\boldsymbol{q}^{mu}_{ikuv} = \mathrm{diag}(\overline{\boldsymbol{q}}_{ikv})\overline{\boldsymbol{\mu}}_{ikuv} - \tilde{\mu}_{iku,v+1}\boldsymbol{q}_{ikv},$$

and $\overline{\boldsymbol{\mu}}_{ikuv} = (\mu_{iku,v+1}, \dots, \mu_{ikuD^s})'$, and the gradient of $R_m(\boldsymbol{\theta})$ with respect to $\boldsymbol{\gamma}$ is

$$\begin{aligned}\frac{\partial R_m(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} &= \sum_{v=1}^{D^s}\sum_{u=0}^{D^r_v}\sum_{k=1}^{m_{n-v-u+1}} N^{rbns}_{(n-v-u+1)ku} \\ &\quad \times \dot{\tilde{\mu}}_{(n-v-u+1)kuv} \\ &\quad + \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} r_{ik}\lambda_{iku}\dot{\tilde{\mu}}_{iku0},\end{aligned}$$

where $\dot{\tilde{\mu}}_{ikuv} = \frac{\sum_{t=v}^{D^s} q_{ikt}\mu_{ikut}\boldsymbol{x}_{ikut}}{\sum_{t=v}^{D^s} q_{ikt}}$.

5. Simulations and real data analysis

Reported in this section include the results from a few small simulations conducted to further investigate M^r. A real data in health insurance was also analysed to show the application of IIM and the accuracy improvement by using IIM with respect to IDM in practice.

5.1. Simulation

In this simulation, the risk exposures associated with every individuals were drawn from the uniform distribution on $[0, 1]$, the covariates were produced by multivariate standard normal distribution and we simulated the random developments of claims for a fixed risk portfolio. In each run, we directly compute M^r according to Equation (24) so that we can know how much accuracy is improved by using IIM with respect to IDM under the fixed risk portfolio.

Because there are only assumptions about mean and variance for reporting developments and payments of claims, we need additional distributional assumptions to generate them, which arise as follows. First, for individual reporting developments N^r_us, we generated them by the additional assumption which says that $\frac{N^r_u}{\phi}$ follows Poisson distribution with mean $\frac{r\lambda_u}{\phi}$. Second, for individual payments Y_l, similarly, we generated them by assuming that $\frac{Y_l}{\phi^p}$ follows Poisson distribution with mean $\frac{\mu_l}{\phi^p}$. Each run in the simulation was conducted

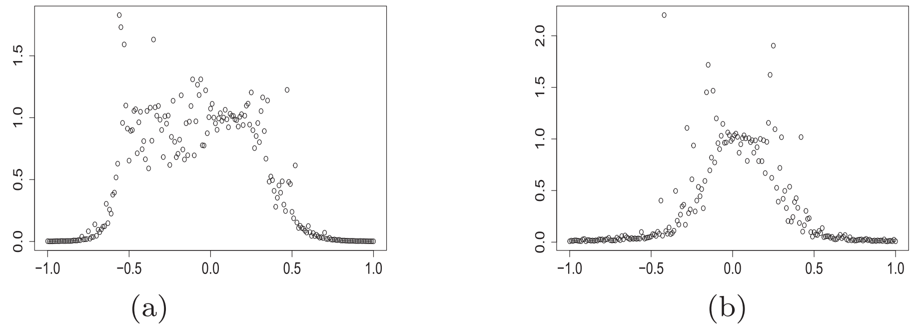

Figure 1. The simulated M^r over varying coefficients of covariates. (a) Example 5.1 and (b) Example 5.2.

with the setting: $n = 5, D^r = 2, D^s = 2$, a risk portfolio size $\mathbf{m} = (10000, 10,000, 10,000, 10,000, 10,000)$, i.e., 10, 000 policies in each year, and any combination of parameters which varied according to the setting in the following two examples.

Example 5.1: Dimension $d = 3$ and the parameters varied in an auxiliary parameter t ranging in $[-1, 1]$ by step 0.01 as

- Parameters for reporting developments:

$$\boldsymbol{\beta} = (-0.5, -t, 2t)', \boldsymbol{\pi}_1 = (1, t, t)',$$
$$\boldsymbol{\pi}_2 = (-1, -t, -2t)',$$

  and $\phi = 2$.
- Parameters for settlement developments: $\boldsymbol{\rho}_1 = (0.1, 0.2t, -0.3t)'$ and $\boldsymbol{\rho}_2 = (-0.1, -0.2t, 0.3t)'$.
- Parameters for payments: $\boldsymbol{\gamma} = (5, 0.2t, 0.4t, 0.1, 0.6, 0.2, 0.8)'$ and $\phi^p = 1.5$.

Covariates were produced by bivariate standard normal distribution in this example.

Example 5.2: Dimension $d = 4$ and parameters varied over t ranging in $[-1, 1]$ by step 0.01 as

- Parameters for reporting developments:

$$\boldsymbol{\beta} = (2, 0.2t, -0.8t, 0.5t)',$$
$$\boldsymbol{\pi}_1 = (2, -t, 3t, -2t)', \boldsymbol{\pi}_2 = (1, 2t, -t, -2t)',$$

  and $\phi = 3$.
- Parameters for settlement developments:

$$\boldsymbol{\rho}_1 = (0.3, 0.1t, -0.5t, 0.2t)',$$
$$\boldsymbol{\rho}_2 = (-0.2, -0.3t, 0.7t, 0.4t)'.$$

- Parameters for payments: $\boldsymbol{\gamma} = (3, 0.6t, -0.2t, 0.7t, 0.3, 0.2, -0.5, 0.4)'$ and $\phi^p = 2.5$.

Covariates were produced by ternary standard normal distribution in this example.

Table 1. The individual information in real data analysis.

