Micro reserving: a look at the one-year risk and a comparison to chain-ladder

Jóhanna Sigmundsdóttir

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2015:4
Försäkringsmatematik
Maj 2015

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stockholm University

Mathematical Statistics
Stockholm University
Master Thesis **2015:4**
http://www.math.su.se

Micro reserving: a look at the one-year risk and a comparison to chain-ladder

Jóhanna Sigmundsdóttir*

May 2015

Abstract

Reserve estimation or reserving is traditionally based on triangulated aggregated loss data, i.e. payments or incurred claims which are summed over origin periods and development periods. The goal of this thesis is to look further into reserving from the viewpoint of survival analysis. These models have been referred to as micro models as they are based on detailed data from individual claims. Each claim is treated as a process of an occurrence, reporting delay, payments and finally a settlement. By modelling the claim's process the complete claim data of the outstanding liabilities is simulated. The model is examined on a dataset from a Swedish insurance company and compared to the reserve estimates obtained by the traditional chain-ladder method. The data displays trend changes through time. The micro model manages to capture these trend changes to a larger extent than the chain-ladder and gives detailed information about the behaviour of the data. The underlying assumptions of the chain-ladder method are not fulfilled leading to a largely different estimate. Further look is taken at the one-year reserve and premium risk, where the one-year risk estimation is a natural extension of the micro model. We conclude that micro models can offer multiple advantages over the traditional methods. Such as reduced uncertainties in estimates and detailed information about different characteristics of the data.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: johannasigmunds@gmail.com. Supervisor: Mathias Lindholm.

Acknowledgements

I would like to thank my supervisor Mathias Lindholm for all his help. And also my parents. And Magnús for finding almost everything interesting.

Contents

1 Introduction

1.1 Background

Reserve estimation or reserving is traditionally based on triangulated aggregated loss data, i.e. payments or incurred claims which are summed over origin periods and development periods. These estimation techniques are for example dealt with by Mack (1993) who has derived the formula for the standard error of the widely used chain-ladder reserve estimates and by England and Verrall (2002) where a wide range of stochastic reserving models for use in general insurance are discussed.

Whereas chain-ladder and other similar methods have been widely used and studied, less has been published on micro reserving (or granular reserving) models where each claim or policy is the basic building block for the model.

Arjas (1989) presented ideas on individual claims reserve estimates based on point processes and Norberg (1993, 1999) and Haastrup and Arjas (1996) have formulated further the mathematical framework. Based on these works Antonio and Plat (2014) set up a likelihood and carried out an extensive case study for which they conclude "the micro-model outperforms the aggregate models under consideration and reveals a more realistic predictive distribution of the reserve". The present study continues from Antonio and Plat (2014) with modifications and different data. Chapter 9 in Mikosch (2009) contains interesting discussions and inference on similar models drawing inspiration from Norberg (1993).

1.2 Objective

The goal of this master thesis is to look further into reserving from the viewpoint of survival analysis based on individual claims. Each claim is treated as a process of an occurrence, reporting delay, payments and finally a settlement. This corresponds to data consisting of the accident date, reporting date, payment and settlement dates along with corresponding payments. Traditional reserving methods (chain-ladder, Bournhatter-Fergusson, separation method) are based on data aggregated over accident and development period, without considering all the detailed data available in modern insurance companies. The present study encompasses the construction of a model that simulates the complete claim data of outstanding liabilities. The model is referred to as a micro reserving model.

The estimated reserve will be compared to the corresponding reserve estimated with chain-ladder method and the result discussed.

A further look is taken into the one-year reserve risk and one-year premium risk where the risk estimation follows naturally from the model construction. The risk will be compared to the risk estimated from one-year risk models based on bootstrapped chain-ladder method.

The model is validated by comparison of the models prediction to observed one-year payments as well as to the chain-ladder prediction. Sensitivity analysis of model assumptions are performed.

2 Model Outline

The claim process is modelled with the marked Poisson process previously discussed by Norberg (1993, 1999) and Antonio and Plat (2014). Antonio and Plat (2014) carried out a case study with a description of the likelihood, parameter estimation and the simulations process. We repeat that case study with a modification of claims which have not received a payment and extensions to one-year premium and reserve risk. We start by giving a short description of the reserving model and its main building blocks before reviewing the theory in the following sections.

2.1 The Main Building Blocks

The model predicts the remaining payments of outstanding claims for historical years. The remaining payments are divided in two parts: RBNS which stands for *reported but not settled* claims and IBNR which are *incurred but not reported* to the insurance company claims. Four main characteristics of the outstanding claims are estimated from historical claims and used to simulate the outstanding payments. These are:

1. **Reporting delay**
   The reporting delay is the time interval between an accident and the time when the claim is reported to the insurance company. With an estimate of this delay we predict the number of IBNR claims yet to be reported and their reporting delay. With the addition of a simulated accident date, the IBNR claims are seen as RBNS in a modelling perspective.

2. **Development process**
   After a claim has been reported it can receive multiple payments before being settled. The settlement either comes with or without a payment. The development process consists of the time-points of these events. We predict the development process of all outstanding claims.

3. **Payments**
   In the development process we simulate time-points of payments. By simulating the payments we can estimate the outstanding liabilities.

4. **Probability of closing a claim without ever receiving a payment**
   Special measures are taken to exclude claims which have been reported but never receive a payment. Some claims are closed without getting any payment and these need to be adjusted for. From here on these claims will be referred to as *zero claims*. The main interest is to estimate the outstanding payments which gives reason to the exclusion of zero claims from the model.

By simulating the payment process the result of the model will be an empirical distribution of the outstanding reserve. One-year reserve and premium risk then becomes a natural extension.

3 The Likelihood of the Occurred Claims

The occurrence process is a non-homogeneous Poisson process with marked distributions. The marked distributions consist of distributions for the delay, time of payments and payments and settlement of the claim. For further theoretical background see Karr (1991) and Norberg (1993, 1999). The outline given here is closer to that of Antonio and Plat (2014).

We will discuss the likelihood from the viewpoint of individual claims. The process of each claim i is divided into three parts. The occurrence time T_i, the reporting delay U_i and the development after the process has been reported to the insurance company, $\mathbf{X}_i$. $\mathbf{X}_i$ consists of the times, V_{ij}, and type of events, E_{ij}, after the claim is reported. Payment events also include a payment, $P_{ij'}$. Here, j runs over all events of the claim and j' runs over events that include a payment.

Claims are observed until time-point τ. The likelihood is constructed for the reported claims where $T_i + U_i + V_{ij} \leq \tau$. The exposure is denoted with $w(t)$.

3.1 The Occurrence Process

Let $N(t)$ be the counting process for the number of claims in $[0, t]$. Further, let $\Delta N(t)$ denote the number of claims on $(t, t + \Delta t)$. The counting process is assumed to be memoryless.

We follow the exposition of Cook and Lawless (2007). The process is adjusted for the reporting delay of each occurrence as well as the exposure at time t, $w(t)$. The intensity of the occurrence process is then defined as,

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(\Delta N(t) = 1) P(U \leq \tau - t) w(t)}{\Delta t}.$$

Assuming that only one event can occur at a given time t, the probabilities for $\Delta N(t)$ are

$$\begin{aligned} P(\Delta N(t) = 1) &= \lambda(t) P(U \leq \tau - t) w(t) \Delta t + o(\Delta t), \\ P(\Delta N(t) = 0) &= 1 - \lambda(t) P(U \leq \tau - t) w(t) \Delta t + o(\Delta t). \end{aligned}$$

We consider the partition $0 = s_0 < s_1 < ... < s_R = \tau$ with $\Delta s = s_{r+1} - s_r$. As R increases $\Delta s \to 0$. The distribution of the counting process is

$$\begin{aligned} P(N|\lambda) &= \prod_{r=0}^{R} P(\Delta N(s_r)) \\ &= \prod_{r=0}^{R} ([\lambda(s_r) P(U \leq \tau - s_r) w(s_r) \Delta s_r]^{\Delta N(s_r)} \\ &\quad \times [1 - \lambda(s_r) P(U \leq \tau - s_r) w(s_r) \Delta s_r]^{1 - \Delta N(s_r)}). \end{aligned}$$

Now, $\Delta N(s_r) = 1$ on the occurrence time-points of each claim $t_1, ..., t_n$ and $\Delta N(s_r) = 0$ otherwise. Using the approximation

$$\begin{aligned} & \log[1 - \lambda(s_r) P(U \leq \tau - s_r) w(s_r) \Delta s_r] \\ \approx & - \lambda(s_r) P(U \leq \tau - s_r) w(s_r) \Delta s_r, \end{aligned}$$

we obtain the likelihood by dividing by $\prod_{r=0}^{R} \Delta s_r^{\Delta N(s_r)}$ and letting $R \to \infty$:

$$L(\cdot) = \left(\prod_{i=1}^{n} \lambda(t_i) P(U \leq \tau - t_i) w(t_i) \right) \exp\left(- \int_0^{\tau} \lambda(s) P(U \leq \tau - s) w(s) ds \right)$$

Note that the limits of the integral do not change around the n events.

The occurrence process of the occurred claims is a Poisson process with intensity $\lambda(t) P(U \leq \tau - t) w(t)$.

3.2 The Reporting Delay

The observed claims all have $T + U \leq \tau$. We denote the density of the distribution $P(U \leq u|T + U \leq \tau)$ with $f_U(u|T + U \leq \tau)$. The likelihood of the observed delays is

$$L(U|T + U \leq \tau) = \prod_{i=1}^{n} f_U(u_i|T_i + U_i \leq \tau)$$

The relationship between the conditional density $f_U(u_i|T_i + U_i \leq \tau)$ and the density of the delays from all claims (observed and unobserved), denoted with $f_U(u_i)$, is discussed further in section 4.2.

3.3 The Development Process

Each claim can have multiple payments before being settled with or without a final payment. Let $E = \{sep, se, p\}$ where p stands for payment, sep for settled with a payment and se settled without a payment.

The development process thereby contains recurrent events of multiple types. The derivation of the likelihood is similar to that of the occurrence process. See Cook and Lawless (2007) for further theoretical background.

Let $N_{ie}(t)$ be the counting process for number of events $e \in E$ on $[0, t]$ for claim i and $\Delta N_{ie}(t)$ the number of events in $(t, t + \Delta t)$.