Covariate	Description	Type	Levels
x_1	Insured's age	Quantitative	
x_2	Insured's gender	Binary	female: 1, male: 0
x_3	Policy type	Binary	general health insurance: 1 serious illness insurance: 0
$x_4 - x_8$	Geographical location	Categorical	regions I-V: one-hot encoding with $(x_4, x_5, \ldots, x_8)$ so that region VI: $x_4 = \cdots = x_8 = 0$

In each run, we estimated loss reserve by both IIM and IDM using the simulated data that is we computed $\hat{R}_{II}$ by Equation (21) as well as $\hat{R}_{ID}$ by (22) and true parameters were used to compute $\mathrm{Var}(R \mid \mathscr{F}^{uo})$ and $\mathbb{E}[R \mid \mathscr{F}^{uo}]$ according to Theorem 4.1. Then we computed M^r by inserting the computed $\hat{R}_{II}$, $\hat{R}_{ID}$, $\mathrm{Var}(R \mid \mathscr{F}^{uo})$ and $\mathbb{E}[R \mid \mathscr{F}^{uo}]$ into Equation (24). At last, we plotted the simulated results in Figure 1.

We obtained the results consistent with Theorem 4.3 from the simulations above.

(1) When the coefficients of $x_1, x_2, \ldots, x_{d-1}$ approach zero, most M^rs are close to real number 1 that is loss reserving by IIM almost has the same accuracy as that by IDM.
(2) When those coefficients are away from zero, M^r tends to be zero that is the prediction accuracy of loss reserving by IIM is greatly improved with respect to IDM.

5.2. Real data analysis

In this section, we analysed a dataset, which was collected by a commercial insurance company in China. The dataset recorded writing and expiring dates of policies, individual information, see Table 1, and developments of reported claims between 1/1/2019 and 8/31/2019.

To visualise the effects of individual information on the developments of claims, for example, the histograms of reporting and settlement delays measured in

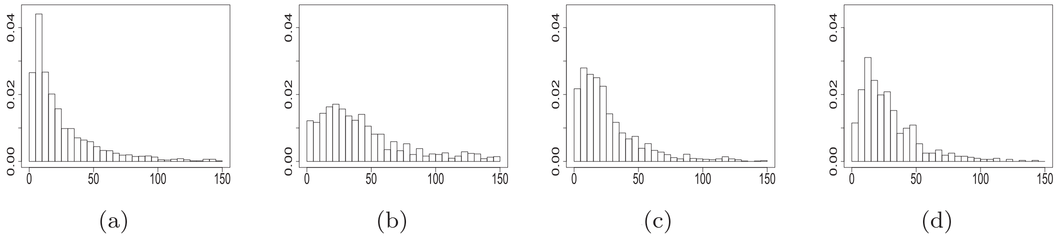

Figure 2. Histograms of reporting delays (in days): (a) Female, Region III, age 9–20; (b) Male, Region I, age 45–50; (c) Male, Region VI, age 20–40; (d) Male, Region III, age > 55.

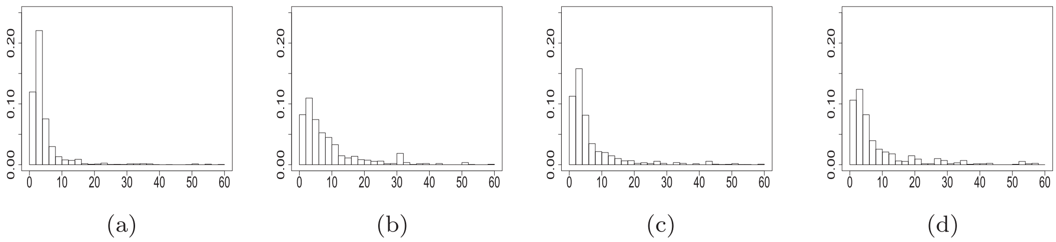

Figure 3. Histograms of settlement delays (in days): (a) Female, Region III, age 9–20; (b) Male, Region I, age 45–50; (c) Male, Region VI, age 20–40; (d) Male, Region III, age > 55.

days were provided under a few combinations of covariate values including gender, geographical location and age, as presented in Figures 2 and 3. It was strongly proposed that the individual information had impacts on the distributions of reporting and settlement delays.In the dataset, all the reporting delays were not more than 150 days (5 months). By China Banking and Insurance Regulatory Commission, the reported claims in health insurance are generally required to be settled within 2 months if no disagreement exists. It is appropriate to take 1 month as the time unit ('accident year' in previous sections). Thus the maximum reporting and settlement delays were safely set to $D^r = 5$ and $D^s = 3$ (the real data supported this assumption).

Table 2. Estimated parameters for reporting developments, their standard errors and *p*-values.

	$\boldsymbol{\beta}$			$\boldsymbol{\pi}_1$		
	Estimate	Std. error	*p*-value	Estimate	Std. error	*p*-value
Intercept	−4.6083	0.0078	0.0000	0.8135	0.0158	0.0000
x_1	0.0307	0.0001	0.0000	0.0044	0.0003	0.0000
x_2	0.0979	0.0109	0.0000	−0.0359	0.0221	0.1044
x_3	1.2725	0.0114	0.0000	−0.1032	0.0231	0.0000
x_4	0.7215	0.0233	0.0000	−0.0277	0.0466	0.5523
x_5	0.6734	0.0327	0.0000	−0.1187	0.0654	0.0697
x_6	0.5786	0.0121	0.0000	−0.6239	0.0244	0.0000
x_7	0.3889	0.0145	0.0000	−0.5281	0.0292	0.0000
x_8	0.3114	0.3114	0.0000	−0.3829	0.0532	0.0000
	$\boldsymbol{\pi}_2$			$\boldsymbol{\pi}_3$		
Intercept	−0.0198	0.0224	0.3762	−0.8734	0.0358	0.0000
x_1	0.0050	0.0005	0.0000	0.0064	0.0008	0.0000
x_2	−0.0200	0.0314	0.5239	−0.0716	0.0507	0.1575
x_3	0.0011	0.0322	0.9712	−0.0313	0.0515	0.5433
x_4	0.0685	0.0581	0.2385	0.0542	0.0917	0.5541
x_5	−0.1122	0.0831	0.1769	−0.4597	0.1543	0.0028
x_6	−0.8899	0.0361	0.0000	−0.7543	0.0549	0.0000
x_7	−0.8522	0.0436	0.0000	−0.8407	0.0703	0.0000
x_8	−0.5350	0.0734	0.0000	−0.7509	0.1305	0.0000
	$\boldsymbol{\pi}_4$			$\boldsymbol{\pi}_5$		
Intercept	−1.6427	0.0525	0.0000	−2.2198	0.0768	0.0000
x_1	0.0057	0.0012	0.0000	0.0090	0.0017	0.0000
x_2	−0.0404	0.0739	0.5843	0.1614	0.1028	0.1164
x_3	0.0716	0.0734	0.3293	0.0059	0.1096	0.9570
x_4	0.4313	0.1220	0.0004	0.1285	0.2004	0.5213
x_5	−0.5048	0.2568	0.0494	−0.1750	0.3074	0.5691
x_6	−0.5444	0.0801	0.0000	−0.5695	0.1152	0.0000
x_7	−0.7191	0.1079	0.0000	−0.8211	0.1620	0.0000
x_8	−0.4864	0.1865	0.0091	−0.3980	0.2528	0.1153

To illustrate the proposed model for loss reserving, evaluation date was set as 8/31/2019. That is, we worked with $n = 8$, $D^r = 5$ and $D^s = 3$ (months). There are four factors organised into eight features $x_1, \ldots, x_8$, as shown in Table 1. Besides, reporting and settlement delays, which were regarded as factors to model claim payments as Assumption 3.2 formulated, were respectively organised into five features $x_9, x_{10}, \ldots, x_{13}$ and three features x_{14}, x_{15}, x_{16}.

The estimated parameters for the reporting developments under IIM, their standard errors and p-values of significance test were displayed in Table 2, while the corresponding estimated results under IDM, i.e., $\hat{\lambda}_u$, $u = 0, 1, \ldots, 5$ in (7) are

$$(0.0218, 0.0341, 0.0133, 0.0058, 0.0034, 0.0022),$$

respectively. Besides, the estimated dispersion parameter $\hat{\phi} = 1.9433$. These results in (2) provide obvious evidence that individual information has effects on the reporting developments of claims in sense that most covariates associated with individual information are significant at significance level 0.05.

Similar results for settlement developments and payments are listed in Tables 3 and 4. These results also provide obvious evidence that individual information has effects on settlement developments and payments of claims. Besides, the estimated dispersion parameter $\hat{\phi}^p = 15467.1$ and the estimates under IDM are

$$(\hat{q}_0, \hat{q}_1, \hat{q}_2, \hat{q}_3) = (0.6435, 0.3121, 0.0363, 0.0081),$$
$$\hat{\gamma}^{ID} = (8.3055, 0.4535, 0.5830, 0.6227, 0.58380.4557, 0.4204, 0.6139, 0.8472)'.$$

In Table 5, the columns with names 'IBNR', 'RBNS' and 'Loss reserving' correspond to estimates of IBNR reserve, RBNS reserve and total loss reserve, respectively. The square roots of approximated conditional

Table 3. Estimated parameters for settlements developments, their standard errors and p-values.

	ρ_1			ρ_2		
	Estimate	Std. error	p-value	Estimate	Std. error	p-value
Intercept	−0.9013	0.0124	0.0000	−2.2347	0.0354	0.0000
x_1	0.0046	0.0002	0.0000	0.0041	0.0008	0.0000
x_2	−0.0307	0.0174	0.0783	−0.1307	0.0510	0.0104
x_3	−0.0246	0.0182	0.1765	0.4631	0.0461	0.0000
x_4	0.0475	0.0368	0.1971	−0.9837	0.1058	0.0000
x_5	0.0230	0.0511	0.6529	−1.2136	0.1596	0.0000
x_6	0.0414	0.0192	0.0311	−0.8547	0.0509	0.0000
x_7	−0.0353	0.0232	0.1281	−1.2330	0.0749	0.0000
x_8	0.0831	0.0412	0.0439	−1.3182	0.1414	0.0000
		ρ_3				
Intercept	−4.8241	0.0809	0.0000			
x_1	0.0011	0.0019	0.5482			
x_2	−0.1864	0.1202	0.1210			
x_3	−0.3856	0.1295	0.0029			
x_4	0.8514	0.2535	0.0007			
x_5	−1.7351	0.9992	0.0824			
x_6	1.0403	0.1047	0.0000			
x_7	0.3729	0.1047	0.0222			
x_8	−0.2783	0.4388	0.5259			

Table 4. Estimated parameters $\hat{\gamma}$ for payments, their standard errors and p-values.

	Estimate	Std. error	p-value
Intercept	9.3727	0.0088	0.0000
x_1	−0.0030	0.0002	0.0000
x_2	−0.1756	0.0132	0.0000
x_3	−2.0128	0.0311	0.0000
x_4	−0.0642	0.0233	0.0058
x_5	−0.6621	0.0409	0.0000
x_6	−0.4123	0.0150	0.0000
x_7	−0.2496	0.0151	0.0000
x_8	−0.3536	0.0306	0.0000
x_9	0.2062	0.0135	0.0000
x_{10}	0.7433	0.0183	0.0000
x_{11}	0.7745	0.0311	0.0000
x_{12}	0.6025	0.0529	0.0000
x_{13}	0.2796	0.0950	0.0032
x_{14}	0.3586	0.0155	0.0000
x_{15}	0.5814	0.0461	0.0000
x_{16}	0.5057	0.1174	0.0000

Table 5. Reserving, accuracy of prediction and accuracy improvement of IIM with respect to IDM.

Model	IBNR	RBNS	Loss reserving	$\sqrt{\widehat{\text{MSEP}}}$	$\hat{M}^r$
IIM	56393880	24858086	81251966	2784005	0.2237
IDM	58017529	28420566	86438096	5886139	

MSEPs under IIM and IDM are in the fourth column of Table 5. The rightmost column in this table showed the computed $\hat{M}^r$ by (28). We can see that loss reserving by IIM provides more stable prediction of outstanding liabilities than that by IDM since the former has smaller conditional MSEP and after incorporating useful individual information into loss reserving, the prediction accuracy is greatly increased by 77.63%.

6. Conclusion

This paper explored the improvement of accuracy in predicting outstanding liabilities, which are incurred by general insurance companies, by incorporating useful individual information into modelling. The reporting developments and payments of individual claims were given weak assumptions about their first two moments and modelled under quasi-likelihood theory, while settlement delays were modelled by multinomial logistic regression. Based on the model specification, loss reserve and conditional variance of outstanding liabilities were derived, which were further used to compute loss reserving and conditional MSEP. It was theoretically proved that loss reserving incorporating useful individual information shows higher accuracy than that under IDM, where the accuracy is measured by the conditional MSEP, when portfolio size is large enough. The conclusion is also supported by the simulations and real data analysis.