The development process is assumed to be memoryless. The hazard rates are defined by

$$h_e(t) = \lim_{\Delta t \to 0} \frac{P(\Delta N_{ie}(t) = 1)}{\Delta t}$$

for all $e \in E$. As for the occurrence process we assume that only one event can happen at a given time-point. Therefore,

$$\begin{aligned} P(\Delta N_{ie}(t) = 1) &= h_e(t)\Delta t + o(\Delta t),\ \forall e \in E \\ P(\Delta N_{i.}(t) = 0) &= 1 - \sum_e h_e(t)\Delta t + o(\Delta t). \end{aligned}$$

With a partition $0 = s_0 < s_1 < ... < s_R = \tau_i$, where $\Delta s = s_{r+1} - s_r$, $\tau_i = \min(\tau - T_i - U_i, V_i)$ and V_i is the time from notification until settlement. We get

$$
\begin{aligned}
P(N|h) &= \prod_{r=0}^{R}\prod_{e} P(\Delta N_{ie}(s_r)) \\
&= \prod_{r=0}^{R}\prod_{e} [h_e(s_r)\Delta s_r]^{\Delta N_{ie}(s_r)}[1-h_e(s_r)\Delta s_r]^{1-\Delta N_{i.}(s_r)}.
\end{aligned}
$$

By using the same approximation as for the occurrence process we get

$$
\begin{aligned}
L(\cdot) &\approx \prod_{i=1}^{n}\left(\left(\prod_{j=1}\prod_{e} h_e(v_{ij})\right)\exp\left(-\sum_{e}\int_0^{\tau_i} h_e(s)ds\right)\right) \\
&= \prod_{i=1}^{n}\left(\prod_{j} h_{sep}^{\delta_{ij1}}(v_{ij})h_{se}^{\delta_{ij2}}(v_{ij})h_{p}^{\delta_{ij3}}(v_{ij})\right)\exp\left(-\int_0^{\tau_i}(h_{sep}(s)+h_{se}(s)+h_p(s))ds\right)
\end{aligned}
$$

where v_{ij} is the time passed from notification until event j and n is the number of claims.

The likelihood of the payments is

$$
L(\cdot) = \prod_{i=1}^{n}\prod_{j'} f_P(u_i + v_{ij'}),
$$

where u_i is the delay and $v_{ij'}$ is the time passed from notification until payment j' and f_P is the density for the payments. This is a slight modification from Antonio and Plat (2014) where the likelihood of the payment is $\prod_{i=1}^{n}\prod_{j'} f_P(v_{ij'})$.

We do not consider the possibility of reopening closed claims. This situation could be approached by adding new claims.

4 Estimation

The complete likelihood of the claims reported before τ is

$$
\begin{aligned}
L(\cdot) &= \left(\prod_{i=1}^{n} \lambda(t_i)P(U \leq \tau - t_i)w(t_i)\right) \exp\left(-\int_0^{\tau} \lambda(s)P(U \leq \tau - s)w(s)ds\right) \\
&\times \prod_{i=1}^{n} f_U(u_i|T_i + U_i \leq \tau) \\
&\times \prod_{i=1}^{n} \left(\prod_{j} h_{sep}^{\delta_{ij1}}(v_{ij})h_{se}^{\delta_{ij2}}(v_{ij})h_{p}^{\delta_{ij3}}(v_{ij})\right) \exp\left(-\int_0^{\tau_i} (h_{sep}(s) + h_{se}(s) + h_p(s))ds\right) \\
&\times \prod_{i=1}^{n} \prod_{j'} f_P(u_i + v_{ij'})
\end{aligned}
$$

Estimation of parameters is carried out via maximum likelihood. No robustness analysis of the parameters was performed but to include parameter uncertainty of the delay and payment distribution we use the asymptotic normality of the maximum likelihood estimator, $\hat{\theta}$, according to Casella and Berger (2002). That is, we use that

$$
\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \nu(\theta)),
$$

and as an approximation to $\nu(\theta)$, the variance of $\hat{\theta}$, we use

$$
Var(\hat{\theta}) \approx -\frac{1}{\hat{I}(\hat{\theta})},
$$

were $\hat{I}(\hat{\theta})$ is the observed information number, i.e. the negative second derivative of the log likelihood. Other possibilities of including parameter uncertainty is to apply bootstrap methods and Bayesian assumptions. Due to how computationally heavy the model is, the normal assumption is applied.

We use the Akaike information criteria, or AIC, to select from the fitted distributions, see Akaike (1973):

$$
AIC = 2k - 2\log(L),
$$

where k is the number of estimated parameters and L is the likelihood of the distribution. The distribution with the lowest AIC is chosen.

4.1 Intensities

The exposure and the intensities for the occurrence process are assumed to be constant on intervals $[d_{l-1}, d_l)$ where $l = 1, ..., m$. Let $N^{oc}(l)$ be the number of occurrences on $[d_{l-1}, d_l)$. Then by maximizing

$$\begin{aligned} L(\cdot) &= \left(\prod_{l=1}^{m} (\lambda_l w_l)^{N^{oc}(l)}\right) \left(\prod_{i=1}^{n} P(U \leq \tau - t_i)\right) \\ &\times \exp\left(-\sum_{l=1}^{m} \lambda_l w_l \int_{d_{l-1}}^{d_l} P(U \leq \tau - s) ds\right), \end{aligned}$$

the maximum likelihood estimator of the intensity for the process of occurred claims becomes

$$\hat{\lambda}_l = \frac{N^{oc}(l)}{w_l \int_{d_{l-1}}^{d_l} P(U \leq \tau - s) ds}.$$

The intensity estimators $\hat{\lambda}_l$ are dependent on the estimation of the delay distribution.

4.2 Delay Distribution

The observed data is biased towards shorter delays since we cannot see the complete data, see Parodi (2014). The observed data comes from the distribution

$$P(U \leq u | T + U \leq \tau).$$

The true underlying distribution, $P(U \leq u)$, is needed for the model. The relation between the densities of these distribution is

$$\begin{aligned} f_U(u | T + U \leq \tau) &= \frac{P(T + U \leq \tau | u) f_U(u)}{P(T + U \leq \tau)} \\ &= \frac{P(T \leq \tau - u) f_U(u)}{P(T + U \leq \tau)}. \end{aligned}$$

We assume that $P(T \leq \tau - u)$ is distributed according to the exposure on the interval $[0, \tau]$, i.e.

$$P(T \leq \tau - u) = \frac{\int_0^{\tau - u} w(s)ds}{\int_0^{\tau} w(s)ds}.$$

The exposure is here assumed to be constant on yearly intervals. One could for example add seasonal effects and make more sophisticated assumptions.

The denominator, $P(T + U \leq \tau)$, is difficult to estimate. However, in this case it is only a normalizing factor so derivation of the function is not needed. The parameters of the delay distribution, $P(U \leq u)$, are estimated by maximizing

$$\prod_{i=1}^{n} \frac{\frac{\int_0^{\tau - u_i} w(s)ds}{\int_0^{a} w(s)ds} f_U(u_i)}{P(T + U \leq \tau)}, \; u < \tau.$$

Numerical methods are used to estimate the parameters of the distribution. Since no delay larger than τ has been observed, $P(U \leq u)$ is undefined for $u \geq \tau$. We will assume the same estimated parametric distribution for all u.

4.3 Hazard Rates

We assume that the hazard rates are constant on intervals, l, of length k on $[0, \tau_i]$, where $\tau_i = \min(\tau - t_i - u_i, v_i)$ and v_i is the time from notification until settlement. The likelihood of the development process is

$$\begin{aligned} L(\cdot) &= \prod_{i=1}^{n} \left(\prod_j h_{sep}^{\delta_{ij}^{sep}}(v_{ij}) h_{se}^{\delta_{ij}^{se}}(v_{ij}) h_p^{\delta_{ij}^{p}}(v_{ij}) \right) \\ &\times \exp\left(- \sum_l \int_{\min(k(l-1),\tau_i)}^{\min(kl,\tau_i)} (h_{sep}(s) + h_{se}(s) + h_p(s))ds \right) \\ &= \prod_l h_{sep}^{N_{sep}^{oc}(l)}(l) h_{se}^{N_{se}^{oc}(l)}(l) h_p^{N_p^{oc}(l)}(l) \\ &\times \exp\left(- \sum_{i=1}^{n} \sum_l \sum_e h_e(l) \int_{\min(k(l-1),\tau_i)}^{\min(kl,\tau_i)} ds \right) \end{aligned}$$

where δ_{ij}^{e} is equal to 1 if event j for claim i is of type e and zero otherwise and $N_e^{oc}(l)$ is the number of events of type e on interval l. The likelihood is maximized with estimates

$$\hat{h}_e(l) = \frac{N_e^{oc}(l)}{\sum_{i=1}^{n} \int_{\min(k(l-1),\tau_i)}^{\min(kl,\tau_i)} ds}.$$

where $\int_{\min(kl,\tau_i)}^{\min(k(l+1),\tau_i)} ds$ can be interpreted as the time under risk for event i on interval l. As further discussed in section 5.1, n changes each simulation to take account for claims which have been reported but have not received a payment. For simplification we take $\hat{h}_e(l)$ to be the mean of all its values for the number of simulations performed. Notice that the hazard rates are assumed independent of the delay.

4.4 Payment Distribution

The payment distribution $P_P(u+v)$ is dependent on the time passed the from accident, $u+v$. The amount paid tends to increase as the claim gets older. In order to keep the model simple, we estimate the parameters of the payment distribution separately on k different time-intervals. For $u+v$ larger than the first $k-1$ periods we assume constant parameters. Other possibilities could include dependence on time passed since the claim was reported or a mixture thereof.

5 Predicting RBNS and IBNR claims

For estimating the reserves we have to estimate the RBNS (reported but not settled) claims as well as the IBNR (incurred but not reported) claims. For the IBNR we need to predict the occurrence process, the corresponding delays and the development process, but for the RBNS claims the occurrence process and the delays are already known.

Open claims that have not received a payment need to be considered first since they could be closed without ever receiving payment (become a zero claim). No special interest is taken in modelling zero claims so we therefore adjust for them separately.

5.1 Logistic Regression

The probability of a claim i closing without a payment is modelled with logistic regression. We let

$$S_i = \begin{cases} 1 & \text{if the claim will ever receive a payment,} \\ 0 & \text{if otherwise.} \end{cases}$$

Further, S_i is assumed to be dependent on the time since the claim was reported, i.e. $\tau - (T_i + U_i)$. The logistic regression model is

$$P(S_i = 1 | T_i + U_i \leq \tau) = \frac{\exp(\beta_0 + \beta_1(\tau - (T_i + U_i)))}{1 + \exp(\beta_0 + \beta_1(\tau - (T_i + U_i)))} + \varepsilon_i,$$

where $\beta_0, \beta_1 \in \mathbb{R}$. For simplicity we assume that $\varepsilon_i \sim N(0, \sigma^2)$. Iteratively re-weighted least squares are used to estimate the maximum likelihood estimators $\hat{\beta}_0, \hat{\beta}_1$, see Nelder and Wedderburn (1972). All claims are used for the estimation.