While the proposed model is basically a parametric model in statistical context, some one may be concerned with the limitation that the model is subjective and thus question its robustness in practical applications. Regarding this aspect, a possible next step is to study this problem under a nonparametric framework. Especially, it is more interesting to model the dependence of claims development on individual information by machine learning (including deep learning).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Natural Science Foundation of China (71771089), the Shanghai Philosophy and Social Science Foundation (2015BGL001), the National Social Science Foundation Key Program of China (17ZDA091) and China Scholarship Council (201906140045).

Notes on contributors

Zhigao Wang is a Ph.D. candidate in Statistics at East China Normal University.

Xianyi Wu is a professor in School of Statistics at East China Normal University.

Chunjuan Qiu is an associate professor in School of Statistics at East China Normal University.

References

Alai, D. H., Merz, M., & Wüthrich, M. V. (2009). Mean square error of prediction in the Bornhuetter–Ferguson claims reserving method. *Annals of Actuarial Science*, *4*(1), 7–31. https://doi.org/10.1017/S1748499500000580

Alai, D. H., Merz, M., & Wüthrich, M. V. (2010). Prediction uncertainty in the Bornhuetter–Ferguson claims reserving method: Revisited. *Annals of Actuarial Science*, *5*(1), 7–7. https://doi.org/10.1017/S1748499510000023

Antonio, K., & Plat, R. (2014). Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, *2014*(7), 649–669. https://doi.org/10.1080/03461238.2012.755938

Arjas, E. (1989). The claims reserving problem in non-life insurance: Some structural ideas. *ASTIN Bulletin: The Journal of the IAA*, *19*(2), 139–152. https://doi.org/10.2143/AST.19.2.2014905

Crevecoeur, J., Antonio, K., & Verbelen, R. (2019). Modeling the number of hidden events subject to observation delay. *European Journal of Operational Research*, *277*(3), 930–944. https://doi.org/10.1016/j.ejor.2019.02.044

England, P. D., & Verrall, R. J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, *8*(3), 443–544. https://doi.org/10.1017/S1357321700003809

Huang, J., Qiu, C., & Wu, X. (2015). Stochastic loss reserving in discrete time: Individual vs. aggregate data models. *Communications in Statistics-Theory and Methods*, *44*(10), 2180–2206. https://doi.org/10.1080/03610926.2014.976473

Huang, J., Qiu, C., Wu, X., & Zhou, X. (2015). An individual loss reserving model with independent reporting and settlement. *Insurance: Mathematics and Economics*, *64*(1), 232–245. https://doi.org/10.1016/j.insmatheco.2015.05.010

Huang, J., Wu, X., & Zhou, X. (2016). Asymptotic behaviors of stochastic reserving: Aggregate versus individual models. *European Journal of Operational Research*, *249*(2), 657–666. https://doi.org/10.1016/j.ejor.2015.09.039

Lindholm, M., Lindskog, F., & Wahl, F. (2020). Estimation of conditional mean squared error of prediction for claims reserving. *Annals of Actuarial Science*, *14*(1), 93–128. https://doi.org/10.1017/S174849951900006X

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. https://doi.org/10.2143/AST.23.2.2005092

Mack, T. (2000). Credible claims reserves: The Benktander method. *ASTIN Bulletin: The Journal of the IAA*, *30*(2), 333–347. https://doi.org/10.2143/AST.30.2.504639

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed). Chapman and Hall.

Norberg, R. (1993). Prediction of outstanding liabilities in non-life insurance 1. *ASTIN Bulletin: The Journal of the IAA*, *23*(1), 95–115. https://doi.org/10.2143/AST.23.1.2005103

Norberg, R. (1999). Prediction of outstanding liabilities II. Model variations and extensions. *ASTIN Bulletin: The Journal of the IAA*, *29*(1), 5–25. https://doi.org/10.2143/AST.29.1.504603

Pigeon, M., Antonio, K., & Denuit, M. (2013). Individual loss reserving with the multivariate skew normal framework. *ASTIN Bulletin: The Journal of the IAA*, *43*(3), 399–428. https://doi.org/10.1017/asb.2013.20

Pigeon, M., Antonio, K., & Denuit, M. (2014). Individual loss reserving using paid-incurred data. *Insurance: Mathematics and Economics*, *58*(2), 121–131. https://doi.org/10.1016/j.insmatheco.2014.06.012

Van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.

Verrall, R. J., Nielsen, J. P., & Jessen, A. H. (2010). Prediction of RBNS and IBNR claims using claim amounts and claim counts. *Astin Bulletin*, *40*(2), 871–887. https://doi.org/10.2143/AST.40.2.2061139

Wahl, F. (2019). Explicit moments for a class of micro-models in non-life insurance. *Insurance: Mathematics and Economics*, *89*(7), 140–156. https://doi.org/10.1016/j.insmatheco.2019.10.001

Wahl, F., Lindholm, M., & Verrall, R. (2019). The collective reserving model. *Insurance: Mathematics and Economics*, *87*(7), 34–50. https://doi.org/10.1016/j.insmatheco.2019.04.003

Wüthrich, M. V., & Merz, M. (2008). *Stochastic Claims Reserving Methods in Insurance*. John Wiley and Sons.

Yu, X., & He, R. (2016). Individual claims reserving models based on marked Cox processes. *Chinese Journal of Applied Probability and Statistics 32(2), 201-219.* http://aps.ecnu.edu.cn/EN/Y2016/V32/I2/201

Zhao, X., & Zhou, X. (2010). Applying copula models to individual claim loss reserving methods. *Insurance: Mathematics and Economics*, *46*(2), 290–299. https://doi.org/10.1016/j.insmatheco.2009.11.001

Appendix

***Proof of Proposition* 3.1:** To derive the following gradient and Hessian matrix, we need the identities $\frac{\partial q_v}{\partial \boldsymbol{\rho}} = q_v(\boldsymbol{\delta}_v - \boldsymbol{q}) \otimes \boldsymbol{x}$, which gives that

$$\frac{\partial \boldsymbol{q}'}{\partial \boldsymbol{\rho}} = (\mathrm{diag}(\boldsymbol{q}) - \boldsymbol{q}\boldsymbol{q}') \otimes \boldsymbol{x}, \quad \frac{\partial \log \boldsymbol{q}'}{\partial \boldsymbol{\rho}} = (\mathbf{I}_{D^s} - \boldsymbol{q}\mathbf{1}'_{D^s}) \otimes \boldsymbol{x}.$$

Then we have the following gradient according to formulas above, which is

$$\begin{aligned}
\nabla Q^{ios}(\boldsymbol{\rho}) &= \left[\sum_{v=0}^{D_i^s} N_v^s(\boldsymbol{\delta}_v - \boldsymbol{q}) + \sum_{u=0}^{D_i^r} N_u^{rbns}\left(\frac{(\mathbf{0}, \bar{\boldsymbol{q}}'_{n-i-u})'}{\bar{Q}_{n-i-u}} - \boldsymbol{q}\right)\right] \otimes \boldsymbol{x} \\
&= \left[\boldsymbol{N}_i^s + \sum_{u=0}^{D_i^r} \frac{N_u^{rbns}}{\bar{Q}_{n-i-u}} \begin{pmatrix} \mathbf{0} \\ \bar{\boldsymbol{q}}_{n-i-u} \end{pmatrix} - \left(\sum_{v=0}^{D_i^s} N_v^s + \sum_{u=0}^{D_i^r} N_u^{rbns}\right)\boldsymbol{q}\right] \otimes \boldsymbol{x}.
\end{aligned}$$

By some algebraic computation, it follows that

$$\begin{aligned}
\frac{\partial^2 Q^{ios}(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}\boldsymbol{\rho}'} &= -\left[\left(\sum_{v=0}^{D_i^s} N_v^s + \sum_{u=0}^{D_i^r} N_u^{rbns}\right)(\mathrm{diag}(\boldsymbol{q}) - \boldsymbol{q}\boldsymbol{q}') \right.\\
&\quad + \sum_{u=0}^{D_i^r} N_u^{rbns} \\
&\quad \left. \times \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}(\bar{\boldsymbol{q}}_{n-i-u}) - \frac{\bar{\boldsymbol{q}}_{n-i-u}\bar{\boldsymbol{q}}'_{n-i-u}}{\bar{Q}_{n-i-u}} \end{pmatrix}\right] \otimes \boldsymbol{x}\boldsymbol{x}'.
\end{aligned}$$

Because $\sum_{v=0}^{D_i^s} N_v^s + \sum_{u=0}^{D_i^r} N_u^{rbns}$ is just the number of those reported claims incurred in accident year i,

$$\mathbb{E}\left[\sum_{v=0}^{D_i^s} N_v^s + \sum_{u=0}^{D_i^r} N_u^{rbns} \,\middle|\, r, \boldsymbol{x}\right] = r\sum_{u=0}^{D_i^r} \lambda_u.$$

Observe further that $N_u^{rbns} = 0$ for $i \le n - D^s$ and $0 \le u \le n - i - D^s$. Therefore,

$$\begin{aligned}
\mathbb{E}\left[\frac{\partial^2 Q^{ios}(\boldsymbol{\rho})}{\partial \boldsymbol{\rho}\boldsymbol{\rho}'} \,\middle|\, r, \boldsymbol{x}\right] &= -r\left[\sum_{u=0}^{D_i^r} \lambda_u(\mathrm{diag}(\boldsymbol{q}) - \boldsymbol{q}\boldsymbol{q}') \right.\\
&\quad - \sum_{u=(n-i-D^s+1)_+}^{D_i^r} \lambda_u \\
&\quad \left. \times \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}(\bar{\boldsymbol{q}}_{n-i-u}) - \frac{\bar{\boldsymbol{q}}_{n-i-u}\bar{\boldsymbol{q}}'_{n-i-u}}{\bar{Q}_{n-i-u}} \end{pmatrix}\right] \otimes \boldsymbol{x}\boldsymbol{x}'.
\end{aligned} \tag{A1}$$

Let $v = n - i - u$ and note that

$$\mathrm{diag}(\boldsymbol{q}) - \boldsymbol{q}\boldsymbol{q}' - \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathrm{diag}(\bar{\boldsymbol{q}}_v) - \frac{\bar{\boldsymbol{q}}_v\bar{\boldsymbol{q}}'_v}{\bar{Q}_v} \end{pmatrix} = \begin{pmatrix} \mathrm{diag}(\boldsymbol{q}_v) - \boldsymbol{q}_v\boldsymbol{q}'_v & -\boldsymbol{q}_v\bar{\boldsymbol{q}}'_v \\ -\bar{\boldsymbol{q}}_v\boldsymbol{q}'_v & \frac{Q_v}{\bar{Q}_v}\bar{\boldsymbol{q}}_v\bar{\boldsymbol{q}}'_v \end{pmatrix}.$$

Then, Equation (A1) gives rise to the desired result. ∎

***Proof of Theorem* 4.1:** By (3), the loss reserve can be computed as

$$\mathbb{E}(R \mid \mathscr{F}^{uo}) = \mathbb{E}[R^{rbns} \mid \mathscr{F}^{uo}] + \mathbb{E}[R^{ibnr} \mid \mathscr{F}^{uo}].$$

According to Assumption 3.2, for a representative policy in year i, given N_u^{rbns} with $n - i - D^s + 1 \le u \le D_i^r$, $(N_{u,n-i-u+1}, \ldots, N_{uD^s})$ follows multinomial distribution with parameters N_u^{rbns} and $\frac{1}{\bar{Q}_{n-i-u}}(q_{n-i-u+1}, \ldots, q_{D^s})$. Then by Assumption 3.3, the RBNS loss reserve is

$$\begin{aligned}
\mathbb{E}[R^{rbns} \mid \mathscr{F}^{uo}] &= \sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{u=(n-i-D^s+1)_+}^{D_i^r} \mathbb{E}\left[\sum_{v=n-i-u+1}^{D^s}\sum_{l=1}^{N_{ikuv}} Y_{ikuvl} \middle| \mathscr{F}^{uo}\right] \\
&= \sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{u=(n-i-D^s+1)_+}^{D_i^r} N_{iku}^{rbns}\frac{\sum_{v=n-i-u+1}^{D^s} q_{ikv}\mu_{ikuv}}{\bar{Q}_{ik,n-i-u}} \\
&= \sum_{v=1}^{D^s}\sum_{u=0}^{D_v^r}\sum_{k=1}^{m_{n-v-u+1}} N^{rbns}_{(n-v-u+1)ku}\tilde{\mu}_{(n-v-u+1)kuv}.
\end{aligned}$$