Each one of the only reported claim has a predicted probability $\hat{s}_i$ of getting closed without a payment. By drawing $u \sim U(0, 1)$ the only reported claims will either be settled directly or become a part of the other RBNS claims.

The number of only reported claims that will continue into the development process will both have an effect on the reserves for the RBNS claims and on N_{IBNR}, and thereby, the reserves for IBNR claims.

As previously discussed, the occurrence process where the zero claims are excluded is assumed to be a Poisson process. This process can be seen as a thinned Poisson process where all claims also make up a Poisson process.

5.2 Occurrence

In section 3.1 we saw that the occurrence process of the observed claims on $[d_{l-1}, d_l)$ is assumed to follow a Poisson distribution with intensities

$$\lambda_l w_l \int_{d_{l-1}}^{d_l} P(U \leq \tau - s) ds.$$

In the same way the occurrence process of the IBNR claims on $[d_{l-1}, d_l)$ is assumed to be a Poisson process with intensity

$$\lambda_l w_l \int_{d_{l-1}}^{d_l} 1 - P(U \leq \tau - s) ds.$$

Since the maximum likelihood estimator of λ is

$$\hat{\lambda}_l = \frac{N^{oc}(l)}{w_l \int_{d_{l-1}}^{d_l} P(U \leq \tau - s) ds},$$

the occurrence process of the IBNR claims becomes

$$N_{IBNR}(l) \sim \text{Poisson}\left(N^{oc}(l) \frac{\int_{d_{l-1}}^{d_l} 1 - P(U \leq \tau - s) ds}{\int_{d_{l-1}}^{d_l} P(U \leq \tau - s) ds} \right).$$

Notice that the process is not dependent on the exposure since it cancels out. The occurrences, $T_i \in [d_{l-1}, d_l)$ where $i = 1, ..., N_{IBNR}(l)$ are assumed to have the distribution

$$P(T_i = t_i) = \frac{P(U_i \geq \tau - t_i)}{\sum_{k=d_{l-1}}^{d_l - 1} P(U_i \geq \tau - k)},$$

on each interval $[d_{l-1}, d_l)$. This is a modification from Antonio and Plat (2014) were the occurrence times are assumed to be uniformly distributed on each interval $[d_{l-1}, d_l)$. Here, $[d_{l-1}, d_l)$ is taken to be one year.

5.3 Delay

The delays of the observed claims come from the distribution

$$P(U \leq u | T + U \leq \tau),$$

while the delays of the IBNR claims come from

$$P(U \leq u | T + U > \tau).$$

For each IBNR claim, with simulated occurrence time t_i we draw $p_i \sim U(0, 1)$. Since we have estimated $P(U \leq u)$ (see section 4.2) we simulate u_i by inverting

$$\begin{aligned}
& & P(U \leq u_i | t_i + U > \tau) = p_i, \\
\Rightarrow & & \frac{P(U \leq u_i) - P(U \leq \tau - t_i)}{P(U > \tau - t_i)} = p_i, \\
\Rightarrow & & P(U \leq u_i) = p_i P(U > \tau - t_i) + P(U \leq \tau - t_i).
\end{aligned}$$

5.4 Development

The development process is the most computationally heavy part of the whole prediction process. For each claim the following steps are carried out:

1. Simulate v_{next}. Then, $t_i + u_i + v_{next}$ is the time-point of the next event.
2. Simulate the type of event, *se*, *sep* or *p*.
3. If event includes a payment, simulate payment.
4. If event is not closed, start again from step 1.

For simulating v_{next}, we start by looking at the distribution of $P(V \leq v_{next})$. The cumulative distribution function can be written as a function of the cumulative hazard function

$$P(V \leq v_{next}) = 1 - \exp\left(-\int_0^{v_{next}} \sum_e h_e(s)ds\right).$$

However, the time-point of the next event depends on c, the time passed since the report date, so we draw v_{next} from the distribution

$$P(V \leq v_{next}|V > c).$$

For the first event of the IBNR claims we therefore have $c = 0$. To simulate from $P(V \leq v_{next}|V > c)$ we invert it by drawing $p \sim U(0,1)$ and get

$$\begin{aligned}
& P(V \leq v_{next}|V > c) = p, \\
\Rightarrow \quad & P(V \leq v_{next}) = P(c \leq V) + pP(V > c), \\
\Rightarrow \quad & 1 - \exp\left(-\int_0^{v_{next}} \sum_e h_e(s)ds\right) = P(c \leq V) + pP(V > c), \\
\Rightarrow \quad & \int_0^{v_{next}} \sum_e h_e(s)ds = -\log(1 - (P(c \leq V) + pP(V > c))).
\end{aligned}$$

Now the equation can be solved for v_{next}. We calculate the vector

$$(v_{next}, \int_0^{v_{next}} \sum_e h_e(s)ds),$$

for a large set of v_{next}. When the events are simulated we draw $p \sim U(0,1)$ and calculate $\log(1 - (P(c \leq V) + pP(V > c)))$. Thereafter, we search for v_{next} in the vector.

The type of event is determined by the probability of each event at that particular time-point

$$\frac{h_e(v_{next})}{\sum_e h_e(v_{next})}.$$

If $e \in \{sep,\ p\}$ a payment is drawn from the payment distribution, $P_P(u_i + v_{ij'})$. The process continues until all claims are settled.

5.5 Best Estimate

Summing over all the simulated payments from the RBNS and IBNR claims we get an estimate of the corresponding reserves, R_{RBNS} and R_{IBNR} respectively. R_{RBNS} and R_{IBNR} are simulated M times to obtain an empirical distribution for the total reserve

$$R_j = R_{RBNS,j} + R_{IBNR,j}.$$

In each iteration before the model is recalculated we:

1. Predict which of the claims that have only been reported will be classified as RBNS claims.
2. Draw new parameters from the delay distribution (normality assumption, see section 4).
3. Draw new parameters for the payment simulation (normality assumption, see section 4).

The best estimate is then given by:

$$\widehat{BE} = \frac{1}{M} \sum_{j=1}^{M} R_j.$$

Standard deviation and other risk measures can be calculated from the empirical distribution of the best estimate. In this process, we have accounted for both process and estimation variance. One could though consider parameter uncertainty to a larger extent by looking at the uncertainty of the hazard rates and the logistic regression. Moreover, we have not discounted the best estimate and not considered inflation. The expense reserve is not included in the best estimate.

6 Risk

From the model we obtain a distribution of the best estimate, see section 5.5. In the Solvency II directive the solvency capital requirement (SCR) for the standard formula is based on the metric set as $99,5\%$ Value-at-Risk (VaR) over a one-year horizon. We therefore look at the one-year reserve risk and the one-year premium risk. We use the definition of one-year reserve risk presented in Ohlsson and Lauzeningks (2009) and apply the same method to the premium risk. See also Merz and Wüthrich (2008) for further discussion.

6.1 One-year Reserve Risk

The reserve risk concerns the risk in the run-off result from outstanding claims from past years. That is, the risk that the ultimo cost will deviate from earlier estimations. The one-year reserve risk measures the deviation from the ultimo cost estimated at time t from the ultimo cost estimated one year later, at $t+1$, for claims occurring up to t.

In one-year the ultimo cost will consist of payments made in $(t, t+1]$, denoted with C^{t+1}, and the reserve evaluated at $t+1$, R^{t+1}. The SCR of the reserve is the $99,5\%$ quantile of the loss distribution

$$E[R^t] - (C^{t+1} + R^{t+1}).$$

We take the expected value of R^t since the opening reserve is not considered stochastic, as in Ohlsson and Lauzeningks (2009). $E[R^t]$ is estimated with $\widehat{BE}$ as in section 5.5. In all following sections, we simplify the notation by not distinguishing between the stochastic variables R^t, C^{t+1} and R^{t+1} and their respective simulated estimates as it is clear from the context which of them is referred to. Both C^{t+1} and R^{t+1} can be calculated from the reserving model by the following 3 steps.

1. We start by estimating R^t as before.
2. From the simulation of R^t we calculate C^{t+1} that consists of all payments which occur before $t+1$. Then we have a new set of open claims and the number of claims observed, N_{oc}.
3. Using new parameters for the delay and payment distributions, the reserve, R^{t+1}, is estimated standing at time-point $t+1$.

In step 3, when estimating R^{t+1}, it would be optimal to re-evaluate the model by estimating all model components again (distributions, intensities, logistic regression etc.) using all information available at $t+1$. We do not re-evaluate the model here due to how computationally heavy it would be. Instead, parameter uncertainty is taken into account by drawing new parameters (from their asymptotic normal distributions, see section 4) for the delay and payment distributions to give more spread in the result. No correlation is assumed between the parameters used to estimate C^{t+1} and R^{t+1}.

The way of calculating $C^{t+1}+R^{t+1}$ is consistent with the calculation of R^t. Iterating M times gives an empirical distribution of $E[R^t] - (C^{t+1} + R^{t+1})$. Now, the one-year VaR can be calculated along with other risk measures and statistics.

6.2 One-year Premium Risk

Premium risk is the risk that the cost from next years written polices and costs from unexpired contracts will be larger than expected. In a one-year horizon this implies that the payments of the first year, C^{t+1} and the outstanding reserves R^{t+1} are not larger than $E[R^t]$.

Since premium risk measures the risk for next year, all claims are IBNR. $N_{oc}(t+1) = 0$ and $N_{IBNR}(t+1)$ needs to be forecasted. We use the approximation

$$N_{IBNR}(l+1) \sim \text{Poisson}\left(\frac{w(l+1)N^{oc}(l)}{w(l)}\left(1 + \frac{\int_{d_l}^{d_{l+1}} 1 - P(U \leq \tau - s)ds}{\int_{d_l}^{d_{l+1}} P(U \leq \tau - s)ds}\right)\right),$$

where $l+1$ denotes the segment $(t, t+1]$. The number of claims is expected to change proportionally with the exposure. One could also use historical data further back to predict $N_{IBNR}(t+1)$, however last years exposure is here assumed sufficient for the prediction. This is just an illustration of the technique and other more sophisticated ways of modelling $w(t)$ would be used in practice. The following steps are carried out to calculate $C^{t+1}+R^{t+1}$. Observe that steps 3-4 are identical to steps 2-3 in the reserve risk calculation.

1. Draw number of claims for next year $N_{IBNR}(l+1)$ from the Poisson distribution.

2. Estimate R^t with the new $N_{IBNR}(l+1)$.

3. From the simulation of R^t we calculate C^{t+1} which consists of all payments that occur before $t+1$. We have a new set of open claims and the number of claims observed, N^{oc}.