It can be easily proved that IBNR claims are independent of historical observation $\mathscr{F}^{uo}$ by Assumption 3.1–3.3. Hence, IBNR loss reserve is computed by

$$\begin{aligned}
\mathbb{E}\left[R^{ibnr} \mid \mathscr{F}^{uo}\right] &= E\left[\sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{u=n-i+1}^{D^r}\sum_{v=0}^{D^s}\sum_{l=1}^{N_{ikuv}} Y_{ikuvl} \mid \mathscr{F}^{uo}\right] \\
&= \sum_{i=n-D^r+1}^{n}\sum_{k=1}^{m_i}\sum_{u=n-i+1}^{D^r}\sum_{v=0}^{D^s} E\left[\sum_{l=1}^{N_{ikuv}} Y_{ikuvl}\right] \\
&= \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} p_{iku} r_{ik} \exp(\boldsymbol{x}'_{ik}\boldsymbol{\beta})\tilde{\mu}_{iku0}.
\end{aligned}$$

According to independence assumptions in Assumptions 3.1–3.3, the developments of RBNS claims are independent of developments of IBNR claims, which results in the independence between R^{rbns} and R^{ibnr}. Then the variance of R given $\mathscr{F}^{uo}$ is

$$\mathrm{Var}(R \mid \mathscr{F}^{uo}) = \mathrm{Var}(R^{rbns} \mid \mathscr{F}^{uo}) + \mathrm{Var}(R^{ibnr} \mid \mathscr{F}^{uo}).$$

First, for $v \ge n - i - u + 1$, we compute

$$\begin{aligned}
\mathrm{Var}\left(\sum_{l=1}^{N_{uv}} Y_{uvl} \mid E^o\right) &= \mathrm{Var}\left(\mathbb{E}\left[\sum_{l=1}^{N_{uv}} Y_{uvl} \mid N_{uv}, E^o\right]\right) \\
&\quad + \mathbb{E}\left[\mathrm{Var}\left(\sum_{l=1}^{N_{uv}} Y_{uvl} \mid N_{uv}, E^o\right)\right] \\
&= \mu_{uv}^2\mathrm{Var}(N_{uv} \mid \mathscr{F}^o) + \phi^p\mu_{uv}\mathbb{E}[N_{uv} \mid \mathscr{F}^o] \\
&= N_u^{rbns}\left[\mu_{uv}^2\frac{q_v(1-q_v)}{\bar{Q}_{n-i-u}^2} + \phi^p\mu_{uv}\frac{q_v}{\bar{Q}_{n-i-u}}\right].
\end{aligned}$$

For $v_1, v_2 \geq n-i-u+1$, compute $\mathrm{Cov}(\sum_{l=1}^{N_{uv_1}} Y_{uv_1l}, \sum_{l=1}^{N_{uv_2}} Y_{uv_2l} \mid \mathscr{F}^o)$ which is equal to

$$\mathrm{Cov}\left(\mathbb{E}\left[\sum_{l=1}^{N_{uv_1}} Y_{uv_1l} \mid N_{uv_1}, N_{uv_2}, \mathscr{F}^o\right], \mathbb{E}\left[\sum_{l=1}^{N_{uv_2}} Y_{uv_2l} \mid N_{uv_1}, N_{uv_2}, \mathscr{F}^o\right]\right) + \mathbb{E}\left[\mathrm{Cov}\left(\sum_{l=1}^{N_{uv_1}} Y_{uv_1l}, \sum_{l=1}^{N_{uv_2}} Y_{uv_2l} \mid N_{uv_1}, N_{uv_2}, \mathscr{F}^o\right)\right],$$

which can be computed as follows:

$$\begin{aligned}\mathrm{Cov}&\left(\sum_{l=1}^{N_{uv_1}} Y_{uv_1l}, \sum_{l=1}^{N_{uv_2}} Y_{uv_2l} \mid \mathscr{F}^o\right)\\ &= \mu_{uv_1}\mu_{uv_2}\mathrm{Cov}(N_{uv_1} \mid \mathscr{F}^o], N_{uv_2} \mid \mathscr{F}^o])\\ &= -N_u^{rbns}\mu_{uv_1}\mu_{uv_2}\frac{q_{v_1}q_{v_2}}{\bar{Q}_{n-i-u}^2}.\end{aligned}$$

Then by independence among policies and Assumptions 3.2 and 3.3, the variance of RBNS loss reserve given $\mathscr{F}^{uo}$ is

$$\begin{aligned}&\mathrm{Var}(R^{rbns} \mid \mathscr{F}^{uo})\\ &= \sum_{i=1}^{n}\sum_{k=1}^{m_i}\mathrm{Var}\left(\sum_{u=(n-i-D^s+1)_+}^{D_i^r}\sum_{v=n-i-u+1}^{D^s}\sum_{l=1}^{N_{ikuv}} Y_{ikuvl} \mid \mathscr{F}^{uo}\right)\\ &= \sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{u=(n-i-D^s+1)_+}^{D_i^r} N_{iku}^{rbns}\left(\frac{\sum_{v=n-i-u+1}^{D^s} q_{ikv}\mu_{ikuv}^2}{\bar{Q}_{ik,n-i-u}^2}\right.\\ &\quad \left. -\left(\frac{\sum_{v=n-i-u+1}^{D^s} q_{ikv}\mu_{ikuv}}{\bar{Q}_{ik,n-i-u}}\right)^2 + \phi^p\frac{\sum_{v=n-i-u+1}^{D^s} q_{ikv}\mu_{ikuv}}{\bar{Q}_{ik,n-i-u}}\right)\\ &= \sum_{v=1}^{D^s}\sum_{u=0}^{D_v^r}\sum_{k=1}^{m_{n-v-u+1}} N_{n-v-u+1,ku}^{rbns}\left(\frac{\tilde{\mu}_{n-v-u+1,kuv}^s}{\bar{Q}_{n-v-u+1,k,v-1}}\right.\\ &\quad \left. - \tilde{\mu}_{n-v-u+1,kuv}^2 + \phi^p\tilde{\mu}_{n-v-u+1,kuv}\right).\end{aligned}$$

Because IBNR claims are independent of historical observation $\mathscr{F}^{uo}$, variance of IBNR loss reserve given $\mathscr{F}^{uo}$ is computed by

$$\begin{aligned}&\mathrm{Var}(R^{ibnr} \mid \mathscr{F}^{uo})\\ &= \mathrm{Var}\left(\sum_{i=1}^{n}\sum_{k=1}^{m_i}\sum_{u=n-i+1}^{D^r}\sum_{v=0}^{D^s}\sum_{l=1}^{N_{ikuv}} Y_{ikuvl}\right)\\ &= \sum_{i=n-D^r+1}^{n}\sum_{k=1}^{m_i}\sum_{u=n-i+1}^{D^r}\mathrm{Var}\left(\mathbb{E}\left[\sum_{v=0}^{D^s}\sum_{l=1}^{N_{ikuv}} Y_{ikuvl} \mid N_{iku}\right]\right)\\ &\quad + \mathbb{E}\left[\mathrm{Var}\left(\sum_{v=0}^{D^s}\sum_{l=1}^{N_{ikuv}} Y_{ikuvl} \mid N_{iku}\right)\right]\\ &= \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\sum_{k=1}^{m_i} p_{iku}r_{ik}\exp(\boldsymbol{x}_{ik}'\boldsymbol{\beta})(\tilde{\mu}_{iku0}^s\\ &\quad + (\phi-1)\tilde{\mu}_{iku0}^2 + \phi^p\tilde{\mu}_{iku0}).\end{aligned}$$

■

***Proof of Theorem* 4.3:** Expand $R_m(\hat{\boldsymbol{\theta}})$ about true parameters $\boldsymbol{\theta}$ by Taylor expansion. Then we have

$$\frac{1}{m}(\hat{R}_{II} - R_m) = \frac{1}{m}\frac{\partial R_m(\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}'}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) + o_p\left(\frac{\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|}{m}\right).$$

One knows that $\dot{\mu}_{uv} := \frac{\partial\mu_{uv}}{\partial\boldsymbol{\gamma}} = \mu_{uv}\boldsymbol{x}_{uv}$. Write $\boldsymbol{\mu}_{uv} = (\mu_{u,v+1}, \mu_{u,v+2}, \ldots, \mu_{u,D^s})'$ and $\dot{\boldsymbol{\mu}}_{uv} = \frac{\partial\boldsymbol{\mu}_{uv}'}{\partial\boldsymbol{\gamma}}$, $u = 0, 1, \ldots, D^r$ and $v = 0, \ldots, D^s - 1$. To compute the partial derivative in the Taylor expansion above, we need the following partial derivatives:

$$\begin{aligned}\frac{\partial[p_u r\exp(\boldsymbol{x}'\boldsymbol{\beta})]}{\partial(\boldsymbol{\beta}', \boldsymbol{\pi}')'} &= \begin{pmatrix}1\\ \boldsymbol{\delta}_u - \boldsymbol{p}\end{pmatrix} \otimes (p_u r\exp(\boldsymbol{x}'\boldsymbol{\beta})\boldsymbol{x}),\\ \frac{\partial\tilde{\mu}_{uv}}{\partial\boldsymbol{\rho}} &= \left[\boldsymbol{0}, \boldsymbol{\mu}_{u,v-1}'\left(\frac{1}{\bar{Q}_{v-1}}\mathrm{diag}(\overline{\boldsymbol{q}}_{v-1})\right.\right.\\ &\quad \left.\left. - \frac{1}{\bar{Q}_{v-1}^2}\overline{\boldsymbol{q}}_{v-1}\overline{\boldsymbol{q}}_{v-1}'\right)\right]' \otimes \boldsymbol{x}_u, \quad v \geq 1,\\ \frac{\partial\tilde{\mu}_{u0}}{\partial\boldsymbol{\rho}} &= \left[(\mathrm{diag}(\boldsymbol{q}) - \boldsymbol{q}\boldsymbol{q}')\boldsymbol{\mu}_{u0} - q_0\mu_{u0}\boldsymbol{q}\right] \otimes \boldsymbol{x}, \quad \text{and}\\ \frac{\partial\tilde{\mu}_{uv}}{\partial\boldsymbol{\gamma}} &= \frac{1}{\bar{Q}_{v-1}}\frac{\partial\boldsymbol{\mu}_{u,v-1}'}{\partial\boldsymbol{\gamma}}\overline{\boldsymbol{q}}_{v-1}, \quad v \geq 1,\\ \frac{\partial\tilde{\mu}_{u0}}{\partial\boldsymbol{\gamma}} &= \left(\frac{\partial\mu_{u0}}{\partial\boldsymbol{\gamma}}, \frac{\partial\boldsymbol{\mu}_{u0}'}{\partial\boldsymbol{\gamma}}\right)(q_0, \boldsymbol{q}')'.\end{aligned}$$

By the law of large numbers, it can be proved that $\frac{1}{m}\frac{\partial R_m(\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}'} \xrightarrow{\text{a.s.}} \boldsymbol{g}$, where $\boldsymbol{g} = (\boldsymbol{g}_1', \boldsymbol{g}_2', \boldsymbol{g}_3')'$, where denoting

$$M_{uv} = \begin{cases}\left(\boldsymbol{0}, \boldsymbol{\mu}_{uv}'\left(\mathrm{diag}(\overline{\boldsymbol{q}}_v) - \frac{1}{\bar{Q}_v}\overline{\boldsymbol{q}}_v\overline{\boldsymbol{q}}_v'\right)\right)', & v > 0\\ \left(\mathrm{diag}(\overline{\boldsymbol{q}}_0) - \frac{1}{\bar{Q}_0}\overline{\boldsymbol{q}}_0\overline{\boldsymbol{q}}_0'\right)\boldsymbol{\mu}_{u0}, & v = 0,\end{cases}$$