4. Using new parameters for the delay and payment distributions, the reserve, R^{t+1}, is estimated standing at time-point $t+1$.

As for the reserve risk the optimal way would be to estimate the new parameters in step 4 from data until time-point $t+1$. To include correlation between reserve risk and premium risk one could simulate C^{t+1} simultaneously for both the reserve risk and the premium risk and then re-evaluate the model. Due to computation time that was not done. The parameters are drawn from their distribution assuming no correlation between R^{t+1} and C^{t+1} or between reserve and premium risk. By iterating steps 1-4 M times we obtain an empirical distribution for $C^{t+1}+R^{t+1}$. Thereafter, risk statistics easily follow.

7 Chain-Ladder

The following section on the chain-ladder method is only supposed to give a short description of how the chain-ladder algorithm works. For further information, see Mack (1993). Classical reserving methods are based on aggregated data triangles. Data is aggregated by accident period i and development period j where $i, j = 1, ..., I$. An example of an aggregated data triangle is

$$\begin{array}{lll} C_{1,1} & C_{1,2} & C_{1,3} \\ C_{2,1} & C_{2,2} & \\ C_{3,1} & & \end{array}$$

where $C_{i,j}$ is the accumulated total payment from accident period i payments until development period j. The accident periods are assumed to be independent. The underlying assumption of the chain-ladder method is that the same percentage of payments are paid out in each development period, i.e.

$$E[C_{i,j+1}|C_{i,1}, ..., C_{i,j}] = C_{i,j} f_j,$$

for $1 < i < I, 1 \leq j \leq I - 1$. The development factors, $\hat{f}_j$, are calculated as

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j} C_{i,j+1}}{\sum_{i=1}^{I-j} C_{i,j}}, \; j = 1, ..., I - 1.$$

By using the development factors we can estimate the final cumulative payment for each claim year,

$$\hat{C}_{i,n} = \hat{f}_j \cdots \hat{f}_{n-1} C_{i,j},$$

and the reserve $\hat{R} = \sum_i \hat{C}_{i,n} - C_{i,i}$.

An underlying variance assumption of the chain-ladder method by Mack (1993) is

$$\mathrm{Var}[C_{i,j+1}|C_{i,1}, ..., C_{i,j}] = \sigma_j^2 C_{i,j} \tag{1}$$

for $1 < i < I, 1 \leq j \leq I - 1$.

The mean square error (MSE) of the estimator of the reserve $\hat{R}_i$ (also called mean square error of prediction, MSEP) is derived analytically in

Mack (1993). The MSE consists of the process variance and the estimation variance,

$$\text{MSE}[\hat{R}_i] = \text{Var}[C_{i,I}|D] + (E[C_{i.I}|D] - \hat{C}_{i,I})^2$$

where $D = \{C_{i,j}|i+j \leq i+1\}$ is the observed data.

Various stochastic versions of the chain-ladder have been proposed, see England and Verrall (2002). Here, we will bootstrap Mack's model to obtain an estimate of the uncertainty, see England and Verrall (2006).

From equation 1 we have for the individual development factors, $f_{i,j} = C_{i,j+1}/C_{i,j}$, that

$$\begin{aligned} E[f_{i,j}|C_{i,j-1}] &= f_j, \\ \text{Var}[f_{i,j}|C_{i,j-1}] &= \frac{\sigma_j^2}{C_{i,j-1}}. \end{aligned}$$

For an estimate of σ_j we use the unbiased estimator from Mack (1993)

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=1}^{I-j} C_{i,j} (\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j)^2,$$

for $1 \leq j \leq I-2$ and

$$\hat{\sigma}_{I-1}^2 = \min(\hat{\sigma}_{I-2}^4/\hat{\sigma}_{I-3}^2 \min(\hat{\sigma}_{I-3}^2, \hat{\sigma}_{I-2}^2)).$$

See Mack (1993) for justification of $\hat{\sigma}_{I-1}^2$. The scaled Pearson residuals of the model are

$$r_{i,j} = \frac{\sqrt{C_{i,j}}(f_{i,j} - \hat{f}_j)}{\sqrt{\hat{\sigma}_j}}.$$

For simplicity we do not consider any bias adjustments as suggested in England and Verrall (2006). Further the residual $r_{1,n}$ is excluded from the simulations, since it is only zero. By bootstrapping the residuals $r_{i,j}^b$ we obtain the individual development factors

$$f_{i,j}^b = r_{i,j}^b \frac{\hat{\sigma}_j}{\sqrt{C_{i,j}}} + \hat{f}_j.$$

From $f_{i,j}^b$ we can estimate the bootstrap development factors

$$\hat{f}_j^b = \frac{\sum_{i=1}^{n-j+1} C_{i,j} f_{i,j}^b}{\sum_{i=1}^{n-j+1} C_{i,j}}.$$

To obtain the predictive distribution we need to add a forecasting step to include the process error. England and Verrall (2006) suggests

$$C_{i,n-i+2}^* | C_{i,n-i+1} \sim \text{Normal}(\hat{f}_j^b C_{i,n-i+1}, \hat{\sigma}_j^2 C_{i,n-i+1}) \tag{2}$$

for $i = 2, ..., n$ and

$$C_{i,j}^* | C_{i,j-1}^* \sim \text{Normal}(\hat{f}_j^b C_{i,j-1}^*, \hat{\sigma}_j^2 C_{i,j-1}^*)$$

for $i = 3, .., n$ and $j = n - i + 3, ..., n$. With a normal assumption there is a possibility for simulating negative claims. When estimating the one-year risk we later assume a log-normal distribution. Therefore, we will instead of the normal assumption assume a log-normal distribution with estimates derived from the assumed mean and variance under the normal assumption.

From the simulated C^* we calculate the bootstrap estimates of the reserve $\hat{U}^b$ and obtain an empirical distribution of the estimated reserve.

An analogous measure of the MSEP is the variance of the bootstrapped simulated distribution, see discussion in England and Verrall (2006).

When estimating the reserve with the chain-ladder method, we take the accident and development period to be one year and use $B = 10\ 000$ simulations.

7.1 One-year Reserve Risk

As in section 6.1 we follow the definition of one-year reserve risk given by Ohlsson and Lauzeningks (2009). We also follow the description of how the one-year risk can be calculated. Again, we want to estimate the distribution of

$$E[R^t] - (C^{t+1} + R^{t+1}).$$

In Ohlsson and Lauzeningks (2009) several possibilities are suggested to simulate the new diagonal, $C^{t+1} = \sum_{i=2}^{n} (C_{n-i+2,i} - C_{n-i+1,i})$. One of them is to simulate next years payment from a log-normal distribution with a mean given by the chain-ladder estimate and the variance given by equation 1.

B estimates of $\hat{f}_j^b$ are simulated by bootstrapping the triangle as described in 7. In each simulations, a new diagonal is simulated using equation 2 giving an estimate of C^{t+1} and with the new diagonal the chain-ladder factors are re-estimated giving the values of R^{t+1}.

7.2 One-year Premium Risk

The premium risk cannot be estimated in the same way as in section 6.2 since R^t cannot be simulated directly with the chain-ladder method. Instead, we follow the simplified premium risk definition of Ohlsson and Lauzeningks (2009),

$$P - E - (C^{t+1} + R^{t+1})$$

where P is the earned premium expected and E is the operating expenses which both are considered non-random. Since P and E are considered constant we only look at the distribution of the predicted claim cost,

$$C^{t+1} + R^{t+1}.$$

In Ohlsson and Lauzeningks (2009), several possibilities are suggested to simulate C^{t+1}. One of them is to use the corresponding loss ratio with volatility estimated as in QIS4. For simplicity, and to keep consistency between the reserve and premium risk estimation, we assume a log-normal distribution for the payments, C^{t+1}. The mean and variance are estimated by fitting linear regression through the historical loss ratios. Unlike the reserve risk the mean and variance of C^{t+1} is constant for all simulations. Having drawn B estimates of C^{t+1} we can calculate R^{t+1} by using the B development factors simulated as in section 7.

8 Data

The claim-data comes from a subset of a portfolio and insurance contracts from a Swedish insurance company. The original claim data is not described further as it is not necessary for this type of analysis. The data consists of claim payments for 9 consecutive years. In the dataset there are around 11 000 unique claims. The claims can both be settled directly as well as receive multiple payments before being settled. In total the dataset contains around 22 000 events. In extensions to these claims we also use information from around 4 000 zero claims. The results are presented in a fictional currency for reasons of confidentiality. Trends in the data are viewed as changes in percentage for the same reasons.

Table 1 includes a short description of the data variables. Corresponding stochastic variables are denoted with capital letters, T_i, U_i etc.

The data has been modified beforehand to adjust for negative payments. If a negative payment occurs and there exists a payment with a matching positive amount the payments are cancelled out. The latest payment with a matching positive amount is cancelled out. If no payment matches the negative amount the amount is subtracted from previous payments. The data is not inflation adjusted.

Table 1: Variables used from claim records. Corresponding stochastic variables are denoted with capital letters, T_i, U_i etc.

Variable	Description
t_i	Time of occurrence of claim i.
u_i	Delay from the time of occurrence until the insurer is notified.
v_{ij}	Time from notification of claim i until event j.
e_{ij}	Type of event j for claim i.
$p_{ij'}$	Payment of event j' for claim i.
$w(t)$	Exposure, gross earned premium.

9 Results

We start by stating the assumptions and distribution choices made for the model and present an analysis of trends in the data. We show the result of a study on the amount of simulations needed in order to reach convergence. The results from the model are compared to results from the chain-ladder method. For further validation of the model we back-test the model by comparing predicted one-year payments with observed payments and perform a sensitivity analysis of the assumptions and model choices made in the beginning. Finally, we present the results of the one-year reserve and premium risk.

9.1 Model Assumptions and Estimation

We make the following model assumptions:

1. **Logistic regression**
   A plot of the logistic regression is displayed in Figure 1. In each simulation, each claim which has only been reported but has not received a payment is either classified as RBNS claim or a zero claim.

2. **The delay distribution**
   For simplification, the only reported claims (which have never received any payment) are not used for the estimation of the delay distribution.

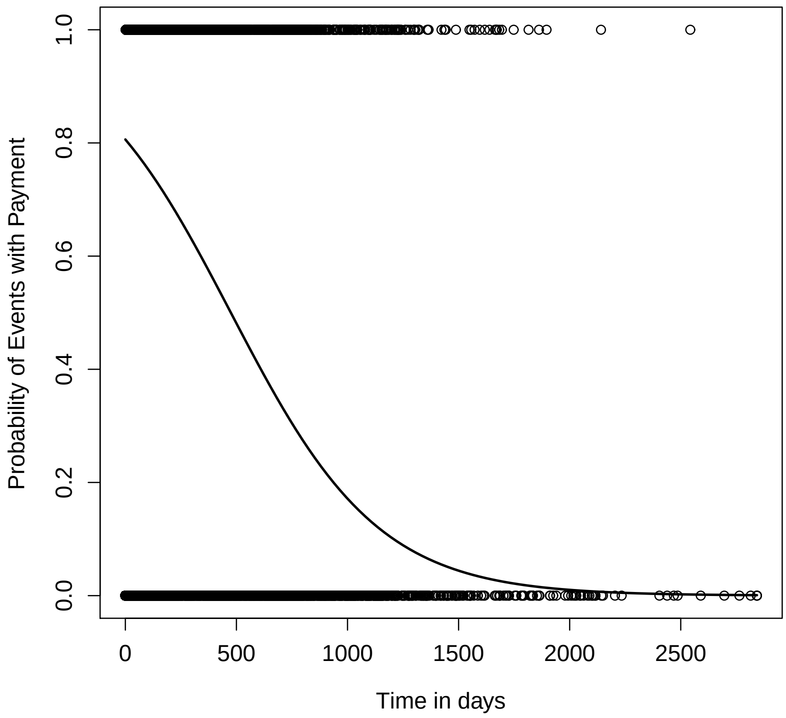


Figure 1: A plot of the logistic regression. The x-axis depicts the time from reporting until either an event including a payment (1) or closing of a claim without a payment (0). The plot also includes the events the logistic regression is estimated from.

The only reported claims are around 3% of the claims in each simulations. The mean and standard deviation of their delays do not show large deviation from the delays of the other reported claims.

We assume a log-normal distribution for the delays. The log-normal distribution gives the lowest Akaike information criterion (AIC) of the Weibull, exponential and gamma distributions, which were the tested distribution. Due to the heavy tail of the log-normal distribution we apply a 30 year cap on the reporting delay.

Numeric problems occur for the fitting of the gamma distribution which lead to exclusion from the distribution choices. A plot of the fitted distribution is displayed in Figure 2.

3. **Intensities**

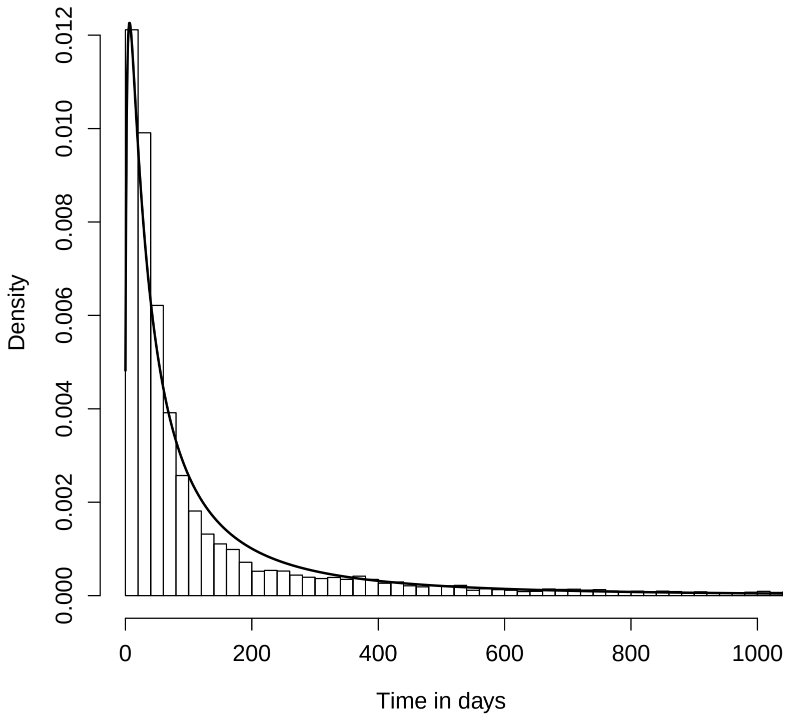


Figure 2: A histogram of the observed claims. The estimated distribution is also plotted. The estimated distribution is, as discussed in section 4.2, adjusted for the bias in the observed data. We here only view the plot for delays shorter than 1000 days to get a clearer plot.

The intensities depend on the delay distribution and the number of RBNS and closed claims. The intensities are stochastic because in each simulation the number of RBNS claims changes to adjust for zero claims and we draw new parameters for the delay distribution from their asymptotic distributions (see section 4). The ratio of the means $\hat{\lambda}_i/\hat{\lambda}_1$ are displayed in Figure 3a.

4. **The payment distribution**
The payments are assumed to be log-normally distributed. Weibull and gamma distributions were even tested but gave a higher AIC. The distributions are constant on intervals of 365 days with the same distribution after 4 intervals. Numerical problems occur with the fitting of the gamma distribution which lead to exclusion from distribution

choices.

5. **Hazard rates**
   The hazard rates are assumed to be constant on intervals of 60 days. After 20 intervals we assume no further changes in the hazard rates. See further discussion on the estimation in section 4.3. The hazard rates are displayed in Figure 3.

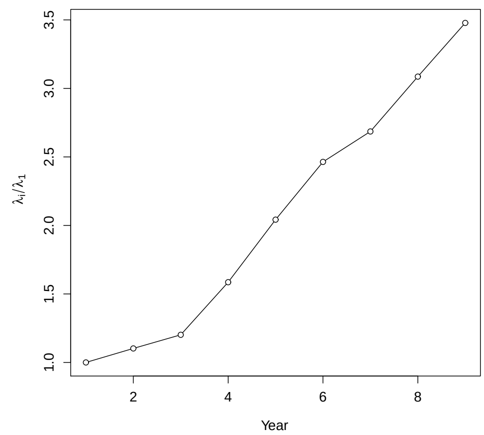
9.2 Trends in the Data

By investigating how the estimated parameters change when the underlying data is constrained to shorter time-intervals, we gain understanding of underlying trends in the data. We look at the delay, hazard rate and payment amount. We also take a look at and their amounts and number of open and closed claims.

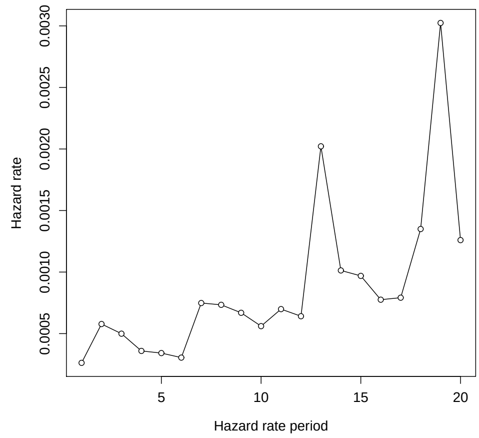
In Table 2 we see the change in mean delay by occurrence and reporting year. The means have been normalized by the mean from the first reporting year of occurrence year x. The average reporting delay has decreased visibly for the claims that are reported the first year. Trends during later years are not obvious.

Table 2: The change in mean of the delay depending on occurrence year and reporting year. The delays have been normalized by the mean from the first reporting year of occurrence year x. The average reporting delay has decreased visibly for the claims that were reported the first year. Trends during later years are not as visible.

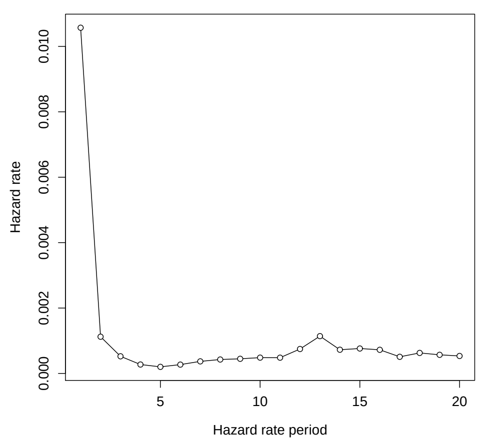
	Reporting year								
Occurrence year	1	2	3	4	5	6	7	8	9
x	1.00	3.39	11.66	17.56	24.32	34.71	37.93	43.53	50.39
$x+1$	1.00	4.29	11.75	18.85	24.00	30.15	36.20	43.90	
$x+2$	1.02	3.47	11.58	17.83	25.12	30.80	37.46		
$x+3$	1.00	3.81	11.58	17.86	23.83	29.58			
$x+4$	0.97	3.78	12.27	17.31	24.03				
$x+5$	0.92	4.20	11.63	18.31					
$x+6$	0.76	3.95	12.14						
$x+7$	0.73	3.53							
$x+8$	0.75								



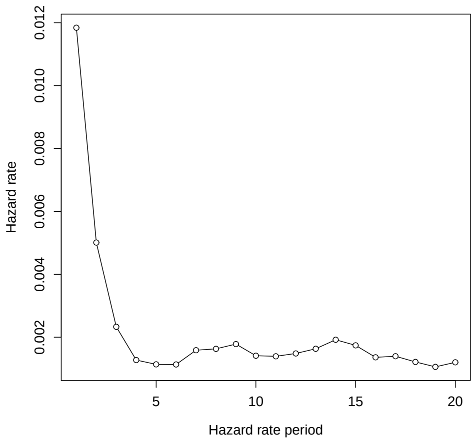
(a) $\hat{\lambda}_i/\hat{\lambda}_1$.



(b) Settled without a payment.



(c) Settled with a payment.



(d) Payment.

Figure 3: Figure 3a displays the ratio of the intensities, $\hat{\lambda}_i/\hat{\lambda}_1$. Figures 3b-3d display the estimated hazard rates. The hazard rates are assumed to be constant on 60 days. After 20 periods we assume no further changes in the hazard rates.

The results from analysing the change in payment amounts are presented in Table 3. The results are normalized in the same way as the delay. We see an increasing trend with longer time passed from the accident. There is not a clear trend across occurrence years.

Table 4 shows the relative changes in the hazard rates. The hazard rates

Table 3: The mean of the payments depending on occurrence year and development year. The payments have been normalized by the mean from the first reporting year of occurrence year x. We see an increasing trend with longer time passed from the accident. There is not a clear trend across occurrence years.