$$\begin{aligned}\boldsymbol{g}_1 &= \sum_{u=1}^{D^r}\sum_{i=n-u+1}^{n}\kappa_i\mathbb{E}\left[\tilde{\mu}_{u0}\begin{pmatrix}1\\ \boldsymbol{\delta}_u - \boldsymbol{p}\end{pmatrix} \otimes (r\lambda_u\boldsymbol{x})\right],\\ \boldsymbol{g}_2 &= \sum_{u=0}^{D^r}\mathbb{E}\left[\left(\sum_{v=0}^{D^s-1}\kappa_{n-u-v}M_{uv} + \sum_{i=n-u+1}^{n}\kappa_i\right.\right.\\ &\quad \left.\left. \times[(\mathrm{diag}(\boldsymbol{q}) - \boldsymbol{q}\boldsymbol{q}')\boldsymbol{\mu}_{u0} - q_0\mu_{u0}\boldsymbol{q}]\right) \otimes (r\lambda_u\boldsymbol{x})\right],\\ \boldsymbol{g}_3 &= \sum_{u=0}^{D^r}\mathbb{E}\left[\left(\sum_{v=0}^{D^s-1}\kappa_{n-u-v}\dot{\boldsymbol{\mu}}_{uv}\overline{\boldsymbol{q}}_v\right.\right.\\ &\quad \left.\left. + \sum_{i=n-u+1}^{n}\kappa_i(\dot{\mu}_{u0}, \dot{\boldsymbol{\mu}}_{u0})(q_0, \boldsymbol{q}')'\right) \otimes (r\lambda_u\boldsymbol{x}_{uv})\right]. \quad \text{(A2)}\end{aligned}$$

It is well known that $\hat{\boldsymbol{\theta}} \xrightarrow{P} \boldsymbol{\theta}$ under some regular conditions and hence $\frac{1}{m}(\hat{R}_{II} - R_m) \xrightarrow{P} 0$. Besides,

$$\frac{\operatorname{Var}(R \mid \mathscr{F}^{uo})}{m} \xrightarrow{\text{a.s.}} V_R$$
$$= \sum_{v=1}^{D^s} \sum_{u=0}^{D^r_v} \kappa_{n-v-u+1} \times \mathbb{E}\left[r\lambda_u \bar{Q}_{v-1} \left(\frac{\tilde{\mu}^s_{uv}}{\bar{Q}_{v-1}} - \tilde{\mu}^2_{uv} + \phi^p \tilde{\mu}_{uv} \right) \right]$$
$$+ \sum_{u=1}^{D^r} \sum_{i=n-u+1}^{n} \kappa_i \mathbb{E}\left[r\lambda_u (\tilde{\mu}^s_{u0} + (\phi - 1)\tilde{\mu}^2_{u0} + \phi^p \tilde{\mu}_{u0}) \right].$$

If individual data model hold true, one can similarly prove that $\frac{1}{m}(\hat{R}_{ID} - R_m) \xrightarrow{P} 0$. Therefore, $M^r \xrightarrow{P} 1$ in this case. If individual information model holds true, we can easily prove that $\frac{\hat{R}_{ID} - R_m}{m}$ is asymptotically biased, which results from the following arguments. The law of large numbers readily gives $\hat{h}_v \xrightarrow{\text{a.s.}} \check{h}_v$ and $\hat{\mu}_{uv} \xrightarrow{\text{a.s.}} \check{\mu}_{uv}$. Further, we have $\hat{q}_v \xrightarrow{\text{a.s.}} \check{q}_v := \check{h}_v \prod_{s=0}^{v-1}(1 - \check{h}_s)$ and $\hat{\tilde{\mu}}_{uv} \xrightarrow{\text{a.s.}} \check{\tilde{\mu}}_{uv} := \sum_{s=v}^{D^s} \check{q}_s \check{\mu}_{us} / \sum_{s=v}^{D^s} \check{q}_s$. We have

$$\frac{\hat{R}_{ID} - R_m}{m} = \frac{\hat{R}_{ID} - \check{R}_{ID}}{m} + \frac{\check{R}_{ID} - R_m}{m},$$

where $\check{R}_{ID} = \sum_{v=1}^{D^s} \sum_{u=0}^{D^r_v} \sum_{k=1}^{m_{n-u-v+1}} N^{rbns}_{n-u-v+1,ku} \check{\tilde{\mu}}_{uv} + \sum_{u=1}^{D^r} \sum_{i=n-u+1}^{n} \sum_{k=1}^{m_i} r_{ik} \check{\lambda}_u \check{\tilde{\mu}}_{u0}$ and then

$$\check{R}_{ID} - R_m = \sum_{v=1}^{D^s} \sum_{u=0}^{D^r_v} \sum_{k=1}^{m_{n-u-v+1}} N^{rbns}_{n-u-v+1,ku} (\check{\tilde{\mu}}_{uv} - \tilde{\mu}_{n-u-v+1,kuv})$$
$$+ \sum_{u=1}^{D^r} \sum_{i=n-u+1}^{n} \sum_{k=1}^{m_i} (r_{ik} \check{\lambda}_u \check{\tilde{\mu}}_{u0} - r_{ik} \lambda_{iku} \tilde{\mu}_{iku0}).$$

Apparently, $\frac{\hat{R}_{ID} - \check{R}_{ID}}{m} \xrightarrow{\text{a.s.}} 0$ and by the law of large numbers and some simple algebra operations, we show that

$$\frac{1}{m}(\check{R}_{ID} - R_m) \xrightarrow{\text{a.s.}} \Delta$$
$$= \sum_{v=1}^{D^s} \sum_{u=0}^{D^r_v} \kappa_{n-v-u+1} \mathbb{E}[r\lambda_u \bar{Q}_{v-1} (\check{\tilde{\mu}}_{uv} - \tilde{\mu}_{uv})]$$
$$+ \sum_{u=1}^{D^r} \sum_{i=n-u+1}^{n} \kappa_i \mathbb{E}[r\lambda_u (\check{\tilde{\mu}}_{u0} - \tilde{\mu}_{u0})].$$

Therefore, if asymptotic bias Δ is not zero, $M^r \xrightarrow{P} 0$. Then we complete the proof. ■