	Reporting year								
Occurrence year	1	2	3	4	5	6	7	8	9
x	1.00	1.76	3.4	3.55	1.76	6.54	1.17	6.96	2.56
$x+1$	0.92	2.05	3.49	3.57	5.52	4.37	5.43	2.69	
$x+2$	1.22	2.35	3.7	3.1	4.7	4.02	3.12		
$x+3$	1.32	2.06	3.29	2.81	4.47	3.07			
$x+4$	1.61	1.96	2.72	3.58	3.59				
$x+5$	1.16	2.19	3.1	2.85					
$x+6$	1.09	2.06	2.74						
$x+7$	1.6	2.34							
$x+8$	1.25								

are estimated with data from the different occurrence years and here we assume no further changes in the hazard rates after 5 intervals (whereas in the model it is 20 intervals). The hazard rates have been normalized by the corresponding hazard rates from occurrence year x and the first hazard rate period. The hazard rates for settled with a payment events have increased while the other types do not show as clear a trend.

Another interesting characteristic of the dataset being analysed is the amount of open and closed claims throughout occurrence and development years. In Table 5 we present an overview of the amount of open and closed claims. We see significant changes between the years. More claims are settled earlier which leads to less open claims. A large deviation is observable in the year $x+8$. Note that the table values have not been adjusted for changes in exposure. Zero claims have been adjusted for (this is only one possible outcome as the number of open claims changes each simulation).

Table 4: Relative changes in the hazard rates. The hazard rates are estimated with data from the different occurrence years and here we assume no further changes in the hazard rates after 5 intervals. The hazard rates are normalised with $h_e(1)$ estimated on occurrence year x. The hazard rates for settled with a payment events have increased while the other types do not show as clear a trend.

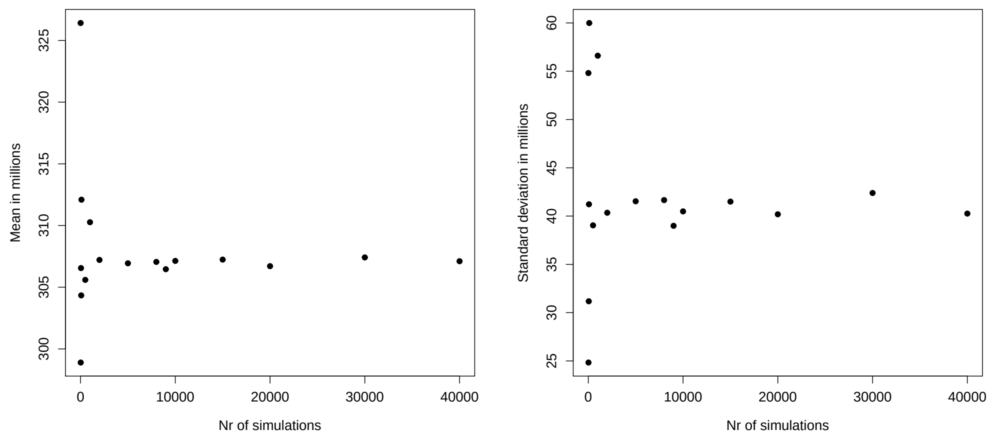
		Hazard rate period				
Type	Occurrence year	1	2	3	4	5
Se	x	1.00	2.04	1.94	1.41	4.98
	$x+1$	1.18	1.95	1.19	1.92	5.14
	$x+2$	1.38	1.51	2.05	3.15	5.06
	$x+3$	1.50	1.69	0.83	1.57	4.65
	$x+4$	1.08	1.32	1.15	1.19	4.18
	$x+5$	0.61	1.89	2.87	1.54	4.54
	$x+6$	0.89	2.55	3.34	1.37	4.28
	$x+7$	1.95	6.56	5.04	3.49	2.93
	$x+8$	4.00	16.89	14.06	4.38	6.54
Sep	x	1.00	0.14	0.08	0.02	0.11
	$x+1$	0.84	0.13	0.08	0.03	0.11
	$x+2$	0.92	0.12	0.06	0.08	0.13
	$x+3$	1.42	0.20	0.12	0.06	0.12
	$x+4$	1.83	0.18	0.09	0.06	0.10
	$x+5$	2.22	0.21	0.07	0.03	0.11
	$x+6$	2.32	0.29	0.14	0.06	0.11
	$x+7$	3.53	0.37	0.20	0.11	0.10
	$x+8$	6.57	0.91	0.40	0.22	0.13
P	x	1.00	0.41	0.19	0.09	0.13
	$x+1$	1.06	0.36	0.16	0.11	0.12
	$x+2$	0.96	0.37	0.15	0.08	0.12
	$x+3$	1.00	0.41	0.18	0.13	0.14
	$x+4$	0.99	0.38	0.18	0.09	0.11
	$x+5$	0.98	0.35	0.19	0.11	0.10
	$x+6$	0.84	0.47	0.22	0.12	0.10
	$x+7$	0.95	0.39	0.17	0.09	0.10
	$x+8$	0.83	0.57	0.30	0.10	0.10

Table 5: The number of settled and open claims. Here, we see a significant change between the years. More claims are settled earlier which leads to less open claims. A large change is observable the year $x + 8$. Note that the table values have not been adjusted for changes in exposure. Zero claims have been adjusted for (this is only one possible outcome as the number of open claims changes each simulation).

	Development year								
Occurrence year	1	2	3	4	5	6	7	8	9
Settled claims									
x	1.00	1.72	1.95	0.93	0.34	0.29	0.43	0.13	0.17
$x+1$	1.18	1.70	2.34	0.97	0.47	0.60	0.19	0.18	
$x+2$	1.08	2.30	1.93	1.41	0.66	0.72	0.20		
$x+3$	2.29	1.81	3.09	1.96	0.93	0.66			
$x+4$	3.00	2.82	2.56	3.32	1.37				
$x+5$	4.75	3.16	3.33	4.19					
$x+6$	5.76	3.89	5.13						
$x+7$	8.69	6.53							
$x+8$	13.41								
Open claims									
x	1.00	0.65	0.34	0.19	0.15	0.11	0.04	0.03	0.01
$x+1$	1.16	0.82	0.38	0.25	0.20	0.07	0.04	0.02	
$x+2$	1.16	0.77	0.44	0.26	0.16	0.04	0.02		
$x+3$	1.37	1.14	0.59	0.29	0.16	0.06			
$x+4$	1.71	1.41	1.00	0.427	0.21				
$x+5$	2.00	1.63	1.07	0.32					
$x+6$	2.06	1.64	0.70						
$x+7$	2.03	1.21							
$x+8$	0.97								

9.3 Determining the Number of Simulations

As previously mentioned, the model is computationally heavy. We aim to have enough simulations for a stable model and to diminish simulation error. We run the model multiple times for an increasing number of simulations, M, and then examine for which M the best estimate and the standard deviation start to stabilize. A new simulation seed is set for each choice of M to eliminate dependent results. Figure 4 shows how the results are stabilized at $M = 10\ 000$. We use 10 000 simulations in all following calculations.



(a) Mean. (b) Standard Deviation.

Figure 4: An analysis of the amount of simulations needed for the study. After 10 000 simulations both the mean and the standard deviation have stabilized.

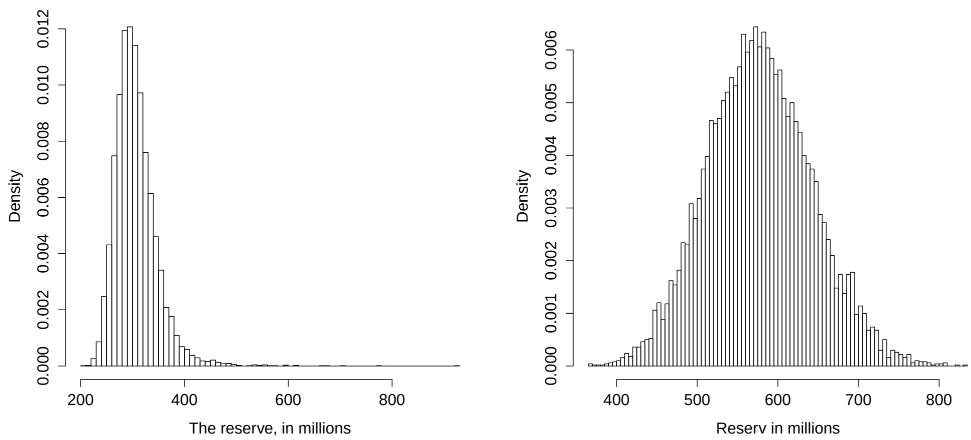
9.4 Comparison to Chain-Ladder

The reserve estimated with the micro model is compared to the reserve estimated with the chain-ladder method. Chain-ladder is widely used for benchmarking. However, in practice one might chose another method for estimating the reserve.

In Table 6 we see the reserve and its standard deviation estimated with both the micro reserving model and chain-ladder method. The chain-ladder estimate is considerably larger than the estimated reserve from the micro reserving model. These two models have large structural differences, which with a closer look explains the different outcomes.

Table 6: The estimated reserve and its standard deviation of the micro reserving method and chain-ladder. There are large structural differences between the methods which lead to the different estimated reserves.

	Micro model	Chain-ladder
RBNS	157 963 799	x
IBNR	148 427 303	x
Reserve	306 391 102	553 542 901
Standard deviation	40 638 609	65 032 603
99.5% quantile	469 387 749	753 818 599



(a) The reserve, micro model.

(b) The reserve, chain-ladder.

Figure 5: Figure 5a shows the empirical density of the reserve estimated with the micro model. The expected value is 306 millions. Figure 5b shows the empirical density of the reserve estimated with the chain-ladder method which has expected value 554 millions. The densities show large differences.

The underlying assumption of the chain-ladder that the same percentage of payments are made each development period between accident periods is not fulfilled. Both reporting time and the time from reporting to settlement have decreased during the observed period. With the micro model the delay distribution and the hazard rates are estimated from individual claims. Volume changes are captured with the micro reserving model, while with the version of the chain-ladder used that is not the case. With the increasing volume, earlier years do not have as a large effect as later years.

As seen in section 9.2 there are large changes in the number of open and settled claims towards the end of the observed period. The chain-ladder only takes into account the total sum paid out, and not how many claims are open. The micro reserving model takes this into account as well as how many claims are settled. The last development year has a large decrease in the number of open claims compared to the other years. This directly affects the size of the RBNS reserve of the micro model. The large number of settled claims increases the number of IBNR claims estimated by the model.

With shortening delays it can be a fact that the main mass of the claims will have a shortened reporting time while the claims with a really long reporting time represent different kind of claims where we do not expect a drastically shortened reporting time. In the micro model we have violated the assumptions of independent observations since there is a downward trend in the reporting time. These trends are smoothed out by using a single fitted distribution. However, separate modelling of the delays increases the flexibility of the model and allows for special adjustments from experts.

Another fact which needs to be taken into consideration is that the hazard rates are assumed to be independent of the delay. To analyse if that is the case we assume that the hazard rates are constant after 5 intervals, and look at the hazard rate estimates estimated from disjoint subsets of the data. The data is split up depending on the reporting day with splitting points being 50, 300 and 800 days. In Table 7 we see the results. We see trends with an increasing delay in the data. However since there are also trends across time in the data it is not possible to conclude whether these are because of shortened settlement time or because of a dependency between delay and payment pattern. The payments, however, are dependent on the time passed since the accident. That is not a case in Antonio and Plat (2014). With this data, it gave poor results to have the payment distributions only dependent on the reporting time. Having the payment distribution only dependent on the reporting time lead to an underestimation of the IBNR reserve when the results were back-tested.

The comparison to the chain-ladder is to some extent naive since one may for example adjust for the trend changes by excluding less relevant years when estimating the development factors or base the analysis on incurred data triangles. The micro reserving method on the other hand is a model which, by estimation of the parameters of the micro reserving method on different time intervals of the underlying data, gives detailed information about the characteristic changes of the reserve. The development factors of the chain-ladder method only give information about the percentage of the reserve paid in a specific period, while the micro model delivers information about reporting delay, payment amounts and payment pattern.

Table 7: The hazard rates are assumed to be constant after 5 intervals and are estimated on disjoint subsets of the data. The data is split up depending on the reporting day with splitting points being 50, 300 and 800 days. We see trends with an increasing delay in the data. However, since there are also trends across time in the data it is not possible to conclude whether these are because of shortened settlement time or because of a dependency between delay and payment pattern. The hazard rates are normalized with $h_e(1)$ estimated on $u \leq 50$.

		Hazard rate period				
Type	Delay	1	2	3	4	5
Se	$u \leq 50$	1.00	1.92	1.61	1.06	3.10
	$50 < u \leq 300$	0.80	2.36	1.85	1.39	3.62
	$300 < u \leq 800$	1.76	3.90	4.28	3.23	4.70
	$u > 800$	2.33	3.95	4.80	4.84	4.21
Sep	$u \leq 50$	1.00	0.09	0.03	0.02	0.04
	$50 < u \leq 300$	0.91	0.08	0.04	0.02	0.05
	$300 < u \leq 800$	0.96	0.19	0.12	0.07	0.08
	$u > 800$	0.95	0.33	0.22	0.12	0.06
P	$u \leq 50$	1.00	0.35	0.16	0.08	0.10
	$50 < u \leq 300$	0.88	0.35	0.17	0.10	0.11
	$300 < u \leq 800$	0.69	0.62	0.28	0.20	0.15
	$u > 800$	0.60	0.70	0.41	0.18	0.12

In Antonio and Plat (2014), the micro reserving model was compared to three versions of the chain-ladder. Antonio and Plat (2014) conclude that the tested aggregate methods tend to overestimate the reserve. However, Antonio and Plat (2014) do not discuss the reason for this or investigate the underlying trends in the data that lead to this result.

9.5 Back-testing

To validate the model further we look at the performance of the model when τ is set equal to prior years and then compare payment outcomes with true outcomes. We want to emphasize that true outcome may have been somewhat unexpected. We test three variants. First we use the whole dataset to estimate the parameters. Secondly, the parameters are estimated with data until τ. As a simplification, the logistic regression is always estimated with the whole dataset. The third variation is the same as the first scenario with the exception of a cap on each payment. The IBNR claims is sensitive to the estimate of the delay distribution since it affects the number of IBNR claims as well as whether they will be reported within one year. The results are presented in Table 8. The micro model gives us an empirical predictive distribution for the RBNS, IBNR and the total one-year payment. In addition to looking at the expected one-year payments from the models we look at in which percentile of the empirical predictive distribution the true outcome and the chain-ladder prediction lie. Notice that the year $x + 7$ had a low one year payment compared to previous years so an extreme percentile is not unexpected. For both the micro model and the chain-ladder method we can see from the second variant that the trends in the data lead to overestimated one-year payments. The predictions are of course closer to the observed payments when the whole dataset is used. A cap on the payments for the micro model has little effect.

9.6 Sensitivity Analysis

The micro model requires certain assumptions about distributions and other parameter choices. A sensitivity analysis examines how the results vary when these assumptions are altered. It gives information about where further analysis is required as well as information about which factors are possible risk drivers of the reserve estimation. In Table 9, we see the results from the sensitivity analysis on the distributions chosen for the delay and the payment distributions. The top line gives the result of the model choices. The reserve vary largely with the different choices. The difference is clearest for the two tested payment distribution.

9.7 Risk

We compare the one-year reserve and premium risk for the micro reserving method and the chain-ladder. When estimating the premium risk we have not estimated R^t with the chain-ladder method. We compare the distribution of $C^{t+1} + R^{t+1}$ instead. In addition, distribution of R^t is compared to the one-year risk. In Table 10 we see the results the reserve risk and in table 11 we see the results for the premium risk. In table 10, the standard deviation of $E[R^t] - (C^{t+1} + R^{t+1})$ are lower for the micro model than the chain-ladder. This is not unexpected since the micro model uses richer data for all estimation. However, VaR is similar for both models. We did not re-estimate the parameters for the micro model, due to how computationally heavy it is, which can lead to an underestimation. We have taken the parameter uncertainty into consideration by drawing new parameters from the asymptotic normal distribution of the delay and payment parameters.

Table 8: Comparison of predicted payments for the next year to true outcomes. We test three variants. 1. The whole dataset is used for parameter estimation. 2. The parameters are estimated with data until τ, with an exception of the logistic regression. 3. The same as 1 except with a cap on each payment. In addition to looking at the expected one-year payments from the models we look at in which percentile of the empirical predictive distribution of the micro model the true outcome and the chain-ladder prediction lie.

True outcomes from the data, $\cdot 10^3$				
τ	RBNS	IBNR	Total	
$x+3$	64 491	18 810	83 301	
$x+4$	76 384	25 272	101 657	
$x+5$	85 754	26 082	111 836	
$x+6$	111 458	40 097	151 554	
$x+7$	92 203	43 688	135 890	
1. All data used for the estimation				
τ	RBNS (%)	IBNR (%)	Total (%)	Chain-ladder(%)
$x+3$	56 843 (81%)	17 981 (67%)	74 824 (80%)	76 700 (64%)
$x+4$	78 862 (51%)	23 214 (74%)	102 076 (57%)	108 254 (72%)
$x+5$	94 535 (32%)	29 218 (37%)	123 753 (27%)	120 075 (48%)
$x+6$	112 706 (55%)	33 744 (85%)	146 450 (67%)	133 253 (26%)
$x+7$	137 213 (0%)	37 463 (84%)	174 676 (1%)	180 945 (68%)
2. Limited data used for the estimation				
τ	RBNS (%)	IBNR (%)	Total (%)	Chain-ladder(%)
$x+3$	79 103 (16%)	18 521 (63%)	97 624 (19%)	93 538 (47%)
$x+4$	107 719 (1%)	23 972 (69%)	131 691 (3%)	126 008 (44%)
$x+5$	120 642 (1%)	30 787 (25%)	151 429 (1%)	133 448 (19%)
$x+6$	133 497 (11%)	34 217 (85%)	167 714 (24%)	137 435 (5%))
$x+7$	152 235 (0%)	39 023 (80%)	191 258 (0%)	196 284 (65%)
3. All data with cap				
τ	RBNS (%)	IBNR (%)	Total (%)	Chain-ladder(%)
$x+3$	55 829 (82%)	17 834 (67%)	73 663 (82%)	76 700 (65%)
$x+4$	77 306 (51%)	22 924 (74%)	100 230 (58%)	108 254 (75%)
$x+5$	92 793 (33%)	28 841 (37%)	121 635 (27%)	120 075 (49%)
$x+6$	110 674(56%)	33 382 (85%)	144 057 (70%)	133 253 (27%)
$x+7$	134 384(0%)	36 905 (84%)	171 289 (1%)	180 945 (72%)

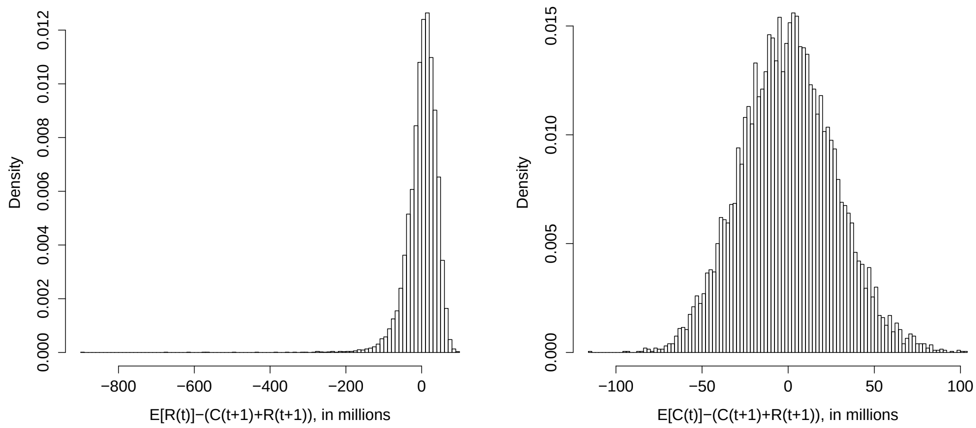
Table 9: A sensitivity analysis of the delay and payment distributions. A log-normal distribution gave the lowest AIC for the both the delay and the payment distribution and was the model choice. The reserve is sensitive to the tested distribution choices.

Delay distribution	Payment distribution	Reserve	sd(Reserve)
Log-normal	Log-normal	306 391 102	40 638 609
Weibull	Log-normal	292 146 747	38 241 106
Exponential	Log-normal	270 914 305	35 775 515
Log-normal	Weibull	230 250 258	13 418 377

Interestingly, the standard deviation of R^t is larger than the standard deviation of $C^{t+1} + R^{t+1}$ for the chain-ladder contrary to the micro model. This is due to how the one-year risk of the chain-ladder is constructed. The predictive distribution of R^t includes a process variance while future payments of the one-year risk include the variability of the re-estimated parameters. The main conclusion is that the risk estimation follows naturally from the micro model since there we simulate from a model while the bootstrapped predictive estimation of the predictive distribution of the chain-ladder method does not as naturally lead to a one-year risk model. In Figure 6 we see the density of $E[R^t] - (C^{t+1} + R^{t+1})$.

Table 10: The table shows the results for the reserve risk. The standard deviation is lower for the micro model than the chain-ladder. This is not unexpected since the micro model uses richer data for all estimation. The VaR is similar for $-(E[R^t] - (C^{t+1} + R^{t+1}))$. However, for the micro model we did not re-estimate the parameters due to how computationally heavy the model is which can lead to underestimation. We have taken the parameter uncertainty into consideration by drawing new parameters from the asymptotic normal distribution of the delay and payment parameters.

	Micro reserving		Chain-ladder	
$E[R^t]$	307 016 029		577 209 139	
	Sd	99.5 %	Sd	99.5 %
R^t	40 691 947	456 611 574	65 032 603	753 818 599
$C^{t+1} + R^{t+1}$	42 485 811	471 814 802	58 883 937	739 545 155
$-(E[R^t] - (C^{t+1} + R^{t+1}))$	42 485 811	-164 798 773	58 883 937	-162 336 016



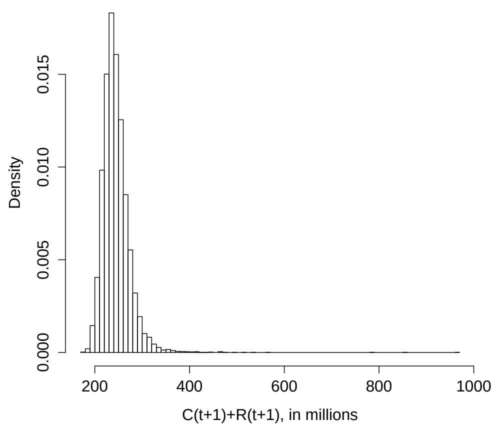
(a) One-year reserve risk, micro model. (b) One-year reserve risk, chain-ladder.

Figure 6: The densities of $E[R^t] - (C^{t+1} + R^{t+1})$. The micro model has a longer lower tail, which means that unique larger losses are simulated.

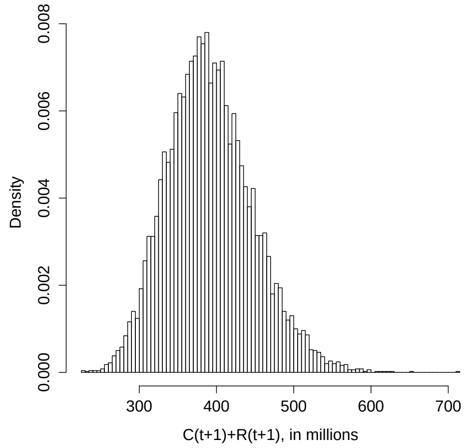
In Table 11, the standard deviation and VaR of $C^{t+1} + R^{t+1}$ is lower for the micro model than the chain-ladder. As for the one-year risk this is not unexpected due to the data usage. Due to the construction of the risk model for the chain-ladder method, no estimates are observed for the variables R^t and $C^{t+1} + R^{t+1}$. Figure 7 shows the densities of $C^{t+1} + R^{t+1}$ and $E[R^t] - (C^{t+1} + R^{t+1})$.

Table 11: The table shows the results for the premium risk. The micro model gives lower estimates of the standard deviation and VaR for $C^{t+1}+R^{t+1}$. Due to the construction of the risk model for the chain-ladder method no estimates are observed for the variables R^t and $C^{t+1} + R^{t+1}$.

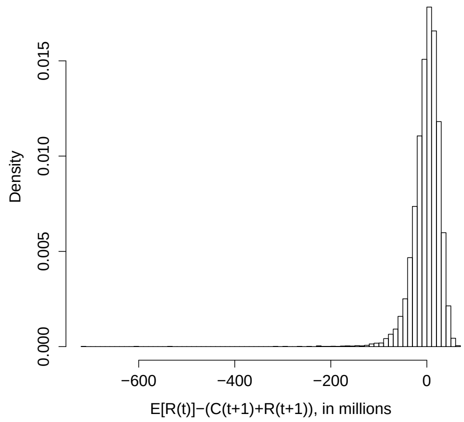
	Micro reserving		Chain-ladder	
$E[C^{t+1} + R^{t+1}]$	244 637 794		391 691 606	
	Sd	99.5%	Sd	99.5%
R^t	27 357 757	348 454 139	-	-
$C^{t+1} + R^{t+1}$	29 568 349	357 143 562	55 489 661	556 570 023
$-(E[R^t] - (C^{t+1} + R^{t+1}))$	29 568 349	-113 605 095	-	-



(a) One-year premium risk, micro model.



(b) One-year premium risk, chain-ladder.



(c) One-year premium risk, micro model.

Figure 7: Figures 7a and 7b show the densities of $C^{t+1} + R^{t+1}$. The density is much wider for the chain-ladder whereas the micro model reaches more extreme values. Figure 7c shows the density of $E[R^t] - (C^{t+1} + R^{t+1})$. A corresponding plot for the chain-ladder is not available for our construction of the chain-ladder risk model.

10 Discussion

The micro reserving model with special consideration for claims which are closed without ever receiving a payment is studied with real data. The results are compared to the classical chain-ladder method. The two methods give largely different results. The data shows an increase in the hazard rate of the events settled with a payment and more claims are settled early. The micro model estimates the RBNS reserve from reported open claims, while the chain-ladder method does not take the number of open claims into consideration. The micro reserving model has therefore a lower reserve with a relatively low RBNS. The underlying assumption of the chain-ladder method is that the same percentage is paid in the same development periods of the accident year, which is clearly not fulfilled for the dataset in question.

The model is validated by one-year back-testing of true payments. The micro model outperforms chain-ladder for the later years, since the model takes into account to a larger extent the time trends in the underlying data. The micro model assumptions are validated with a sensitivity analysis. The model is sensitive to the choice of distributions. However, it also gives the freedom to use expert knowledge for the choice of distribution or use the different distribution for various stress analysis of the estimated reserve.

Risk estimation follows naturally from the micro reserving model. The outstanding payments are simulated from a model which gives an empirical distribution of the reserves. The best estimate, standard deviation and other risk estimators are then easily calculated. In the Solvency II directive the time horizon is one year. We therefore take special look at the one-year reserve and premium risk. A natural way to estimate the one-year risk would be to simulate the payments made the following year and then re-estimate the reserve. The distribution of the original reserve, the payments and the reserve estimated one year later gives the one-year risk. Since the model is computationally heavy we did not re-estimate the parameters but draw new parameters from their corresponding asymptotic normal distributions. That does not give as large a spread in the resulting risk estimation but is an approximation. The standard deviation is lower for the micro model, however the risk models are largely different as the one-year risk does not as naturally follow from the bootstrapped chain-ladder method.

Micro modelling offers large possibilities for reserving. Using a model to simulate the reserves from is an attractive way to obtain predictive distributions and detailed information about different characteristics of the data. Compared to the traditional triangle methods, micro models in general have the advantage of using the detailed data available. Thus, reducing the uncertainty in the estimates. Though the model presented here may be more complicated and computationally heavy, there are large possibilities to develop micro models for use in practice.

10.1 Future Work

In the presented model, no robustness analysis was performed on the distribution assumptions. To use the model in practice one would have to include formal tests on the assumptions made. A usual problem in reserving is that the data does not contain enough history to estimate the reserve directly from the data. One could investigate how different stresses on the estimated parameters affect the reserve as well as the effect of adding a tail to the distributions.

In the model we have, like Antonio and Plat (2014) assumed that the hazard rates are independent of the delay. Due to poor performances the payments are assumed to be dependent on the time passed since the accident date. There are indications of the hazard rates being dependent on the delay, however due to changes in the claims that was not confirmed. Modifications to include the possible dependence would therefore be interesting.

Hazard rates are not easily interpreted and it can be difficult to relate to their estimates. Using hazard rates to predict the future payments is also computationally heavy. It could therefore be interesting to see if one could avoid using them to model the payment pattern, perhaps by simulating a final amount and afterwards spread out the payments to a less detailed extent.

Adding the probability of claims to reopen is a possible extension of the model. That could for example be adjusted for by increasing the intensities of the IBNR process.

Due to how computationally heavy the present model is, simplifications were made when calculating the reserve and premium risk. Future work may include a further look at the one-year risk where the model parameters are re-estimated after the one-year payments.

Accident years are taken as independent in the current model. That may be an unrealistic assumption, hence including trend estimation to the model would be an interesting way to continue the research. Time dependent delay distribution and hazard rates is one way to continue. Another would be to include more covariates in the model and to examine whether parameter estimation performs better when they are dependent on those covariates. To base the reserving model on claim data combined with policy data has been suggested by Gustafsson et al. (2012). Applying such a method on real data and comparing to results from other micro models would be highly interesting.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Petrov, B. N. and Csaki, F., editors, *Second International Symposium on Information Theory*, pages 267–281, Budapest. Akadémiai Kiado.

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

Arjas, E. (1989). The claims reserving problem in non-life insurance: some structural ideas. *ASTIN Bulletin*, 19(2):139–152.

Casella, G. and Berger, R. L. (2002). *Statistical inference*. Duxbury advanced series in statistics and decision sciences. Thomson Learning.

Cook, R. J. and Lawless, J. F. (2007). *The statistical analysis of recurrent events*. Statistics for biology and health. Springer, New York, London.

England, P. D. and Verrall, R. J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, 8:443–518.

England, P. D. and Verrall, R. J. (2006). Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science*, 1:221–270.

Gustafsson, E., Lagerås, A., and Lindholm, M. (2012). Simulation based claims reserving in general insurance. Research Report 2012:9, Mathematical Statistics Stockholm University, Stockholm.

Haastrup, S. and Arjas, E. (1996). Claims reserving in continuous time; a nonparametric bayesian approach. *ASTIN Bulletin*, 26:139–164.

Karr, A. (1991). *Point Processes and Their Statistical Inference, Second Edition,*. Probability: Pure and Applied. Taylor & Francis.

Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN Bulletin*, pages 213–225.

Merz, M. and Wüthrich, M. V. (2008). Modelling the claims development result for solvency purposes. *Causality Actuarial Society E-Forum*, Fall 2008:542–568.

Mikosch, T. (2009). *Non-Life Insurance Mathematics: An Introduction with the Poisson Process*. Universitext. Springer Berlin Heidelberg.

Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A, General*, 135:370–384.

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

Norberg, R. (1999). Prediction of outstanding liabilities II. Model extensions variations and extensions. *ASTIN Bulletin*, 29(1):5–25.

Ohlsson, E. and Lauzeningks, J. (2009). The one-year non-life insurance risk. *Insurance: Mathematics and Economics*, 45(2):203 – 208.

Parodi, P. (2014). Triangle-free reserving: A non-traditional framework for estimating reserves and reserve uncertainty abstract of the London discussion. *British Actuarial Journal*, 19:219–233.